Optical Computing

19950911 094

DTIC QUALITY INSPECTED 5

Bibliography of Proceedings of Recent Optical Computing Conferences

1993

Optical Computing Technical Digest, OSA, Washington DC, 1993, vol. 7, pp.334.

1992

Proc. ICO Topical Meeting on Optical Computing, Minsk, Eds. A.M. Goncharenko, F.V. Karpushko, G.V. Sinitsyn and S.P. Apanasevich, SPIE, vol. 1806, pp.680

1991

Optical Computing Technical Digest, OSA, Washington DC, 1991, vol. 6, pp.317.

1990

Optical Computing Conference Record, Jap. Soc. Appl. Phys., pp.422.

1989

Optical Computing Technical Digest, OSA, Washington DC, 1989, vol. 9, pp.420.

1988

Proc. Int. Conf. on Optical Computing, Toulon, Eds. J.W. Goodman, P. Chavel and G. Roblin, SPIE, vol. 963, pp.734.

DTIC
SELECTED
SEP 1 3 1995
F

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

OPTICAL COMPUTING

139

Optical Computing

Proceedings of the International Conference, Heriot-Watt University, Edinburgh, UK, 22–25 August 1994

Edited by B S Wherrett, Department of Physics, Heriot-Watt University, Edinburgh, UK and P Chavel, Institut d'Optique, Paris, France

Accesion For
NTIS CRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification .....
By .....
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1

Institute of Physics Conference Series Number 139
Institute of Physics Publishing, Bristol and Philadelphia

Copyright ©1995 by IOP Publishing Ltd and individual contributors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the written permission of the publisher, except as stated below. Single photocopies of single articles may be made for private study or research. Illustrations and short extracts from the text of individual contributions may be copied provided that the source is acknowledged, the permission of the authors is obtained and IOP Publishing Ltd is notified. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency under the terms of its agreement with the Committee of Vice-Chancellors and Principals. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients in the USA, is granted by IOP Publishing Ltd to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of $19.50 per copy is paid directly to CCC, 27 Congress Street, Salem, MA 01970, USA.
0305-2346/95 $19.50+.00

CODEN IPHSAC 139 1–660 (1995)

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library.

ISBN 0-7503-0126-0

Library of Congress Cataloging-in-Publication Data are available

Published by Institute of Physics Publishing, wholly owned by The Institute of Physics, London
Institute of Physics Publishing, Techno House, Redcliffe Way, Bristol BS1 6NX, UK
US Editorial Office: Institute of Physics Publishing, The Public Ledger Building, Suite 1035, Independence Square, Philadelphia, PA 19106, USA

Printed in the UK by J W Arrowsmith Ltd, Bristol

Contents

Conference papers presented but not submitted for publication xvii

Preface xxi

Organizing personnel xxiii

Sponsoring and cooperating organizations xxiv

I. DIGITAL OPTICAL PROCESSORS: Algorithms, Architectures and Implementations

Quantum-statistical restrictions on the information transmitting/processing rate in electronic and photonic channels
F V Karpushko and M A Khodasevich 3

Fault-tolerant design in digital optical computing
K Wagner 7

On parallel algorithms for optical image processors
P Chavel and P Lalanne 11

Duality between algorithms and optical implementations: actual examples
J P Heer and P Pellat-Finet 17

Simulation and benchmarking of a new algorithm for the optical cellular logic image processor
M P Y Desmulliez, B R Gillies, J F Snowdon and B S Wherrett 21

A methodology for optical architectures and modelling
M Guizani 25

Direct twos complement parallel array multiplication: algorithm and optical implementation
Guoqiang Li, Liren Liu, Lan Shao and Zhijiang Wang 29

Optical parallel arithmetic using a recoded signed-digit number system
M S Alam 33

Digital optical computing architectures for compute intensive applications
P S Guilfoyle 37

Parallel optoelectronic computing systems and applications
M Ishikawa 41

Free-space holographically interconnected counter
R J Feuerstein, D C O'Brien, A Fedor, C C Mao and L H Ji 47

Versatile compact image processor with optical feedback using photopolymer and ferroelectric liquid crystal on amorphous silicon
P Cambon, J Sharpe and K M Johnson 51

Cascaded optical data transfer through a free-space optical perfect shuffle
M W Derstine, S Wakelin and K K Chau 55

Optoelectronic implementation of tomographic sorting
P A Mitkas and L J Irakliotis 59

Sorting with an optoelectronic recirculating architecture
F R Beyette Jr, P A Mitkas, S A Feld and C W Wilmsen 63

Photothyristor-based stochastic elementary processor
G Prémont, P Lalanne, P Chavel, P Heremans and M Kuijk 67

The optical computing of the National 863 High Technology Programme in China
Yi-Mo Zhang, Shi-yong Liu and Rong-jun Pan 71

An optoelectronic hybrid parallel multiprocessor system
Yi-Mo Zhang, Wenyao Liu, Ge Zhou, Yong Wang, Dongjun Sun and Xiaoqing He 75

Control strategy for an optically reconfigurable architecture
M Murdocca, H R Nahata and Yu Zhou 79

Optical array logic network architecture
J Tanida and Y Ichioka 83

Optoelectronic multiport associative memory for data flow computing architecture
V B Fyodorov 87

HPT/VCSEL array implementation of an optoelectronic data filter
R D Snyder, F R Beyette Jr, S A Feld, K M Geib, L J Irakliotis, P A Mitkas and C W Wilmsen 93

Design of symbolic substitution systems for micro integration
C Passon, K-H Brenner and W Eckert 97

Optoelectronic 3D architectures and appropriate algorithms for TFLOP computing
D Fey, W Erhard and K Zürl 101

Planar–free-space optical circuits based on the 'transverse lock-and-clock' architecture
G V Sinitsyn, F V Karpushko, S P Apanasevich, A V Lyakhnovich, A S Yasukevich and M A Khodasevich 105

2D all-optical shift registers: numerical simulation
S P Apanasevich, A V Lyaknovich and G V Sinitsyn 109

Image processing by a system based on an optoelectronic active medium with controlled connections
Yu I Balkarey and M I Elinson 113

Parallel image processing by media of resonantly excited oscillators
Yu I Balkarey, A S Cohen, W H Johnson, M I Elinson, M G Evtikhov and P I Perov 117

Polarization-optical stacked integration scheme of a cellular logic two-layer image processor
Liren Liu, Yaozu Yin, Haifeng Peng, Bingquan Wang, Wei Xiong, Ning Wang, Feng Liang and Liying Zhao 121

Optical digital data processing based on the photon echo phenomenon
E A Manykin and S M Zakharov 125

II. OPTICAL INTERCONNECTS AND SWITCHING NETWORKS. OPTICAL TECHNIQUES

Free-space optical switching using FET-SEED smart pixel arrays
F B McCormick, A L Lentine, R L Morrison, J M Sasian, T J Cloonan, R A Novotny, M G Beckman, M J Wojcik, S J Hinterlong and D B Buchholz 131

3D free-space transmission polarization-based high-throughput photonic switching systems: architectural issues
V B Fyodorov 137

Architecture of a terabit free-space photonic backplane
T Szymanski and H S Hinton 141

A FET-SEED smart-pixel-based optical backplane demonstrator
D V Plant, B Robertson, H S Hinton, W M Robertson, G C Boisset, N H Kim, Y S Liu, M R Otazo, D R Rolston, A Z Shang and L Sun 145

Experimental results of a 64-channel, free-space optical inconnection network for massively parallel processing
I Redmond and E Schenfeld 149

Free-space WDMA optical interconnects using mesh-connected bus topology
Yao Li, S B Rao, I Redmond, Ting Wang and A W Lohmann 153

Free-space multicast non-blocking interconnection using multidimensional multiplexing concepts
Yao Li, Ting Wang, Z G Pan and J Sharony 157

Optical network for a general purpose massively parallel optoelectronic architecture: interconnection pattern, physical layout and experiment
S Kocon, P Churoux, M Fracès, J P Bouzinac, D Comte, N Hifdi and J Y Rousselot 161

Performance analysis of optical multistage networks and integrated holographic interconnection
I G Voitenko and T Yatagai 165

Systems partitioning and placement in optoelectronic MCM design
J Fan, S H Lee and C K Cheng 169

Design and implementation of a ring array optical interconnect for SIMD machines
J-M Wang, E Kanterakis, A Katz, Y Zhang and Y Li 173

Spatial-light-modulator-based routing switches
W A Crossland, R J Mears, S T Warr and R W A Scarr 177

Development of an optical free-space crossbar
H J White, G M Proudley, C Stace, N A Brownjohn, A C Walker, M R Taghizadeh, B R Robertson, C P Barrett, W A Crossland, R J Brocklehurst, M J Birch, M Snook and D Vass 183

A compact holographically routed optical crossbar using a ferroelectric liquid-crystal over silicon spatial light modulator
D C O'Brien and D J McKnight 187

Low-loss and high-speed optical switching modules for 1.3 μm wavelength using active-matrix ferroelectric liquid-crystal devices
S Shirai, T Serikawa, S Kohda, N Kakuda, M Okamura and N Yamauchi 191

Free-space optical interconnections using liquid-crystal microprism arrays
K Hirabayashi, T Yamamoto and M Yamaguchi 195

Demonstration of optically controlled data switching using quantum-well bistable devices and modulators
P Koppa, P Chavel, J L Oudar, R Kuszelewicz, J P Schnell and J P Pocholle 199

Parallel processing architectures with dynamic optical interconnections using spatial light modulators
N McArdle, J F Snowdon and M R Taghizadeh 203

Optical interconnect using pixellated spatial light modulators
J Gourlay, S Heddle, A O'Hara, S Samus and D G Vass 207

An optical package for a reconfigurable free-space optical interconnection network for multiprocessor systems
Mingcui Cao, Hongpu Li, Xianjun Zhao, Fengguang Luo, Xu Jun, Ronghan Wu and Wenzi Gao 211

A distributed, reconfigurable free-space optical interconnection network for massively parallel processing architectures
I Redmond and E Schenfeld 215

Reconfigurable optical interconnections in $Bi_{12}TiO_{20}$ photorefractive fibre
A A Kamshilin and T Jaaskelainen 219

Design of a lenslet-array-based free-space optical backplane demonstrator
B Robertson, G C Boisset, H S Hinton, Y S Liu, N H Kim, M R Otazo, D Pavlasek, D V Plant and D Rolston 223

An approach to active alignment of free-space optical interconnects
G C Boisset, B Robertson and H S Hinton 227

A novel programmable liquid-crystal lens device for adaptive optical interconnnect and beamforming applications
N A Riza and M C DeJule 231

Integration of fibre holder and microlens arrays by deep proton irradiation
M Kufner, S Kufner and P Chavel 235

Packaging tolerance constraints in the optimization of Fresnel computer-generated holograms for optical interconnections
B Kress, M Schillinger and P Meyrueis 239

Multilevel diffractive phase elements with trapezoidal-shaped pulses
P Blair, H Lüepken, M R Taghizadeh and F Wyrowski 243

Demonstration and discussion of an interlaced fan-out interconnect
N Collings and C Berger 247

Investigations into the use of lenslet arrays in optical signal processing and computing
M F Lewis and R A Wilson 251

Optical implementation of a crossover optical interconnect network using a Dammann grating with 65 × 65 spot arrays
Fengguang Luo, Mingcui Cao, Lingqiang Huang, Hongpu Li, Jun Ai, Jun Xu and Zhaiguang Li 255

Demonstration of a 3D integrated refractive microsystem
J Moisel and K-H Brenner 259

Planar integration of free-space micro-optical systems with refractive elements
W Eckert, K-H Brenner and C Passon 263

Direct laser writing of planar Fresnel elements for optical interconnects
M T Gale, M Rossi, R E Kunz and G L Bona 267

Strategies for realizing large computer-generated holograms from the calculation of sub-holograms
R I Young, A E MacGregor and C W Slinger 271

Three-dimensional distribution of light generated by a diffractive element
R Piestun, B Spektor and J Shamir 275

III. ANALOGUE OPTICAL PROCESSING

Theoretical results on accuracy limitations in analog optical processors
D A Timuçin, J F Walkup and T F Krile 281

Space-variant filtering in fractional Fourier domains
H M Ozaktas, B Barshan, D Mendlovic and H Urey 285

A novel form of incoherent optical correlator
M F Lewis 289

Experimental implementation of a joint transform correlator providing rotation invariance
L Bigué, M Fracès and P Ambs 293

Associative memory for rotation pre-processing and projection invariant pattern recognition
G Lebreton 297

Adaptive pattern recognition with rotation, scale and shift invariance
E Silvera, T Kotzer and J Shamir 301

Application of optical multiple correlation to recognition of road signs: the ability of multiple correlation
K Matsuoka, M Taniguchi and Y Mokuno 305

Parametric design of correlation filters for multiple-object discrimination
M Taniguchi and K Matsuoka 309

An analog retina for edge detection
Chunyan Wang and F Devos 313

Optoelectronic implementation of a phase-retrieval Vander Lugt correlator
S Vallmitjana, A Carnicer, E Martín-Badosa and I Juvells 317

Performance of an acousto-optic joint transform correlator
C D Reeve, A W Houghton and F D Bryant 321

A zeroth-order non-heterodyning space-integrating acousto-optic correlator
A W Houghton and C D Reeve 325

Edge enhancement in photorefractive joint transform correlators
O Daniel, J-M C Jonathan and G Roosen 329

Optical correlators for the analysis of images with the computer processing of the results
A V Korolkov and Y S Mosyakin 333

Optical coordinate transformations: a new architecture
Ronglong Tian, M Itoh and T Yatagai 337

Breaking symmetry in the binary-phase-only matched filter
T D Wilkinson and R J Mears 341

Universal computation and optical correlators
M Montes-Usategui, I Juvells, J Campos and J R de F Moneo 345

A filter design to improve the recognition of defocused images
J Sallent, M Montes-Usategui, J Campos and S Bosch 349

Optical Haar wavelet transform using a time-division-muliplexed shadow-casting system
Ju-Seog Jang and Shin Il Jeong 353

Fractional Fourier transforms, imaging systems and correlators
L M Bernardo, F J Marinho and O D D Soares 357

Programmable 2D optical wavelet transform in the space domain
Dong-Xue Wang, Ju-Wei Tai, Yan-Xin Zhang and Yu Ma 361

Low-resolution Fresnel encoded lenses applied to pattern recognition
E Carcolé, I Juvells and J Campos 365

Region of support to improve discrimination of a phase-only filter
E Ahouzi, J Campos and M-J Yzuel 369

Modified filter-synthetic discriminant functions for pattern recognition
R K Wang, D M Budgett and C R Chatwin 373

Multi-distance, multi-projection, parallel projection method
T Kotzer, N Cohen, J Shamir and Y Censor 377

Fourier optics and holography of temporal signals
Yu T Mazurenko, A G Spiro, S E Putilin and A G Beliaev 381

'Spatial brightness amplification' for increasing the signal-to-noise ratio in 2D optical information processing
G V Sinitsyn, S P Apanasevich, F V Karpushko and A V Lyakhnovich 385

Optical heterodyning noise-protected with stochastic resonance
M I Dykman, G P Golubev, I Kh Kaufman, D G Luchinsky, E A Zhukov, P V E McClintock, N D Stein, and N G Stocks 389

IV. NEURAL NETWORKS

Optical implementations of a stochastic neural system
W A Crossland, T J Hall, J S Shawe-Taylor and M van Daalen 395

An optoelectronic neural network with on-line learning
D Pignon, T J Hall, L Q Xu, K B Russ, S R Cherry, N C Roberts, P Prewett and P Hallowell 399

Optical lateral inhibition networks using self-linearized SEEDs
P Horan 403

Optical spatial/frequency filtering and adaptive neural networks based on fractional Fourier transforms
Soo-Young Lee 407

All-optical self-organizing map applied to character recognition
K Heggarty, J Duvillier, E Carpio Pérez and J L de Bougrenet de la Tocnaye 411

Evaluation of a fast-access phase-encoded photorefractive memory
C Alves, L Gérard, G Pauliat and G Roosen 415

Photorefractive implementation of neural 'self-correcting' learning algorithms
T Galstyan, G Pauliat, A Villing and G Roosen 419

Dynamic effects in volume interconnects with feedback
C Slinger and P Conway 423

Holographic associative memory using a learning pattern method
Ho Hyung Suh, Chong Hoon Kwak, Gyungock Kim and El-Hang Lee 427

Optical associative memory with bipolar edge-enhanced learning
Xu-Ming Wang, T J Hall and Jian Wang 431

Single-rail optical associative memory
A Bergeron, H H Arsenault and D Gingras 435

High-efficiency photorefractive storage for multimode phase-conjugate resonators
C C Chang and D R Selviah 439

Optical learning neural network with a reversal-input superposing technique
M Mori, M Kondo, Y Hayasaki, I Tohyama and T Yatagai 443

Feature extraction pre-processing and an optical neural network
I Tohyama, M Kondoh, T Yatagai, M Mori and Y Hayasaki 447

Optical digit recognition based on Kohonen maps
M Barge, R Chevallier, E Curatu and A Maruani 451

Overcoming noise in a fast neutral network
R P Webb 455

Optical self-routing complex neural network using mixed negabinary numbers
Liren Liu, Guoqiang Li, Lan Shao and Yaozu Yin 459

Neural networks with unipolar weights and normalized thresholds
J S Bródka and B Macukow 463

Experimental investigation of the performance of the two-layer neural network based on an optical vector-matrix multiplier
N N Evtihiev, R S Starikov, B N Onyky, V V Perepelitsa and I B Scherbakov 467

Features of application of holographic memories in PC networks and optical neural computers
V N Shahgedanov 471

Optoelectronic fuzzy control of an inverted pendulum using light-emitting-diode arrays and position-sensing devices
H Itoh, B Houssay, S Mukai, T Yamada and S Uekusa 475

V. ARRAY DEVICES: Sources, Smart Pixels and Spatial Light Modulators

Independently addressable vertical-cavity surface-emitting laser diode arrays
E Zeeb, B Möller, T Hackbarth, H Leier and K J Ebeling 481

Non-uniformity tolerance and its price for different modes of operation of FET-SEED smart pixel arrays
L M F Chirovsky, A L Lentine, T K Woodward, G Livescu and G D Boyd 485

Optical, algorithmic and electronic considerations on the desirable 'smartness' of optical processing pixels
M P Y Desmulliez, B S Wherrett, J F Snowdon and J A B Dines 489

Logic and layout design of optically interconnected self-routing exchange/bypass node arrays
M P Y Desmulliez, F A P Tooley, G J Crowder, B S Wherrett, P W Foulk, S Ashcroft, M Coy and R A Novotny 493

Architecture of a field-programmable smart pixel array
T H Szymanski and H S Hinton 497

The Scottish Collaborative Initiative on Optoelectronic Sciences (SCIOS)—devices and demonstrators for free-space digital optical processing
A C Walker, D J Goodwill, B S Ryvkin, M McElhinney, F Pottier, B Vogel, M C Holland and C R Stanley 501

Smart pixels with VCSELs: potential and demonstration system
K Zürl, E Gluch, B Schäfer, J Schwider, D Fey and W Erhard 507

A smart-pixel optical neural network design using customized error propagation
A J Waddie and J F Snowdon 511

Dynamic studies of GaAs/AlGaAs multiple-quantum-well modulators using photon-counting techniques
G S Buller, J S Massa, S J Fancey and A C Walker 515

The monolithic integration of quantum-well high-electron-mobility field-effect transistors, and asymmetric Fabry–Perot optical modulators
G O'Sullivan, T Aherne, E McCabe, J Hegarty, P Horan and B Corbett 519

Contactless quantum confined Stark modulator based on GaAs/AlAs quantum filters
N T Pelekanos, B Deveaud, P Gravey, F Clérot, J M Gérard, J Hebling and J Kuhl 523

Smart pixel arrays: challenges of manufacturing and testing
G Livescu, L M F Chirovsky, L A D'Asaro, S Hui, B Tseng, R A Novotny, A L Lentine, R E Leibenguth, M W Focht, J M Freund, K G Glogovsky, G D Guth and G Przybylek 527

Monolithic integrated optoelectronic smart pixels
U Kehrli, D Leipold, K Thelen, O J Homan, H P Schweizer, P Seitz and B D Patterson 531

Liquid-crystal over silicon spatial light modulators
D J McKnight, M A Follett and K M Johnson 535

High-speed liquid-crystal on silicon spatial light modulators
C C Mao, D J McKnight and K M Johnson 539

Electronically addressed ferroelectric liquid-crystal over silicon spatial light modulators
D C Burns, I Underwood, A O'Hara and D G Vass 543

Sensitivity transition of LAPS-SLM and its application for optical neural networks
N Kato, N Kasama, R Sekura, K Matsusita, Y Mitsuoka and T Iwaki 547

VI. OPTICAL STORAGE AND SWITCHES. GUIDED WAVE TECHNIQUES

All-optical dynamic memories
M P Petrov 553

A novel information storage device utilizing photostimulated luminescence in Eu-doped KCl phosphors
H Nanto, F Endo, Y Hirai, M Ikeda, S Nasu and N Takeuchi 561

AlAs/GaAs multilayered structures as low-intensity optical media
S Knigge, M Wicke, D Niederland and D Jäger 565

Large reduction of saturation carrier densities by strain in InGaAs/AlGaAs quantum wells
M Moloney, J Hegarty, L Buydens and P Demeester 569

Dynamic behaviour of opto-optical logic switching devices employing n-i-p-i-based smart pixels
P Kiesel, M Kneissl, P Riel, K Reingruber, K H Gulden, E Greger, B Knüpfer, X X Wu, J S Smith and G H Döhler 573

Optical bistability above the bandgap energy in a GaAs Fabry–Perot
C Bagnoud and N Collings 577

An optical set–reset flipflop semiconductor laser with two mutually complementary outputs
M Watanabe, S Mukai and H Yajima 581

Resonant-tunnelling triangular-barrier optoelectronic switch (R-TOPS) for optical logic
H Sakata, K Utaka and Y Matsushima 585

Triangular-barrier optoelectronic switch (TOPS) for optical logic
H Sakata, K Utaka and Y Matsushima 589

All-optical bistable devices (ORION) using laser diodes coupled to interference filters of narrow spectral bandwidth
Y Maeda, M Migitaka, Hong K Kim, Yabo Li, Cheng C Li and D W Langer 593

Carrier-heating-induced optical bistability in degenerate semiconductors and quantum-well structures
V I Tolstikhin and M Willander 597

Photonic devices based on excited-state non-linear absorption
Chun-Fei Li and Miao Yang 601

Polychromatic all-optical gate based on second-order non-linearity
V P Torchigin and A E Kobyakov 605

Diffraction kinetics of electronic and thermal transient gratings in GaAs epilayers and GaAs/GaAlAs multi-quantum wells
N Gouaichault, R Grac, L Legratiet, E Bedel, C Fontaine, J L Iehl, M Pugnet and J H Collet 609

Design examples of wavelength conversion and filtering components
J-i Mizusawa 613

Waveguide-type rotating-phase plate as an optical frequency shifter
H Inoue, S Nishimura, T Kanetake and H Takano 617

Sequences of bifurcations and transition to chaos in an optical processing element
A González-Marcos and J A Martín-Pereda 621

New optical effects stimulated by anisotropy in information processing schemes
A M Goncharenko, N A Khilo, N N Korneyev, V N Belyi and A G Khatevich 625

Architectures for space-wavelength photonic switching using collinear guided-wave acousto-optics
V V Proklov, A U Bashlakov and V A Birjukov 629

An acousto-optic FM demodulator with inherent AM and laser noise rejection
P Brooks and C D Reeve 633

Integrated-optic commutator for a high-speed local communication network with cross-talk minimization in optical directional couplers
I G Voitenko and T Yatagai 637

Optical characteristic of photorefractive three-dimensional waveguides in lithium niobate
K Itoh, O Matoba and Y Ichioka 641

Fibre-optical loop memory structure for an optoelectronic computer
V A Pilipovich, A K Esman, I A Goncharenko, V S Posedko and I F Solonovich 645

All-optical arithmetic operations by means of spatial soliton interactions
F Garzia, C Sibilia and M Bertolotti 649

New types of optical soliton-like structures and their possible applications to optical computing
A V Fedorov, S V Fedorov, G V Khodova, N N Rosanov, V A Smirnov and N V Vyssotina 653

Asymmetric light bullet dragging logic
R Mcleod, S Blair and K Wagner 657

CONFERENCE PAPERS PRESENTED BUT NOT SUBMITTED FOR PUBLICATION

Architecture design and implementation issues for massively parallel processors
S Nelson

Creative uses of redundancies in optical computing
H J Caulfield

Digital optical computing demonstration systems
F A P Tooley, S Prince, D Baillie, D J Goodwill, M P Y Desmulliez and M R Taghizadeh

Programmable logic array with differential pairs of pnpn photothyristors: an experimental assessment
A Kirk and H Thienpont

Optical fibre parallel logic processor
M R Wahiddin and S Abdul Halim

Parallel information recording and processing using lD hologram technology
A L Mikaelian

Discrete incoherent shadow-casting: theoretical investigation
V Laude, P Chavel and P Réfrégier

Optical free-space sliding tandem Banyan architecture for self-routing switching networks
M W Haney and M P Christensen

Diagnostic techniques for investigating free-space photonic switching systems
R L Morrison, F B McCormick, T J Cloonan, A L Lentine, R A Novotny, J M Sasian, M G Beckman, M J Wojcik, F A P Tooley, E Kerbis and S J Hinterlong

A cell-based algorithm for designing beam array generators
R L Morrison and M J Wojcik

Fabrication of fibre arrays for optical computing and switching systems
J M Sasian, R A Novotny, M G Beckman, S L Walker, M J Wojcik and S J Hinterlong

Integration of free-space interconnects using selfoc lenses: optical properties of a basic unit
K Hamanaka, K Nakama, D Arai, Y Kusuda, T Kishimoto and Y Mitsuhashi

Hybrid lens implemented using an afocal microlens doublet
S M Prince, F A P Tooley and M R Taghizadeh

Microlens array for perfect shuffling fabricated by a laser writing system
S H Song and S-D Jung

Focal length considerations in the design and manufacture of lenses for optical computing
C P Barrett and D T Neilson

Design and fabrication of wide-angle thin-film components for optical computing applications
G S Buller, E C Smith, D T Neilson and G MacKinnon

Optoelectronic parallel processing system with reconfigurable diffractive interconnections
A Kirk, T Tabata, T Ishida and M Ishikawa

Optical detection filters and algorithm fusion
D Casasent

Optimal non-linear filtering for pattern recognition and optical implementation
P Réfrégier, B Javidi, V Laude and J-P Huignard

Time-integrating correlation using a fibre-optic delay line processor
R A Athale and G W Euliss

Optical implementation of signal feed-forward and error back-propagation with a ground glass
H-J Lee, S-Y Lee and S-Y Shin

A normalizing detector array for an adaptive neural network
J Sharpe, I Bar-Tana and K Johnson

Two-layered bidirectional optical neural network based on the holography correlator: variation of the attractor position to achieve new solution types
A V Pavlov

Optoelectronic sorting neurochip
V R Grigor'ev and S P Naumov

Large-scale integration of LEDs and GaAs circuits fabricated through MOSIS
A C Grott, D Psaltis, K V Shenoy and C G Fonstad Jr

Optoelectronic gain of phototransistor/surface-emitting-laser smart-pixel logic gates
S A Feld, F R Beyette Jr, K M Geib, R D Snyder and C W Wilmsen

Quantum-well smart pixels for optical switching and processing
D A B Miller

Arrays of field-effect-transistor self-electro-optic-effect-device (FET-SEED) differential transimpedance amplifiers for two-dimensional optical data links
R A Novotny, M J Wojcik, A L Lentine, L M F Chirovsky, L A D'Asaro, M W Focht, G Guth, K G Glogovsky, R Leibenguth, M T Asom and J M Freund

Saturation effects and wavelength optimization of quantum-well modulators in smart pixels
G D Boyd, L M F Chirovsky, A L Lentine and G Livescu

High-speed parallel switching of symmetric self-electro-optic-effect devices (S-SEEDs)
D J Goodwill, D A Baillie and F A P Tooley

Photothyristor smart pixels in MIMD optical computing architectures
T Van de Velde and H Thienpont

Operation of 16 × 16 array of strained MQW InGaAs/GaAs S-SEEDs for 1047–1064 nm
D J Goodwill, A C Walker, M McElhinney, F Pottier and C R Stanley

Optoelectronic switch operating with 0 2 fJ μm^{-2} at 15 MHz
M Kuijk, P Heremans, R Vounckx and G Borghs

Optical circuitry for data transcription and digital optical logic based on photothyristor differential pairs
H Thienpont, T Van de Velde, A Kirk, W Peiffer, M Kuijk, W Stevens, J Fernandez, I Veretennicoff, R Vounckx, P Heremans and G Borghs

The completely depleted PnpN optoelectronic switch
P Heremans, M Kuijk, R Vounckx and G Borghs

Silicon-implanted oxide-based optical logic gates for use in silicon smart-pixel circuitry: device characterization, modelling and optimization
W Peiffer, H Thienpont, I Veretennicoff, C de Tandt, W Ranson, R Vounckx and A Koster

Fractal dimension as a new method of the precision evaluation on a non-linear device
A González-Marcos

Preface

The 1994 International Conference on Optical Computing took place at the Edinburgh Conference Centre, Heriot-Watt University, UK on 22–25 August.

The conference series is organized under the auspices of a steering committee established by the International Commission for Optics. Previous meetings include those at Minsk, Belarus (1992), Kobe, Japan (1990) and Toulon, France (1988). In alternate years the conference is held in the USA, as a Topical Meeting of the Optical Society of America—Palm Springs (1993), Salt Lake City (1991, 1989). The subject itself, however, can be traced back to the early 1960s, and many meetings, some with published proceedings, were held prior to the ICO and OSA series.

The particular topics covered in Optical Computing conferences have varied over the years, as new optical techniques have become recognized and more especially as new devices have become available. For example one of the earliest meetings, the 1964 Symposium on Optical and Electro-Optical Information Processing, was held in Boston, Massachusetts, just four years after the demonstration of the laser. The meeting included a wide range of topics from coherent optical processing and synthetic aperture radar processing through to optical storage and ideas for optically bistable devices. Fourier and acousto-optic-based analogue processing were to dominate the subject through the 1970s; optical correlators continue to evolve and new developments may be found in the papers in Part III of the present Proceedings. Acousto-optic- (Bragg-cell-) based matrix–vector processors also began to appear in the mid-1970s. These topics still formed the mainstream for conferences through to 1983. Topics such as tomography and radioastronomy also appeared briefly during the 1970s meetings, but have since been treated in conferences of their own, as has much of integrated optics.

By 1984 the advantages of optical interconnections within computing, using fibres, waveguides or free space, came under the umbrella of optical computing. The guided-wave work has tended recently to be covered in Photonic Switching meetings, but free-space interconnection and associated switching networks formed a major part of the 1994 meeting—Part II herein.

1985 saw the introduction, at the Lake Tahoe conference, of sessions on optically bistable devices and systems. Low-power optical bistability had been researched for five years, and optical logic and volatile memory devices were being studied. The possibilities for all-optical computing stemmed from this work and formed a significant part of meetings through to the 1990s. Work continues in this area; new devices are covered in Part VI. The trend, however, has been toward 'smart-pixel' technologies, with optically interfaced arrays of electronic cells. These were first reported at the 1991 Topical Conference at Salt Lake City; they are treated here in Part V.

With the injection of further enthusiasm, and new architectural opportunities, provided by smart pixels, digital optical computing remains one of the major themes of the subject. Part I of this Proceedings contains papers on algorithms, architectures and a range of implementations of digital optical schemes for parallel processors.

In addition to bistability, 1985 saw the development of spatial light modulators (Part V) and their use initially as inputs to Fourier processors. Optical neural networks (Part IV) were also first introduced at the '85 meeting; from the first digital Hopfield demonstrators advantage has been taken of Fourier (holographic) techniques and the reconfigurability of interconnections offered by advances in photorefractive materials. The significance of both device developments and of materials advances has been long recognized in optical computing. The 1990 Kobe conference was especially strongly influenced by these areas, and the recent smart-pixel possibilities owe much to developments in semiconductor growth and fabrication techniques.

The purpose of the Optical Computing conference series is both to enable the communication of state-of-the-art information among delegates from around the world, and to foster interactions and friendships among members of the scientific community. The Heriot-Watt conference was attended by over 250 delegates, originating from 24 countries. There were 9 extended invited talks, 52 contributed talks and 135 poster presentations during the 4-day meeting, including 17 post-deadline presentations. The potential applications for optics in computing were discussed throughout, and as a foil an invited talk on 'Architecture design and implementation issues for massively parallel processors' was presented by Steve Nelson of Cray Research Inc. There was a healthy mix of attendees from both the academic and industrial research communities, with papers covering topics from fundamental information theory in electronics and photonics, through optical phenomena, devices, architectures and processor implementations, to integration and packaging issues.

It is my belief that both the technical and social aspects of the meeting were appreciated by all delegates. I must thank the personnel of the Edinburgh Conference Centre at Heriot-Watt for their professionalism, which established the framework for the smooth organization of the meeting. The conference could not have taken place without the backing of the Department of Physics at Heriot-Watt, and the willing assistance of many members of the academic, technical and secretarial staff, and graduates of the department. It goes without saying that the efforts over a long period of the Local Organizing Committee (Andy Walker, Mo Taghizadeh, Frank Tooley, John Snowdon, Jimmy Smith and Janice McClelland) are thoroughly appreciated. My thanks to all.

For assistance with the production of this Proceedings I should particularly like to thank Janice McClelland and Florence Jensen.

Brian Wherrett
1994

Organizing Personnel

General Chairman

Brian S. Wherrett	Heriot-Watt University, UK

International Advisory Committee

Pierre Chavel	Institut d'Optique, Paris, France (ICO representative)
Joseph W. Goodman	Stanford University, USA
Satoshi Ishihara	AIST, MITI, Japan
Fedor Karpushko	Academy of Sciences, Belarus
S. Desmond Smith	Heriot-Watt University, UK
Andrew C. Walker	Heriot-Watt University, UK

Technical Programme Committee Co-Chairs

Joseph W. Goodman	Stanford University, USA (ICO representative)
Yoshiki Ichioka	Osaka University, Japan
S. Desmond Smith	Heriot-Watt University, UK

Members

J.-L. de Bougrenet	ENST de Bretagne, France
Karl-Heinz Brenner	University of Erlangen, Germany
Vsevolod S. Burtsev	Academy of Sciences, Moscow, Russia
Pierre Chavel	Institut d'Optique, Paris, France
H. Scott Hinton	McGill University, Canada
Kristina M. Johnson	University of Colorado, USA
Soo Young Lee	Korea Advanced Institute of Technology, South Korea
Roger A. Lessard	Laval University, Canada
Chun-Fei Li	Harbin Institute of Technology, P.R.China
David A.B. Miller	AT&T Bell Labs., Holmdel, USA
John A. Neff	University of Colorado, USA
Alexander Sawchuk	University of Southern California, USA
Joseph Shamir	Technion, Israel
George Sinitsyn	Academy of Sciences, Belarus
Kunio Tada	University of Tokyo, Japan
Roger Vounckx	Vrije University of Brussels, Belgium
Toyohiko Yatagai	University of Tsukuba, Japan

Proceedings Editors

Brian S. Wherrett	Heriot-Watt University, UK
Pierre Chavel	Institut d'Optique, Paris, France

Local Organizing Committee

The Department of Physics, Heriot-Watt University was pleased to provide the local organization for OC'94.

Brian Wherrett	Chairman
Andy Walker	Treasurer
Mohammad Taghizadeh	Coordinator (Publicity)
Frank Tooley	Coordinator (Registration)
John Snowdon	Conference Secretary
Jimmy Smith	Social Events Organiser
Janice McClelland	Committee Secretary

Sponsoring and Cooperating Organizations

International Commission for Optics
Optical Society of America
SPIE
IEEE/Laser and Electro-Optics Society
The Japan Society of Applied Physics
European Physical Society
European Optical Society
IEE, UK
Institute of Physics, UK

In addition to expressing our appreciation of the sponsorship provided by the Cooperating Societies, the organizers of OC'94 gratefully acknowledge the considerable support and financial assistance provided by the following organizations and companies:

Lothian and Edinburgh Enterprises Ltd.
The City of Edinburgh
The European Research Office of the US Army
BNR Europe Ltd.
Pilkington Technology Management Ltd.
Taylor & Francis Ltd.
Sharp Laboratories of Europe
The International Science Foundation
The Royal Society of London

I. DIGITAL OPTICAL PROCESSORS

Algorithms, Architectures and Implementations

Inst. Phys. Conf. Ser. No 139: Part 1
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Quantum-statistical restrictions on the information transmitting/processing rate in electronic and photonic channels.

F.V.Karpushko and M.A.Khodasevich

Division for Optical Problems in Information Technologies, Academy of Sciences of Belarus, P.O.Box 1, 220072, Minsk, Republic of Belarus

Abstract. The behaviour of noisy electronic and photonic channels are significantly different. This is due to the fact that bosons and fermions obey different statistics. For one dimensional signals quantum effects restrict the information capacity of channels at input data rate $\geq 10^{12} nit/s$. Below this rate both types of channels behave classically. At higher input data rates the energy required for transmission of information is 2 – 4 orders less in one-dimensional boson channel as compared with the corresponding one-dimensional fermion channel. Three-dimensional optical information channels are free of quantum limitations up to information rates $\sim 10^{19} nit/cm^2 s$.

1. Introduction

The basic results in information theory concerning the channel capacities, transmitting and processing rates, etc. have been obtained by use of classical physics methods. The investigation of the information features of the electromagnetic (photon) channels in the presence of noise, treated on the basis of Bose-Einstein statistics [1], shows quantum effects to be significant at certain conditions. However, at low signal-to-noise ratios photonic channels behave classically and are described well by the Shannon information theory. The latter does not differentiate between photon and electron channels from the point of view of the quantum nature of the information carrying agent.

Since photons and electrons obey different statistics, the relationships between signal and noise are different for these two cases. For photons the power of the total radiation (signal and noise) can be considered as an independent sum of the corresponding values of the signal and noise powers; because the occupation of any microstate either by signal photons or by thermal (noise) photons is statistically independent. But this is not so for electrons; owing to the Pauli principle a definite microstate of noise restricts the variety of possible microstates of signal and vice versa. Obviously this difference leads in the quantum limit to different possibilities for transmitting information through electronic and photonic channels with presence of noise. This was underlined in [1]. However, to date we know of no corresponding consideration for the case of fermion channels. Meanwhile, such a consideration becomes more necessary as the speed of information processing systems increases in line with decrease of the physical extent (micro- and nano-structures) of electronic and opto-electronic devices that leads to high energy densities of the data flows. In the present short length paper we summarize our results

on the general information features of the fermion (electron) channels with an equilibrium noise studied on the basis of Fermi-Dirac statistics. The corresponding properties of the fermion and boson (all-optical) channels are also compared. The results are interesting for electronics alone and give a fundamental basis for the comparison with other technologies.

2. Formulation of the problem

Let us consider an information channel consisting of a transmitter, an information transmission agent or physical carrier with a certain space-time extent, i.e. the channel in itself, and a receiver. It is assumed that the role of the physical carrier is played by a quantum ensemble governed by either Fermi-Dirac or Bose-Einstein statistics. The transmission of data in such a channel means the formation of a certain microstate of ensemble, say, in terms of occupation numbers assigned to energy levels of the channel degrees of freedom. We also assume that the physical carrier of information is affected by some noise sources. Due to this noise the microstate of the physical carrier determined by the receiver at the channel output will be different from that introduced by the transmitter at the input.

Abstracting ourselves from possible physical limitations of both transmitter and receiver, we aim to find out the limitations related only to the information channel in itself with respect to its quantum properties and the existence of noise. As an example of fermion channel we consider here an electron channel which is one-dimensional due to its nature.

Accounting for the specific of the problem, the energy distribution over the signal degrees of freedom is considered in the "time-frequency" phase space. For the case of one-dimensional (1D) stationary temporal signal, its energy is represented by its Fourier harmonics. We introduce the quantum properties into the information channel allowing the energies of Fourier harmonics (channel's degrees of freedom) be quantized.

Since we deal with an ensemble (of particles, signal microstates, degrees of freedom, etc.) we must speak of the mean occupation numbers of the ensemble elements. Thus, from more general point of view the choice for the "time-frequency" phase space of the physical carrier of information instead of the "coordinate-momentum" phase space means no more than a replacement of the mean over the ensemble by the mean over the time (Ergodic theorem).

In accordance with the definition of information, this is the average measure of uncertainty for the prevision of the result of completely random nature (see, for example, original work of Shannon [2] and the interpretation on the basis of the Brillouin negentropy principle [3]). The mathematical expression for the value of information has the form of an entropy equation in thermodynamics. The only difference is that in statistical thermodynamics the entropy is understood as a measure of disorder.

The information channel capacity is the maximum amount of information transmitted by a channel per unit time and is given as:

$$C = \frac{1}{\tau}\left[H_{max} - H_0\right] \tag{1}$$

where $H_{max} - H_0$ is the so-called entropy defect of the physical system considered and τ is the signal duration. The entropy value H_0 corresponds to the case when the channel input is influenced by a determinate signal. In contrast, H_{max} is the maximum possible

value of entropy of the system which corresponds to its thermodynamic equilibrium, i.e. when probabilities of the system microstates are given by the Gibbs distribution. Such an equilibrium channel state arises under the excitation of the channel completely at random. In other words, the entropy defect shows how far the determinate choice of the signal deviates an information channel (as a physical system) from its thermodynamic equilibrium. Thus, the evaluation of the information capacity of a particular channel is related to the problem of computation of the above H_{max} and H_0 values of the entropy.

For the case when information is communicated through a channel with noise, the total energy E_J received by the receiver from the channel output during the time interval τ consists of the energy of the signal E_S introduced to the channel input and an accumulated noise energy E_N:

$$E_J = E_S + E_N \,. \tag{2}$$

This gives a formal basis to evaluate the above maximum entropy values H_{max} for both quantum types of channels, boson and fermion, by using the corresponding distribution functions. This leads to the unique establishment of the corresponding effective temperatures T_J of the channels and allows to express the values of the joint energies E_J and, consequently, the maximum entropies H_{max} as functions of these effective temperatures: $E_J = E(T_J); H_{max} = H(T_J)$, at least implicitly.

For the entropy value H_0, however, the problem is not so trivial. For a boson channel, the signal and noise joint distribution function can be considered with no restrictions as the sum of signal and noise distribution functions. Hence, in the presence of a determinate signal (this means that the signal probability equals 1) the total energy transmitted through the channel during time τ is the sum of the signal and noise energies, but the entropy flux remains unchanged because the microstate of the signal is completely determined and it does not add any part to the entropy. The latter is then the entropy of noise due to the channel interaction with an environment at temperature T_0. Hence we can take the value H_0, which we are interested in Eq.(1), as the boson noise entropy: $H_0^B = H^B(T_0)$.

For the fermion case, however, introducing a "determinate" signal to the channel input is questionable. Noise occurring in the channel restricts the occupation of certain channel degrees of freedom. This is due to the fact that the same energy state can not be occupied by more than one element in the case of a fermion ensemble: fermions obey the Pauli principle.

As a result of this statistical interdependence of signal and noise, we can introduce a particular signal with, at the best, some maximum probability but less than 1, to a fermion channel. Hence it is not possible to reduce the entropy value H_0^F to the fermion noise entropy $H^F(T_0)$. However, as we are not able to compute the value H_0^F at this stage explicitly, we may replace it by $H^F(T_0) < H_0^F$ in Eq.(1). This allows us to evaluate the upper limit of a fermion information channel capacity for an arbitrary signal energy E_S, in the same way as for a boson channel.

3. Information features of the electron and photon channels and their comparison

Based on the above approach we have simulated some information features of the fermion and boson channels and compared them with assumption that they are in the same

conditions of interaction with their environment. Fig.1 shows some examples of the behavior of the information capacities for both types of channels. More complete results of the study undertaken are to be published [4].

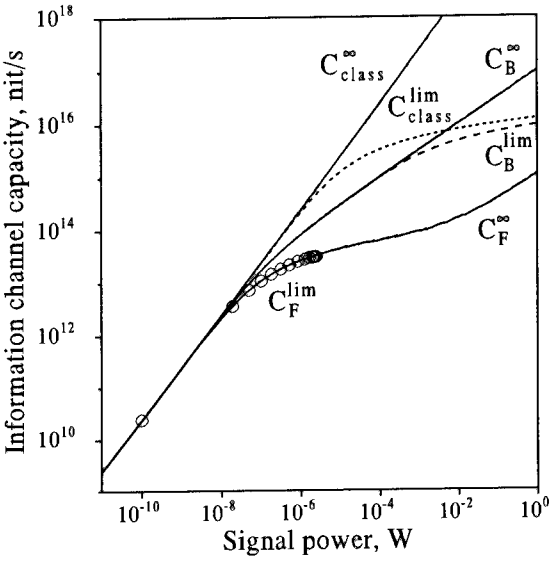


Fig.1. Information capacities for one-dimensional fermion and boson channels at $T_0 = 300K$ as functions of the signal power. C_{class} are classic (Shannon) capacities, C_F and C_B are capacities for fermion and boson channels respectively. The superscripts "∞" and "*lim*" indicate the broad and finite band cases respectively. The finite band fermion channels (a case of the bandwidth of 10^{15} Hz shown by broken off sequence of circles o) have the maximum information capacities which correspond to the maximum input signal power allowed. (1 nit = ln2 bit)

The obtained results demonstrate that quantum-statistical effects become significant for both types of channels when the signal power is essentially represented by its harmonics with frequencies which exceed the effective boundary frequency of the noise. For 1D channels the beginning of the influence of quantum effects corresponds to input data rate of about $10^{12} nit/s$. This is the *fundamental limit for electronic systems.*

For finite band fermion channels their information capacities are limited, in principle, whilst for finite band boson channels are not.

3D optical channels are free of quantum statistical restrictions up to input data rates of about $10^{19} nit/cm^2 s$ [4].

In the region of quantum effect influence, the data transmitting and processing are no longer the linear processes in the sense of the independence of the information features of channels on signal powers/energies. This implies the beginning of the *nonlinear dynamics of the information processes.*

References

[1] Lebedev D S and Levitin L B 1966 *Information and Control* **9** 1
[2] Shannon C E and Weaver W 1949 *The mathematical theory of communication* (Univ. of Illinois)
[3] Brillouin L 1960 *Science and information theory* (2d Edition NY: Academic)
[4] Karpushko F V 1994 *J. Appl. Phys.* to be published

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Fault-Tolerant Design in Digital Optical Computing

Kelvin Wagner
Optoelectronic Computing Systems Center, Dept of ECE, box 425
University of Colorado, Boulder CO 80309-0425 USA
1-303-492-4661 kelvin@boulder.colorado.edu

Abstract

Fault tolerance can be incorporated in digital optical computers using a distributed redundancy technique called quadding, at an expense of quadruplicating the hardware, and fan-in and fan-out increasing to 4.

A key requirement in a practical digital optical computing scheme is the ability to recover from transient and permanent device faults and signal errors. This is especially critical when using emerging technologies such as optical switch arrays which are bound to have large numbers of device errors, and for optically interconnected systems which may suffer transient errors such as that due to dust particles floating through the optical beams. Redundancy is the most common technique utilized to endow a system with a limited degree of fault tolerance and increase the system reliability beyond that given by the product of the probabilities of correct operation of the components. However, in many redundant systems, a voter is required to resolve conflicts between the redundant components, but a fault in the voter still produces erroneous outputs. Although multiply redundant voters can be incorporated, a mechanism must be included that eliminates and replaces faulty elements from the circuits, or else errors can propagate. These techniques can become quite complex and may be inappropriate for optical implementation. Another approach is to distribute the voter throughout the circuit using the technique of quadded logic.[1-3] In quadded logic, 4 copies of the circuit are produced, then interconnected in a permuted fashion that allows isolated errors within a quadded block of elements to be detected and corrected within the next few layers. This is often considered an expensive approach to fault tolerance, since it multiplies the hardware by a factor of 4, and doubles the fan-out and fan-in of the elements. However, the optical implementation of quadded logic in a regularly interconnected system has some attractive features as shown in Figures 1 and 2. The depth of the circuit is not increased, just the width, so no additional delay or speed penalties are imposed. The interconnections between the quadded circuits are reasonably regular and appear to be amenable to optical implementations. These interconnection topologies are reminiscent of optical crossovers, and might be implemented with a similar technique. One possibility is to interleave the original circuits on rows separated by 4, and interleave the quadded duplicates at the intervening positions. The same basic architecture of shuffles or crossovers within the rows can be performed, and holographic interconnections within a quadded set of 4 rows might not overly increase the system complexity.

The operation of a quadded system based on a regular logical structure is illustrated in Figure 1. Device errors are indicated as signals in boldface type. For example $x = 1$ should be represented at the input with four 1's, however the last input is in error. The quadded interconnects in the first layer mix the signals from the first two and last two redundant inputs, represented as (12,34). This corrects the NOR gate subcritical error $1 \rightarrow 0$ in one step. On the other hand, the input error for $\bar{x}$ is a critical $0 \rightarrow 1$ NOR error, which is converted by the first layer quadded interconnect to two subcritical errors, shown in italics. As long as the next layer of quadded wiring mixes the correct and erroneous signals, these subcritical errors will be corrected in the next layer, and this is accomplished by mixing the odd redundant signals and the even redundant signals, represented as (13,24). Similarly, errors introduced throughout the logical structure can be corrected, unless they appear too close to another error. As a result error free outputs are produced.

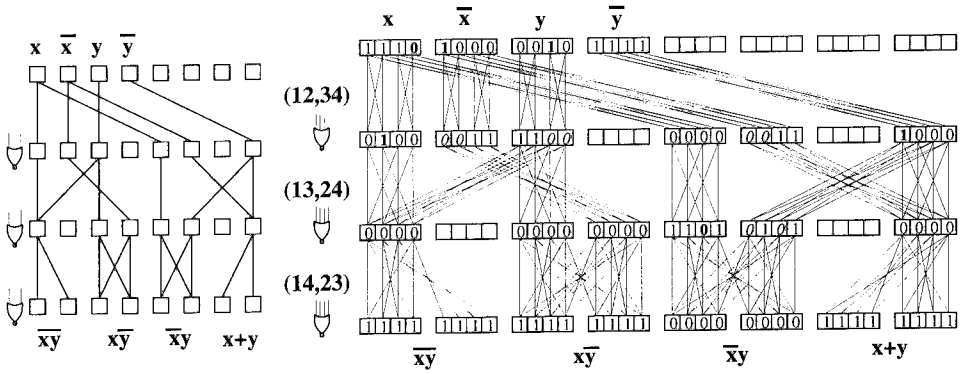


Figure 1: Error correction operation of a quadded regularly interconnected system.

The regular, layered structure of the conventional digital optical computer interconnects is eminently suited for quadding without the complexities encountered in random logic. In addition, the quadded interconnects take advantage of the capabilities of the optical interconnects to accommodate complex wiring topologies. Device defects often occur in clusters and this can nullify the benefits of local quadding as in Figure 1 since two critical errors within a quad can be uncorrectable. One powerful feature of optical quadding is the ability to avoid cluster defects through the use of the long range capabilities of optical interconnections. As an example Figure 4 shows a two shuffle interconnection and its long range quadded equivalent. Clustered defects would only affect one of the four members of each quad, allowing successful error correction.

The isomorphism between the (12,34) and (14,23) quadding patterns with crossover interconnects should allow efficient implementation of quadding in crossover networks, and in these regular structures only two alternating quadded patterns are needed. Alternatively, by placing the redundant quadded logic devices in a sparse topology then a shift-invariant interconnection can be utilized. The example shown in Figure 3 shows that the (12,34) and (13,42) quadding patterns can be implemented with a fanout of 3, light efficiency of 66% and device packing densities of 66% and 50% respectively, while the quadding pattern (14,23) requires a fanout of 5, and achieves a light efficiency of only 40% and a packing density of only 44%. Quadding as 2 by 2 blocks allows 2-D shift invariant interconnections with 44-50% packing density. For these shift invariant topologies the (12,34) and (13,42) quadding patterns should be alternated. A space invariant implementation of a programmable two shuffle with quadding for error correction implemented with a holographic interconnection is shown schematically in Figure 5. The two shuffle is programmed by splitting the two parts of the shuffle vertically where they can be blocked or passed under the control of a mask or SLM. These two halves are then combined, for example using walkoff in an off-axis anisotropic plate. A 2:1 anamorphic magnification is required in x to replicate the input NOR gate topology, which can be accomplished with either cylindrical optics or prism beam expanders. The output of the odd layers of (12,34) quadded two-shuffle logic must be quadded with a different pattern in the intervening even layers, and these quadding patterns must alternate. This can be accomplished with a different shift invariant hologram that implements the programmable two shuffle and the (13,24) quadding interconnect in the even layers. This output can be fed back to the original input with a 1 row vertical shift to implement a many layered fault-tolerant logical system with programmable space-variant operations, yet utilizing space invariant interconnections.

The error performance of a simple multilayered regularly-interconnected NOR gate based logical system is compared with its quadded equivalent in Figure 6, as a function of input probability of error and additive detector noise at the threshold logic NOR gate inputs. The original circuit has an output error probability that increases linearly with input error probability

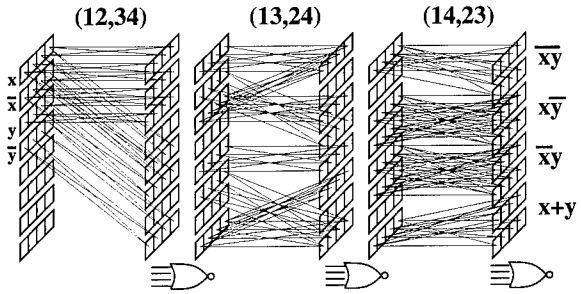


Figure 2: One possible 3-D topology of quadded regular interconnections.

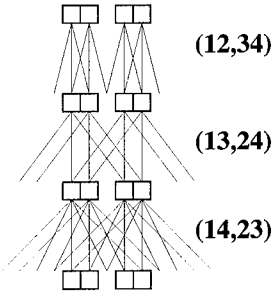


Figure 3: Sparse device layout for shift invariant quadding.

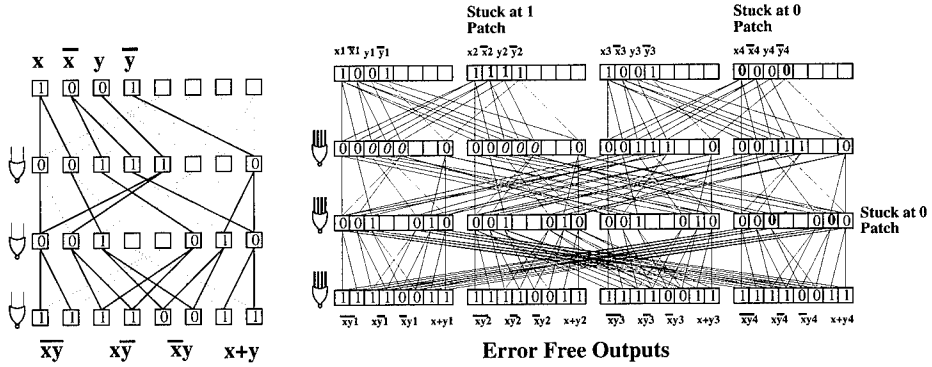


Figure 4: Two-shuffle interconnect and quadded equivalent using long range interconnects for quadding to avoid clustered defects.

and detector noise worsens the situation. This indicates the lack of error correction in the original circuit. The quadded version of this circuit demonstrates an output error probability that increases slower and quadratically with input error probability, and appears to have a large region of completely error free performance as desired. This data is for the 4 level circuit shown in Figures 1 and 4, and deeper circuits should have even more tolerance to errors. For low probability of device errors, the general tendency is for the system to achieve an error performance approaching the square of the device error performance, so a system of 10^5 devices with individual probability of error of 10^{-10} should lead to a system probability of error of 10^{-15}.

This approach to redundant fault tolerant digital optical computing may allow the utilization of devices with increased probabilities of failure without an unacceptably large system reliability penalty. Such an approach may be required in order to make practical and reliable digital optical computers out of simple arrays of unreliable switching elements.

The author acknowledges support of the NSF young investigator program ECS 9258088.

References

1. J. G. Tyron, Quadded Logic, in R. H. Wilcox and W. C. Mann, Eds., *Redundancy Techniques for Computing Systems*, Spartan Books 1962.
2. W. H. Pierce, *Failure-Tolerant Computer Design*, ch 5 Academic 1965.
3. Z. Kohavi, *Switching and Finite Automata Theory*, ch 8.6, Mcgraw-Hill 1970.

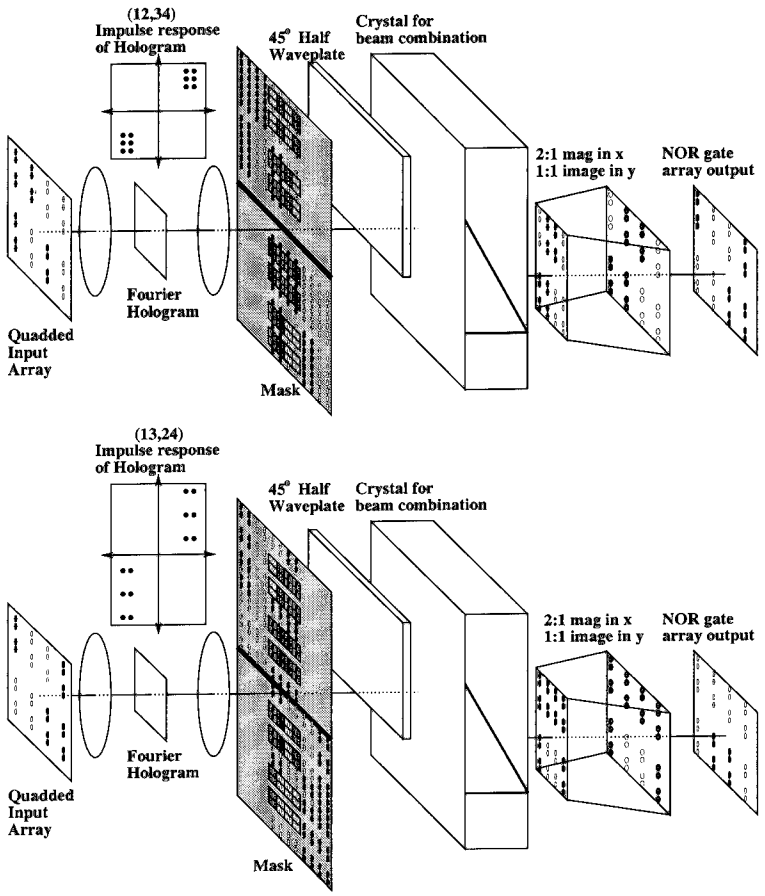


Figure 5: Alternating layers of a fully shift-invariant holographically-interconnected implementation of a programmable two-shuffle with quadding for distributed fault tolerance and error correction.

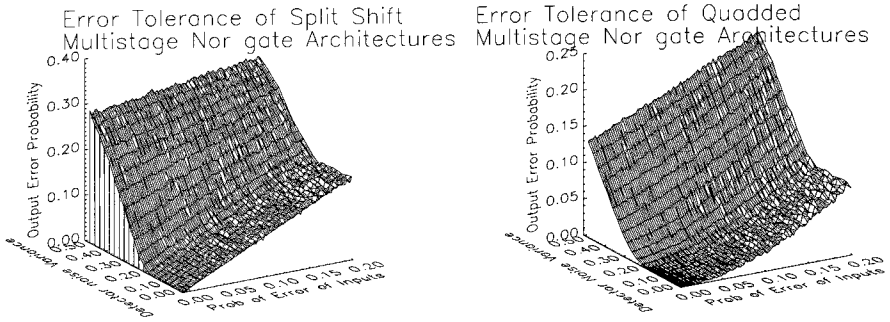


Figure 6: Error probabilities for original circuit of Figure 1 and quadded version as input probability of error and detector noise are varied.

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

On Parallel Algorithms for Optical Image Processors

Pierre Chavel, Philippe Lalanne

Institut d'Optique (CNRS)
B.P. 147, 91403 Orsay cedex, France, tel. (1)69 41 68 41

Abstract Fine grain parallel optoelectronic processors for dedicated tasks may deserve further consideration. The main case is machines for "real time" vision ; stochastic algorithms working on images considered as Markov fields are one possible approach to tackle real images.

1. Introduction

The processing of real images is one area where optical computing may come up with competitive solutions. The basic reason, as usual, is the number of interconnects that are needed before any sensible processing task on an image is completed ; the number of interconnects required is particularly large in parallel implementations, which is where free space optics is attractive. The word interconnect should be understood in a broad sense, including data input and output and the provision of various control signals to the processing elements as well as exchange of information among the processors themselves. We are interested in "optical scale parallelism", where the processor consists in a number of processing elements (PE's) equal to the number of pixels in the image, typically at least 10^4 : this puts the heaviest weight on interconnects but alleviates considerably programme and data transfer control. For easy image format input, all PE's should fit together on a chip a few square centimetres on a side. But technology will allow to integrate only fairly weak PE's on such a small area. We are then faced with a double challenge : devise optoelectronic architectures that make the best use of optical interconnects to maximise the computing power and find algorithms that can use it for meaningful image processing tasks. Starting with the well known architectures of optical correlators (i.e. convolvers) and optical cellular automata, we shall devote the major part of this work to the application of optical computing to optimisation problems. When approached using parallel versions of the class of stochastic algorithms known as simulated annealing, these could in some reasonable future benefit from a suitable combination of optical and microelectronic functions.

2. From optical convolutions to optical cellular automata

The simplest and most famous application of the above concept is the optical convolution, where the PE reduce to photodetectors and where weighted optical interconnects that define the impulse response do all the processing. The main application is pattern recognition. The double challenge then takes on the following form : can convolution be helpful in real pattern recognition problems and if yes, can optics implement such convolutions ? Progress on filter reprogrammability, adaptivity to the input signal and invariances has been fast in

the last few years [1] and will be reported in several papers in this conferences. Also, nice implementations of rugged and compact optical convolvers have been published [2].

Convolution nevertheless shows only limited generality and it is important to seek broader classes of our image processors. The next simple case is cellular automata [3], that have been investigated in some detail for some years under various names, including symbolic substitution and mathematical morphology [4]. Their operation cycle consists in the combination of one convolution and one point nonlinearity. The main role of optics here is to implement the convolution part, while optoelectronic or nonlinear optical devices located at every pixel respond nonlinearly to its result.

The invited communication by Casasent in this conference develops mathematical morphology applications for a number of image processing tasks.

3. Optimisation problems in image processing

One further step is to introduce optimisation problems on images into the realm of optical computing. In an optimisation problem, an energy function $E(\underline{x})$ is introduced as a measure of the departure of an image $\underline{x}$ from an ideal goal to be reached by the processing. The definition of function E incorporates all relevant knowledge, i.e. the input data but also, for example, the sources of degradation to be removed, a priori information on the class of object, and the features of interest. The processing minimises the energy with respect to $\underline{x}$. Typical applications include edge, texture and motion detection, as well as higher level tasks such as pattern classification [5].

Previous algorithmic work, notably by Geman et al. [6], has demonstrated energy functions that can detect, for example, texture, edge or motion in fairly difficult situations.

However, the computational load is usually extremely heavy because a small change in the image can generate a large change in the energy - the "energy landscape" is said to be wild - so that secondary minima will prevent deterministic descent algorithms from reaching the desired minimum or even an acceptable suboptimal solution. With most energy functions, the problem is non-polynomial in time, i.e. the optimal image $\underline{x}_o$ can be found only through exhaustive search in the space of all possible images, whose size increases exponentially with the number of pixels. As a consequence, it is impossible in practice to find the absolute minimum.

But the situation is even worse than that : efficient suboptimal procedures are themselves hardly practicable. Let us take the example of simulated annealing, that was advocated by Geman in the work cited above and will be developed in the next sections. Finding a good suboptimal solution will typically require to loop through a procedure of energy updating a quite large number of times, typically of the order of a few million times the number of pixels. This is still impractical unless some way can be found to implement them in parallel : as we shall illustrate now, their parallel implementation is where, in our opinion, optics may have a new role to play. In conclusion to this discussion, among the powerful algorithms that have been found to progress in the processing of real images, one subset may be open to "optical scale parallelism" and this is what we are investigating.

4. Simulated annealing : one example

In the following, for the sake of specificity, we shall discuss one particular example of an optimisation technique. While it may not be completely general, it has the advantage of simplicity, without being trivial, and it allows to introduce all the relations that we see

between parallel stochastic algorithms and optical computing. This example addresses the so-called "weak string problem" [7], which is the one-dimensional version of image restoration preserving discontinuities. In this problem, starting from the observed, real valued, noisy image $\underline{x}$ as input data, two images are estimated : a real-valued smoothed image $\underline{\hat{x}}$ and a binary border image $\underline{\hat{b}}$, equal to 1 where there is a border and to 0 everywhere else. Figure 1 shows the spatial structure of the two estimated images.

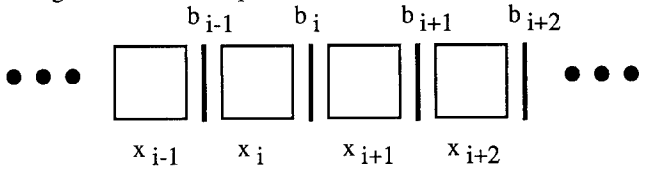


Figure 1. Data structure for the "weak string problem".

The energy is heuristically defined as :

$$E(\underline{\hat{x}},\underline{\hat{b}},\underline{x}) = \sum_{\text{all pixels } i} \left[(\hat{x}_i - x_i)^2 + \lambda^2 (\hat{x}_i - \hat{x}_{i+1})^2 (1 - \hat{b}_i) + \beta \hat{b}_i \right] \quad (1)$$

The first term enforces similarity between the estimated image and the input data, the second term is a smoothness constraint applied to adjacent pixels that have no border pixel between each other, and the third term is a border penalty to avoid uncontrolled growth of the borders. λ and β are adjustable parameters.

In this case, simulated annealing consists in iterating the following steps :

- 1 - select one border pixel i for "updating",
- 2 - evaluate the energy change δE when $\hat{b}_i$ is changed from 0 to 1,
- 3 - set $\hat{b}_i = 1$ with probability such as

$$p(\delta E) = \frac{1}{1 + \exp\left(\frac{\delta E}{T}\right)} \quad (2)$$

where T is a parameter, called temperature by analogy with statistical physics,

- iterate the updating (steps 1-3), slowly decreasing T according to some suitable "annealing schedule" (see the reference by Geman cited above).

Notes :

- the fact that this procedure is a good heuristic procedure for estimation problems is documented in the literature ; we shall be content here with the comment that the initial high temperature allow to accept basically any energy change with a significant probability and therefore avoid being trapped in local minima of the energy landscape, while the final low temperature are useful for finding local minima in a neighbourhood where the energy is already quite low ;
- in other cases, the change of $\hat{x}_i$ to $\hat{x}_i + \delta x$ should also be made in a stochastic mode but this is not so useful here because the energy function is <u>everywhere concave</u> in the variable $\underline{\hat{x}}$; it is therefore relatively simple to find its minimum ; for a more complete description of this intricate technical point, see reference 8.

As already mentioned, we are interested in the parallel implementation of simulated annealing. In the present case, the energy change related to one border pixel i depends on $\hat{b}_i$, but not on the $\hat{b}_j$, $j \neq i$. Parallel updating of all border pixels is therefore possible. More generally, parallelism of a degree linear in the number of pixels is legal if the energy change at one point depends only of a finite neighbourhood of this point. Asynchronous operation is an appealing but more complex issue that we shall not approach here.

5. Optical computing solutions for parallel simulated annealing

In this context, optical computing can provide three functions : convolution for the calculation of δE, production of the random numbers requested to implement the stochastic decision part of simulated annealing, and optoelectronic thresholding.

5.1 - Optical convolution :

Firstly, the calculation of energy variations δE often implies convolutions, and we are back to the first section. Specifically, let us give two examples :

• from equation 1, it is straightforward to derive

$$\delta E = -\left[\lambda^2(\hat{x}_i - \hat{x}_{i+1})^2 - \beta\right], \qquad (3)$$

which may be implemented in part using a coherent optical convolution on the $\underline{\hat{x}}$ image followed by the usual quadratic detection of optics. In this simple example, the convolution reduces to a straightforward difference between the grey value of two nearest neighbour pixels. Quite likely, it is much easier to implement in hardwired microelectronics than using optical convolution. The same is not true, however, for more complex energy functions that involve larger neighbourhoods — texture detection is one example.

• If the $\underline{x}$ dependency of the energy function is quadratic but not everywhere concave, then, as already mentioned, energy minimisation with respect to $\underline{x}$ is a difficult task and simulated annealing over $\underline{x}$ may be useful. Differentiation of a quadratic form will obviously always lead to a linear form, i.e. a matrix product, that in turns simplifies to a convolution if the energy model assumes stationarity of the image properties.

5.2. Parallel generation of random numbers :

The parallel generation of a large amount of random numbers of a good statistical quality is a problem. The requirement here is to provide independent random numbers with the appropriate statistics to all processing element, i.e. every pixel, at every energy update operation, which means around 10^{10} random numbers per second on a microelectronic chip. We have suggested to use a physical random number generator for this purpose and investigated the use of laser speckle projected onto an array of photodiodes. The electronic part of our parallel processor therefore consists of a "smart pixels" chip with at least one photodetector per image pixel being processed. Of course, with suitable sequencing, the same photodetector may be used for image input, for the input of δE, and for the speckle input.

Specifically, we have shown that speckle statistics can be easily moulded into exactly the required form of probability law, and that the simulated temperature can be controlled directly by the average speckle brightness, i.e. the laser power [9]. Let us just summarise the basic principle in a few lines.

A fully developed speckle integrated over the area of a photodiode obeys known statistics that depend on the number of speckle grain over the detector area [10]. We have experimentally demonstrated the possibility of producing 10^{10} random numbers per second by projecting speckle from a suitably moving diffuser onto a 1 cm^2 silicon photodetector array 11. Figure 2 illustrates how to obtain the probability law required by equation (2) : two speckle photodetectors are used instead of one. An analogue adder combines the signal from the first photodetector with the energy difference δE. The result is sent to the positive input of a thresholding gate, while its negative input receives the second speckle photodetector

signal. Analysis shows that, with suitable speckle parameters, the resulting probability that the positive input exceeds the negative input quite well approximates equation (2). Temperature is emulated by the average speckle intensity.

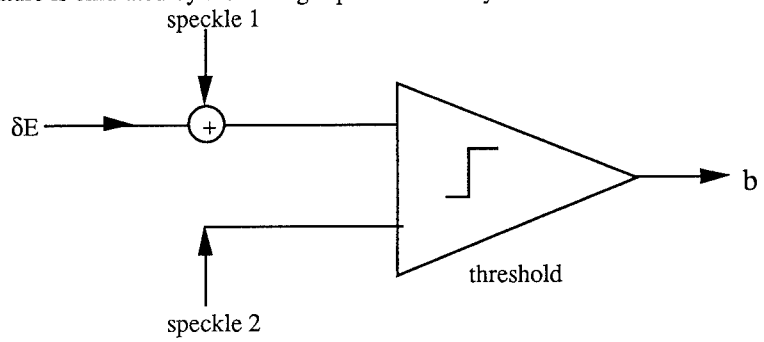


Figure 2. Generating the probability law of equation 2 with two speckle samples.

5.3. Optoelectronic thresholding :

Finally, novel optoelectronic or nonlinear optical arrays such as SEEDs or pnpn photothyristors may be used to make the required decision. For example, as shown in the contribution by Prémont et al. in this conference [12], the role of the thresholding gate of figure 2 can be played by a differential pair of optical photothyristors. The output, i.e. border pixel b_i, is then available in the form of an optical signal for some further processing step or for the output of results.

More generally, integrated circuits with a complexity depending on the particular energy function will be required in association with optics ; globally, only the combination of electronic and optical functions can open the way to compact, massively parallel integration of image processors for such algorithms.

Conclusion

The favourite operation of analogue optical processing, convolution, may provide a solution for a certain number of image processing problems. But we believe that it could well be combined with other readily available optical functions and with integrated circuit microtechnology into video-real-time systems for a significantly wider class of vision problems, in particular optimisation problems that can be handled by parallel stochastic algorithms.

We acknowledge the contributions of E. Belhaire, F. Devos, P. Garda, L. Garnero, G. Prémont D. Prévost and J.C. Rodier to the research leading to this work.

References

[1] See for example Javidi B. and Réfrégier P., editors, 1994,*Optical Pattern Recognition, Euro-American Workshop* (Bellingham, Washington : SPIE Press), or Horner J.L., and Javidi B., editors, *Special section on Pattern Recognition, Opt. Engin.* **33**, 1751-1862.

[2] Rajbenbach H. et al., 1992, "Compact photorefractive correlator for robotic applications," *Appl. Opt.* **31** 5666-5674.

[3] Taboury J. et al., 1988 and 1989, "Optical cellular processor architectures," part I, *Appl. Opt.* **27** 1643-1650, part II, *Appl. Opt.* **28** 3138-3147. Chavel P. and Taboury J., 1990, "Binary optical cellular automata : concepts and architectures," *SPIE Proc.* **CR35** 245-265.

[4] Huang A., 1983, "Parallel algorithms for optical digital computers", *IEEE 10th International Optical Computing Conference*, 13-17. Huang K.S., Jenkins B.K., Sawchuk A.A., 1989, "Binary image algebra and digital optical cellular image processor design," *Computer Vis. Graph. Image Proc.* **45** 295-345.

[5] Bouthemy P. and Lalande P., 1993, "Recovery of moving object masks in an image sequence using local spatiotemporal contextual information," *Opt. Engin.* **32** 1205-1212.

[6] Geman S and Geman D, 1984, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", *IEEE Trans Pattern Anal. Mach. Intel.* **PAMI6** 721-741.

[7] A. Blake and A. Zisserman, 1987, *Visual Reconstruction* (Cambridge, Massachusetts : the MIT Press).

[8] Prévost D, Lalanne Ph, Garnero L and Chavel P, 1994, "Quasi-static algorithm for image restoration preserving discontinuities", *SPIE Proc.*, in press.

[9] Lalanne Ph., "Progress towards the optoelectronic implementation of stochastic artificial retinae," submitted for publication.

[10] Goodman J.W., 1975, "Statistical properties of laser speckle patterns," in *Laser speckle and related phenomena*, J.C. Dainty, editor (Heidelberg : Springer), 9-74.

[11] Lalanne Ph et al., 1993, "Optoelectronic devices for Boltzmann machines and simulated annealing," *Opt. Engin.* **32** 1904-1914.

[12] Prémont G. et al, 1994, "Optical thyristor based stochastic elementary processor," *in this volume.*

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Duality between algorithms and optical implementations : actual examples

J. P. Heer, P. Pellat-Finet

ENST de Bretagne, BP 832, 29285 Brest, France

Abstract. Considerations on number representations are introduced to make use of the advantage of optics in computer architectures. Specific optical systems are proposed for CORDIC algorithms and for residue arithmetic with applications to discrete transforms.

1. Introduction

The recent developments of bistable and fast switching optoelectronic components (combination of laser diodes and photodiodes, S-SEED) do not necessarily imply that optics could be an alternative technology for high speed parallel computers. There is indeed a need for a duality between algorithms and optical implementations and this duality is not generally considered by scientists working in optical computing.

As an example of this duality, we will focus on some number representations and their corresponding optical processing architectures. We will only distinguish two well known number representations: the position number and the residue number representations. More precisely we will show that the choice of appropriate number representations and optical architectures leads to actual implementations of more general algorithms for fast parallel computing which should reinforce fast component capabilities.

2. An optical implementation for CORDIC algorithms

2.1. The modified signed digit representation and its optical implementation

The binary number representation is a position representation, widely used in electronic computers. Nevertheless it appears to be imperfectly suited to optical implementations because of its sequential feature. An alternative is the signed digit number representation, introduced by Avizienis [1]: it allows carry free parallel processing with a number of steps independent of the number of digits. Attention has been paid to the modified signed digit (MSD) representation [2] and an optoelectronic MSD adder has been built [3]. We will see in the next section how this adder can be inserted in a larger architecture in order to perform CORDIC algorithms. We only emphasize at this stage that the adder provides an example of the necessary adaptation (duality) between an algorithm and an optical implementation, taking advantage of parallel processing capabilities of optics.

2.2. Application to CORDIC-like algorithms

CORDIC (COordinate Rotations on a DIgital Computer) algorithms were introduced by Volder [4] to compute trigonometric functions by performing only additions and shifts. According to Walther [5, 6] an extended algorithm is as follows. We introduce a decreasing sequence (e_n) (n is a positive integer) and a number m which depend on the function to be computed (examples will be given below). Consider the sequences (x_n), (y_n) and (z_n) such that :

$$x_{n+1} = x_n - m d_n y_n 2^{-n} \quad ; \quad y_{n+1} = y_n + d_n x_n 2^{-n} \quad ; \quad z_{n+1} = z_n - d_n e_n \qquad (1)$$

where (d_n) is a sequence depending on sequences (y_n) and (z_n) and on the function to be computed.

It can be shown [5,6] that for $e_n = \tan^{-1} 2^{-n}$, $m = 1$, $x_0 = [\prod_0^\infty (1 + 2^{-2n})]^{-1/2}$, $y_0 = 0$, and $d_n = \text{sign} z_n$, the sequence (x_n) tends to $\cos z_0$ and the sequence (y_n) to $\sin z_0$ for infinite n. For $e_n = 2^{-n}$, $m = 0$, $d_n = \text{sign} z_n$ and $y_0 = 0$, the limit of y_n is $x_0 z_0$. For $e_n = 2^{-n}$, $m = 0$, $d_n = -\text{sign} z_n$ and $z_0 = 0$, the limit of z_n is y_0/x_0. Other initial values lead to compute other functions such as $\tan^{-1}$ and hyperbolic functions. Similar (CORDIC–like) algorithms can be used to compute the exponential function [7]. Hence the algorithm allows the computation of classical elementary functions, multiplication and division.

In eq.(1) we perform multiplications of x_n and y_n by 2^{-n} (because $|m d_n| = |d_n| = 1$), that is, if the binary representation is used, shifts on n digits. Moreover, it can be shown that CORDIC–like algorithms are compatible with the MSD representation of numbers [7]. This means that additions in eq.(1) can be achieved carry free with a parallel processing. In this way the above mentionned adder can be used and only needs to be completed with a digit shifter. An optical solution to the digit shifter has been proposed and built and a general optical architecture for CORDIC–like algorithms can be deduced [7]. The adaptation of an optical implementation to CORDIC algorithms is an example of the necessary duality between algorithms and implementations.

3. An optical implementation for discrete transforms

3.1. Optical processor for residue arithmetic

We will describe a residue arithmetic optical processor that we recently built [7]. Residue arithmetic will be assumed to be known [8,9]: it reduces the operation complexity by dividing the number representation into smaller independent integers, allowing parallel processing.

Although the actual processor deals with numbers modulo 15, we will explain its principle with 5 as modulus. The processor whose design is given by fig.1 has two entries which represent the numbers to be processed. An entry is a luminous source, called an *input source*, and made up of five LED rods. Each rod is assigned to a value in the set $\{0, 1, 2, 3, 4\}$ and only one rod among the five is activated according to the value to be represented. One input source is made up of horizontal rods and the other of vertical ones. The shape of the area spanned by the five rods is a square.

A beam splitter mixes light beams coming from the input sources, the plane of which are symmetrical to each other with respect to the beam splitter separating area.

A spherical lens focuses the input source images on an output plane in which there is a square array of 25 detectors. Among the 25 detectors, 16 do not receive any light, 8 receive light from one rod and only one receives light from two rods which belong to the two input sources. The latter detector indicates on which input values the operation has to be performed. In order to distinguish the response of this detector among the others, a thresholding has to be accomplished and since it is a non-linear processing, this is electronically achieved. Since only three illumination levels have to be distinguished the system is highly insensitive to noise.

Behind the detector array, there is another light source, called *output source*, which is designed like an input source. Each of the 25 detectors is linked to a rod of an output source, the choice of the rod depending on the operation to be performed. In practice the electric current coming from a detector has to be amplified to be high enough to control a LED rod. The thresholding and amplification processing and link between detectors and the output source are included in what is denoted by *electronic link* on fig.1 (where only one output is designed). An example is given on fig.1 and deals with 1+2=3. The inversions due to the lens and the beam splitter have not been taken into account. Shaded areas indicate activated LEDs and detectors. Only the activated electronic link has been designed.

Such a system does not carry out any calculation but only finds the result in a memory: the input source and the beam splitter help finding the address and the electronic link assigns the result value. It should be clear that the system would run at the same speed whatever the modulus. The built processor works at a 20 MHZ clock rate. A more complete system should be made up of various processors working with different moduli. Notice that two outputs can be obtained with a unique beam splitter and then two operations can be simultaneously performed.

3.2. Quadratic residue arithmetic

Quadratic residue arithmetic is useful to handle complex numbers [8]. The equation $x^2 \equiv -1 \bmod m$ admits a solution, say i, if m is equal to 2 or is a prime number of the form $4n+1$ where n is a positive integer. Let $x = x_r + ix_i$ be a complex number where x_r and x_i are integers. Let $x_r \equiv a \bmod m$ and $x_i \equiv b \bmod m$. Then x is represented by $(a+ib, a-ib)$, that is, by two integers modulo m on which algebraic operations are performed separately. Notice that a and b can easily be deduced from $a+ib$ and $a-ib$.

3.3. Optical architecture for discrete transforms

We explain now how the optical residue processor can be included in an architecture for discrete transforms. We will choose the fast Fourier transform with the usual butterfly architecture on N points. We consider a residue system with M moduli m_j $(1 \leq j \leq M)$ and let $[x]_j$ denote the residue of x modulo m_j. If $w = \exp(-2i\pi/N)$, the discrete Fourier transform of $x(k)$ $(k = 1, \ldots, N)$ is $\hat{x}(n) = \sum_k x(k) w^{kn}$ and split in M transforms modulo m_j defined by [10]:

$$[\hat{x}(n)]_j = \sum_k \left[[x(k)]_j [w^{kn}]_j \right]_j \tag{2}$$

Since w is a complex number, eq.(2) must be written according to the quadratic residue system. The optical implementation of an elementary cell of the butterfly structure is

given on fig.2. Only the beam splitter is designed: it represents indeed an optical processor as in paragraph 3.1. Additions and multiplications by w^{kn} are performed according to quadratic residue arithmetic. A general butterfly architecture can be deduced [7].

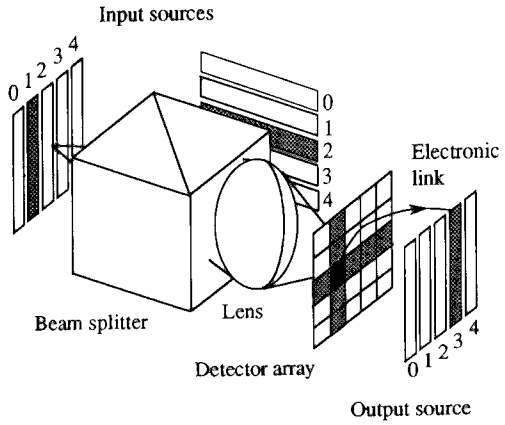


Fig.1

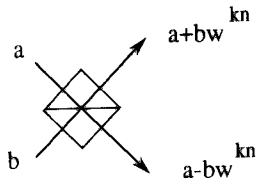


Fig.2

References

[1] Avizienis A 1961 *IRE Trans. on Elec. Computers* **EC-10** 389–400

[2] Drake B L, Bocker R P, Lasher M E, Patterson R H and Miceli W J 1986 *Opt. Eng.* **25** 38–43

[3] Pellat-Finet P and Armand-Guerineau I 1993 *Opt. Comm.* **98** 250–256

[4] Volder J E 1959 *IRE Trans. on Elec. Computers* **EC-8** 330–334

[5] Walther J S 1971 *Sprint joint Computer Conference* **38** 379–385

[6] Muller J M 1989 *Arithmétique des ordinateurs* (Paris: Masson)

[7] Heer J P 1994 *De la représentation des nombres aux processeurs optiques de calcul numérique* (Thesis Université de Paris XI Orsay)

[8] Niven I and Zuckerman H S 1966 *An Introduction to the Theory of Numbers* (New York: John Wiley & Sons)

[9] Huang A Tsunoda Y., Goodman J. W. and Ishihara S. 1979 *Appl. Opt.* **18** 149–162

[10] Truong T K, Chang J J, Hsu I S, Pei D Y and Reed I S 1986 *IEEE Trans. on Computers* **C-35** 1008–1012

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Simulation and benchmarking of a new algorithm for the optical cellular logic image processor

M.P.Y. Desmulliez[1], B.R. Gillies, J.F. Snowdon, B.S. Wherrett

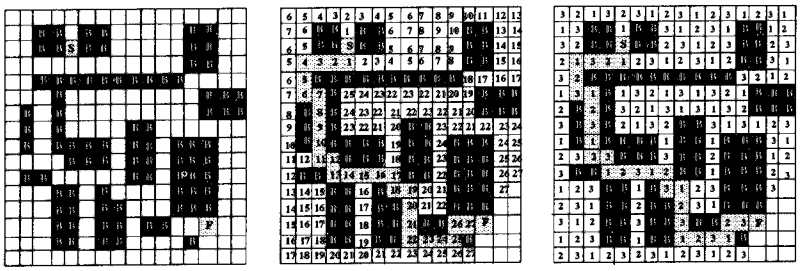
Department of Physics, Heriot-Watt University, Riccarton EH14 4AS, Scotland, U.K.

Abstract. The Optical Cellular Logic Image Processor, O-CLIP, was developed to demonstrate the functional capabilities of optically interconnected logic arrays. The O-CLIP can in principle implement any binary image algebra tasks. As an example, we present here, a new shortest path hunt algorithm that precisely maps onto the architecture of the constructed O-CLIP. The main phases of the routine are explained. The efficiency of the algorithm is compared to its serial implementation.

1. Introduction.

In 1961, Lee developed a maze algorithm capable of finding the shortest path(s) between two given points within a rectangular grid filled with obstacles [1]. For this algorithm, the nearest neighbours of the starting point (marked S in figure 1a) are labelled. The labelling is repeated in increasing order to the next nearest neighbours until the finish point (marked F) is reached (figure 1b). The shortest path is then found by tracing back from the finish point the decreasing sequence of the labels generated by the expansion process. This algorithm, also called the shortest path hunt algorithm, was later derived for single instruction, multiple data (SIMD) computers within the image logic algebra (ILA) [2, 3]. The optical implementation necessitates however, at certain stages, a fan-out corresponding to the number of elements of the grid. The new version presented here utilizes a four-nearest neighbours interconnect. All the information as to the labelling of the elements can be held on just three labels planes. Three different labels are in fact the minimum number needed to distinguish the character of a non-monotonic number sequence (figure 1c).

[1] E-mail : marc@phy.hw.ac.uk



(1a) (1b) (1c)

Figure 1a-1c. Lee and modified Lee routines.

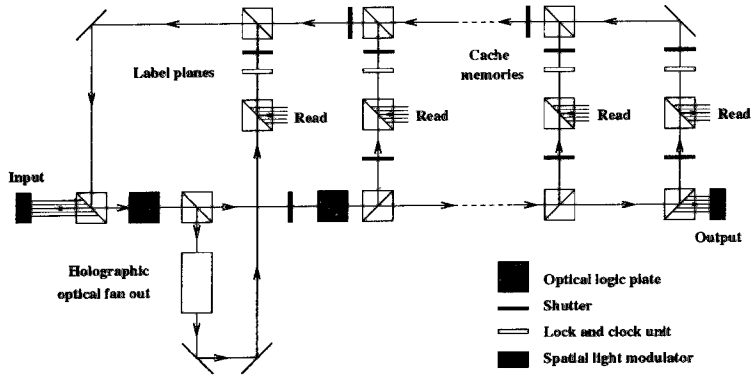


Figure 2. Schematic of the O-CLIP layout.

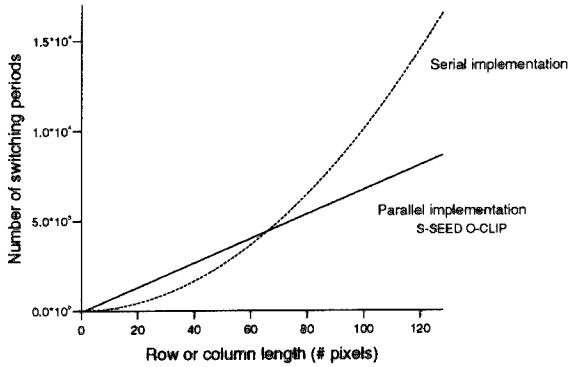


Figure 3. Parallel and serial implementation of the Lee routine. Benchmark.

2. The O-CLIP implementation.

The algorithm is to be utilized for a four nearest-neighbours interconnect system such as in the O-CLIP [4]. The different processing planes are Symmetric Self-Electrooptic-Effect Device (S-SEED) arrays. The devices can be programmed as NAND/NOR logic gates [5]. They can also be used as inverting buffers. A schematic of a possible layout is shown in figure 2. Besides the three different S-SEED arrays which store the label sequence, three cache memories are used to perform the intermediate processing steps. Label planes and cache memories form individual optical loops within the O-CLIP architecture. The loops ensure the optical feedback of a previously processed array. This array is compared with data from the input spatial light modulator (SLM) or with the information from another loop. The result of this comparison is then fanned out, thresholded and sent to one of the label planes or the cache memories. This is performed by the fan-out and threshold units and the optical logic plates. The input information is provided by an SLM.

3. Explanation of the algorithm.

The algorithm has been run on a distributed array processor (DAP). This parallel (electronic) computer simulates the algorithm as if implemented in the S-SEED O-CLIP. The two main phases of the algorithm are given below : the expansion of the labelling process (section A) is carried out in a diamond-like shape and the tracing back of the shortest path(s) (section B) is performed up to the starting point. The expansion phase checks that the expanded wave does not intersect the obstacles and the other label planes. The tracing phase ensures that the shortest path(s) is (are) found by following the decreasing sequence starting from the finish point. The path(s) is (are) constructed, in real time, in one of the cache memories. The steps of the algorithm are presented below, along with the number of S-SEED switching periods required for each step.

Expansion phase. Section A.

Until expansion reaches finish point **Do**

FAN-OUT (cache 1) $\mapsto$ thresholder $\mapsto$ fan storage unit	(5 steps)
NAND (obstacles, fan storage) $\mapsto$ cache 3	(4 steps)
NOR (cache 3, label plane j-1) $\mapsto$ cache 1	(4 steps)
NOT (cache 1) $\mapsto$ cache 3	(3 steps)
NOR (cache 3, label plane j+1) $\mapsto$ label plane j	(4 steps)
increase j by 1	

End Do

Start with label plane containing finish point (assume label j)

Path back phase. Section B.

Until path reaches starting point **Do**

FAN-OUT (cache 1) $\mapsto$ thresholder $\mapsto$ fan storage unit	(5 steps)
NAND (label plane j-1, fan storage) $\mapsto$ cache 3	(4 steps)
NOT (cache 3) $\mapsto$ cache 1	(3 steps)

NOR (cache 2, cache 1) $\mapsto$ cache 3 (4 steps)
NOT (cache 3) $\mapsto$ cache 2 (3 steps)
decrease j by 1
End Do

Although conditional statements are explicitly written above, the algorithm allows the different loops to be run indefinitely. Alternatively, the rapid fill algorithm can be used to test the validity of the conditional statements [4].

4. Efficiency of the algorithm

The total number of switching periods to implement the Lee algorithm depends on the size of the problem and on the nature of the logical devices. The parallel implementation of the Lee routine is carried out with S-SEEDs arrays in this example [5]. The clocking procedure of this device is performed usually in three consecutive steps. The SEED is preset by an optical beam to program the logical operation. The signal beams write onto the devices which are subsequently read to propagate the information onto the neighbouring array. The worst case in terms of numbers of switching periods has been analysed for both the parallel and serial implementation. The number of switching periods for the whole algorithm is less than the sum of switching periods at each step because different operations can be time multiplexed in the various arrays of the O-CLIP. The Lee routine necessitates, as a consequence, only 68x(N-1) switching periods for an NxN array. This is to be compared to Nx(N+1) for the serial implementation (figure 3). The break-even occurs for a 66x66 array.

5. Conclusion.

A new algorithm for the S-SEED O-CLIP has been presented. A DAP simulation has been carried out and the main parts of the algorithm were explained with reference to the DAP simulation. The parallel implementation has been proved to be more efficient than its bit-serial implementation for array sizes larger than 66*66.

References

[1] C. Lee *IRE Trans. Electron. Comput.* **EC-10** 346 (1961).

[2] M. Fukui and K. Kitiyama *Appl. Opt.* **31** 581 (1992).

[3] M. Fukui and K. Kitayama *Appl. Opt.* **31** 4645 (1992).

[4] B.S. Wherrett and J.F. Snowdon *SPIE Optical Computing 88* **963** 15 (1988).

[5] A.L. Lentine, H.S. Hinton, D.A.B. Miller, J.E. Henry, J.E. Cunningham, L.M.F. Chirovsky *IEEE Journ. Quant. Electron.* **25** 1928 (1989).

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A Methodology for Optical Architectures Modeling

Mohsen Guizani

Computer Engineering Department
College of Computer Science and Engineering
King Fahd University of Petroleum and Minerals
Box # 1969, Dhahran 31261, Saudi Arabia.

Abstract. In this paper, we devise a methodology for optical component modeling. We present the general description of single-port and multi-port optical components. Then we use signal flow graphs to analyze optical architectures prior to the implementation phase. It illustrates the behavior of the optical data signal at the output of the optical system.

1. Introduction

Development of optical components have witnessed great interest in the last few decades [1]. These components are successfully used for optical computing. Experimental and practical results are available in the literature [2,3,4]. To our knowledge, only [5] discussed modeling of some optical components in a classical manner. Whereas [6] developed a methodology for a special purpose optical architecture.

The objective is to develop a new methodology that can be used to model widely used optical devices. The method should be easily extendible to large optical architectures that use these devices.

2. Methodology of the Component Modeling

We discuss first the general aspects of the modeling techniques. In practice there are three generic classes of optical network architectures. Those that use fiber-optic interconnects (fiber-optic networks), others that use waveguides (integrated-optic networks), and third that use free-space interconnects (free-space optic networks). The analysis of the methodology presented here follow the analysis of the microwave networks. the similarities stems from our concern with the traveling waves, in which the so-called *scattering parameters* are useful. These parameters constitute the *S-matrix* (scattering matrix). Similarly, with optical components, we can talk about the S-matrix. Except that when we have optical components we should introduce two single virtual ports instead of each physical port. Therefore, our optical scattering matrix is composed of elements of 2x2 matrices rather than regular numbers. The S-matrix of some common optical devices such as

the one-port, two-port components and directional couplers are available in the literature [5].

3. General Modeling Concepts

As mentioned earlier, virtual ports have to be introduced for the optical components in addition to the existing physical/real ports. This is necessary in order to account for the reflections or losses of the data signals traveling through the optical components. This might add to the complexity of the analysis but the use of signal flow graphs simplifies the results.

3.1. Single-Port Component

We can describe the single-port component with an equation relating the output to the input as:

$$A' = SA + C$$

where A' represents the output port (*sink*), A is the input port, S is a complex 2x2 matrix, and C is a complex two-element vector. The vector element C has a significant value only in active components (e.g. optical sources and optical amplifiers). It is normally neglected in passive elements. The S matrix usually represent the reflection of the optical component. Therefore, to represent an optical source which has no reflections, we should set S to zero. This gives

$$A' = C$$

In a signal flow graph, we can represent both of these equations as shown in figure 1 which represents the general configuration of the single-port connection. Whereas, the representation of the optical source is the same except that the connection of the input port disappears. Therefore, we can represent graphically a source without reflections with a branch of unity transmission by a single output node. To represent an optical detector graphically, we should realize that a detector is a passive single-port component. An optical detector with no reflection can be represented in a similar fashion as that in figure 1.

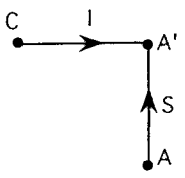


Figure 1. One-port component signal flow graph.

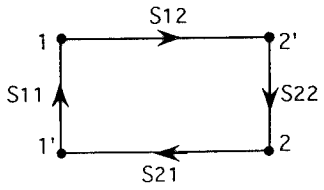


Figure 2. Two-port component.

3.2. Two-Port Component

In examining how to represent the two-port component, we follow the general description of the N-port component which is briefly described as follows. Each node, usually regarded as a generator, has to have a sink. So for node j, we introduce sink j' and for node k, we

introduce sink k'. Therefore, for a two-port component, we associate two nodes with each port. Every input node is connected with every output node by a directed branch. The branch values are the component S-matrix elements shown by the value $S's$ in figure 2.

4. Optical Architecture Modeling

In this section, we present a simple architecture, shown is figure 3, composed of one- and two-port components and use the above methodology to analyze the output attributes of the architecture. The architecture is composed of a source, mirror, a polarized beam splitter, and an interference filter. For simplicity, and with no loss of generality, we ignore the quarter wave plates since they do not affect the data signal except in changing its polarization. Assume also that a polarized beam splitter is an ideal two-port optical component.

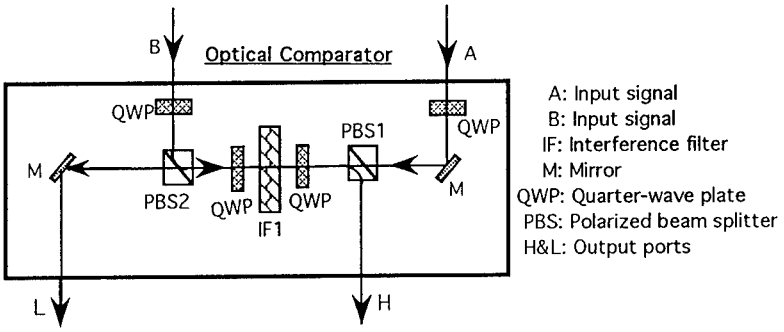


Figure 3. Optical architecture to be modeled.

The derivation of the signal flow graph corresponding to the optical architecture shown in figure 3 is described in figure 4. Note that for an ideal two-port optical component, there are no reflections. This implies that S_{11} and S_{22}, in figure 2, are both equal to zero. Furthermore, a null source can always be removed without affecting the graph, such as node 2' in figure 4 (b). Likewise, a sink, such as node 7, may be removed since its presence does not affect the value of any other node. Note that when a node is removed from a single flow graph, all the branches connected to it are removed as well. The final signal flow graph is thus given in figure 4 (c). We then use standard signal flow graph reduction to obtain the transfer function that gives the relationship between the input and output. The reduction result of figure 4 is found to be

$$I^2 S_{12} T_{67'} T_{3'6} (I - T_{33'} T_{66'} T_{3'6})^{-1} T_{23'}$$

In a practical case, the values of **T**'s of the above equation are known, then we can compute the information required at the output, such as power (intensity), phase, and others. Work is in progress to find the required mathematical model to solve for the attributes of the outputs. In dealing with matrices (such as S-matrix), a good mathematical package such as Mathematica or MathLab is required to handle matrix inversion especially for large optical systems.

5. Conclusion

We have developed a new methodology that can be used to model optical architectures based on widely used components. Extension of this methodology to large optical architectures is fairly simple. The objective is to provide a tool to the experimenter to deal with the architecture before its implementation. Work is in progress for the validation of this methodology.

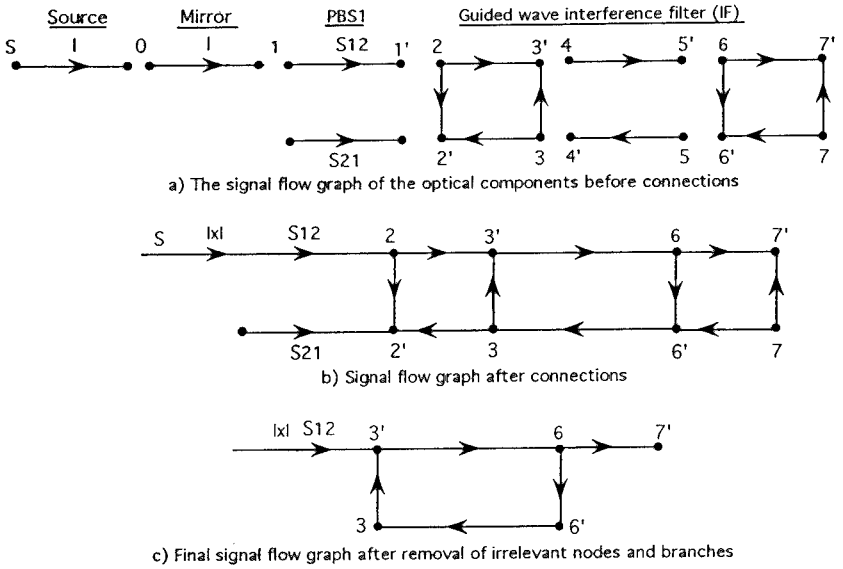


Figure 4. Reduction of the graph of the optical architecture of figure 3.

Acknowledgment

The author would like to acknowledge the support of King Abdulaziz City for Science and Technology under project number KACST AR-13-12. Special thanks to S. Ghanta for valuable discussions.

References

[1] McCormick B,Tooley A, Cloonan J, Brubaker L, Lentine L, Morrison L R, Hinterlong S J, Herron M J, Walker S L and Sasian J M 1992 *Applied Optics* **31** 5431-5446
[2] Guizani M 1993 *Applied Optics* **32** 1587-1599
[3] Heuring V P, Jordan H F and Pratt J P 1992 *Applied Optics* **31** 3213-3224
[4] McAulay A D 1991 *Optical Computer Architectures: The Application of Optical Concepts to Next Generation Computers* (John Wiley & Sons Inc., New York)
[5] Weissman Y 1992 *Optical Network Theory* (Artech House)
[6] Murdocca M 1990 *A Digital Design Methodology for Optical Computing* (The MIT Press).

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Direct twos complement parallel array multiplication: algorithm and optical implementation

Guoqiang Li, Liren Liu, Lan Shao and Zhijiang Wang

Information Optics Laboratory, Shanghai Institute of Optics and Fine Mechanics, Academia Sinica, P.O.Box 800-211, Shanghai 201800, P. R. China

Abstract. A direct 2's complement parallel array multiplication architecture is proposed. The algorithm has overcome the problems encountered in the conventional one. Using a two-stage-array, complex multiplication is achievable. Correspondingly, a hybrid optoelectronic system with free-space interconnection is suggested.

1. Introduction

The conventional twos complement algorithm[1-2] has the following defects: (1) The operands must be encoded with the same bits as needed for their product, so the range of the result must be known in advance; (2) The upper half bits of the product are discarded, hence this is a severe waste of the space bandwidth product(SBWP); (3) The intermediate mixed 2's complement result is unable to be weightedly summed. Although a modified version[3] requires no sign-extension bits and permits all processor channels to be used for the value encoding, its postprocessing needs diverse sign decisions and irregular additions. To solve these problems, we propose a direct 2's complement parallel array multiplication algorithm[4,5].

2. Modified direct twos complement parallel array multiplication algorithm

In 2's complement representation, any analog number x encoded by a N-bit string (x_{N-1}, x_{N-2}, ···, x_1, x_0) takes the value

$$x=-x_{N-1}\cdot 2^{N-1}+\sum_{i=0}^{N-2}x_i\cdot 2^i, \tag{1}$$

and the additive inverse of x equals

$$-x = -\bar{x}_{N-1}\cdot 2^{N-1} + \sum_{i=0}^{N-2}\bar{x}_i\cdot 2^i + 1,\ \bar{x}_i=1-x_i\ (i=0, 1, \cdots, N-1). \tag{2}$$

With the mixed 2's complement representation, addition of two numbers is

performed by summing the corresponding digits of both strings without carries, and subtraction is performed by summing the minuend and the additive inverse of the subtrahend.

Based on these knowledge, the direct parallel array multiplication algorithm can be derived. The product of two N-bit numbers u and v is given by

$$p = u \times v = \left(-u_{N-1} \cdot 2^{N-1} + \sum_{i=0}^{N-2} u_i \cdot 2^i \right) \left(-v_{N-1} \cdot 2^{N-1} + \sum_{i=0}^{N-2} v_i \cdot 2^i \right) \tag{3}$$

$$= (u_{N-1} v_{N-1} \cdot 2^{2N-2} + \sum_{i=0}^{N-2} \sum_{k=0}^{N-2} u_i v_k \cdot 2^{i+k}) - (\sum_{i=0}^{N-2} u_{N-1} v_i \cdot 2^{i+N-1} + \sum_{i=0}^{N-2} v_{N-1} u_i \cdot 2^{i+N-1}) \ .$$

Since

$$- \sum_{i=0}^{N-2} v_i \cdot 2^i = -1 \cdot 2^{N-1} + \sum_{i=0}^{N-2} \bar{v}_i \cdot 2^i + 1 \ , \tag{4}$$

and

$$- \sum_{i=0}^{N-2} u_i \cdot 2^i = -1 \cdot 2^{N-1} + \sum_{i=0}^{N-2} \bar{u}_i \cdot 2^i + 1 \ , \tag{5}$$

there exists

$$p = (u_{N-1} v_{N-1} - u_{N-1} - v_{N-1}) \cdot 2^{2N-2} + \sum_{i=0}^{N-2} \sum_{k=0}^{N-2} u_i v_k \cdot 2^{i+k} + \sum_{i=0}^{N-2} (u_{N-1} \bar{v}_i + v_{N-1} \bar{u}_i) \cdot 2^{i+N-1} \tag{6}$$

$$+ (u_{N-1} + v_{N-1}) \cdot 2^{N-1}.$$

The coefficient $(u_{N-1}v_{N-1}-u_{N-1}-v_{N-1})$ can be transformed to

$$(u_{N-1} v_{N-1} - u_{N-1} - v_{N-1}) \cdot 2^{2N-2} = (1 + \bar{u}_{N-1} \bar{v}_{N-1}) \cdot 2^{2N-2} - 1 \cdot 2^{2N-1}. \tag{7}$$

Therefore the uniform additive formula is advanced:

$$p = 1 \cdot (-2^{2N-1}) + (1 + \bar{u}_{N-1} \bar{v}_{N-1}) \cdot 2^{2N-2} + \sum_{i=0}^{N-2} \sum_{k=0}^{N-2} u_i v_k \cdot 2^{i+k}$$

$$+ \sum_{i=0}^{N-2} (u_{N-1} \bar{v}_i + v_{N-1} \bar{u}_i) \cdot 2^{i+N-1} + (u_{N-1} + v_{N-1}) \cdot 2^{N-1}. \tag{8}$$

The above convolution can be realized in an parallel array architecture(Fig. 1). The digits of the two numbers are spread orthogonally. All the bit products are formed in parallel by AND-ing the inputs crossing at each cell. Meanwhile, all the

partial products along each diagonal are equivalently weighted and are added separately without carries, resulting in a new 2N-digit-long mixed twos complement sequence. The advantages are as follows: (1) The weighted sum constitutes the product directly. (2) We need not know the number of bits to represent the product before the operation is conducted. (3) The present algorithm requires only (N+2) channels, and all the bit products contribute to the result. These have greatly saved the SBWP. In addition, it is also feasible to just use N channels if the last term in Eq.(8) is supplemented by postprocessing. (4) Since each digit of the mixed product is weighted as in 2's complement system, in principle, the architecture is cascadable.

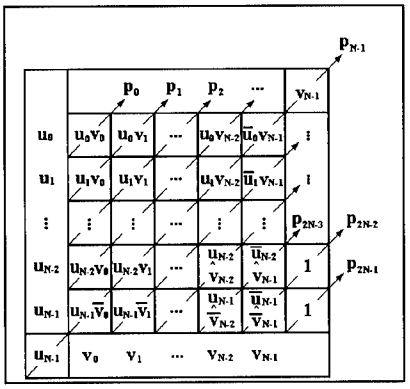


Figure 1

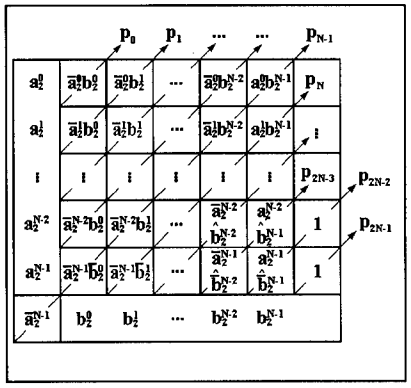


Figure 2

3. Two-stage array complex multiplication

Assume two complex operands: $a=a_1+ja_2$, $b=b_1+jb_2$, where a_1, a_2, b_1, b_2 are real numbers and are each encoded by N bits in 2's complement. We use superscripts to symbolize the digit significance. In the product

$$axb = (a_1b_1-a_2b_2) + j (a_1b_2 + a_2b_1),$$

the multiplications a_1b_1, a_1b_2, a_2b_1 are performed in the preceding way, and $(-a_2b_2)$ can be realized by a different parallel bit product array shown in Fig. 2 according to the following equation, except the terms $b_2^i 2^i$ (i=0, 1, ···, N-2) left for postprocessing:

$$(-a_2)\times b_2 = 1\cdot(-2^{2N-1}) + (1+a_2^{N-1}\overline{b_2}^{N-1})\cdot 2^{2N-2} + \sum_{i=0}^{N-2}\sum_{k=0}^{N-2}\overline{a_2}^{i} b_2^{k}\cdot 2^{i+k}$$

$$+ \sum_{i=0}^{N-2}(\overline{a_2}^{N-1}\overline{b_2}^{i} + b_2^{N-1}a_2^{i})\cdot 2^{i+N-1} + (\overline{a_2}^{N-1}\cdot 2^{N-1} + \sum_{i=0}^{N-2} b_2^{i}\cdot 2^{i}). \qquad (9)$$

The complex product can be generated by a two-stage 2x2 array in perfectly parallel(see Fig.3). In the $(-a_2b_2)$ subarray, the inputs related to a_2 are negated in

comparison with the left a_2b_1 subarray.

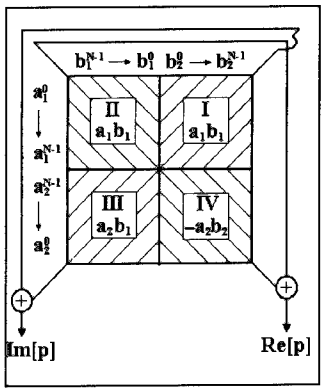


Figure 3

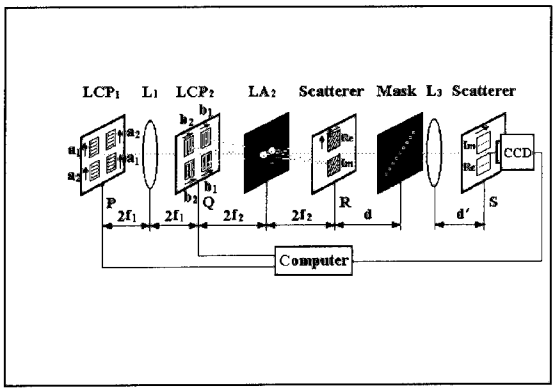


Figure 4

4. Optical implementation

The two-stage array architecture has been verified by experiment. The optical construction, which comprises three parts, is illustrated in Fig. 4. The first imaging part(P-Q) forms the four product subarrays. The second multiple-imaging system[6](Q-R) carries out addition of two of the relevant subarrays. The incoherent correlator[7](R-S) integrates the diagonal elements to get both real and imaginary parts of the result. The system is interfaced with a PC computer for pre- and post-processing.

Acknowledgment

The financial support of the National Natural Science Foundation of China is gratefully acknowledged.

References:

[1] R. P. Bocker, S. R. Clayton and K. Bromley, Appl. Opt. **22** (1983) 2019-2021.
[2] A. P. Goutzoulis, Appl. Opt. **23** (1984) 4095-4099.
[3] B. K. Taylor and D. Casasent, Appl. Opt. **25** (1986) 956-961.
[4] K. Hwang, Computer Arithmetic: Principles, Architecture, and Design (Wiley, New York, 1979), Ch. 6.
[5] E. Swartzlander, Appl. Opt. **25**, (1986) 3021-3032.
[6] L. Liu and X. Liu, Opt. Commun. **82** (1991) 446-452.
[7] L. Liu, Opt. Lett. **14** (1989) 482-484.

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical Parallel Arithmetic Using Recoded Signed-Digit Number System

Mohammad S. Alam

Purdue University at Fort Wayne, Department of Engineering, Fort Wayne, IN 46805

Abstract: A new technique for high-speed recoded trinary signed-digit (TSD) arithmetic using optical symbolic substitution is proposed. This technique performs multi-digit carry-free addition and borrow-free subtraction in constant time using only 39% of the minterms required in the most recently reported TSD arithmetic technique.

1 Introduction

The problems of achieving massively parallel optical computing have been investigated by many authors using the residue number system, signed-digit number system, and recoded signed-digit number systems [1-4]. Using residue number system, one can perform parallel arithmetic in constant time using symbolic substitution (SS) [5], but the size of the truth table required increases rapidly with the increase of the operand length. The redundant signed-digit number system [1] allows parallel arithmetic with fewer carry propagation steps. Higher radix signed-digit number systems allow higher information storage density, less complexity, fewer system components, and fewer cascaded gates and operations. Among the higher radix number systems, the trinary signed-digit number system appears to be the most promising in terms of storage density and processing elements [6].

Recently, a two-step symbolic substitution (SS) technique for TSD arithmetic has been reported [6]. This technique performs carry-free addition and borrow-free subtraction by checking a pair of reference digits from the next lower order digit position and requires 58 four-variable minterms for each output digit. Later on, a higher-order TSD SS technique has been designed [4] which enhances the computation speed at the expense of increasing the number of six-variable minterms per output digit. A simpler carry-free addition scheme [7] where the binary signed-digit numbers are recoded before performing the addition has recently been proposed. Awwal [8] implemented the recoded binary signed-digit arithmetic technique using an opto-electronic implementation which requires sixteen four-variable minterms per output digit. Thus, in any SS scheme, the most important objective is to minimize the number of minterms (substitution rules) and computational steps while incorporating the optimum information in fewer digits. To achieve these objectives without sacrificing the processing speed and parallelism of optics, we propose herein the recoded TSD system for high-speed multi-digit carry-free addition and borrow-free subtraction.

2 Recoded TSD Arithmetic

A TSD literal may be represented as a member of the set $\{\overline{2}, \overline{1}, 0, 1, 2\}$ with $\overline{2}$ and $\overline{1}$ representing -2 and -1 respectively. When two TSD numbers are added, a carry will be generated for the digit combinations $\overline{2}\,\overline{2}$, $\overline{2}\,\overline{1}$, $\overline{1}\,\overline{2}$, $1\,2$, $2\,1$, and $2\,2$. For carry-free addition, the aforementioned

digit combinations must be eliminated from the augend and addend operands. By exploiting the redundancy of the TSD numbers, it is possible to eliminate the occurrence of the above mentioned digit combinations through a recoding truth-table as shown in Table I. An n-digit TSD number $P = P_n P_{n-1} P_{n-2} \cdots P_1 P_0$ when recoded using Table I results in an $(n+1)$-digit recoded TSD number $Q = Q_{n+1} Q_n Q_{n-1} Q_{n-2} \cdots Q_1 Q_0$ such that both P and Q are numerically equal and the digits in Q satisfies the condition: $Q_i \times Q_{i-1} \notin \{2, 4\}, 0 < i \leq (n+1)$. In Table I, four TSDs are investigated to output a recoded TSD. With four TSDs, it is possible to generate a number over the range of -80_{10} ($\bar{2}\bar{2}\bar{2}\bar{2}_3$) to $+80_{10}$ (2222_3). For space limitation, in Table I, we have shown only those input combinations within the range -80 to -1 (referred to as the negative category) which yields non-zero outputs. The remaining entries representing the range $+1$ to $+80$ (referred to as the positive category) may be obtained by complementing the corresponding negative category entry. For illustration, if the first entry ($\bar{2}\bar{2}12_3 = -75_{10}$) of Table I is complemented, we obtain the corresponding digits ($22\bar{1}\bar{2}_3 = 75_{10}$) of the positive category entry. Thus, in Table I, the input minterms and their corresponding output digits have a complementary relationship. Also, the Q_i output of Table I does not include the 2 and $\bar{2}$ literals. Therefore, the recoding operation maps the TSD set $\{\bar{2}, \bar{1}, 0, 1, 2\}$ into a smaller set $\{\bar{1}, 0, 1\}$. The addition truth-table corresponding to the recoded TSD output of Table I is shown in Table II, where A_i, B_i and S_i represent the addend, augend and sum operands, respectively. Table II represents a simplified TSD addition truth-table since it does not incorporate any minterm involving the 2 and $\bar{2}$ literals. Note that it is necessary to pad three zeros trailing the least significant digit and one zero preceding the most significant digit in order to apply Table I to recode a TSD number. To show the application of Table I and Table II, consider the following example for TSD addition:

Operand type		TSD representation		Recoded TSD representation		Decimal representation
↓		↓		↓		↓
Addend	=	$(\emptyset 221\bar{2}\bar{2}\bar{1}1\,\emptyset\emptyset\emptyset)_3$	=	$(10\bar{1}001\bar{1}1)_3$	=	$(1951)_{10}$
Augend	=	$(\emptyset\bar{1}121\bar{2}\bar{2}2\,\emptyset\emptyset\emptyset)_3$	=	$(00\bar{1}\bar{1}01\bar{1}\bar{1})_3$	=	$(-319)_{10}$
Final sum	=		=	$(10\bar{2}\bar{1}02\bar{1}0)_3$	=	$(1632)_{10}$
Final carry	=		=	$(000\,00000)_3$	=	$(0)_{10}$

where $\emptyset$ indicates a padded zero. The above results clearly illustrate that the addition of the recoded TSD numbers is carry-free.

3 Optical Implementation

Among the various SS architectures proposed for optical implementation of the binary and signed-digit arithmetic [2-4, 6, 9-12], the direct truth-table content-addressable memory (CAM) based approach that precalculates and stores the processing results [11] appears to be the most promising in terms of the computation speed. To design a system with the optimum memory requirement but relatively high speed, either holographic or nonholographic CAM based optical implementation as detailed in References 10 and 12 may be used. The first step to design a CAM is to minimize the truth table and obtain the minterms corresponding to the non-zero entries of the desired output. The reduced minterm sets of the recoding truth table are $Q_1 \in \{d_{1\bar{2}}\bar{2}2d_{\bar{1}012}, d_{1\bar{2}}\bar{1}\bar{1}d_{\bar{1}012}, 012d_{\bar{1}012}, 02\bar{1}d_{\bar{1}012}, d_{\bar{1}2}\bar{2}1d_{10\bar{1}\bar{2}}, d_{1\bar{2}}11d_{10\bar{1}\bar{2}}, d_{\bar{1}2}\bar{1}\bar{2}d_{10\bar{1}\bar{2}}, d_{1\bar{2}}2\bar{2}d_{10\bar{1}\bar{2}}, d_{1\bar{2}}\bar{1}d_{012}d, d_{1\bar{2}}1d_{0\bar{1}\bar{2}}d, d_{\bar{1}2}\bar{2}d_{0\bar{1}\bar{2}}d, 02d_{012}d, d_{1\bar{2}}\bar{2}12, d_{1\bar{2}}\bar{1}\bar{2}2, d_{1\bar{2}}122, d_{1\bar{2}}2\bar{1}\bar{2}, d_{\bar{1}2}\bar{2}\bar{2}2, d_{\bar{1}2}\bar{1}\bar{1}\bar{2}, d_{1\bar{2}}0dd, 02\bar{2}2, 0112\}$ and $Q_{\bar{1}} \in \{d_{\bar{1}2}2\bar{2}d_{10\bar{1}\bar{2}}, d_{\bar{1}2}11d_{10\bar{1}\bar{2}}, 0\bar{1}\bar{2}d_{10\bar{1}\bar{2}}, 02\bar{1}d_{10\bar{1}\bar{2}}, d_{1\bar{2}}2\bar{1}d_{\bar{1}012}, d_{1\bar{2}}11d_{\bar{1}012}, d_{1\bar{2}}12d_{\bar{1}012}, d_{1\bar{2}}\bar{2}2d_{\bar{1}012}, d_{\bar{1}2}1d_{0\bar{1}\bar{2}}d, 02d_{0\bar{1}\bar{2}}d, d_{\bar{1}2}\bar{1}d_{012}d, d_{1\bar{2}}2d_{012}d, d_{\bar{1}2}2\bar{1}\bar{2}, d_{\bar{1}2}12\bar{2}, d_{\bar{1}2}\bar{1}\bar{2}2, d_{\bar{1}2}\bar{2}12, d_{1\bar{2}}2\bar{2}2, d_{1\bar{2}}112, d_{\bar{1}2}0dd, 0\bar{2}\bar{2}\bar{2}, 0\bar{1}\bar{1}\bar{2}\}$, respectively. Thus, 42 four-variable ($P_i P_{i-1} P_{i-2} P_{i-3}$) minterms are required to generate the 1 and $\bar{1}$ outputs of Q_i. Notice that the minterms for yielding the 1

Table I. Recoding Truth-Table for TSD Numbers.

P_i	P_{i-1}	P_{i-2}	P_{i-3}	Q_i
$\bar{2}$	$\bar{2}$	1	2	1
$\bar{2}$	$\bar{2}$	2	$d_{\bar{1}012}$	1
$\bar{2}$	$\bar{1}$	$\bar{2}$	2	1
$\bar{2}$	$\bar{1}$	$\bar{1}$	$d_{\bar{1}012}$	1
$\bar{2}$	$\bar{1}$	d_{012}	d	1
$\bar{2}$	0	d	d	1
$\bar{2}$	1	$d_{0\bar{1}\bar{2}}$	d	1
$\bar{2}$	1	1	$d_{10\bar{1}\bar{2}}$	1
$\bar{2}$	1	1	2	$\bar{1}$
$\bar{2}$	1	2	$\bar{2}$	1
$\bar{2}$	1	2	$d_{\bar{1}012}$	$\bar{1}$
$\bar{2}$	2	$\bar{2}$	$d_{10\bar{1}\bar{2}}$	1
$\bar{2}$	2	$\bar{2}$	2	$\bar{1}$
$\bar{2}$	2	$\bar{1}$	$\bar{2}$	1
$\bar{2}$	2	$\bar{1}$	$d_{\bar{1}012}$	$\bar{1}$
$\bar{2}$	2	d_{012}	d	$\bar{1}$
$\bar{1}$	$\bar{2}$	$d_{0\bar{1}\bar{2}}$	d	1
$\bar{1}$	$\bar{2}$	1	$d_{10\bar{1}\bar{2}}$	1
$\bar{1}$	$\bar{2}$	1	2	$\bar{1}$
$\bar{1}$	$\bar{2}$	2	$\bar{2}$	1
$\bar{1}$	$\bar{2}$	2	$d_{\bar{1}012}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$d_{10\bar{1}\bar{2}}$	1
$\bar{1}$	$\bar{1}$	$\bar{2}$	2	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{2}$	1
$\bar{1}$	$\bar{1}$	$\bar{1}$	$d_{\bar{1}012}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	d_{012}	d	$\bar{1}$
$\bar{1}$	0	d	d	$\bar{1}$
$\bar{1}$	1	$d_{0\bar{1}\bar{2}}$	d	$\bar{1}$
$\bar{1}$	1	1	$d_{10\bar{1}\bar{2}}$	$\bar{1}$
$\bar{1}$	1	2	$\bar{2}$	$\bar{1}$
$\bar{1}$	2	$\bar{2}$	$d_{10\bar{1}\bar{2}}$	$\bar{1}$
$\bar{1}$	2	$\bar{1}$	$\bar{2}$	$\bar{1}$
0	$\bar{2}$	$d_{0\bar{1}\bar{2}}$	d	$\bar{1}$
0	$\bar{2}$	1	$d_{10\bar{1}\bar{2}}$	$\bar{1}$
0	$\bar{2}$	2	$\bar{2}$	$\bar{1}$
0	$\bar{1}$	$\bar{2}$	$d_{10\bar{1}\bar{2}}$	$\bar{1}$
0	$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{1}$

Table II. Addition Truth-Table for Recoded TSD Numbers.

Addend A_i	Augend B_i	Sum S_i
$\bar{1}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	0	$\bar{1}$
0	$\bar{1}$	$\bar{1}$
0	0	0
$\bar{1}$	1	0
1	$\bar{1}$	0
0	1	1
1	0	1
1	1	2

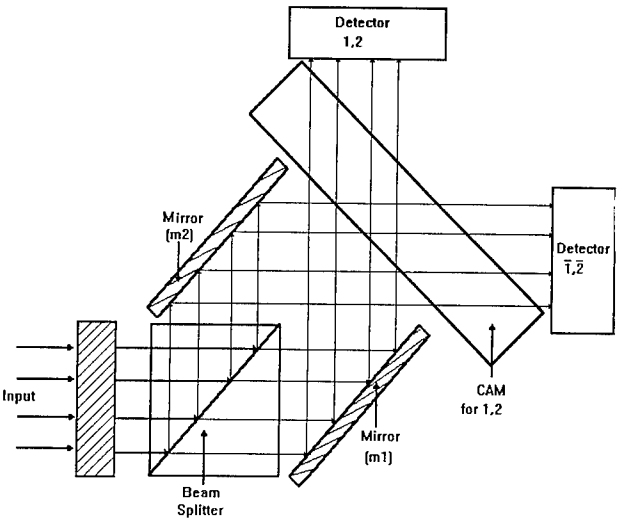


Fig. 1. Optical implementation for the recoded trinary signed-digit arithmetic using CAM.

and $\bar{1}$ outputs of Q_i are exact complements of each other. For illustration, when the reduced minterm $d_{1\bar{2}}\bar{2}\,2\,d_{\bar{1}0\,1\,2}$ for generating the 1 output of Q_i is complemented, we obtain the corresponding minterm $d_{\bar{1}2}\,2\,\bar{2}\,d_{10\bar{1}\bar{2}}$ for generating the $\bar{1}$ output of Q_i. For the second processing step using Table II, only two minterms 01 and 10 ($0\,\bar{1}$ and $\bar{1}\,0$) are required to generate the 1 ($\bar{1}$) output, and only one minterm 11 ($\bar{1}\,\bar{1}$) is required to generate the 2 ($\bar{2}$) output of S_i. Also, the minterms for $S_i = 1$ (2) are digit-by-digit complement of the corresponding minterms for $S_i = \bar{1}$ ($\bar{2}$). Thus, 21 minterms are required to generate the 1 output of C_i, and 3 minterms are required to generate the 1 and 2 outputs of S_i. In the quad-rail coding scheme, a TSD $\overline{t_i}$ is 180° rotated version of t_i. Accordingly, the minterms corresponding to the $\bar{1}$ output of Q_i, and the $\bar{1}$ and $\bar{2}$ outputs of S_i can be generated by using a slightly different geometric configuration employing mirrors. Therefore, the whole system can be implemented with 21 four-variable minterms for recoding and 3 two-variable minterms for generating each final output digit. It may be mentioned that a two-variable minterm occupies only half of the encoding space occupied by a four-variable minterm. The most recently reported two-stage implementation for TSD addition employed 58 four-variable minterms [6]. Thus, the recoded TSD technique requires only 39% of the minterms used in Reference 6 for TSD arithmetic.

The proposed technique can be implemented using a holographic or nonholographic CAM-based SS system [10, 12]. A holographic CAM requires coherent light and suffers from its inconvenient recording problems. On the other hand, a nonholographic CAM can be implemented with coherent or non-coherent optical processing. Because of the ease of alignment and implementation, in comparison to a holographic CAM, a nonholographic CAM approach may be used for the proposed technique. The optical implementation for the proposed technique is shown in Fig. 1 where the inputs are recoded electronically and the addition step is performed optically to generate the final output. Note that only the minimized minterms for the 1 and 2 outputs of the addition step are stored in the nonholographic CAM while the minterms for the $\bar{1}$ and $\bar{2}$ outputs are generated by the beam splitter and mirror (m2) combination.

4 Conclusion

We have proposed an efficient trinary signed-digit arithmetic technique for performing parallel carry-free addition and borrow-free subtraction in a constant time in two steps independent of the digit string length. This technique requires the minimum number of minterms when compared to the previously reported trinary signed-digit arithmetic techniques [4, 6]. Finally, a CAM based optical implementation is suggested for the proposed scheme.

References

[1] Avizienis A 1961 *IRE Trans. Electron. Comp.* **EC-10** 389-400

[2] Hwang K and Panda D K 1992 *Opt. Eng.* **31** 2422-2433

[3] Drake B L, Bocker R, Lasher M, Patterson R and Miceli W1986 *Opt. Eng.* **25** 38-43

[4] Alam M S, Karim M A, Awwal A A S and Westerkamp J J 1992 *Appl. Opt.* **31** 5614-5621

[5] Huang A 1983 *Proc.10th International Optical Computing Conference* (IEEE Computer Society, Los Angeles) 13-17

[6] Awwal A A S, Islam M N and Karim M A 1992 *Appl. Opt.* **31** 1687-1694

[7] Parhami B 1988 *IEEE Trans. Comput* **37** 1470-1476

[8] Awwal A A S 1992 *Appl. Opt.***31** 3205-3208

[9] Cherri A K and Karim M A 1988 *Appl. Opt.* **27** 3824-3827

[10] Li Y and Eichmann G 1987 *Appl. Opt.* **26** 2328-2333

[11] Mirsalehi M M and Gaylord T K 1986 *Appl. Opt.* **25** 2277-2283

[12] Li Y, Kim D H, Kostrzewski A and Eichmann G 1989 *Opt. Lett.* **14** 1254-1256

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Digital Optical Computing Architectures for Compute Intensive Applications

Peter S. Guilfoyle

OptiComp Corporation, P.O. Box 10779, Zephyr Cove, Nevada 89448

Abstract

Two digital optical computing architectures are being developed for compute intensive applications. A 32-bit digital optical processor has been implemented in hardware. The machine is called DOC II (digital optical computer). The platform implements the N^2 (parallel) architecture at an interconnect density of 8192 and a clock rate of 100 MHz resulting in an gate-interconnect-bandwidth-product of 0.8192×10^{12}. Development efforts are underway to implement the N^4 (global) smart pixel architecture within miniature high performance optoelectronic computing (HPOC) modules. Objectives include the fabrication of 10^6 interconnect modules at clock rates up to 1 GHz.

1.0 N^2 parallel architecture

The architecture for DOC II is referred to as the N^2 architecture. It is a Boolean vector/matrix multiplication unit as shown in Figure 1. Equation (1.) relates the Boolean summation of an **I** element vector, $\mathbf{x}_i$, (I→N) against a control microcode matrix ξ_{ij} which contains on the order of N^2 terms **I*J**.

$$f_j = \overline{\sum_{i=1}^{I} x_i \xi_{ij}} = \prod_{i=1}^{I} \overline{x_i \xi_{ij}} \qquad (1.)$$

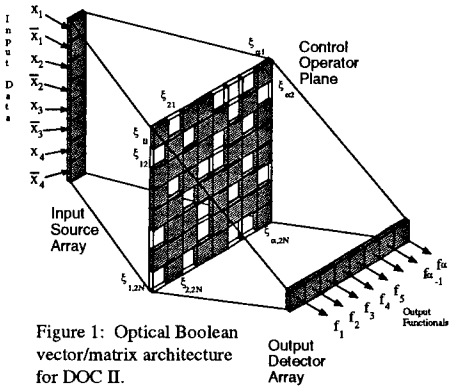


Figure 1: Optical Boolean vector/matrix architecture for DOC II.

In actuality **i** goes to **I** which is on the order of **N** but **j** is arbitrary depending on the number of detectors and/or the size or availability of the control operator hardware. If **i≠j** then the gate interconnect bandwidth product for this architecture is $\mathbf{B*I*J}=\gamma\mathbf{*B*N^2}$, where γ is a proportionality constant depending on **J** (the limit of **j** which equals α in Figure 1.), and **B** is the bit rate. Since the summation on the detector is strictly Boolean, i.e. a threshold is performed strictly between the zero light state and a FAN-IN of 1, then by DeMorgan's theorem[1.], the output is represented by the product/multiplication as indicated by equation (1.). Consequently, α Boolean minterms or functionals on a given data input vector $\mathbf{x}_i$ may be formed on the detector array at each clock cycle. A second pass through the system allows for the summation of the functionals to form partial or complete instructions depending on the complexity of the instruction.

The hardware consists of two primary assemblies (Figures 2 and 3): the illumination assembly (Train A) and the modulation relay assembly (Train B). Both trains, which are optimized for maximum throughput at 837 nm, are positioned on a 36" x 48" optical table which is packaged in a optoelectronic cabinet. During each clock cycle the hardware

provides up to 128 functionals on a given input vector of up to 32 bits (dual rail) in length. This is provided by 64 independently modulated input edge emitting laser diodes arranged in groups of eight elements. The control mask is clocked into the spatial light modulator (64 x 128 input control data matrix). The spatial light modulator consists of a 64 channel GaP acousto-optic Bragg cell. Each channel is centered at 800 MHz with a bandwidth of 400 MHz and a time aperture of 2.56 μs. Only an aperture of 1.28 μs is utilized. The output consists of a 128 element linear avalanche photodiode array followed by a transimpedance receiver and ECL thresholds. For further reading on this hardware the reader is referred to references 2 and 3. Thus, the processor can achieve an optical data carrier fan-out of 1:128 and a control logic fan-in of 64:1.

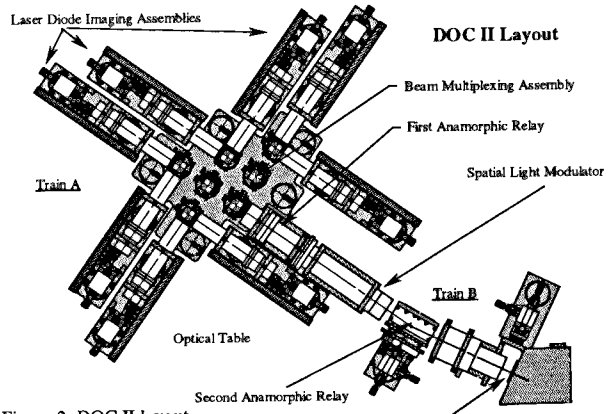


Figure 2: DOC II layout

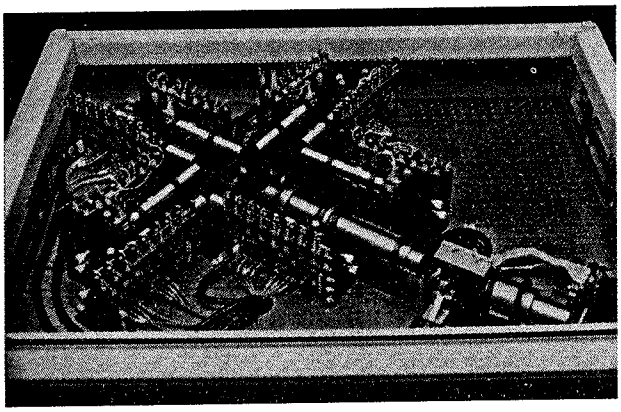


Figure 3: Photograph of DOC II layout

Current compute intensive applications including RISC emulation, full text data searches, and multimedia database development have been demonstrated.

2.0 N^4 global architecture

The architecture for DOC III is referred to as the N^4 architecture. It is a Boolean matrix/tensor multiplication unit as shown in Figure 4. Equation (2.) relates the Boolean summation of matrix, $\mathbf{x}_{ij}$, which contains on the order of N^2 terms, **I*J**, (**i→I, j→J**) against a control microcode tensor $\xi_{ij,kl}$ which contains, in optical memory, on the order of N^4 terms, **I*J*K*L**.

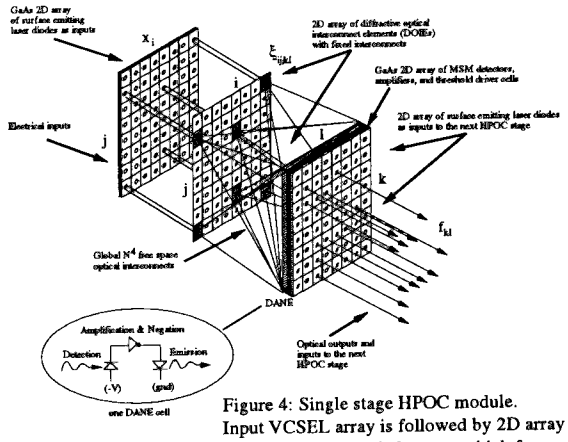


Figure 4: Single stage HPOC module. Input VCSEL array is followed by 2D array of diffractive optical elements which form N4 interconnects to subsequent receiver amplifier arrays and VCSEL re-emitters.

$$f_{kl} = \overline{\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}\,\xi_{ij,kl}} = \prod_{i=1}^{I}\prod_{j=1}^{J} \overline{x_{ij}\,\xi_{ij,kl}} \qquad (2.)$$

In actuality **i** goes to **I** which is on the order of **N** but **k** is arbitrary depending on the number of detectors and/or the complexity or availablilty of the control operator hardware. In addition, **j** goes to **J** which is on the order of **N** but **l** is arbitrary depending on the number of detectors and/or the complexity of the control operator diffractive optical element. If **i≠k** and **j≠l** then the gate interconnect bandwidth product for this architecture is $\mathbf{B*i*j*k*l=\gamma*B*N^4}$, where γ is a proportionality constant depending on **I, J, K,** and **L**, where **B** is the bit rate. Since the summation on the detector is again strictly Boolean, i.e. a threshold is performed strictly between the zero light state and a FAN-IN of 1, then by DeMorgan's theorem, the output is represented by the matrix/ tensor product as indicated by equation (2.). Consequently **K*L** Boolean minterms or functionals on a given data input matrix $\mathbf{x_{ij}}$ may be formed on the detector array at each clock cycle.

The cascade of a second HPOC module, as shown in Figure 5, allows for the summation of the functionals to form partial or complete instructions depending on the complexity of the instruction.

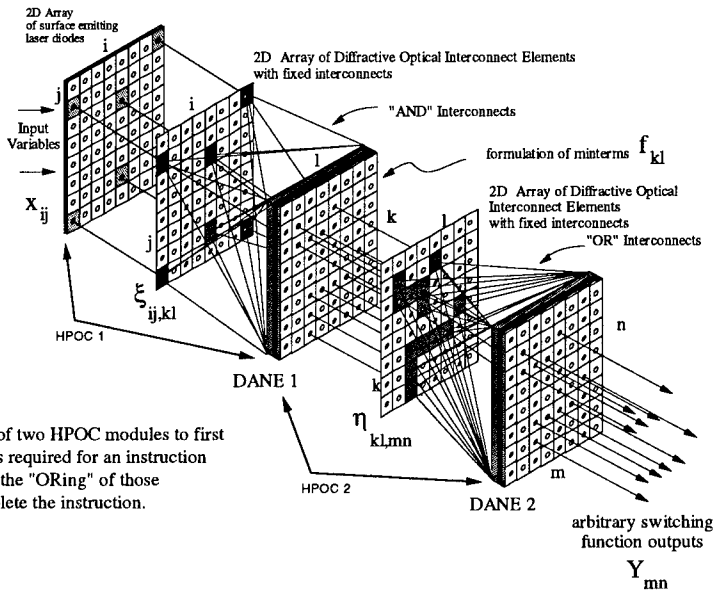


Figure 5: Cascade of two HPOC modules to first form the functionals required for an instruction and second to form the "ORing" of those functionals to complete the instruction.

The second HPOC module provides for the double summation of the functionals computed from the first HPOC module. The output of the double stage cascade, $\mathbf{Y_{mn}}$, represents complete or partial instructions computed depending on the length of the instruction encoded into the diffractive optical elements within each HPOC. As can be seen in equation (3.) the double stage cascade produces a four dimensional Boolean matrix/ tensor multiply followed by a Boolean matrix/tensor addition.

$$Y_{mn} = \sum_{k=1}^{K} \sum_{l=1}^{L} \left\{ \prod_{i=1}^{I} \prod_{j=1}^{J} \overline{x_{ij}\, \xi_{ij,kl}} \right\} \eta_{kl,mn} \qquad (3.)$$

This is required to fulfill the sum-of-products formulation as required by Shannon's generalized digital computation theory. [4.]

The hardware being developed for HPOC module implementation is shown Figures 5 and 6. Each module is being developed as a hybrid structure consisting of five elements. The transmitter arrays consist of vertical cavity surface emitting laser diodes (VCSELs). Initially these GaAs arrays are electrically connected at the edges through electrical FAN-IN from the receiver arrays via wire bonds. The laser arrays are epoxied to a ceramic substrate. The receiver arrays are designed in GaAs with metal-semiconductor-metal (MSM) detectors followed by an E/D MESFET transimpedance amplifier and laser driver stage. The diffractive optical elements are mounted above each of the GaAs structures and physically bonded to beam folding optics which allows for semi-planar implementation as shown in Figure 6. Here a 3x3 array of four-stage cascade HPOC modules are mounted on a board. Each four-stage HPOC island consists of two arrays of VCSELs and two arrays of detection, amplification, negation and driver stages. More complete descriptions of this hardware may be found in references [5-7].

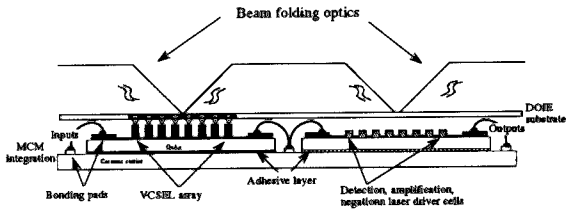
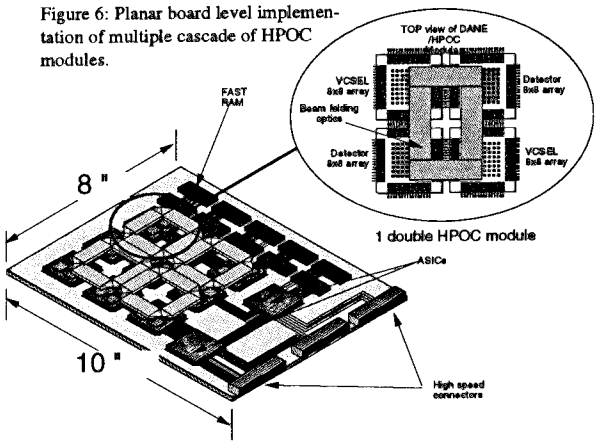


Figure 5: Single stage HPOC module hardware configuration. Beam folding optics provide semi-planar implementation of hybrid GaAs transmitter/receiver structures coupled by diffractive optical interconnect elements to provide N^4 free space optical interconnects.

Figure 6: Planar board level implementation of multiple cascade of HPOC modules.



The author wishes to express its gratitude for the funding provided for this work under ONR/ ARPA contract N00014-93-C-0134, ONR/ROME Lab/BMDO contract N00014-89-C-0266, and ONR/BMDO contract N00014-91-C-0253.

References

[1.] J. Millman, C. Halkias, **Integrated Electronics: Analog and Digital Circuits and Systems**, McGraw-Hill, New York, SBN 07-042315-6, 1972, Pgs. 173-175.

[2.] P.S. Guilfoyle, et. al., "Comprehensive Digital Optoelectronic Computing Program for High Performance Applications," **Proceedings IEEE Micro**, Dec. 1994 (accepted, to be published).

[3.] Stone, et. al., Digital Optical Interconnects for Photonic Computing, **SPIE, OE/LASE, Optoelectronic Interconnects, 2153-06,** January, 1994, Los Angeles, CA.

[4.] C. E. Shannon, "A Symbolic Analysis of Relay and Switching Circuits," **AIEE Transactions**, Vol. 57 1938, pp 714-723.

[5.] P.S. Guilfoyle, R.V. Stone, F.F. Zeise, J.M. Hessenbruch, W. J. Miceli; "Low Power, 4D (Global) Optical Interconnects for Wide Word Processing"; **Proceedings of the SPIE**, 2297-41, Optoelectronic Enhancements to Digital Computing Technology; San Diego, CA, July, 1994.

[6.] P. S. Guilfoyle, et. al., "Diffractive Optical Interconnect Design for Digital Switching"; **SPIE, OE/LASE, 2152-20**, Diffractive and Holographic Optics Technology, January, 1994, Los Angeles, CA.

[7.] P. S. Guilfoyle, et. al., "Smart Optical Interconnects for High Speed Photonic Computing", **OSA, Optical Computing**, March, 1993, Palm Springs, CA.

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Parallel optoelectronic computing systems and applications

Masatoshi Ishikawa

University of Tokyo, Bunkyo-ku, Tokyo 113, Japan
e-mail: ishikawa@k2.t.u-tokyo.ac.jp

Abstract: A massively parallel processing system with a reconfigurable shift-invariant optical interconnection among electronic general purpose processing elements (PE's) is described. Each PE is so compact that more than 4,000 PE's can be integrated into one chip for directly coupling with array type optical devices in parallel. The optical interconnection is constructed by using a suface emitting laser diode array and a phase modulation type spatial light modulator on which optimized computer generated holograms are written. In this paper, the design concept of the system and the PE, configration of an experimental system, and algorithm for the parallel optoelectronic computing system with some applications are shown.

1. Introduction

Integrated optical devices such as surface emitting laser diode arrays, photo detector arrays and OEIC's promise high performance parallel processing. Although these devices have two-dimensional parallelism for pattern information processing, their potential capabilities are not utilized in conventional parallel processing systems. The "I/O bottleneck" between processing element (PE) array and I/O devices sets the limit to the processing speed.

In order to overcome the I/O bottleneck, two-dimensional interconnections between them are required. In other words, it is expected that high performance parallel computing systems can be realized by using such two-dimensional optical devices and fine grain electronic PE arrays. Although processing architectures for utilizing such devices are intrinsically based on massively parallel processing with parallel optical input and output, the interconnects can not be implemented by using conventional macro-scale wiring technology. Only the integrated wiring technology such as VLSI, flip chip bonding, and GaAs on silicon can implement the interconnection. This means that the PE should be so compact to be implemented with integrated optical devices. Integrated PE and optical device arrays realize fully parallel processing with high computing performance for some application fields such as high speed visual image processing, robot vision, automated visual inspection, and pattern recognition.

In this paper, the design concept of the system, configuration of an experimental system, and algorithms for the parallel optoelectronic computing system are shown.

2. Parallel optoelectronic processing systems

Conventional microprocessors can not be used as the PE of the system, because the number of gates is too large that many PE's cannot be implemented into small area on VLSI's. Keys of the design are how generality of processing is kept by using small number of gates and how the interconnection is implemented.

Based on this design concept, Ishikawa et al. have proposed a fully parallel hierarchical processing architecture and a compact PE architecture for general purpose processing [1]. A conceptual diagram of the architecture is shown in Fig. 1. Each layer in this architecture consists of the same module with the two-dimensional PE array and optical I/O devices and is optically interconnected with neighbor layers. Since the processing of each layer is independent of the processing of other layers, different programs can be carried out at each layer.

Since the architecture is difficult to be implemented as an experimental system because the system requires many devices, an equivalent artitecture by time sharing algorithm which carries out the processing of each layer in a time-sequential order has

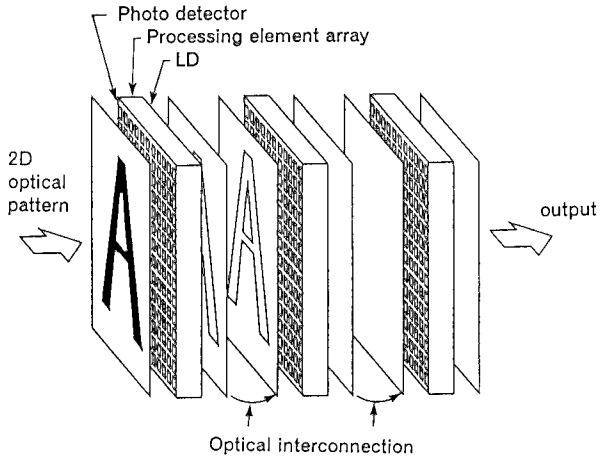


Fig. 1 Hierarchical optoelectronic processing system

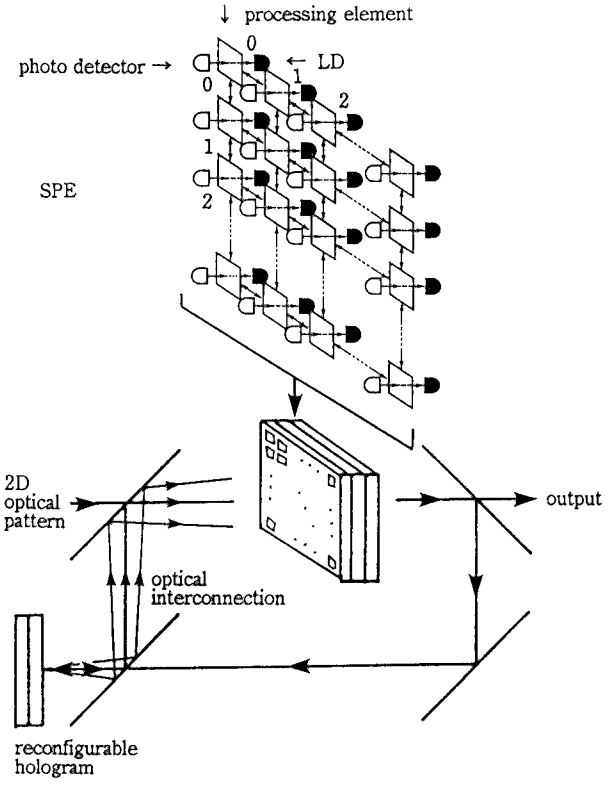


Fig. 2 Feedback type system

been proposed. A conceptual diagram of the feedback type architecture is shown in Fig. 2. Although the PE has electrical interconects with the four nearest neighbours, reconfigurable optical interconnects are used for realizing wade range of interconnection in order to maintain the generality of interconnection.

2.1 Processing Element (PE)

In order to implement the optoelectronic parallel processing system shown in Fig. 2, Ishikawa et al. designed a compact and general purpose PE. The block diagram of the PE is shown in Fig. 3 [1].

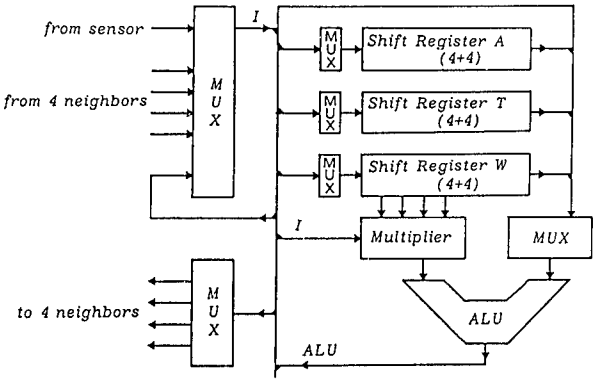


Fig. 3 Block diagram of processing element

The architecture of the PE has following features to be compact; 1) direct connection between PE and optical I/O, 2) SIMD type parallel processing, 3) controlled by micro-instruction, 4) bit serial processing, 5) restricted electrical interconnects (4 neighbours).

Each PE has three 8bit registers (A:Accumulator, T:Template, W:Weight), one arithmetic logical unit (ALU, 1bit), and one 4-bit multiplier as shown in Fig. 3. Processing architecture of the ALU is based on bit serial processing which is slow in comparison with bit parallel processing, but dose not require so many number of gates, then it has major advantages for integration and variable bit length processing. Functions of ALU include AND, OR, Exclusive OR, addition, subtraction, multiply (4bits × 4bits) and combinations of these basic functions such as weighted sum for calculation of correlation.

Most important specification of the PE is the number of gates. In the result, the PE is implemented by using 337 gates. The number of the gates is quite small. Processing cycle time is typ. 44ns, max. 87ns.

2.2 Scale up model : SPE-4k

The PE have been designed to be so compact that more than 64 × 64=4096 PE's may be integrated into one chip by using present VLSI technology. An experimental scale up model with 64 × 64=4096 PE's have been also developed. The system is named SPE-4k (Sensory Processing Elements - 4k). The system is regarded as a scale up model of one chip optoelectronic processing layer shown

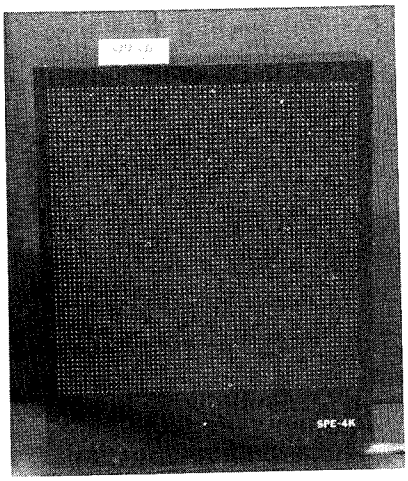


Fig. 4 SPE-4k

in Fig. 1 and Fig. 2. The overview of SPE-4k is shown in Fig. 4. PE's which are implemented by using gate arrays are arranged between the front LED array and the back PTR array.

Considering the 8bit integer addition as a basic operation of processing for the evaluation of the speed of SPE-4k, total 32MOPS (Mega Operations Par Second) by the present system and 3.2GOPS (Giga Operations Par Second) at the maximum speed of the system are obtained.

2.3 Reconfigurable optical interconnection

The PE has electronic interconnects with four neighbors, but other interconnects require iterative operations which takes a long processing time, when only electronic interconnects are used. In order to directly connect between arbitrary PE's. Kirk et al. have designed a reconfigurable optical interconnection [2] and showed basic experimental results [3] using a phase modulation type spatial light modulator, PAL-SLM (Parallel Aligned Spatial Light Modulator) developed by Hamamatsu K.K. [4].

The interconnection subsystem is based on computer generated hologram (CGH) optimized by using simulated annealing algorithm. It realizes shift invariant interconnection for SIMD type parallel processing. Reconfigurability of the interconnection is the most important for realizing generality of processing. It is implemented by rewriting the CGH on the PAL-SLM by using liquid crystal display controlled by a personal computer. Consequently, practical functions of interconnection are ob- tained and demonstrated.

Recently, a surface emitting laser diode (SELD) array for integration of the optical

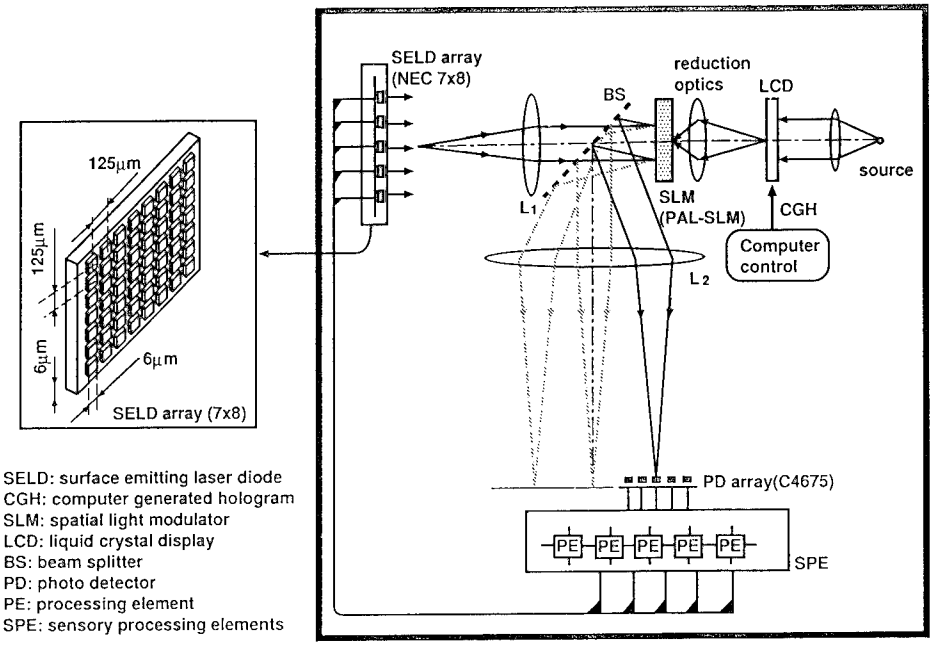


Fig.5 Parallel processing system with reconfigurable optical interconnection

interconnection is examined in the system and the computing capabilities of the whole system for parallel processing including usage of the reconfigurable inter- connection are demonstrated. The block diagram of the system is shown in Fig. 5. The SELD array used in the system, which is developed by NEC Corp. [5], is based on VC-VSTEP technology and has 7 × 8 LD's (wavelength: 975nm, output power: 1mW, pitch: 125 μ m) in a chip. In order to match with the SELD array, the PAL-SLM used in the system is developed for that wavelength. The PD array is 16 × 16 PIN photo diode array which has parallel amplified output. The sensitivity of the PD array for the wavelength is about 0.15A/W. By using these optoelectronic devices, fully parallel operation without scanning operation has been implemented. The photograph of the system is shown in Fig. 6. In the system, 16 × 16 = 256 PEs are implemented as parallel PEs in the SPE. The part of optical system is shown in Fig. 7. The PAL-SLM modulates the light from the SELD array.

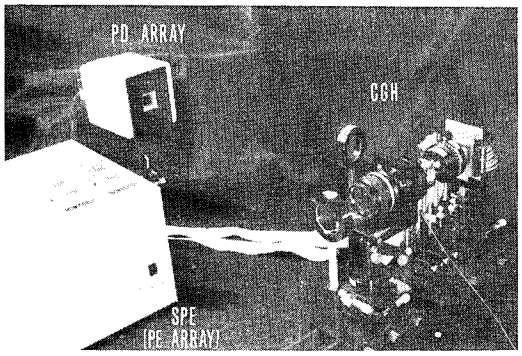


Fig. 6 Experimental system

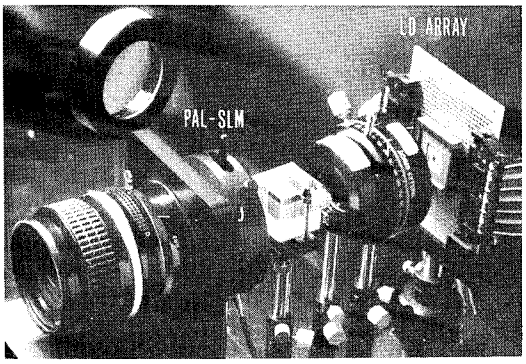


Fig. 7 Optical subsystem

A liquid crystal computer display (640 × 400) for OHP is used as the display r the CGH. It can be controlled in pixel wise by a computer. An example of interconnection pattern is shown in Fig. 8. This pattern is one-to-four interconnects from four LD's. One column, which includes four spots, is interconnected with the same LD.

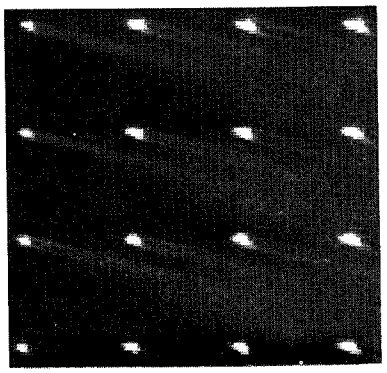


Fig. 8 Example of interconnection

2.4 Parallel processing algorithm

In order to evaluate processing performance of the system, parallel computation of matrix-vector products as an example have been carried out on the system. In the case that data have been already loaded to the PE's, a flowchart of the calculation is shown in Fig. 8. The calculation requires three interconnection patterns shown in Fig. 8. In the present system, changing time for

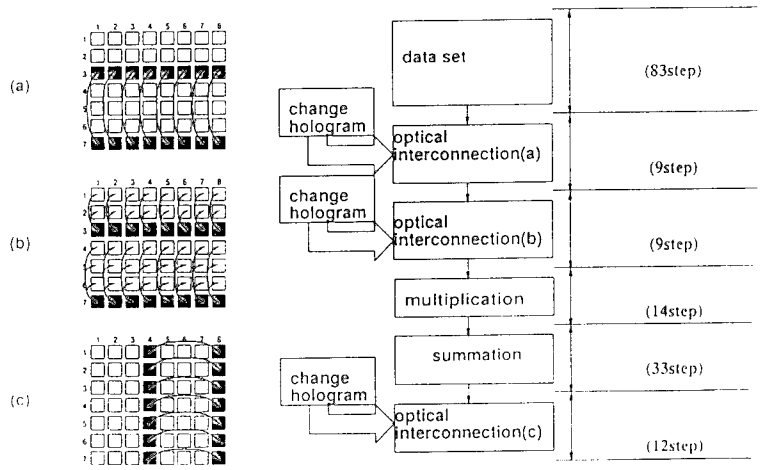


Fig. 9 Algorithm of matrix–vector products and the inter connection patterns

interconnection pattern is not so fast because of the speed of the LCD. Except this point, the speed is limited by the speed of the PD array (about 19kHz) because of high gain amplifiers. The nember of processing steps is decreased to the half of the case without the optical interconnection.

3. Conclusion

Parallel processing architecture for realizing an integrated optoelectronic system and some experimental results are shown. Reconfigurable diffractive interconnects using a surface emitting laser diode array and a phase only modulation type spatial light modulator are implemented and demonstrated. The system has general purpose computing capabilities using the compact programmable processing elements and the reconfigurable inter-connection. The design concept of the system can lead to an integrated optoelectronic parallel computing system that is a real smart pixel system

The author thanks A.Kirk of Vrije Universiteit Brussel for his contribution when he was a research fellow at University of Tokyo, and also thanks T.Tabata, T.Ishida, M.Naruse, H.Yamamoto, Y.Nakabo, and N.Terada of University of Tokyo for their assistance. As for the devices in the system, the author thanks K.Kasahara and his research group of NEC Corporation for their kind support in supplying SELD array, and thanks Y.Suzuki, T.Hara, and their group of Hamamatsu Photonics K.K. for their assistance in setting the PAL-SLM.

References

[1] Ishikawa, M., Morita, A. and Takayanagi, N., 1993, Optical Computing Technical Digest 1993 (Optical Society of America, Washington, D.C.), **7**, 272-275.

[2] Kirk, A., Tabata, T., Ishikawa, M. and Toyoda, H., 1994, Opt. Comm., **105**, 302-308.

[3] Kirk, A., Tabata, T. and Ishikawa, M., 1994, Appl. Opt., **33**, 1629-1639.

[4] Yoshida, N., et al., 1993, Spatial Light Modulator and Applications Technical Digest 1993 (Optical Society of America, Washington, D.C.), **6**, 97-99.

[5] Kajita, M., et al., 1994, Jpn. J. Appl. Phys., **33**, 859-863.

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Free Space Holographically Interconnected Counter

R.J. Feuerstein[1], D.C. O'Brien, A. Fedor, C.C. Mao, L.H. Ji

Optoelectronic Computing Systems Center, University of Colorado Boulder, CO 80309-0525 Phone: 303-492-7077, 303-492-3674 (FAX)

Abstract. We have constructed a simple 4-bit counter using holographic interconnects, microlens arrays, linear arrays of vertical cavity surface emitting lasers (VCSELs), and a CMOS detector array. It was packaged using Spindler and Hoyer microbench components. Design and performance of this prototype system are discussed.

1. Introduction

We are working on various architectures for three-dimensional optoelectronic computers [Neff]. They all share free space holographic interconnects from one planar optoelectronic processing board to another. Systems of this type avoid the interconnection bandwidth bottleneck of systems with strictly in-plane electronic interconnects. We will describe the results of an experiment to construct a system using this technology. This is also the first step in the development of a testbed for evaluation of source and detector arrays, holograms and the necessary optomechanics.

2. Overview

Figure 1 is a schematic of the experiment. The system implements a simple 4-bit counter using two optical assemblies with associated electronics. The optical assemblies are designed to demonstrate both fan-out and fan-in operations. The upper arm has 8 inputs which fan-in to 4 outputs, performing an optical OR operation , and the lower fans-out 4 inputs to 8 outputs. The electronics perform the Boolean expressions required for the counter, with the right half of the system counting the odd numbers and the left the even. On each clock, or increment signal the left counts 1, 3, 5, 7, 9, 11, 13, 15, 1... and the right 0, 2, 4, 6, 8, 10, 12, 14, 0...etc. In this paper we describe the optical system. There are two separate optical arms, with 8 and 4 input channels respectively. VCSEL arrays are used as the input sources, and the light from these is collimated using a microlens array. The collimated beams illuminate an array of Binary phase Computer Generated

[1] E-mail:robertf@fred.colorado.edu

Holograms (CGHs), one for each input channel. The light from the holograms is then Fourier transformed with an achromat, and an array of photodetectors is placed at the Fourier plane to detect the output from the system. The components are mounted with standard Spindler and Hoyer microbench parts. Both the source and receiver assemblies are mounted on x-y adjustable mounts to allow for alignment of the light beams with the detector pads. The complete optical path, using a 10mm focal length achromatic transform lens, is less than $2cm$ long.

3. System Design

The source assembly consists of the VCSEL array, microlens and holograms (Figure 2). The VCSELs are arranged in a linear array on a $250\mu m$ pitch . They operate at approximately $850nm$ and emit a close to Gaussian beam with a (e^{-2}) half-angle of 0.1 radians. The peak output power is approximately $1.8mW$ ($V = 5.5V$ $I = 9.0mA$). The beams are then collimated using a commercially available microlens array, which has a focal length of $560\mu m$ and a $250\mu m$ pitch. A spacer is glued to the VCSEL surface so each VCSEL lies at the focus of a microlens when the array is glued to the spacer. The lens array is then aligned with the VCSEL array using an active technique: The VCSELs are turned on and the output beams monitored with a CCD camera. The lens array is placed on the spacer, then translated until the best optical field is obtained. A UV curing epoxy is used to attach the array. The hologram array is spaced approximately $300\mu m$ from the microlens array using a spacer ring glued to the VCSEL package (this allows reuse of the collimated array).

There are two spatially variant CGHAs, one with 4 holograms and the other with 8. The first CGHA performs fan-out of a single input beam to two output spots while the latter performs fan-in of two beams to one spot (Figure 3). The holograms were encoded using a simulated annealing algorithm [Feldman, Yoshikawa] . The simulated annealing algorithm was used to encode 64×64 pixel images which were replicated to create a 128×128 array. We used an optical lithography system to fabricate a master set of holograms with a resolution of $2.2\mu m$. The holograms are then copied into a phase medium by contact copying the master holograms onto a quartz substrate and etching using a reactive-ion etching system [Zhang]. This produces a two-level phase hologram with a theoretical best efficiency of 40%. Measured efficiencies of these CGHAs varied from $20-26\%$. This difference may be due to roughness of the quartz where it is etched. The quartz CGHA is aligned to the VCSEL beams using an active alignment technique similar to that described above. Figure 4 shows the measured spot pattern for the fan-in CGHA when four of the eight lasers are on.

A 2×8 photodetector array was used to detect the output from the system. This was designed locally and fabricated in $2\mu m$ CMOS by the MOSIS fabrication service. The array consists of a high sensitivity 1×8 array and a 1×8 array with lower sensitivity. Signals from the detectors are amplified and thresholded to provide TTL level outputs from the detector IC [Mao]. The high sensitivity array was used to form the output of the system and required $> 40mW$ of optical power to generate a TTL level output pulse. The detectors are $200 \times 200\mu m^2$ on a $250\mu m$ pitch. A typical light spot size at the detector for this system is approximately $100\mu m$, so the large detector area allows for shifting of the spots due to wavelength variations, as well as minor alignment errors and distortions caused by the optical system.

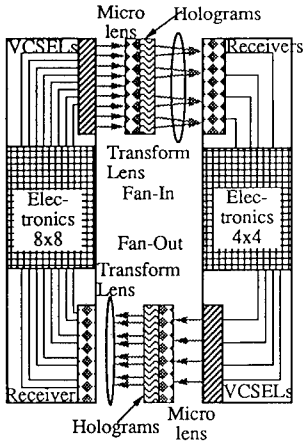


Figure 1. Overview of complete holographically interconnected counter.

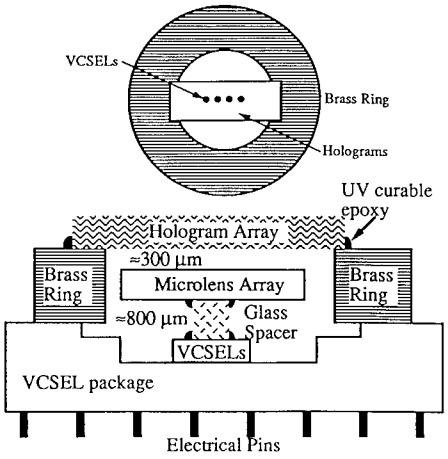


Figure 2. Packaging of one of the VCSEL arrays with the microlens and hologram arrays.

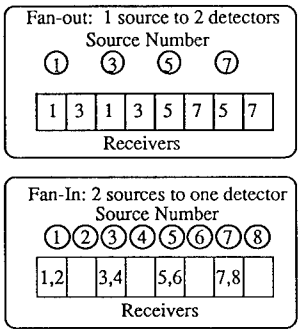


Figure 3. Hologram design patterns for fan-out and fan-in.

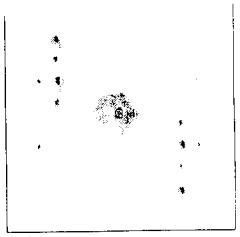


Figure 4. Measured spot pattern of the Fan-in CGHA with four lasers on. Note both +1 and -1 order as well as 0 order are visible.

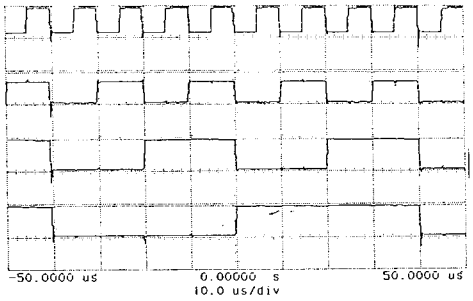


Figure 5. Oscilloscope trace of the four bit output counter for even numbers. Bit_0 is not shown as it is 0 all the time. The upper trace is the clock or increment signal for the counter. The next three traces are the count bits.

4. System operation

The system speed is limited by the minimum power required for the receiver chips to generate a TTL level output pulse. This in turn depends on the efficiency of the holograms, and the VCSEL output powers. Figure 5 shows the trace of the four bit output counter for even numbers. Bit 0 is not shown as it is 0 all the time. The upper trace is the clock or increment signal for the counter. One can see all even numbers shown for this 100 kHz clock rate. This was the highest speed achievable for stable operation of the counter. There were a number of VCSELs whose power dropped so much that they were no longer able to trigger the receiver chip. The appropriate signals were hard wired electronically to get around this problem. The system is quite stable, probably due to its large beam position tolerance.

5. Conclusions

We have demonstrated a simple 3-D holographically interconnected counter using VCSELs. The system can be used for further evaluations of design algorithms and fabrication technologies for CGHAs, as well as receiver arrays and VCSELs by simply interchanging the appropriate parts. This can be easily accomplished since standard optomechanical mounts and electronic sockets are used throughout. This valuable first step exposed many of the issues involved in the design, packaging, alignment and assembly of a complete 3D system.

Acknowledgements

This work was supported by the National Science Foundation and the Colorado Advanced Technology Institute. We would like to acknowledge the assistance of the following people: V. Morozov, C. Stirk, E. Tang, K. Geib, S. Feld, C. Wilmsen, H. Temkin, J. Neff, D. McKnight, K. Malone, M. Young, and M. Popovic. In addition we would like to thank Photonics Research Inc. for their generous loan of a VCSEL array.

References

[Feldman] M.R. Feldman and C.C. Guest, "Iterative encoding of high-efficiency holograms for generation of spot arrays", Opt. Lett., 14, 479-481 (1989).

[Mao] C.C. Mao, L.H. Ji, D. McKnight, R.J. Feuerstein, J. Neff, "CMOS 2×8 photoreceiver array for free space holographically interconnected counter", IEEE/LEOS Summer Topical Meeting on Smart Pixels, July 11-13, 1994, Lake Tahoe, Nevada.

[Neff] John A. Neff, "Planned development of a 3D computer based on free space optical interconnects", Optoelectronic Interconnects II, Proc. SPIE, Vol. 2153, SPIE, Bellingham, WA USA (1994).

[Yoshikawa] Nobukazu Yoshikawa and Toyohiko Yatagai, "Phase optimisation of a kinoform by stimulated annealing", Appl. Opt., 33, 863-868 (1994).

[Zhang] J. Zhang, C. Xu and S. Gao, "Optical interconnetion with phase-only masks by means of reactive ion etching", Opt. Eng., 32, 1866-1870 (1993).

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Versatile Compact Image Processor with Optical Feed-Back using Photopolymer and Ferroelectric Liquid Crystal on Amorphous Silicon

Pierre Cambon

GOSC, Télécom Bretagne, Département d'Optique
BP 832, 29285 Brest Cedex, France.
Fax +33 98 00 10 25, E-Mail: cambon@gosc.enst-bretagne.fr

John Sharpe and Kristina M. Johnson

Optoelectronic Computing Systems Center
University of Colorado, Boulder, Colorado, 80309-0525
Fax +1 303 492 3674

Abstract. An optical image processor with optical feedback and gray scale capability compactly organized around bistable binary amorphous silicon and ferroelectric liquid crystal devices and a one lens correlator using reflective multiplexed photopolymer hologram is briefly presented. First experiments relative to the interconnection holograms are reported.

1. introduction

A broad class of efficient and widely used low level image processing operations are the non-linear filters such as ranked-order filtering and morphological operations. They can be performed on binary images by convolving them with a binary or gray scale kernel and thresholding the output. Extension to gray scale images is performed by considering them as a finite collection of binary slices, processing each slice as previously and finally summing the processed slices to obtain the processed gray scale image. [1].

2. Proposed architecture

Optically addressed spatial light modulators (OASLM) using Ferroelectric Liquid Crystal (FLC) are considered here. VLSI backplane OASLM [2] could also be used. The device is a mosaic of nine independently driven aSi:H/FLC bistable OASLMs with fixed high

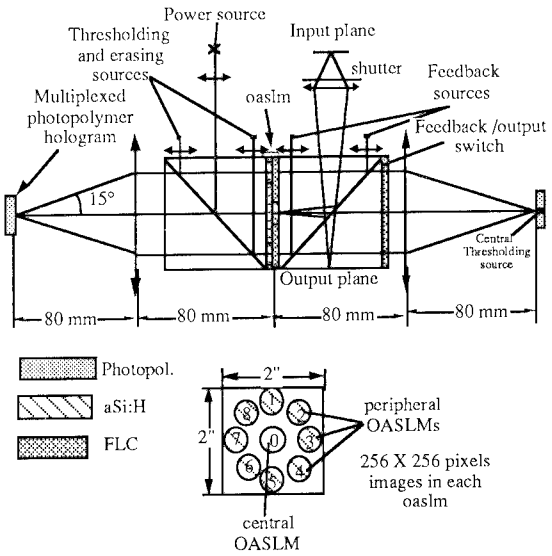


Figure 1. *The processor architecture*

threshold [3] (250 $\mu W/mm^2$ at 633 nm and 2.5 kHz frame rate) realized on the same 2" x 2" optical flats . It is illuminated on both sides through the polarizing beam splitters (see Figure 1).

The aSi:H/FLC interface is reflective (pixelated aluminium). The central OASLM faces the opposite direction compared to the peripheral ones. The right side performs the feedback by using a 80 mm focal length symmetrical lens , a multiplexed reflective phase hologram and independently controlled feedback laser diodes (1 mW, 670 nm). The left side performs correlation from the central OASLM to the peripheral one's by using a power source (10 mW - 633 nm HeNe Laser) a variable light biasing for threshold control and erasing (equally 1 mW - 670 nm laser diodes) and a lens identical to the right one . The multiplexed reflective holograms recorded at 514 nm on a separate bench use recent Du Pont photopolymer [4] materials and color tuning films [5] to shift replay wavelength to 633 nm for left hologram or 670 nm for right hologram . The gray level input 256 x 256 pixels image in the input plane is formed by an input lens on the 8 x 8 mm useful aperture of the central OASLM. The biasing light from the central right thresholding / erasing diode through the right hologram is incoherently added to the image to determine the threshold level and the result thresholded and stored as a binary image in the FLC. Using the power source this slice is convolved simultaneously with the eight kernels recorded in hologram chosen for the application. One or a few of the convolved images are stored on a peripheral OASLM while the previously stored images on the other OASLMs are retained. The other slices of the images are successively formed, processed in the same way and stored in peripheral OASLMs. The final summing operation is performed by illuminating simultaneously with the feed-back diodes the OASLMs where the previously processed slices are stored. The input image is obstructed using an input

shutter. The gray level output image is directed onto the Output plane by activating an FLC quarter waveplate which rotates the polarization of the reflected beam by ideally 90 deg.. Binary image processing can be performed in the same way but additionally in that case operations are easily cascaded by using the feedback diodes.

3. Holograms

3.1. Photoplolymer hologram

Reducing the power of the main optical source requires to use high reflection efficiency hologram. So phase holograms are used and photopolymer material are chosen because of their low cost, very simple process, performances on replay and stability over time. A drawback of these material is that they may not be recorded in-situ in practice because of their too low sensitivity (15 to 30 mJ/cm^2). The process consists just a 1 min. UV light exposure and a $1^{1/2}$ hour baking at 100 Celsius deg.. During the baking, the index wave in material slightly shrinks (about 2%). The photopolymer material is available on non rigid mylar with a removeable cover sheet. The tacky photopolymer is laminated on a 2" x 2" glass substrate to be easily handled. Some experiments have been performed to measure reflection efficiency versus external incidence of a quasi unslanted grating recorded in HRF 600 material 20 μm thickness from Du Pont.

The recording was performed using a Argon Ion laser at 514 nm with 60 mJ/cm^2 on each beam (6 mW/cm^2 during 10 s). The hologram is rotated from the bissector plane of the two beams by the angle required to achieve a 1.5 deg. external reflection angle. The beam angle is 150 deg.; it has been computed using a Kogelnik based program so that assuming a 2 % shrinking during the heat treatment, the maximum reflection efficiency at the same wavelength occured for a normal incidence with the external reflection angle of 1.5 deg.. This program also computes how much must the hologram should be rotated from the bissector plane. An external reflection efficiency more than 85 % is obtained around normal incidence so the expected high reflection efficiency and shrink compensation is verified. Taking into account the normally reflected amount of power on the hologram surface (4%), internal reflection efficiency of 90% can be estimated .

3.2. Hologram multiplexing

Sequential exposure multiplexing has been first chosen because it requires only simple optical set-up. The recording optical set-up is the same as discussed above, the hologram is just rotated by 45 deg. around its axis between two successive exposures. Considering two gratings engraved the gratings vectors are labelled for first exposure $\mathbf{K_1}$ and second $\mathbf{K_2}$. Observing the reflected beams, the $main_1$ beam reflected by $\mathbf{K_1}$ is only 24% of incident and $main_2$ beam reflected by $\mathbf{K_2}$ is 12%, the reflection angle δ is 2 deg..

Although the exposure times (2.5 s) were equal for both gratings, the firstly recorded give twice more reflection efficiency than the second one. This is due to a non uniform photopolymerisation of monomer during the time [6, 7]. Brightness balancing could be reached by increasing exposure time from first to last records. The main part of the remaining energy is located on two parasitic beams that are in the same plane as the main ones. The $main_2$ reflected beam is also reflected by the $\mathbf{K_1}$ grating that is not off Bragg enough. Then this beam traveling to the back side of the

material is reflected all along the travel by $\mathbf{K_2}$ for the same reason as previously giving the $cross - reflected_{12}$ beam. The same process occurs for the $main_1$ beam giving $cross - reflected_{21}$. The $cross - reflected_{12}$ beam is as strong as $main_2$ (12% reflection efficiency) and $cross - reflected_{21}$ beam is 4% Other very weak beams can be observed and depicted as back-reflected, 2nd order, cross-cross-reflected, back-back-reflected.

The 1st order cross-reflected beams must be lowered as much as possible. Qualitatively this can be done by using a large δ angle and increasing the material thickness. The analytical Kogelnik theory can't be used to model the coupling between a main reflected beam and its adjacent gratings vector because the theory is not valid when grating vector is not lying in the incidence plane. The three dimensional numerical theory [8] from Moharam and al must be used. It describes rigorously this problem of conical diffraction. Some experiments not reproduced here have been done by using various δ angle 10, 20 and 30 deg. . The angle must be as high as possible and not below 10 deg.. 15 deg. is a good compromise between this constraint, the compactness and feasibility of lens.

4. Conclusion

An architecture for an image processor based on a symmetrical multichannel looped correlator using photopolymer, a-Si:H and FLC is presented with possible dimensions. In addition to iterative binary image processing, gray-scale processing using threshold decomposition with possibly gray-scale different slice-kernels is shown to be easily implementable. Use of Du Pont photopolymer materials to produce the required reflective multiplexed phase holograms is experimented. The processor may be feasible with low cost and available devices.

References

[1] J.W. Hereford, W.T. Rhodes *"Non linear optical image filtering by time-sequential threshold decomposition"* Optical Engineering, vol 27, No 4, (1988), pp 274

[2] K.M. Johnson, D.J. McKnight and I. Underwood *"Smart spatial light modulators using liquid crystal on silicon"* IEEE J. of Quantum electronics, vol 29, No 2, (1993), pp 699

[3] J.B. Chevrier, P. Cambon, R.C. Chittick, B. Equer *Use of nipin aSi:H structure for bistable OASLM"* J. of Non-Crystalline Solids, 137 and 138, (1991), pp 1325-1328.

[4] R.W. Brandstetter, N.J. Fonneland *"Photopolymer elements for an optical correlator system"* Proceedings SPIE on Photopolymer device physics, chemistry and applications II (1991) vol 1559, pp 308.

[5] S. A. Zager, A. M. Weber *"Display holograms in Du Pont's Omnidex films"* Proceedings of SPIE on Practical Holography V, vol 1461, (1991), pp 58-67

[6] K. M. Johnson, L. Hesselink and J. W. Goodman *"Holographic reciprocity law failure"* Appl. Opt. vol 23, (1984) , pp 218

[7] K. Curtis and D. Psaltis " *Recording of multiple holograms in photopolymer films"* Appl. Opt. vol 31, No 35, (1992), pp 7425-7428

[8] M. G. Moharam and T. K. Gaylord *"Three dimensional vector coupled-wave analysis of planar-grating diffraction "* J. Opt. Soc. Am. vol 73,No 9, (1983), pp 1105-1112

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Cascaded optical data transfer through a free space optical perfect shuffle

M W Derstine, S Wakelin and K K Chau

Optivision, 4009 Miranda Ave., Palo Alto, CA 94304, USA

Abstract. We describe the design, fabrication and testing of an optical interconnect system constructed from a two dimensional array of symmetric self-electro-optic effect devices and a free space perfect shuffle module.

1. Introduction

The perfect shuffle has been proposed as an efficient interconnect that optics can do well[1]. However, to our knowledge, no free space optical system has been constructed that tests the operation of a shuffle between optical logic elements. Most previous work has focused on the development of the perfect shuffle optics without regard to interfacing to the optical logic elements.

An optical system designed to be used to implement sequential logical processing has been fabricated and used to demonstrate transfer of optically encoded data from an optical logic array though a perfect shuffle interconnect and back onto the logic array. An array of self-electrooptic effect devices (S-SEEDs) [2] serve as the logic array and the perfect shuffle is similar to that developed at Heriot-Watt University[3].

2. System Design

The basic architecture is shown in Figure 1. The input, output and power supply optics are not shown. The primary optical circuit is implemented in a loop. The S-SEED array is divided up so that each column corresponds to stage of the interconnect. The output of the S-SEED array is routed through a one-dimensional perfect shuffle module and then imaged back onto the S-SEED array displaced by one column. This produces a multistage interconnect with perfect shuffles between each stage.

The S-SEED is arranged to so that these are used as 8 columns of 16 devices each. This size was selected to match the field of a commercial laser objective used to image the S-SEED. Spots to read out the array are generated by two lasers, one for the even columns and one for the odd columns. The necessary spot pattern is generated by a binary phase diffractive element.

The perfect shuffle employs the basic split-and-shift approach to implement a theoretically lossless interconnect[3]. The other major advantages are that it can be constructed from readily available components, it could be aligned in stages, it has identical loss (theoretically zero) for all the beams through the system, it could be enhanced to provide internal fanout, and would work with data in differential representation.

A schematic of the hardware used for the perfect shuffle is shown in Figure 2. The collimated p polarized beam is transmitted through the polarizing beam splitter (PBS) and quarter wave plate (QWP) at the upper left of the figure. The transform lens focuses the beam,

and the focal plane array is split into two halves with a knife edge mirror placed perpendicular to the direction of propagation. Half of the array is reflected, the other half passes by the edge and is recollimated and transmitted through another QWP. The beams are directed using total internal reflection in a 45° prism so that beams from both portions of the array are on the same axis but propagating in opposite directions. Transform lenses are used to focus each of the half-arrays onto opposite sides of a second mirror patterned with stripes, at which plane the arrays are interlaced and then exit the module in the same polarization at the lower right PBS.

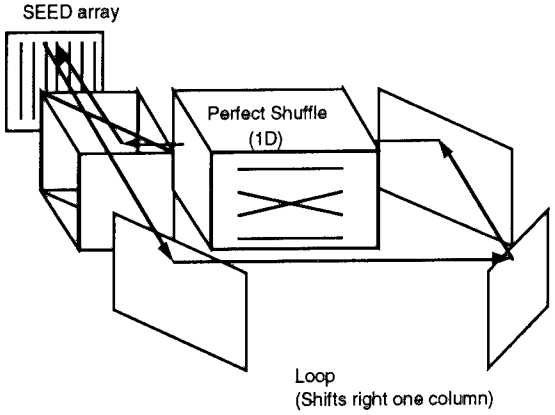


Figure 1. Interconnect architecture

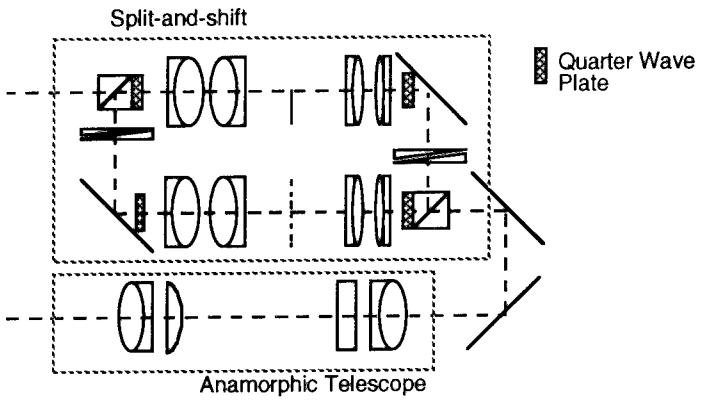


Figure 2. Schematic of hardware used in perfect shuffle interconnection module.

The design of the shuffle module incorporated a factor of $\sqrt{2}$ magnification in the imaging operation. The reason for this is due to the 2:1 magnification that occurs between input and output array pitch in one dimension when unity magnification is used. It is considerably easier to design an anamorphic telescope that will magnify the ratio of the object sides by a factor of two, rather than strictly 2:1 magnification. Hence, a magnification of $\sqrt{2}:1/\sqrt{2}$ used in the anamorphic telescope determined the relay lenses that were used in the shuffle implementation. The anamorphic telescope used compound lenses constructed from a cylindrical lens and spherical lens. The two compound lenses are then oriented with cylinder axes perpendicular. The configuration of the anamorphic telescope was designed to have telecentric input and

output, i.e. the pupils of the spherical/cylinder lens combination and the single spherical lens were coincident. The alignment task is considerably easier than if a non-telecentric design had been used.

The input to the cascaded of data is provided by an electrically addressed linear array of surface emitting lasers. It is used to set the first column of S-SEEDs. The VCSEL array is directly driven with an open-collector TTL circuit that is also used to sequence the column and programming lasers.

Figure 3 shows a schematic of the basic system. Light from the even and old column laser is combined and then the array of spots is generated. These spots are then directed through a beam combining system and focused onto the S-SEED array. The reflected light from the S-SEED is shuffled and then has the spot pitch corrected by an anamorphic telescope before being combined and imaged onto the S-SEED array. The entire system is doubly telecentric for improved alignment tolerance, except at the intermediate focus of the anamorphic telescope. The system was constructed using the milled baseplate approach[4].

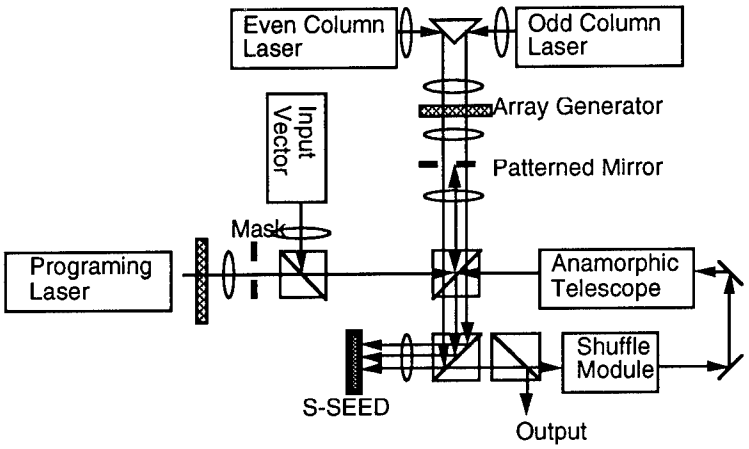


Figure 3. Block diagram of the cascaded system.

3. Operation

The system has been assembled and aligned and demonstrated cascaded data transfer with 100% correct operation with a switching time of 1.5 μs. Figure 4 shows a computer enhanced picture of the light reflected from the S-SEED plane. The system is being clocked at 100kHz and the data from the first column of S-SEED devices is cascaded through the succeeding columns. Notice that because S-SEEDs are differential devices, each column is made up of two spots. The first column was set with the upper eight devices in one state and the lower eight in the other. This light is shuffled in one dimension and cascaded to the second column. The light reflected from the second column of S-SEEDs shows the results of the the shuffle and the invertion operation of the S-SEED. Light from the third and subsequent columns are cascaded and inverted in a similar manner. Note that after 4 (= $\log_2 16$) cascades the data has been returned to the original configuration.

For this system the key limiting factor for the system speed was not variation in cascaded beam powers[3], but rather crosstalk between the readout and cascaded arrays of beams. The readout light was generated with a binary phase grating designed to illuminate alternate

columns. However, some light from orders that were supressed still reached the S-SEED plane. These orders were aligned with the other columns of S-SEEDs; the columns where the light from the perfect shuffle module was focused.

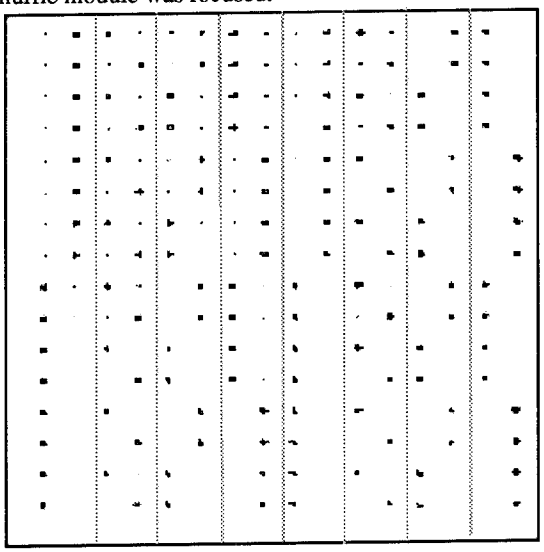


Figure 4. Computer enhanced image of the focal spot array on the device plane, showing leftmost column shuffled at each subsequent column.

4. Conclusions

In conclusion, we have demonstrated the first, to our knowledge, free space optical system that cascades optical data through a perfect shuffle between logic elements. We have shown its 100% correct operation and found the operating speed to be limited by crosstalk on the array. This system demonstrates that extensive free space optical systems can be constructed, aligned and demonstrated. However, difficulty cascading back to the same image plane suggests that interconnect architectures that can be more easily implemented, such as the butterfly or Banyan, are preferable to the perfect shuffle for free space applications.

5. Acknowledgments

We wish to thank Robert A. Morgan of AT&T for providing the laser array. This research is supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Air Force Office of Scientific Research under Contract No. F49620-92-C-0050. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

References

[1] Lohmann AW, Stork W, Stucke G, May 1986*Applied Optics*, Vol. 25, No. 10

[2] Lentine A L and Miller D A B, 1993 *IEEE J. Quant. Electron.* **29**, 655-669

[3] Wakelin S., 1992 "Design and construction of a free-space digital optical information processor," Ph.D. Thesis, Heriot-Watt University

[4] Brubaker J et al., 1991*Proc. SPIE* **1533** 88-96

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optoelectronic Implementation of Tomographic Sorting

Pericles A. Mitkas and Leo J. Irakliotis

Optoelectronic Computing Systems Center and the Department of Electrical Engineering
Colorado State University, Fort Collins, CO 80523, USA

Abstract. A parallel sorting algorithm and its implementation on a 3-D multistage optoelectronic sorting network are presented. The network operates on bit-slices of words and combines regular structure and simple interconnections between stages.

1. Introduction

One of the advantages of parallel optical storage systems, such as volume holographic memories, is the ability to store and retrieve two-dimensional blocks of data in parallel. Any attempt to process these blocks in a traditional fashion would result in bottlenecks between the memory and the processing units of the electronic computer. To fully exploit the inherent parallelism of optical memory, we must design new processing architectures capable of handling large blocks of data in parallel. Such architectures are expected to be based on multiple stages of fine-grain optoelectronic processing elements [1].

Sorting is a common operation in a large range of applications and can be extremely time–consuming for long lists of words [2]. In this paper we present tomographic sorting, a parallel pipelined algorithm that can be mapped directly on a three-dimensional computer architecture that uses optical interconnections to propagate data and control bits from one processing plane to another. The algorithm is a combination of the odd-even transposition sort and the standard radix sort [2]. We begin our discussion by describing the algorithm and the structure and functionality of the tomographic sorting network. Then we discuss how this network can be implemented as a multistage architecture similar to the three-dimensional optoelectronic computer under development at the Optoelectronic Computing Systems Center in Colorado. Tomographic sorting can serve as a test application for this computer.

2. Sorting Network

The tomographic sorting network operates on a list of N words with M bits each and has M processing stages. A two-dimensional version of the network is shown in Figure 2. The N words are stored as M bit slices (thus the term tomographic) with the k-th slice containing the k-th significant bit of every word. All the bit slices are pipelined into the network, and remain there for a fixed number of clock cycles until they are sorted and ready to be transferred out and into a buffer. Sorting is performed by rearranging bits on the same slice so that, eventually, words are reshuffled and end up in ascending order from top to bottom. After the bit slices are loaded into the sorting unit, only control signals are transmitted from one stage to another to indicate any necessary bit exchanges.

Two types of cells are present in the tomographic sorting network: data and control cells. Each cell contains a flip-flop to hold the value of a bit. Rows of data cells are interleaved with rows of control cells as shown in Figure 2. The control cells are basically bitwise compare-and-exchange modules with some additional control lines. Optical signals propagate along the horizontal direction, while electronic signals traverse the vertical direction.

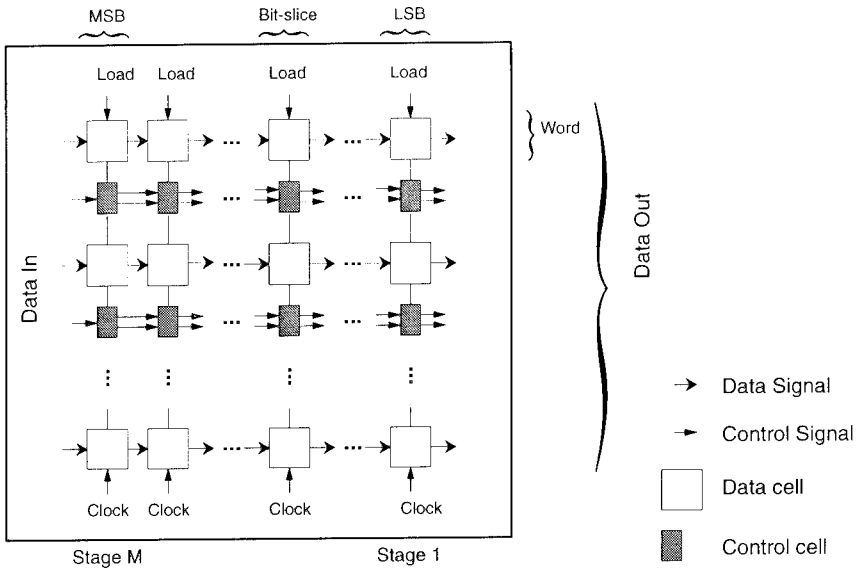


Figure 1. Tomographic Sorting Network in two-dimensions.

A column of data and control cells constitutes a stage. The data cells in every stage will receive and hold the contents of the corresponding bit slice. Notice that all signals flow from left to right and that the stages are numbered in descending order from M to 1 (from left to right) to coincide with the bit order in a word. The number of data cells on each stage is equal to N, while the number of control cells is N–1. The number of stages that are required is equal to the length of each word in bits. The interconnection pattern between any pair of stages is a straightforward one-to-one mapping. Another characteristic of this network is its regular structure, which also contributes to a simple implementation and permits easy scaling up to accommodate larger sets of words.

3. Algorithm Description

As mentioned earlier, the sorting algorithm is a combination of the odd/even transposition and the radix sort algorithms. The process is initiated by looking at the most significant bit slice and trying to partition it into two groups: the 0-group at the top and the 1-group at the bottom. Every pair of adjacent bits that are not equal allows a decision on the relative magnitude of two words to be reached. If the two words must be exchanged, the process is completed in a pipelined fashion with an unconditional exchange control signal gradually forcing all the lower significance bits of the two words to exchange positions. If a decision cannot be reached at the MSB slice because the two bits under consideration are equal, the decision is deferred to the next most significant bit slice by sending a conditional exchange control signal to the next stage. This will force the next control cell to perform a bit comparison to determine the relative magnitude of the two words. While an unconditional or conditional exchange signal ripples through all the stages of the network, a new set of tests can be initiated at the MSB slice. The outcome of these tests and the subsequent execution of any action required will not affect the completion of the previously initiated exchanges that may still continue to occur towards the final stages of the pipeline.

Three distinct phases are present in the tomographic sorting operation: (i) data loading, during which the bit-slices are pipelined into the sorting stages; (ii) sorting, and (iii) data unload-

ing, during which the bit slices of sorted words exit the pipe. During the first phase, bit-slices are pipelined into the sorting unit starting with column 1. In one cycle, an entire column is transferred in parallel from stage k to stage k–1. After M cycles, all the M bit slices will be loaded into the data cells of the sorting unit. At the second phase, the operation begins by examining pairs of bits in the control cells of the MSB bit slice in a way similar to that in the bubble sort algorithm. If two bits are found to be out of order, they are exchanged and a forced exchange control signal is generated and pipelined to the remaining stages. Since the words are sorted in ascending order from top to bottom, the combination 1,0 in a pair of bits will initiate this process. If two bits are equal, no decision can be made on the relative magnitude of the words, so the control cell will issue a compare-and-exchange signal to the corresponding control cell in the next stage. If two bits are in the right order (i.e., 0,1), no action will take place and both control lines to the next stage will be low. At the end of this phase, the words will be stored in ascending order from top to bottom, i.e., the minimum value will be located in the first row of data cells, while the maximum value may be found in the bottom row of the array. Finally, during the third phase, the bit slices of the sorted words can be flushed out of the pipe, one slice at a time, in a total of M cycles. Since data move from left to right, the first and the last phases can overlap when repeated sorting operations must be performed.

4. Discussion

Despite the fact that special purpose numerical sorting processors and networks have relaxed the burden of sorting, when the size of these networks becomes large, the latency of the data input/output process and the complexity of the interconnections between stages become serious bottlenecks that render their implementation impractical. An optoelectronic sorting network with straight pass optical interconnections for interstage communications, such as the one presented in this paper, lends itself to a simpler implementation.

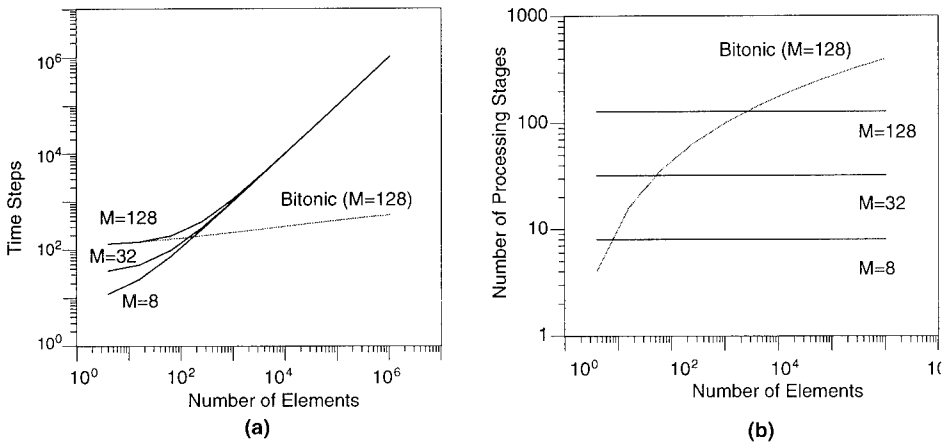


Figure 2. (a) Time complexity and (b) space complexity of the bitonic sorting network for M=128 versus three tomographic sorting networks of different sizes.

Tomographic sorting can be completed in O(N) time steps. Although the tomographic sorting network is not faster than the bitonic sorting network, as shown in Figure 2, it requires fewer and simpler stages as the number of elements increases. An optical implementation of the bitonic sorting network [3] becomes prohibitively complex when the list to be sorted exceeds

100 elements. The reduced space complexity of the tomographic sorting network is its main advantage over other sorting networks.

The tomographic sorting network can be expanded in three dimensions. In the 3D implementation, the linear bit slices are replaced by two-dimensional bit slices stored as pages in a parallel optical memory. At the end of the operation, the words are arranged in ascending order along a serpentine path from the upper left to the lower right corner of the page. The processing stages here are two-dimensional arrays where interstage communications are carried out optically and intrastage communications are electronic. A general view of a 3D tomographic sorting network based on transmissive stages is shown in Figure 3. This design uses heterostructure phototransistors as the receiving elements and vertical cavity surface emitting lasers as light sources. Data and control logic remains electronic.

We are currently working towards a VLSI implementation of the 2D version of the tomographic sorting network. We have also designed and are implementing a recirculating optoelectronic architecture for tomographic sorting that uses two processing stages and combines silicon detectors, CMOS electronics, and vertical cavity surface emitting laser arrays [4].

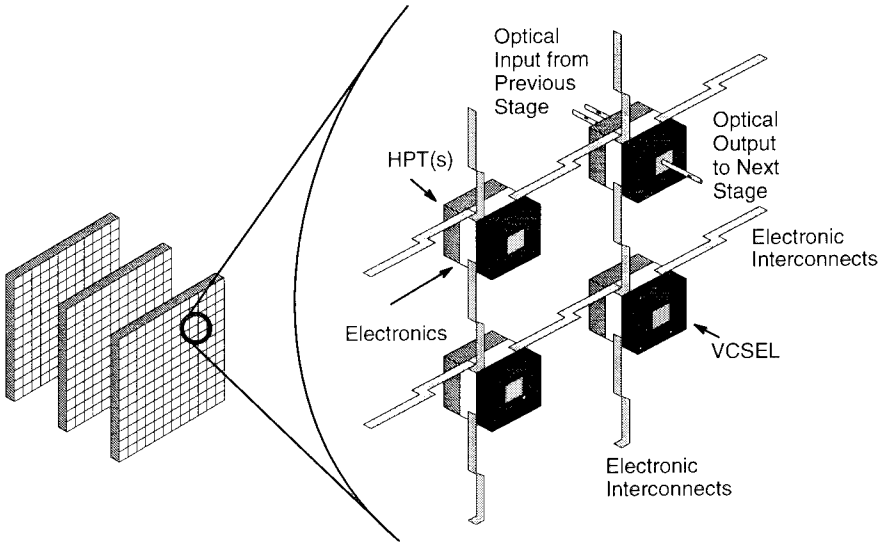


Figure 3. Schematic of the 3-D optoelectronic architecture. HPT: Heterostructure Phototransistor; VCSEL. Vertical-Cavity Surface-Emitting Laser.

Acknowledgement

This work was partially supported by NSF-ERC grant ECD9015128, the Optoelectronic Computing Systems Center, the Colorado Advanced Technology Institute, and the Colorado Commission on Higher Education.

References

[1] P. A. Mitkas, F. R. Beyette, Jr., S. A. Feld, L. J. Irakliotis, and C. W. Wilmsen, "Optoelectronic Parallel Processing with Straight Pass Interconnections and Smart Pixel Arrays," Proc. of the First Intern. Workshop on Massively Parallel Processing using Optical Interconnections, April 26–27, 1994, Cancun, Mexico, pp. 170–181.

[2] S. G. Akl, *Parallel Sorting Algorithms*, Academic Press Inc., Orlando, 1985.

[3] C. W. Stirk and R. A. Athale, "Sorting With Optical Compare-and-Exchange Modules," *Applied Optics*, Vol. 27, No. 9, pp 1721–1726, May 1988.

[4] F. R. Beyette, Jr., P. A. Mitkas, S. A. Feld, and C. W. Wilmsen, "Optoelectronic Bitwise Compare–and–Exchange Modules based on a Silicon / Vertical Cavity Laser Hybrid," IEEE/LEOS Summer Topical Meeting Digest on Smart Pixels, Lake Tahoe, NV, July 11–13, 1994, pp. 70–71.

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Sorting with an Optoelectronic Recirculating Architecture

F R Beyette, Jr., P A Mitkas, S A Feld, and C W Wilmsen

Optoelectronic Computing Systems Center and the Department of Electrical Engineering
Colorado State University, Fort Collins, CO 80523, USA

Abstract: An optoelectronic implementation of bitonic sorting is presented which uses a recirculating architecture to reduce the number of required stages to two. This architecture decreases the mechanical complexity at the cost of system throughput. In addition to the system layout, a bitwise compare-and-exchange module suitable for this architecture is presented.

1. Introduction

With sorting applications varying from knowledgbase and database manipulation to telecommunications switching, it appears certain that efficient and practical methods of implementing sorting will continue to be important. While a wide variety of sorting algorithms have been explored for implementation in both hardware and software, [1] only a few implementations have been proposed which utilize the parallelism associated with optoelectronic processing arrays. [2,3] In both of these previous works, the implementation is based on a straight forward construction of the bitonic sorting network shown in figure 1. Each stage of the sorting network is comprised of latching compare-and-exchange (C&E) modules which perform a bitwise comparison on two input words. The stages are connected with optical perfect shuffle interconnection networks. In theory, the capacity (defined as the maximum number of input words or channels) of these sorting networks is unlimited. However, the number of processing stages required to implement the bitonic sorting algorithm is dependent on the system capacity. Therefore, the realization of a large capacity optoelectronic sorter may be limited by the mechanical complexity required to build and maintain such a large multistage network.

Previously, it has been suggested that the mechanical complexity of a multistage network can be reduced if the data are recirculated through the same hardware. [4] In the early seventies, Stone [5] suggested a recirculating architecture for bitonic sorting. This

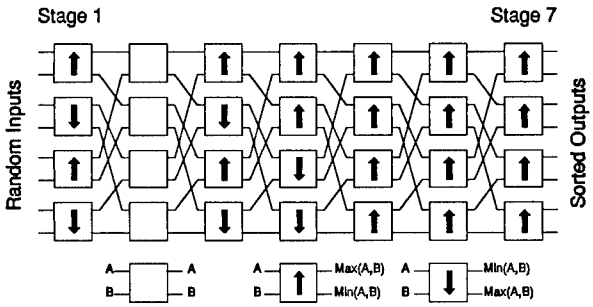


Figure 1. Bitonic sorting network comprised of compare-and-exchange processing stages interconnected with a perfect shuffle interconnection.

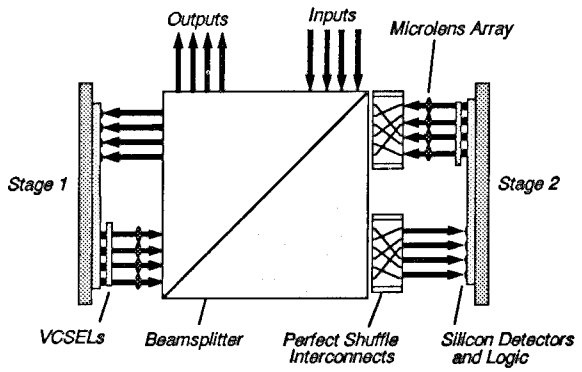


Figure 2 Schematic layout of a recirculating bitonic sorter using only two stages and two perfect shuffle interconnections.

recirculating architecture utilized an N element array of latches connected via a perfect shuffle interconnection network to an $N/2$ element array of C&E modules. The outputs of the C&E modules were then connected in a one to one fashion to the latch inputs. In this paper we propose an alternative recirculating architecture which utilizes arrays of optoelectronic bitwise C&E modules.

2. C&E modules and System Layout

Figure 2 shows the system layout of the proposed optoelectronic recirculating sorter. It is comprised of two stages positioned on opposite sides of a beam splitter. This arrangement is compact and the beam splitter provides a means for convenient optical input and output. Each stage of the recirculating system consists of C&E modules which compare adjacent elements of the input set. The optical outputs of one stage provide optical inputs to the other stage. The optical data signals of each stage are imaged onto multifaceted holograms using a micro lens array. The multifaceted hologram deflects the optical signals to form a perfect shuffle interconnection.

This recirculating architecture requires that each array hold all of the data bits for each of the words to be sorted. Thus, the data are available for word-wise C&E modules. However, the simplicity of bitwise C&E modules, which operate on pairs of input bits, is desirable. In order to compare m-bit words with bitwise C&E modules, m-bitwise C&E modules are electrically chained together as shown in figure 3a. The electrical connection are used to communicate control signals such as the exchange status (conditional exchange, CE, or unconditional exchange, UE) and exchange criteria (up/down, UD).

Figure 3b shows how the bitwise C&E module can be implemented with simple combinatorial logic. A logical one(zero) on the UD input indicates that if an exchange is required, the larger word should appear on the H(L) output lines. A logical one on the CE input initiates a comparison of the two data bits based on the exchange criteria established by the UD input. A logic one on the UE input forces an unconditional exchange of the input data. If both the CE and UE inputs are a logic zero then the data are passed through the C&E module without exchange. This state of the control signals provides the functionality required for the C&E modules with no arrows in figure 1. The edge triggered latches shown in figure 3b are used to synchronize the flow of data and control signals within the system.

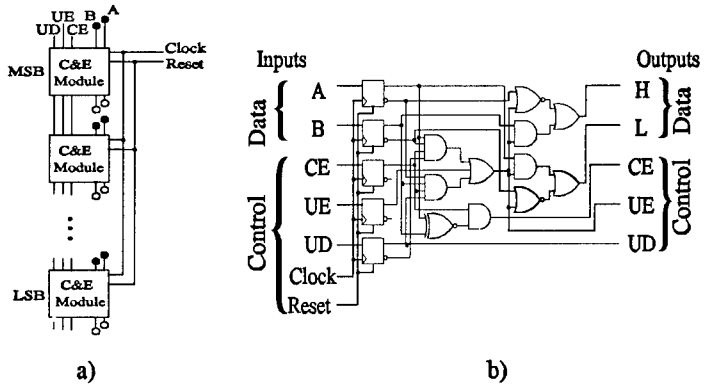


Figure 3 a) Bit parallel word comparison using bitwise C&E modules. b) Combinatorial logic required to implement bitwise C&E modules.

3. System Capacity and Scalability

The proposed recirculating architecture reduces the number of required processing stages to two. However, the system capacity is reduced unless the size of each stage is increased. Thus, there is a hardware trade off between the pipelined and recirculating architectures. The former requires many stages while the latter requires two large stages. While it is possible to increase system capacity by utilizing larger arrays, available fabrication technology will set a maximum achievable limit on array size. In order to overcome this limitation on system capacity, it is possible to implement each stage of the recirculating architecture with multiple arrays. However, this increases the mechanical complexity of the system. Figure 4 plots the number of arrays required to implement a bitonic sort as a function of maximum array size (and sorter capacity) for both architectures. For comparison it is assumed that system capacity is the same for both architectures. The shaded portion of figure 4 indicates the region where the recirculating architecture requires fewer arrays than the pipelined architecture. Note that for word lengths shorter than 64 bits, the number of arrays required for the recirculating system is always less than the number required for the pipeline system. For word lengths greater than 64 bits, there is a region where the pipeline

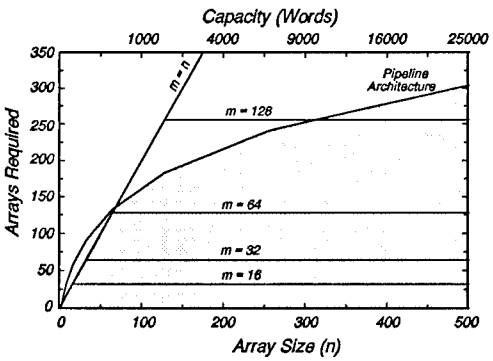


Figure 4 Arrays required to implement both architectures as a function of array size for word lengths of m = 16, 32, 64, 128, and n.

system requires fewer arrays. However, even for larger word lengths, the number of arrays required to implement very large capacity systems can be reduced with a recirculating architecture.

4. System Throughput

As shown in figure 4 the recirculating architecture may require significantly fewer arrays than the pipelined architecture. While this reduction in hardware components reduces the mechanical complexity of the recirculating architecture, this advantage is obtained at the cost of system throughput, which is defined as the number of sorting operations performed per unit time. For the pipeline architecture, the most significant bits of a data set follow the least significant bits of the previous data set. Thus, a new set of sorted data is output every m clock cycles. Therefore, the throughput of a pipeline sorter is not dependent on the system capacity. Conversely, the recirculating architecture cannot load a new set of input data until the previous set of data has been sorted. The number of clock cycles required to sort a set of data is dependent on both the system capacity and the word length. Therefore the throughput of a reciruclating sorter is dependent on system capacity. For a reasonable capacity, the pipeline system will have a larger throughput than recirculating systems. Thus, a trade off between mechanical complexity and system throughput exists.

The optimization of this trade off is certainly application dependent. Since telecommunications applications are throughput sensitive, the added mechanical complexity required to increase throughput may be warranted. However, computing applications such as database manipulation may not be as strongly dependent on throughput. Therefore, the reduction of mechanical complexity may justify the corresponding reduction in throughput. Ultimately, the acceptable degree of mechanical complexity for a system will be limited by the cost and reliability of the implementation. Thus, regardless of requirements for high throughput, a recirculating architecture may be necessary for sorting very large data sets.

5. Conclusions

We have presented an optoelectronic recirculating architecture which decreases the hardware required to implement a bitonic sorter. Along with a system layout, a arrangement of bitwise C&E modules was presented which performs a bit parallel word comparison. The bit parallel data organization used in this recirculating system results in decreased system throughput. Thus, the recirculating architecture reduces the mechanical complexity of the system by reducing the number of stages required at the cost of reducing the system throughput. The optimiation of this tradeoff is dependent on the specific applicaton of the sorting system.

6. Acknowledgments

This work was funded partially by NSF/ERC grant ED9015128 the Optoelectronic Computing System Center and the Colorado Advanced Technology Institute

References

[1] S. G. Akl 1985 *Parallel Sorting Algorithms,* (Orlando FL: Academic Press, Inc.)
[2] E. Gluch, H. Koboll, K. Zurl, N. Streibl, and J. Schwider 1993 *J. Mod. Opt.* **40**(10) 1857-1869
[3] C. W. Stirk and R. A. Athale 1988 *Appl. Opt.* **27**(9) 1721-1726
[4] L. Cheng, J. M. Wang and A. A. Sawchuk 1993 *Appl. Opt.* **32**(32) 6482-6492
[5] H. S. Stone 1971 *IEEE Trans. on Comp.* **C20**(2) 153-161

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Photothyristor based stochastic elementary processor

Guy Prémont, Philippe Lalanne, Pierre Chavel

Institut d'Optique Théorique et Appliquée, CNRS, B.P. 403, 91403 Orsay Cedex, France

Paul Heremans, Maarten Kuijk

IMEC, Kapeldreef n°75, B-3001 Leuven, Belgium

Abstract. We describe the implementation of a sigmoid updating probability required for Monte Carlo computations and simulated annealing. It relies on the simultaneous detection of two samples of the same speckle field. The temperature of the sigmoid curve produced is entirely controlled by the mean intensity of the incident speckle field. An experimental demonstration is described using a differential pair of *pnpn* photothyristors.

1. Introduction

The simulated annealing algorithm [1] is one of the most widely used stochastic optimisation techniques. It is based on a random exploration of the state space. Formally, changing one variable of the energy function E, resulting in an energy gradient ΔE, will be accepted with the sigmoid probability

$$P(\Delta E) = \frac{1}{1 + \exp(\Delta E/T)} \qquad (1)$$

where the parameter T is called the temperature and should decrease slowly over time (annealing) to reach global energy minimisation. We show that, when combined with speckle illumination, a simple processing element (namely a differential pair of *pnpn* photothyristors) can perform this stochastic operation.

Many different optoelectronic devices could conceivably be used in the elaboration of this new technique: SEEDs, optically activated VCSEL, etc. We use *pnpn* photothyristors for our demonstration. These devices provide amplification and allow the implementation of cascading operation particularly well-suited for iterative computation. Combined in a differential pair, two *pnpn* photothyristors act as a simple maximum detector [2]. This mode of operation is obtained by connecting two photothyristors in parallel. The voltage source applied to the pair is a three-step voltage sequence corresponding to reset, light detection and light emission. These devices can be described simply as comparators: the photothyristor which has detected the highest intensity switches on and emits light.

2. Differential detection of random speckle

Laser speckle, as a random number generator, offers great flexibility and easy implementation [3]. By differential detection of two independent speckle samples from the

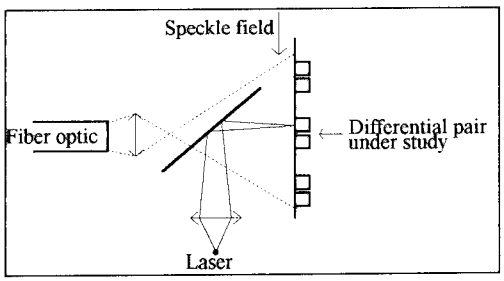


Fig. 1 Optical inputs applied to the differential pair.

same speckle field, a random signal of zero mean is created and thus an easy implementation of a symmetric probability function is allowed.

The speckle intensity I incident on a detector can be statistically described with a gamma probability function where the only parameters are the mean intensity of the speckle field $\langle I \rangle$ and the number of degrees of freedom M in the detected speckle [4]:

$$p(I) = \frac{1}{\Gamma(M)} \left(\frac{M}{\langle I \rangle} \right)^M I^{M-1} e^{-\frac{M}{\langle I \rangle} I} \qquad \text{with} \quad I \geq 0 \tag{2}$$

The parameter M is also linked to the variance of the speckle and can be shown to be the square of the optical signal to noise ratio: $M = \langle I \rangle^2 / \sigma^2$. To take an intuitive approach in this paper, we consider the gamma probability functions to be approximately equal to gaussian functions. This approximation is valid when the number of degrees of freedom M is large (in practice greater than 10).

In order to demonstrate the effect of differential detection, we suppose a differential pair of *pnpn* photothyristors is illuminated with a spatially homogeneous random speckle pattern. The photothyristor detecting the most intense speckle will switch on. The switch-on probability can be calculated by associating it to the speckle probability function. This is simply the probability that the first incident speckle intensity I_1 is larger than the second one I_2: $p(ON) = p(I_1 > I_2)$. We also include the energy gradient ΔE of Eq. 1 as an illumination on the second photothyristor of the pair with an additional laser beam having an intensity ΔE (see Fig. 1). The switch-on probability for the first photothyristor is the probability that the speckle intensity received on the first photothyristor is greater than that on the second photothyristor, which is the combined intensity of the speckle and of the laser.

$$p(ON) = P(\Delta E) = p(I_1 > I_2 + \Delta E) = p(I_2 - I_1 < -\Delta E) \tag{3}$$

The random variable $I_2 - I_1$ results from the difference between two known gaussian random variables having the same statistical parameters. The probability distribution of this new random variable is a gaussian probability distribution of zero mean and double variance. Therefore, the probability that the photothyristor to which the additional beam ΔE is applied will switch on is

$$P(\Delta E) = \frac{1}{\sigma\sqrt{4\pi}} \int_{-\infty}^{-\Delta E} \exp\left[-\frac{1}{2} \left(\frac{x}{\sigma\sqrt{2}} \right)^2 \right] dx \tag{4}$$

This is simply an Erf function. It is equal to the sigmoid of Eq. 1, with a high degree of accuracy, provided that

$$T = \sigma\sqrt{\pi/4} = \langle I \rangle \sqrt{\pi/4M}\,. \tag{5}$$

The temperature is found by matching the slopes of the sigmoid and of the Erf function for $\Delta E = 0$. Consequently, a differential pair of photothyristors illuminated with a speckle pattern

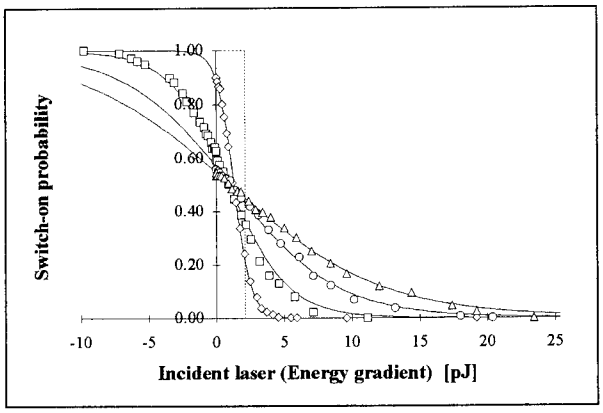


Fig. 2 Switch-on probabilities for one *pnpn* of the differential pair as a function of the incident laser energy, for many values of the mean speckle field.

should implement the stochastic updating operation of Eq. 1. Although we limited ourselves to gaussian speckle probability functions, we have analytically shown that the gamma statistics of the speckle allows accurate quasi-sigmoid operation with arbitrarily small M values.

3. Experimental results

3.1. Set-up

We tested the differential operation described above with a differential pair of *pnpn* photothyristors. The two photothyristors were identical squares of 50x50 μm^2 separated by 15 μm. Two laser illuminations were used. First, a speckle pattern created by the modal noise of a step index fibre was used to randomly illuminate the *pnpn* integrated circuits. The spatial independence of the speckle samples detected by the two photothyristors is satisfied by an appropriate magnification of the output face of the fibre. Also, the successive generation of time-independent speckles was ensured by inserting a rotating diffuser between the laser source and the fibre input, thereby randomising the phase of the input laser beam. As a second input, a laser diode was imaged onto one of the photothyristors of the pair, thus acting as the energy gradient on one of the photothyristors.

Light emission of the photothyristor not illuminated by the laser diode was recorded with a photodiode. We synchronised the detection with the voltage alimentation of the pair. In this way we detected the optical output of one of the *pnpn*s and were able to calculate its probability to switch on from the number of times it emitted over the total number of observations.

3.2. Results

The experimental results are shown in Fig. 2 and 3. The data points of the quasi-sigmoid curve are computed switch-on probabilities, each corresponding to the average over 5000 intensity measurements. Four probability curves, corresponding to four different temperatures, are shown. To demonstrate the obvious sigmoid characteristic of the response curves, the sigmoid updating probabilities of Eq. 1 were fitted to the experimental data points. In accordance with Eq. 5, these curves were obtained with four different values of the mean speckle intensity $\langle I \rangle$,

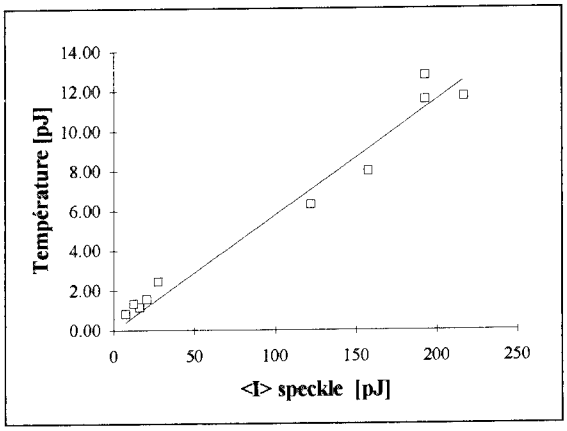


Fig. 3 Temperature of the sigmoid as a function of the mean speckle intensity.

respectively 0.1 mW, 0.4 mW, 0.6 mW and 1.4 mW. Since the temperature determines the amount of randomness in the updating operation, large values of T correspond to large values of $\langle I \rangle$ and the results correspond perfectly with the theoretical expectations. Furthermore, the temperature characterising each of the sigmoid curves is directly proportional, as expected, to the corresponding mean speckle energy. This is shown in Fig. 3 where we plotted the experimental temperature as defined in Eq. 5.

We have noted that all the sigmoid curves are shifted to the right. This is the effect of using real elements instead of perfect comparators. The shift corresponds to the smallest energy required for a correct switch-on operation. This behaviour is caused by the intrinsic asymmetry of the *pnpn* photothyristor pair. It corresponds to an energy of about 2 pJ. We can minimise the effect of this undesirable phenomena by using large sigmoid temperatures, thus working with energy gradients one order of magnitude larger than the shift itself.

An integration time of 60 μs was used. It was limited by the available laser power and not by the *pnpn* photothyristor technology. Operating frequencies in the range of 100 kHz are achievable.

4. Conclusion

We have presented a simple way to implement a stochastic processing element with sigmoid updating probabilities. We use speckle to generate spatially and temporally independent random numbers that we combine with pairs of *pnpn* photothyristors for a thresholding comparison. The result shows obvious sigmoid behaviour for the switch-on probabilities. Thus, it suggests a possible use in a parallel optoelectronic simulated annealing device.

[1] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, *Optimization by simulated annealing*, Science **220**, pp. 671-680, (1983).

[2] P. Heremans, M. Kuijk, R. Vounckx, G. Borghs, *Fast and sensitive two-terminal double-heterojunction optical thyristors*, Microelectronic Eng., **19**, p.49, 1992.

[3] P. Lalanne et al., *2-D generation of random numbers by multimode fiber speckle for silicon arrays of processing elements*, Opt. Comm., **76**, p. 387, 1990.

[4] J.W. Goodman, "Statistical Properties of Laser Speckle Patterns", in *Laser Speckle and related phenomena*, J.C. Dainty ed., Chap. 2 (1975).

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

The Optical Computing of the National 863 High Technology Program in China

Yi-mo Zhang*, Shi-yong Liu and Rong-jun Pan

Group of Optoelectronic Devices and Integration, National 863 High Technology Program, Beijing 912 Mail Box, 100083 Beijing, P.R. China
*Institute of Optoelectronics and Precision Engineering, Tianjin University, Tianjin 300072, P.R. China

Abstract. Some of the achievements of optical computing in the National 863 High Technology Program in China are summarised.

1. Introduction

The National 863 High Technology Program, which was proposed and drawn up by scientists, has been implemented in China since March 1986. Under the program there are several groups in different scientific fields. There is Optical Computing in the Group of Optoelectronic Devices and Integration, which incorporates more than twenty universities and institutes that have made achievements in optical computing. The several areas of optical computing are being implemented as follows.

2. Devices

2.1 Investigation of vertical cavity surface emitting lasers (VCSELs) (Beijing Institute of Semiconductors). The optimum structure of VCSEL has been designed and fabricated by MBE. By comparison of the theoretical with the experimental results of reflection spectra, x-ray double-crystal diffraction and photoluminescence, the match between the centre wavelength of the DBR high reflection band, the resonant frequency of the F-P cavity and the stimulated wavelength of the quantum well are realised. By means of proton bombardment the gain waveguide type of the device has been developed. The lowest threshold current under pulsed conditions at room temperature is less than 10 mA; it works at single-transverse and single longitudinal mode.

2.2 GaAs/GaAlAs multi quantum well reflectance modulator and self electro-optic effect devices (SEEDs) (Beijing Institute of Semiconductors). The analysis is emphasised on the combined behaviour of the Quantum Confined Stark Effect (QCSE), DBR and the asymmetric F-P cavity (ASFP) effect. Experimental results include: (i) A fabricated modulator array, with contrast ratio 7db ~ 10db, and its application to free-space optical switching and interconnection. (ii) Demonstration of bistability of a S-SEED array with optical switch energy less than $10fJ/(\mu m)^2$.

2.3 8 × 8, 0.88 μm LED matrix module (JiLin Institute of Physics). A 8 × 8, 0.88 μm LED matrix module of specific design with a black ceramic material base has been fabricated by a

hybrid integration technique. This device can be used for optical interconnection or 2D optical computing.

2.4 An optical addressed LCLV device and an electrically addressed CRT-LCLV (Xian Optical and Fine Machine Institute). The technical functions are as follows:

	LCLV	CRT-LCLV
Size of image plane	ϕ 50mm	ϕ 50mm
Resolution (0.5MTF)	50 lp/mm	35 lp/mm
Contrast	100:1	50:1
Response time	30 ms	30 ms
Sensitivity	3.7 erg/cm^2	

2.5 A ferroelectric liquid crystal light valve (FLCLV) (Shanghai Institute of Optical Instruments). A FLCLV has been developed. The technical performance is as follows:

Size of image plane	$\geq \phi$ 20mm
Resolution (0.5MTF)	$\geq$ 35 lp/mm (centre)
	$\geq$ 25 lp/mm (off-centre)
Response time	= 1~5 ms
Lifetime	$\geq$ 1 year

2.6 CdS, ZnS F-P device array for thresholding, feedback and gain (Changchun Institute of Optics and Fine Machines). CdS, ZnS F-P device arrays has been developed. Its threshold is implemented with the intensity sum of signal and holding beams. When their intensity sum exceeds the threshold of the device, the holding beam as a feedback beam will pass through the device array. The device array work-area is 10×10mm^2, 125×125 elements; the element size is $5\mu m \times 5\mu m$; the space between two elements is $3\mu m$. The ZnS device array is suitable for a working wavelength of 514.5nm; the operation energy density is 200nJ/μm^2. The CdS device array is suitable for a working wavelength of 632.8nm; the operation energy density is 20nJ/μm^2.

2.7 Nonresonant low power real-time storage (Jilin University). Two materials for real-time storage are developed: one is a push-pull azobenzene compound doped polymethyl methacrylate (PMMA) thin film. It is shown that this polymer has appreciable third-order optical nonlinearities with a nonresonant $\chi^{(3)}$ of about 4×10^{-4} esu for a He-Ne laser at 632.8nm. Another is biphoton real-time storage with methyl orange (MO) and ethyl orange (EO) sensitised polyvinyl alcohol (PVA) film. Experimental results: Ar^+ laser (514.5nm) pump power: 0.6Wcm^{-2} and He-Ne laser lowest write power 0.2Wcm^{-2}.

3. Architecture and system

3.1 An optoelectronic hybrid parallel multiprocessor system (Tianjin University). An optoelectronic hybrid parallel multiprocessor system is established. It consists of two stages of processor arrays (2DPA). Its architecture is a 3D pipeline. The interconnection in the 2DPAs are electronic, and it can be easily reconfigured into various topologies. The optical interconnection provides the communication of two PEs located at two 2DPAs, respectively, in the third dimension.

3.2 Hybrid morphological processor (Tsinghua University). An optoelectronic real-time morphological processor is established. A Dammann grating produces multiple images, $LCTV_2$ situated at the spectrum plane of the Dammann grating is used to display the structure

element. When the light illuminates $LCTV_1$, the convolution between the input image and the structure element is performed. A CCD camera collects the convolution result into a computer for thresholding. Thus morphological operations, such as dilation erosion can be obtained. Union and complement are also realised.

3.3 Optical cellular logic image processor (Shanghai Institute of Optics and Fine Mechanics). A one-operation image algebra was developed, in which the major operator consists of a parallel binary logic operation followed by a binary morphological dilation operation. Combined with threshold-decomposition and sum-superposition, various morphological functions for binary and gray-tone image processing can be realised by the simple iterative use of the operator. A four-layer architecture of cellular logic array was suggested to execute the algorithm for both binary and gray-tone image processing. The CPU is the cascade of a parallel binary logic processor and a binary dilation processor. Three systems were constructed. In the first system, the logic processor is a multi-imaging system with dual-rail spatial coding and the dilation processor is simply a defocussed unit with coded aperture. In the second system, the logic processor is a prism-optics assembled module and the dilation processing is realised by a Dammann grating. As the third alternative, an optical module scheme using 3-D stacked integration of polarisation elements was proposed. This packaged module is compact and effective. In all the configurations, the feedback is carried out by a PC computer.

4. Neural network processors

4.1 A fully-bipolar neural network with 32 × 32 neurons. A coaxial architecture and 2D optically-bipolar neural network with 32 × 32 neurons is developed. The IWM (Interconnection Weight Matrix) is implemented with a transparent SLM or a slide illuminated by a parallel beam from a fluorescent lamp with spectra λ_1, λ_2 and λ_3. The IWM is placed in the front focal plane of a lenslet array. The light passing each submatrix is coming to the liquid crystal switch array (LCSA) through the relevant lenslet and the imaging lens L_1. The LCSA is the conjugate plane of the IWM. For geometrical optics image property, N × N images of each submatrix will overlay together on the LCSA. Because the plane of the detector array is conjugated with the back focal plane of the lenslet array, N × N sharp bright spots which represent the weighted summations will be obtained on the detector array and the positions of three spots will shift with the direction of the parallel beam illuminating the IWM. Therefore, on the surface of each detector array element, we can obtain three spots corresponding to λ_1, λ_2 and λ_3, due to dispersion of a wedge. By proper colour encoding for the polarities of weights and neuron states, we can get $I_{\lambda_1}, I_{\lambda_2}, I_{\lambda_3}$ corresponding to positive and negative weighted summations. The operations of thresholding and feedback are fulfilled by a computer for the next iteration.

4.2 Optoelectronic hybrid real-time four-channel associative memory system (Tianjin University and Harbin University). This system includes three parts: input system, processing unit and output system. The proceeding courses are as follows:

1. Pre-processing of an image: A large number of reference images are stored in external storage carried out by a computer disk, after being processed by software under the control of a computer.

2. Input of images: As a four-channel system, the monitor screen is divided into one of 2×2 units by software. Four reference images and a recognised object are displayed on proper units of the monitor.

3. Associative recognition: Firstly, four different reference images are recorded in four photorefractive crystals by means of a Fourier transform (FT) hologram. Then the same

four displayed images of one unknown object are input into the processing unit, where they will do the correlating operation with the reference images simultaneously. The results are the signal beams that control on-off states of liquid crystal switches. If there is one reference image auto-correlating with the unknown one, there will be a maximum output. As a result, the corresponding switch is on, the holding beam transmits it and addresses the photorefractive crystal in this channel.

4.3 Optoelectronic hybrid system for image information extraction and associative recognition (Beijing Institute of Physics). This system extracts the image information with cross-scanning; a vector is set up with 32 bits of information, from an input image. The associative recognition result is obtained by the inner-product of the vector and a stored matrix.

4.4 Optoelectronic implementation of a compression-attraction-sphere associated memory using an optical pattern-encoding method (Tsinghua University). A compression-attraction-sphere associative memory model is proposed. An optoelectronic implementation of its inner-product architecture using an optical spatial pattern-encoding method is demonstrated by computer simulation and preliminary experiments.

5. Some applications in pattern recognition

5.1 Four channel real-time hybrid joint transform correlator **(JTC)** (Tianjin University and Changchun Institute of Optics and Fine Mechanics). A four-channel real-time JTC is proposed. Two optical wave-front-division multiplexers (OWFDM) divide the aperture of the system into four channels. In this system, the readout light energy as well as the physical area of the liquid crystal light valve (LCLV) devices can be fully utilised. When this system is used with only one OWFDM in front of the system, it can work at a speed four times as fast as a conventional one-channel system.

5.2 3-D spatial object recognition using analogue optical computing (Huadong University of Science and Technology). The spatial object is recognised by using a correlation operation in a coherent optical filtering system. The correlation result is received by an optoelectronic detector array which is placed at the output plane; thus the recognised result is given through computer comprehensive judgement. A small bank of filters is made up with a series of filters. The key step in this method is to design and make a bank of filters.

5.3 Hit-or-miss optical pattern recognition system (Tsinghua University). This system is based on the morphological hit-or-miss operation for multi-target recognition. It has the shift invariant, size invariant and rotation invariant property.

5.4 An optoelectronic implementation of a neural network for recognising handwriting characters (Tsinghua University). This system consists of two sub-networks. The input layer and hidden layer comprise the first sub-network; the second sub-network implements Boolean logic (AND-OR) operations. It combines the classification discrimination curved surface with a super-plane formed by the first sub-network.

6. Conclusions

(1) Under the united organisation all the studies are cooperatively going on.

(2) There are some differences between the National Nature Academic Fund, which mainly supports the studies of areas such as new ideas, new architectures and new kinds of devices, etc., and the National 863 Program, which lays particular emphasis on the creative researches in practical applications.

(3) The first stage of this program will end by 2000. When the time comes, there will be some available optical computing systems.

Inst. Phys. Conf. Ser. No 139: Part 1
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

An optoelectronic hybrid parallel multiprocessor system ★

Yimo Zhang, Wenyao Liu, Ge Zhou, Yong Wang, Dongjun Sun, Xiaoqing He

Institute of the Optoelectronics and Precision Engineering
Tianjin University, 300072, Tianjin, P. R. China

Abstract: In this paper the architecture of a 3D O–E hybrid multiprocessor is proposed. An O–E hybrid interconnection network is used for this system.

1. The architecture of the optoelectronic hybrid parallel multiprocessor system

An optoelectronic hybrid parallel multiprocessor system is building up, in which there are two stages of 2D processor arrays (2DPA); the optical interconnections are employed between these arrays, as shown in Fig. 1. Its architecture is a 3D pipeline.[1]

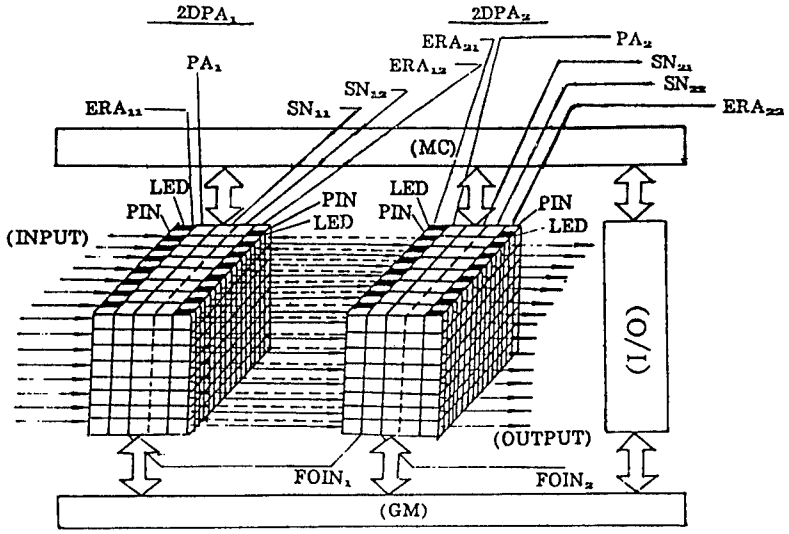


Fig. 1 The scheme of O-E hybrid parallel multiprocessor system
(2DPA)–2D O–E hybrid Processor Array, (PA)–Processing element Array, (SN)–Switch Network, (ERA)–Emitter and Receiver Array, (FOIN)–Fiber Optic Interconnection Network, (FC)–Fiber–optic Channel.

Every two dimensional O–E hybrid processor array comprises N × N processing elements (PEs). The interconnection in 2DPA is electronic. It can be easily reconfigured into various topologies.

★This research is supported by national high technology 863 program.

Eventually, there are two such O-E parallel multiprocessor system connected by multiple bidirectional optical channels. Fig.2 shows the diagram of this O-E parallel processor array system.

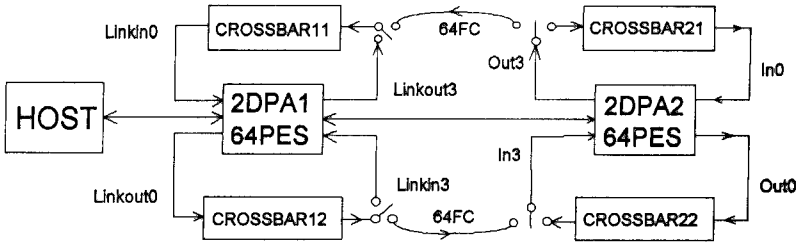


Fig.2 The schem of physical fiber optical interconnection network (FOIN), FC-Fiber Channel

2. O-E interconnection network

The optical interconnection provides the communication of two PEs located at two 2DAPs, respectively, in the third dimension. In our system, the fiber optical interconnection network (FOIN) shown in Fig.1 has 64 bidirectional fiber optical channels (FC), which consists of a emitter array with 64 LED modules and a receiver array with 64 PIN modules. The global transmission rate of the FOIN is 1.28Gbits/s. The test results show that the probability of error of data transmission is less than 10^{-12}.

In the practical system, the switch network (SN) shown in Fig.1. is the crossbar with 64 × 64 (I/O) ports in FOIN. The schem of interconnection network is shown in Fig.2. It can be seen that, there are four crossbars (i. e. SNs in Fig.1.), which play important roles in either the optical interconnection or topological reconfiguration of 2DPA. There are 6 processing elements in each crossbar, the programmes are stored in the processing elements.Once the communication between two 2DPAs is needed, the programmes will make the four crossbars connect with two 64 FC module, for instance, crossbar11 connects with crossbar21 by 64 FC, crossbar12 connects with crossbar22 using another 64 FC simultanously. The manner may be discribed as following equation:

$$(2\mathrm{DPA})_n (\mathrm{PE}_{ij}) \Leftrightarrow (2\mathrm{DPA})_m (\mathrm{PE}_{kl})$$

where n, m= 1, 2, i, k=1 ~ 8, and j, l=1 ~ 8.

3. Implementation of topological reconfiguration

The realization of topological reconfiguration of each 2DPA relies on two crossbars with 64 × 64(I/O) ports. The diagram of 2DPA is shown in Fig.3. The crossbar with 64 × 64(I/O) ports in our system is combined using six crossbars with 32 × 32(I/O) ports.[2] This combined crossbar is easily programmable to connect one input with any one output channel.

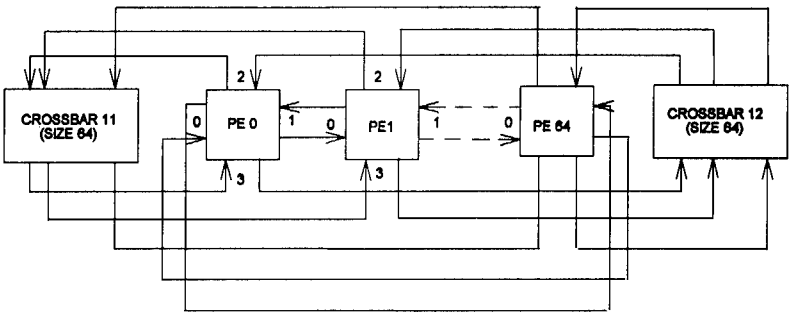
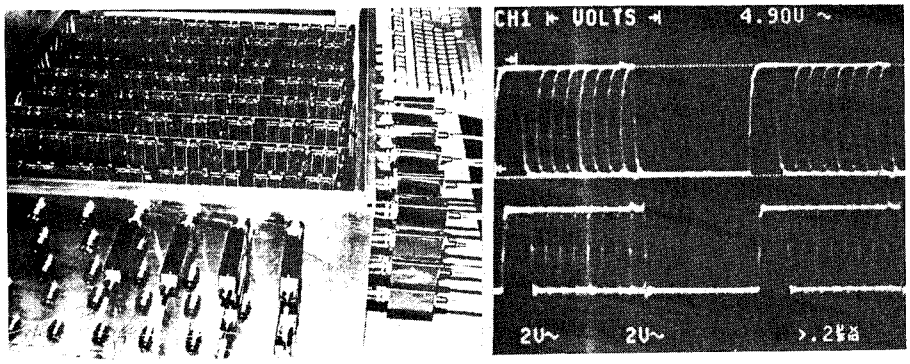


Fig.3 The schem of interconnection in 2DPA.



(a) (b)

Fig.4 (a) The photograph of a 2DPA built up with (FOIN).
(b) The photograph of waveform of data transmission in the optical fiber channel (FC).

It can be seen, from Fig.3, a processing element in 2DPA has four pairs of high speed links, two pairs of them connect other two elements directly. The other two pairs of links are connected with crossbar11 and crossbar12 respectively to form complete connectivity. By programming such two crossbars, the interconnection among the 64 PEs may be reconfigured. Therefor, various topologies such as mesh, tree and hyper cube, etc., may be reconfigured depending the computing algorithm.

The hardware of a 2DPA built up is shown in Fig.4(a) which involves the fiber optical interconnection network (FOIN). Having passed various tests shows that this system has very fine performances. Fig.4(b) is a photograph of data transmission signal of FC.

4. Conclusions

1) A O-E hybrid multiprocessor system is accomplished.

2) A O-E fiber optic interconnection network is used in this system successfully.

3) This system is used for pattern recognition in parallel and the results are obtained much quickly.

References

[1] Y. M. Zhang, W. Y. Liu, SPIE Vol 1983, 1993/429-431.
[2] Y. M. Zhang, et al, proc. of LEOS'93, 1933, 73.

Inst. Phys. Conf. Ser. No 139: Part 1
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Control Strategy for an Optically Reconfigurable Architecture

Miles Murdocca, Hans Raj Nahata, and Yu Zhou
Department of Computer Science, Rutgers University, Hill Center, New Brunswick, NJ 08903,
murdocca@cs.rutgers.edu (Murdocca), hnahata@paul.rutgers.edu (Nahata),
yuzhou@remus.rutgers.edu (Zhou)

Abstract. A two-dimensional array of microlasers controls a mating array of optical modulators in a model of a reconfigurable optical architecture. Free-space connections are enabled through the configuration of the microlasers. A problem in scalability is addressed in which arbitrary microlaser patterns are created by using a sequence of subpatterns that require only row and column control lines, rather than using an individual control line for each microlaser. Although pattern composition with this method may require several cycles, the principle of *functional locality* ensures that reconfiguration takes place infrequently. The result is a control method that scales gracefully as the size of the architecture increases. An example of a mapping of an arbitrary interconnection problem to the reconfigurable architecture is presented.

1. An Optically Reconfigurable Architecture

Traditional computer architectures are static in the sense that the connectivity among the internal components does not change, except when the hardware is physically rewired or circuit boards are changed. Despite the static nature of traditional computer architectures, widely varying behavior can be achieved as a result of modifying the software that the architectures execute. The underlying connectivity among the components, however, remains unchanged.

The instruction set in a static architecture is fixed, so performance improvements are typically made before the architecture is finalized by optimizing the execution of the most frequently occurring instructions [1]. For a given program, however, only 20% of the instruction set is used, and an even smaller percentage (about 10%) accounts for the bulk of instruction execution (about 85%) over a period of time [2]. This property of *functional locality* can be exploited to improve performance if the architecture is optimized for the instructions that are currently needed (the *working set*) rather than using a static architecture that is optimized for the average mix of instructions. In order to optimize the architecture for the execution of the working set, a mechanism for reconfiguring the connectivity among the architecture components is needed that does not introduce significant latency in the clock cycle.

Although reconfiguration can be implemented with a variety of electronic approaches, a significant latency is introduced when components are interconnected with an electronic multistage interconnection network (MIN). Further, a significant amount of time is required to reconfigure communication channels. We address the latency problem by using a shallow free-space optical implementation, and we address the reconfiguration problem by narrowing the control stream.

One way that the latency issue can be addressed is through the use of beam-steering, which

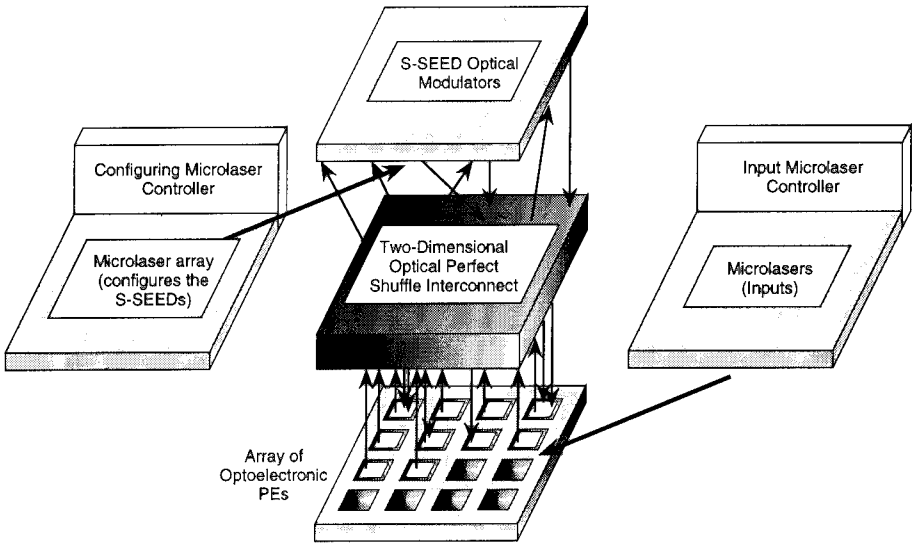


Figure 1: *Model of reconfigurable optical architecture.*

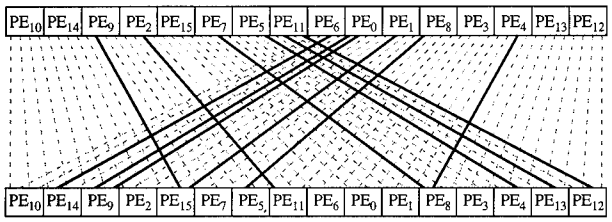
is analogous to physically rewiring an electronic processor. An alternative approach that we explore here is to provide a dense array of free-space connections, and selectively enable the connections that are needed. The underlying architecture for this approach is shown in Figure 1. A two-dimensional array of processing elements (PEs) communicates through optical input/output (I/O) ports. Optical signals travel in free-space through a static optical interconnect from the PE array to an array of optical modulators (S-SEEDs [3] for this case) and continue back to the PE array. An array of configuring microlasers sets the states of the modulators in parallel to be reflective or absorptive. A second microlaser array provides optical inputs to the PEs.

A similar system (without the PEs) was recently demonstrated in the U.S. Air Force's Photonics Center at Rome Laboratory / Griffiss AFB [4] in which vertical cavity surface-emitting lasers (VCSELs) [5] configure the S-SEEDs and also serve as inputs. In the Photonics Center system, modulated beams (the signal beams that are reflected from the S-SEED array back to the PE array) are generated by discrete edge emitting lasers and binary phase grating array-generators. This method of generating an array of beams can be replaced with a microlaser array, leading to the model shown in Figure 1.

A problem in scalability for this style of processor is that the configuring microlaser array needs to be controlled in such a way that only a narrow control stream is needed between the configuring microlaser controller and the microlaser array. Section 3 addresses the control stream issue in the context of an example presented in Section 2, which shows how a circuit is mapped onto the architecture.

2. Mapping a Circuit onto the Architecture

We are using a method of designing systems in which PEs are interconnected with a single two-dimensional optical perfect shuffle. The goal is to achieve all needed connections using a single pass through the perfect shuffle, although multiple passes may be needed. A simplified form of the problem is illustrated in Figure 2a, in which 16 PEs that each have four optical



(a)

From PE	To PE
0	9
0	9
1	15
2	11
4	8
5	4
6	14
7	8
8	5
9	15
11	12
11	13

(b)

Figure 2: *(a) Perfect shuffle interconnected PE array; (b) netlist for the circuit.*

inputs and four optical outputs are interconnected with a 64-channel one-dimensional perfect shuffle. Architecturally, the one-dimensional perfect shuffle corresponds to an unfolded version of a two-dimensional perfect shuffle. The unfolded version can take many different forms, and the form shown in Figure 2a is used for its visual regularity.

Using software developed for this application, a connection list among components (a *netlist*) is automatically translated into a placement of PEs for the perfect shuffle. The netlist shown in Figure 2b, which was generated randomly, led to the mapping shown in Figure 2a. More specific applications have also been mapped such as various types of binary adders [6]. Although a successful mapping of a netlist to a single stage of a two-dimensional perfect shuffle is possible for the case shown, not all encountered mappings can be satisfied with this approach. For these cases, multiple passes through the network are required. A key to making successful mappings for those that can be satisfied with a single pass is to use only a fraction of the available connections. For this example, only 12 of the 64 available connections are used.

3. A Scalable Method of Controlling the Microlaser Array

Once a mapping from a netlist to the shuffle-connected PE array is obtained, a control sequence needs to be created that will enable the connections. The 64 channels of the one-dimensional perfect shuffle shown in Figure 2a are folded in raster fashion into an 8×8 two-dimensional matrix in which 1's and 0's represent enabled and disabled connections, respectively, as shown in Figure 3a. For instance, the second row from the top in Figure 3a corresponds to connections originating from the PE_9,PE_2 pair.

For an array of *individually addressable* microlasers, each microlaser has an independent control line. Ideally, we would like to apply the full 64 control bits for the matrix shown in Figure 3a to the configuring microlaser array all at once. As the size of the matrix increases, however, the number of control bits quickly increases to sizes that are impractical to implement using an electronic controller. While the control problem might be solved optically, a more scalable electronic approach is possible in which only row and column addressing is needed rather than independent control of each microlaser. This is referred to as *matrix addressable* control of the microlaser array.

The matrix addressable configuration is more scalable than the individually addressable configuration because the number of control lines scales by only $N+M$ for an $N\times M$ array, unlike an individually addressable array in which the number of control lines scales by $N\times M$. Arbitrary patterns cannot always be implemented with a single row/column control vector, how-

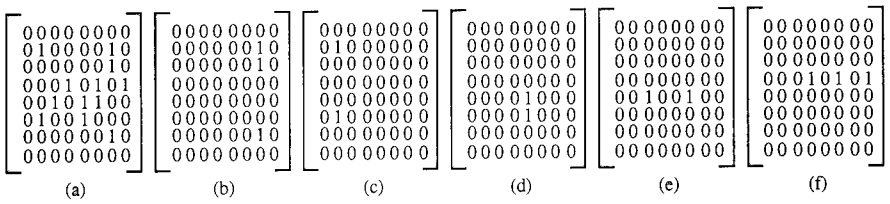


Figure 3: *(a) Microlaser configuration for the connection pattern shown in Figure 2a; (b-f) matrix addressable decomposition of the control pattern shown in Figure 3a.*

ever. For these cases, the target pattern must be decomposed into a number of subpatterns that are applied in sequence, which has the effect of increasing the reconfiguration time. In the worst case, the number of subpatterns equals the number of rows (or columns, depending on the orientation), but no greater, since the columns and rows are independently modulated.

Using decomposition software developed for this problem, the target pattern shown in Figure 3a is decomposed into five patterns shown in Figures 3b-f, which is less than the worst case of eight subpatterns (actually, a row-by-row decomposition would result in six subpatterns). The savings may be small for small matrices such as this, but can be substantial for larger matrices. Fewer subpatterns may also result if the S-SEED element at each row-column intersection is toggled rather than being set to an absolute state. In this scenario, the entire S-SEED array is initially enabled, and then the subpatterns disable S-SEED elements rather than enable them.

4. Conclusion

A model for a reconfigurable optical architecture makes use of a two-dimensional microlaser controlled array of optical modulators interconnected by a two-dimensional perfect shuffle. Mappings of arbitrary netlists to the architecture can be made for a number of cases with only a single pass through the perfect shuffle. The control stream that enables perfect shuffle channels is decomposed into matrix addressable subpatterns through a scalable process. The main conclusion is that a flexible reconfigurable optical architecture can be achieved with a shallow static interconnect and a narrow control stream.

The work reported in Section 3 was jointly supported by the Air Force Office of Scientific Research and the National Science Foundation (NSF) on NSF grant ECS 93-12625. The remainder of the work was supported by NSF on grant (MIP 92-24707).

5. References

[1] Hennessy J L and Patterson D A 1990 *Computer Architecture: A Quantitative Approach* Morgan Kaufmann Publishers, San Mateo, CA

[2] Murdocca M J and Gupta V, Mar. 1993 "Architectural implications of reconfigurable optical interconnects," *Journal of Parallel and Distributed Computing* **17** no. 3 pp. 200-1

[3] Lentine A L, Hinton H S, and Miller D A B 1988 "Symmetric self-electro-optic effect device: Optical set-reset latch" *Appl Phys Lett* **52** 1419-21

[4] Battiato J, Stone T, Murdocca M, Bussjager R, and Cook P "Free-space optical system based on vertical cavity surface-emitting lasers and self-electro-optic effect devices," submitted to *Applied Optics*

[5] Jewell J L, Scherer A, McCall S L, Lee Y H, Walker S J, Harbison J P, and Florez L T, 1989 "Low threshold electrically-pumped vertical cavity surface-emitting microlasers" *Electron Lett* **25** 1123-4

[6] Murdocca M, Battiato J, Berger D, Bussjager R, and Stone T, "Reconfigurable Architecture Based on Selective Enabling of Microlasers," *Int. J. Opt. Comp. and Proc.*, (submitted for a special issue on reconfigurability)

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical array logic network architecture

J Tanida and Y Ichioka

Osaka University, Department of Applied Physics, 2-1 Yamadaoka, Suita 565, JAPAN

Abstract. A new concept called optical array logic network architecture (OAL-NA) is proposed for effective construction of an optoelectronic hybrid computing system. With the help of optical array logic (OAL), not only data communication but also global data processing is implemented in the interconnection network. In this paper, the concept of the OAL-NA and possible operation of the architecture is described.

1. Introduction

Digital optical computing is a promising scheme for massively parallel computation in the near future because of the common foundation with the current computer science. According to the concept, a lot of techniques including optical array logic (OAL) [1] have been proposed. However, quantitative estimation of processing capability shows that a processing style in which all operations are executed by optical techniques is inefficient and it is difficult to achieve processing capability comparable to electronic computing systems. Current electronic computers offer tremendous computational power, so that massive parallelism and ultra fast operation must be exploited in the optical computing systems to overcome the electronic ones. Considering the status of researches on optical computing, we conclude that at least in near future optoelectronic hybrid computing systems are promising and practical. However, even for optoelectronic hybrid computing system, sophisticated architecture is required to fully utilize the capabilities of optical computing techniques. In this paper, we propose a concept of optoelectronic hybrid computing system, in which OAL is used for a high functionality interconnection network.

2. Optical array logic

OAL is a paradigm for optical parallel logical processing based on spatial encoding and discrete digital correlation as shown in Fig. 1. Since OAL is composed of simple operations on image data, various optical techniques can be used to implement it and large processing capabilities can be accomplished. In addition, OAL is based on a conventional logic system, so that accumulated resources in computer science can be utilized. Until now, various application fields, such as image processing [2], numerical processing [3], emulation of parallel processor [4], inference [5], and database management [6], have been

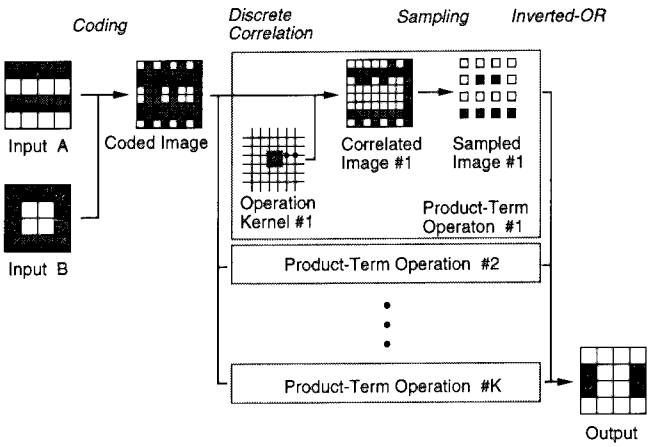


Fig. 1 Processing procedure of OAL

investigated with OAL.

However, quantitative estimation of the processing capability of the developed algorithms in OAL suggests that such processing cannot provide capability comparable to the current electronic computers. For example, to achieve 200MIPS (million instructions per second), the performance of a medium level microprocessor, by OAL with 1000 x 1000 pixels, the frame rate must be 60kHz for 16bit addition and 7.2MHz for 16bit multiplication. Although these values are not unachievable ones, it is difficult to find advantages of the OAL scheme with this processing style. Thus, we have investigated the developed algorithms and the characteristics of the OAL scheme to overcome the problem.

As the characteristics of the OAL scheme, the following are obtained. 1) Single complicated operation is more effective than a sequence of simple operations. 2) Template matching and image shifting are good examples of effective processing. 3) MIMD (multiple instruction stream multiple data stream) processing is difficult to implement with simple optical techniques. 4) Programmability of OAL is an important feature to obtain controllability in optical systems. Considering these characteristics, the optical array logic network architecture (OAL-NA) is developed.

3. Concept of OAL-NA

The OAL-NA is a conceptual architecture of an optoelectronic hybrid computing system as shown in Fig. 2. Essentially, multiple electronic processors execute arithmetic and logical operations in parallel and they communicate with each other through an optical interconnection network. The most important feature of the OAL-NA is that the optical interconnection is implemented by an OAL processor and provides high functionality: data transfer, data test, data processing, and system control.

Figure 3 shows a schematic diagram of an example system of the OAL-NA. In the system, multiple processing elements (PE's) with local memories are located on one plane (PE array) and a set of a global-sharing-memory and input/output processors are arranged onto another plane (MIO array). The two arrays are connected by an OAL network processor with symmetric data flow.

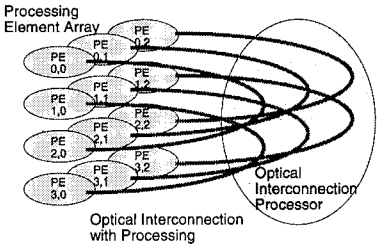


Fig. 2 Concept of the OAL-NA

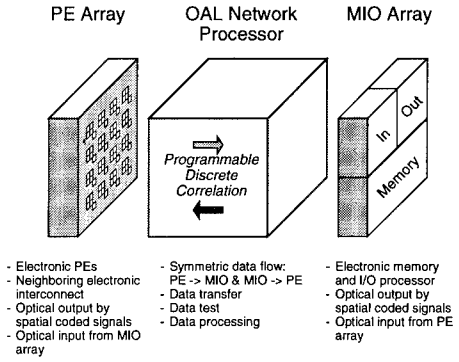


Fig. 3 Example system of the OAL-NA

Output signals from one array are used as inputs to the following OAL processor and the output of the OAL processor is transferred to the other array. Electronic modules on the array complete the encoding required in OAL and generate a coded image by composition of individual elements of the modules. Once the coded image is generated, the output signals from the electronic modules are processed by OAL. In the reverse data transfer, the same procedure is adopted. For inter-PE communication, round-trip routing is established. Specifying the OAL processor, we can achieve versatile functions.

4. Operations in OAL-NA

The OAL-NA provides various operations in the interconnection level. Typical operations are 1) data transfer (unconditional/conditional broadcasting, bi-directional shift [4], data exchange, token propagation [5]), 2) data test (conditional search, data validity check, convergence check), 3) data processing (differentiation, logical operation for same/neighbor PE's), 4) system control (command sending, arbitration), and so on. These operations are described by OAL programs and easy to executed by an OAL processor. Figure 4 shows the examples and corresponding OAL programs. Note that the left-hand images are the outputs of the PE array and that the pixel corresponds to the output of an individual processing element. The right-hand images are transfered into the following MIO or PE array. Accumulated resources in OAL research can be fully utilized for the OAL-NA.

5. Discussion

The OAL-NA has various advantages with respect to system construction and system capability. Use of highly developed microprocessor and smart pixel technologies for PE's is practical and matches with the current trend of massively parallel architecture in computer science. It has been pointed out that optical interconnection at interchip level is advantageous over electronic. In addition, interconnection merging with processing, the key idea of the OAL-NA, is expected to provide more effective communication and system operation than the conventional simple optical interconnection. OAL implements SIMD (single instruction stream multiple data stream) parallel processing, which is suitable

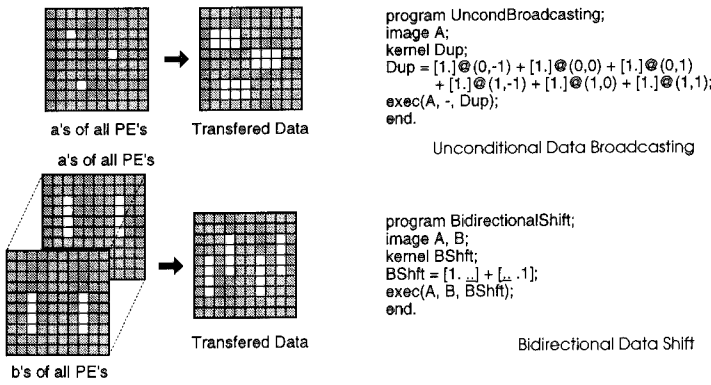


Fig. 4 Example operations in the OAL-NA

for global search and test on the output data from PE's, memory, and input/output processors. In terms of processing complexity, global search and test are appropriate for the OAL scheme and the advantages of OAL can be effectively utilized. Although an encoding device must be prepared even for a simple OAL processor, the encoding process can be combined into each electronic modules and no additional device is required in the OAL-NA.

The operational sequence of the OAL network processor consists of two phases: 1) kernel setup phase and 2) signal transfer phase. The transmission rate *T.R.* of the OAL network processor is estimated by Eq. (1).

$$T.R. = \frac{b}{a\,(t_{setup} + b\,t_{trans})} \qquad (1)$$

where a and b are the required number of kernel setups and the bit number of the transfer data, respectively; t_{setup} and t_{trans} are the setup time and the transfer time. Equation (1) indicates that *T.R.* is strongly affected by a and that the overhead of kernel setup is reduced as b increases. Fortunately, for most of the operations, a can be set to unity, so that the OAL-NA is expected to be operated without serious performance degration.

6. Summary

In this paper, we have proposed a new concept called optical array logic network architecture for effective construction of an optoelectronic hybrid computing system. With the help of optical array logic, not only data communication but also global data processing are implemented in the interconnection network. As a result, effective data communication and system operation can be expected.

References

[1] Tanida J and Ichioka Y 1990 *Int. J. Opt. Comput.* **1**, 113-28
[2] Tanida J and Ichioka Y 1988 *Appl. Opt.* **27**, 2926-30
[3] Tanida J, Fukui M and Ichioka Y 1988 *Appl. Opt.* **27** 2931-9
[4] Fukui M, Tanida J and Ichioka Y 1990 *Appl. Opt.* **29** 1604-9
[5] Iwata M, Tanida J and Ichioka Y 1992 *Appl. Opt.* **31**, 1093-102
[6] Iwata M, Tanida J and Ichioka Y 1993 *Appl. Opt.* **32** 1987-95

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optoelectronic multiport associative memory for data flow computing architecture

V B Fyodorov

Scientific Computer Centre, Russian Academy of Sciences, Leninsky pr. 32 A, Moscow, 117334, Russia

Abstract. A new multiport associative memory based on optoelectronic principle of data processing is suggested. This memory enables M users to execute simultaneously and independently a parallel content-based keys search and data retrieval into a common memory of N stored words by M search arguments, as well as a random-access writing of keys and data. The main parameters of such an associative memory are evaluated and its hardware implementation is discussed.

I. Introduction

An associative memory (AM) is of primary importance for data flow supercomputers [1-2] and processing neural networks. Attempts to create a fast high-capacity AM based on various physical principles have been repeatedly made almost since the first computers appeared. However, all these attempts either proved to be unsuccessful for fundamental or technical reasons or the implemented AM models could not be used in computers because of their poor relative economical efficiency.

In the present paper the creating principles and optical hardware setup of the new high-speed, high-capacity optoelectronic multiport AM for parallel systems with data flow organization of computing process are suggested. The main parameters of such an AM are evaluated and its hardware implementation is discussed. The possibility of using the suggested AM as an optoelectronic multiport random-access memory is considered. The principle of data flow computing organization based on using such memories is discussed.

2. Multiport associative memory organization

The multiport AM is a storage that enables M users to execute simultaneously and independently a parallel content-based key search and date retrieval into common memory array of N words by M search arguments, as well as a random-access writing of keys and data. The block diagram of the M-port AM is represented in Fig. 1. The data search process for such a memory for the m-th port (m = 1,2, ... , M) consists of the following. A binary code a_1a_2 ... a_L of a key to be found is written in the m-th search argument register and compared with keys $b_{1n}b_{2n}$ ... b_{Ln} (n = 1,2, ... , N) of all the stored words. The results of a comparison are fixed in the m-th response store. The addresses of coincidences that occur are revealed in turn by a priority multiple-match resolver and converted in AM address codes according to which the stored in the memory cells can then read out.

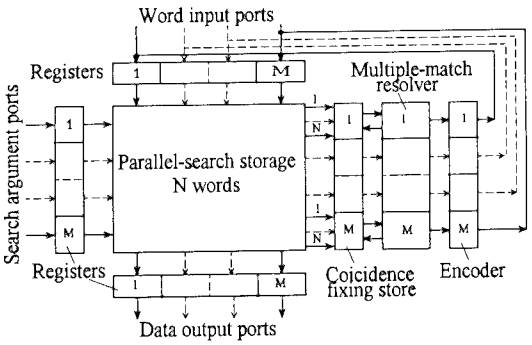


Fig. 1. The organization of an associative memory.

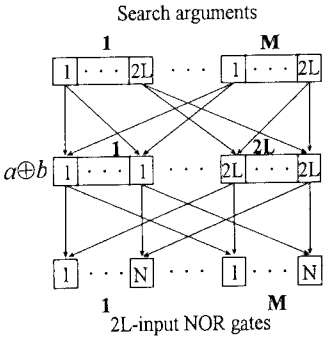


Fig. 2. Graph of associative search.

The most difficult problem is to provide a solution to three correlating tasks: fast simultaneous transfer of the produced search arguments to each memory cell, bit-wise comparison of each search argument with all stored keys (the logical function $a \oplus b$) and revealing of coincidence occurrences with subsequent information retrieval according to the located physical addresses of keys. This problem is difficult to solve mainly because of the necessity of realising a great many interconnections between memory cells and other multiport AM devices.

3. Optoelectronic implementation of the M-port associative memory

The proposed M-port AM has two distinguished features. Firstly, to bypass the problem of a high-speed light-sensitive reversible recording medium, writing, erasing, and storing information in such an AM are fulfilled by electronic methods, whereas the key search is realized by optical means. For this purpose the memory array is implemented in the form of two blocks: optical block of the associative search (SAM) on the basis of a hybrid memory board (HMB) with optoelectronic individually electrically addressed VLSI/SLM smart pixels (see, as an example, [3]) in which the compared bits of all the words are stored and block of the random-access retrieval on the basis, for example, of a semiconductor M-module random-access memory (RAM) or an optoelectronic M-port memory to store and address read out the data bits of the words. The HMB is a 2-D memory cell array, each memory cell consisting of spatial light modulator (SLM) and MOS-transistor-based trigger placed near electrically connected to the SLM and with the RAM.

Secondly, the like bits of the N keys are grouped together on HMB in 2L fields: $b_{1,1}b_{1,2} \ldots b_{1,N}$, ... $b_{2L,1}b_{2L,2} \ldots b_{2L,N}$, that allows one to use for an optical system of the AM the lenses of an available size, since N>>2L as a rule. A graph of the parallel-by search argument, parallel by-bit content-based search of stored keys for such a bit arrangement on HMB is depicted in Fig. 2. If optical data processing methods are implied it is preferable to use logical circuits based on threshold light inverters (multi-input NOR gates) to reveal coincidences.

The scheme of a block for the parallel optical associative search of keys' addresses is represented in Fig. 3 for a particular case for M=4, L=4, and N=16. The image multiplication and superimposition in such an optical system is implemented by using the spherical lenses only (see reference [4]). The stored keywords are registered on the HMB as transparent and opaque areas (ON and OFF spatial light modulators) corresponding to '1' or '0' of the binary bit b_{ln}. The functions of masked search argument registers are accomplished by the electrically controlled point light source arrays. The availability or absence of the light beam corresponds to '1' or '0' in the binary bit a_{lm} of search argument, respectively. The physical addresses c_{mn} of coincidence occasions (light absence) are fixed by photodetector arrays

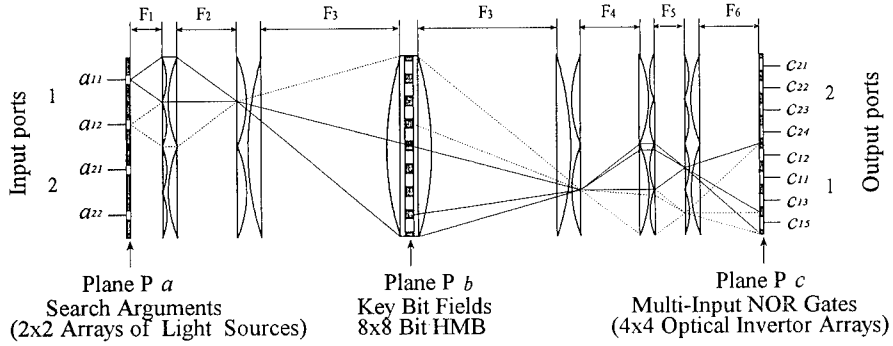


Fig. 3. The optical setup for an optoelectronic multiport associative memory.

which serve as threshold optical invertor arrays. The binary bits a_{1m}, a_{2m} ... a_{Lm} of all (m = 1,2. ... , M) the search arguments are encoded by a dual-rail code and binary bits b_{1n}, b_{2n} ... b_{Ln} of all (n = 1,2, ... , N) the keywords are stored in an inverse dual-rail code. The optical system projects each search argument image (plane P_a) onto all the keys stored on the HMB. Images of the like bit fields (plane P_b) transmitted through the HMB are superimposed L-times in each output port (plane P_c). As is seen in Fig. 3 the optical threshold inverter denoted by entry c_{mn} fixes the fact of coincidence of the m-th search argument with the n-th key, i.e. the inquiry answer of m-th user appears always in m-output port.

The optical system depicted in Fig. 3 ensures the associative search in the HMB in light transmission mode. However, if light polarization plane modulators are employed for the HMB it can be modified to operate in reflected light mode by using the polarizing beam splitter cube placed before the lens of the HMB. It is significant that the optical system represented in Fig. 3 is reciprocal in respect to optical inputs and outputs, and hence it can function as an multiport optoelectronic random access memory, if $N^{0.5}xN^{0.5}$ photoreceiver arrays replace by $N^{0.5}xN^{0.5}$ laser source arrays and $L^{0.5}xL^{0.5}$ laser source arrays replace by $N^{0.5}xN^{0.5}$ photoreceiver arrays.

4. Key coding and optical one-match resolver

In the optoelectronic AM there are enough reasons leading to a spread of logical levels for '1' and '0' light signals. Its influence on the search operation reliability can be substantially diminished by using a redundant code. In this case, the minimum difference between two arbitrary code combinations will be defined by the code distance d >1, and hence the level of optical signal distinguishing the coincidence of the search argument with a keyword from noncoincidences is increased (see Fig. 4).

Code combinations with the minimum code distance d > 4 can be realized by representing binary words in Reed-Müller codes [5]. Such codes are characterised by the following parameters: code length $k = 2^q$; number of key bits $l = \Sigma_i^h{}_{=o}\ C_q^i$; minimum code distance $d = 2^{q-h}$, where q 3 is a positive integer, h q is the code order. The analysis shows that the codes with parameters l, k, d respectively equal to (26, 32, 4), (42, 64, 8), (64, 128, 16) and (130, 512, 64) should be selected from many Reed-Müller codes. In this case d/k=1/8 with a relatively small redundancy $\alpha = k/l \cong$ 1.3, 1.5, 2.0 and 3.9 as well as sufficiently significant decrease in requirements by the factor $\beta = dl/k \cong$ 3.3, 5.3, 8.0 and 16 to the contrast ratio of smart SLM pixels and a light power spread and variation.

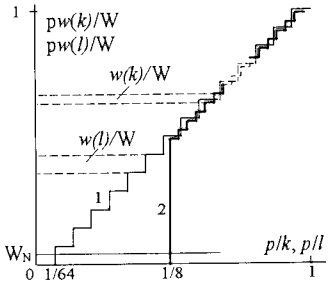


Fig. 4. Idealised dependence of the light level when executing nonequivalence operation versus the number p of noncoinciding bits in l-bit keys in the case of non-excess 1 and excess 2 Reed-Müller coding with $d = k/8$: $w(l)$ and $w(k)$ - one bit energy for 1 and 2, respectively; W_N - a noise level).

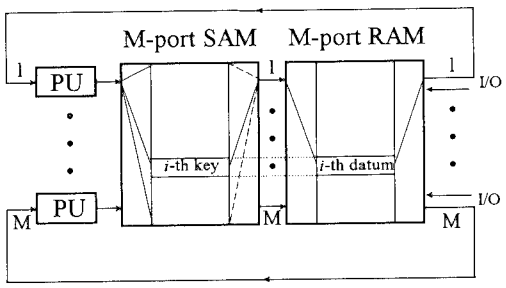


Fig. 5. The simplified block diagram of data flow architectire: PU - information processing unit, I/O - input/output ports.

In a data flow computing architecture only an one-fold coincidence of the search argument with all the stored keys is possible. This characteristic property allows one essentially to simplify the optical implementation of the multiple-match resolver and encoder (see reference [6]).

5. Possible performances of the M-port optoelectronic associative memory and its application for a data flow architecture

In the case of optoelectronic AM implemented according to the scheme in Fig. 3, the values of 2LN and M can be connected by the following relation: $2LMN = (2u/3\lambda)^2 S$ if it is supposed that a centre-centre pixel spacing corresponds to 3-fold Rayleigh criterion, all lenses with focal lengths F_1, F_2, ..., F_6 have equal numerical apertures u and hence $F_1 = F_5$, $F_2 = F_4 = F_6$, as well as cross sections of the optical system are of the same aperture size S = DxD. In accordance with this relation for the concrete values S = 8x8 cm, $u = 0.25$ and the light wavelength $\lambda = 0.84$ μm the value $2LMN = 2.5 \times 10^8$ bits. This means that the optical block of the optoelectronic AM can provide, in particular, the following utmost characteristics: the number of port M = 256 and the memory capacity being $N = 10^4$ 50-bit dual-rail coded keys. As this takes place, a centre-centre spacing of pixels in the HMP, point light sources and photoreceivers in arrays are equal ~80 μm, 500 μm and 50 μm, respectively, overall length of such an optical system is equal to about 40 cm and the aperture size of the smallest lenses $d_5 = D/N^{0.5} = 0.5$ mm ($u = 0.25$). Fabricating of array of such lenses for forming of 100x100 resolvable sports will be no problem [7].

The associative access time t_a is determined by many factors and circumstances. For the optoelectronic AM, two factors of principle, related to the finite optical signal speed and quantum nature of light, limit the minimum achievable associative access time. The first factor limits t_a by value of $t_{a,1} \approx$ 1-2 ns. Such time is necessary for passing the light signals from plane P_a to plane P_b in free-space optical system of an optoelectronic AM (see Fig. 3). The second factor of principle determines the minimum number of photons (or the minimum energy W_{th}) that must be recorded for the reliable discovering of a coincidence and hence allows to find the utmost value of $t_{a,2}$ from an energy consideration. In the case of coherent radiation the threshold sensitivity of the arrays containing a great number of photoreceivers $W_{th} \approx 10^{-14}$ J [8]. The $t_{a,2}$ value may be evaluated by the following expression: $t_{a,2} \geq (k/d)MNW_{th}/P$, where P is the average radiation power of all the light sources. For the above-

mentioned example of AM ($N=10^4$, M=256) the value of $t_{a,2} \approx 2.5$ ns, if the Reed-Müller code with (42, 64, 8) parameters are employed and P = 80 W. Hence the maximum achievable rate of key retrieval limited by the value $M/t_{a,2} \approx 10^{11}$ keys s^{-1}. As this takes place, the radiation power of each laser is equal to 5 mW; the admissible heat release in light source arrays (Q_l), HMB (Q_s) and photoreceiver arrays (Q_r) is equal to $Q_l = (1-\eta)P/\eta \approx 8.5$ W cm^{-2} and $Q_s \approx Q_r = 0.5P/S \approx 0.6$ W cm^{-2}. Here, it is suggested that the coefficient of performance of GaAs lasers $\eta = 0.15$, as well as aperture and absorption photon losses are neglected.

Since for a HMB data storage and rewriting are accomplished by electronic circuitry, memory cell modulators do not need to have a threshold light modulation characteristic that is obligatory for a light-sensitive reversible recording medium. This means a wide class of materials can be used for SLM smart pixels. We think that it is possible to reach in the direct addressing mode a nanosecond cycle for data rewriting in HMB cells by use of such SLMs and fast MOS transistor circuits fabricated by means of integrated circuit technology.

The exploitation of such a multiport AM in data flow digital computer is illustrated by the block diagram in Fig. 5. Any result coming out of the *m*-th processing unit (PU) searches by special code (search argument) for a data set in the AM to be transferred to the next operation for the execution. When the optoelectronic multiport AM contains the searched key, the data stored in the optoelectronic multiport RAM are read out and united data set (two as a rule) passes for execution by the next operator in the *m*-th PU. If the searched key is not found, the key together with the datum coming out of the PU is written to an empty place in the *m*-th block of a block-oriented (for writing only!) electronic memory array.

Since the keys are unique, the architecture shown in Fig. 5 eliminates the possibility of memory and PU collisions during the data processing. Therefore, there is no need to incorporate MxM truly nonblocking interconnection networks and multiword buffer memories in data flow hardware, as it done, if M modules of the one-port AM are used [2].

6. Conclusion

Our study showed that it has become possible to create compact high-performance M-port optoelectronic associative and random-access memories for the storage of N L-bit words by an integration of optical and electronic methods of information processing.

The total number of global free-space interconnects realized in such memories by optical means is at least ~ $2LMN = 10^8$-10^9. It is quite clear that the physical implementation of such a number of independent interconnects in a limited volume is impossible by electrical means.

References

[1] Burtsev V S and Fyodorov V B 1991 *Optical Computing & Processing* **1** 275-278
[2] Burtsev V S and Fyodorov V B 1992 *Optical Computing & Processing* **2** 161-172
[3] Johnson K M, D J McKnight and Underwood 1993 *IEEE J. of Quantum Electronics* **29** 699-714
[4] Fukui M and Kitayama K 1992 *Appl. Optics* **31** 5542-5547
[5] Chetagurov Y A and Rudnev Y P 1974 *An Enhancement of Digital Systems Reliability by Means of Redundant Coding* (Moscow: Energiya)
[6] Verbovetskii and Fyodorov V B 1980 *Kvantovaya electronika* 7 1769-1777 (in Russian)
[7] Ida K 1990 *Proc. SPIE* **1319** 486-490
[8] Gusev V K, Roslova M L, Fyodorov V B and Shilov I A 1982 *Radiotechnika* **37** no. 6, 21-30 (in Russian)

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

HPT/VCSEL Array Implementation of an Optoelectronic Data Filter

R.D. Snyder, F. R. Beyette, Jr., S. A. Feld, K. M. Geib, L. J. Irakliotis, P. A. Mitkas, and C. W. Wilmsen

Optoelectronic Computing System Center and the Department of Electrical Engineering
Colorado State University, Ft. Collins, Colorado 80523, USA

Abstract: The design and implementation of an optoelectronic data filter are discussed. Design requirements of the smart pixel arrays and system optics are presented. Preliminary system testing demonstrates viability of an optoelectronic data filter.

1. Introduction

In order to take full advantage of the massive transfer rates of 2-dimensional and 3-dimensional optical memories as well as the 2-dimensional processing nature of optical systems, new optoelectronic components with a high degree of parallelism must be developed. These components must be fast, perform a logic function, be integratable into dense arrays and optically addressed. This smart pixel approach can then be used to perform parallel operations on a full page of data, thereby greatly increasing the data throughput. Since optoelectronic processing should not compete with electronic computing but rather complement it, we propose to fabricate optoelectronics interfacing processors which preprocess data in order to decrease the input load on the electronic host computer. One system which has these characteristics is an optoelectronic database filter which is a preprocessor situated between a high speed optical memory and an electronic processing unit. This paper describes an implementation of this filter based on arrays of smart pixels containing vertical cavity surface emitting lasers (VCSELs) and heterojunction phototransistors (HPTs).

2. Optoelectronic Data Filter System

The data filter accepts data in a paged format and carries out relational operations such as projections and selections by comparing the data with search arguments. These two operations produce a subset of the input data which is then fed to the electronic processing unit via an optoelectronic RAM. The two operations are carried out within separate modules of the filter. These modules make up the filter as shown in Figure 1 below; each module contains arrays of logic elements. There is an AND array in both the projection and selection modules along with an XOR array in the selection module. A description of the data filter's operation is given in [1].

2.1. Approach

In designing the system, we kept the following goals in mind. We wanted to provide a viable platform on which to test the VCSEL/HPT optoelectronic elements in an array format, and demonstrate that such arrays can be used along with optical elements to perform computing operations on data sets in a highly parallel system.

Our first method of implementation is to build the system around arrays of hybrid devices (i.e., the HPT and VCSEL arrays are on separate chips). This allows us to test the concepts of the database filter and identify possible design problems before implementing the filter using the more complex and costly monolithic design. The gate configurations used in the filter are shown in Figure 2. The AND gate requires two HPTs which drive a VCSEL. The VCSEL provides the gate output and is "on" only when both inputs are "on." There are four HPTs in the XOR gate with two in parallel and two in series. Each input is directed to one parallel and one series connected HPT pair. The VCSEL is "on" when only one of the inputs is "on." For the XOR gate, the design of the HPTs must ensure that when both A and B are on, the series HPTs can shunt enough current to drop the VCSEL current below threshold.

2.2. Array Design

There were three main requirements in the design of the XOR and AND gate arrays. They are: maximizing gate gain, minimizing pixel size, and providing large optical input windows. From Figure 1, it is apparent that there will be optical signal loss due to the beam splitters as well as losses in the lenses and devices. Therefore, in a first order analysis there must be gate gains of at least four to overcome the losses due to the optical signals that must pass through a maximum of two beam splitters. Maximizing gate gain in VCSEL/HPT-based systems has previously been investigated [2]. Minimizing the pixel size provides a denser, more cost-efficient array. Finally, providing large input windows reduces losses since larger portions of the optical signals enter the HPT active regions.

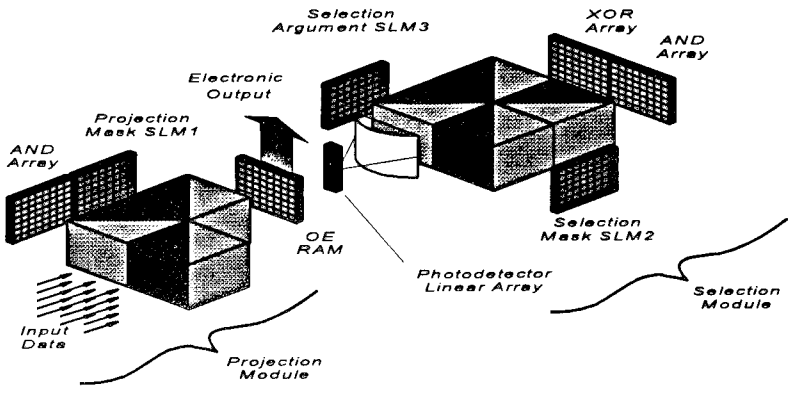


Figure 1: Schematic of database filter showing projection and selection modules[1]

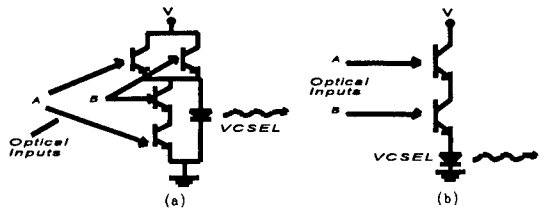


Figure 2: Gate design (a) XOR array (b) AND array

A diagram of the HPT portion of the XOR pixels is shown in Figure 3. As required by the circuit in Figure 2a, each optical signal illuminates two HPTs. Unfortunately, the hybrid implementation shown here requires that each pixel have an off-chip connection to a VCSEL. Thus, the hybrid approach is not scalable; however, the monolithic approach is scalable since the inputs and outputs are strictly optical.

3. Results

The hybrid XOR arrays have been fabricated from InGaP/GaAs HPTs grown by GSMBE on a semi-insulating substrate. The VCSELs were fabricated on a separate AlGaAs/GaAs wafer grown by GSMBE on a conducting substrate. After fabrication, the XOR array was tested with the initial system setup, shown in Figure 4. This initial setup was used to identify the difficulties in building the filter with bulk optics. Later versions will use lenslet arrays for compactness and simplicity.

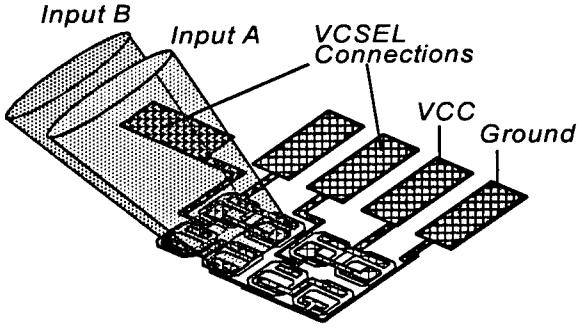


Figure 3: Layout of XOR HPT arrangement and full XOR array

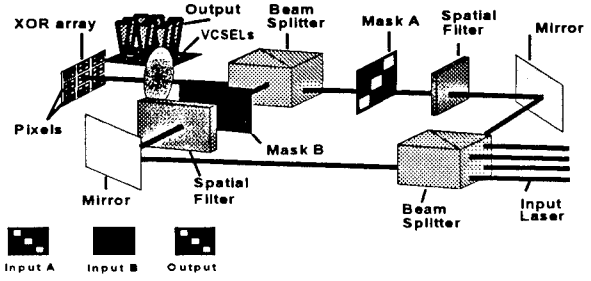


Figure 4: Initial system setup

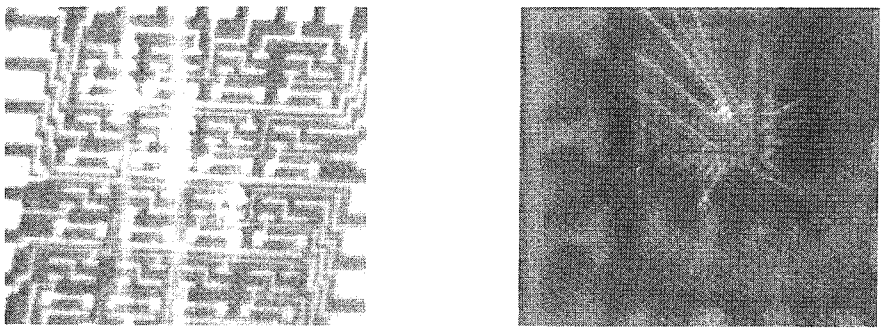


Figure 5: Photograph of (a) input pattern on XOR gates; (b) output pattern on VCSEL array

The optical input came from an edge emitting laser at 850-nm, and was split into two paths by an initial beam splitter. The two optical signals were directed with mirrors onto two masks, generating input patterns A and B. The second beam splitter combines the patterns before they are focused by a single lens onto the input windows of the XOR gates. The VCSEL outputs of the array are on a separate chip. As observed in Figure 5, the VCSEL array gives the expected output pattern. Effort is now concentrated on building the separate modules of the system. Once they operate correctly, the entire filter will be tested.

4. Conclusions

We have proposed on optical database filter which is based on HPT/VCSEL smart pixel arrays. Initial system testing has suggested that the filter concepts and design are viable. We are currently implementing the filter using bulk optics and intend on utilizing microlenses in the future as well as developing the system using monolithic arrays.

5. Acknowledgments

This work was funded partially by NSF/ERC grant ED9015128, the Optoelectronic Computing System Center, and the Colorado Advanced Technology Institute.

References

[1] P. A. Mitkas, L. J. Irakliotis, F. R. Beyette, Jr., S. A. Feld, and C. W. Wilmsen 1994 *J. Appl. Optics*, **33**(8) 1345-1353

[2] S. A. Feld, F. R. Beyette, Jr., K. M. Geib, R. D. Snyder, H. Temkin, and C. W. Wilmsen 1994 Submitted to Proc. IEE-J Optoelectronics

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Design of symbolic substitution systems for micro integration

C. Passon, K.-H. Brenner and W. Eckert

Universität Erlangen-Nürnberg, Physikalisches Institut,
Lehrstuhl für Optik, Staudtstraße 7/B2, 91058 Erlangen, Germany

Abstract. The design of micro optical systems has to consider the features of micro optical components. We introduce design concepts for micro integrated symbolic substitution systems.

1. Introduction

Systolic arrays represent a concept for developing highly parallel computer systems using regularly interconnected simple processor arrays [1]. We have recently demonstrated that systolic arrays can be easily mapped to symbolic substitution rules [2] and thus can be implemented optically. We have constructed an optical pipeline adder based on systolic arrays, which was realised with macroscopic optical components.

This adder consisted of an array of 8x8 half adders, performing a full addition of 8-bit, dual rail coded numbers in a pipeline within 8 iterations. The active array consisted of 16x16 pixels. From space-bandwidth considerations one can derive that imaging of such an array requires only lens-diameters of a few hundred microns. Consequently the size of the whole system can be reduced into the submillimeter range using micro optical components. A concept for miniaturising free-space optical systems was recently presented [3,4]. With this stacked approach, the packing density can be increased and the connectivity of three-dimensional optical systems can be utilised better than with planar integration.

2. Design Concepts.

Here we try to build on this integration concept in order to first realise a concept for a miniaturised version of the optical pipeline adder. Then we will generalise this architecture to implement general systolic array architectures. The concept takes the features of micro optical components into account and is thus compatible with the fabrication constraints.

The optical pipeline adder consists of a recognition and a substitution part. Each part is realised by a sequence of two multiple imaging systems (MIS). Active components (NOR- and OR-gate arrays) are located at the entrance/exit of each part. Each MIS (fig. 1) consists of two Fourier transform stages, which are constructed as light pipes, because this configuration offers best light efficiency and resolution [5].

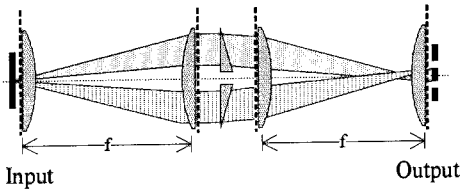


Fig. 1: Multiple imaging system

The complete sequence of components, needed to realise a substitution system is shown in fig. 2. This stretched out version could in principle be realised by a stack of optical components. There are, however, a series of issues that have to be addressed

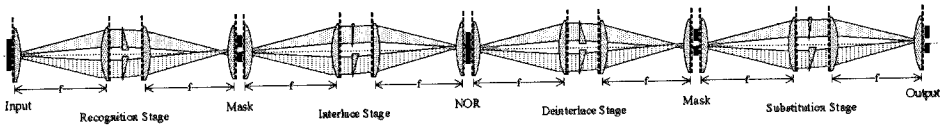


Fig. 2: Complete symbolic substitution stage

First, the number of layers, to be stacked, is very large in this version. This may cause problems in aligning the system.

Second, the active device in the centre of this sequence would have to be realised in transmission mode, which is undesirable both for thermal and for connectivity reasons. A more realistic design should place the active components at the ends of the stack. Thus a folding of the system is necessary.

3. Folded Systems

The first version of folded systems, shown in Fig. 3 has deflecting mirrors in the Fourier domain. The deflecting prisms for the multiple imaging operation have to be positioned between the two mirrors. Here also a pupil has to be placed to prevent vignetting, which results from the distance between the 2 Fourier systems.

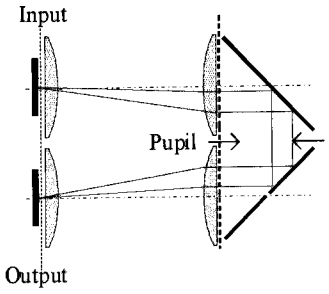


Fig. 3: Fourier folded MIS

This is not compatible with a layered structure, since the pupil and the prisms would have to be oriented perpendicular to the substrate surface, in order to be located in the centre between the mirrors.

Another limitation of this configuration is that the location a of the pupil has to satisfy the condition $D \leq a \leq f/2$, where D is the lens diameter. The largest numeric aperture (*N.A.*) is achieved for minimal a. Thus the best configuration $D = a = f/2$ results in a maximum *N.A.* of 0.5. The light efficiency is determined by the diameter of the pupil $D_p = D(1-a/f)$.

In our second approach (fig. 4) we insert the mirrors into the light pipes. Here all components are arranged in separate layers and no vertically oriented components are necessary. In this approach, we move the image plane away from the input data plane and the N.A. is consequently also limited by the increased distance resulting from the double mirror reflection. The beam splitters required in this approach can be implemented by the LIGA technique [6]. The minimum distance b between lenses is limited by :

- the size of the folding mirrors/beamsplitters (max. D necessary)
- the thickness of the lens layer d_{lens}
- the thickness of the prism layer d_{prism}

$$\Rightarrow b = D + 3d_{lens} + d_{prism}$$

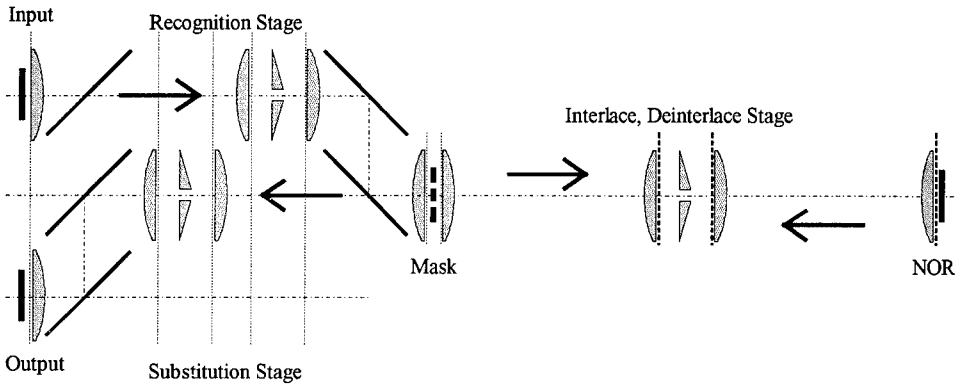


Fig. 4: Fully folded substitution design with active components at the ends of the layer structure

In a next step we can exploit the fact that the required shifts in two successive stages are identical. Thus the same hardware can be used in the forward and in the backward direction if a reflective array is placed between these stages (fig. 4, right side). This scheme can be replicated infinitely above and below. Thus the initially stretched out system can be arbitrarily cascaded to achieve complex systems and also make full use of the available substrate area.

4. Space Bandwidth Considerations

We have recently introduced a generalized Fourier system [5] based on the light pipe, which allows an additional distance a between input/output and the lenses, thus reducing the distance b between the lenses. Introducing a system factor $L = b/f$, $L = [0 \,..\, 1]$ the generalized Fourier system [5] can be classified with L:

$L = 0$: 2f-system $\qquad$ $L = 1$: light-pipe

The dependencies of a, b and the space bandwidth product N (number of pixels in the image plane) are:

$$a = f + \frac{f}{2 - L} \qquad b = L \cdot f \qquad n = \frac{D}{6.9 \cdot \lambda \cdot f\# \cdot (2 - L)}$$

Asuming $\lambda = 0.65\mu m$, $f\# = 4$ and a lens diameter of $250\mu m$ the figures 5,6 show the graphs for a, b and $n = \sqrt{N}$ (number of pixels in one dimension) depending on the system factor L.

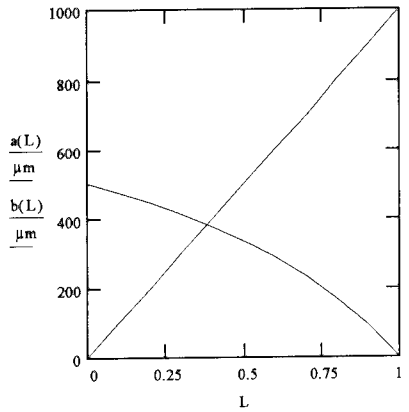


Fig. 5: Distance parameters a and b as function of the system factor L

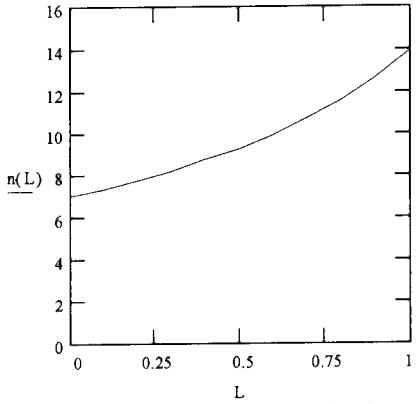


Fig. 6: One-dimensional SWP as function of the system factor.

Assuming a thickness of the prism layer $d_{prism} = 250\mu m$, the minimum distance of a can be $a_{min} = 125\mu m$, thus resulting in $n = 12$ pixel. It has to be noted that the pixel configuration assumes a pixel pitch of two times the pixel size.

5. Conclusion:

Here we investigated the design of folded symbolic substitution systems by considering the constrains of microoptical components. The system design is suitable for the mapping of Systolic Arrays to symbolic substitution rules. The folding of the systems is done by mirrors, beamsplitters in the Fourier domain or within the light pipe. Using common micro lens arrays with the system offer only a limited space bandwidth product ($N = 12^2$ pixel). For multiple imaging architectures with multiple lenses can work around these limitations [7, 8].

[1] W. Erhard, D. Fey *Parallele digitale optische Recheneinheiten*, B.G.Teubner, Stuttgart, 1994
[2] D. Fey, K.-H. Brenner *Digital optical arithmetic based on Systolic Arrays and Symbolic Substitution Logic*, Optical Computing **1**, 153-167, 1990
[3] K.-H. Brenner, Techniques for integration of 3D-optical systems, SPIE Proc. Vol 1544, Miniature and Micro-Optics, San Diego, p. 263-270, 1991
[4] K.-H. Brenner *Three dimensional microoptical integration techniques*, SPIE Vol. 1806, Proc. of the 16th Congress of ICO 1992, Topical Meeting on Optical Computing, Minsk, p. 98-104, 1992
[5] K.-H. Brenner, W. Eckert, C. Passon *Demonstration of a systolic array optical adder based on symbolic substitution*, Optics & Laser Technology, accepted Jan. 1994
[6] K.-H. Brenner et. al *Applications of LIGA components in three-dimensional microoptics*, Appl. Optics, **32** No. 32, 6464-6469, 1993
[7] J. Moisel, K.-H. Brenner *Demonstration of 3D-integrated refractive microsystems*, aural talk WD4 at OC94, Edingburgh
[8] W. Eckert, K.-H. Brenner, C. Passon Considerations for free space microoptical systems in layer architecture, poster WP17 at OC94, Edingburgh

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optoelectronic 3-D Architectures and appropriate Algorithms for TFLOP Computing

D. Fey, W. Erhard, K. Zürl*

University of Jena, Department of Computer Architecture and Communication
Ricarda-Huch-Weg 16, D - 07743 Jena
*University of Erlangen, Department of Applied Optics, Staudtstr. 7/B2, D - 91058 Erlangen

Abstract. A 3-D computer architecture with smart pixels is presented exploiting dense optical interconnections and appropriate algorithms to combine parallel processing techniques as pipelining and array processing. Simulation results specify the needed hardware requirements to enter the TFLOP range.

1. Introduction

The challenge to build parallel computers with a TFLOP peak performance requires not only fast processing elements (> 1ns) but high communication bandwidth (> 1 TBit/s). High integrated electronic logic and high-dense optical interconnections in smart pixel systems fulfil in principle these requirements. 3-D computers with smart pixels offer a large potential of computing performance [1], [2]. Our architecture concept allows the combination of the two fundamental techniques in parallel computing, pipeline and array processing.

An important question in this context is, how complex a simple smart pixel processing element should be? We show that small processing elements are to prefer because they offer more parallelism than larger processing elements. The use of small processing elements does not mean inevitably a loss of computing performance if pipeline mechanisms are consequently exploited. This needs appropriate algorithms and a special smart pixels architecture as we present it in the next two sections. In our investigation we assumed a smart pixels system consisting of in silicon integrated smart detectors and an externally realised sender array with surface emitting microlasers.

2. Bit algorithms

In numerical processing the computation of complex functions as floating-point divisions, sine, cosine, logarithm and exponential function is absolutely necessary. This can be performed with simple processing elements if so-called convergence or bit algorithms [3] are used. These algorithms have already successfully been used on a massively parallel computer like the DAP (Distributed Array Processor). Architecture concepts similar to that of the DAP have often been developed for digital optical and optoelectronic computing [4], [5].

On the DAP the computation time for the above mentioned functions was in the range of 1-3 floating-point multiplications instead of 10-12 if a "traditional" calculation scheme as Taylor had been used [6]. This algorithm class is well suited for smart pixels because only simple operations like shifting, conditional addition and memory access to predefined bits are necessary. This makes a simple setup of the smart pixels possible. The principle of a bit algorithm is the iterative computation of a pair of real numbers (x_i, y_i). The iteration process stops if x_i has reached a known value. Then y_i became automatically the desired function value. This principle is shown in (1) for the example of a division.

$$\frac{Y}{X} = \frac{Y \cdot a_1 \cdot a_2 \cdot \ldots \cdot a_n}{X \cdot a_1 \cdot a_2 \cdot \ldots \cdot a_n} = Q \qquad \begin{aligned} X_{i+1} &= X_i \cdot a_i = X_i \cdot (1 + 2^{-i}) \\ Y_{i+1} &= Y_i \cdot a_i = Y_i \cdot (1 + 2^{-i}) \end{aligned} \quad X_{i+1} \le 1$$

$$X \cdot a_1 \cdot a_2 \cdot \ldots \cdot a_n \to 1 \qquad X_{i+1} = X_i \qquad X_{i+1} > 1 \tag{1}$$
$$Y \cdot a_1 \cdot a_2 \cdot \ldots \cdot a_n \to Q \qquad Y_{i+1} = Y_i$$

By multiplying nominator and denominator with the same multiplying factors a_i, the resulting denominator converges toward unity, and the resulting numerator converges automatically toward the desired quotient. The multiplying factors a_i are selected in such a way that the multiplication is reduced to shift operation and addition. Similar methods are available for the computation of square root, reciprocal values, sine, cosine, logarithm and exponential [3]. The iterative calculation scheme for logarithm and exponential function shows (2).

$$y_n = e^x \qquad\qquad y_n = \ln x$$
$$y_0 = 1, x_0 = x, 0 \le x < 1 \qquad\qquad y_0 = 0, x_0 = x, 0 \le x < 1$$

$$\begin{aligned} x_{i+1} &= x_i - \ln(1 + 2^{-i}) \\ y_{i+1} &= y_i + (y_i \text{ shr } i) \end{aligned} \quad x_{i+1} \ge 0 \qquad \begin{aligned} x_{i+1} &= x_i - \ln(1 + 2^{-i}) \\ y_{i+1} &= y_i + (y_i \text{ shr } i) \end{aligned} \quad x_{i+1} \le 1 \tag{2}$$

$$\begin{aligned} x_{i+1} &= x_i \\ y_{i+1} &= y_i \end{aligned} \quad x_{i+1} < 0 \qquad \begin{aligned} x_{i+1} &= x_i \\ y_{i+1} &= y_i \end{aligned} \quad x_{i+1} > 1$$

The values $\ln(1+2^{-i})$ are precalculated constants. Four basic operations are fundamental for the calculation of a bit algorithm: *addition, multiplexing* to select the right operand pair for the new iteration, *check the condition* $x_{i-1} < 0$ and $x_{i+1} > 1$ respectively, and *access to memory* which stores the constant values and the data shifted to the right about i bits (operator *shr* in eq. 2).

3. A 3-D smart pixels architecture based on bit algorithms

Figure 1 displays the scheme of the optoelectronic 3-D architecture based on bit algorithms. SEL, ADD and FLAG are processing layers, which we mapped onto smart pixel arrays. These smart pixel arrays perform the elementary operations of the bit algorithms. The layer SEL selects the right operand pair (x_i, y_i) or (x_{i-1}, y_{i-1}) for the next iteration. The layer ADD works as a 1-bit full adder. The task of the layer FLAG is to check the above described conditions. A 3-dimensional organised data memory (DATA) is connected to the processing layers. In the i^{th} row of the memory the i^{th} bit of the operand is shifted out first to realise the necessary shift operation of the bit algorithm.

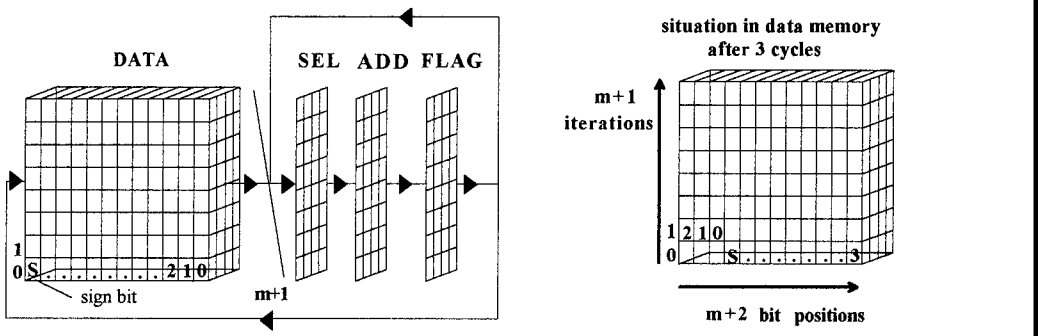


Figure 1: Diagram of the optoelectronic 3-D architecture

The architecture processes data bit-serial and word-parallel. All layers operate in pipeline mode. Each row of the 3-D architecture contains a pipeline. In every pipeline one iteration of the algorithm is executed. If m is the number of the mantissa bits the results appear at the top of the 3-D memory after m+1 iterations. If the architecture is completely filled with data along the vertical axis the results are produced in the same time in which one iteration is executed. Not only one result but so many results are simultaneously generated as pipelines or rows can be integrated along the depth of the 3-D system. Hence, each layer contains an array of pipeline stages. The maximum number of parallel pipelines is determined by the most complex smart pixels array. It turned out that this was the adder layer with 45 transistors, one flip-flop and 2 optical inputs and outputs per one smart pixel.

4. Specifications for the hardware requirements

With the software tool HADLOP [7], a hardware description language that we especially developed for simulation and evaluation of digital optical and optoelectronic architectures, we carried out simulations of the architecture described above to find out specifications for the needed hardware requirements.

Figure 2 shows the result for a setup of the architecture in which the smart detectors and sender arrays are hybridly integrated on a silicon waferboard of 15x15 inch size and the optical interconnects are running through a glass plate mounted on top of the wafer. The left vertical axis shows the computing performance we can expect for the processing of the mantissa bits versus the integration density in the smart detector. The curves correspond to various delay times t_{gate} through one smart pixel array. In this delay time two gates and the optoelectronic interfaces at the input and output side have to be passed. The right vertical axis shows the power loss in the microlaser array for two different assumed power losses in one laser diode. If we assume an integration density of 7000 transistors/mm^2 (state of the art in the DEC ALPHA microprocessor) and a 1-2 ns delay time for one smart pixel array more than 200 GFLOPS would be achievable. Problems arise by the heat transfer of the laser diodes that would be 12 W/cm^2 if 1mW dissipated power per laser diode is assumed.

If a further reduction in power consumption per laser diode is achievable it would be possible to set up a space-saving optoelectronic TFLOP computing system on only 5 wafer boards. Then probably the heat problem lies more in the power dissipation of the smart detector if data rates up to 1 GHz are intended. With sophisticated cooling techniques this problem should be solvable.

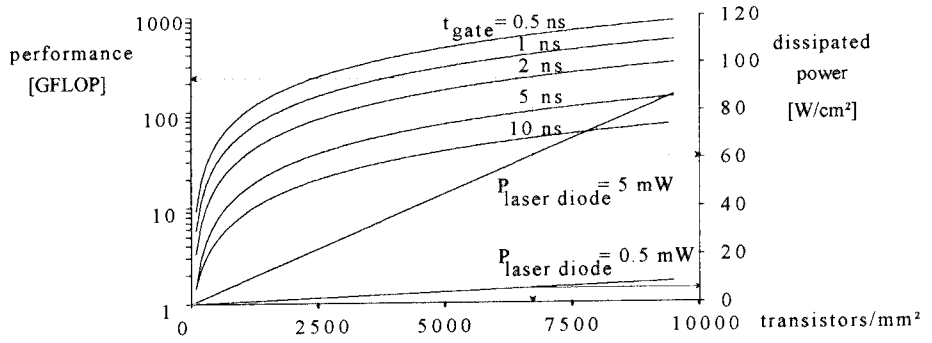


Figure 2: Performance analysis for massively parallel smart pixels architecture

5. Conclusion and future outlook

We presented a concept for a powerful architecture for an optoelectronic massively parallel computer. If the hardware parameters we specified can be fulfilled it is to expect that the performance of such an architecture is higher than in pure electronic parallel systems. This will not only possible by better devices but also by the right combination of algorithm, architecture and intensive exploitation of optical interconnects, as we have shown it here.

The result of our performance estimation to first order shows that it is worthwhile to carry out further investigations towards general purpose architectures. Our smart pixel needs only a small area. Hence, there is enough space to increase the complexity towards programmability and larger local storage capacity. This will lead to 3-D computers with a great communication potential because of optical interconnects. Such architectures are especially suited for solving 3-D problems like for example parallel volume rendering, processes in fluid mechanics and computer simulations using particle models. These are possible application fields for smart pixel computers we will investigate in the future. For an efficient mapping onto real smart pixels systems a close co-operation between device technologists and computer architects is necessary to lead optoelectronic computing to success.

References

[1] J.A. Neff, "3D-Computers with smart pixels", IEEE/LEOS Summer Topical Meeting 92 on Integrated Optoelectronics, Santa Barbara, August 1992

[2] K. Zürl et. al., "Smart Pixels with VCSELs: Potential and Demonstrations System", Conference Proceedings IOC'94, Edinburgh, August 1994

[3] T.C. Chen, "Automatic Computation of Exponentials, Logarithms, Ratios and Square Roots," IBM Journal Res. and Dev., July 1972, pp. 380-388

[4] G. Stucke, "Parallel architecture for a digital optical computer", Applied Optics, vol. 28, no. 26, 1989

[5] M. Ishikawa, "Parallel Optoelectronic Processing Systems and Applications", Conference Proceedings IOC'94, Edinburgh, August 1994

[6] W. Erhard, "Parallelrechnerstrukturen", Teubner 1990

[7] D. Fey, "Computer-aided design of digital opto-electronic systems with HADLOP", in: S.D. Smith, R.F. Neale (eds.), "Optical Information Technology", Springer 1993.

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Planar – free space optical circuits based on the "transverse lock–and–clock" architecture

G V Sinitsyn, F V Karpushko, S P Apanasevich, A V Lyakhnovich, A S Yasukevich, and M A Khodasevich

Division for Optical Problems of Information Technologies,
Academy of Sciences of Belarus,
P.O.Box 1, 220072 Minsk, Republic of Belarus [1]

Abstract. An approach of "transverse interconnection optical processor architecture" is proposed which allows flexible optically programmable data exchange between individual beam channels without changing their relative dispositions.

1. Introduction

Using the light as information carrier implies usually that the direction, in which data are transported in free space or in waveguide, coincides with the light beam direction. However, there is a number of nonlinear optical phenomena such as transverse effects in optical bistability (OB) that allow to direct the information flow perpendicularly to the light beam. When studying such nonlinear transverse phenomena it has been found that there are different steady profiles of the output beam/image intensity for the same distribution of input light interacting with a bistable interference layer. The selection of a desired profile as well as transitions between them can be easily controlled optically. It gives a basis, for instance, for all-optical implementation of 2D–data shift. In this method all the stages of the information transfer are performed in the plane of matrix of optical elements. It provides an opportunity for organizing the interconnections and information exchange between neighbour logic elements/pixels within 2D–array leading to new optical computer architectures.

2. Basic method

In this paper the method of "transverse lock–and–clock" processing [1] is extended to a "planar - free space" optical interconnections and circuits. The basic idea of the

[1] E-mail: dopit%bas02.basnet.minsk.by@demos.su

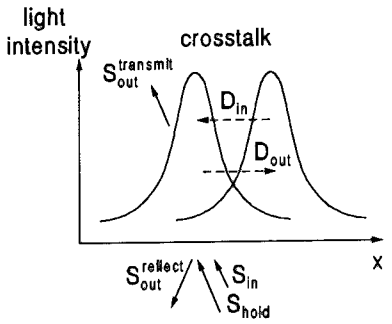


Fig. 1. Base element of the shift processor with longitudinal S_{in}/S_{out} and transverse D_{in}/D_{out} input/otput

"planar – free space" architecture is to combine the parallel data transfer between sequentially located (within loop parallel processor) matrices of switching/bistable elements and planar transverse interconnections and processing within the plane of each matrix.

In such an approach an ensemble of light beams (a digital image) preserves the topology of their relative disposition (with no shuffle, etc.) when propagating in free space between two sequentially located nonlinear matrices, in contrast to other flexible interconnection architectures. The flexibility of the information connections between different parts of the image is reached due to the optically programmable transverse transport and redistribution of the different logical states of transmission/reflection within the matrices of nonlinear elements. As a result, the data in different individual channels are exchanged without changing the spatial positions of the beams that carry the data along the processor loop.

The basic element of such shift processor is a pair of coupled neighbour optical cells (pixels) shown in Fig. 1. In comparison with traditional OB–elements, here arises an additional channel for input of information into a pixel of 2D optical processing matrix due to data transfer from neighbour pixels.

The main features of the shift processor are:

* possibility of transverse data shift simultaneously in counter directions without interaction between the counter propagating data flows ("transparent mode");
* logical operations with the neighbouring digits of 2D matrix;
* input and storage of several 2D data pages as well as of interim results of logical operations in the same OB–plate;
* possibility to design multilayer processing structures with the lateral input/output.

3. Design and experimental results

The schematic layouts of all–optical serial–parallel and parallel–serial data converters, stack memory, multiplexer–demultiplexer, non-blocking crossbar, planar loop circuits and regular networks are given in Fig. 2. Experimental modeling of data converters and

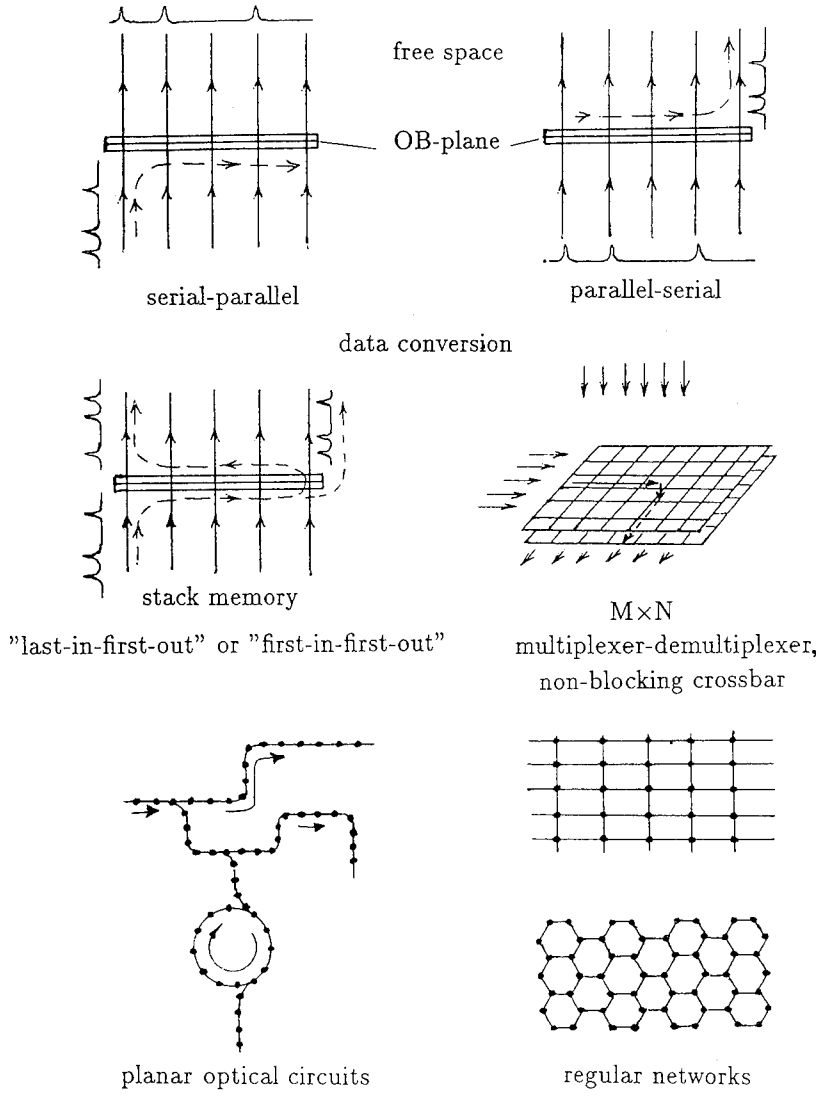


Fig. 2. Optical "planar – free space" logic circuits

stack memory has been performed [1] using 2D bistable thin–film Fabry-Perot interferometer [2] with optical aperture of 36 *mm*.

For practical realizations of "planar – free space" circuits OB–matrices based on media with electron nonlinear mechanisms are more promising from the view point of operation rate. We report here preliminary experimental results on detecting switching waves and measuring their speed in one of such media. In detail these results will be presented later in a separate publication.

All–epitaxial GaAs/GaAlAs optically bistable Fabry-Perot etalons were used that are similar to those described in [3]. Switch–on/off time for single pixel has been mea-

sured as ~ 40 ns [3]. Nonlinear changes in our experiments were monitored in reflected beam. The pulsed generation of a tunable color-center laser at a wavelength 891 nm in the vicinity of interference maximum was focused on the etalon's surface into the spot with diameter of about 200 μm. By measuring local intensities of reflected light in different points along the beam radius (similar to the technique [4]) we estimated the speed of switching wave in the on-axis area of the beam as ~ 1 $\mu m/ns$. Therefore, one can expect that in such OB–etalons parallel transverse shift of information between neighbour pixels could take about 10–50 ns.

For media with picosecond relaxation times this estimation can be smaller by 2–3 orders of magnitude. In this case the fraction of cycle time required for transverse modification (patterning) in OB–matrix of parallel loop processor will be comparable with that required for longitudinal transfer of information between OB–matrices.

4. Conclusion

Approach to architecture design is proposed which combines parallel free space data transfer between sequentially located (within the optical loop parallel processor) OB – matrices with planar transverse interconnections and data processing within the plane of each matrix using transverse effects in optical bistability. Some simple "planar – free space" circuits are modelled. Experimental estimations for speed of switching waves in all–epitaxial GaAs/GaAlAs etalons are made.

Acknowledgments

This work has been supported by the Foundation for Fundamental Research, Republic of Belarus

References

[1] Sinitsyn G V, Apanasevich S P, Lyakhnovich A V, Karpushko F V 1993 *Optical Computing, Proc. SPIE* vol 1806 ed A M Goncharenko *et al* (Washington: SPIE — The International Society for Optical Engineering) p 559–65

[2] Apanasevich S P, Karpushko F V and Sinitsyn G V 1988 *Laser Optics of Condensed Matter, Proc. 3 USA-USSR Symposium* ed J L Birman *et al* (New York: Plenum Press) p 475–80
Sinitsyn G V 1988 *Phys. Stat. Solidi (b)* **150** 455–64

[3] Goodwill D J, Walker A C, Kean A H and Stanley C R 1992 *Electronics Letters* **28** 1599–600

[4] Apanasevich S P, Karpushko F V and Sinitsyn G V 1985 *Sov. J. Quant. Electr. (USA)* **15** 251–3; transl. of: 1985 *Kvant. Elektronika (USSR)* **12** 387–90

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

2D all-optical shift registers: numerical simulation

S P Apanasevich, A V Lyakhnovich and G V Sinitsyn

Division for Optical Problems of Information Technologies,
Academy of Sciences of Belarus,
P.O.Box 1, 220072 Minsk, Republic of Belarus [1]

Abstract. Operation of shift register based on propagation of switching waves in distributed nonlinear media is studied by numerical simulation.

1. Introduction

The concept of lock-and-clock architecture for optical data processing has been recently modified to take into account two transverse degrees of freedom and all-optical shift register based on this concept has been demonstrated experimentally [1]. Information shift in the register takes place in the plane of matrix of optical bistable elements. For further development of this concept an adequate theoretical description would be useful.

2. Theoretical model

In present contribution we describe theoretical model of planar all-optical chip consisting of an array of bistable elements that can be coupled using transverse interconnections. Analysis of the model is based on solving coupled equations for light field in optically nonlinear spacer of Fabry–Perot interferometer and equations that describe phenomenologically diffusion of medium's nonlinear parameters. In general form these equations can be written as

$$\rho_i c_i \frac{\partial U_i}{\partial t} = \kappa_i \nabla^2 U_i + Q_i[I_{in}(r,t)F(U_j), U_j] - \rho_i c_i \frac{U_i}{\tau_i} \qquad i,j = 1,2,\ldots \tag{1}$$

$$F = \frac{(1-R)(1+Re^{-\alpha h})}{(1-Re^{-\alpha h})^2 + 4Re^{-\alpha h}\sin^2\{\pi[\varepsilon + \frac{m+\varepsilon}{n_0}\delta n_i(U_i)]\}} \frac{(1-e^{-\alpha h})}{\alpha h} \tag{2}$$

where U_i – are parameters of the medium which contribute to nonlinear change of refractive index, for instance, carrier density (electrons, holes, excitons), temperature, pressure, etc.; ∇^2 – Laplacian; ρ_i, c_i, κ_i – physical characteristics of the medium, in particular, κ_i determine diffusion of correspondent parameters U_i; Q_i – some source functions for

[1] E-mail: dopit%bas02.basnet.minsk.by@demos.su

U_i; $I_{in}(r,t)$ – incident light intensity distribution; F – function giving the relationship of light field within the spacer of Fabry-Perot interferometer and incident light field; τ_i – time constants describing the relaxation of U_i; R – power reflectivity of mirrors; α – absorption factor; h – thickness of spacer; m – order of interference; ε – initial detuning; n_0 – refractive index; δn_i – change of refractive index associated with change of U_i.

Initial and boundary conditions are adequate to those in typical experimental realizations of OB–layers.

The main assumption of this theoretical model is that all changes of medium's parameters U_i contributing to optical nonlinearity are much slower than light field build-up inside interferometer. This statement is quite obvious, since for the spacer of $1st$–$3rd$ order of interference as in calculations internal light field reaches its steady–state value in $\sim 10^{-14}$–$10^{-12}s$. Dealing with thin-base Fabry-Perot interferometers one may also consider that change of refractive index does not depend on longitudinal space variable and is determined by mean values of parameters U_i averaged over the thickness of spacer.

Since the model is applied to numerical simulation of real experimental conditions where thin-film interferometers (TFI) with Fresnel number of about 10^3 are used, diffraction is negligible and therefore not taken into account.

The approach can deal with various types of nonlinearity, such as band–filling or nonlinearity due to free carrier generation. For each case one have just to define correctly the functions Q_i which are responsible for the rise of nonlinearity.

In particular, on the basis of the above model we have performed numerical simulations for nonlinear thin–film interferometers [2] where nonlinear mechanism is thermo–optical effect. In this specific case U is temperature of the spacer and thermo–induced change of its refractive index can be found by averaging over the thickness of spacer.

$$\delta n_T(T) = \frac{n_T}{h} \int_0^h T(z)dz$$

Function Q has then physical sense of light power absorbed in TFI and appearing as heat, κ is thermal conductivity, ρ – density, c – specific heat, etc..

3. Results and discussion

The model allows to study dynamics of light intensity redistribution in 2D spread OB–plates including well–known transverse effects in OB such as the emergence and propagation of switching waves, formation of steady profiles, etc. For example, Fig. 1 gives evidence of switching waves for uniform distribution of input intensity. The waves are

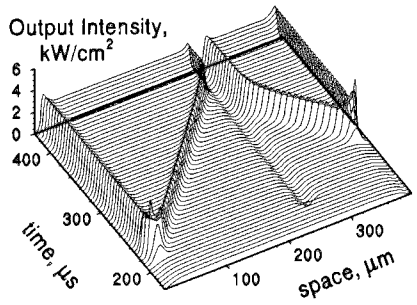


Fig. 1. Switching waves in spread homogeneous OB–layer at uniform distribution of holding input intensity.

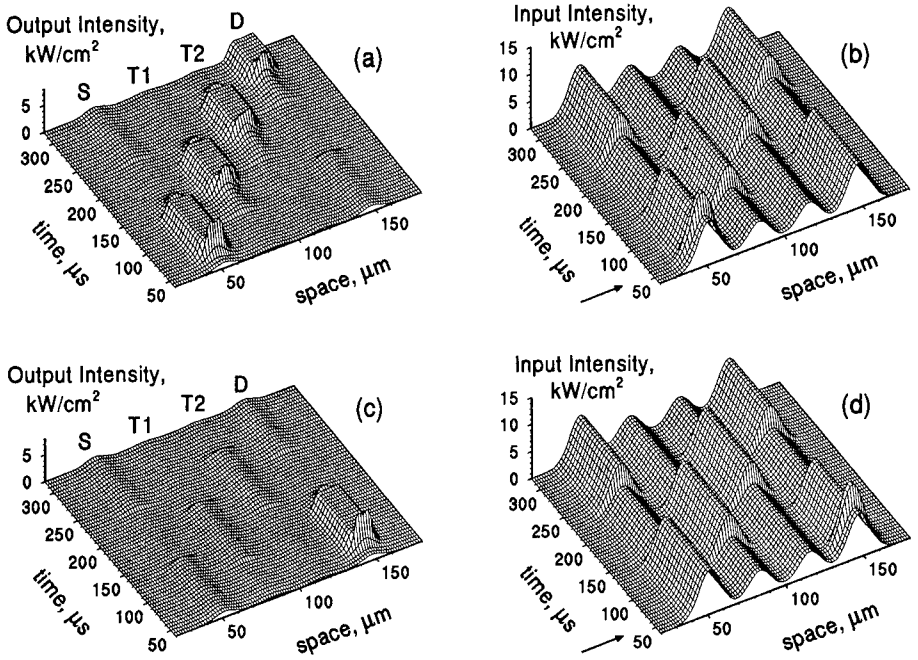


Fig. 2. Spatial and temporal dynamics of shift register during operation cycle. (a, c) – output and (b, d) – corresponding input intensity distributions along the register. Upper plots show "1"→"0" shift operation, lower ones – "0"→"1". Arrows on plots for input intensity distributions point to the stage of recording initial information into main pixels.

triggered by local rises of input intensity near the edges of uniform holding beam. Local drop of input intensity in its turn leads to stoppage of waves' fronts and prevents switching waves from further spreading. Namely this circumstance allows to isolate the information states of neighbour OB-pixels in "transverse lock-and-clock architecture" devices. Speed of stationary switching wave for the particular conditions is $\sim 1.3\ \mu m/\mu s$.

Nonuniform and time–dependent distribution of input intensity results in more complicated behaviour of the system. In particular, using the described theoretical model we have performed numerical simulation of optical data shift in the simplest 4-pixel shift register studied experimentally in [1]. Fig. 2 shows the evolution of input and output intensity distributions along the register. Input intensity is a combination of 4 independently controlled gaussian beams. Notation of pixels is the same as in [1]. From Fig. 2,a one can see that "1"-state of source pixel S is shifted through intermediate transfer pixels T1 and T2 to destination pixel D by means of corresponding intensity modulation in all gaussian beams.

If pixel S is initially in "0"-state, the same modulation sequence shifts this information bit to pixel D which is in "1"-state at the beginning of shift cycle (Fig. 2,c). All other possible combinations of binary data stored in pixels S and D have also been subjected to rightward shift operation. Leftward shift has been performed as well by exchanging modulation sequences of intermediate pixels T1 and T2. For the shown disposition of pixels (radius of each gaussian beam 10 μm, distance between them 32 μm) the entire

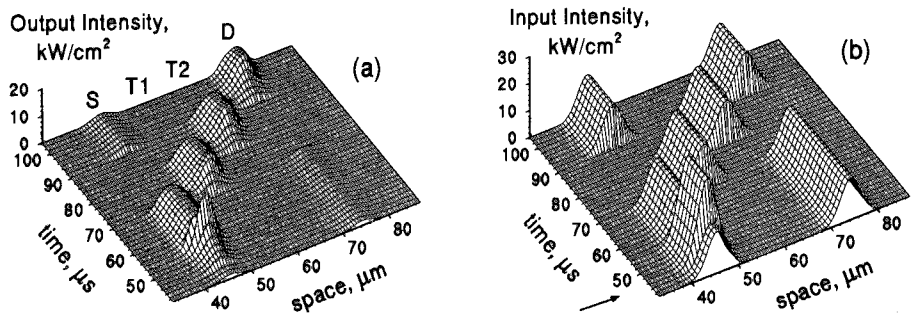


Fig. 3. Dynamics of output (*a*) and input (*b*) intensity profiles along the shift register during "1"→"0" data transfer for the case of sharper focusing of input gaussian beams.

shift cycle takes about 200 μs. This time depends obviously on size of pixels and it is possible to achieve faster operation by placing pixels closer to each other.

Fig. 3 shows all stages of shift cycle "1" → "0" for more sharply focused input beams (radius of 3 μm), distances are made smaller by the same ratio. In this arrangement rise of temperature in pixel S can directly influence the information state of pixel D, therefore range of intensity modulation in all pixels should be also changed to provide the correct operation of shift register. However when the proper parameters of input intensities are found, shift cycle can be reduced to $\sim 30\mu s$, exhibiting nearly square dependence on pixel's size.

By introducing additional gaussian beams into the input intensity distribution, one can translate the simplest 4-pixel shifting chain along transverse dimension. Linear circuits of pixels arranged in such a way may be regarded as multi-digit shift registers. Their operation as well as behaviour of more complicated planar circuits can be also studied numerically on the basis of the above theoretical model.

Acknowledgments

This work has been supported by the Foundation for Fundamental Research, Republic of Belarus.

References

[1] Sinitsyn G V, Apanasevich S P, Lyakhnovich A V, Karpushko F V 1993 *Optical Computing, Proc. SPIE* vol 1806 ed A M Goncharenko *et al* (Washington: SPIE — The International Society for Optical Engineering) pp 559–65

[2] Apanasevich S P, Karpushko F V and Sinitsyn G V 1988 *Laser Optics of Condensed Matter, Proc. 3 USA–USSR Symposium* ed J L Birman *et al* (New York: Plenum Press) pp 475–80

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Image Processing by System Based on an Optoelectronic Active Medium with Controlled Connections.

Yu. I. Balkarey, M. I. Elinson

Institute of Radio Engineering and Electronics, Russia Academy of Sciences,
11 Mokhovaya St., Moscow, 103907, Russia.
Tel: (095) 404 6175; Fax: (095) 203 8414
E-mail: LAB191@IRE.UUCP.FREE.NET

Abstract. New abilities of image processing appearing from the presence of controlled local connections of an inhibitory type between cells in an optoelectronic trigger medium are discussed.

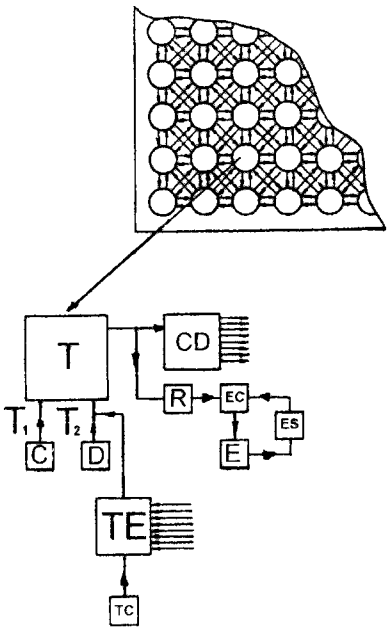
1. Introduction

An optoelectronic memory usually consists of *independent* cells with optical inputs. We consider new abilities of optoelectronic memory arising from the presence of *local connections* between cells. The memory medium is able to extract (process) the informative features of recorded images when we use specific types of connections. Resolution and memory capacity are determined by the microelectronic technology possibilities. We will refer to a memory medium with connected cells as MCC. The possibilities of creating solid- state continual media and information processing by them are discussed in [1-3]. Microelectronics analogue realizations of homogeneous media of active elements with local connections, which were called Cellular Neural Networks (CNN), were proposed in [4].

2. Design of the cell

Consider the binary optoelectronic memory shown overpage. Every cell comprises a trigger T with an optical input and output. The image projected onto the medium is detected by photocell (C), and is recorded by trigger flip- over from state T_1 to state T_2 ($T_1 \rightarrow T_2$). The circuits (photocells) for reading (R) and erasing (D) are switched off during recording. Uniform illumination of all R turns on light emitters (E) of all cells where trigger T is in the state T_2; the photocells C and D are turned off. It restores recorded images. The picture is destroyed by flip- over $T_2 \rightarrow T_1$ by uniform illumination of all D.

Every cell is connected with its nearest eight neighbors (here we discuss the square array). The connections are of inhibitory type; they block up connected cells, i.e. prevent $T_1 \rightarrow T_2$ flip- over. A special threshold element (TE) (coupled with D to the trigger arm T_2)



combines all connections from neighbouring cells. The threshold control unit (*TC*) sets the number of connections (*N*) necessary for a signal to appear at the *TE* output. This signal switches the memory cell off ($T_2 \rightarrow T_1$) in the same way as the *D* signal does.

The system also includes a connection distributor (CD) and light emitter control unit (*EC*). When *E* is turned on, after the delay time the *ES* unit switches the *E* off for the whole image analysis cycle.

The *main feature* of MCC with the inhibitory connections is that the memory state of the chosen cell depends only on the number of illuminated neighbour cells which prevent $T_1 \rightarrow T_2$ flip- over by their connections. Of course, there is also a reverse inhibitory effect of the chosen cell on its neighbors.

There are the obvious temporal restrictions: if τ_1 is the internal time of the whole cell, τ_2 is the time of connecting of cells, and τ_3 is the characteristic time of the image changing, then $\tau_1 << \tau_2 << \tau_3$.

3. Image processing by medium with local connections

We assume that the minimal size of a picture element is of the same order as of a cell size and the distance between cells is substantially smaller than the linear cell size.

When N = 1, the memory records only isolated elements of a size of the order of a cell, with at least a one cell spacing. It is not possible to record other picture elements, because even one inhibitory connection prevents such a process. The addition of a new neighbouring illuminated element to the one- cell spot leads to local information destruction.

When N = 2 the memory extracts both 1- cell and arbitrary 2- cell spots. If N = 3 it is also possible to record 3- cell spots and arbitrary lines of a one cell thickness, when the neighbouring elements are connected only by two inhibitory connections. With the increasing

of N the minimal size of recorded spots is increased, and cells in the off state can appear inside illuminated spots.

Large- pattern contours can be extracted when N = 5: inner cells are inhibited, whereas contour cells (which have no illuminated neighbours outside the pattern) are not inhibited.

Analysis of pictures, maps etc. is best carried out by "step by step" cycle increasing N from 1 to 8 with exclusion of fragments extracted at the (N - 1) th step. In fact MCC can be used for decomposition of the initial picture into fragments: small spots, contours of large spots, lines, line crossings, angles, etc.

4. Extraction of moving fragments

An important possible application is extraction of moving fragments. The initial picture is projected onto the MCC and then all corresponding *E* are turned off by unit *ES*. Then the initial picture with shifted elements is projected again. Obviously, only emitters of the cells corresponding to recorded shifted elements are turned on. The subsequent projecting operations make it possible to determine the directions and velocities of moving elements. This method is also suitable for extraction of motionless but distinct parts of patterns.

5. Possible development

The possible applications of MCC are expanded if every cell is connected not only to the nearest cells or not only by inhibitory connections or if not only square arrays are discussed. Activating excitatory connections (with inhibitory connections switched off) for a certain time it is possible to enlarge the size of extracted spots and the width of lines, and to spline indented edges.

6. Acknowledgments

The authors are grateful to the Russian Fund of Fundamental Researches for financial support (project No. 94-01-01538).

References

[1] Balkarey Yu I, Elinson M I and Nikulin M G 1985 In: *Problems of Modern Radio Engineering and Electronics.* Ed. V.A. Kotelnikov (Moscow: Nauka) 293- 318

[2] Balkarey Yu I, Evtikhov M G and Elinson M I 1991 *Proceedings SPIE* **1621** 238- 249

[3] Krinsky V I, Biktashev V N and Efimov I R 1991 *Physica D* **49** 247- 253, 1991

[4] Chua L O and Yang L 1988 *IEEE Transactions on Circuits and Systems* **CAS 35** 1257- 1272

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Parallel Image Processing by Media of Resonantly Excited Oscillators

Yu.I. Balkarey, A.S. Cohen, W.H. Johnson*, M.I. Elinson, M.G. Evtikhov, P.I. Perov.

Institute of Radio Engineering and Electronics, Russia Academy of Sciences,
11 Mokhovaya St., Moscow, 103907, Russia.
Tel: (095) 404 6175; Fax: (095) 203 8414
E-mail: L191@IRE.UUCP.FREE.NET
* Physics and Engineering Department, Suffolk University,
41 Temple St., Boston, Massachusetts, 02114 USA
Tel: (617) 573 8602 Fax: (617) 573 8513
E-mail: W.JOHN@ACAD.SUFFOLK.EDU

Abstract. A system of oscillators with optical inputs and outputs transforms a projected image into spatial distribution of oscillator frequencies. Resonant excitation of different groups of oscillators is discussed as a new method of image processing.

1. General idea

The main idea of the discussed approach is the following [1,2]. Suppose there is a matrix of identical quasiharmonic noninteracting oscillators with optical inputs and outputs. The image I(X,Y) is projected onto the matrix and it leads to the change of oscillators' frequencies ω_0.

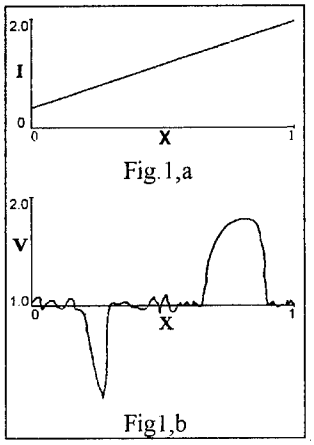


Fig.1,a

Fig1,b

$$\omega_0 \Rightarrow \omega_{loc} \equiv \omega_0 + \alpha I(X,Y)$$

Hence the image is encoded into the spatial distribution of frequencies.

Then the homogeneous pumping is projected onto the medium through a wide aperture modulator with a harmonic regime of modulation (frequency Ω). The pumping excites only such parts of the medium which correspond to the condition: $\omega_{loc}(X,Y) \approx \Omega$. The parametric resonance can be used by the same way.

Our results do not depend on concrete model which describes the oscillators. Hereinafter we use the Van- der-Pol like model in the regime of quasi harmonic damping

oscillations. Variable V describes the amplitude of the medium response.

The parametric excitation of one dimensional medium with a gradient I(X) (Fig.1,a) by two frequency pumping illustrates our method in Fig.1,b. The widths of resonant stripes depend on Ω as well known in the theory of parametric resonance. Varying Ω we can provide the extraction of regions of equal intensity. If oscillators' inputs are sensitive with respect to the light wavelength it is possible to extract regions of certain color.

Example of a problem which can be solved by the proposed method. A sheet of paper contains a set of patterns drawn with different intensities (or colors). Projecting the sheet onto the medium and using a set of appropriate Ω it is possible to extract every pattern.

2. Abilities of resonance method

We have studied different modifications whereas the additional elements are used: passage of analyzing pictures through a modulator; periodic shift of picture along the medium; using of the inhomogeneous background light pumping. Such complications allows us to realize the following procedures: extraction of certain parts of picture; removing small scale distortions from patterns; extraction of contours, points of extrema; correlative comparison of images; determination of common and distinct parts of images, moving elements, etc.

Below several possibilities are discussed in details.

2.1. Extraction of common and distinct parts of images

Two binary pictures are projected onto the medium (onto the same place) In the common parts we have $\omega_{0,1} = \omega_0 + 2\alpha J$, in the distinct parts- $\omega_{0,2} = \omega_0 + \alpha J$, J is unity level. By resonance excitation one can extract the common parts of images if $\Omega = \omega_{0,1}$, when $\Omega = \omega_{0,2}$ the distinct parts may be excited.

2.2. Extraction of image contour and moving fragments

With projecting the image and its copy enlarged a little by optical system and adjusting to resonance at $\Omega = \omega_{0,2}$ it is possible to extract the contour. The method can be generalized easily for analysis of gray- toned images and moving fragments. In the latter case one has to project the same picture onto the medium in a set of different moments and use resonance excitation for the extraction of shifted elements only.

2.3. Halftoning

If Ω is chosen in such a way that the medium extracts only the most intensive parts (peak 1 in Fig. 2,a) the image encodes into a binary one (Fig.2,b). For Ω corresponding to the resonance with peak 2 the result of binarization is shown in Fig. 2,c.

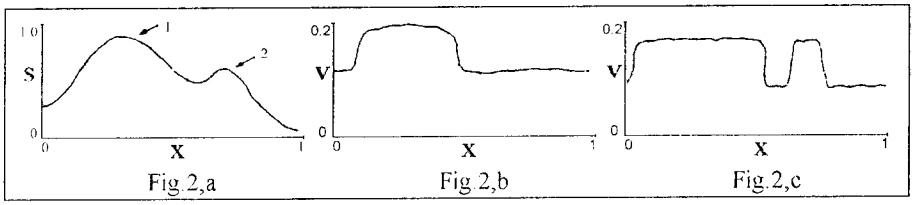
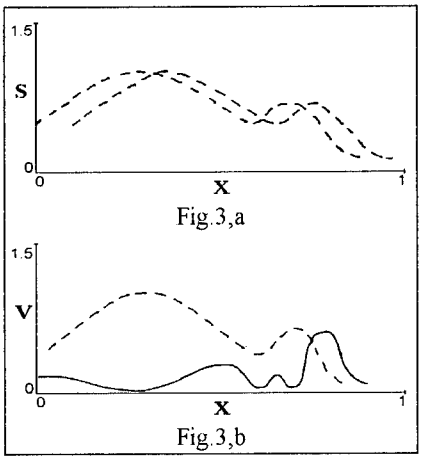


Fig.2,a Fig.2,b Fig.2,c

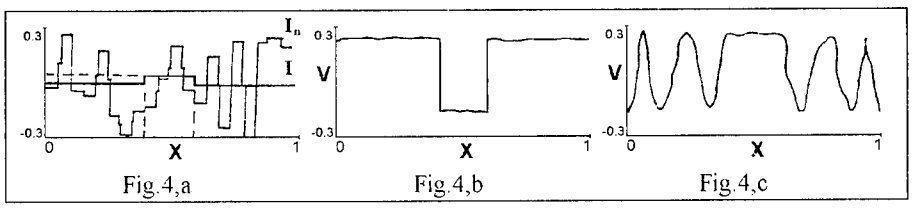
2.4. Extraction of extrema



A one dimensional image S is projected onto the medium through a modulator and simultaneously the projection oscillates in space along the medium (Fig. 3,a). There is a phase shift $\pi/2$ between two types of oscillations and their frequencies are equal. It is possible to demonstrate that in Van- der- Pol like model the amplitude of medium response is proportional to $\frac{dI}{dx}$. In the extrema $\frac{dI}{dx} = 0$ and amplitude is equal to zero too, therefore these points may be extracted. The result of modeling is shown in Fig. 3,b. Analysis of two dimensional functions I(X,Y) is possible by the same way changing the direction of the picture tremble step by step followingly.

3. Image processing and self sustained oscillations (extraction of images from dynamic noise)

Also it is possible to use non resonant methods for image processing. Consider unconnected oscillators in a self- oscillating regime. For simplicity we discuss the one- dimensional case. The oscillatory medium is excited by a weak binary image I(X) and by intensive noise fluctuations I_n (X,t) (Fig. 4,a). For many periods of self- oscillations the image I(X) (which locally changes the frequency ω_0) gives a large accumulation of inhomogeneous local phase shifts $\Delta\varphi(X)$. The noise gives local frequency shifts of different signs and can be fully averaged during many periods. Therefore in the medium response noise can produce only a constant homogeneous phase shift determined by the statistical properties of the noise. The result of numerical modeling for the certain moment (when $\Delta\varphi(X) \sim \pi$) is shown in Fig. 4,b. After the information signal I(X) is switched off, the medium becomes a memory since the self- oscillations do not damp. Very weak signals I(X) can be extracted by this method of phase accumulation but the time of extraction becomes very large. We modeled also image extraction by a medium of self- oscillators connected by diffusion. When the image I(X) was switched on for a long time, the local phase shifts became spread and spatio- temporal oscillating halo around the image were formed (Fig. 4,c). This feature can be useful for some applications, for example, for making an enlargement of small isolated picture elements. After switching the light off the phase distribution relaxes due to diffusion. All mentioned earlier is also valid for two dimensional case.



4. Realization of media of oscillating elements

Media of oscillating elements can be realized in several ways.
a) Matrices of microcavity lasers. There are structures in which every microcavity laser is integrated with a photocell for noncoherent light [3]. Informative light flow is percepted by the photocells and changes the frequency ω_0 of laser electron- photon resonance ($\omega_0 \approx 10^9$ s^{-1}). Existing electrooptical modulators can excite this resonance. Matrices of microcavity lasers without photocells can also be used when the coherent informative light flow is used.
b) Matrices of passive optical bistable elements in regime of optical oscillations [4] can be used in the same way.
c) Analog microelectronic matrices containing oscillating circuits with optical inputs and outputs are convenient too. For example, circuit which converts light intensity to signal frequency, is discussed in [5].

5. Acknowledgments

Authors (Yu.I.B., A.S.C., M.I.E., M.G.E., P.I.P.) are grateful to the Russian Fund of Fundamental Researches for financional support (projects No. 94-01-01538).

References

[1] Balkarey Yu I, Cohen A S, Johnson W H, Elinson M I, Evtikhov M G, and Perov P I 1994 *Russian Microelctronics* **24** 3- 17
[2] Balkarey Yu I, Cohen A S, Johnson W H, Elinson M I, Evtikhov M G, and Perov P I 1994 *Proceedings of IEEE World Congress on Computational Intelligence*. Sect. Machine Vision II, Paper Number 28. IEEE Press
[3] Chan W K, Harbison J P, von Lechmen A C, Florez L T, Nguen G K, and Schwarz S A 1992 *Appl. Phys. Lett.* **58** 2342- 2344
[4] Ikeda K and Akimoto O 1982 *Phys. Rev. Lett.* **48** 617- 620
[5] Tanaka K, Ando F, Taketoshi K, Ohishi I and Asari G 1993 *Jpn. J. Appl. Phys.* **32 pt.1** 5002- 5007

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Polarization-optical stacked integration scheme of cellular logic two-layer image processor

Liren Liu, Yaozu Yin, Haifeng Peng, Bingquan Wang, Wei Xiong, Ning Wang, Feng Liang, and Liying Zhao

Information Optics Laboratory, Shanghai Institute of Optics and Fine Mechanics, Academia Sinica, P.O.Box 800-211, Shanghai 201800, P.R. China

Abstract. A 3-D polarization-optical stacked integration technique is developed, by which optical processors are packaged to solid-state modules from the cascading of building blocks of birefringent crystals and spatial light modulators. Designs and constructions of the modules of morphological processor, self-programming logic processor, beam array generator, and their cascaded combination to cellular logic image processor are demonstrated.

1. Introduction

Free-space optics offers the potential for building optical processors. For the practical use, the packaging of optical components and devices into functional compact solid-state modules is necessary. Optical interconnects in the free space can be implemented by geometric optics[1], diffractive optics[2], and micro-optics[3]. There were many suggestions for compact packaging. However, each packaging technique assembles the modules for a special-purpose application. So, it is interesting to develop a universary packaging technique to assemble processor from unique building blocks.

In this paper, we propose a new packaging technique that to fabricate functional building blocks of polarization-optical components for interconnection and of devices for processing and interface, then to assemble compact solid-state modules of processors by stacking different building blocks, and, further, to construct complex modules by cascading processor modules. The designs and experiments of a cellular two-layer logic image processor module is shown, which is a cascaded integration of the modules of morphological processor, self-programming logic processor, and beam array generator. The features are: (1) The crystal building blocks are simple plates. (2) The stacked modules are compact and insensitive to environment. (3) Calcite and other crystals have high extinction and transmission radios. (4) The interference is minimized with orthogonal polarizations. (5) The alignment of crystal blocks is not critical. (6) Cascaded stacking of processor modules is possible.

2. Basic building blocks

Calcite plates, quartz polarization rotators, waveplates, and polarizers are used as the building blocks for optical interconnection. The function of a calcite plate for beam splitting or combination and of a quartz rotator are seen in Fig.1. Using calcite, the maximum packaging density is about 30(spots/mm). Solid-state electro-optical SLMs with electric dual-rail addressing are developed as shown in Fig.2. The other SLM used is the PROM which can be used as either optical interface device or thresholding device.

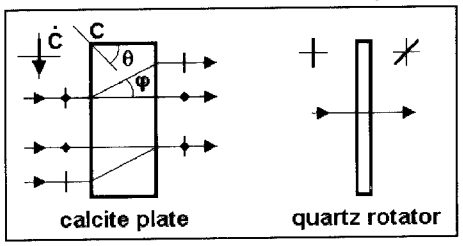


Figure 1

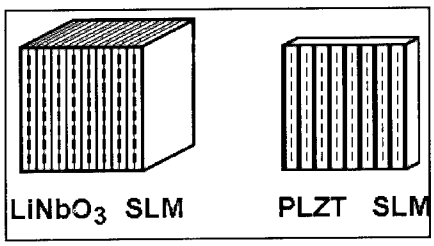


Figure 2

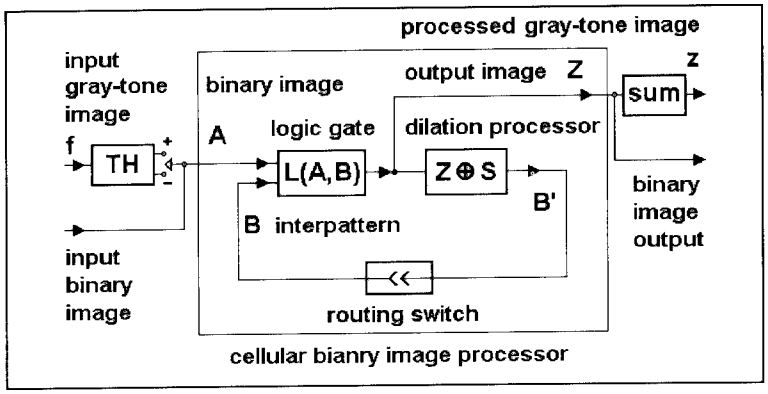


Figure 3

3. Architecture of cellular two-layer logic image processor

Cellular logic image processors[4] have become an important architecture for optical computing[5-8]. Based on mathematical morphology[9], we have proposed an one-operation image algebra and a cellular two-layer logic binary image processor[10]. All kinds of binary morphological processing and pattern-recognition functions can be realized simply by the logic programming in an iterative mode. By using threshold decomposition and sum-superposition various morphological gray-tone image processing functions can be achieved in a double iterative mode[11]. Fig.3 shows the architecture. The elementary CPU is a cellular two-layer logic binary image processor which consists of a parallel binary logic processor followed by a morphological binary dilation processor.

4. Optical morphological processor module

The morphological processor model is shown in Fig.4, which consists of three calcite plates of the thickness ration 1:1:$\sqrt{2}$. An input binary image will be splitted into four images. This is equivalent to a structuring element of {0,1;1,0;0,-1;-1,0}. A thresholding level of 1 or 4 results in dilation or erosion.

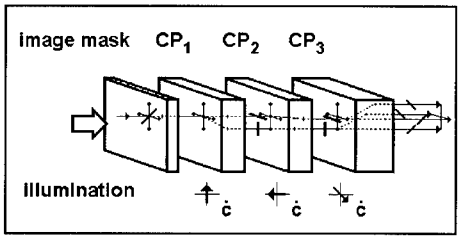


Figure 4

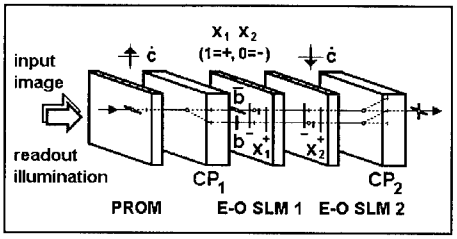


Figure 5

5. Module of self-programming logic processor

Based on dual-rail coding of multiple variables, we suggest the self-programming logic. The two coded patterns are a pair of an image and its negation (B_1 and B_2) and a pair of programming patterns (X_1 and X_2), which may be the other image, its negation, a bright background, or a dark background. The different access of X_1 and X_2 leads to all the sixteen logic operations. Fig.5 shows the module. A PROM is used to code an input image into a collimated pattern with dual-rail format. Two dual-rail SLMs are used to enter x_1 and x_2. The two accessing beams pass sequentially through the two SLMs, and then combined together by the calcite plate CP_2.

6. Array illuminator module

A fundamental stage consists of two identical calcite plates but with the optical axes orthogonal, and an incident beam will become 2x2 beams. Thus, N stages with multiplied thickness of calcite can produce an array of 2Nx2N beams.

7. Module of cellular two-layer logic image processor

Fig.6 gives the configuration of the solid-state system module assembled by the cascaded stacking of the individual modules and beam splitters. An array of photodetector converts the 2-D light signals to electrical signals. A PC computer executes the dilation thresholding and the dual-rail coding for the E-O SLM in the logic processor module.

8. Design and experimental observation

Two series of calcite plates are fabricated: for 1mm or 2mm separation the

thickness is 9.2mm or 18.4mm. The apertures are 22mmx22mm. For a quartz 45° polarization rotator, the thickness is about 2.4mm. The size of $LiNbO_3$ slabs is 1mmx16mmx8mm, the measured half-wave voltage is nearly 400V.

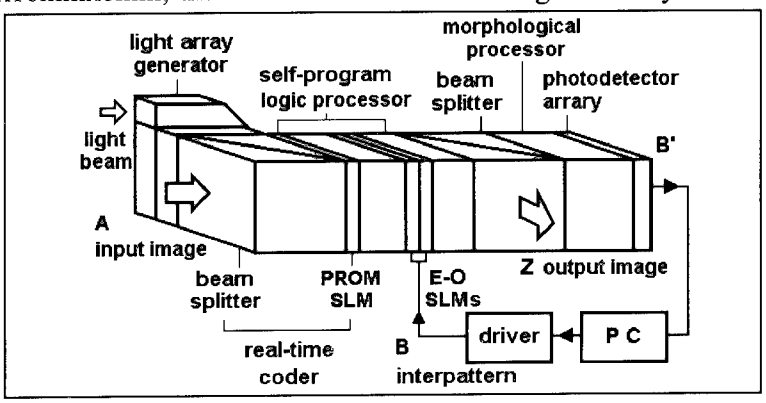
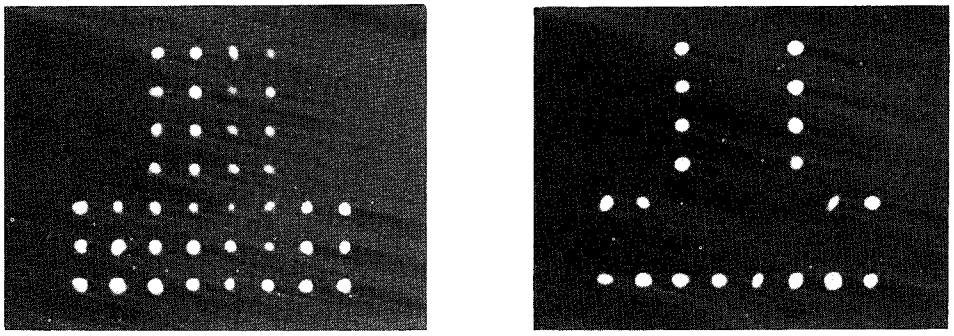


Figure 6

An input image of 8x8 pixels and the processed pattern of edges are seen as:



Acknowledgement

This work was supported by the National Natural Sciences Foundation of China, the Bureau of High Technique of China, and Chinese Academy of Sciences.

References

[1] A.W.Lohmann, 1986, Appl. Opt. **25**, 1534.
[2] J.W.Goodman, F.J.Leonberger, S.Y.Kang, and R.A.Athale, 1989, Proc.IEEE, **72**, 805.
[3] K.Iga, M.Oikawa, S.Misawa, J.Banno, and Y.Kokubun, 1982, Appl. Opt. **21**, 3456.
[4] Jr.K.Preston, M.J.Duff, S.Leviadi, P.E.Norgren, and J.Toriwaki, 1979, Proc.IEEE, **67**, 826.
[5] R.G.A.Craig, B.S.Wherrett, A.C.Walker, F.A.P.Tooley, and S.D.Smith, 1991, Appl.Opt. **30**, 2297.
[6] F.A.Tooley, and S.Wakelin, 1993, Appl. Opt. **32**, 1850.
[7] L.Liu, X.Liu, and B.Cui, 1991, Appl. Opt. **30**, 943.
[8] K.S.Huang, A.A.Sawchuk, B.K.Jenkins, P.Chavel, J.-M. Wang, A.G. Weber, C.-H.Wang, and I.Glaser, 1993, Appl.Opt. **32**, 166.
[9] P.Maragos, and R.W.Schafer, 1990, Proc.IEEE, **78**, 690.
[10] L.Liu, Z.Zhang, and X.Zhang, 1994, J. Opt. Soc. Am. A, (in press).
[11] L.Liu, X.Zhang, and Z.Zhang, 1994, Appl. Opt. **33**, 4383.

Inst. Phys. Conf. Ser. No 139: Part I
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical digital data processing based on the photon echo phenomenon

E.A.Manykin,

Russian Research Center "Kurchatov Institute", 123182, Moscow, Kurchatov Square, Russia

S.M.Zakharov

Institute for High-performance Computer Systems of the Russian Academy of Sciences, 117334, Moscow, Leninsky prospect, 32A, Russia

Abstract. Principles for optical digital data processing based on photon echo phenomenon in resonant media are considered. The various schemes of optical processors realization with the pixel structure and of holographic type based on using the digital multiplication by the analog convolution (DMAC) algorithm is suggested. Methods of optical data flow switching of informational channels on the basis of the photon echo phenomenon are discussed. Interconnection switching schemes in sequential in time and parallel in space codes are suggested

1. Introduction

In recent years the photon echo (PE) phenomenon has attracted attention as not only the method of kinetic relaxation processes studying in a resonant media, but also as the mean for multifunctional data processing presented as optical images formed by the excitation pulses at different time moments. Primary proposals of PE using concerned in the main of analog processing methods [1-3]. These methods are characterized by very high-speed processing and can be used in the problem of pattern recognition [4,5]. However there is interest in development of optical information processing methods based on the PE phenomenon for the realization of digital data processing and for optical digital processors design.

In this paper we analyze the optical processing for realization of vector and vector-matrix algebra operation as intrinsic (scalar) product calculations, vector-matrix and matrix-matrix multiplication.

As the main principle of optical digital information processing the *digital multiplication by analog convolution* (DMAC) algorithm is usually used [6,7]. The DMAC algorithm have got a wide extension in acousto-optical processing [8]. In accordance with this method the arbitrary number can be represented in the binary form and as a result the multiplication looks as in mixed binary form. Then it is necessary to convert the mixed binary number into conventional binary representation. Binary words of the mixed digits are added after being shifted with respect to each other.

2. Photon echo based digital processor for vector-matrix multiplication

Due to unique dynamical in space and in time properties of PE the coherent responses can be successfully applied to the data processing systems. To utilize the DMAC algorithm in the PE phenomenon it is necessary to make binary incoding in time or in space. We consider two types of optical schemes with using this algorithm in time-domain and space-domain fields [9].

2.1 PE pixeled vector-matrix multiplier with spatial integration

In the first case the binary representation of the multiplied numbers is accomplished as temporal sequential code and the processing takes place in the independent space structural regions (pixels) (fig.1). In each pixel of the resonant medium the convolution or correlation functions corresponding to two pulses in time are calculated. On the whole, the spatial-temporal light modulator (SLM) has to form different images, the number of which is equal to the significant digit of using representation for under the time action of exciting light pulses. To work in this way the SLM should have electrically independent addressing of each pixel. The adding operation of the different components is accomplished by cylinder lens.

2.2 PE holographic vector-matrix multiplier with spatial integration

In the second case (the so-called scheme of holographic type) all numbers are introduced in parallel code and DMAC algorithm is realized in space. A compound optical scheme accomplishes one-dimensional Fourier transformation on one coordinate and other coordinate forms optical image and compensates the phase distortion (fig.2). In this case the resonant medium is a spectral plane of multichannel dynamic hologram with the one-dimensional filtering of spatial frequencies. As a result, the data of products of different vector components A_iB_i in the mixed binary representation is formed in output plane. The cylinder lens executes the addition operation and the signal that is proportional to the scalar multiplication of two vectors **A** and **B** is arrised on the linear array of photodetectors.

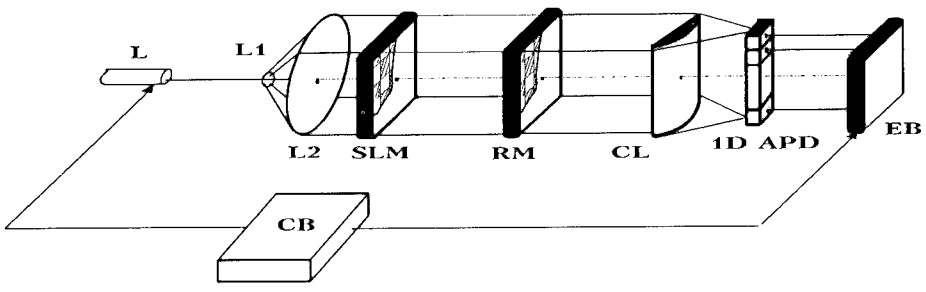


Figure 1. PE pixeled vector-matrix multiplier with spatial integration. L - laser, L1 and L2 - lenses to form plane wave front, SLM - spatial light modulator, RM - resonance medium, CL - cylindrical lens for spatial integration, 1D APD - 1D array of photodetectors, EB - electronical block for data processing, CB - control block.

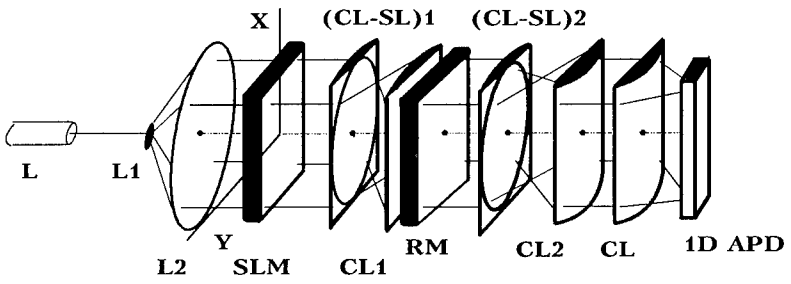


Figure 2. A scheme of PE holographic vector-matrix multiplier. L1 and L2 - collimating lenses to form plane wave front, SLM - spatial-temporal light modulator, (CL - SL)1,2 - combination of spherical and cylindrical lenses to make the optical image over Y and Fourier transform over X, CL1,2 - cylindrical lenses for phase compensation, RM - resonance medium, CL - cylindrical lens for spatial integration, 1D APD - 1D array of photodetectors combined with a data processing electronic block.

If during the action of some of the incident laser pulse the SLM generates a set of different frames that corresponds to the assignment of vector and matrix data, then the output data stream from the resonance medium provides series of signals in time.

Thus, the digital data processing principle is based on filtering in a resonant medium general time frequencies in processor with the pixel structure and in processor of holographic type - on multichannel one-dimensional filtering of spatial frequencies.

3. Interconnection data flow switching based on the PE phenomenon

Optical interconnection elaboration is a key problem in implementation of informational network, possessing by the higher passing capacity. In the particular case of digital data processing the data flow switching may act as a control data transmission, that is information channel switching.

Here we offer some principles based on the using PE phenomena. Data flow switching can be easily implemented on the principle of vector-matrix multiplication. In part 2 we considered optical processing for vector-matrix conversions. In particular, two schemes of binary representation of inputting numbers with using of sequential in time and parallel in space codes are proposed.

In the first scheme the setting of vector components is initiated by the specification of images with identical row and column numbers being equal vector dimension, and addition operation is accomplished by the cylinder lens.

In accordance with the principle of vector-matrix interconnection switching in definition of matrix **A** with the only element differed from zero, for example, $A_{ik} = 1$ for i=k, and the other $A_{ik} = 0$, the control data flow switching of channels with numbers i and k occurres (information role of which plays the vector components) and input data of channel number k fall to the channel i on output. In this case each vector **B** of the image row is introduced in a sequential code and thus each component of vector **B** corresponds to definition of

information channel k. As a result on input system the number of information message is equal to the number components (dimension) of vector **B**, and the number of corresponding column of image of identical vectors coincides with the channel number. The photodetector sequence on output of system forms the set of output channels in each of which the only message from B_k (k=1,2,...N) falls.

In the scheme of holographic type each image formed by the spatial-temporal light modulator represents the vector with the component number on one coordinate equal to dimension of defined vector and with the number of binary digit on the other coordinate. Thus, so far as matrix **A** defines under the time action of one of light pulses in the form of M images (that corresponds to the definition of M vectors or data matrix), that is in the form of different vector sequence in time then the channel numeration on input must correspond to definition of number sequence in time.

We shall call the sequence being normally ordered when channel number with large numerical number is introduced in time with more delay. Therefore to accomplish the input channels it is necessary to define the control matrix in the form of image sequence with the only transmission window.

As a result on linear photodetector array in a parallel code sequentially in time components B_k of vector **B** will arise. Besides this sequence in time will not already normally ordered and numbers of output information messages will correspond to position of transmission windows of matrix **A**.

Thus, if in the first case (in optical scheme with the pixel structure in temporal data coding) the channel switching is occur in space, in the second case (in optical scheme of holographic type in spatial data coding) - in time.

4. Conclusion

Thus, the possibilities of using PE phenomenon in optical information processing systems may be substantially extended by means of digital data processing. It will be useful also for simultaneous using the optical parallel processing in combination with the accuracy and high speed.

References

[1] Zakharov S M and Manykin E A 1988 In: *"Problemy kvantovoi optiki"* (OIYaI, Dubna) 44-48 [in Russian]
[2] Manykin E A and Zakharov S M 1989 *Izv. Ak. Nauk SSSR, ser. fiz.*, **53** 2281 -2286
[3] Xu E Y, Kroll S, Huestis D L, Kachru R and Kim M K 1990 Opt. Lett. **15** 562 -564
[4] Shen X A and Kachru R 1992 Opt. Lett. **17** 520 -522
[5] Shen X A and Kachru R 1993 *Appl. Opt.* **32** 5810-5815
[6] Babbitt W R, Bai Y S and Mossberg T W 1986 *Proc. SPIE* **639** 240-247
[7] Manganaris D, Talagala P and Kim M K 1992 *Appl. Opt* **31** 2426-2429
[8] Kulakov S V, Kulakov V S, Preslenev L N and Tigin D V 1988 *Zarubezhnaya elektronika* **12** 30-40
[9] Zakharov S M 1994 *Izv. RAN, ser. fiz.* **58** 123-128

II. OPTICAL INTERCONNECTS AND SWITCHING NETWORKS

OPTICAL TECHNIQUES

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Free-Space Optical Switching Using FET-SEED Smart Pixel Arrays

F. B. McCormick*, A. L. Lentine, R. L. Morrison, J. M. Sasian, T. J. Cloonan, R. A. Novotny, M. G. Beckman, M. J. Wojcik, S. J. Hinterlong, and D. B. Buchholz.

AT&T Bell Laboratories, 263 Shuman Blvd., Naperville, IL 60566
*This work was done while F. B. McCormick was with AT&T Bell Laboratories. His current address is University of California, San Diego, Department of Electrical and Computer Engineering, 9500 Gilman Drive, La Jolla, CA 92093-0407, email: fmac@ece.ucsd.edu

Abstract

High capacity (hundreds of channels with >100 Mb/s per channel) switching systems frequently encounter difficulties in interconnect packaging within the switching fabric. These interconnection problems may be alleviated through the use of surface normal optical interconnections using optoelectronic smart pixels. Two recent experiments with FET-SEED smart pixels culminated in the operation of a 5 stage free-space switching network at a 155 Mb/s channel rate. The network used embedded control techniques for network control and the smart pixels consisted of GaAs FET logic with MQW SEED detectors and modulators. The system also incorporated external cavity lasers, bulk, micro, and diffractive optics, two-dimensional fiber bundles, and novel optomechanics. At 155 Mb/s, 77 of the 80 total pixels in the system and 31 of the 32 input fibers were functional. Two of the network paths have carried digital video at 105 Mb/s for over 6 months without readjustment. Error rate measurements on these paths have shown a short term BER of 10-10.

1. Introduction

Smart pixel optoelectronic device arrays represent a potential solution to the interconnection "bottleneck" problem present in many high performance digital systems. By utilizing both high speed electronics and surface-normal optoelectronics in 2-dimensional arrays, smart pixels enable the interconnection and communication advantages of optics to complement the processing power of electronics in computing and switching applications. A recent system prototype incorporated GaAs multiquantum well (MQW) FET-SEED smart pixels to implement the multi-stage switching network shown in Fig. 1.[1] This 32-input, 16-output multi-stage switching fabric used 5 stages of 4x4 smart pixel (2,1,1) node arrays.[2] In addition to the smart pixels, the system used computer generated holograms realized as binary and multi-level phase gratings (B/MPG) [3,4], 2-D fiber bundles [5], external cavity semiconductor lasers [6], high resolution bulk [7] and microoptics [8], and custom optical bench optomechanics. [9] Network control and call-load generation were done by a personal computer, custom high-speed electronics, and a programmable multichannel digital pattern generator. Results of the initial experiment included operation of the 5 stage network at 50 Mb/s per channel with 15 of the 32 inputs populated. In this paper, we describe enhancements to the initial system experiment that have allowed us to exercise the entire fabric (32 inputs), and increase the fabric's bit rate from 50 Mb/s to 155 Mb/s per channel.

2. Modifications to the original System5 experiment.

Several system modifications have been made for this experiment, as shown in Table 1.

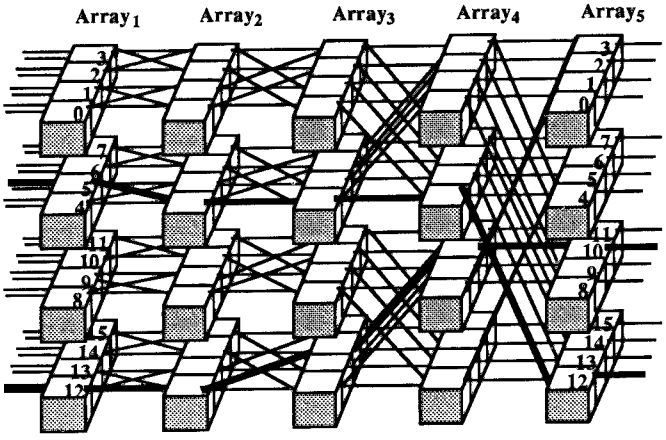


Figure 1. 5-stage Banyan interconnection network using embedded control and 4x4 switching node arrays.

The use of larger modulator windows (10x10 mm^2) reduced clock/power supply beam clipping. By removing the binary phase grating (BPG) that implemented the inter-node connection in our earlier system, [1] the signal power per receiver was increased by a factor of approximately 3.7. This enabled higher bit rate operation and improved the overall system stability. A new input fiber bundle, phase gratings, and diagnostic/analysis software have been added to the system

Other changes have been made to the smart pixel circuit design. [2] The node circuit of Fig. 2 consists of an optical receiver, a control memory to store a routing bit extracted from the data stream, and a multiplexer/modulator driver. The extracted control bit determines whether each node's multiplexer selects either its own optical receiver's output, or the output of another node's optical receiver (connected via inter-node metallizations). Design and processing modifications have increased FET currents, although the smart pixel array's performance was partially limited by lower than expected threshold voltages.

An important system issue is the lack of spatial uniformity of FET and pixel performance within each array, and between arrays. This non-uniformity can lead to a dramatic difference between the operational bit rate (or required optical energy for a given bit rate) of a single device and that of an array (or multiple arrays of devices). This non-uniformity is mainly due to variations in the receiver inverter threshold voltages across the array. In the initial experiment [1], these variations were mainly due to level-shifting and clamping diode voltage drop variations, bias voltage variations from voltage drops along power supply metallization, and FET threshold variations. Optimally matching a single pixel's characteristics requires the optimization of bias voltages, incident wavelength, optical power levels, etc.. In a practical system, these system parameters must be fixed to some "average" value, and if pixel uniformity is lacking, most of the pixels will operate sub-optimally. Individual smart pixels of the arrays used in the first experiment operated at 50 Mb/s with only 40 fJ of incident differential optical energy, but simultaneous operation of an entire array of pixels at 50 Mb/s required about 300 fJ of differential optical energy. In the second generation of these smart pixels, the MQW level shifting diodes were replaced by Schottky barrier diodes and thicker metallization lowered the voltage drop along the power supply traces. Individual node circuits on these new arrays operated at 155 Mb/s with 94 fJ of incident differential optical signal energy without system parameter (e.g.- bias voltage) optimization. In contrast to the first generation devices, only small (< +/-135 mV) variations of the receiver inverter threshold voltages across the arrays were noted.[1] This significantly eased the system parameter tolerances and eliminated the repetitive parameter adjustments required in the previous experiment.

Attribute	System$_5$	System$_{5\text{-prime}}$
Detector Area, Spacing	7x7 μm^2, 35 μm	7x7 μm^2, 40 μm
Modulator Area, Spacing	7x7 μm^2, 35 μm	10x10 μm^2, 40 μm
Pixel Area	210x210 μm^2	240x240 μm^2
Interconnect Technique	Binary Phase Grating	Interconnect Metalization
BPG Efficiency	60%	73%
Laser Modulation	50% duty cycle square	30% duty ≈ triangle
# of Equipped Inputs	15 fibers	32 fibers
Fiber Bundle Type	SM @ 810 with μoptics	SM @ 1300 w/o μoptics
Fiber / Input Rcvr Pitch	250 μm / 70 μm	500 μm / 120 μm
Pixel Switching Energy	40 fJ (≈300 fJ array oper.)	94 fJ (array operation)
Available Diff. Sig. Pwr.	15 mW (ave.)	55 mW (ave.)
System bit rate	50 Mb/s	155 Mb/s
Sytem Path Availability	Unmeasured	45%
System Bit Error Rate	Unmeasured	10^{-10}

Table 1. Component and system comparison for the two System5 experiments.

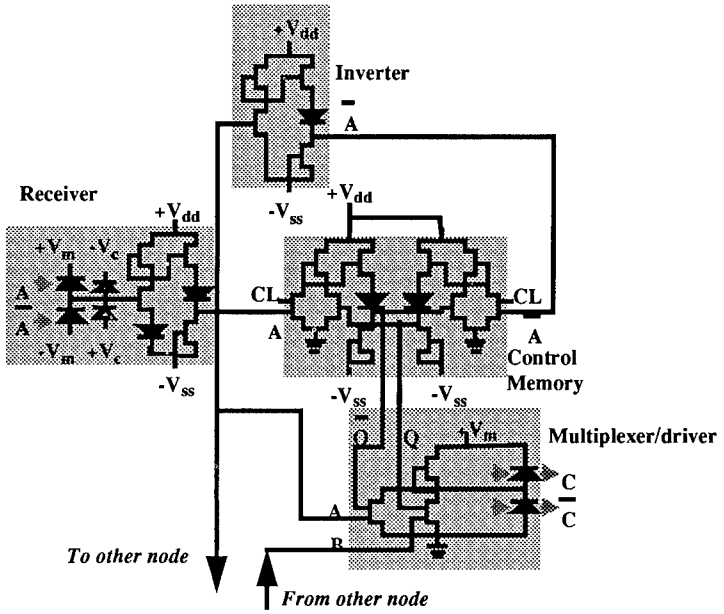


Figure 2. The (2,1,1) node schematic for the nodes used in the second experiment.

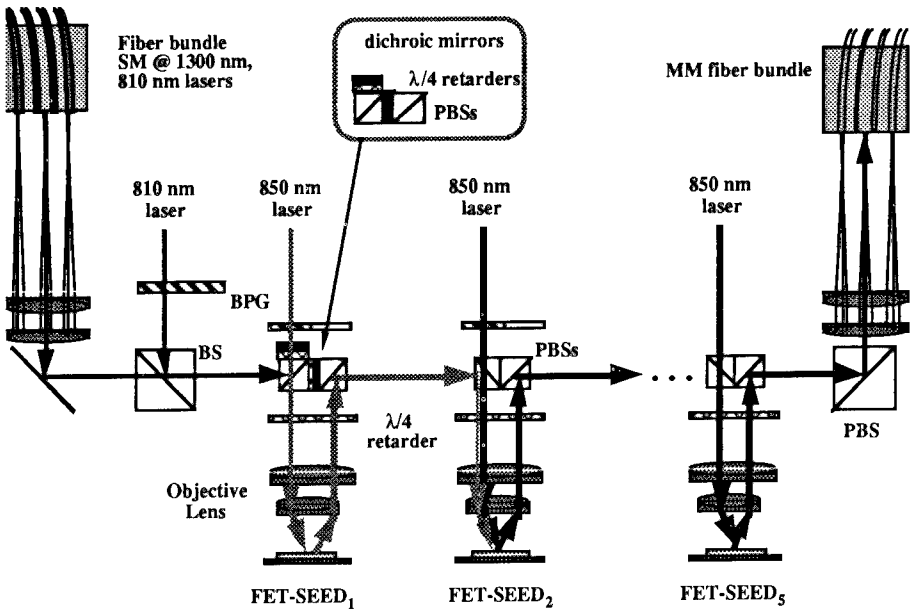


Figure 3. Optical schematic of the prototype free-space switching network, with arrows indicating the beam propagation direction.

3. System hardware design

The system's infinite conjugate optical design is shown schematically in Fig.3. The modulators were supplied with approximately 20 mW of average optical clock power at 850 +/- 0.2 nm from a collimated frequency stabilized external cavity laser. The laser's light output had a 30% duty cycle at 155 MHz. The modulation of the laser's output was somewhat triangular, due to high frequency roll-off in its drive circuitry. The use of non-separable design BPGs [10] [4] for spot array generation provided increased light efficiency. The BPGs had a power uniformity of 93%, and a measured efficiency (neglecting Fresnel losses) of 73%, compared to 60% in the initial experiment. The use of 1300 nm SM fibers in the input fiber bundle (versus the 850 nm SM fibers used in the first experiment [1]) eased the alignment of the data input lasers, and active alignment of the fibers during the bundle assembly ensured less than 2 μm fiber positioning errors. System inputs are generated by 810 nm lasers attached the input fibers, which are arranged in an 8x4 bundle on a 500 mm pitch. This fiber matrix is imaged onto the first node array through a dichroic beam combination system similar to the PBS assemblies of the other stages, except that PBS_1 has λ/4 a retarders and dichroic mirrors (transmit 850 nm, reflect 810 nm) attached so that the unpolarized light from the input fibers is directed onto the first node array regardless of its polarization. This obviated the need for polarization maintaining fiber for the input fiber bundle. To limit the total number of fibers, input lasers, and laser drivers required, the fiber bundle provides single-ended inputs to the 2 optical receivers of each node in the first array. The first laser (also at 810 nm) provides a reference level to convert the single-ended inputs to differential format. The beam from this laser is split into an array of beams by the BPG, and these beams are combined with the fiber input signals at a 50%:50% beam splitter. The combined arrays of beams propagate though the dichroic combiner and are focused onto the first FET-SEED array. Due to the 120 μm pitch of the first stage's optical receivers, a 0.24x demagnification is needed between the fiber and first node arrays is necessary to match pitches. This introduces clipping at the objective lens (since the beam numerical apertures at the fiber and FET-SEED planes are approximately equal) and results in a 10 dB power loss. The use of 1300 nm SM fibers improved the efficiency and

uniformity of the laser coupling to the fibers, but since they support 4 modes at 810 nm, a -2.2 dB modal noise optical power penalty was incurred. Both of these issues may be reduced by the use of a microlens array in a hybrid imaging system, as in Ref. 1. Fabrication errors in the alignment of the input beam splitter assembly's dichroic mirrors have been corrected by the addition of a birefringent wedge arrangement in the optical path before the assembly. The MM output fiber bundle used 100 μm core-diameter fibers on a 400 μm pitch.

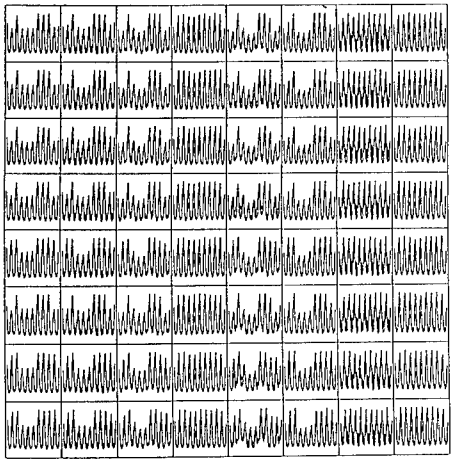


Figure 4. Ocillographs for the network connections between inputs 0 through 15 (rows) to outputs 0 through 7 (columns), for system operation at 155 Mb/s. Each plot is composed of 2 overlaid oscillographs, corresponding to the 2 inputs to the given first-stage node.

4. System testing and results

An automated call-load generation program was developed for this system to identify the effects of faulty nodes on the network. This system automated the sequential testing of all 512 paths through the 5-stage network at various channel rates, by configuring the network and storing oscillographs of each channel's output in computer memory. Figure 4 shows one quarter of the oscillographs recorded by this test system during a system exercise at 155 Mb/s. Subsequent data reduction analyzed the faulty paths to deduce the faulty nodes within the network. At 155 Mb/s, there were 3 non-functional pixels (2 in stage 1 and 1 in stage 3) and 1 low-powered input fiber. Due to the embedded control structure of the network, a faulty node (or a low powered input fiber) in one stage will prevent control information from reaching later stages, thus lessening the number of available paths. Thus, the 3 bad nodes and 1 bad fiber limited the network path availability to 45%. The system maintained this functionality within our lab (+/- 1 degree C, Newport table) for over 4 weeks. Two inputs were equipped with video CODECs operating at 105 Mb/s and these 2 paths have remained operating for four months without adjustment. The short term BER was 10^{-10} at 105 Mb/s for one of these video paths, and its eye-diagram is shown in Fig. 5. Although some channels have higher BERs, it is possible that further optimization of FET-SEED bias voltage, clock beam power, and operating wavelength can lower the BERs. The long term BERs on these channels were higher (10^{-7}), due to amplitude instability of the stage 5 clock laser.

The development of this prototype would not have been possible without the combined efforts of the authors listed in references 1, 2 and 8.

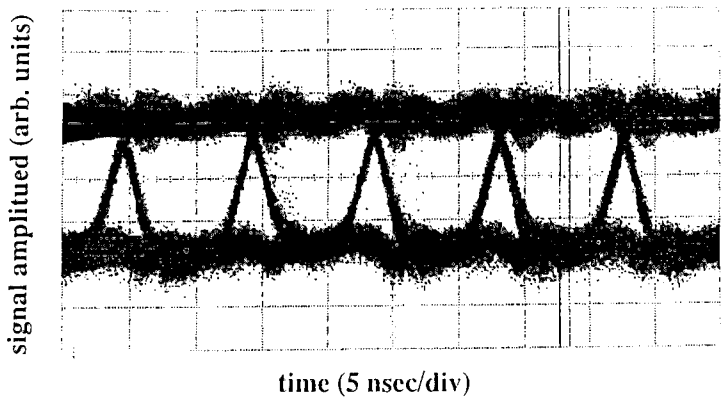


Figure 5. Eye diagram of the detected signal output (from Stage5) of one interconnection path operating at 105 Mb/s.

References

[1] F. B. McCormick, T. J. Cloonan, A. L. Lentine, J. M. Sasian, R. L. Morrison, M. G. Beckman, S. L. Walker, M. J. Wojcik, S. J. Hinterlong, R. J. Crisci, R. A. Novotny, and H. S. Hinton, "A 5-Stage Free-Space Optical Switching Network Using FET-SEED Smart Pixel Arrays," *Applied Optics, special issue on Optical Computing*, **33**(8), 1601-1618, (1994).

[2] A. L. Lentine, R. A. Novotny, T. J. Cloonan, L. M. F. Chirovsky, L. A. D'Asaro, G. Livescu, S. Hui, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, K. G. Glogovsky, T. K. Woodward, "4 x 4 arrays of FET-SEED embedded control 2 x 1 optoelectronic switching nodes with electrical fan-out," accepted for publication in *IEEE Photon. Technol. Lett.*.

[3] R. L. Morrison and S. L. Walker, "Design issues for beam array generation," in *Optical Computing Technical Digest, 1993*, (Optical Society of America, Washington, D.C., 1993), **7**, pp. 250-253.

[4] A. Vasara, M. R. Tagizadeh, J. Turenen, J. Westerholm, E. Nopponen, H. Ichikawa, J. M. Miller, T. Jaakkola, and S. Kuisma, "Binary surface-relief gratings for array illumination in digital optics," *Applied Optics*, **31** (17) 3320-3336 (1992).

[5] J. M. Sasian, R. A. Novotny, M. G. Beckman, S. L. Walker, M. J. Wojcik, and S. J. Hinterlong, "Fabrication of fiber arrays for optical computing and switching systems," accepted for publication in *Optical Engineering*, (1994).

[6] J. M. Sasian, R. L. Morrison, T. J. Cloonan, M. G. Beckman, M. J. Wojcik, and S. J. Hinterlong, "Frequency control, modulation, and packaging of an SDL (100 mW) laser diode," in Technical Digest of *OSA Topical Meeting on Optical Design for Photonics*, March 22, 1993, pp. 167-170.

[7] J. M. Sasian, F. B. McCormick, R. Webb, R. J. Crisci, and K. O. Mersereau, "Design, assembly, and testing of an objective lens for a free-space photonic switching fabric," *Optical Engineering*, **32** (8), 1871-1878 (1993).

[8] K. Mersereau, C. R. Nijander, A. Y. Feldblum, and W. P. Townsend, "Fabrication and measurement of fused silica microlens arrays, " *Microoptics*, Proceedings of SPIE, **1751**, (1992).

[9] F. B. McCormick, "Optomechanical Packaging Issues in Free-Space Optically Interconnected Systems," accepted for publication in the*International Journal of Optical Computing and Processing*, special issue on packaging aspects of optics for computing, (1993).

[10] R.L. Morrison and M.J. Wojcik, "Selective cell-based algorithm for designing high efficiency beam array generators," Proceedings of Diffractive Optics: Design, Fabrication, and Applications OSA Topical Meeting, June, (1994).

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

3-D free-space transmission polarization-based high-throughput photonic switching systems: architectural issues

V B Fyodorov

Scientific Computer Centre, Russian Academy of Sciences, Leninsky pr. 32 A, Moscow, 117334, Russia

Abstract. It is demonstrated how various types of optical polarization-based link blocks and polarization switch arrays make possible an effective architecture for $2N^2x2N^2$ three-dimensional compact high-performance optoelectronic nonblocking multistage interconnection networks for switching images. The main characteristics of such networks are evaluated.

I. Introduction

The multistage interconnection networks (MIN) containing $O(N log_2 2N)$ binary switching elements suit better for networks of a large size. However, the practical implementation of the polarization-based MINs, which transfer 2-D images through connected pair of optical channels, encounters several problems. One of them consists in creating such an optical system that ensures equal optical path lengths for both *p*- and *s*-polarized light components in any input- output interconnection patterns and has a sufficient numerical aperture to transfer images consisting of a large number of elements. The impact of the optical leakage of the signal into its orthogonal polarized component upon the quality of the picture transmitted is poorly known. Unlike the loss from photon absorption, such an optical leakage loss results in the fundamental limitation on parameters of the optoelectronic MINs.

In this paper several versions of compact optoelectronic multistage MINs for image switching between $2N^2$ input and $2N^2$ output channels are suggested. The switching stages of the MIN use arrays of light polarization plane modulators (LPM). The regular optical interconnections of link stages are implemented by using polarized-sensitive elements and free-space optics. The influence of crosstalk due to light leakage of signal into its orthogonal component upon the contrast ratio and the intensity of output image is analysed. The main parameters of such MINs are estimated.

2. Polarization-based optical systems for link stages

The three types of the polarization-based building blocks with M inputs and M outputs, that can be exploited for designing of link stages are depicted in Fig. 1. The polarising beam splitter (PBS) cubes and PBS prisms are the main components of these blocks. According to experimental and theoretical investigations at certain thicknesses of the dielectric layers of polarizing coating the angular aperture of the PBS cube can be as large as $\pm 10^0$ for *p*- and *s*-polarization components with the very good transmittance, beam splitting ratio, and depolarisation factor [1]. Theoretically [2], because of nonplanar wavefronts, depolarisation

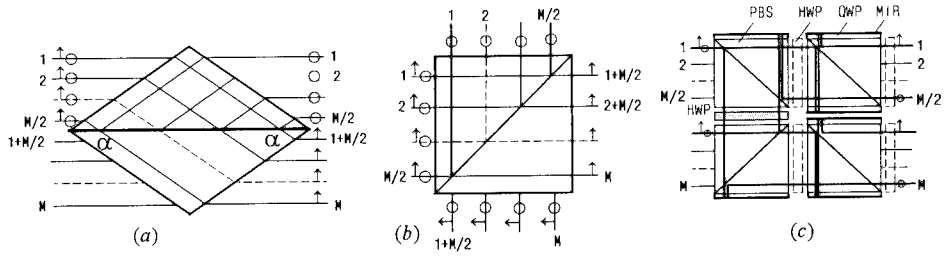


Fig. 1. Three types of optical link blocks, implemented on the base of PBS prism A, PBS cube B and four PBS cubes C. PBS - polarizing beam splitter, HWP and QWP - half- and quarter-wave plates, MIR - mirror.

factor for the angle of ray incidence $\theta < 10^{o}$ is less than $\gamma_d \approx 0.25 sin^4\theta \approx 2\times10^{-4}$. This value is comparable with the ultimate optical leakage in the orthogonal polarization for ideal LPM at the same angle of ray incidence q equals to $\gamma_m = sin^2(\pi sin^2\theta/2\ n_m^2)$, where $n_m = 1.5$ is some average of ordinary and extraordinary reflective index for LPM [2]. If angle an of the PBS prism with the refraction index n satisfies the relation $cos = 2^{-0.5}\ (1+1/8n^2)^{0.5} + 1/4n \approx 2^{-0.5}+1/4n$, inputs and outputs with the same name are placed on the horizontal optical axis that passes through them. For $n = 1.73$ the angle $\alpha = 30^{o}$. In this case angle $90^{o}-\alpha$ equals to Brewster angle. In the link setup C the half-wave plates (HWP) depicted by dashed lines can be absent if the output PBSs are turned by 90^{o} about the horizontal axis [3, 4]. The setup of each building block guarantees equal optical path lengths of communication links between input and output channels and allows one to apply optical means with the ultimate numerical aperture $u_A = nsin\alpha/(1+cos\alpha)M$, $u_B = n/M$ and $u_C = n/4M$ for blocks A, B, and C, respectively. The greatest numerical aperture of optical channels can be for building block C. The numerical aperture u_A for n=1.5–1.8 is about 1.3 time less than that of block u_C. The pattern connections for these link stages (see Fig. 2) are topologically equivalent to other more commonly used ones such as Banyan and perfect shuffle connections.

3. Architectural issues for multistage image interconnection networks

To design an nonblocking image MIN of size 2N×2N with pattern connections insured by building blocks of types A, B, and C it is necessary to use the blocks with M = N, N/2, N/4 ... 2. The numerical aperture of such a MIN will be limited by the value u of the block with the maximal value of M=N. As an example, the implementation of 2-D optoelectronic MIN of size 8x8, in which the blocks of type B are used, presents in Fig. 3. The optical path lengths of connections in such a MIN are equalised by the use of compensators shown as grey triangles. The size of compensator may be chosen out of the condition $l_c = l_p\ /(n_c - 1)$,

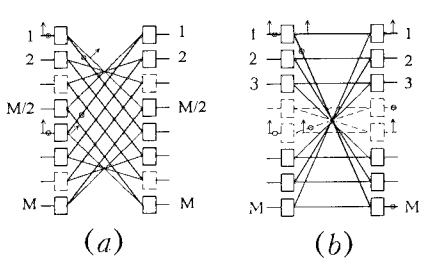


Fig. 2. Interconnection patterns for p- and s-polarized beams passing throw the link block A or B (a) and link block C (b).

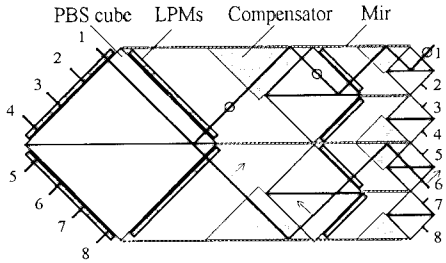


Fig. 3. A setup of 8x8 optoelectronic nonblocking multistage interconnection network with a link stage of the type B.

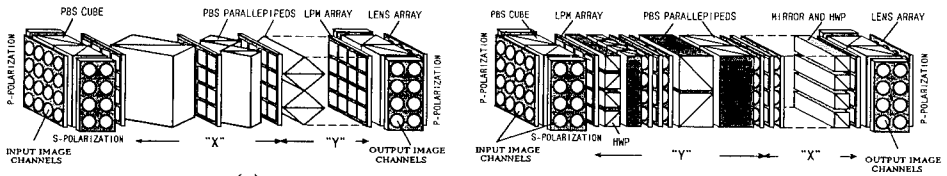


Fig. 4. Implementation of 32x32 polarization-based free-space image interconnection network with link blocks of the type A (*a*) and type C (*b*).

where l_c and l_p are the lengths of contacted sides of the compensator and the PBS cube, respectively, and n_c is refractive index of the compensator. If the compensator is made of glass with n_c=1.8 (heavy flint) then l_c=1.25l_p. The mirrors alter the interconnection pattern realized by block B. However, it remains topologically equivalent to the crossover network.

3-D compact optoelectronic image MIN of size $2N^2x2N^2$ is obtained with square LPM arrays, when $N=2^k$, k=1,2,3 ... . Examples of such nonblocking multistage MINs of size 32x32 (N=4, K=5) are exhibited in Figs. 4. The image MIN with blocks B provides for the greatest numerical aperture of optical channels, but the mutual composition of horizontal and vertical interconnection units in such a structure represents a difficult enough design task. Furthermore, the optoelectronic image MIN with blocks B cannot allow one to use square arrays of lenses and LPMs. The optoelectronic image MIN with blocks C is the most compact. If the size of the cross-section of all stages is the same, the length of each link stage in this MIN is approximately 2 or 1.25 time less than the necessary when using blocks A or B.

4. Performances of the optoelectronic image interconnection networks

The image contrast at the output of the channel MIN depends on both the connection pattern and the kind of picture transmitted through other channels. Some estimation of the image contrast ratio (signal to noise ratio) in the output channels of K stage Banyan MIN for the worst and the best cases can be made under the following assumptions. First, all light signals that form images are incoherent, and therefore, there is no interference between signals of the same polarization. Second, the decrease in the intensity of a data light signal is the result of the leakage of a part of the light signal into the orthogonal component provided that the total intensity is not changed. The total leakage loss γ in the LPM and polarization-based elements of the link stage is the same for both ON and OFF state of the LPM ($\gamma = \gamma_{on} = \gamma_{off}$) and does not depend on polarization. Other factors which lead to a change in the intensity are neglected. Third, all light signals that form the input image are of the same intensity. The input intensity A_O of *p*-polarized signal equals to the input intensity B_O of *s*-polarized signal.

If this conditions are fulfilled, it can be shown that for the worst case the intensity ($A_{K,S}$), corresponding to logical level of the signal '1' and the image contrast (Γ_{out}) at the output of the channel are equal to

$$A_{K,S} = A_O\,(1)(1-\Gamma)^K[1+(\Gamma_{min}\Gamma_{in})^{-1}] \quad (1)$$

$$\Gamma_{out} = \Gamma_{min}\,[1+(\Gamma_{min}\Gamma_{in})^{-1}]/[1+\Gamma_{min}\Gamma_{in}^{-1}] \quad (2)$$

$$\Gamma_{min} = A_{K,S}/A_{K,n} = (1\text{-}\gamma)^K \,/\, 1\text{-}(1\text{-}\gamma)^K \quad (3)$$

where Γ_{min} is the output image contrast when the contrast ratio input image $\Gamma_{in} \gg 1/\gamma$ and $A_{K,n}$ is the intensity crosstalk (noise) corresponding to the logical level of the signal '0'.

For the case of the best contrast ratio the change of levels of the dual-rail signal, associated with the partial signal leakage into the orthogonal polarized component, is

compensated by the leakage of signals transmitted by this orthogonal component. Thus, the input image goes through the MIN without any changes of the intensity $A_{K,S}$ and $A_{K,n}$, as well as contrast ratio , that is $\Gamma_{min} = \Gamma_{in}$

Therefore, we can make a conclusion that in K stage Banyna interconnection network the optical signal leakage leads to the spread in the output image intensity $A_{K,S}$ of the data signal from the input image value of A_O to the value of $A_{K,S} \leq 0.5\ A_O$, where $A_{K,n}$ is determined by (1), and to the spread in the output image contrast from $\Gamma_{out} = \Gamma_{in}$ to the value determined by (3). As it follows from (2) and (3), with relatively high requirements to the output image contrast ratio $\Gamma_{min} > 10$, the creation of Banyan MIN of a large size (for example 512 x 512) is possible only with the use of the polarization-based elements and LPMs, which have the small total leakage $\gamma < 0.01$: the overall light leakage γ through the LPM and polarization-based elements of link stage must be less than ~0.04-0.025 to make sure the image contrast ratio Γ_{min} = 3-5 in MIN of size 128x128 what is possible for the practical development.

The maximal pixel density ρ_{max} can be evaluated by the following expression: $\rho_{max} = (2u_C/3\lambda)^2 = (n/6\lambda)^2$ According to this expression and taking into account inevitable aberrations of the optical system and requirement of its simple alignment as well opportunities for realizing lasers and photodetector arrays, it can be considered that for a practical design when λ=0.9 μm and n=1.5, it is possible to form, to transfer and to register images with the maximal number of pixels $m \times m \approx 10^5$–10^4 on a centre-centre spacing of about 50-100 μm for array size of L×L ≈ 1 cm^2.

In the case of a laser array with the total light power P = Qη≈1 W (η = 0.1) and L^2 =1 cm^2, under heat release in the laser array Q_l =10 W and in photodetector array Q_r =1 W and the threshold sensitivity of photodetectors ~ 10 fJ [5] the rate of data transfer $V=QL^2/E$ can reach a value of V ≈ 0.1 Pbit s^{-1}. To maintain such a rate the image-repetition frequency must be F =1–10 GHz for the number of pixels $m \times m \approx 10^5$–10^4. As this takes place, the peak total data capacity W = NV ≈ 10 Pbit s^{-1}. The reconfiguration time t_r is the total time which need to change completely the state of all LPMs for establishing a new interconnection pattern. When a ferroelectric smectic liquid crystal LPM arrays are used the time t_r is equal, at best, to a few microseconds. The judgement about overall dimensions of the optoelectronic image MIN can be obtained from the construction sketched in Fig. 4 *b*. The length of 128 input and 128 output MIN is equal to about 20 cm under the cross-section of 2.5×2.5 cm.

5. Conclusion

Our study has showed that now it is possible to create compact bulk high-performance optoelectronic interconnection networks for data transmitting through any pair of connected optical channels by images or parallel binary codes. These interconnection networks use arrays of light polarization plane modulators and link stages implemented by classical optical techniques. Such interconnection networks may exceed electronic ones in the total data throughput and density of data flow (bit cm^{-2}c^{-1}) on a few orders.

References

[1] Mitsai V N and Fyodorov V B 1980 *Kvantovaya Electronica* 7 1028-1036 (in Russian)
[2] Shamir J Caulfield H G and Johnson R B 1992 *Appl. Optics* **28**, 311-324,
[3] Fyodorov V B 1992 *Optical Memory & Neural Networks*, **1** 303-313
[4] Fyodorov V B 1993 *Optical Computing & Processing* **3** 123-128
[5] Gusev V K, Roslova M R, Fyodorov V B and Shilov I A 1982 *Radiotechnika*, **30** no. 6 22-30 (in Russian)

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Architecture of a Terabit Free-Space Photonic Backplane

Ted Szymanski and H. Scott Hinton

Department of Electrical Engineering, McGill University
Montreal, PQ, Canada, H3A 2A7
email: teds@macs.ee.mcgill.ca, hinton@photonics.ee.mcgill.ca

Abstract. *The architecture of a "universal" photonic backplane is described. The architecture contains a large array of "programmable" smart pixels which can be configured into a few basic states. By setting pixel states appropriately any network including crossbars, hypercubes, meshes, and hypermeshes can be embedded into the backplane.*

1. Introduction.

A photonic backplane consists of a large number of parallel optical channels, typically 1,000 to 10,000, spaced a few hundred microns apart [1][2]. Each optical channel operates at between 100 Mbit/sec and perhaps 1 Gbit/sec, and the peak capacity of the photonic backplane is between 0.1 Terabit/sec and perhaps 10 Terabit/sec. Each printed circuit board contains one or more smart pixel arrays which could access the photonic backplane through an Optical Extender Card (OEC), as shown in Figure1 [1]. This paper proposes a high performance photonic backplane architecture called the "*HyperPlane*".

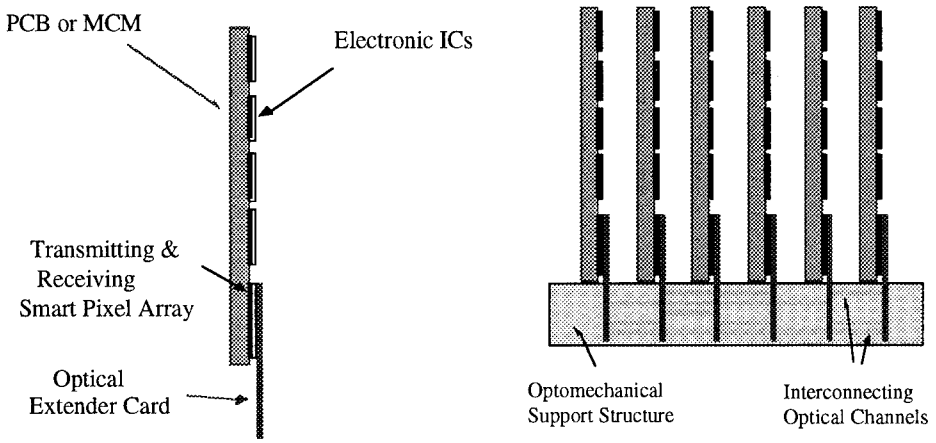


Figure I: (a) PCBs with Optical Extender Card. (b) Photonic Backplane based on OECs.

2. The Photonic *HyperPlane*.

The *HyperPlane* architecture can be modeled as a collection of N nodes partitioned among M printed circuit boards (PCB) or multi-chip modules (MCM) that are interconnected through a large number of optical communication channels provided by the photonic backplane, as shown in Figure II. Each node is composed of one or more processing elements (*PE*) and a message processor (*MP*). The processing elements within the node

could be either a general purpose computing processors or specialized processors such as an ATM switching nodes.

The message processor has the responsibility of controlling the communication between the nodes through access to a subset of the Z optical communication channels $\{C_1, C_2, \ldots, C_Z\}$. The *MP* has access to the communication channels through both X injector and Y extractor access channels labeled $\{I_1, I_2, \ldots, I_X\}$ and $\{E_1, E_2, \ldots, E_Y\}$ respectively, where $X \leq Z$ and $Y \leq Z$. The injectors provide the capability of injecting signals into a selected subset $\mathbb{I}$ of communication channels while the extractors are used to extract information from another subset $\mathbb{E}$ of communication channels. The node-to-node connectivity associated with the combination of both the communication and access channels creates a multi-dimensional interconnection design space which will be called the *"HyperPlane"*. Using this model, all N node interconnection networks of degree-K can be embedded onto the *HyperPlane*. By allowing the *MPs* to dynamically change the node access to the backplane communication channels the effect of a *dynamically programmable interconnection network* is achieved. Such a programmable interconnection network could implement crossbars, meshes, hypercubes and other networks as required for each desired software application. The backplane communication channels can be implemented through space, wavelength or time-division multiplexing or through combinations of these basic approaches.

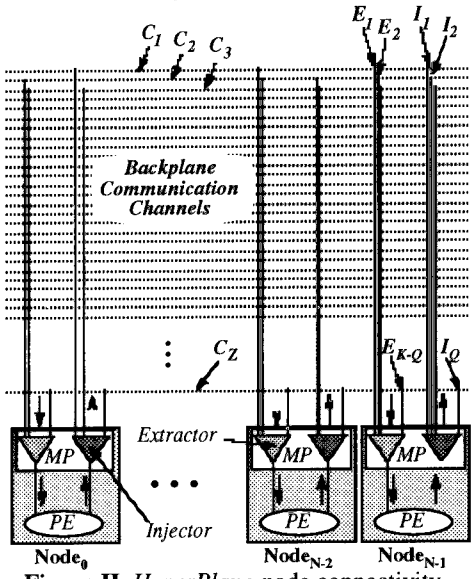
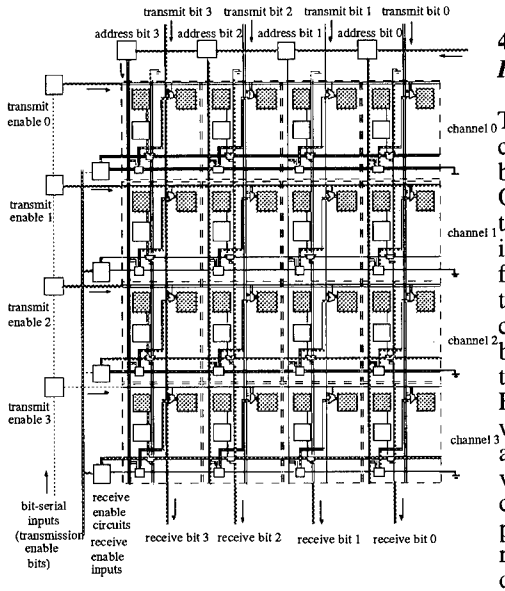


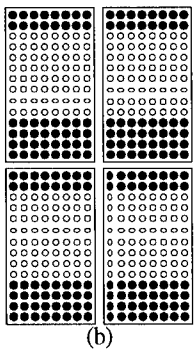
Figure II. *HyperPlane* node connectivity.

3. The Smart Pixel Arrays.

The smart pixel arrays for a *HyperPlane* based upon space-division multiplexing have been described in [4]. Pixels can be programmed to be in one of four states, the *"transparent"*, *"transmitting"*, *"receiving"* and *"transmitting-and-receiving"* states. The state of a pixel can be programmed by down-loading a bit-stream from the associated message-processor. The pixels can also be programmed to recognize and receive messages for any destination by down-loading the appropriate address bits. The pixels are logically arranged into an array called a *"communication slice"* as shown in Figure III, where each row represents a logical communication channel which is many bits wide. Each slice interconnects α electronic channels from the *MP* to β optical channels in the backplane where $\beta \geq \alpha$. The ratio of electrical to optical bandwidth of a slice can be adjusted by varying α and β. The data for configuring the slice is loaded in bit-serially; parallel data to be transmitted enters from the top and parallel data being received exits from the bottom. A smart pixel array containing a 32-by-32 array of pixels can be organized in various formats, i.e., with one 32-by-32 slice, with four 16-by-16 slices, or with sixty-four 4-by-4 slices. The ratio of electrical to optical bandwidth of a die may be adjusted by varying the number of slices on the die and by varying α and β within the slices. These organizations allow the architect to vary the architectural aspects of the *HyperPlane*.



(a)



(b)

Figure III. (a) A single ommunication slice and a (b) large smart pixel array with multiple slices.

4. Graph Embeddings in the *HyperPlane*.

The *HyperPlane* can embed conventional interconnection networks by programming the pixels accordingly. Optimal embeddings would minimize the length of the longest optical channel in the backplane. Optimal embeddings for linear arrays, 2D and 3D meshes, toroids, hypercubes, crossbars, dilated crossbars, hypermeshes and shuffle-based networks have been identified. A typical embedding is shown in figure V. Each box represents a backplane node with one or more smart pixel arrays and associated message processors. Each vertical line represents an electrical channel between an *MP* and its smart pixel array. Each bold horizontal line represents an optical channel between different nodes which is implemented by programming the pixels at each end as transmitters and receivers while those in between remain in the default transparent state. The dark squares represent transmitters and the clear circles represent dynamic receivers. Multiple circles aligned along a vertical line represent a concentration function. The embeddings can be changed in real-time by down-loading appropriate control bits to the smart pixel arrays.

A *Circular HyperPlane* is a variation of the basic *HyperPlane* where the optical communication channels are organized in rings. One physical implementation of a large scale *Circular HyperPlane* is shown in Figure IV. The *Circular HyperPlane* may offer improved performance over the conventional linear *HyperPlane* since the lengths of the longest optical channels of embedded networks may be reduced. This implementation is also scalable to very large capacities spanning large numbers of printed circuit boards.

The *HyperPlane* architecture can also be designed to exploit transparent long distance optical transmissions. By appropriately designing the optical extender card it is possible to allow long-distance optical transmissions to by-pass intermediate smart pixel arrays as they travel down the backplane [1]. This "*Transparent HyperPlane*" may improve performance by eliminating the delays associated with passing through intermediate smart pixel arrays. Since the intermediate smart pixels can no longer transmit or receive over the long-distance channels which bypass them, there may be a restriction on the types of embeddings possible.

5. Acknowledgments.

This research was supported by NSERC Canada Grant OPG 0121601 and the BNR-NT-NSERC Chair in Photonic Systems at McGill University.

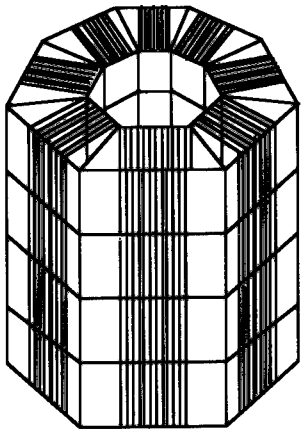


Figure IV: A physical implementation of a *Circular HyperPlane*.

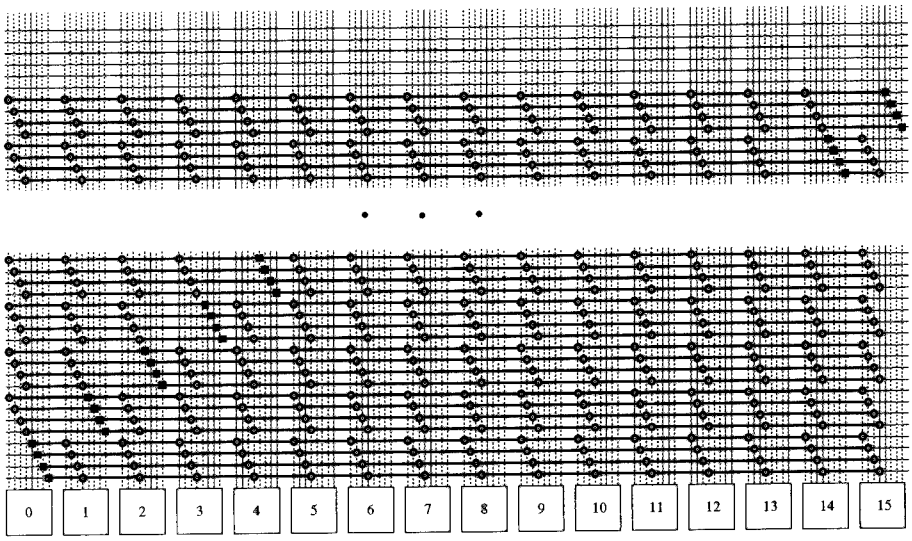


Figure V. Embedding of a crossbar network.

VI. References.

[1] H.S. Hinton, Canadian Institute for Telecommunications Research - Research Program 1993-94, pp. 143-156, 1993.

[2] K. Hamanaka, "Optical Bus Inter-connection using Selfoc Lenses", Optics Letters, Vol. 16, No. 16, pp. 1222-1224, 1991.

[3] L.A. D'Asaro et al, "Batch Fabrication and Operation of GaAS-Alx-Ga1-xAs Field-Effect Transistor Self-Electrooptic Effect Device (FET-SEED) Smart Pixel Arrays", IEEE Journal Quantum Electronics, Vol. 29, No. 2, Feb. 1993, pp. 670-678.

[4] T.H. Szymanski and H.S. Hinton, "Graph Embeddings in a Free Space Photonic Backplane", IEEE Int. Conf. Applications of Photonics in Technology, Toronto, June 21-23, 1994.

[5] T.H. Szymanski and H.S. Hinton, "Smart Pixel Arrays for a Dynamic Free-Space Optical BackPlane", IEEE Summer Topicals - 94, Lake Tahoe, Nevada, July 11-13, 1994.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A FET-SEED Smart Pixel Based Optical Backplane Demonstrator

by

D.V. Plant, B. Robertson, H.S. Hinton*, W. M. Robertson**, G.C. Boisset, N. H. Kim, Y. S. Liu, M.R. Otazo, D.R. Rolston, A. Z. Shang, and L. Sun
Department of Electrical Engineering
McGill University
Montreal, PQ, Canada
H3A 2A7

Abstract

We have demonstrated a representative portion of an optical backplane using free-space optical channels to interconnect printed circuit boards which employ FET-SEED based smart pixel arrays. Results of system demonstrator performance are discussed.

Introduction

Future digital systems such as ATM switching systems and massively parallel processing computer systems will have large board-to-board connectivity requirements that will not be met by conventional electronic backplanes. Free-space optical interconnects represent a solution to the needs of these future connection-intensive digital systems. Such an interconnect is capable of providing greater connectivity via an optical backplane. This optical backplane can be established using two-dimensional arrays of passive, free-space, Parallel Optical Channels (POCs) which optically interconnect electronic Printed Circuit Boards (PCBs). Such a backplane could be capable of supporting terabit/second aggregate capacities with connectivity levels on the order of 10,000 input/output channels per PCB.

Based on these needs, we are currently developing the optics, optomechanics, and PCB optoelectronic packaging to demonstrate these high bit-rate optical backplanes. As part of this program, we have constructed a representative portion of an optical backplane capable of interconnecting two printed circuit boards using diffractive optics to establish the POCs and FET-SEED smart pixel transceiver arrays for processing. Figure 1 schematically represents the two sided demonstrator system.

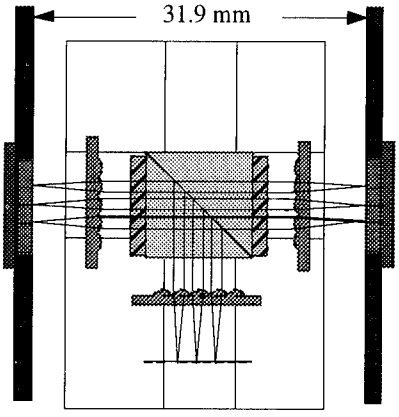


Figure 1: Schematic of the Optical Backplane

FET-SEED Smart Pixel Arrays

The FET-SEED transmitter and receiver smart pixel circuits were designed and fabricated using a batch fabrication process.(1) Figure 2a shows a schematic of the transmitter circuit. At the input, the drive FET modulates the voltage across a Multiple Quantum Well (MQW) modulator pair resulting in differentially modulated output light. The electrical input impedance was designed for 50 ohms to ensure efficient coupling of high frequency signals, which resulted in high speed operation of the optical modulators. Figure 2b shows a schematic of the receiver circuit. Here, the high speed optical modulation was detected using the MQW diode pair, fed to an inverting amplifier section, and then amplified using power FETs (375 μm gate width) designed to drive 100 ohm transmission lines on a PCB. Both the 4 x 4 transmitter and receiver array optical windows were 25 x 25 μm, separated by 50 μm, with the pixel to pixel pitch being set at 200μm.

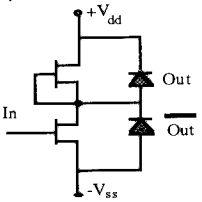


Figure 2a: Transmitter Array

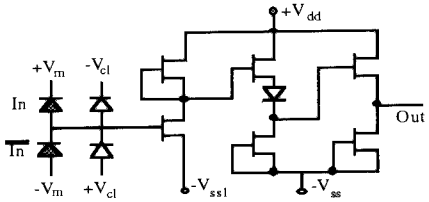


Figure 2b: Receiver Array

Figure 2: FET-SEED Smart Pixel Circuits

Each smart pixel array was bonded into a high speed quad flat pack that was then installed onto a printed circuit board via solderless connectors. These connectors permitted impedance matching of the smart pixel array input/output impedances to the 50 ohm printed circuit board transmission lines. Measurements of the rising edges of the packaged transmitter and receiver smart pixel array circuits yielded 0.811 ns and 2.57 ns respectively, in good agreement with device and circuit models.(2) These circuits were designed to run at 155 MBits/sec in parallel.

Optics and Optomechanics

The optics and optomechanics implemented in the demonstrator were designed to create optical microchannels using diffractive lenslet arrays and binary phase gratings. The board-to-board interconnection was achieved using a two-sided optical backplane approach(3), and the optomechanics were based on a modified baseplate approach in which magnesium was used as the baseplate material.(4)

Figure 3 shows a schematic of the optical system. The optical power supply was achieved using light from the output of an argon laser pumped Ti:Sapphire laser, which was delivered to the baseplate and collimated using collimating optics. A binary phase grating and a

40.34 mm Fourier lens were used to provide spot array generation. Risely beam steerers were used for fine positioning of the beam. The spot array was relayed using 2 lenslets to form a 4F relay at the focus of the Fourier lens to the modulator smart pixel array. The modulated light was then relayed from the transmitter board to the receiver board using a second 4F relay created by two lenslet arrays. Polarizing optics are used to direct the beams, with vertically polarized light being reflected off a polarization beam splitter, directed through a quarter waveplate, and focused onto the transmitter array. Modulated light was then reflected back towards the adjacent printed circuit board through the quarter waveplate which horizontally polarized the beams, causing them to propagate over to the receiver board. The modulated light was detected, and the signals were directed off the PCB for appropriate processing.

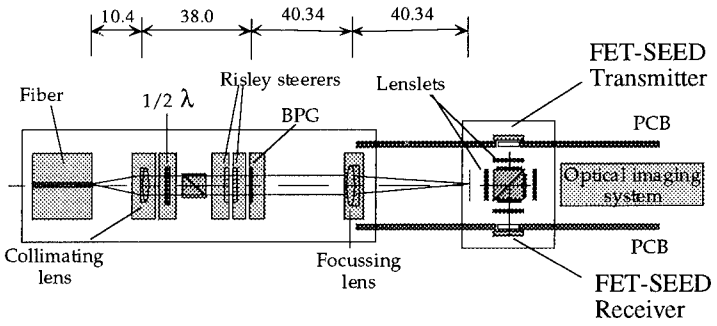


Figure 3: Optical System Layout

System Performance

The system was operated in two configurations. Based on the 600 μm center-to-center spacing of the lenslets arrays, in the first configuration the 4 corners of the modulator/receiver smart pixel arrays were interconnected optically. Figure 4 shows the results of transmitting data from board to board over one of the optical channels. In this configuration, each lenslet is supporting on dual rail optical channel.

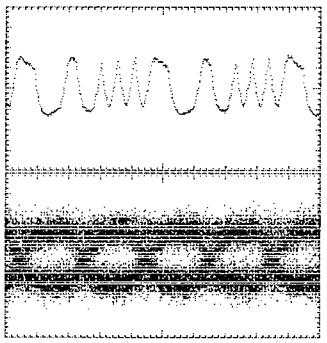


Figure 4: 16 Bit patterns and a PRBS at 50MBit/sec - individual channel performance

In the second configuration, a single lenslet was aligned to support four dual-rail differential optical channels in a Super Pixel configuration. Based on the efficiency of the optics, we could interconnect a 2 x 2 modulator with a 2 x 2 receiver pitched at 200 μm yielding an effective spatial channel density of 1111 channels/cm^2. Figure 5 shows a typical recording of the output of the system with all four channels being driven simultaneously. To our knowledge, this is the first time a single lenslet has been used to support the

interconnection of not only one dual-rail optical channel but also four dual rail optical channels.

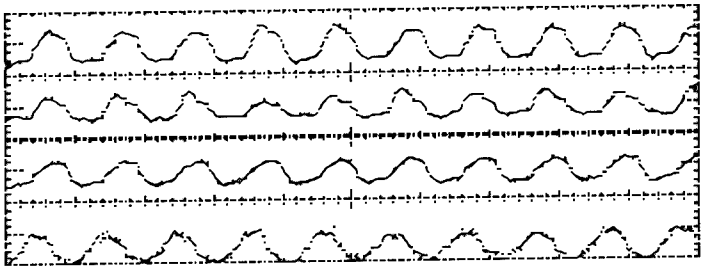


Figure 5: Four dual rail optical channels operating at 25 MBit/sec through a single lenslet relay.

Conclusions

In conclusion, we have constructed an optical backplane demonstrator capable of PCB-to-PCB optical interconnection. The free-space optical channels were defined using multilayer diffractive lenslet arrays and polarizing optics which interconnected FET-SEED smart pixel arrays. Future demonstrators will expand on these results by implementing sophisticated architectures such as the Hyperplane which utilize the increased connectivity provided by an optical backplane.(5)

Acknowledgments

This work was supported by the BNR-NT/NSERC Chair in Photonics Systems and the Canadian Institute for Telecommunications Research. In addition, DVP acknowledges support from NSERC (#OPG0155159), FCAR (#NC-1415), and the McGill University Graduate Faculty.

The authors gratefully acknowledge M.R. Taghizadeh of Heriot-Watt University for providing the diffractive lenslet arrays.

*Present address: Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309.

**Permanent address: National Research Council of Canada, Institute for Information Technology, Ottawa, Canada.

References:

1) L.A. D'Asaro et. al., IEEE Journal of Quantum Electronics, QE-29, no 2, pp. 670-677, 1993.
2) D.V. Plant et. al, IEEE/LEOS Topical Meeting on Smart Pixels, July, 1994, Paper M3.4.
3) H.S. Hinton, Canadian Institute for Telecommunications Research, Research Program 1993-94, pp. 143-156, 1993.
4) F.B. McCormick et. al., Applied Optics, vol. 32, no. 26, pp. 5153-5171, 1993.
5) T. Szymanski and H.S. Hinton, Optical Computing, August, 1994, paper WD2.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Experimental results of a 64 channel, free-space optical interconnection network for massively parallel processing

Ian Redmond and Eugen Schenfeld

NEC Research Institute, Inc., 4 Independence Way, Princeton, NJ 08540 USA

Abstract. We report on an experimental 64 channel, high data rate, free-space interconnection network for massively parallel processing. It uses VCSEL emitter arrays, fiber arrays for detection and a passive optical routing network. It forms the global network between clusters in the Interconnection Cache architecture [1].

1. Introduction

Our long term objective is to develop a system capable of achieving:

- Interconnection of 1000s processors in an interconnection cache architecture [1].
- Processor-to-processor data rates of 1-10Gbit/sec.

Many free-space optical interconnection schemes have been suggested, but often have difficulties of very demanding optics (SEED systems [2], or require diffractive elements (and hence wavelength control), or use large central switches (such as SLMs [3] and photorefractive crystals [4]) which do not scale well. We have attempted to make the best use of the optical and electronic components currently available. The optics is used only as an efficient method for communication, while signal generation, routing and detection is all done using electronics. Beam combination and splitting is done by simple mirror surfaces (no space-variant redirection), and relaying is done by simple optics, on-axis wherever possible, to reduce the build-up of aberrations.

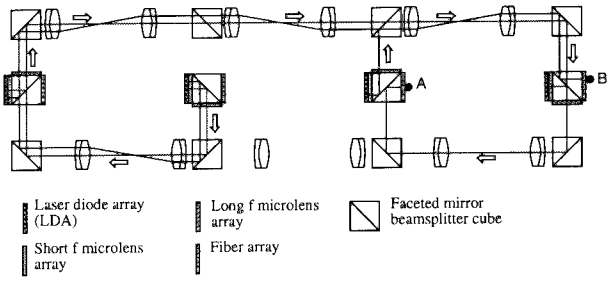


Figure 1. Schematic of the optical system.

2. System Overview

Our experimental system (figure 1) presents the one-way paths of the optical channels. All possible routings to 2 PEs, A and B, are shown. The experiment has 4 'columns' each with one simulated 'board' of 16 processors driving an 8x8 element VCSEL array. Data is received at the board by a 4x4 fiber array and led away to a separate receiver for each PE.

Data is carried on arrays of parallel slow diffracting gaussian beams from the VCSELs and relayed by microlens and micromirror arrays and simple bulk lenses. The experimental system is constructed on a slot-rail system which allows modular assembly, minimal alignment and good stability, similar to slot-plate systems already demonstrated [2].

Each optical channel is fixed and terminates at the receiving photodiode of one processor, so that there are N channels for N processors. Each processor selects, via electronics, the appropriate laser that transmits onto the channel of the target receiver fiber and PD.

Since a large number of processors must have access to the same channel, an optical fan-in problem exists. Fan-in occurs at each beamsplitter cube. To optimize the fan-in power efficiency, the reflectivity at each beamsplitter must be controlled, so that, beginning at the most distant processor from the target, the coupling efficiencies onto the channel are 1, $\frac{1}{2}$, $\frac{1}{3}$, .. $\frac{1}{M}$, for a fan-in of M. This gives a $\frac{1}{M}$ efficiency for every connection. However, because all the channels within a combining cube are generally at different distances from their targets, the coupling efficiencies for each channel within a single cube are also different. Hence, micromirror arrays are used at the center of each cube, and the channels are kept spatially separate. This is the main reason for the slow-gaussian microbeam approach. The approach also enables easy routing of individual beams onto and off columns.

The key advantages of this system are:

1)Distributed nature - no central switch or control needed.
2)Very few high speed i/o's to each VCSEL array, independent of network size.
3)The microbeam approach is scalable.
4)Only one laser per LDA is on, so low LDA power dissipation, reliable.
5)Speed limited by LD/PD technology, not by optics technology.

3. Optics

Diffraction limits the dimensions and hence the total number of distinct microbeams that can be supported within a faceted splitter/combiner cube. The relation giving this maximum number, N_{max}, is

$$N_{max} = \frac{\pi n D}{4\lambda m^2} \tag{1}$$

where n is the cube refractive index, D the cube dimension and $2mr_0$ the beam pitch (where r_0 is the $\frac{1}{e^2}$ radius) in the cube. Figure 2 shows the relation between N_{max} and λ for $m = 1.5$ at various cube sizes. It illustrates the high numbers of channels that can be used with compact ($\leq$30mm) optics.

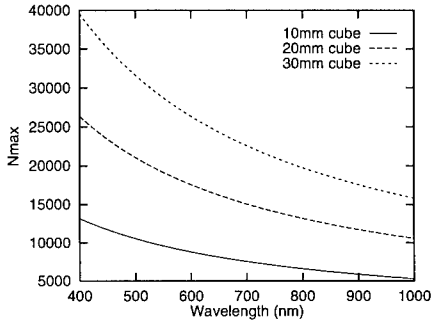


Figure 2. Max. number of channels for $m = 1.5$.

In our experiment, the 850nm beams emitted from VCSEL arrays have a 250μm pitch, and are collimated by ion-exchange planar microlenses. Some micromirror arrays are conformal HOEs made in dichromated gelatin, with controlled efficiencies and optimized for 850nm, and some are evaporated patterned metal.

Board-to-board relaying (20-50mm) within a column is done by $f = 14$mm, 16 level diffractive microlens arrays with a typical efficiency of 90%. The microlens arrays are bonded to the surface of each cube. The distance from column to column is larger ($\approx$ 300mm for a board width), and conventional bulk doublets can be used in a 4-f imaging arrangement to relay. Individual beams have extremely small numerical apertures (NAs), easing imaging requirements.

The theoretical aberrations of this system are very low, because NAs are very low (< 0.01), and much of the imaging is on-axis. Exact ray-tracing simulations on the worst case path of the experimental system shows negligible aberrations (strehl ratio ≈ 1). The telecentric 4-f bulk lens imaging system adds only significant spherical aberration, which, in our case, results in a beam pointing and position error. Simulations show that for our standard 40mm lenses at full 7mm field, the pointing error is only 0.025° and position 0.69μm. For a full size system using 80mm lenses and 20mm cubes, full 15mm field has only 0.028° and 2.1μm errors. After 8-f errors are 0.057° and 8.6μm. Optimized lenses would give improved performance, potentially allowing greater separation.

4. Experimental

Our system could operate a maximum of 48 'slow' channels (50Mb/s max.) driven from a logic analyzer, and 1 fast channel (3Gb/s, driven from a bit-error tester) simultaneously, with the particular channels used chosen by hardware connections.

Figure 3(a) shows a state analysis diagram of 24 channels operating without error at 20Mb/s. 50Mb/s operation was limited by the performance of the fiber receivers used and all but 4 possible channels (not lasing, or high threshold) operated correctly. Figure 3(b) shows a 1Gb/s eye diagram from the error tester. All tested fast channels operated with an error rate of $< 10^{-13}$, and most operated up to 1.6 Gb/s, limited by driver circuitry design and impedance mismatching.

Limitations were due to polarization instability, heating and variation in the VCSELs properties. Also, lossy metal micromirrors and use of diffractive microlens arrays limited signal powers. Future larger systems will require better VCSELs and more efficient components. The experimental system nevertheless has indicated the feasibility of the approach for large networks, particularly based on the interconnection cache architecture.

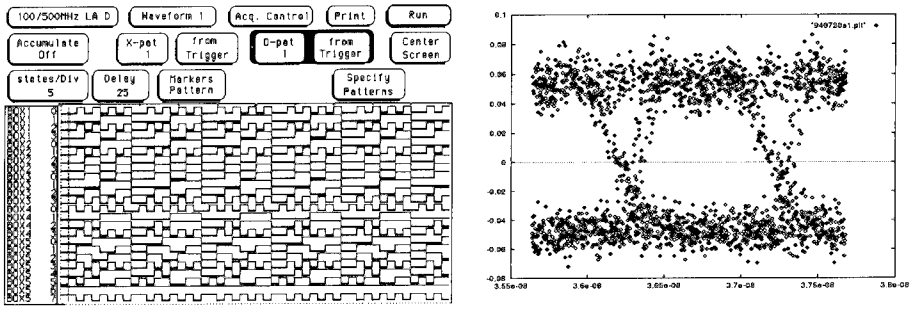


Figure 3 (a) 20Mb/s state diagram (24 channels) (b) 1.0 Gb/s eye diagram, typical channel.

References

[1] Redmond I and Schenfeld E 1994 *Proc. Optical Computing 94, Edinburgh* (Institute of Physics Publishing, London)

[2] Cloonan T J et al 1990 *IEEE Photon. Technol. Lett.* **2** p 438–440

[3] Walker A et al 1993 *Proc IEE Fourth International Conference on Holographic Systems, Components amd Applications (Conf. Publ. No. 379), Neuchatel, Switzerland* p 31

[4] Chiou A and Yeh P *Appl. Opt.* **29** 8 p 1111–1117

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Free-space WDMA Optical Interconnects Using Mesh-connected Bus Topology

Yao Li, Satish B. Rao, I. Redmond, Ting Wang
NEC Research Institute, Princeton, NJ 08540, U.S.A.

Adolf W. Lohmann
Angewandte Optik, University of Erlangen
Erlangen, Federal Republic of Germany.

A mesh-connected bus networking topology is proposed for implementing the three-stage Clos network and is experimentally demonstrated using a WDMA technology.

We consider networks for interconnecting parallel computers that are parameterized by their diameters (number of switching stages) and degree (fan-out) of switches which are used to build the network. Clos used switch fan-out to describe a family of networks that range from the complete network that has maximum fan-out and minimum switching diameter to the "Benes" network with nearly minimum fan-out and larger networking diameter. Electronic networks tend towards the small fan-out/large diameter end of the spectrum due to limitations on the fan-out of electronic switches. The fan-out of systems of optical switches is anticipated to be significantly higher than electronic switches. Thus, it has been suggested that optics can be used to build the complete network which is the most desirable of Clos's family of networks. Even optics, however, has practical limits on the fan-out of switches in any given configuration. In this paper, we describe an implementation of a network with significantly smaller degree and slightly larger network diameter than the complete network. The network that we implement is called the mesh-connected bus network which is the second most desirable network defined by Clos, which is widely referred to as the Clos network.

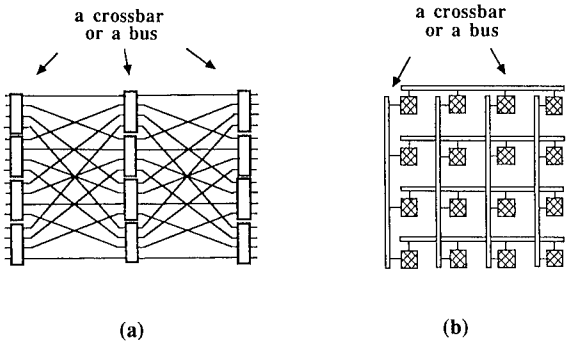


Fig.1. (a) a 16-node Clos network and (b) its MCB topology.

The Clos network is often depicted in its linear node distribution format, such as the one shown in Fig.1(a). We remap this 1D format to a 2D linking topology of Fig.1(b) which can be referred to as the mesh-connected bus (MCB) topology. The three switching stages are embedded into a permutation along a row, followed by along a column, and finally along a row in a grid structured 2D node array. It can be shown depending on the number of switches used in each row and column, the MCB offers at least a rearrangable non-blocking switching environment. For some interconnect topologies, efficient embeddings leading to a two-stage or even a single stage routing are possible. It has also been shown that the MCB can deliver comparable performance under randomized routing environments. Details of other important embedding, off-line as well as on-line communication features of the MCB network can be found in Ref.[1].

To optically implement the MCB interconnect, we propose to use the wavelength-division multiple access (WDMA) concept which allows to communicate among N users with only $\sqrt{N}$ wavelength channels. To support the WDMA communications, a critical passive component is a star-coupler which serves to broadcast the wavelength coded signals. A star-coupler is typically fabricated based on a butterfly or a back-to-back tree connections of individual 2x2 fiber-couplers. It can be seen that as many as 3 $\sqrt{N}$ such composite star-couplers are needed in order to establish the MCB networking using a WDMA scheme. One aspect of this research is to try to conceptualize optical interconnect schemes which can greatly reduce the overall hardware complexity of a network. To accomplish this goal for implementing a WDMA MCB interconnect, simple free-space optical components such as cylindrical optical lenses, mirrors, or 1D gratings, etc. for handling the star-coupling functions of the WDMA signals

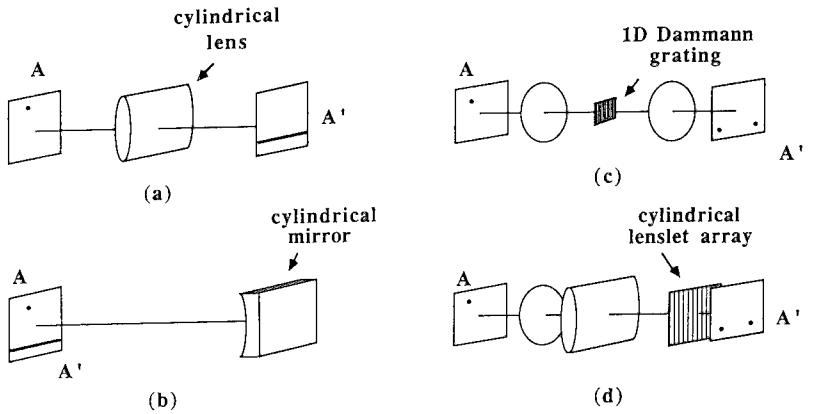


Fig.2. (a) - (d), four possible ways of making free-space star-coupler arrays.

were proposed (see Fig.2(a) - 2(d)). Each approach can supply a parallel array of $\sqrt{N}$ star-couplers necessary for one stage of the WDMA-based MCB routing. While the first two approaches allow only the so-called continuous broadcasting, the last two approaches can perform the discrete broadcasting thereby reducing optical power lost.

The proposed system concepts are experimentally verified through routing both low (10 MHz) and high (1.25 GHz) bandwidth optical signals (see Fig.3 for our system connection diagram) [2]. In the low bandwidth situation, some base-band video, audio and RS-232 data

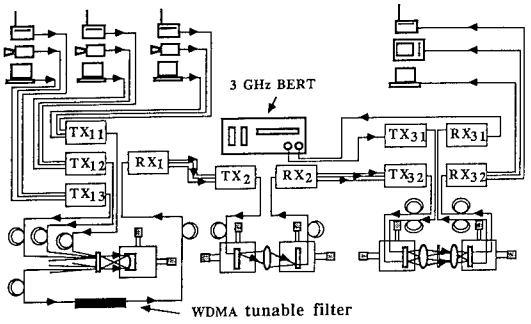


Fig.3. Flow chart of our demonstration setup.

are multiplexed in the square-wave frequency modulation format occupy a 20 MHz bandwidth. Conventional Fabry-Perot lasers operating in the range of 1290 - 1340 nm were used. Our power measurement confirmed that a fan-out to as many as 36 channels could be established along each bus making a network linking 1,296 nodes power-wise possible. Through 3-stages of free-space routing and associated electric-to-optical and optical-to-electric conversions, an overall 49 dB signal-to-noise ratio for video signal reception was still

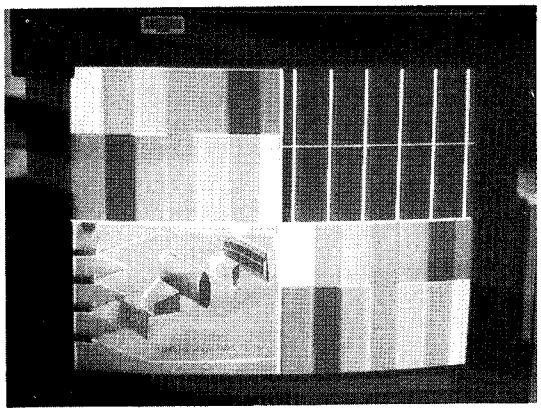


Fig.4. A typical WDMA video transmission and selection quality.

maintained. The images shown in the four quadrants of Fig.4 are the three wavelength coded video input images and one selected output image through WDMA filtering. The sound channels can also be added or dropped at each stage electronically. Our high bandwidth data transmission experiment involved using high-quality DFB lasers operating around 1310 nm. The corresponding optical system power measurement illustrates that at the transmission rate of 1.25 Gb/s, our WDM bus array can accommodate 16 channels per bus or equivalently 256 nodes in the network with a guaranteed receiving bit-error-rate of lower than 2 x 10^{-13} (see Fig.5 for the bit-error rate measurement results). The photograph of the overall system is shown in Fig.6 where the TV monitor shown at the right-hand side displays four identical images at its f our quadrants. They are the input image, the received images after being transmitted through the first, the first and second, and all three stages of our MCB network. The majority of space on the breadboard was occupied by electronics for handling signal multiplexing/demultiplexing, modulation/demodulation, electric-to-optical conversion, and optical-to-electric conversion.

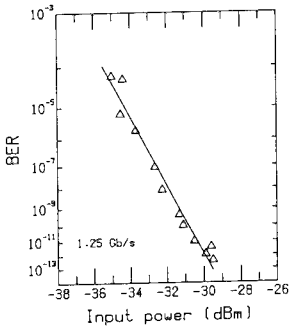


Fig.5. Bit-error-rate test result of 1.25 Gb/s data transmission in MCB.

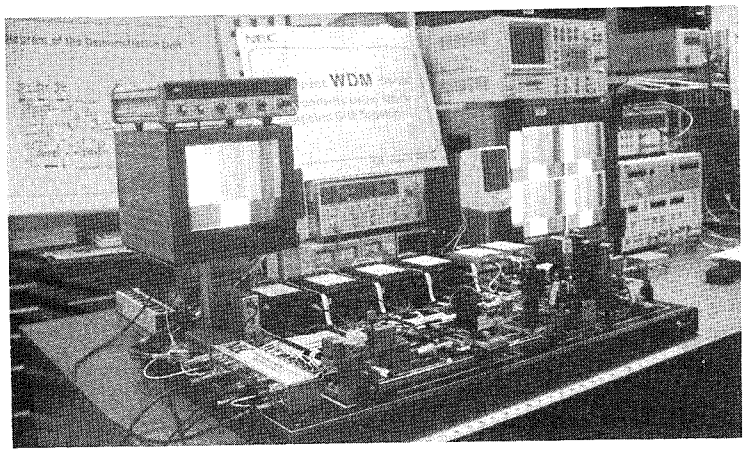


Fig.5. A picture of our demonstration system.

References

[1] Li Y, Lohmann A W, and Rao S B 1993 Appl. Opt. **32** 6425-6437

[2] Li Y, Lohmann A W, Pan Z G, Rao S B, Redmond I, and Wang T to appear in IEEE Proceeding (Nov. 1994).

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Free-space Multicast Non-blocking Interconnection Using Multidimensional Multiplexing Concepts

Yao Li, Ting Wang, Z. G. Pan
NEC Research Institute, Princeton, NJ 08540, U.S.A.

Jacob Sharony
Department of Electrical Engineering, Columbia University
New York, NY 10025, U.S.A.

Free-space optical schemes to minimize the number of switches for non-blocking multicast and broadcast interconnect applications are introduced and experimentally demonstrated based on the multidimensional multiplexing concept.

A non-blocking generalized switching network of N nodes is the one where each input can broadcast or multicast its message to any combination of output nodes without experiencing any internal blocking. In case that one input is to be multicast to j, $j < N$, outputs the remaining N-j output nodes can still be accessed by any of the remaining input nodes without experiencing internal blocking. Using this criteria, most popular point-to-point switching networks do not qualify for handling non-blocking communications for the generalized switching applications. Using the space switching concept a cross-bar network is about the only one which handles strictly non-blocking general purpose networking. A cross-bar has to use an order of N^2 basic, i.e. 2x2, switches making the hardware implementation of a large network a very difficult task. The reduction of switching complexity for a general-purpose strictly non-blocking network is not possible using a purely 1D switching technique since such a network has to make available, at each of its N output nodes, all its N input signals so that any output node can independently select any input signal.

In our approach, instead of a purely 1D switching the multiple multiplexing techniques are used [1]. Here, the multiple multiplexing implies that the explicitly multiplexed signals in the time, wavelength, or space domain need to be multiplexed one on top of the other again. Let us assume for a general k-tuple multiplexed system of N nodes with $N = \mu_1 \mu_2 \ldots \mu_k$; each has a unique identity composed of k indices. Since each of the N receiving nodes receives entire inform from all the input nodes, to select an input at an output node, the receiver has to access each of the k dimensions and to select (by demultiplexing) the corresponding channels. Selecting a specific sub-channel in dimension d is equivalent to a complexity of $O(\mu_d\text{-}1)$, i.e. the number of basic 2x2 switches is $O(\mu_d\text{-}1)$, and for N users the total switch count is $O[(\mu_d\text{-}1)N)$. Thus, the total complexity of the network is $O[(\mu_1 + \mu_2 + \ldots + \mu_k\text{-}k)N]$. Since all μ_i $(i = 1,2 \ldots k)$ are symmetrical, a minimum complexity will occur when all μ_i are identical. Thus, a minimum complexity of $kN(N^{1/k}\text{-}1)$ is obtained for $\mu_1 = \mu_2 = \ldots = \mu_k = N^{1/k}$. The overall complexity can be reduced by increasing k, however, since

the minimum number of sub-channels in any dimension is two, no more than log_2N dimensions can be used for a total of N signals. Thus, for $k = log_2N$, the complexity approaches the absolute limit of $Nlog_2N$ which ties the information lower bound pointed out by Shannon [2]. Using electronics, a $k=2$ system (time, and space) can be implemented. The use of optical fiber components allows $k=3$ with an additional wavelength dimension. When the free-space optics is used, angular dimensions in two perpendicular planes as extra degrees of freedom can be included making a $k=5$ system feasible. One additional advantage of the multiple multiplexing is that instead of squeezing all complexity into a single dimension making the system working at its extreme limit, each dimension handles its fair share of switching complexity at its comfortable range. In Fig.1, switching complexity comparisons are shown for various multicast non-blocking circuit switches. The standard cross-bar and the tree-based cross-bar switches, each of which consumes about N^2 basic switches, are represented by the top curves, while the theoretical minimum complexity is denoted by the bottom curve which labels $Nlog_2N$. Between them are the two curves labelled $k=2$ and $k=5$, which correspond to the two mentioned multidimensional multiplexing cases. It can be seen that for large scale network, the proposed concept offers a significant hardware reductions.

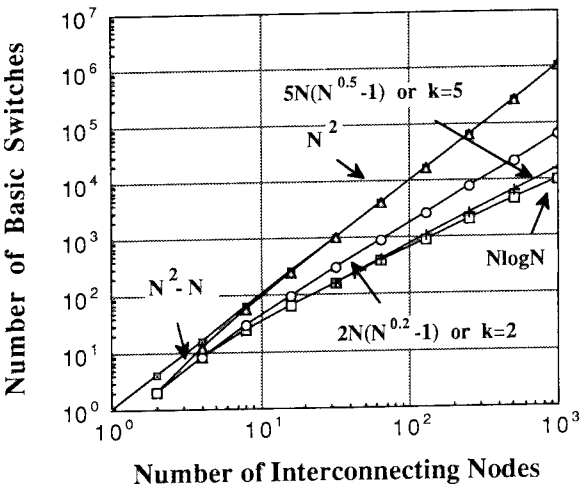


Fig.1. plots of network switching complexity

To confirm the proposed multi-dimensional switching concept, an N=9, k=2 system utilizing two mutually orthogonal angular dimensions is sketched in Fig.2 for its top and side views. For the broadcast and select operations, respectively, a 2D Dammann grating and two 1D optical deflector arrays are employed. The Dammann grating splits an input beam into 9 angularly encoded beams and multiplexes them with the help of a spherical lens into 9 spatial spots to be selectively switched by the two beam deflector arrays. The first and second deflector arrays select one out three angles along the XZ and YZ planes, respectively. In this way, instead of using 81 switching states, only 2x9x3 = 54 switching states are used. The larger the N, the more savings of the switching states.

The angular selective switching can be implemented using a multichannel A-O deflector array. Since the number of angular channels is fundamentally related to the spatial and temporal frequency resolutions, various technological and fundamental limitations restrict the network interconnect capacity. We have performed some first-order calculations of limitation on such a capacity based on parameters of the multichannel A-O deflector array. Fig.3 plots network capacity limits due to the A-O deflector's time bandwidth product, the grating's space-bandwidth product, and the grating to A-O cell interfacing numerical aperture, where the parameters are f_1, the focal length of the lens; f_c, the central frequency of the A-O deflector; λ, the optical wavelength, z and h, the thickness and width of the transducer of each deflector, and s, the spacing between two consecutive deflectors. More than 32 angular channels can be used in each dimension making a system interconnecting more 1,024 nodes possible.

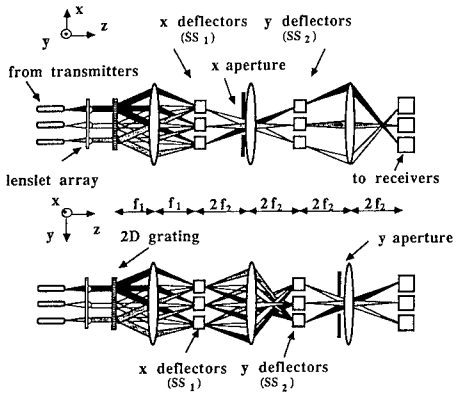


Fig.2. Top and side views of the proposed switching setup.

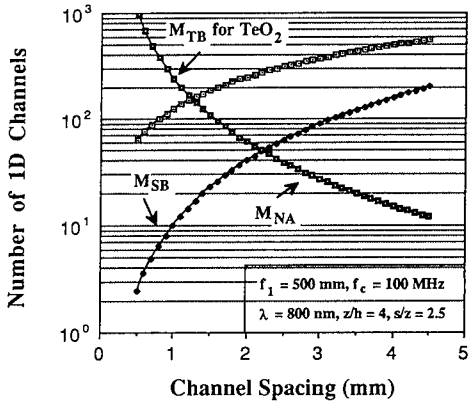


Fig.3. Limit on the number of multiplexible angular channels.

Experimental verification of the proposed k=2 system was performed using two 8-channel TeO_2 A-O deflector arrays fabricated by Brimrose [3]. Because of using a 8x1 and a 1x8 deflector arrays in cascade instead of two ideal 8x8 deflector arrays, our proof-of-principle experiment was designed to demonstrate one row of the non-blocking selection of $N = 8x8 = 64$ node switching system. The experimental setup is photographed in Fig.4 where a folded system (by two mirrors to reduce the longitudinal dimension) of geometry of Fig.2 is used. A 2D mask containing 8x8 round holes (see Fig.5(a)) was used. To better track this input pattern in the system, eight holes in the diagonal direction was blocked. The duplication of the single

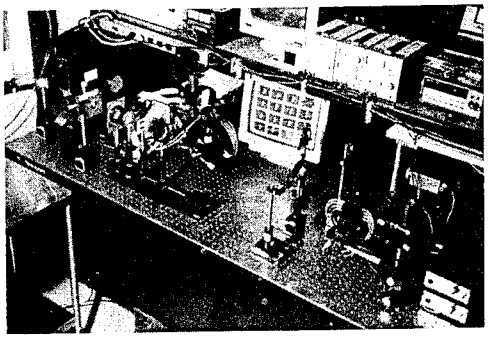


Fig.4. A photo of the demonstration system.

focused pattern into 8 patterns needed for the deflector array was performed by the Dammann grating. In Fig.5(b), a photograph containing three simultaneously deflected patterns from the 1st, 2nd, and 8th Bragg deflectors is shown. Each of the 3 deflected 8x8 patterns was then individually tuned to allow only one row of the 8 Bragg degenerated spots to pass a narrow slit attached to the second lens (see Fig.5(c)). This slit and lens combination was used to perform an unit magnification imaging of the light distribution at the first Bragg deflector array to the second Bragg deflector array. Fig.5(d) shows the photograph taken at the second Bragg deflector array plane where only 3 spots appear due to our slit selection. Since each of the 3 spots still contains 8 angularly multiplexed signals, the tuning of the second stage Bragg deflector could further select a single spot from each of the three 8 spot arrays using still another slit-lens combination with the slit perpendicular to the first one. In our case, since all the 3 spots focused onto a single Bragg deflector, they were deflected by the same angular quantity. By arranging the deflecting angle so that one out of 8 spots in each of the three first stage selected beam arrays passed through the slit-lens combination, we obtain the final selection result at the output plane (see Fig.5(e)). Thus, using a cascade of two Bragg deflector stages, the operational principle of the proposed novel complexity reduction scheme useful for the non-blocking multicast photonic network was confirmed.

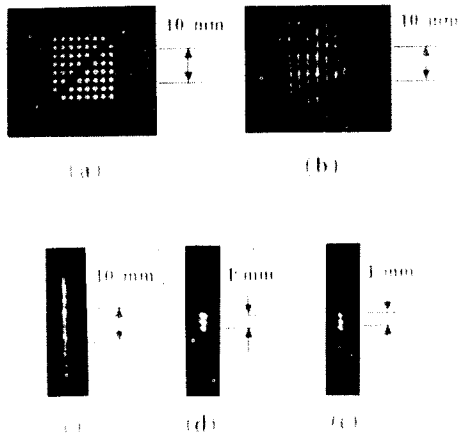


Fig.5. Experimental results of switched signals.

References

[1] Sharony J, Stern T, and Li Y to appear in *IEEE/ACM Tran. Network*

[2] Shannon C E 1950 *Bell System Tech. J.* **29** 343-49

[3] Li Y, Wang T, Pan Z G, and Sharony J 1994 *Opt. Lett.* **19** 515-17

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical network for a general purpose massively parallel optoelectronic architecture : interconnection pattern, physical layout and experiment.

S. Kocon, P. Churoux, M. Fracès, J.P. Bouzinac, D. Comte, N. Hifdi, J.Y. Rousselot.

ONERA -CERT
2, avenue Edouard Belin - B.P. 4025, 31055 TOULOUSE CEDEX
Téléphone : (33) 62.25.25.86 - Télécopie : (33) 62.25.25.88

An interconnection network applied to a general purpose massively parallel computing architecture is discussed. Several criteria enable us to choose a specific network. An optoelectronical system is proposed. Experimental results and simulations are presented. The network is based on a 2D perfect-shuffle optical interconnection pattern coupled to an electric mesh network. Light is propagated in free space and deflected by a microprim array.

1. Presentation

We have studied in our laboratory a processing architecture called OEDIPE (OptoElectronic DIgital PEs). It is a general purpose massively parallel optoelectronic SIMD/SPMD (Single Instruction/Process Multiple Data) architecture. OEDIPE is composed of a N^2 electronic processing element (PE) array, an electronic planar memory, a global control unit and optical interconnections. Data are organized in a N^2 matrix. Different kinds of applications can work on this architecture, such as data base, image processing and matrix operations.

2. Network structure

Because of the difficulties to implement a crossbar network [1] with a level of parallelism we want as high as possible, it is prefered here to select a communication kernel extracted from the needs of data motions required by the aimed applications. This kernel is composed of neighbourhood data motions, row and column shifts, row and column broadcasts and matrix transposition.

Bidimensional optical multistage networks are attractive. According to the optical SBWP (Space Bandwidth Product), these networks can be implemented with a high level of parallelism. A multistage network is composed of several stages of fixed long link patterns combined with dynamic short link patterns (the switch arrays). In our approach, the fixed long link patterns are optical while the switch arrays are implemented with electronic technology which is efficient for neighbouring connections.

Some of the multistage networks can not produce a given data motion. For the others, the way of running depends on the kind of network considered. We have chosen a multistage network which can respond to the following criteria :

- minimizing the number of feed-back loops through the network for the desired communications ;
- achieving all the communication kernel with a single configuration order in each stage. This order is broadcasted over all the switches of a stage. Then, all the switches of a same stage are set in an identical state. This property is very important to reconfigure as quick as possible the connection pattern.

Classical multistage networks such as crossover [2] or shuffle-exchange [3, 4] are not well fitted with the previous constraints. So, we have developped a specific one, the 2D Enhanced Shuffle-Exchange (ES-E) multistage network (figure 1). It is a bidimensional shuffle-exchange network with additional neighbouring connections between nods in a same stage. These new connections are represented in dotted lines in figure 1. For N^2 inputs, ES-E network contains $\log_2(N)$ stages. The fixed long link patterns between two consecutive stages are identical and correspond to 2D perfect-shuffle schemes. This network can be reduced to the last stage over which are performed $\log_2(N)$ loops.

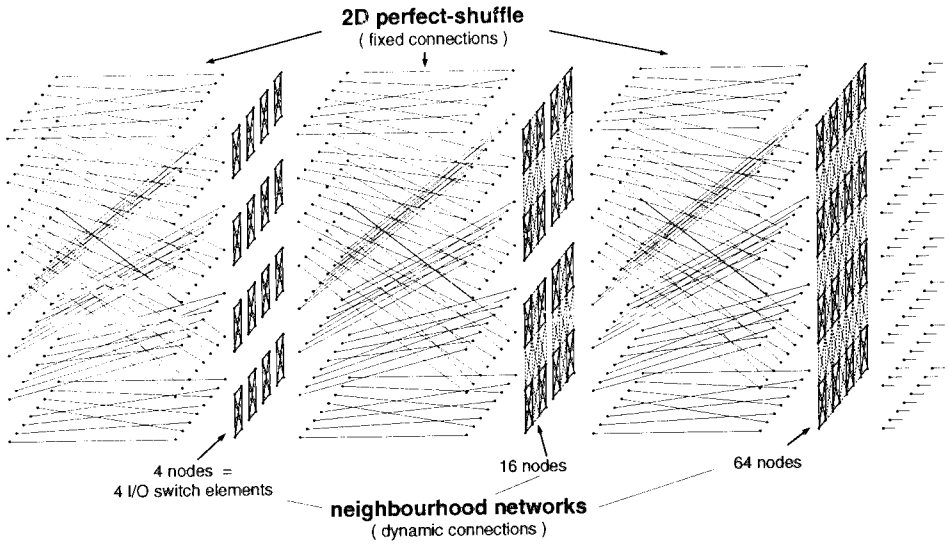


Figure 1 : a 2D Enhanced Shuffle-Exchange multistage network with $N^2 = 8 \times 8$ inputs.

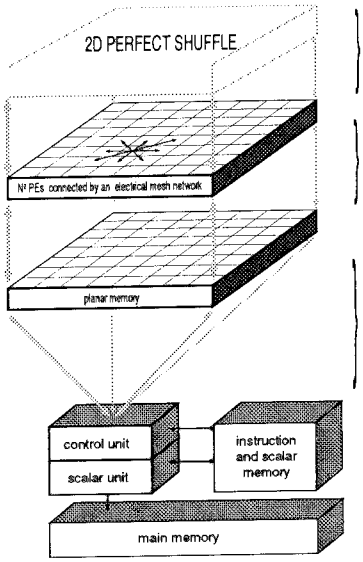


Figure 2 : OEDIPE architecture design.

An implementation of OEDIPE architecture is proposed on figure 2. This architecture is composed of three optical fixed networks :

- a 2D perfect-shuffle network connected to the PE array ;
- a duplex bus between the PE array and the planar memory ;
- a duplex broadcasting network between the global control unit and the PE array through the planar memory.

The ES-E network implementation is a single stage one. PEs are interconnected with an electrical mesh network so that they simulate switch arrays.

3. Experimental approach

An optical solution to implement the ES-S network is presented now.

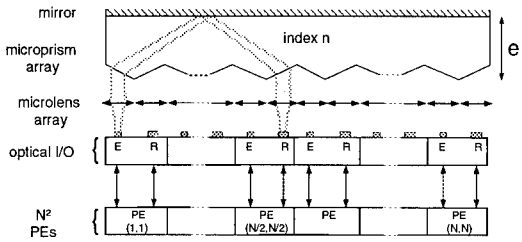


Figure 3 : an optical 2D perfect-shuffle implementation using a microprism array.

The I/O stage of the PE array is composed of a 2D array of pairs of emitters and receivers. All beams are collimated by a microlens array. After crossing optical components which perform the desired deflection pattern, all the beams are focused by the same microlens array on the corresponding receivers

Among the technological solutions which can be imagined to achieve the deflection function, we can mention : a microprism array ; a micrograting array ; a 2D CGH which can realize the collimation and the deflection functions by the use of off axis Fresnel microlens.

We have first studied the microprism solution (figure 3). Our goal is to estimate the deflection limitations of this technic and to verify if it is able to produce a highly parallel 2D perfect-shuffle pattern.

The experimental setup is described on figure 4. It is composed of the following elements.

- An optical source which is a pigtailed laser diode. The fiber has a 5 µm core diameter which permits to simulate a VCSEL (Vertical Cavity Surface Emitting Laser diode).
- Two NPL (National Physical Laboratory) 103x103 250 µm diameter microlens arrays.
- A BK7 glass flat which can be rotated to simulate different prism angles.
- An acquisition system for detection.

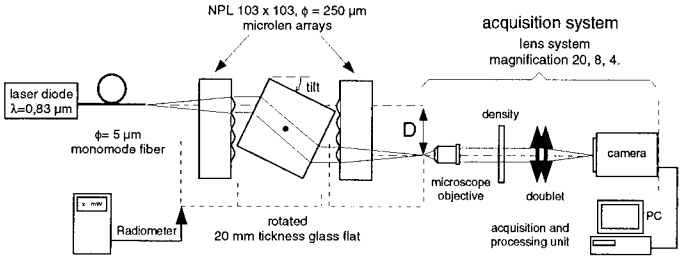


Figure 4 : experimental setup to study the deflection system.

To define the characteristics of the network, we have imposed to the Fresnel reflection loss on the microprism face to be less than 10 %, and to the collecting microlenses to troncate at $1/e^2$ the beams which are supposed gaussian.

So, the maximum tolerable value for the deflection angle is 63°and the glass flat thickness is 20 mm.

To evaluate the performance of the system we have studied the two extreme values of deflection angle : 0° and 63°.

4. Results

Theoretical simulations and experimental measurements were done to evaluate the following parameters : - the shift D (see figure 4) which limits the level of parallelism, N^2 ;

- the optical power budget between the emitter and the receiver ;
- the output image quality which will define the receiver size.

Experimental results are compared with simulations in table 5 and figure 6.

Table 5 : comparison between experimental results and simulations				
	SIMULATION		EXPERIMENT	
	0° tilt	63° tilt	0° tilt	63° tilt
Shift D	0 mm	11,184 mm	0 mm	11,18 ± 0,06 mm
Connection transmittance	70,6 %	61,1 %	64 % ± 6	54 % ± 5

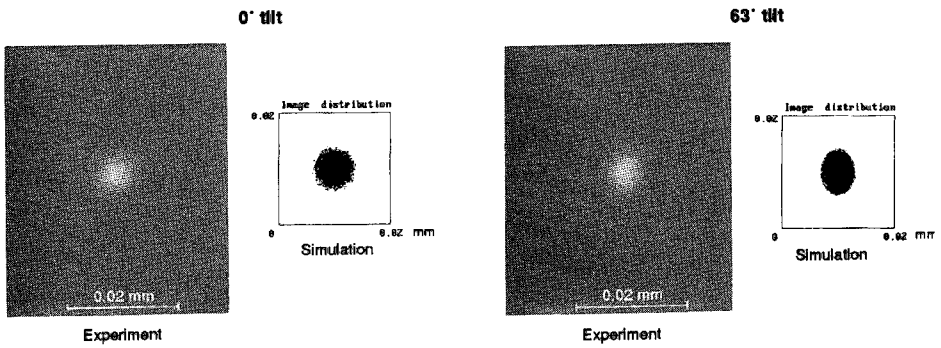


Figure 5 : comparison between experimental output images and simulations.

As we can observe it, the shift D, the transmittance and the output image are in good agreement with simulations. Moreover, we can see the image quality and the tranmittance are quite convenient.

Characteristics projected by simulation for a folded 2D perfect-shuffle network (see figure 3) in OE-DIPE architecture using SF11 microprisms are presented in table 7.

Table 7 : characteristics projected by simulations for a folded 2D perfect-shuffle pattern		
	Case 1 200 µm microlens diameter	Case 2 400 µm microlens diameter
Number of inputs, N²	128 x 128	256 x 256
PE array area	~ 2 x 2 inch²	~ 8 x 8 inch²
Microprism array tickness, e (see figure 3)	3 cm	12 cm
Maximal single link attenuation loss	# 3 dB	# 3 dB

5. Conclusion

We have presented a general purpose optoelectronic massively parallel architecture. Processing is left electronic and communications are optical. A specific optical network was developped for an over 100 x 100 parallel architecture including a very compact single folded stage network.

This network uses microlens and microprism arrays. Further improvements will use off axis Fresnel microlens arrays. These components are well fitted with planar electronic integrated technology.

References

[1] Fracès M. and all 1991 *Proceeding of 7ᵉ Journ. d'étude des fonctions opt. ds l'ordinat.* (Brest september 91)

[2] Murdocca M 1990 Int. J. of Optoelectronics **5** (2) 191-203

[3] Cheng L and Sawchuck A A 1992 *Appl. Opt.* **31** (26) 5468-5479

[4] Stone H S 1971 *IEEE Trans. on computers* **20** (2) 153-161

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Performance analysis of optical multistage networks and integrated holographic interconnection

Igor G.Voitenko and Toyohiko Yatagai*

Institute of Applied Optics, Byelorussian Academy of Sciences, 212793 Mogilev, Belarus
*Institute of Applied Physics, University of Tsukuba, Tsukuba, Ibaraki 305, Japan

Abstract- A bit sequence can be routed properly by a banyan network using integrated-optic directional couplers and the possible optical interpretations of this network is investigated. In this paper we consider holographic interconnections between nodes and stages for some types of the optical interconnection networks and represent experimental results for holographic interconnections in photorefractive waveguides formed by titanium and iron indiffusion in $LiNbO_3$.

1. Introduction

One of the most promising applications of optical computing techniques is interconnections for VLSI systems. The communications crisis in the area of VLSI circuits and systems is very serious. Problems of clock distribution and data communication inside VLSI chips are solved by using optical interconnections with optical fibers and holographic techniques [1]. Volume holograms in waveguides offer a straightforward means of interfusing dynamically reconfigurable interconnections with integrated optoelectronic devices. Previous uses of thick holograms in waveguides have included grating couplers and distributed feedback lasers [2]. Holograms for dynamic applications have also been considered, especially in photorefractive crystals. In this paper we consider holographic interconnections between nodes and stages for some type of the optical interconnection networks and represent experimental results for holographic interconnections in photorefractive waveguides formed by double indiffusion in $LiNbO_3$.

It is shown that there exists a topologically equivalent class of multistage interconnection networks which include the indirect binary n-cube network, and that the multistage interconnection networks in the class can be developed based on the n-cube network[3].

In this paper we investigate the possible optical interconnections and propose non-blocking optical interconnections and the use of integrated holographic devices in the proposed networks.

2. Motivation of holographic interconnections

In many interconnection networks in use, regularity is observed. A hologram, with the aid of several conventional optical components, can be useful to realize complex and massive interconnections among processors. Additionally, thick volume holograms provide very high diffraction efficiency and low crosstalk for interconnections. Elementary holographic gratings can be used to implement interconnection links between individual processing elements of two distinct planes in a multi-layered optical network. For example, a 5-grating hologram can perform a 2-D

mesh interconnection network of any size, and a 2^n-node binary hypercube interconnection network can be easily realized by an n-grating volume hologram.

In addition, planar multiplexed holograms have been suggested for coupling between totally internally reflected beams propagating in a dielectric substrate. Holograms have also been demonstrated which couple light from the evanescent field of a guided mode in a waveguide.

The principal advantage of using holographic elements in photorefractive materials for interconnect applications is the ability to write the hologram with a single exposure and to erasure by the double-exposure technique with a variable phase shift between the two recording gratings.

3. Waveguide photorefractive hologram

Consider a planar waveguide fabricated by diffusion of Fe in x-cut $LiNbO_3$. The y-axis is defined to be the principal axis of propagation in the waveguide. The x-axis is normal to the plane of the waveguide and the y-axis is transverse to the z-axis in the plane of the waveguide.

After the first stage of Fe in-diffusion (T= 960° C, diffusion time=4 hours, and thickness Fe-film=60nm), we repeated the procedure of diffusion with other parameters (T=1020° C, diffusion time = 6 hours, and thickness Fe-film = 60 nm). The mechanism of the photorefractive effect in the Fe:$LiNbO_3$ structure is based on $Fe^{2+} \leftrightarrow Fe^{3+}$ interconversion and the space charge set up by the action of light is attributed to a charge redistribution between divalent and trivalent impurities.

After ion doping under reducing conditions, new characteristic absorption bands are seen. As result, we have to restrict the maximum of the current in the photoconductive spectra. For the preparation of photoconductive cell, photolithographic technique is utilized for making Al interdigital electrodes pattern on the clean surface of $LiNbO_3$, and on the Fe:$LiNbO_3$ in-diffusion waveguide. The sizes of the electrode finger and gap are 8 and 4 μm, respectively. These devices are used to examine the relation between the photoconductive current and the light wavelength. The photorefractive effect involves the photoexcitation of mobile carriers from impurity centers (Fe^{2+} impurities) to the conduction band, the sensitivity is expected to be strongly correlated with the absorption spectra of the substrate crystals[4].

Two methods for holographic gratings were studied. In the first method, a classical setup is used. The substrate with the waveguide is placed at the intersection of two coherent light beams. One of the beams may be considered the reference beam, and the other the signal beam. We assume that the bisector of the two beams is normal to the crystal surface, and we choose the z-axis in the plane of incidence normal to the bisector. Reffering to Fig. 1 we consider a scheme of two-wave mixing for the study of photorefractive dynamic holograms. An input optical beam is split by a beam splitter into a pump beam and a signal beam which interact inside a photorefractive crystal. In this case the energy efficiency is defined as the ratio of the optical power of the amplified signal beam to that of the input beam in the Bragg condition. In addition, we can measure the diffraction efficiency using an output prism of the waveguide.

Typically the incident power in each beam was 200mW, and the gratings were written in 0.5 sec exposure. These parameters were adjusted to produce approximately a 50:50 splitting ratio, when a He-Ne guided wave light was incident at the appropriate Bragg angle.

The Bragg angle for the grating was given by

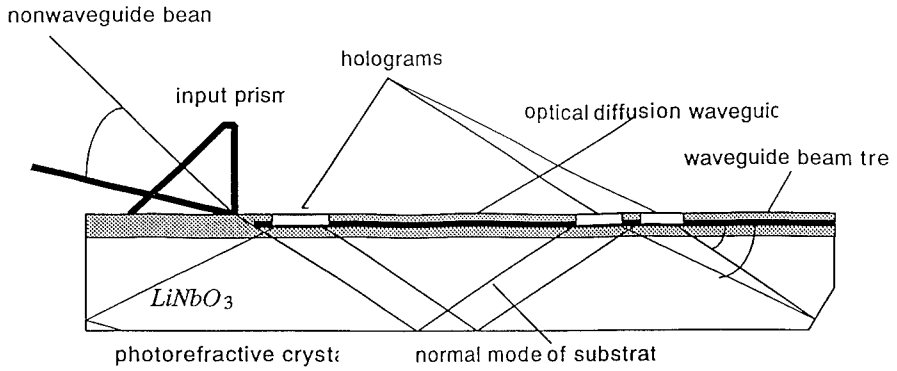


Fig.1 Photorefractive waveguide for optical interconnects.

$$\theta_B = \sin^{-1}\frac{\lambda_0}{2\Lambda N} \tag{1}$$

where θ_B denotes the Bragg angle, λ_0 the wavelength, Λ the grating pitch, N the refractive index of the guided mode. For the given values of λ_0 = 633nm, Λ = 1.4μm and N =2.2241, we have the Bragg angle $\theta_B = 6.2^\circ$. The optical beam in the waveguide was reduced to approximately 150μm width by focusing externally, and the interaction length L varied from 1 to 2 mm. The deflection efficiency is defined in this case as the power coupled out in the deflected beam divided by the sum of the powers coupled out in the deflected and the transmitted beams.

Diffraction efficiencies for the guided beams were as high as 8 %. By rotating the grating in the z-y plane, we were able to find the optimum Bragg condition. In each case of measuring the diffraction efficiencies, we found a difference between the calculated and measured ones because of the effect of spatial dephasing between the absorption and the phase gratings. This is indeed true for most photosensitive materials since the same physical or chemical processes are usually responsible for both properties. The variation of the diffraction efficiency was approximately 12%.

In the second method, we have used an acousto-optic control. The grating was written into waveguide. In our experiment, an acousto-optic Bragg cell is utilized at 910 MHz, so that a wide band range of angles from 2.2° to 4.6° . This allows us to use holographic gratings made with small phase shifts or different periods. This system requires synchronization in order to stop the recording of the second grating when the efficiency reduces to its minimum. We extended the double-exposure technique to the case of time-averaging exposure. Extension to time-averaging holographic exposure simplifies the proposed interconnection system.

4. Optical interconnection applications

The implementation of a large space switch requires the interconnection of many smaller switches used as building blocks. For point-to-point networks, the interconnection of these building blocks to construct a large switching system can be done with a Berns, Cross, banyan or

mesh interconnection network or a hypercube network.

For optical interconnection applications, the signal beam (in the first method) is expanded through a binary matrix to carry the interconnection pattern. Depending on the experimental configuration, such a mask can be used to realize a 1-to-NxN interconnection or NxN cross-bar switch. To achieve maximum energy efficiency, we need to match the beam profile spatially at the photorefractive substrate. The decompositions of a 2-D mesh interconnection network and the decompositions of an 8-node binary hypercube interconnection network with holographic implementation are considered.

A mesh interconnection network is used to communicate to four nearest neighbors for any node in the array. An 8 node hypercube network can be decomposed as 6 regular interconnect patterns.

It is clear that a hologram with five stored gratings can easily realized a mesh interconnection network of any size, and a hologram array with three stored gratings in each holographic element can implement the 8-node hypercube interconnection network. We have generated gratings by using a Fourier transform configuration of interconnection patterns.

5. Conclusion

We have described holographic interconnections and demonstrated the dynamic formation of holographic gratings in photorefractive double Fe-in-diffusion LiNbO3 waveguides. The results we have obtained shows the validity of the holographic techniquefor fabricating grating in photorefractive waveguides.We have proposed the double-exposure technique with combining optical and acousto-optical controls. Combined with the large storage capacity available with holographic recording, this double-exposure technique could be suitable for the optical implementation of learning networks and some types of interconnection networks. The advantage of holographic interconnections is that different interconnection networks are addressable for the same multi-computer systems.

Acknowledgements

This work was partially supported by the Fund of Fundamental Researches of Belarus under Project F56-411.

References

1. D. J. Brady and D. Psaltis, "Holographic interconnections in photorefractive waveguides," Appl. Opt., **30**, 2324-2333 (1991).
2. R. G. Hunsperger, *Integrated Optical: Theory and Technology*, (Springer-Verlag, Berlin, 1985).
3. M. J. Narasimha, "The Batcher-Banyan self-routing network: universality and simplification," IEEE Trans. on Comm., **36**, 1175-1178 (1988).
4. D. von der Linde and A. M. Glass, "Photorefractive effects for reversible holographic storage of information," Appl. Opt., **8**, 85-100 (1975).
5. E. Y. B. Pun, K. K. Wong, I. Andonovic, P. J. R. Laybourn, R. M. De Rue, "Efficient waveguide Bragg-deflection grating on LiNbO3," Elect. Lett., **18**, 470-472 (1982).

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Systems Partitioning and Placement in Optoelectronic MCM Design

J. Fan, S. H. Lee and C. K. Cheng
Dept. of ECE, University of California at San Diego, La Jolla, CA 92093-0407.

Abstract. *This paper introduces the Computer Aided Design (CAD) approach for partitioning and placement in designing and packaging opto-electronic systems into Opto-Electronic Multichip Modules (OE MCM). We will first discuss the design issues of speed, power dissipation, area and fabrication limits in optoelectronic system design with free space optical interconnects. We will then define the formulations for OE MCM partitioning and chip placement. New algorithms are described for optimizing this partitioning based on the minimization of the power dissipation and placement based on fabrication limit. The result of an example of a multistage interconnect network is given and more than 50% reduction in power consumption and the maximum interconnect distance can be achieved.*

1. Introduction

Free-space optical interconnections (FSOI) have been shown to have speed or power advantage over electronic interconnections for long distance interconnections[1]. By introducing FSOI into new MCM designs, it is possible to create an OptoElectronic MCM (OE MCM). However, incorporating optics into electronics presents new challenges in computer-aided design, fabrication, and packaging. To integrate electronics and optics at the system level, we need new partitioning and placement algorithms. In this paper, we will try to address these important CAD issues.

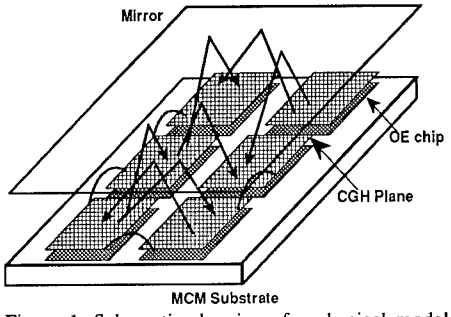


Figure 1. Schematic showing of a physical model of OE MCM package in reflective configuration. Only a few of the many actual interconnects are shown.

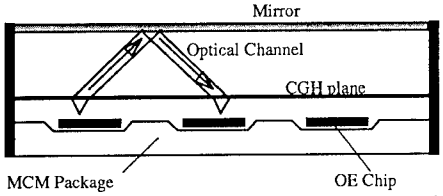


Figure 2. Cross-Section of an Opto-Electronic MCM package.

Figure 1 shows a physical model of an OE MCM. There are three layers in the packaged system. The lower layer consists of multiple chips of electronics and optoelectronic devices on MCM substrate. Each chip contains a number of switching elements (SEs) or processing elements (PEs). The middle layer consists of diffractive optics (DOEs) or computer generated holograms (CGHs). The upper layer is a mirror. The cross section of this system is shown in Figure 2. The optical transmitters, e.g. surface emitting lasers, illuminate an off-axis lens where the beam is collimated and deflected into the desired direction. The beam then reflects off a mirror and is focused by another off-axis lens onto a receiver (detector) on another chip.

There are four aspects of optoelectronic MCM design that are key to the system performance: interconnection speed, heat

dissipation, chip size and MCM size, and maximum optical interconnect distance. Currently, optical sources (such as modulators or laser diodes and their drivers) generate heat in densities much higher than that of VLSI. Based on speed, switching energies, and power budgets, electrical and optical interconnect technologies can be compared and a break-even interconnection length, say l_{be}, can be defined[1]. When l_{be} is approximately the lateral size of the chips, optical interconnection technology is preferred for chip-to-chip interconnects where the signals must propagate distances longer than l_{be}. For interconnects among SEs within chips where signal propagating distances is less than l_{be}, electronic interconnection technology is preferred. It is also known that the minimum feature size of the fabrication technology limits the maximum interconnect distance [2]. Under these technological constraints, partitioning and chip placement can be optimizd by CAD methods.

2. System Partitioning of OE MCM

Optoelectronic system partitioning is performed: (a) to divide the total number of SEs in the system into fewer number of chips; and (b) to choose the best technologies for each interconnect required by the netlist of a certain computing architecture, so that one of the system performance parameter will be minimized while the remaining three performance parameters mentioned above be constrained. For example, given a total of N SEs, the task is to divide the N SEs into the M partitions so that the constraints are satisfied and the goal is achieved. The size constraint will be defined as NC SEs per chip. The connection is defined by the netlist of the system. Within the chips, connections will be electrical, while among chips the interconnection will be optical. The goal is to minimize the power consumed by the whole system, including both optoelectronic chips and

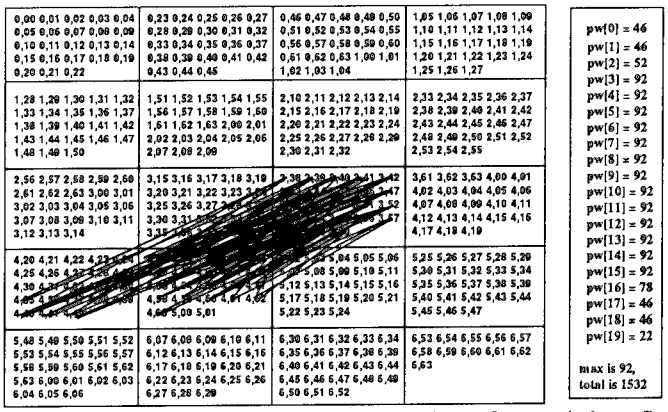


Figure 3. The straight forward partitioning of a twin-butterfly interconnection network into a 4×5 OE MCM (left). The numbers of optical interconnects on each chip are listed (right). The pw[i] is the number of optical interconnects for the ith chip. The 1st index of each item is the stage index, and the 2nd index is the SE index in the stage.

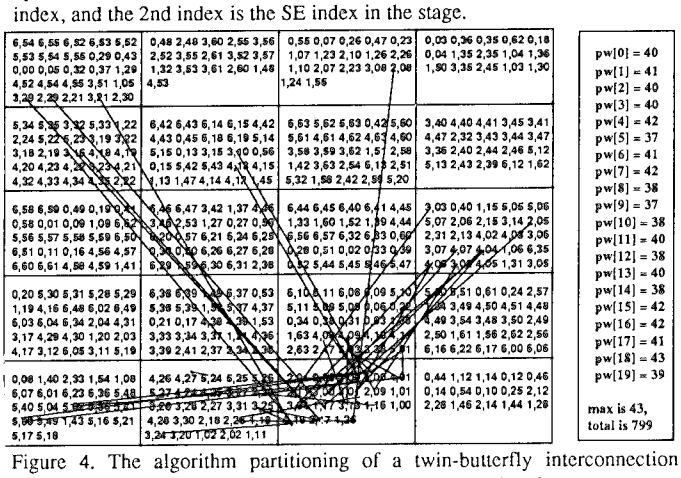


Figure 4. The algorithm partitioning of a twin-butterfly interconnection network into a 4×5 OE MCM (left). The numbers of optical interconnects on each chip are listed (right).

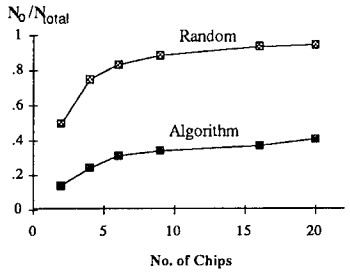


Figure 5. Comparison of the power dissipation of the interconnects between random partitioning and algorithm partitioning.

optical interconnects.

To solve this OE MCM partition problem, we present an algorithm based on a modified weighted matching algorithm, which is based on labeling techniques[3], and the Burkard's heuristic[4] adjusted to solve our optoelectronic MCM partitioning problem. The modified weighted matching algorithm is used here to minimize the power consumption. By virtue of using a matching algorithm, the size constraint is automatically satisfied. The optimized partitioning is achieved by rearranging the PEs among different chips to reduce the number of interchip interconnections, and therefore to reduce the power dissipation. To evaluate the algorithm, we compare the result with that of a straight forward partitioning which follows 2-D raster ordering inside each chip (as shown in Figure 3 for a 20 chip MCM with a maximum 25 SEs/chip). Figure 4 shows the result for the same setup done by the algorithm. The hottest chip in the 1st case dissipates 2 times more heat than the 2nd case. For various numbers of chips, Figure 5 gives a comparison of the results of random partitioning with the results of the CAD algorithm on an irregular interconnection network example[5]. A reduction in power of over 50% is achieved in many of the cases studied[6].

3. SE Placement in OE MCM

In the OE MCM design, placement is the design phase in which the exact physical location for each SE in the chips is determined. There are many ways to determine the placement of a system. For example, one can select the location of each SE by random ordering which we call a random placement, or by the natural 2-D raster ordering which we call a straight-forward placement. A good placement can result in reduced fabrication requirements for the CGHs in allowing larger OE system to be layout and built for the same fabrication requirement and cost. Since the fabrication technology increases with maximum interconnect distance, it is the maximum interconnection distance that should be minimized in the OE system placement.

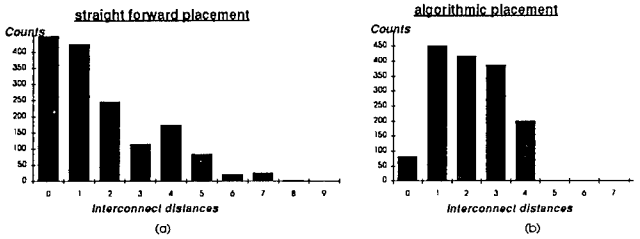


Figure 6. Comparison between two placement methods. (a) Result of the straight forward method. (b) Result after applying the placement algorithm. The maximum interconnect distance is reduced by 50%.

For example, in the case of placeing a multistage interconnection network onto an OE MCM where each stage with many SEs are placed on to one chip, an iterative matching algorithm can be used[5]. The physical model of such case has been shown in Figure 1. The algorithm starts with a random placement of SEs within chips and reduces the maximum interconnection distance through a fixed number of iterations by rearranging PEs within chips. For example, Figure 6 shows the histogram results of the straight-forward placement and of the placement by the iterative matching algorithm for comparison, for the same network example used in section 5. Figure 7 shows the actural layouts of the 1st stage of the interconnection network. Figure 7(a) is the result of the straight-forward placement method, and Figure 7(b) is the result of the algorithm placement. In both cases, the 1st stage is the worst-case for the whole network, and the worst case interconnects are

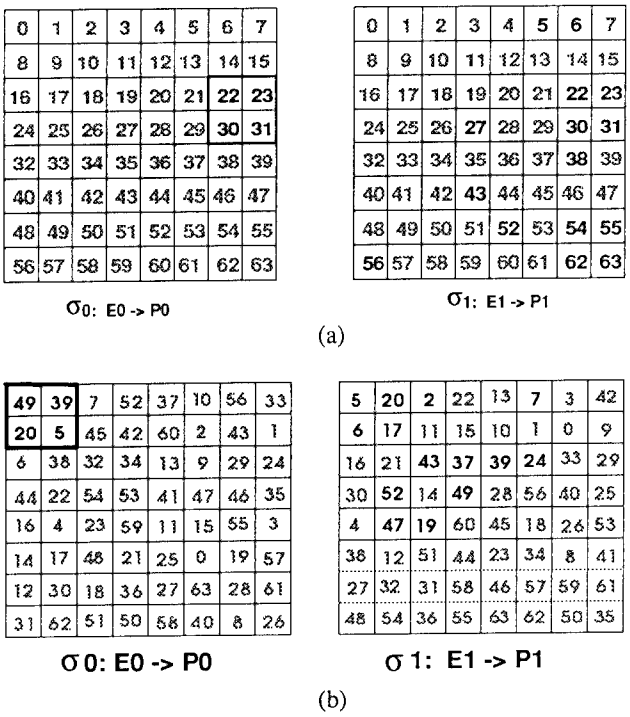


Figure 7. 2-butterfly OE MCM. Only the 1st stage is shown here. (a) Straight-Forward Placement. The connections for node 22,23,30 and 31 are marked. Worst case is 8.602 from 23 to 56. (b) Algorithm Placement. The connections for node 49, 39, 20 and 5 are marked. The worst case is 4.24 from 49 to 49. The improvement is about 50%.

marked down. More than a 50% reduction in the maximum interconnect distance (or on the packaging volume of the system) can be achieved when we applied the algorithms to the OE-MCM design.

5. Conclusion.

In this paper, we discussed optoelectronic MCM models and CAD algorithms for partitioning and placement. Results show that more than 50% reduction in power consumption and the maximum interconnect distance can be achieved by applying these CAD algorithms to an OE MCM design.

6. Acknowledgment

We would like to acknowledge the helpful discussions with B.Catanzaro, D.Zaleta, and V.H.Ozguz. We also thank the financial support of ARPA monitored by ONR.

7. References

1. M. R. Feldman, S. C. Esener, C. C. Guest, and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations", Applied Optics, 27, 1742-1751, 1988.
2. S. H. Lee, "Diffractive optics and computer generated holograms for optical interconnects", Critical Review conference on Diffractive and Miniaturized Optics, paper CR-14, San Diego, July 1993.
4. E. L. Lawler, "Combinatorial Optimization: Networks and Matroids", Holt, Rinehart and Winston, N.Y., Section 5.7, 1976.
3. R. E. Burkard and T. Bonniger, "A Heuristic for Quadratic Boolean Programs with Applications to Quadratic Assignment Problems", European Journal of Operational Research, 13, 372-386, 1983.
5. J. Fan, D. Zaleta, C.K. Cheng and S.H. Lee, "Physical Models and Algorithms for Optoelectronic MCM Layout", To appear in IEEE Transactions on VLSI Systems, 1994.
6. J. Fan, B. Catanzaro, V. H. Ozguz, C. K. Cheng, and S. H. Lee, "Design Considerations and Algorithms for Partitioning Opto-Electronic Multichip Modules", Proc. ACM & IEEE First International Workshop on Massivelly Parallel Processing Using Optical Interconnections, April 1994, Cancun, Mexico

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Design and Implementation of a Ring Array Optical Interconnect for SIMD Machines

J.-M. Wang, E. Kanterakis, A. Katz, Y. Zhang, Y. Li*

InterDigital Telecom, INC.
Division of InterDigital Communication Corp.
833 Northern Boulevard, NY 11021
Phone: 516-767 2254

* NEC Research Institute, INC.
4 Independence Way
Princeton, New Jersey 08540

Abstract: In this paper, the design and implementation of a high speed optical ring topology based free space optical interconnect is described. This interconnect system operates at 500 MHz and consists of 16 laser transmitters, a four channel free space interconnect, and a fast speed receiver. A Nearest Neighbor (NN) interconnect has been successfully demonstrated. At the data rate of 500 MHz, the total system throughput is 8 Gbps. The system can easily be operated at much higher data rates since the rate was only limited by the electronic circuitry. This interconnect is very promising in the implementation of ultra fast massively parallel SIMD machines.

1. Introduction

Electronic interconnections have been recognized as the bottlenecks of high performance computing systems. Because of their 3-D nature and their matched impedance, optical free space interconnects are the best alternative to electronic counterparts. Massively Parallel Machines (MPMs) require that the interconnect latency between processors be kept as small as possible. For this purpose, we have developed optical interconnects using ring topologies [1]. This novel interconnect arranges the processors on a ring instead of a rectangular grid configuration used on most of current systems. It relies on imaging, uses only spatial invariant elements and completely eliminates the boundary effect which exists in systems based on rectangular topologies. Operations required by the ring topologies can be implemented with simple image rotations using Dove prisms. Because of its interconnection regularity schemes, the proposed interconnect is appropriate for inter-processor communications of Single Instruction Multiple Data-stream (SIMD) machines [2]. As a continuation of our previous work, here we present the experimental implementation of a 500 MHz free space interconnect based on the ring topology. Section 2 briefly describes the ring topology. Section 3 describes the system setup, while the experimental results are shown in Section 4.

2. Ring topology

The principle of a Nearest Neighbor ring topology is described here. Assuming an array of N=nxn processors in a two-dimensional plane as shown in Fig.1(a). The NN interconnect can mathematically be expressed by

$$NN_{+1}(i) = i+1, \quad (1)$$
$$NN_{-1}(i) = i+1, \quad (2)$$
$$NN_{+n}(i) = i+1, \quad (3)$$
$$NN_{-n}(i) = i+1. \quad (4)$$

This is a variation of a mesh architecture that allows wraparound connections. The major problem is that the interconnection length between processors is not the same on the boundary as in the central region. This unequal interconnection length will cause system latency. The problem becomes severe when the total number of processors is increased. By arranging the processors in a ring as shown in Fig.1(b), this boundary effect can be eliminated. Other rectangular grid topologies such as PM2i and Hypercube[3], can be implemented in a similar fashion. In addition, operations required by the ring topologies can be implemented by performing simple image rotations.

3. *Optical system design*

For the implementation of the ring topology based interconnect system, we imposed the following constraints.

1. Keep system complexity down by minimizing the number of elements.
2. Only rotationally invariant optical elements were allowed.
3. Maintain a constant high interprocessor communication rate.
4. Maintain identical latency.

The schematic of the experimental system is shown in Fig.2. It consists of a data generator, an input ring assembly, a four-channel free space interconnect, and an output ring assembly. The input ring assembly consists of 16 laser transmitters; each being individually modulated by a 500 MHz PN data sequence generated by a data generator. In the four channel interconnect, each channel is designed to perform one of the four operations required by the NN interconnect. DP1-DP4 are Dove prisms. Each Dove prism is prerotated to a direction to perform the required imaging rotation operation. To demonstrate channel switching operations, a polarization mechanism is applied. BS1-BS6 are polarizing beamsplitters. Liquid Crystal (LC) switches (s1-s5) are employed in each channel. The selection of the interconnect channel can be done by controlling the settings of the LC switches. A single GaAs receiver is employed at the position of the output ring assembly. Data sequences for the four interconnect channels are received sequentially by setting the proper state of each LC switch. The received data sequence is then compared with that of the input.

4. *Experimental Results*

We have used the above system design to implement the NN interconnect. As mentioned above, the implementation of the NN interconnect requires four different operations: i+1, i+4, i-1, and i-4 connections. To perform these operations, the Dove prisms in all four channels were pre-rotated to certain angles so that each channel could realize one of the four operations required by the NN interconnect. In the experiment, channels 1 through 4 were assigned to perform the i+4, i+1, i-4, and i-1 connections, respectively. The data generator was set to generate a 500 MHz PN data sequence. During the experiment, set the LC switches so that channel 1 to channel 4 were sequentially selected. The experimental results are shown in Fig.3. The experimental results showed that the data sequences corresponding to the NN interconnect were received

correctly for all 16 transmitters. We also demonstrated that other topologies such as PM2i and Hypercube can be implemented by performing different rotation operations. This novel interconnect, using space invariant optical elements, having identical interconnect latency, and allowing network reconfigurability, will be suitable for high speed Massively Parallel Machines such as SIMD machines.

References

[1] Y. Li B. Ha T. Wang S. Wang A. Katz, X. J. Lu E. Kanterakis 1992 *Appl. Opt.***31** 5548-5558

[2] H. J. Siegel 1979 *IEEE Trans. Computers* **28** 907-917

[3] A. H. Wester 1972 *Proc. Saganoe Comput. Conf.* 29-40

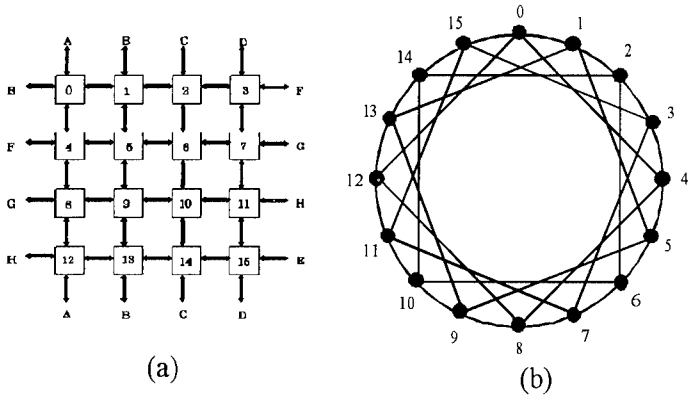


Fig. 1 (a) Nearest Neighbor (NN) interconnect configured on a rectangular grid.
(b) The same interconnect but with a ring topology.

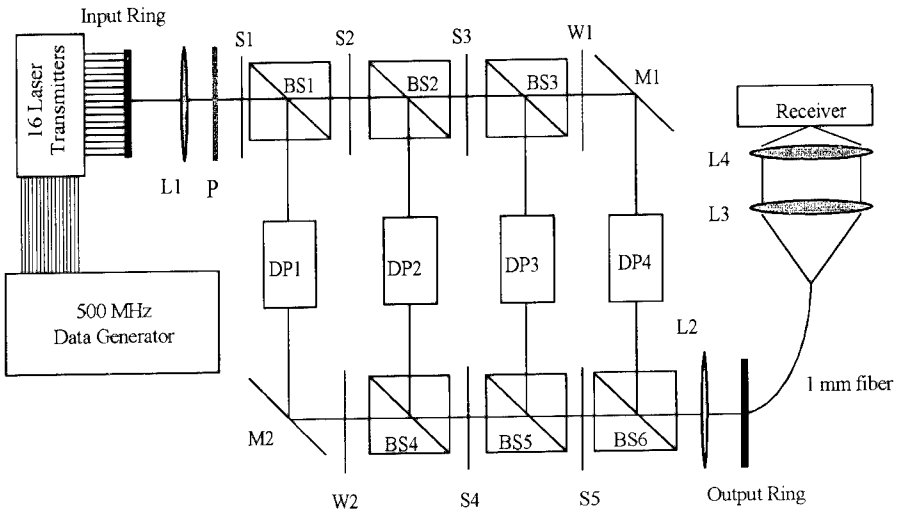
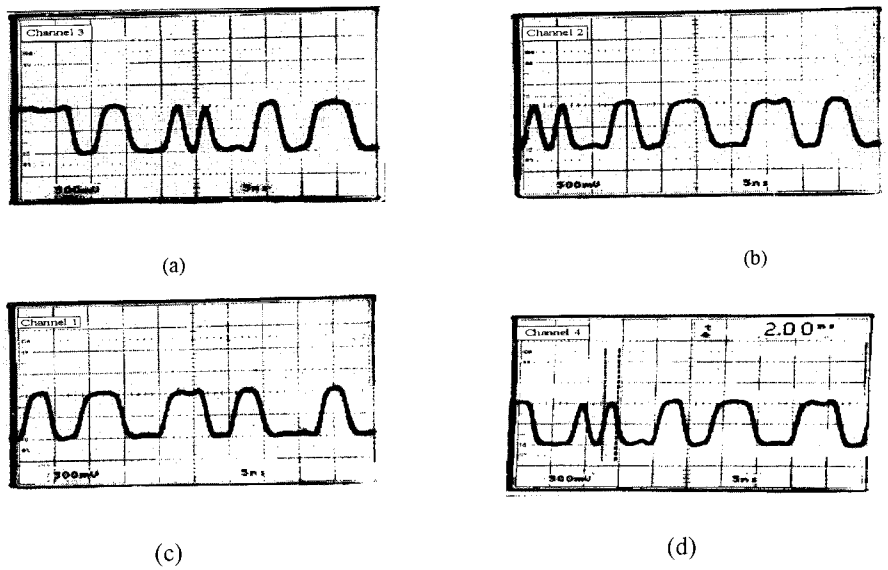


Fig. 2 Schematic of the experimental setup



(a) (b)

(c) (d)

Fig.3 Experimental results of the Nearest Neighbor interconnect. The pulse width is 2 ns. (a) corresponds to i+4 interconnect; (b) corresponds to i+1 interconnect; (c) corresponds to i-4 interconnect; (d) corresponds to i-1 interconnect..

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Spatial-Light-Modulator Based Routing Switches

W.A. Crossland, R.J. Mears, S.T. Warr, R.W.A. Scarr†

Department of Engineering, Trumpington St., Cambridge University, CB1 2PZ, (0223) 330264. †Consultant.

Abstract. Spatial light modulator based routing systems offer potential advantages over conventional electronic and waveguide-based systems. Integration of ferroelectric liquid crystal over silicon smart pixel modulators permits high levels of switch parallelism to be achieved.

1. Introduction

Free-space optical interconnection enables much greater fan-out and fan-in than can be accessed by electronics. These features make it appropriate to re-examine single-stage crossbar architectures. The purpose of this review is to summarise the technology of spatial light modulator routing systems and to consider scalability, control and arbitration issues.

2. Device technology

The spatial light modulators (SLMs) in the systems proposed here are intended to be silicon integrated circuits with an overlay of chiral smectic liquid crystal [1]. Very large electro-optic effects are observed in these birefringent ferroelectric materials – on inverting the polarity of the voltage across the liquid crystal layer (whose thickness is normally 2μm or less) the crystal optic axis rotates through an angle 2θ, which may vary from a few degrees to close to 90^o. The liquid crystal switching speed varies strongly with this angle.

In reflective devices built over silicon die, the factor limiting the overall attenuation ratio is the quality of the aluminium pads that form mirrors on the silicon die. Efforts are underway to improve this and to increase the optical quality of the modulators to approach that of devices in all-glass cells [2]. Attenuation ratios of several 100 have been measured for silicon backplane devices (compared with values up to 2000 for glass cells), and this is expected to increase with improvements in device technology and planarisation.

Currently optimised materials [3] with switching angles of approaching 45^o and wide phase temperature ranges are approaching 10μs switching speeds at voltages around 10V at 45°C, *figure 1.* In addition, very high switching angle materials ($\simeq 90^o$) such as Chisso 2004, can now provide polarization *insensitive* phase modulation of the input light, by removing the need for polarisers either side of the cell. However smaller switching angles (perhaps only a few degrees) and thus higher switching speeds may be useful in some

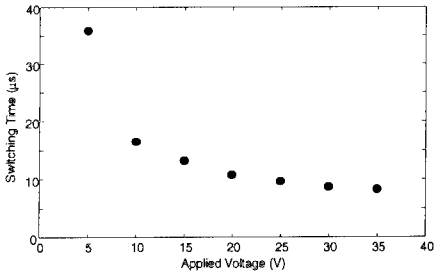


Figure 1: Switching time vs. voltage at 45°C for material DRA2

applications. In such devices it seems likely that switching speeds of the order of 100ns will be possible.

The very large electro-optic effects that are observed enable high fan-out/fan-in switch structures to be constructed, and the electronic functionality of the silicon backplane, capable of electronically controlled mirrors and sensitive photodetectors, forms a powerful 'smart pixel' technology.

3. Basic Switch Architectures

3.1. Matrix-matrix crossbar

An optical vector-matrix processor for fast Fourier transform calculations was proposed in 1978 [4], later identified as a single-stage, non-blocking space-switch, and then extended to the general matrix-matrix crossbar [5]. These architectures passively fan-out each optical input towards every output. The replications of the inputs are then 'shadowed' by means of a reconfigurable shutter array and hence selective fan-in is achieved at the output plane, *figure 2.*

Electrically addressed liquid crystal SLMs operating in a binary transmission mode are suitable high-speed switching arrays that can be used as the shutter plane. Each input replication must be optically resolved through a single shutter such that any arbitrary interconnection pattern may be formed, including broadcast, multicast and multiple-input fan-in. Reconfiguration of the switch simply involves closing any (single) shutters corresponding to completed calls and opening any new paths required.

The intrinsic replication of the optical inputs which is associated with this crossbar leads to a power loss of $\frac{1}{N}$ per input, given N-to-N routing. However by *imaging* the replications into the shutter plane, very little of the SLM needs to be optically active, allowing control and arbitration circuitry to be easily placed around the pixels. A passive diffractive phase plate or hologram is well suited to performing this fan-out operation. Fan-in can then be achieved using (multiple) lens arrays which could also be diffractive.

The argument of reciprocity, leads to the conclusion that if the input and output devices have the same numerical aperture, there must also exist a $\frac{1}{N}$ *fan-in* power loss in addition to the fan-out replication loss. An optically transparent matrix-matrix crossbar

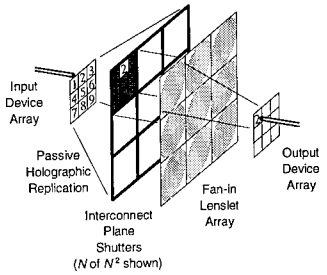


Figure 2: Matrix-matrix crossbar concept

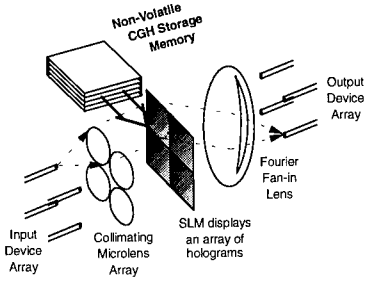


Figure 3: Holographic crossbar

with N inputs, placed in a single-mode fibre network, therefore exhibits a $\frac{1}{N^2}$ total power loss per input, plus any additional loss due to the optics, and will probably require pre- and post-amplification stages using optical amplifiers. However the fan-in loss may be avoided by using high numerical aperture devices at the output such as multi-mode fibres.

Due to crosstalk effects caused by the incomplete extinction of light through those pixels which are 'off', the output signal-to-noise (SNR) ratio and hence the switch scalability is essentially determined by the contrast ratio, C, of the liquid crystal shutters. Given N inputs, the SNR at any output port is of the magnitude $\frac{C}{N-1}$.

3.2. Dynamic holographic crossbar

The generation and use of computer-generated holograms is well documented, e.g., [6]. Such holograms are usually the coarsely quantized representations of the sampled 2-D Fourier transform of a certain desired output image or spot. The principle of operation of the holographic crossbar is the use of holograms to deflect as much optical power as possible from the inputs to the outputs by eliminating the initial fan-out operation associated with the generic matrix-matrix architectures. Thus the interconnect plane is divided into N routing areas and each routing area is filled with one from a set of base holograms which are possibly stored in a non-volatile memory behind the interconnect plane, *figure 3*. Each hologram acts as an independent diffraction grating and has an associated quantization noise limited diffraction efficiency, η, which will generally be less than unity [7]. Input broadcasting or multicasting may be achieved simply by designing a routing hologram to produce more than one output peak. The holograms are spatially invariant so fan-in from multiple inputs to a single output may also be achieved by placing the same hologram in more than one routing area.

An electrically addressed SLM operating in a binary *phase* mode produces a more efficient system than an amplitude mode device because the zero-order (i.e., non-diffracted path) can be eliminated in the output plane by ensuring there are equal pixel numbers of each phase state. However both configurations lead to a redundant rotational symmetry in the output plane because of the *binary* nature of the Fourier plane, which can only be removed by utilisation of more than two phase levels in the hologram.

A key scaling issue of the holographic interconnect is the number of pixels required per routing hologram to provide acceptable noise characteristics. Discrete Fourier trans-

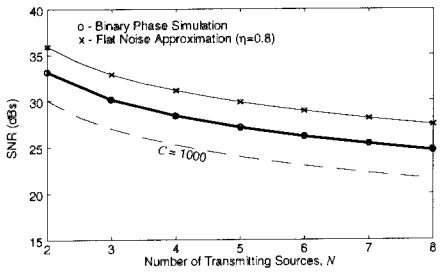


Figure 4: Signal-to-noise ratio at a single output port

forms have the property that to be able to resolve the N output ports, each routing hologram must contain mN pixels, where $m \geq 2$ because of the binary redundant symmetry. As an approximation, if we assume that m is sufficiently large that the Gaussian tails of the output peaks do not coincide and that the diffraction noise is uniformly distributed, the SNR at any output port will be of the order of $\frac{m\eta}{2(1-\eta)}$. As with the matrix-matrix crossbar, the holographic switch exhibits a $\frac{1}{N}$ SNR dependency, *figure 4*. The curves shown are for 44×44 pixel routing areas, compared to a matrix-matrix SNR curve with a contrast ratio of 1000.

4. Data throughput, switch control and arbitration

Free-space photonic switches based on ferroelectric liquid crystals offer high levels of parallelism and reconfiguration times may approach a microsecond. For optically transparent fibre optic switches the data throughput is determined by the optical power budget and can be very high. The reconfiguration time is considerably longer than the bit period of the data streams, but this is acceptable in many applications and any potential data loss during switch-over can be either accepted or can be avoided by scheduling the reconfiguration.

The routing algorithms for non-blocking single-stage switches are simple compared with those for multi-stage blocking switches since all paths through the switch are mutually independent. For the switches based on shuttering light beams it is simply necessary to open the shutter corresponding to the requested path through the switch. For holographic routing, once the path is requested, the required hologram pixel pattern is ascertained via a look-up table, and the control circuitry writes this pattern to the pixel array. There is such a hologram for every path through the switch. In both cases there is a group of pixels onto which the appropriate pattern must be written. For the matrix-matrix switch this is the fan-in group serving each output port. For the holographic switch it is the array of pixels serving each input port. An objective would be that the speed of arbitration and addressing via the silicon backplane should allow this to be accomplished in a time that is smaller than the switch reconfiguration time.

Control information must be communicated to the SLM in the interconnect plane of the switch. This information may be separated from the data paths through the switch,

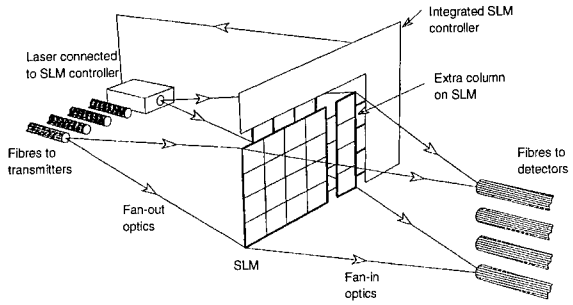


Figure 5: Overlay control network

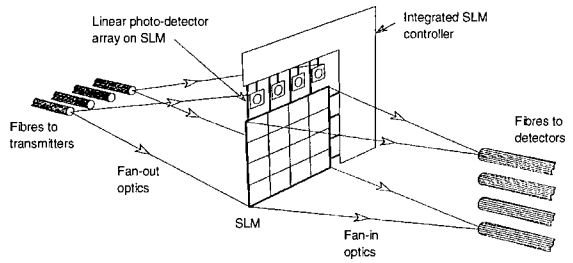


Figure 6: Embedded control network

carried on an overlay electronic or optical network (depending on the switching rate) linking the switch inputs and the SLM, *figure 5*, or it may embedded in the main data paths using the existing free-space optics and photo-detectors integrated on the SLM silicon backplane, *figure 6*.

In a space switching context, the two types of free-space photonic switches discussed here are strictly non-blocking single-stage designs. The detection of contention external to the switch and the taking of the steps necessary to stop multiple channels being directed towards the same output port are also considerably simplified in these structures (as opposed to multistage switches). Arbitration facilities for resolving contention can be built directly into the switch control electronics, and special circuitry to accomplish this can be placed in the silicon backplane of the interconnect SLM.

In a matrix-matrix switch, contention is avoided by simply only allowing one pixel to be opened in each 'fan-out' or 'fan-in' group. In this case, major simplifications to both arbitration and switch control result from the fact that there is only one unique path between switch input and switch output rather than a choice of paths as there are in multistage switches, and each such path is associated with only a single electronically accessible smart pixel.

In dynamic hologram switches, collision detection is best situated at the look-up table where all routing requests are received. This facility could perhaps be situated on the interconnect SLM, however separating the data and control paths and using an overlay

network for control is more likely in this case because of the extra difficulty in free-space communication with the interconnect SLM which is in a Fourier plane.

5. Conclusions

The data throughput of optically transparent fibre optic switches is determined by the optical power budget and can be very high. The use of smart pixel systems therefore provides a viable switch technology for connection of at least several hundred users. Both systems display good SNR and in principle could be made lossless and/or polarisation insensitive through use of appropriate materials.

The holographic architecture could potentially form a very low loss 1:1 interconnection switch. The output SNR is dependent on the number of hologram pixels used in each routing area and can be very high. The matrix-matrix switch requires fewer pixels and may operate faster because each *pixel* represents a connection path through the system, rather than a whole holographic image. The SNR is limited by the contrast ratio of the liquid crystal shutters.

Using FELC/VLSI technology, control and arbitration electronics can be associated with either SLM pixels in matrix-matrix switches, or arrays of pixels for holographic architectures, resulting in 'smart pixel' or 'smart hologram' systems.

References

[1] N. Collings, W.A. Crossland, P.J. Ayliffe, D.G. Vass, and I. Underwood. Evolutionary development of advanced liquid crystal spatial light modulators. *Appl. Opt.*, 228(22):4720–4747, November 1989.

[2] A. O'Hara, J.R. Hannah, I. Underwood, D.G. Vass, and R.J. Holwill. Mirror quality and efficiency improvements of reflective spatial light modulators by the use of dielectric coatings and chemical-mechanical polishing. *Appl. Opt.*, 32(28):5549, October 1993.

[3] We are grateful to Dr. D. McDonnell at DRA Malvern, UK, who developed this liquid crystal mixture based on a difluoroterphenyl host system.

[4] J.W. Goodman, A.R. Dias, and L.M. Woody. Fully parallel, high-speed incoherent optical method for performing discrete Fourier transforms. *Opt. Lett.*, 2(1):1–3, January 1978.

[5] A.G. Kirk, W.A. Crossland, and T.J. Hall. A compact scalable free space optical crossbar. In *Proc. Third International Conference on Holographic Systems, Components and Applications*, pages 137–141, Edinburgh UK, 16th September 1991. IEE London. Cong Publication No. 342.

[6] M.P. Dames, R.J. Dowling, P. McKee, and D. Wood. Efficient optical elements to generate intensity weighted spot arrays: design and fabrication. *Appl. Opt.*, 30(19):2685–2691, July 1991.

[7] D.C. O'Brien, W.A. Crossland, and R.J. Mears. A holographically routed crossbar: Theory and simulation. *Optical Computing and Processing*, 1(3):233–243, September 1991.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Development of an Optical Free-Space Crossbar

H.J.White[1], G.M.Proudley[1], C.Stace[1], N.A.Brownjohn[1], A.C.Walker[2], M.R.Taghizadeh[2], B.R.Robertson[2], C.P.Barrett[2], W.A.Crossland[3], J.R.Brocklehurst[4], M.J.Birch[4], M.Snook[5], D.Vass[5]

1 British Aerospace Sowerby Research Centre, PO Box 5 Bristol BS12 7QW, UK. Tel. 0117 936 3532.
2 Dept. of Physics, Heriot-Watt University, Riccarton, Edinburgh, UK. Tel. 031 451 3036.
3 Engineering Department, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, UK. Tel. 0223 330264.
4 THORN EMI CRL. Smectic Technology, Dawley Road, Hayes, Middlesex, UN3 1UK. Tel. 081 848 6445.
5 Dept. of Physics, Edinburgh University, The King's Buildings, Mayfield Road, Edinburgh, EH9 3JZ, UK. Tel. 031 650 1000.

Abstract A free-space single-stage optical crossbar switch has been designed and fabricated. A liquid crystal on silicon SLM is used as the active re-routing element. The initial results of the switch components are reported indicating the scalability of this technology.

1 Introduction

The use of free-space optics to perform high bandwidth connections for computing and communications applications has been widely studied over the recent years [1][2]. In this paper we present the initial practical results from a free-space, single-stage, optical crossbar switch which has been developed as part of a collaborative project entitled OCPM (Optically Connected Parallel Machines). Previous work from this project has been reported elsewhere [3]; here we present details of a 64×64 crossbar switch. The particular area that this switch addresses is where a high data bandwidth (>1Gbit/s) and large number of nodes (>16) are required together with a medium length reconfiguration latency ($\sim 10\mu s$).

The origins of the optical crossbar switch using the vector-matrix multiplier architecture have been detailed previously [4] and a number of designs have been described. The performance and compactness of space switches of this type are enhanced by adopting a two dimensional crossbar architecture [5] as shown in figure 1. This matrix-matrix structure can be implemented using free-space optics, based on binary phase gratings and spatial light modulators (SLMs), based on ferroelectric liquid crystals integrated with silicon VLSI (FELC/VLSI)[6]. These SLMs allow large arrays of shutters (or optical crosspoints) to be made, each with a large contrast ratio. For switches with modest reconfiguration times, these features remove the main motivations for using multi-stage re-arrangeable networks and may allow switch control functions to be simplified by the use of the single-stage crossbar architecture. The control and arbitration functions can

be integrated into the silicon VLSI backplane of the crosspoint SLM to form a compact and eventually self-routing system.

If intrinsic fan-in loss is to be avoided, the collection channel must have a greater numerical aperture (N.A.) and/or area than that of the input channels. If the input is from N mono-mode fibres, then the output could be collected into a large cross-section multi-mode fibre that supports at least N modes. By totally filling the N.A. of the multi-mode fibre the number of input channels which may be used before intrinsic fan-in loss occurs is given by:

$$N = \left(\frac{r_2(N.A.)_2}{r_1(N.A.)_1}\right)^2 \approx \left(\frac{4r_2(N.A.)_2}{\lambda}\right)^2 \quad (1)$$

where r_1 is the radius of the beam emitted from the mono-mode input fibres and r_2 is the beam radius accepted by the multi-mode output fibre. For mono-mode fibres $(N.A.)_1 r_1 = \frac{\lambda}{4}$.

2 Description of 64×64 switch

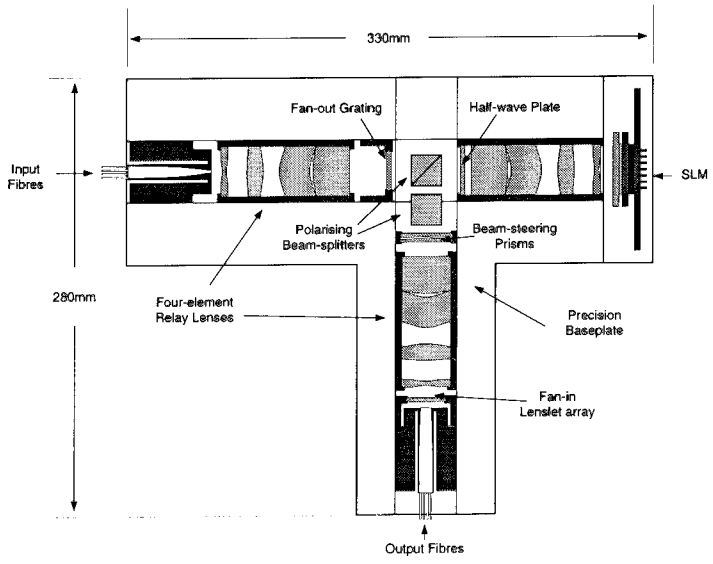


Figure 1: Plan of the 64×64 crossbar baseplate

Figure 1 shows the layout of the 64 input, 64 output optical system developed. A matrix-matrix crossbar design has been used and data inputs to the switch are provided by 790nm laser diodes pigtailed to polarisation preserving fibres. To create the most compact system, the design uses one-to-one imaging and hence the input fibres are arranged in an array that matches the pixel spacing of the SLM. These are 80μm cladding diameter polarisation preserving fibres which are arranged in an 8×8 square array with a 120μm pitch. Each fibre has its polarisation axes orientated in the same direction. This

fibre array was fabricated in an array of holes formed by excimer laser material ablation of a Kevlar substrate [7].

To obtain a compact and rugged system, the optical elements have been mounted on a baseplate [8] which allows fine alignment of the system. The light from the input array is fanned out by means of a binary phase grating based on a non-separable two-dimensional design. A half-wave plate is used to align the polarisation of the channels with the liquid crystal axes of the SLM. The reflected outputs of the SLM are then re-imaged onto the holographic fan-in array which focusses the light into the multi-mode output fibres. The output fibres are 300μm core diameter, step-index with an N.A. of 0.22 arranged to form an 8×8 square on a 960μm pitch. The substrate used for this array is two layers of silicon with arrays of wet-etched holes [7]. Putting the specifications of this fibre into equation 1 implies that a fan-in of approximately 28,000 is achievable before any intrinsic fan-in losses occur. In practice, it is less than this due to misalignments in the fibre arrays, but is still much greater than the 64 of the present system. The fan-out and fan-in elements are such that each input can address all the output fibres, the state of the SLM determining which channels are open at any given time.

3 Initial results

The laser diodes and photo-diodes are being incorporated into transceiver units which will enable high bandwidth signals to be sent through the switch. Initial results from the transceiver units indicate a performance of 500Mbits/s at a Bit Error Rate (BER) of 10^{-10} at -23dBm. Table 1 indicates the current losses expected from the 64×64 switch when fully assembled. With the 6dBm sources to be used in the system we expect to achieve these data rates with this prototype system. The liquid crystal SLM is operating in the book-stack geometry, providing a high contrast latching device, with reconfiguration speeds of the order of 20μs. Control and arbitration of the switch is performed through a P.C. interface during this evaluation stage, however, it is intended to incorporate this functionality on the silicon backplane of the SLM.

Table 1: 64×64 crossbar losses

Component	Losses	
	transmission	dB
Fusion splice (laser to fibre array)	0.98	-0.10
Intrinsic fan-out	0.0156	-18.06
Fan-out element	0.75	-1.25
SLM	0.70	-1.55
Reflection losses	0.79	-1.00
Fan-in element	0.50	-3.01
Coupling into output fibre	0.95	-0.22
Detector Coupling	0.96	-0.20
TOTAL	0.00289	-25.39

4 Conclusions

The development of free-space optical crossbar switching technology has been achieved. The 64×64 matrix-matrix single-stage switch described here is designed to have no intrinsic fan-in losses and therefore is scalable to larger sizes dependent upon the inherent fan-out losses. The performance of the present components indicate that this 64×64 switch can sustain a data rate of 500Mbits/s at a BER of 10^{-10} and will be extendible to a 128×128 switch at 2.5GBits/s. The single stage nature of the switch simplifies the required contention and arbitration issues which become a major factor for a large switch.

References

[1] M.Fraces J.P.Bouzinac P.Churoux M.Laug D.Compte P.Siron X.Thibault. In J.W.Goodman P.Chavel G.Roblin, editor, *Optical Computing 88, Proc. SPIE 963*, pages 223–231, 1988.

[2] P.J.DeGroot J.R.Noll. Reconfigurable bipolar analog optical crossbar switch. *Applied Optics*, 28(8):1582–1587, 15 April 1989.

[3] H.J.White J.R.Brocklehurst W.A.Crossland I.M.McLaren D.M.Monro and A.C.Walker. Optically Connected Parallel Machines. In *JFIT Technical Conference Digest*, pages 189–196, March 1993.

[4] J.W.Goodman A.R.Dias L.M.Woddy. Fully parallel, high-speed incoherent optical method for performing discrete Fourier transforms. *Optics Letters*, 2(1):1–3, Jan. 1978.

[5] A.G.Kirk W.A.Crossland T.J.Hall. A compact and scalable free-space optical crossbar. In *Proc. of the 3rd International Conference on Holographic Systems, Components & Applications* , pages 137–141, 1991.

[6] N.Collings W.A.Crossland P.J.Aycliffe D.G.Vass I.Underwood. Evolutionary development of advanced liquid crystal spatial light modulators. *Applied Optics*, 28(22):4740–4747, 15 November 1989.

[7] G.M.Proudley C.Stace and H.J.White. Fabrication of two-dimensional fiber optic arrays for an optical crossbar switch. *Optical Engineering*, 33(2):627–635, February 1994.

[8] F.B.McCormick F.A.P.Tooley J.L.Brubaker J.M.Sasian T.J.Cloonan A.L.Lentine R.L.Morrison R.J.Crisci S.L.Walker S.J.Hinterlong M.J.Herron. In *Optical Applied Science and Engineering, Proc. SPIE 1533*, page paper 12, 1991.

The authors would like to acknowledge the invaluable work of D.J. McKnight, I. Redmond, G. Smith and D. Neilson in the development of the crossbar switches. The OCPM project (Optically Connected Parallel Machines) is a collaboration between British Aerospace (Sowerby Research Centre), BNR Europe, Heriot-Watt University, the University of Bath and THORN EMI CRL. The project is coordinated by British Aerospace and is funded in part by the DTI and EPSRC.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A compact holographically routed optical crossbar using a ferroelectric liquid-crystal over silicon spatial light modulator

D.C.O'Brien, Douglas J. McKnight, Optoelectronic Computing Systems centre, Campus Box 525, University of Colorado, Boulder, CO 80309., USA. (tel 303 492 3330)

Abstract. A demonstration of a 16 channel holographic crossbar switch is presented. The switch uses a 4x4 array of vertical cavity surface emitting lasers (VCSEL's) as individual channel inputs. Light from each laser illuminates a portion of a ferrolectric liquid crystal spatial light modulator(SLM), which displays a computer generated hologram(CGH) in a binary phase mode. Holograms are designed to direct the appropriate source to the desired output port of the crossbar. Simulation and experimental results are presented.

1. Introduction

Free space interconnects have wide application in optical and optoelectronic computing and information processing. Such interconnects are routinely implemented using binary and multilevel phase CGH techniques (see for instance [1]) and excellent results are reported. However, most holographic interconnects are fixed, which limits their flexibility. In this paper a reconfigurable holographic interconnect structure is presented, based on the crossbar geometry [2].

2. Overview

Figure 1 shows a schematic of the crossbar switch. An array of collimated beams, one for each input port, illuminates an SLM. Each input illuminates a separate Fourier computer generated binary phase subhologram, which is designed to route light to the centre of the desired output detector. The Fourier lens provides the necessary transform and fans-in light from all the input channels. For a 16 channel crossbar there are 16 base holograms, each of which routes light to a particular output port. The state of the switch is set by displaying the correct hologram in front of the correct illuminating beam.

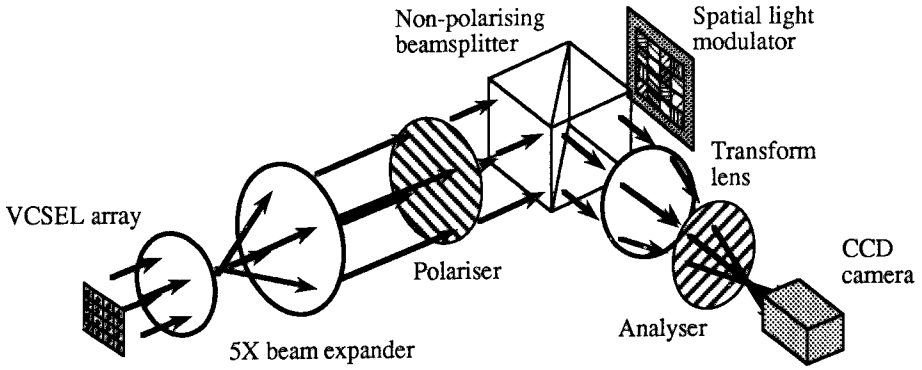


Figure 1. Schematic of holographic crossbar

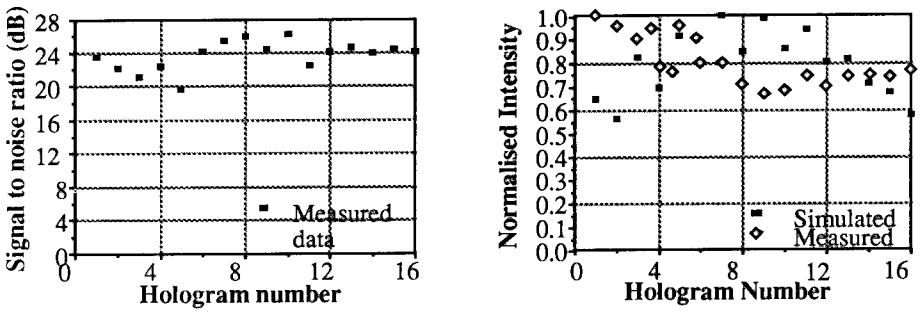


Figure 2. Graphs showing peak uniformity and hologram signal to noise ratio

An array of vertical cavity surface emitting lasers (VCSELs) is collimated using a microlens array. The array of beams then passes through a beam expander (approximately 5X) so the VCSEL array matches the hologram pitch on the SLM. The SLM displays a 4x4 array of subholograms in a binary phase mode. Light is reflected off the SLM, and an achromat takes a Fourier transform to create a hologram replay at the output plane. A CCD camera is placed at the output plane of the switch. Each component is described more fully below;

2.1. Sources

The crossbar uses a 4x4 array of VCSELs, manufactured by PRI[3]. The complete array is 8x8 in extent, on a pitch of 250 μm. The lasers emit a close to Gaussian beam with a numerical aperture of $\approx$0.1, at a wavelength of $\approx$840nm. The output power is up to 2mW. A fused silica spacer is glued to the surface of the laser array and a microlens is glued to the spacer. The thickness of the spacer is designed so that the sources lie at the focus of the microlenses, creating collimated beams.

2.2. Spatial light modulator (SLM)

The SLM used in the experiment is a 256 x256 liquid crystal over Silicon SLM, with 21.6 μm pixel pitch, and a total aperture of 5.53mm. It is described fully in [4]. The device is used in a binary phase mode, the phase states being obtained thus: Each pixel can be considered to be a waveplate which can be electrically switched so that its optic axes rotate through an angle 2θ in the plane of the pixel, where θ is the tilt angle of the liquid crystal. The input light to the device is linearly polarised, the polarisation oriented so it bisects the angle the optic axes of the liquid crystal switches through, and the output analyser is placed orthogonal to the polariser. The two possible output states are of equal amplitude and π radians apart. There is a loss penalty, which can be removed if $\theta = \frac{\pi}{4}$ (normally $\theta = \frac{\pi}{8}$ for intensity modulation) and the pixel is a half waveplate. In this particular case the phase modulation is polarisation independent.

A preliminary experiment was undertaken to examine the phase performance of the SLM. A series of horizontal stripes is written to the SLM and these are viewed in a phase mode. This creates a grating which should have no zero order component in replay. The ratio of the first to zero orders was measured to be $\approx$ 35:1, which compares favourably with other ferroelectric devices [5] and shows the device has a robust binary phase mode.

2.3. Holograms

The subholograms are 64x64 pixels in extent and are designed using a simulated annealing technique. They are designed to route to a 2x8 array of output ports, corresponding to a Silicon photodetector array which will form the output plane in future crossbar demonstrations.

3. Experimental Investigation

Successful operation of the crossbar requires that (i) the light from the hologram peaks are sufficiently uniform and that (ii) the signal to noise ratio (SNR) of the holograms is large enough so that outputs which should be dark do not receive an 'on' level. The worst case crosstalk (defined as the ratio of the highest received dark level to the lowest received on level) should be less than one for correct operation.

Each base hologram was displayed on the SLM and the replay field recorded. A plot of the hologram spot intensity is shown in Figure 2 along with the measured SNR. Together these can be used to simulate the received power at each detector for all the potential states of the switch. This gives a worst case crosstalk of $\approx$ 0.1 which indicates the crossbar should function with good noise margin.

Figure 3 shows an image of the SLM and the output plane with all ports illuminated. The increased zero order is due to the crystal alignment varying over the surface of the SLM. The spot pattern should form a regular grid, but surface curvature of the particular SLM used causes the mismatch at the replay plane. Compensation for this can be built into the hologram design process. The SLM transmits about 0.4% of the light an ideal device would pass and this is where the majority of the losses within the system lie. However, the SLM used was optimised for shorter illumination wavelengths. Using a device optimised for this wavelength and a more suitable liquid crystal would

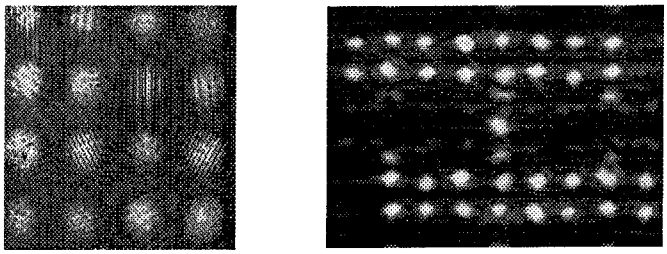


Figure 3. Image of SLM Hologram replay (note symmetrical 2x8 output arrays with zero order between)

be expected to improve this figure considerably. The VCSELs are attractive as sources, as they are compact, easily modulated and emit relatively high power. However, they have two orthogonal output polarisations, and oscillate between them unpredictably. In the experiment a waveplate was used to rotate the VCSEL polarisation so that the SLM input polariser bisected the two orthogonal outputs, so that all sources were roughly equal intensity. More suitable solutions would be to optimise the SLM for polarisation independent modulation, or to control the VCSEL polarisation state.

4. Conclusions

The experiment demonstrates the the principles and some of the attributes of this type of architecture. It has no intrinsic fan-out loss and alignment is straightforward. It also appears possible to design holograms which compensate for non-ideal system components. SLMs with high space bandwidth product are required but such devices continue to become available.

Dynamic holographic routing has many advantages and a modest improvement in performance should make such routing structures attractive for optical processing applications.

References

[1] In *4th Int. Conf. on Holographic Systems, Components and Applications.* IEE, September 1993.

[2] J.W.Goodman, A.R.Dias, and L.M.Woody. Fully parallel, high speed incoherent method for performing discrete Fourier transforms. *Optics Letters*, 2(1):1–3, 1978.

[3] Photonics Research Incorporated, 2402 Clover Basin Drive, Suite A,Longmont CO80503 U.S.A.

[4] Kristina M. Johnson Douglas J. McKnight and Roylnn A Serati. 256x256 liquid-crystal-on-silicon spatial light modulator. *Applied Optics*, 33(14):2775–2784, 1994.

[5] D.C.O'Brien, T.D.Wilkinson, R.J.Mears, and W.A.Crossland. Dynamic holographic interconnects that use ferroelectric liquid-crystal spatial light modulators. *Applied Optics*, 33(14).

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Low-loss and High-speed Optical Switching Modules for 1.3 μm Wavelength Using Active-matrix Ferroelectric Liquid Crystal Devices

Seiiti Shirai, Tadashi Serikawa, Shigeto Kohda, Nobuhiko Kakuda, Masamichi Okamura and Noriyoshi Yamauchi

NTT Interdisciplinary Research Laboratories
3-9-11, Midori-cho, Musashino-shi, Tokyo 180 Japan
Tel: +81-422-59-2942, Fax: +81-422-59-3870

Abstract. Highly efficient optical coupling system using microlens arrays for free-space optical network was demonstrated. High speed (<100 μs) and high extinction ratio for the 1.3 μm wavelength were obtained by a proposed ferroelectric liquid crystal driving method.

1. Introduction

Two-dimensional optical switching devices consisting of liquid-crystal cell arrays and birefringent plates have many advantages in making large-scale, transparent optical switching networks. We have demonstrated that an 8-stage optical concentrator using 1024-input-port optical beam shifter modules operate effectively for visible light [1, 2]. To use this type of module in an optical communication network, good performance for long-wavelengths, highly efficient fiber-to-fiber coupling and high-speed switching operation are required. This paper describes a new optical design for low-loss interconnection and a new driving method for high-speed ferroelectric liquid crystal (FLC).

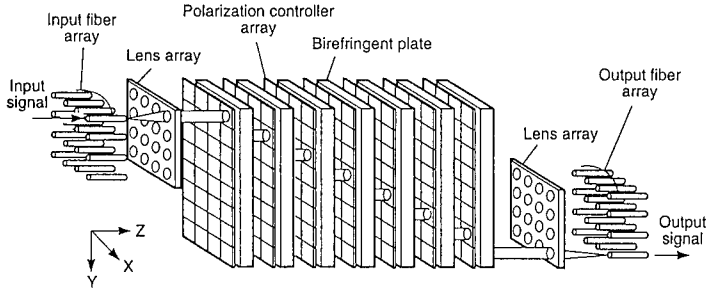
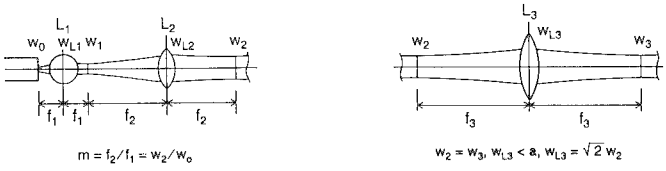


Fig.1 N × N free-space optical switching networks[3].

2. Optical design for fiber interconnection

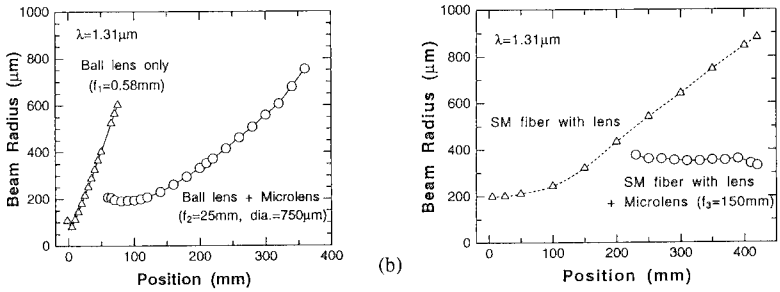
Figure 1 shows an example of an N × N optical switching network [3]. The distance between input and output fiber arrays is several tens of centimeters, because the system consists of cascaded optical beam shifter modules. Thus, beam expansion due to diffraction results in very large increases of the coupling loss and limits the size of the switching network.

Figure 2 shows the newly proposed configuration of lenses for low-loss propagation. Figure 2(a) shows the confocal lens system, which consists of a very short focal lens and a long focal lens, for obtaining the long collimation distance. Figure 2(b) shows the relay lens with long focal length for compensating beam expansion. The number of relay lenses depends on the size of the switching system. This relay lens can compensate the beam expansion. Figure 3(a) shows the changes in beam radius along the propagation direction in Fig. 2(a) using a ball lens (f_1=0.58 mm) and sputter-liftoff microlens (f_2=25 mm). A long collimation distance was obtained by long focal length microlens. Figure 3(b) shows the compensation of beam divergence by relay lens (f_3=150mm). The radius of the beam from single mode fiber with collimation lens was reduced to about 40% at the position of 400mm. These long focal length microlens were fabricated by sputter deposition using the shadowing effect. Figure 4 shows a photograph of the sputter-liftoff microlens array (32 × 32, dia.=750μm).



(a) Confocal lens system (b) Relay lens system

Fig.2 Microlens configuration for long-span beam propagation.



(a) (b)

Fig.3 Propagation profiles for (a) collimation lens system and (b) relay lens system.

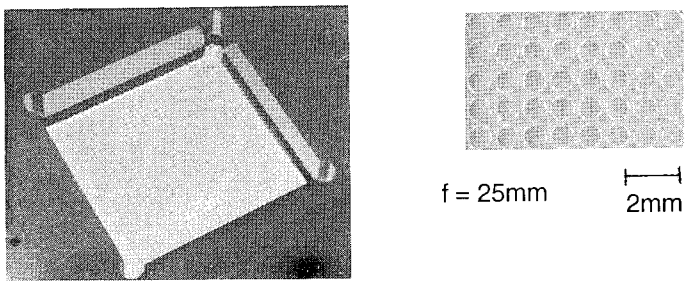


Fig.4 Photographs of sputter-liftoff microlens array (32 × 32, dia.=750μm).

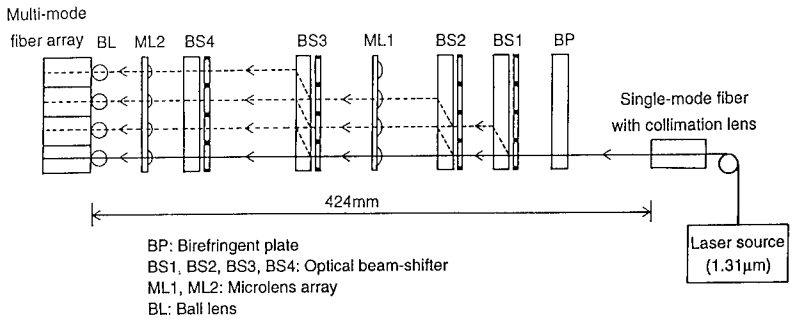


Fig.5 Experimental set-up for fiber-to-fiber coupling using 1 × 4 optical switching network.

Figure 5 shows the experimental configuration of a polarization independent 1 × 4 switching system. On the input side, an input beam is separated into two components, an ordinary-ray and an extraordinary-ray. These beams go through the two adjacent cells in each optical beam shifter. At the final stage, the polarization of each beam is controlled and the beams are combined into one beam with non-polarization. Optical beam shifter modules with 990 μm cell size and 350 μm aperture radius are used. The output fiber array consists of 4 × 4 stacked ceramic ferrules [4]. Two long focal lenses, ML1 (f=150 mm) and ML2 (f=25 mm), are very effective for highly efficient optical coupling between an input fiber and an output fiber array. Very low total insertion-loss of 8.7 dB was obtained for four output ports. This value proves the accurate coupling to output fibers in spite of the very long propagation span of 424 mm.

3. High speed optical switch for 1.3 μm light using FLC

Ferroelectric liquid crystal is a very attractive material for high-speed operation (two to three orders faster than a twisted-nematic liquid crystal). There are, however, two main problems to solve in adapting it to optical switching devices for long wavelength. To obtain the high extinction ratio, the cell thickness must be twice as thick as for visible light. This thicker cell thickness reduces the FLC memory effect. The ordinary bipolar driving method is not adequate for optical switching because of the short break of the optical signal by the reset pulse.

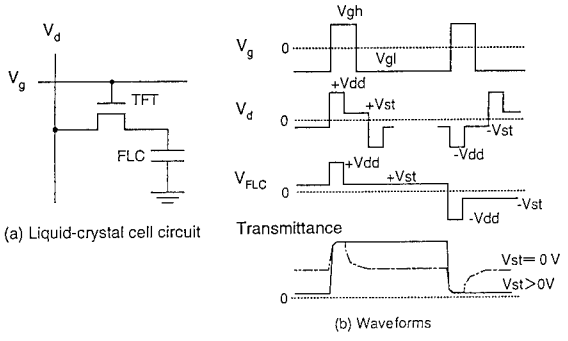


Fig.6 New driving method for ferroelectric liquid crystals.

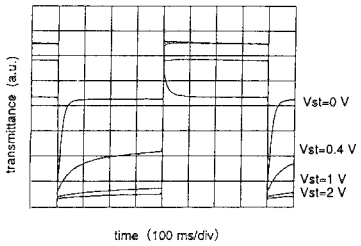


Fig.7 Switching waveforms of transmitted light.

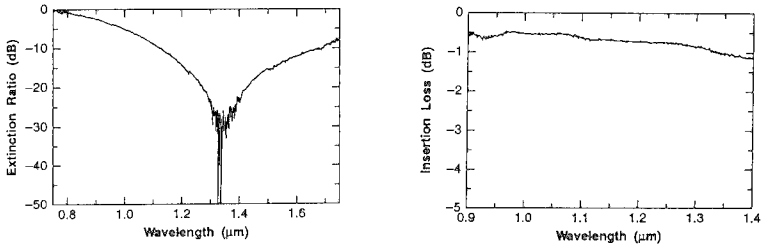


Fig.8 Dependence of extinction ratio and insertion loss on wavelength.

To overcome these problems, we developed a new driving method for the FLC optical switch. An active-matrix driving method with the voltage waveforms shown in Fig. 6 was used. The unipolar data driving pulse can prevent the short signal break. The low voltage part of the pulse can sustain the cell data without loosing reliability and the high voltage part results in high speed switching and a high extinction ratio. Figure 7 shows the switching waveforms of transmitted light for various sustaining voltages V_{st}. The high-speed switching operation (<100 μs) was confirmed. This value is two orders faster than for a twisted-nematic liquid crystal. As shown in Fig.8, high extinction ratio (>30 dB) and low insertion-loss (<1 dB) were obtained at the 1.3 μm wavelength.

4. Conclusions

We proposed new microlens system configuration for highly efficient optical coupling. Experiments demonstrated the system's low propagation loss. The long-span microlens system was shown to be very useful for large-scale optical switching networks. An FLC driving method for optical switches in optical communication networks was proposed. High-speed operation and high extinction ratio were obtained by this new method. This optical switching modules are applicable to large scale free-space optical switching networks.

We would like to thank Dr.Tomoyuki Toshima for his encouragement throughout the course of this work.

References

[1] Shirai S, Kimura K, Irita Y, Naito N, Serikawa T, Kohda S, Yamauchi N and Yamaguchi M 1992 *SPIE Photonic Switching* **1807** 328-334

[2] Yamaguchi M, Matsunaga T and Shirai S 1992 *SPIE Photonic Switching* **1807** 435-437

[3] Yamaguchi M, Matsunaga T, Shirai S and Yukimatsu K 1994 *IEICE Trans. Commun.* **E77-B** 163-173

[4] Koyabu K, Ohira F, Yamamoto T and Matsuno S 1993 *OFC/IOOC'93 Technical Digest* TuB2

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Free-space optical interconnections using liquid-crystal microprism arrays

Katsuhiko Hirabayashi, Tsuyoshi Yamamoto, and Masayasu Yamaguchi

NTT Communication Switching Laboratories, 3-9-11, Midori-cho, Musashino-Shi, Tokyo 180, Japan

Abstract: Liquid-crystal microprism arrays are shown to be very useful for electrically controlled alignment of optical beams and free-space optical interconnection. They can deflect closely spaced micro-optical beams individually to any position with a high transmittance of 95% and a large deflection angle. Various optical interconnections can be made simply by changing the voltage applied to each microprism .

1. Introduction

One problem with free-space interconnection is the complicated mechanically controlled alignment. Using many mechanical micropositioners for alignment within several micrometers is hard work and increases the module size. Furthermore, for optical interconnections among two-dimensional (2D) optical switches, various optical interconnections such as crossing, parallel, and concentration are necessary. Hologram cells, prisms, and lenses provide fixed optical interconnections. Thus for techniques electrically controlled alignment employing optical beams and deflectors able to deflect in arbitrary directions are desired for optical interconnection. Various kinds of optical beam deflecting device, such as holographic diffraction cells[1-3], have been developed for free-space optical interconnection. However, they have drawbacks such as large loss, low steering angle, undesired high-order diffraction peaks, and a limited wavelength range. Sato et al. [4] explained that LC Fresnel lenses change not only the focal length but also the propagation direction of the optical beams. We started with these devices and improved their performance for free-space optical interconnection. We call them LC microprism arrays. This paper investigates the application of these arrays to optical interconnection for electrically controlled alignment of optical beams and various interconnections.

2. Device structure and principle of beam deflection

The structure of the LC microprism array is shown in Fig. 1. This is a homogeneously aligned nematic LC cell, in which microprisms with a pitch of 250 μm are fabricated on the bottom transparent glass plate.

When no voltage is applied, the LC molecules align parallel to the glass plate. The input optical beam is refracted at the prism plane according to Snell's law. When voltage is applied, the LC molecules are realigned, and the refractive index decreases from n_e to n_o. The deflection angle thus changes with applied voltage. The deflection angle is determined by the refractive indices of the LC and the transparent plate, and by the apex angle. A larger deflection

angle can be obtained by using an LC with larger refractive anisotropy. We used a Merck BL-009 LC because it has the largest refractive anisotropy: 1.8181 (n_e) and 1.5266 (n_o).

3. Experiments and Results.

We prepared LC microprism arrays with apex angles of 10°, 20°, 30°, and 40°. The striped ITO electrodes of the LC microprism were connected in parallel by a flexible cable to variable resistors (5 kΩ). The applied voltages are controlled by these variable resistors.

First, to clarify whether LC microprism arrays are feasible for optical interconnection, we measured their basic performance using a single microbeam with a beam waist of about 100 μm. The beam profiles of the deflected beams are shown in Fig. 2 for the apex angle of 30°. The applied voltage dependence of the deflection angles is shown in Fig. 3 as a parameter of the apex angle. The maximum deflection angle was 15°. The optical beams were deflected continuously by the LC microprism up to 2.8 V. Transmittance was no less than 95%. However, above 2.8 V, the optical beams were deformed and split into two peaks. Furthermore, when the apex angle was more than 40°, the optical beams were deformed even with no applied voltage. This is because the LC molecule alignments are not homogenous. Thus, for optical interconnections, the apex angle should be smaller than 30° and the applied voltage should be lower than 2.8 V. Thus available deflection angle was 8°. The response on-time was about 1-2 seconds, independent of the deflection angle and the apex angle. The response off-time was very slow (10 seconds -40 seconds), and increased with the apex angle.

In applying the LC microprism arrays to optical interconnections among 2D switches, as shown in Fig. 4(a), we demonstrate deflections of optical beams from a fiber array with the collimating microlens arrays shown in Fig. 4(b). An 8 x 8 fiber array with a 250-μm pitch was used; it had a planar microlens array with the same pitch as that of the fiber array for the collimating light beams. The apex angle of the LC microprism array we used was 20°. The propagation directions of the beams from the fiber array with the collimating lens were not perfectly the same, but a little random even if the LC microprism array was not used. This is because the fiber arrays were assembled with a slight positioning error (a few micrometers). By changing the resistors connected to the LC microprism, and monitoring the spot positions with the CCD camera, the propagating direction of each beam could easily be controlled, and any optical interconnection was possible, as shown in Fig. 5. Next, we demonstrate various optical interconnections using the crossed two-LC microprism arrays, which can deflect optical beams in two directions X and Y. The outputs of the CCD camera are shown in Fig. 6. Each of eight optical beams from the fiber array could be deflected in arbitrary directions.

Finally, we show the concept of a 2D optical-switch network using the LC microprism arrays in Fig. 7. It consists of 2D optical switches, microlens arrays, LC microlens arrays, and an LC macrolens. They are cemented together and mounted in square holders, then mounted on an L-shaped block. Optical interconnections between the optical switches are achieved by tuning the variable resistors connected to the LC microprism arrays. Complicated, precise, mechanically controlled alignment is not necessary to arrange the optical components because the output beam directions can be controlled electrically with precise positioning. This module is very compact because it does not use macrolenses or diffract devices.

4. Summary

LC microprism arrays were shown to be very useful for free-space optical interconnection and electrically controlled alignment of optical beams: these arrays can deflect closely spaced optical beams individually to any point with high transmittance (95%), at a high deflection angle, and a low voltage (< 2.8 Vrms). Various optical networks can be constructed by simply changing the voltage applied to each microprism array.

Acknowledgments

We wish to thank Dr. Susumu Sato of Akita University for his helpful discussions.

References

[1] McCormick F B, Cloonan T J , Tooley F A, Lentine A L L, Sasian J M, Brubaker J , Morrison R L, Walker S L, Crisci R J , Novotny R A, Hinterlong S J, Hinton H Sand Kerbis E, 1993 Appl. Opt., **32**, pp. 5153-5171,.

[2] Pauliat G , Herriau J P, Delboulbe A, Roosen G, and Huignard J P, 1986 J. Opt. Soc. Am. **B2**, pp. 306-314 .

[3] Tervonen E, Friberg A T, Westerholm J, Turunen J, and Taghizadeh M R, 1991 Opt. Lett., **16**, pp. 1274-1276 .

[4] Sato S, Nose T, Yamaguchi R and Yanase S, 1989 Liquid Crystals, **5**, pp. 1435-1442 .

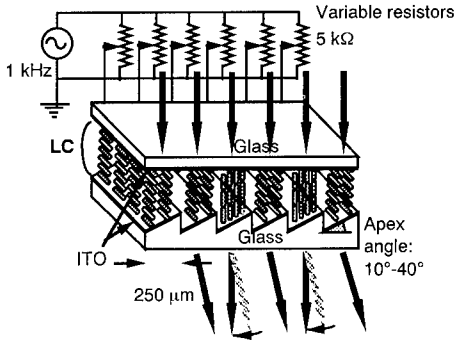


Fig. 1 Structure of LC microprism

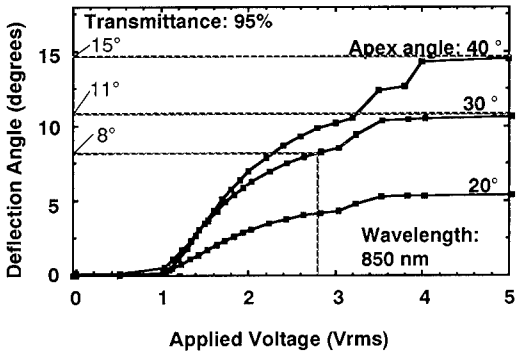


Fig. 3 Deflection angle versus applied voltage

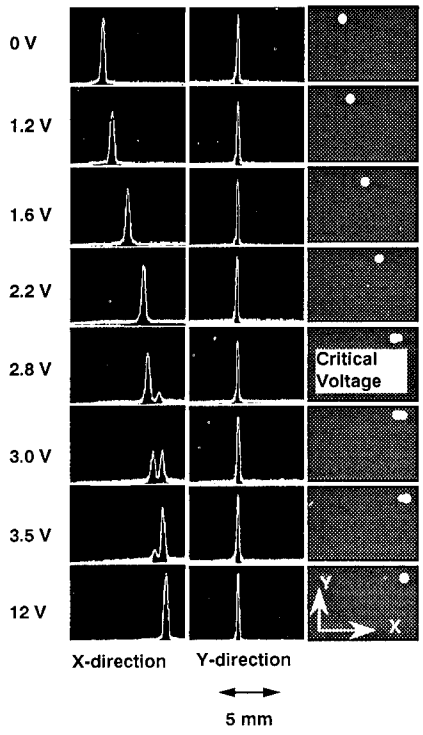


Fig. 2 Profiles of deflected beams

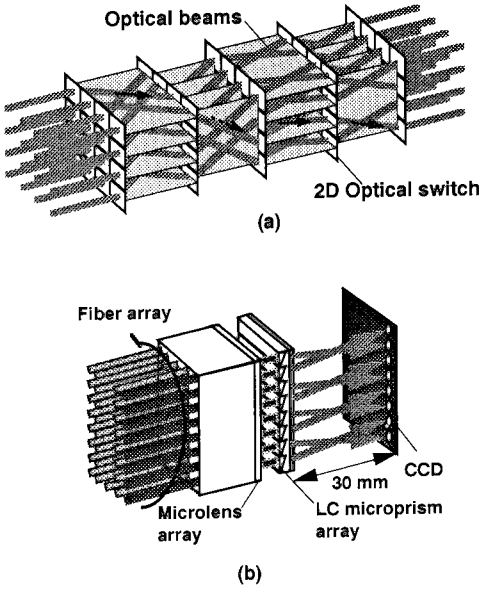


Fig. 4 (a) Optical interconnections between 2D optical switches and (b) interconnection experiment using an LC microprism array

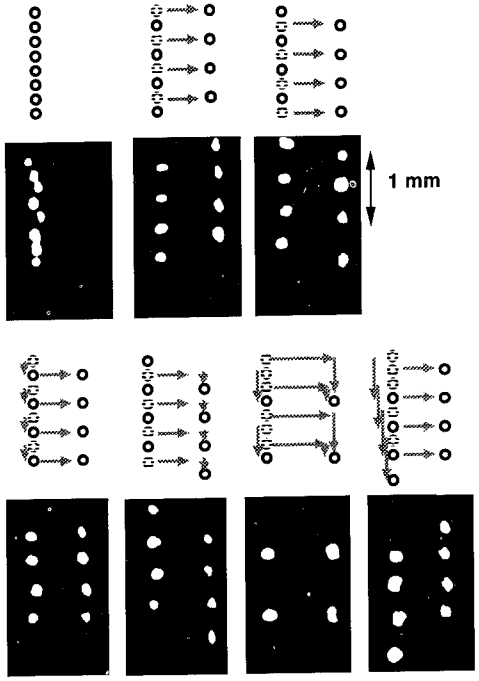


Fig. 6 Optical interconnections using cascaded two-LC microprism arrays

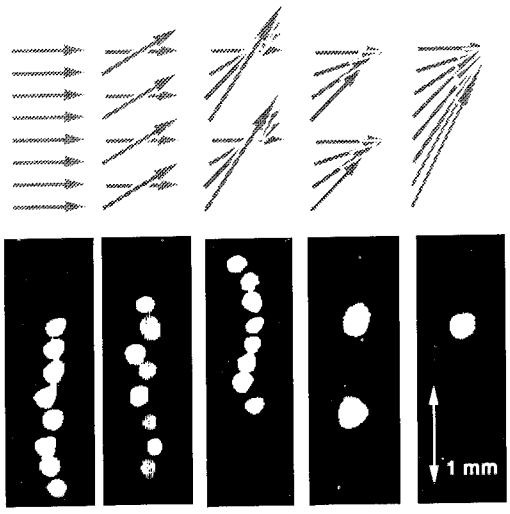


Fig. 5 Optical interconnections using a single-LC microprism array

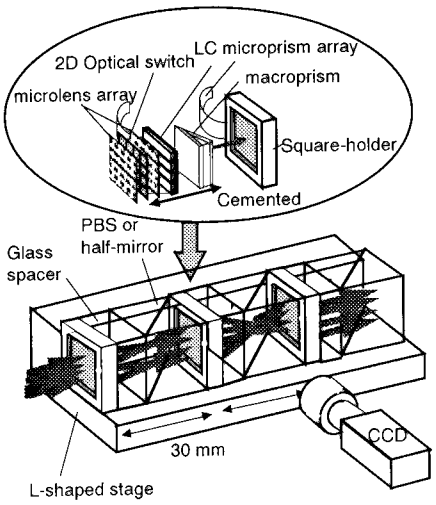


Fig. 7 Concept of a 2D optical switch network using LC microprism arrays.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Demonstration of optically controlled data switching using quantum well bistable devices and modulators

P. Koppa, P. Chavel (Institut d'Optique-CNRS, BP.147 - 91403 Orsay Cedex, France)
J.L. Oudar, R. Kuszelewicz (France Telecom - CNET , BP.107 - 92225 Bagneux Cedex, France)
J.Ph. Schnell, J.P. Pocholle (Thomson CSF. LCR, Domaine de Corbeville, 91404 Orsay, France)

Abstract
First experimental results on a 64 channel free-space photonic switching system are presented. Two control schemes are demonstrated : direct optical addressing with potential signal amplification and self-routing operation acting on data packets. A brief study on uniformity and noise effects is given, which are found to be the major limiting factors of the system.

1. Introduction

We present a 1 to 64 free-space optical switch which can be operated both in packet switching and circuit switching environments. The system associates a spatial light modulator (1) (SLM) and an optical bistable device array, both based on GaAs/GaAlAs Multiple Quantum Well (MQW) structures. The SLM is an 8x8 array of individually addressable transmission type electro-absorption modulators. The second active element, the optical bistable device (2) consists of a non-linear Fabry-Perot etalon. To obtain an 8x8 array of bistable devices, pixels are simply defined by separated incident light spots on the same cavity. When maintained in the bistable region of operation by a holding beam array, the devices exhibit a two state memory effect for about 1 ms, illustrated by points A and B in figure 1(a). With a slight modification of the detuning from the Fabry-Perot resonance, this hysteresis loop can be transformed into a simple thresholding curve (fig. 1(b)).

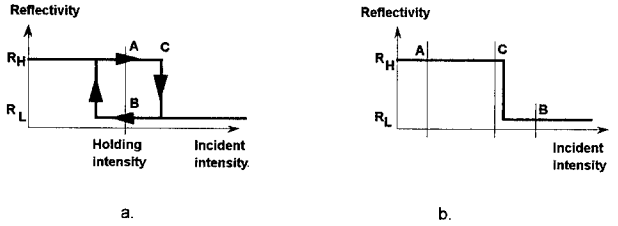


Fig. 1. : Non-linear Fabry-Perot device response curves. a): bistable, b): simple thresholding operation

2. Packet switching operation

In packet switching mode self-routing commutation is implemented by optical decoding of the header address preceding each data packet. This address decoding (3) consists of fanning-out the signal in 64 channels, comparing bit by bit the address of the packet to the address of each output channel, and cutting all channels where the identification of one or more bits fails. The following two bit coding technique is used to encode the binary addresses : each bit is followed by its complementary value, so bit one is encoded by a high-low sequence, bit zero by a low-high sequence. By the use of this code, address comparison becomes a simple

bitwise multiplication of the packet address by the inversely coded channel address, introduced as the modulated transmission of the SLM pixel. Any mismatch between the two addresses is turned into the transmission of a high level address pulse by the modulator element. This pulse then switches down the corresponding bistable element from its initial reflective state (point A) into its blocking state (point B on fig 1/a). Consequently, the only bistable element staying in the reflective state will be that of the destination channel, where address matching is perfect. After the data packet has passed through this channel, all bistable devices are reset to the initial high reflection state to receive the next packet.

Figure 2 shows output signals obtained by single channel operation of the system. The first oscilloscope trace shows a packet addressed to the observed output : after the initial synchronizing pulse, there is no high level pulse due to address mismatch, so data are transmitted by the system. The last three traces represent packets with different addresses which do not match the address of this output. We can distinguish the high level address pulse which immediately switches off the bistable device. The bistable devices used in the experiment have a contrast ratio of 2 between high and low reflection states, so theoretically blocked packets arrive at the outputs with non-negligible intensities. The total input power to the set-up is 10 mW, giving a peak intensity of 3.5 mW on the bistable element. The clock frequency of the system is fixed at 20 MHz due to our experimental means (device response times being of the order of ns). The operation of the system at address bit periods considerably shorter than 50 ns would require higher intensity levels. The data bit rate can be increased independently of device limitations because the system is entirely transparent to data. System operation is demonstrated for packet lengths up to 8 kbits with a typical address recognition rate exceeding 99 %. The system stability is limited by parasitic commutations provoked by amplitude fluctuations of the input laser beam.

We have also demonstrated simultaneous operation of three neighbouring channels of the set-up. The stability in this experiment was quite poor, correct operation could not be obtained for data string lengths exceeding 400 bits. As we show in section 4, the number of simultaneously operating channels and the error rate of address decoding is mainly limited by the non-uniformity of active and passive devices and the amplitude fluctuations arising from the laser source.

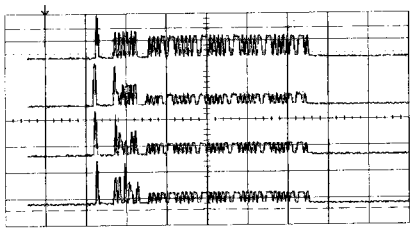


Fig. 2 : Output signals of the self-routing operation

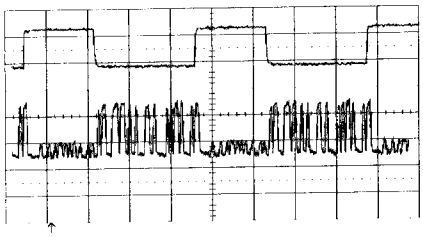


Fig. 3 : Output obtained by direct addressing

3. Switching using direct optical addressing

In circuit switching mode holding beams are directly used to control bistable device reflectivity, by setting the holding beam intensity at a lower or higher level than the switching threshold, using the non-bistable operation mode of the nonlinear Fabry-Perot cavity (fig. 1/b).

Figure 3 shows one channel operation of this scheme: when the control (or holding) beam is off (point A of figure 1/b), the bistable device is reflective, so the signal is routed to the output ; when the control beam is on, the device is in the low reflection state, so the signal

beam is mostly absorbed. In this operation mode the contrast ratio can be improved up to a value of 4, while the total input intensity can be as low as 3 mW. It means that switching threshold is about 1 mW, versus 3.5 mW in the bistable mode. In this operation scheme the control beam intensities (points A and B on Fig. 1/b) can be placed relatively far from the switching threshold. This results in an increased robustness against noises and non-uniformities if the signal beam is weak compared to the control power. Our experimental results confirm this expectation: 37 channels out of 64 could be operated simultaneously, without any observable noise-induced instability.

This operation mode can be transformed into active switching, by setting the control beam intensity to point C of figure 1/b (i.e. very close to the switching threshold). If the additional signal intensity exceeds the threshold, the device is switched by each data bit change, so modulation of the signal is transmitted onto the control beam. As the control beam is in general more intense than the signal beam, data amplification can be obtained. The actual gain depends on the device response curve and the signal to noise ratio. In our system an amplification factor of nearly two was obtained, but with other bistable devices higher gains have also been demonstrated(4). This operation scheme presents the advantages of data reshaping and amplification, but it requires data-rate switching of the bistable element, limiting the data rate by its operation speed. Rise- and fall-times of the order of 10 ns were observed for our actual devices at 1 mW threshold intensities.

4. Tolerances to noises and non-uniformities

For the simultaneous and stable operation of the channels a certain number of working conditions have to be fulfilled for all channels and during the whole operation. In packet switching mode there are three criteria: the sum of the holding beam intensity and the low level signal beam intensity must be inferior to the higher switching threshold [1], the holding beam with the signal beam at high logic level have to exceed it [2], while the single holding beam has to exceed the lower threshold [3]. These conditions can be written in the following form:

$$\alpha \cdot I_{contr} + \sigma \cdot \alpha \cdot I_{sign} < \frac{I_{th}}{\beta} \qquad [1]$$

$$\frac{I_{contr}}{\alpha} + \frac{I_{sign} \cdot C}{\sigma \cdot \alpha} > I_{th} \cdot \beta \qquad [2]$$

$$\frac{I_{contr} \cdot C}{\alpha} > \frac{I_{th} \cdot \beta}{L} \qquad [3]$$

Where I_{contr} , I_{sign} and I_{th} are respectively the mean values of the holding beam, the low level signal beam and the higher threshold intensities. The contrast ratio C and the hysteresis width L is defined such that $I_{sign} \cdot C$ is the mean high level signal intensity and I_{th}/L is the lower threshold. The Greek letters (σ, α, β) represent the magnitude of the spatial and temporal variations of these intensities, with the common definition that all changes of a parameter I with a variation γ have to lie within the interval $[I/\gamma, I\gamma]$, where $\gamma \in (\sigma, \alpha, \beta)$. So the mean intensities defined previously are simply the geometrical means between the maximal and minimal values of the given fluctuating parameter set. Variations arising from the bistable devices, the modulator array and the passive components are described by β, σ and α respectively. Introducing the ratios $G=I_{contr} / I_{sign}$ and $T=I_{th}/I_{contr}$ and finding the optimal setting for T and G, we obtain the following relationship between these parameters :

$$1 < \alpha \cdot \beta < \sqrt{\frac{L \cdot C}{C + (L-1) \cdot \sigma^2}} \qquad [3]$$

Similar expressions were derived for the two other switching schemes. Figure 4 plots the tolerance regions for the three schemes. To meet all these tolerance limits $\alpha \cdot \beta < 1.05$ and $\sigma < 1.07$ are needed which means that all spatial and temporal changes must be within ± 7 % for the SLM transmission, and typically ± 2.5 % for the bistable device thresholds and the passive components. We to note that the so called passive device variations include the amplitude fluctuations arising from the laser as well as all non-uniformities and noises arriving from devices others than the two key elements. The performance of the devices used in this first series of experiments fall completely outside of the tolerance regions, so they do not allow the simultaneous operation of all 64 channels. Non-uniformity compensation by individual tuning of the holding beam intensities can be envisioned.

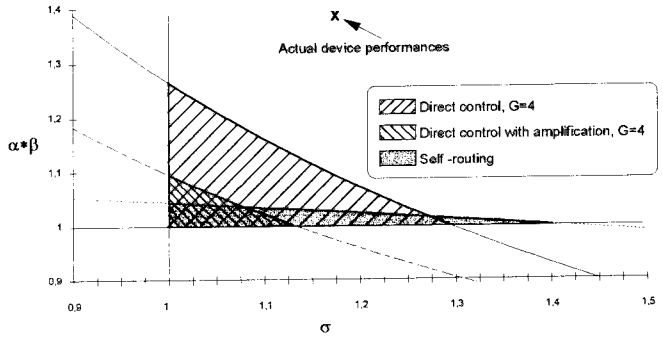


Fig. 4: Tolerance regions for the three control modes (shaded areas meet the requirements of parallel operation)

5. Conclusion

In conclusion, first experimental results on the operation of the 1 to 64 free-space optoelectronic switch was presented both in packet switching and in circuit switching mode. Uniformity and noise effects are found to be the major limitations of the system operation, hence an improvement of device characteristics and non-uniformity compensation are needed. Another challenging study will be the miniaturisation of the system.

This work is supported by the French Ministry of Research under the project MOTS/OSTI 91.

References :

1. J. Ph. Schnell, J. Raffy, J. P. Pocholle, A. Delboulbe, G. Dutrey, J. Lehoux, M. Werner, M. Papuchon, J. P. Huignard ; Technical Digest of Spatial Light Modulators 1990, Vol.14, pp 60-64
2. J. L. Oudar, R. Kuszelewicz, B. Sfez, D. Pellat, R. Azoulay; Superlattices and Microstructures, **12**, p.89 (1992)
3. P.Chavel, P.Koppa, J.Taboury, J.L.Oudar, R. Kuszelevicz, J .Ph. Schnell, J. Raffy, J.P. Pocholle, M. Papuchon ; Topical Meeting on Spatial Light Modulators, Palm Springs, California, 1993, OSA 1993 Technical Digest Series Volume 6, p.81
4. B. Sfez, J. L. Oudar , R. Kuszelewicz , J.C.Michel, R. Azoulay ; Proceedings of Nonlinear Dynamics in Optical Systems, Afton, US, June 1990

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Parallel Processing Architectures with Dynamic Optical Interconnections using Spatial Light Modulators

N. McArdle, J.F. Snowdon, and M.R. Taghizadeh

Department of Physics, Heriot-Watt University, Edinburgh, EH14 4AS, UK

Abstract. We discuss the advantages of dynamic optical interconnections in the context of parallel processing. After describing algorithms where reconfiguration is useful, we describe an optical system using a nematic liquid crystal spatial light modulator acting in a phase-only modulation mode. Binary phase structures displayed on the device provide programmable interconnection and fan-out topologies.

1. Introduction

The increasingly important roles of optical interconnections in communication and high speed processing systems is due to the many advantages of optical connections over their electronic (metallic) counterparts. These include the high degree of three-dimensional parallel connectivity, the ability of beams to cross in free-space without crosstalk, and the high bandwidth of optical communication channels.

A lens system can perform a simple 1:1 interconnection between planes of processing elements. More complex interconnection topologies can be performed by holographic elements, microlens arrays, mirrors and prisms. However, unless some mechanical system is employed the interconnection topology is fixed. In this paper we discuss some of the cost and performance benefits which may be achieved if the interconnection system is reconfigurable, i.e. when the interconnection topology can be dynamically programmed under computer control. Various devices have been utilised for such systems[1][2][3]. Here we describe a reconfigurable optical interconnection system based on a nematic liquid crystal spatial light modulator (LCSLM) which operates in a phase-only modulation mode. This device has been chosen for its high resolution, low cost, and ease of availability. By displaying binary-level phase structures on the LCSLM we can implement a variety of interconnection and fan-out topologies which can be dynamically reprogrammed.

2. Role of Reconfigurable Optics

Reconfigurable interconnections have several advantages over fixed interconnection stages. In particular they allow the maximisation of the performance gains of non-local optical interconnections over a range of algorithms inherent in practical image processing tasks. They may also be used to provide acceleration of individual algorithms by changing the processor interconnection harness during run-time of a program.

We have investigated how reconfigurable interconnects may be employed over a small range of tasks including image correlation, noise removal and enhancement, FFTs and sorting.

In this study we have used such interconnects as we have experimentally configured, including nearest neighbour, next-nearest neighbour, image dilation and various shuffle connections. In particular we have built a set of optimised tasks which may be expected to run on our own demonstrator hardware containing a reconfigurable plane. This enables us to generalise the functionality of our optical processing modules while maintaining the performance advantages of the optics, thus increasing the overall performance of the system. Numerical results obtained by simulation of our reconfigurable architecture on a distributed array processor (DAP), allow comparison with non-reconfigurable computing schemes.

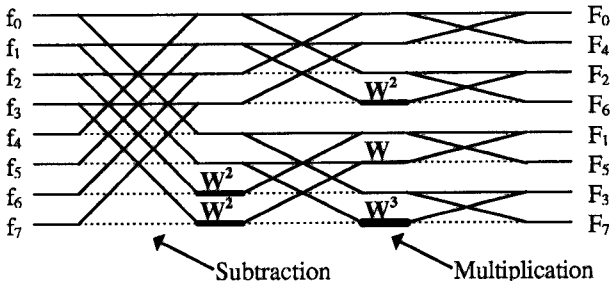


Figure 1. Radix-2 Cooley-Tukey FFT algorithm (N=8 inputs). A single reconfigurable stage could replace the fixed interconnection stages.

For example, Figure 1 shows schematically the data flow connections for a standard radix-2 Cooley-Tukey FFT algorithm. Consider only the interconnection stages labelled 1,2, and 3 which may be connecting arrays of electronic, optical or hybrid processing elements. Each stage requires a fan-out of two: one path undeviated and the other deflected by a distance depending on the depth of the stage. However, in stage 1 the top four data paths require a different connection geometry to the bottom four paths. Nevertheless it is possible to implement these connections in a space-invariant optical system by performing a fan-out of three and providing a masking function (either electronic or physical) at the following processing stage. So for N=8 inputs the three fixed interconnection stages could be replaced by a single reconfigurable stage. The hardware saving is of the order of $\log_2 N$.

3. Demonstration

The liquid crystal (LC) panel used here is one of three panels obtained from a television projection system manufactured by Seiko-Epson (VPJ2000). This panel has a larger number and smaller pixel size than that used in any previously reported experiments. The panel is a twisted nematic device designed for intensity modulation when placed between crossed polarisers. The plastic polarisers have been replaced with high quality polarisers in rotation stages. The panel has physical dimensions of 26.9×20.2 mm^2 and contains 480×440 pixels. The pixel active area is 31×31 μm^2 with a centre-to-centre spacing of 56 μm and 46 μm in the horizontal and vertical directions respectively. The thickness of the LC layer is 4.5 $\pm$ 0.5 μm and its birefringence is 0.091. The panel uses active matrix addressing and has a polycrystalline Si thin film transistor (TFT) circuit at each pixel. We have interfaced the projector to an IBM compatible computer for displaying arbitrary patterns on the panel.

Due to addressing lines between pixels, the effective clear aperture of the panel is 37%. With Fresnel reflection losses at the air/glass/liquid crystal interfaces and absorption losses in the LC layer we have measured an absolute transmission of 28%. Another source of loss is diffraction into higher orders due to the diffraction grating behaviour of the pixel structure when illuminated by coherent light. After spatial filtering, the proportion of incident power left in the zero order was approximately 10%. To overcome these losses we have investigated

the use of a microlens array (MLA) to focus the incident light through the pixel windows. An eight-level diffractive MLA attached to one of the panels increases the transmission of the panel from 28% to 51%. Even better performance is expected in future designs. The interconnection results presented here were obtained using an unmodified panel.

Several techniques have been described to obtain phase-only modulation from twisted nematic (TN) panels. These usually involve accurate setting of the bias voltage [4] or operating the panel in a double pass reflection configuration [5]. Unfortunately, these techniques are not suitable in our case and an alternative technique is necessary. The simple view of the operation TN panels where the polarisation of the incident light "follows" the twist of the LC molecules is only true in the case of thick LC layers. In general the operation is more complex and a full understanding requires detailed analysis of the panels by Jones calculus. Such analysis allows the derivation of an expression for the phase delay of the exiting light which depends on the angles of the polariser and analyser as well as the voltage applied across the panel. It is possible to find particular orientations of the polariser and analyser for which the panel acts as a phase modulator while introducing negligible intensity modulation [6]. It is in this configuration that we operate our panel.

One of the simplest interconnection patterns to implement is the fan-out of a single beam to two identical diffracted orders (1×2). This is achieved by displaying a binary stripe grating where the phase depth is alternately 0 and π radians. The interconnection distance can be varied by altering the grating period. This phase structure has a theoretical diffraction efficiency of $8/\pi^2$ (81%) into the +1 and -1 orders. Table I summarises the theoretical and experimental performance of the interconnection patterns described in this paper. If the grating phase depth is less than π radians then some power is left in the undiffracted zero order. By correctly choosing the phase depth the power in the +1, 0, and -1 orders can be equalised which provides a fan-out to three (1×3) interconnect. This is the interconnect pattern required for an optical implementation of the FFT algorithm. The necessary phase depth was calculated to be 0.64π and can be obtained by adjusting the voltage signal applied to the pixel.

A commonly used interconnection in image processing functions is the neighbourhood interconnection which connects a pixel to its neighbours either by (i) fanning out to nearest neighbours (NN) in the up, down, left and right directions, or (ii) by fanning out to the next-nearest-neighbours (NNN) in the diagonal directions. The phase structure required for the NN interconnection is obtained by crossing the phase structure for the 1×2 fan-out described above. This produces a checkerboard phase pattern which, when rotated by 45° and adjusting the grating period, provides the NNN interconnection. These structures have a theoretical diffraction efficiency of $(8/\pi^2)^2$. Measured efficiencies agree well with theoretical predictions.

Dammann gratings are commonly used to generate arrays of equally intense spots. However, higher diffraction efficiencies can be achieved by nonseparable phase structures where the phase transition points are independently optimised in the horizontal and vertical dimensions. We have generated trapezoidal stripe-geometry binary phase structures [7] to

Table I. Theoretical and experimental performance of various interconnection patterns. (NN=nearest neighbour. NNN=next nearest neighbour). Uniformity error is expressed as ratio of standard deviation to average power.

Interconnect Pattern	Theoretical Diffraction Efficiency	Measured Diffraction Efficiency	Measured Uniformity Error
1x2	81%	≈80%	6%
1x3	86%	≈81%	7.4%
1x4 NN	65.7%	≈60%	9.5%
1x4 NNN	65.7%	≈60%	7.2%
4×4	77%	≈69%	10.7%
8×8	72%	≈72%	32.9%

Figure 2. (a) Binary phase structure for non-separable trapezoidal stripe-geometry design to generate a 4×4 spot array, and (b) the spot array at the Fourier plane generated by this structure. (λ=632.8nm)

generate a variety of spot arrays including 4×4, 4×8, 2×4, and 8×8. These phase structures have theoretical diffraction efficiencies of typically >70% and the measured efficiencies agree well with predictions (Table I). Figure 2 shows an example array and the phase structure which generated it.

4. Conclusions

We have described the advantages of a reconfigurable optical interconnection system in the context of parallel optical computing. Using a commercially available nematic liquid crystal television projection panel, operating in a phase-only modulation mode, we demonstrated a variety of dynamic interconnection topologies with a reconfiguration time of 20-30 ms. These include simple fan-out in one dimension, nearest-neighbour and next-nearest-neighbour interconnections, non-local interconnections and array generation. Measured diffraction efficiencies are close to theoretical predictions. The relatively large uniformity errors on some of the arrays are due to the fact that we are displaying pixellated versions of the original phase structure designs. Improved performance would be achieved in future by using higher resolution panels. The next generation of LC panels will have pixel sizes of the order of 10 μm and an order of magnitude increase in the number of pixels. Although the panels used here are twisted devices designed for intensity modulation, we have described a technique of obtaining phase-only modulation. However, the phase structures are limited to binary levels of 0 and π radians. If untwisted (parallel aligned) devices were more readily available then multilevel phase modulation would be possible which would allow the implementation of kinoform phase structures which have much higher efficiencies and would not require the polariser and analyser components.

5. References

1. Barnes T H, Eiju T, Matsuda K, Ichikawa H, Taghizadeh M R, and Turunen J 1992 *Applied Optics* **31** 5527-5535
2. Ichikawa H, Barnes T H, Taghizadeh M R, Turunen J, Eiju T, and Matsuda K 1992 *Optics Communications* **93** 145-150
3. Kirk A, Tabata T, Ishikawa M, and Toyoda H 1993 *Optics Communications* **105** 302-308
4. Ohkubo K and Ohtsubo J 1993 *Optics Communications* **102** 116-124
5. Yamauchi M and Eiju T 1994 Optical Review to be published
6. Ichikawa H, Turunen J, Snowdon J F, Taghizadeh M R, and McArdle N 1993 *Proceedings SPIE* **1983** 491-492
7. Vasara A, Taghizadeh M R, Turunen J, Westerholm J, Noponen E, Ichikawa H, Miller J M, Jaakkola T, and Kuisma S 1992 *Applied Optics* **31** 3320-3336

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical interconnect using pixellated spatial light modulators

J.Gourlay, S.Heddle, A.O'Hara, S.Samus and D.G.Vass
Applied Optics Group, Department of Physics and Astronomy,
The University of Edinburgh, The Kings Buildings, Mayfield Road, Edinburgh EH9 3JZ
tel: 031 650 5270, fax: 031 650 5220, email: jg90@castle.ed.ac.uk

Abstract

Two approaches to optical interconnect using pixellated, electronically addressed spatial light modulators as reconfigurable optical elements are presented: one analytical described using a low resolution device; the other using computer generated holograms and a state-of-the-art planarised silicon backplane device.

1 Introduction

Reconfigurable optical interconnection has been identified as a crucial function in optical computing. One of the more attractive techniques involves the use of an electronically addressed spatial light modulator (EASLM), in a Fourier transform optical system to give holographic routing. For example, the optical information to be routed can be input to a 4-f coherent optical processor, witl the EASLM placed at the Fourier plane and the information destination is at the output plane. By the correct selection of the pattern on the EASLM, the information can be routed to the required regions in the output plane. The patterns on the EASLM are generally gratings, Dammann gratings or computer generated holograms (CGH). A very attractive generic type of EASLM is the Ferroelectric Liquid Crystal over Very Large Scale Intergration (FLC/VLSI) SLM [1]. This device has an FLC cell fabricated on top of a custom designed VLSI silicon backplane. The backplane contains an array of pixel memory elements, pixel mirrors and addressing circuitry. The controllable pixels in the SLM can modulate the relative phase by exactly 0 or π due to the switchable uniaxial nature of the FLC structure, as well as perform amplitude modulation. In EASLMs there is a trend towards more pixel complexity and more intelligent modulator arrays, which is the primary advantage of the VLSI backplane based devices. Techniques such as data compression can increase the reconfiguration time, because the main information 'bottleneck' of the EASLM is the transfer of electronic information to the modulator array from 'off-chip' sources.

Generally, pixellation is inherent in spatial light modulators and this has major consequences for their use in Fourier transform optical systems. The pixel transmission function determines the 'sinc' envelope function in the Fourier plane. A 100% fill-factor in the pixel transmission function results in no replication at the Fourier plane, only the zeroth order. Any fill-factor lower than 100% conventionally results in higher order replications. Through planarisation techniques, the pixel fill-factor can be increased towards 100%, but seperation is always required between the pixel modulation elements to avoid inter-pixel cross talk.

2 Manipulation of Replications

The pixellation can be exploited to allow routing through the manipulation of the replicated orders. In a coherent optical processor a conventional pixellated Fourier plane filter will give rise to replication of the information at the output. This is due to the output being equal to the input rotated by 180^0 convolved with the Fourier transform of the filter. Previous work [2] has shown that the delta-function-like spikes in the power spectrum of the filter which give rise to the replication may be attenuated or removed by selection of a specific pixel pitch:size ratio and the introduction of a specific pixel position randomisation scheme which applies a displacement of the pixel centre from its regular position. This generalises the expression for the power spectrum of

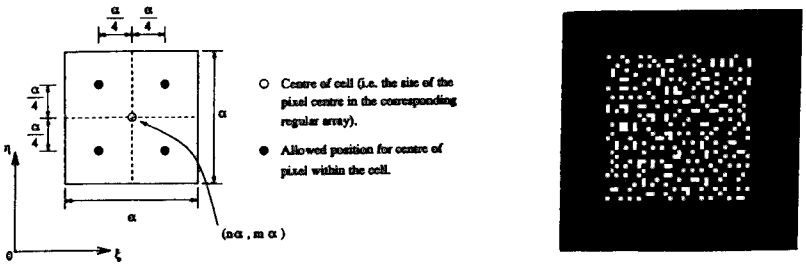


Figure 1: a) Possible pixel positions within $\alpha \times \alpha$ cell b) A resultant 16x16 array of pixels

the Fourier plane filter ($|t_2(x,y)|^2$) with all pixels 'on' to:

$$\langle |t_2(x,y)|^2 \rangle = |P(x,y)|^2 \times Q^2 \left[1 - |\rho(x)|^2 |\rho(y)|^2 + \frac{1}{Q^2} \frac{\sin^2(\pi Q \alpha x)}{\sin^2(\pi \alpha x)} \frac{\sin^2(\pi Q \alpha y)}{\sin^2(\pi \alpha y)} |\rho(x)|^2 |\rho(y)|^2 \right] \quad (1)$$

where $P(x,y)$ is the Fourier transform of the single pixel transmission function, $\rho(x)$ and $\rho(y)$ are the Fourier transforms of the pixel position probability distribution functions, α is the pitch of the underlying regular array, and the array consists of $Q \times Q$ pixels. The spectral orders are placed at $(n/\alpha, m/\alpha)$, $n, m = 0, \pm 1, \pm 2, ...$ With square pixels of side $\alpha/2$, $P(x,y) = \frac{\alpha^2}{4} sinc^2(\frac{\alpha}{2}x, \frac{\alpha}{2}y)$ which has zeroes at the positions of all the even numbered spectral orders other than the zero order. Further, choosing $\rho(x) = \cos(\frac{\pi\alpha}{2}x)$ and $\rho(y) = \cos(\frac{\pi\alpha}{2}y)$, we find that the allowed displacements of the pixel from the regular position are such that each pixel can take one of four allowed positions in an $\alpha \times \alpha$ square, (figure 1) and most importantly the odd numbered spectral orders in both x and y directions are eliminated by zeroes of $\rho(x)$ and $\rho(y)$. If however we retain periodicity in either the $x-$direction ($\rho(x) = 1$) and/or the $y-$direction ($\rho(y) = 1$), the first orders on the respective axes will remain. The power in the attenuated orders is redistributed into a diffuse background $\sim 1/Q^2$ times the zero-order intensity ($\sim 0.4/Q^2$ times the first order intensity).

The suggestion here is that a means exists whereby information can be routed via certain of the first order replicas, dependent on the distribution of the transmitting subpixel in each cell of the array- with the discretely randomised array described all first orders would be 'off'; by using arrays with periodicity with respect to the axes the first orders on the x-axis, or y-axis (or both, if the array is regular) would be 'on'. (Figure 2).

There are several points to note. First, the idea lends itself to low resolution SLMs: using the 16x16 silicon backplane SLM designed by Underwood [3], a 2×2 block of pixels could be used to code each of the α by α cells, giving an 8x8 array of cells with one pixel 'on' in each cell. This would then give a background illumination due to the 'off' state of the first order $\sim$ 26 times less than the 'on' state, a difference in states which should be simple to threshold. Further, the SLM would be capable of routing 8x8 pixel arrays when used in the Fourier plane of a 4-f optical processor, offering the benefits of parallelism and a modest degree of image fanout. These properties are scaleable for larger arrays [1], with the added benefit that the signal (intensity of the 'on' first order replicas) to noise (the diffuse background) is proportional to the inverse square of the array size Q. Also, all the 'on' first order replicas have the same intensity regardless of whether there are zero, two or four of them (to a good approximation): this would not be the case if the routing were achieved simply by writing stripe patterns on the SLM to approximate diffraction gratings, as the total number of transmitting pixels would vary depending on the desired number of replicas. Lastly, this process has been described in terms of amplitude modulating SLMs. Phase modulating SLMs are equally applicable and offer the prospect of eliminating the zero order.

3 Computer Generated Holograms

CGHs allow reconfigurable routing within the orders themselves. This is presented using a high performance silicon backplane device, which has been planarised to give a large phase-flat fill-factor. The backplane used is a 176 × 176 pixel DRAM backplane. The pixels have a 30μm pitch and the pixel mirror before planarisation is 14μm × 22μm i.e. $\approx$ 30% fill-factor. The device frame rate is 1kHz, limited by the interfacing electronics. When CGHs were displayed on the device before

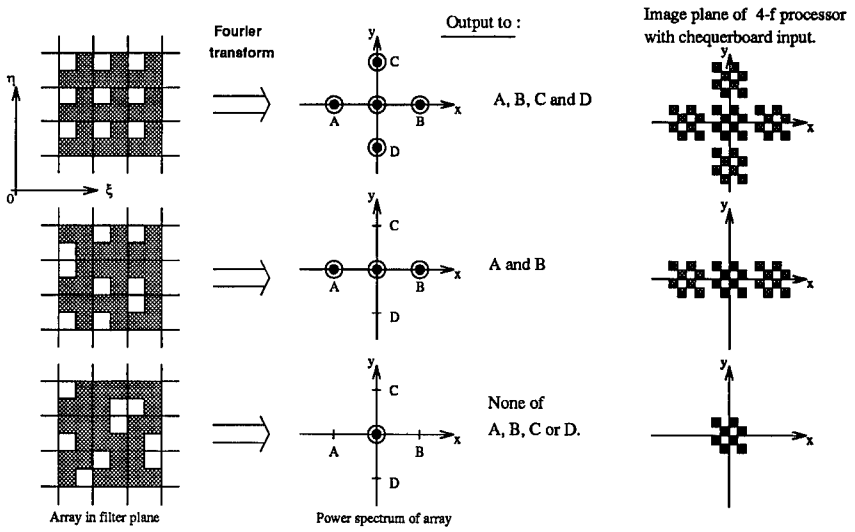


Figure 2: Selective removal of aliasing to provide selective fanout/routing

Figure 3: Diffraction from a simple grating (a) unplanarised EASLM, (b) planarised EASLM.

planarisation, poor results were obtained [4]. This was mainly due to the low pixel fill-factor, but also the poor FLC alignment caused by the low metal quality. The mirrors are fabricated using a standard VLSI aluminium metal layer, so the layer is deposited with electrical conductivity, rather than optical quality, as the main criterion. The mirrors are not optically flat and hence scatter light to give noise in the reconstruction. The rough surface profiles of the mirrors also prevent uniform high quality alignment of the FLC over the array. The low fill-factor resulted in very prominent replicated orders and a very high d.c. spot. Hence, the diffraction efficiency or the quantity of light measured where it was intended, was very low.

Planarisation techniques have been used to overcome the problems associated with the VLSI fabricated circuitry and mirrors by burying them beneath a polished dielectric film, depositing high quality metal mirrors on the surface and providing small interconnection vias through the dielectric layer between the original and new layers [5]. These techniques have been applied to the device described above and have increased the phase-flat fill-factor to 81% [6].

The resulting improvement to the performance of the device in a Fourier transform system was examined. A FLC cell was fabricated on top of a planarised device and a non-planarised device under identical conditions. A simple grating structure (with an 8 pixel pitch) was displayed on both devices and the resulting Fourier transforms are shown in figures 3(a) and 3(b). The planarised device has a very large d.c. term (central peak), very low diffraction peaks and very prominent replications (not shown). The planarised case shows considerable improvement, with the d.c. term much reduced and the diffracted peaks the most prominent.

The display of CGHs on the planarised FLC/VLSI EASLM can produce interconnection or fanout. The CGHs have been designed by simulated annealing to give arrays of spots from a plane

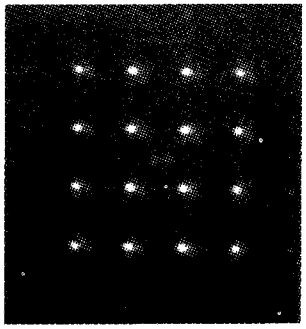


Figure 4: A 4 × 4 fanout pattern produced by a CGH on the EASLM.

wavefront [4]. In the 4-f coherent processor, the input information would be routed to the position of the spots in the output plane. By changing the CGH pattern on the EASLM, the position and number of the spots can be altered and hence route the input information as required by the system. A CGH pattern, designed through simulated annealing to produce a 4 × 4 fanout, was displayed on the EASLM. The resulting fanout obtained after Fourier transformation by a simple lens (f=140mm) is shown in figure 4. The d.c. spot is very low, and the diffraction efficiency (**light in the 16 peaks/total light in the Fourier plane**) is ≈ 65%. Replicated orders are still evident as the pixel function has < 100% fill-factor, but they are very low as the efficiency shows.

The VLSI/SLM is a very attractive device for this application, due to the potential level of addressing complexity available, which allows very high reconfiguration times. Future devices will have smaller pixel pitches (more compact systems), a higher space-bandwidth product (a higher number of pixels) and improved optical quality. There is also the potential for storage of sets of CGH patterns 'on-chip', i.e. designed into the backplane for a specific system requirement, which would increase considerably the reconfiguration time, as only 'local' data transfer need be performed

4 Conclusions

Both of the techniques are complementary, allowing the manipulation of the replicated orders, and of information inside the orders. The pixellation of the EASLM has been shown to be a very important parameter in terms of both the performance and capabilities. Control of the pixellation is crucial in future devices for high performance in optical computing systems. These conclusions are being considered as the next generation of EASLM devices are being developed.

1. K.M.Johnston, D.J.McKnight and I.Underwood, IEEE Journal of Quantum Electronics, 29, pp699-713, (1993).

2. S.Heddle and R.M.Sillitto, J. Mod. Opt. vol. 40(2), pp299-313 (1993).

3. I.Underwood, D.G.Vass and R.M.Sillito, IEE Proceedings, 133, Pt J, pp77-82, (1986).

4. J.Gourlay, S.Samus, P.McOwan, D.G.Vass, I.Underwood and M.Worboys, Accepted for Appl. Opt.

5. A.O'Hara, J.R.Hannah, I.Underwood, D.G.Vass and R.J.Holwill, Appl. Opt., 32, pp5549-5556, (1993)

6. I.Underwood, D.G.Vass, A.O'Hara, D.C.Burns, P.W.McOwan and J.Gourlay, Appl. Opt., 33, pp2768-2774 (1994)

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

An Optical Package for a Reconfigurable Free—Space Optical Interconnection Network For Multiprocessor Systems

Mingcui Cao, Hongpu Li, Xiangjun Zhao, Fengguang Luo, Xu Jun

(National Lab. of Laser Technology, Huazhong University of Scirnce and Technology, Wuhan, 430074 P. R. China)

Ronghan Wu, Wenzi Gao

(Nation OEIC Lab. ,Institute of Semiconductors, Chinese Academy of Science, Beijing, 100083 P. R. China)

Abstract. An optical package for a reconfigurable free—space optical interconnection network based on a novel (2,2,2) switching node for the multiprocessor system is presented. A novel photo—electronic hybrid switching node which is consist of SEED and S—SEED devices is described. The optical package of one stage in a reconfigurable interconnection network is demonstrated.

1. Introduction

The interest has recently grown in reconfigurable optical interconnects and optical switching networks for multiprocessor and telecommunication systems [1-3].

In this paper, we describe an optical package of reconfigurable free-space optical crossover interconnection network for multiprocessor systems. The link interinterconnection stages are implemented by some monolithic integrated free-space digital optics. A novel photo-electronic hybrid node type of (2,2,2) is used for a switching node, in which the hardware comprises two monolithic self-electro-optic-effect devices, one electronic addressing (SEED) and one S-SEED. The optical package of one interconnection stage in reconfigurable interconnection network is demonstrated. Some preliminary results for the functionalities of the switching network are reported.

2. The architecture of optical interconnection network package

A free-space multistage interconnection network consists of the node stages and link stages. In our experiment system, an opto-electronic hybrid (2,2,2) node with electronical signal control as shown in Fig. 1 is presented for the switching node of the interconnection network; its optical hardwares consist of two electronically addressing SEEDs and one S-SEED. The electronic addressing signals that must be sent from the controlling processor of the network are offered for controll signals of data

paths, one SEED controlls straight paths, another SEED controlls exchange paths. The S-SEED device operating in a state of OR logic is used as cascading.

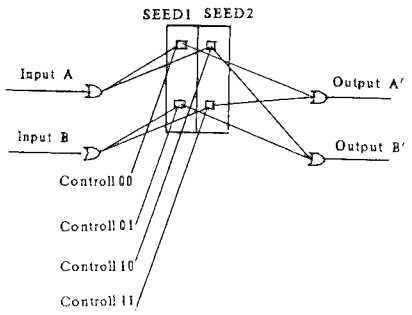


Fig. 1, The opto-electronic hybrid (2,2,2) switching node.

An optical crossover network based on photo-electronic hybrid (2,2,2) switching nodes is presented as a reconfigurable optical interconnection network as shown in Fig. 2. A reconfigurable optical switching network with N ports must be a non-blocking switching network system that comprises of 2 Log_2N - 1 stages, in which optical systems in each stage have the same hardware components except for different periods of prismatic mirror arrays. The optical package of each stage as shown in Fig. 3 consists of phase fresnel lenslet arrays (F), a Prismatic grating mirror array (PG), the beam splitters (BS), the quarter waveplates, SEED devices, and S-SEED devices. The optical system is based on the principle of focal plane imaging. The phase fresnel lenslet array which splits one collimated laser beam with homogeneous power distribution in some region to a light spot array on the focal plane, also is used for imaging lenslet array which images the light spot on each mesa of SEED in each channel on each window of S-SEED substrates, or vice versa. The optical system of the switching network shown in Fig. 3 consists of three parts. The first block which includes BS_1,F_1, PG and M is called the optical link stage. The two connection types

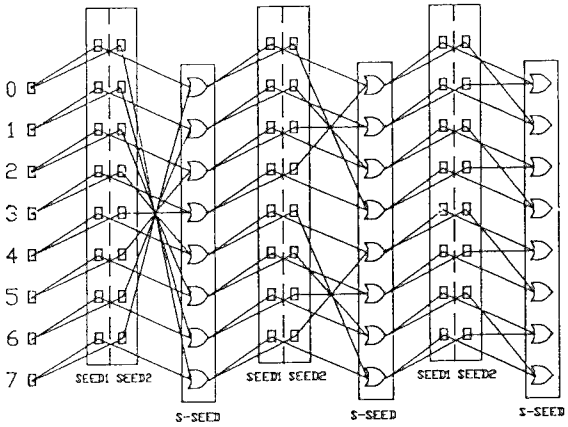


Fig. 2 An optical crossover interconnection network based on photo -electronic hybrid (2,2,2) switching node

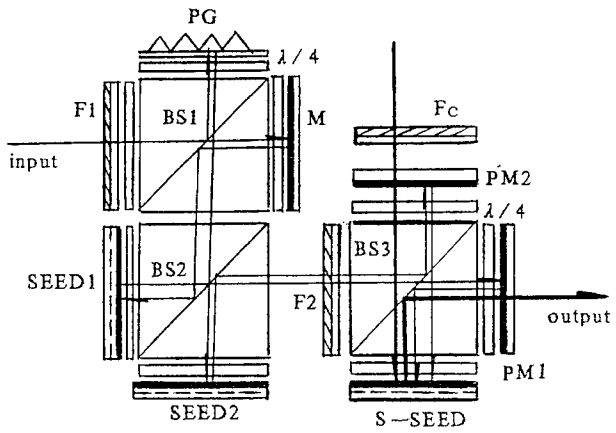


Fig. 3 The layout of the optical package of free-space crossover interconnection network

of the straight and the exchange in the crossover network are implemented in the first block. The second block which consists of BS_2, $SEED_1$ and $SEED_2$ is called the control stage of routing. Whether the straight or the exchange channels would be routed from the input to the output port at a time is determined by applied electronic signals on the pixels of $SEED_1$ and $SEED_2$. The third block includes BS_3, F_2, F_3, PM_1, PM_2, and S-SEED is called the cascading stage. The S-SEED device operates in an optical set-reset-latch that can be made to simulate an OR logic gate. The input signals are amplified and routed to the next stage. As viewed from the operation function, the second and third blocks are the optical hardwares of the optical switching nodes as shown in Fig. 1.

The optical network has two connection typies of the straight and the exchange. The straight connections are implemented by the route of F_1-BS_1-M-BS_2-$SEED_1$-F_2-PM_1-BS_3-S-SEED. The exchang connections are implemented by the route of F_1-BS_1-PG-BS_2-$SEED_2$-F_2-BS_3-PM_2-S-SEED. The 4-level phase fresnel lenslet arrays F_1, and F_2 are imaging lenslet arrays. F_1 images the light spots from each mesa of S-SEED device in the previous stage on each mesa of $SEED_1$ and $SEED_2$ substrates. The F_2 images the pattern of $SEED_1$ and $SEED_2$ substrates on the windows of the S-SEED. The function of 4-level phase fresnel lenslet array F_c is the generation of light spot array. It splits a collimated laser beam with homogeneous power distribution in some region into light spot array on the focal plane which is located at the window plane of S-SEED. The light spot array is called the clock light array to read out the states S-SEED to the next stage. In the optical interconnection network of 8 x 8, F_1 and F_2 are phase fresnel lenslet array of 8 x 8, F_c is a lenslet array of 8 x 16. It is obvious that route of seleetion is controlled with electronic addressing signals applied on the pixels of $SEED_1$ and $SEED_2$. Then the data signals are transfered to S-SEED device, read out by the clock light array and are amplified to the next stage. The eletronic addressing signals come from personal computer which perform path hunting. Based on the architecture detailed above, we fabricated the optical package of one stage in a reconfigurable free-space optical switching network of 8 x 8. The space between adjacent channels is 0.4 mm. All 8 x 8 channels are arranged in the areas of 3.2 x 3.2 mm^2. The insertion losses of the optical switch device SEED is about 40 %. The switch time of SEED is about 10 ns when device are operated in

normal-on state. When the S-SEED device was replaced by the testing array whose response time is about 2 ns, some basic operation performances of the optical switching node (2,2,2) and optical interconnection network were demonstrated. Data packets which include 8 bits of data signals come from a personal computer are parallelly trasported through the optical package of interconnection network from the input ports to the output ports.

3. Conclusion and discussion

The optical systems based on the principle of focal plane imaging in which a phase fresnel lenslet array is used as the imaging lenslet array, this optical system has the following characteristics, the focal longth of phase fresnel lonslet array can be made very short, the design of optical system is very compact, in which all optical components can be fixed by optical glue to form a package architecture. By comparison with an infinite conjugate, telecentric, 4f imaging system, this architecture avoids aberration accumulation and alignment difficulty. Thereby the architecture of the optical package has some features, such as small physical size, more tolerance of alignment, shorter alignment time, high stability, and high reliable. The optical power loss, the density and the amount of optical interconnection channels are perhaps the major limitations of this optical package architecture. There are several ways for decreasing the optical power loss in the system,decreasing the loss in the SEED, increasing the high state reflectivity of the SEED, increasing the precision of photoolithographic and processing. It is a important issue in the current system that how to get a collimated laser beam with homogeneous power distribution in the required size for clock beam. The best way is that the vertical cavity surface emitting lasers is used as the reading beams, because the surface emitting microlaser not only can provide high power, but also can be easy to form a collimated and homogeneous laser beam.

We gratefully acknowledge the support of National Deference Science Foundation and the National 863 High Technology Program.

Reference

[1] T. Sakano, et al, 1991, Applied Optics, Vol. 30, No. 17, PP. 2334-2343.
[2] H. S. Hinton and D. A. B. Miller, 1992, AT & T Technical Journal PP. 84-92.
[3] F. B. McCormick, et al, 1993, Applied Optics, Vol. 32, No. 26,pp. 5153-5171.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A distributed, reconfigurable free-space optical interconnection network for massively parallel processing architectures

Ian Redmond and Eugen Schenfeld

NEC Research Institute, 4 Independence Way, Princeton, NJ 08540, USA

Abstract. We present a free-space network architecture, supporting communication rates of greater than 1 Gbit/Sec per port. Using today's technology, our network scales to thousands of ports and is useful for massively parallel processing architectures. A 64-ports prototype was demonstrated.

1. Introduction

The interconnection network plays an important role in the overall performance of parallel processing systems. Optics has been suggested as a potentially superior technology over electronic, for interconnecting thousands of processing elements (PEs) in a massively parallel processing system (MPP). In an MPP system, a PE may be built around a state-of-the-art, off-the-shelf processor (Digital "Alpha", MIPS R4400, etc.). To serve the interconnection needs of today's and near future processors ("Intel 2000 CPU"), that are being and will be used for building MPP systems, a high performance network is needed. Such processors, having a performance rate of one to two thousands MIPS (million instructions per second) on words of information that are 64 to 128 bits long, generate a total throughput of 128 to 256 Gbit/Sec. Not all of this traffic needs to be handled by the network. A fraction of this (e.g., 10 GBit/Sec per PE) may cause difficulties in large systems with thousands of PEs.

A network can be characterized by several parameters: Number of ports; bandwidth capacity per port; the network's structure; latency delays: from one port to another, without arbitration (i.e., circuit switching or reconfigurable mode of operation); and routing and switching delays (caused by arbitration and other traffic specific issues, such as blocking, and also depends on the network structure).

It is generally accepted that for large networks (thousands of ports), optics is better than electronics in providing high bandwidth connections. However, it is also commonly accepted that today, optics cannot compete with electronics for general purpose processing because of higher cost and the immaturity of the optics technology.

2. A reconfigurable free-space network

Our proposed optical network is based on several basic principles and observations on the needs and possible tradeoffs of an optical MPP network:

1. "Logic-Less" optical operation: Optical technology is used for point-to-point high bandwidth connectivity and not to do dynamic routing or arbitration.

2. Distributed Layout and Control: No centralization in terms of one common element (e.g. a lens) or active device (e.g., one array of VCSELs). No central arbitration and control is needed. The optical network scales in proportion to the number of PEs (although for practical reasons, one would choose to use fewer, but larger active devices such as VCSEL arrays).

3. Reconfiguration is not needed very often as many parallel applications exhibit switching locality: i.e., a PE will only need to communicate with a small number of other PEs for a period of time. Choosing between these PEs can be done by the electronic switches that acts as "interconnection caches" to the initial reconfiguration of the optical network.

3. Folded Clos network

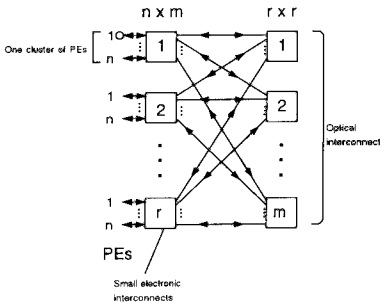


Figure 1. A folded Clos.

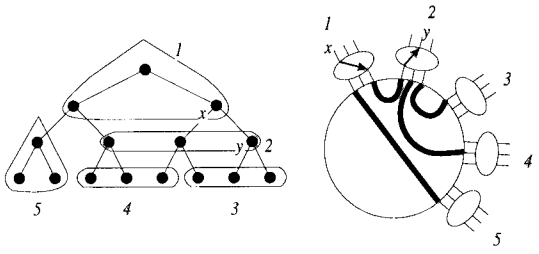


Figure 2. Embedding a binary tree.

Figure 1 presents the structure of our network. The middle stage is the optical network, made of "logic-less" switches (i.e., it is a circuit-switching operation) using VCSEL arrays. The first and third stages (folded into one) are made of small electronic crossbar switches.

Such a network can be used to map parallel applications that have switching locality. We have investigated the mapping issue previously and come to the conclusion that many classical communication patterns (such as trees, 2-D and 3-D mashes, hypercubes, pyramids, mesh of trees etc.), as well as non-symmetrical communication patterns, can be well embedded in our structure [1, 2].

4. Mapping and embedding

Figure 2 is an example of the embedding of a tree into our network. Note that the clustering means that nodes in the same cluster will be mapped to boards that share a common small electronic crossbar switch. The connections between clusters are made by reconfiguration of the optical middle layer of the Clos network.

5. System overview

Figure 3 represents an overall view of our system with a distributed optical network. PEs are grouped into boards, each having a fast electronic crossbar switch. Boards are arranged into columns that are connected using optical free-space multiple buses. Each board can have optical connections to few other boards (8 to 16). The selection of these boards is done by reconfiguration of the optical interconnections. This mode of operation is useful in many parallel applications that exhibit switching locality or/and have "phases" of computation with specific interconnection topologies. An example of such applications using the reconfigurable network can be found in [3, 4].

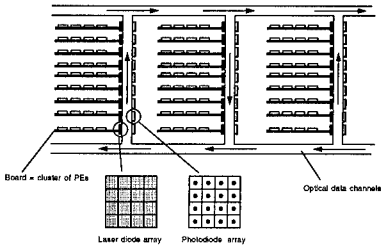


Figure 3. Optical network.

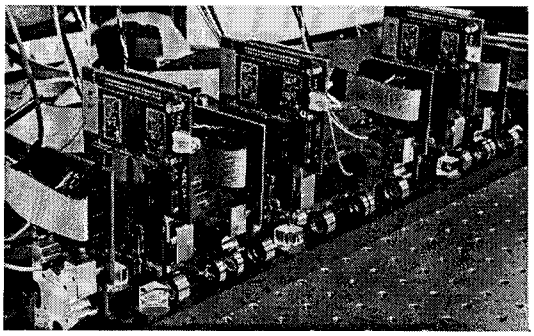


Figure 4. 64-ports prototype.

In our experimental system, we used VCSEL arrays that are arranged as 8×8 devices with individual access to each VCSEL. Such individual access is not needed for our architecture, but these devices were available off the shelf. However, by having access to individual VCSELs, it was possible to use only four 8×8 devices, each subdivided into 16 separate zones each having 2×2 devices. Thus each PE can reconfigure its subdivision to connect to one of 3 other boards. Hence the system simulated 4 boards of 16 PEs each.

Figure 4 is a picture of a 64-ports prototype we have built based on the above network architecture. A detailed description of this prototype and its characteristics appears in another paper in the proceedings [5].

6. Scalability and future issues

Future systems may have larger numbers of PEs. One critical element is the size and the electrical connections for the VCSEL arrays to be used. Our prototype is an example

of the foreseen maximum needed (i.e., 16 connections per VCSEL array). We anticipate fewer electrical connections in future VCSELS arrays. Such devices can be made as an X-Y matrix or other structures to control each sub-division on the array for each PE. Thus individual VCSELs can be selected to work and transmit data. By selecting a certain VCSEL, the connection to that corresponding Board, can be made by the optical layer of the Clos network. The number of high bandwidth connections that need to be made to the package of such an array, will be limited to at most 16. In addition, low speed connections used for choosing the appropriate VCSEL in each sub-division are needed. These are low speed connections since the reconfiguration is expected not to be done very often. One important issue to address is that of the power budget. Currently, we have employed variable reflectivity mirror elements to maximize the power available. In addition to this, it may be necessary to add amplification or regeneration of the optical data, in certain location, for a high bandwidth and low noise communication. The performance of such a system was found to be close to an "ideal" crossbar (i.e., one that has the same communication delays while it scales to large sizes) [6].

7. Summary

We have presented the system structure of a prototype for an optical interconnection network to be used for MPP architectures. Our system benefits from the high connectivity and throughput optics can offer while limiting the processing needed for routing to electronic technology. Many parallel applications can use such a reconfigurable network as they naturally exhibit switching locality in their communication needs. We hope to extend our system in the future to accommodate thousands of PEs in a multi-Gigabit/Sec channels.

References

[1] Y.-D. Lyuu and E. Schenfeld, *"Parallel Graph Contraction with Applications to a Reconfigurable Parallel Architecture"*, ICPP, St. Charles, IL, August 15-19, 1994.

[2] V. Gupta and E. Schenfeld, *"Combining Message Switching and Reconfiguration in the Interconnection Cached Network"*, ISPAN'94, December 14-16, 1994, Kanazawa, Japan.

[3] Y.-D. Lyuu and E. Schenfeld, *"Total Exchange on a Reconfigurable Architecture"*, SPDP, pp. 2-11, Irving, Texas, December 1-4, 1993.

[4] Y.-D. Lyuu and E. Schenfeld, *"New Algorithms for Matrix Operations with Applications to a Reconfigurable Parallel Architecture"*, PDCS, October 6 - 8, 1994, Las Vegas, Nevada

[5] I. Redmond and E. Schenfeld, *"Experimental Results of a 64 Channel, Free-Space Optical Interconnection Network for Massively Parallel Processing"*, OC'94, Edinburgh, Scotland, August 22-25, 1994 (this proc.).

[6] V. Gupta and E. Schenfeld, *"Performance Analysis of a Synchronous, Circuit-Switched Interconnection Cached Network"*, ICS'94, Manchester, U.K., pp. 246-255, July 11-15, 1994.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Reconfigurable Optical Interconnections in $Bi_{12}TiO_{20}$ Photorefractive Fibre

A A Kamshilin and T Jaaskelainen

University of Joensuu, Väisälä Laboratory, Department of Physics, P. O. Box 111, SF-80101, Joensuu, FINLAND

Abstract. Reconfigurable optical interconnections by using double phase conjugate mirror in photorefractive $Bi_{12}TiO_{20}$ fibre are demonstrated. Recyclable phase conjugate mirror with conversion efficiency up to 8% has been recorded in the fibre in a wide range of incidence angles of the pumps at λ = 632.8 nm. Practically it means that no adjustment is needed to create interconnection by using mutually incoherent beams, except introducing the pump beam into the fibre. The response time was measured to be 5 seconds for pump beam intensity of 1 mW/mm^2.

1. Introduction

Photorefractive crystals (PRC) have been proved to be significant in parallel optical information processing owing to their real-time, high density, recyclable volume holographic recording and phase conjugation. One of the most interesting and promising effects recently discovered in PRC is double phase conjugation of mutually incoherent pump beams. For the PRC illuminated on opposite faces by two pumps, two independent counterpropagating beams build up, originating from the beam fanning. They give rise to a common grating, which is reinforced by positive feedback and finally results in the presence of a pair of phase conjugate beams through a cross-readout process [1]. This double phase conjugate mirror (DPCM) can be used as reconfigurable interconnections between two sets of fibres [2].

Experimental studying of the DPCM has been concentrated mainly on the research of bulk crystals. However, recently, the double phase conjugation in a fibre sample has been demonstrated in our laboratory [3,4]. Compared with bulk crystals, fibres are much easier in production owing to the Laser Heated Pedestal Growth (LHPG) technique, which allows to grow a single crystal fibre of better quality and much faster. Moreover, the physical mechanism of beam coupling in the fibre sample differs from that in widely studied photorefractive bulk crystal. In this presentation, the advantages of the $Bi_{12}TiO_{20}$ fibre application to reconfigural optical interconnections are demonstrated.

2. Experimental configuration

2.1. Optical scheme for interconnection

As an example, interconnection between one fibre and a set of three fibres is shown on the Fig.1. Light emitted from the fibre (Input 1) is focused by lens 1 on the PRC's face. At the same time, the light emitted from the set of fibres (Inputs 2-4) is focused by lens 2 on the opposite side. After a response time τ, common holographic grating is recorded in the PRC, producing a pair of phase conjugate beams, which are automatically introduced into the fibre from one side and into the set of fibres from the other side. Phase conjugate beams can be extracted by using beamsplitters BS. Therefore, the signal from the Input 1 achieves all the Receivers R2-R4 and the sum of the signals from the Inputs 2-4 achieves Receiver R1. In

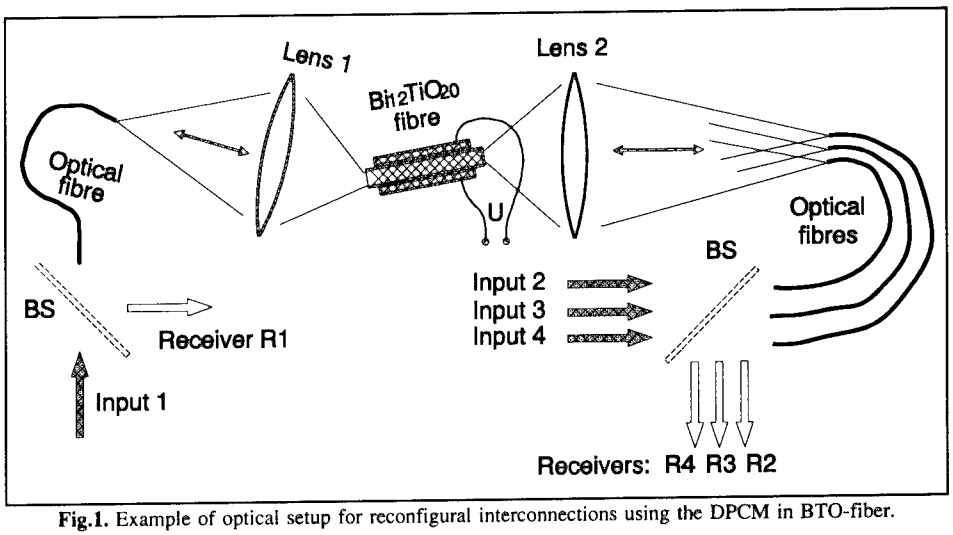


Fig.1. Example of optical setup for reconfigural interconnections using the DPCM in BTO-fiber.

this example the Input 1 is interconnected to all the three Inputs 2-4. Interconnections can be simply changed by introduction of light only into two fibres (3 and 4, for example). In that case after the response time τ, a new common holographic grating will be recorded in the PRC and new interconnections between Input 1 and Inputs 3,4 will be established.

The performance of reconfigural interconnection depends on three most important parameters of the photorefractive DPCM. They are the efficiency of phase conjugation, the response time, and the spatial bandwidth. All these parameters have been measured in our experiments.

2.2. Photorefractive fibres

Experiments were carried out on fiber-like $Bi_{12}TiO_{20}$ (BTO) samples. This crystal belongs to the same structural class (sillenite) as $Bi_{12}SiO_{20}$. However, BTO has larger electro-optic coefficient and lower optical activity compared with other sillenites, and is more sensitive compared with $BaTiO_3$ and SBN photorefractive crystals. The fibres were cut out from the bulk crystal along the [110] crystallographic axis and were glued between two electrodes. Their length is 8-18 mm, and they have rectangular cross-section with approximately 1 mm size. The alternating bipolar electric field of square-wave form with pulse duration of 14 ms and with front duration of 2.5 ms could be applied to the crystal so that either the plane {110} or plane {111} would be perpendicular to the electric field vector. Later orientation is known to produce in BSO-type crystals the highest energy exchange effect along the applied field vector [5]. The electrodes were made 1.2 mm shorter than the sample length and covered completely with epoxy glue to prevent electric breakdowns via air and crystal's surfaces. Therefore, as high voltage as 5 kV (which corresponds to a 50 kV/cm bias electric field) could be applied to the fibre.

It is known [6] that the hologram recording in the BSO-type photorefractive crystals under ac electric field results in the formation of the refractive index grating, which is 90° phase shifted in respect to the illuminating interference pattern. This condition is essential to produce strong energy exchange between the pump beam and the scattered light, resulting

on the appearance of strong fanning light, which is the origin of double phase conjugation.

Four longer faces of the fibre were optically polished, but end-faces were slightly grounded to produce a scattering of transmitted beam into the angle of approximately 4 degrees. It was especially made to introduce pump beams of complex (speckle-like) wavefront into the fibre, which allowed us to avoid conical degeneracies in the Bragg matching angle on a given volume grating. Two independent linearly polarized helium-neon lasers (λ = 632.8 nm) were used to pump the BTO-fibre. The input polarization angles were chosen experimentally to produce as high phase conjugate reflectivity as possible.

3. Experimental results

3.1. Conversion efficiency

It is very easy to record the DPCM in photorefractive fibre. To this aim it is enough to introduce into the fibre from its opposite end faces two mutually incoherent (derived from independent lasers) pump beams. We have found that both types of fibres (with external field being parallel to <110> and <111> axis) possess high enough conversion efficiency to produce phase conjugation of complex images [4]. The spatial resolution of phase conjugated images as estimated from the experimental photographs is approximately 50 lp/mm. The experiments with phase conjugation of images have been made with <110> orientated fibre.

To the other hand, we have found that the DPCM recorded in BTO fibre of <111> orientation has much wider band of pump incidence angles than it was in our previous works [4,7]. Experimental results are shown in Fig.2, where a conversion efficiency in steady state, η is plotted in form of a contour map as a function of incidence angles of the pump #1 (α_1) and the pump #2 (α_2). Here we define η as S_1/P_2, where S_1 is the power of the beam, which

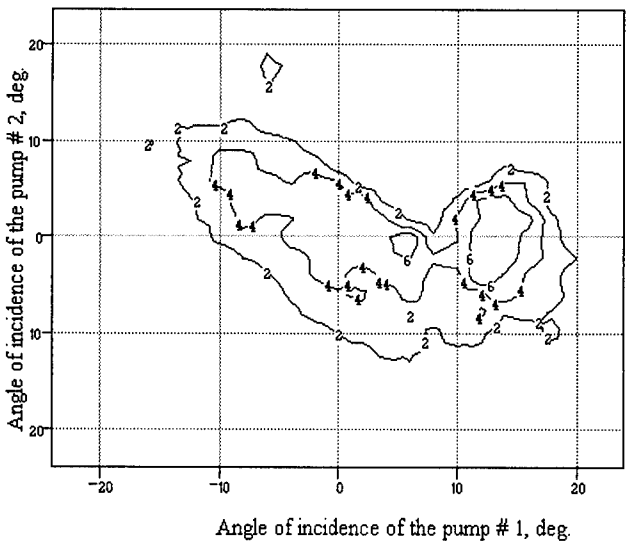


Fig.2. The DPCM conversion efficiency as a function of pumps incidence.

is a phase conjugate replica of the pump #1 and $\mathbf{P_2}$ is the input power of the pump #2. We should emphasize that our definition of η does not take into account the high losses of light power (20% for each reflection), owing to high refractive index of BTO crystal ($\mathbf{n}$ = 2.58 at λ = 632.8 nm). Therefore, antireflection layers evaporated on the fibre end-faces can significantly improve the efficiency of the DPCM. As one can see from Fig.2, the DPCM with efficiency more than 2% can be obtained while changing the pump #1 incidence from -15° to 18° and the pump #2 incidence from -12° to 12°. The numbers on the lines of Fig.2 show the efficiency in per cent. The maximal conversion efficiency is 8%. Practically it means that no adjustment is needed to get an effective interaction of mutually incoherent beams, except introducing the pump beam into the BTO fibre.

3.2. Response time

In spite of rather uniform steady-state value of the DPCM conversion efficiency, its temporal behaviour varies significantly with incidence angle α_1 and α_2 of both pumps. Usually, the response time is inversely proportional to the intensity of pump beam, but it was found that it also depends on the power ratio of pump beams [7], approaching a minimal value for equal intensity of pump beams. It takes 5 sec to reach the maximum of conversion efficiency with pump beam intensity of 1 mW/mm^2.

4. Conclusion

The application of $Bi_{12}TiO_{20}$ photorefractive fibre to reconfigural optical interconnections has been proposed. The BTO fibres can be assembled in a variety of configurations, isolating each stack of interconnections from other stack in adjacent fibres. Moreover, they can be easily integrated with semiconductor lasers. This can provide realization of great amount of interconnected fibres with a possibility of independent reconfiguration of different groups of them.

References

[1] Yeh P, Chang T Y and Ewbank M D 1988 *J. Opt. Soc. Am. B* **5** 1743-50
[2] Weiss S, Segev M, Sternklar S and Fischer B 1988 *Appl. Opt.* **27** 3422-8
[3] Kamshilin A A, Ravatinen R, Tuovinen H, Jaaskelainen T and Prokofiev V V 1993 *Opt.Commun.* **103** 221-6
[4] Kamshilin A A, Liu S, Tuovinen H, Prokofiev V V and Jaaskelainen T 1994 *Opt.Lett.* **19** 907-9
[5] Stepanov S I, Shandarov S M, and Hat'kov N D 1987 *Sov. Phys. Solid State* **29** 1754-6
[6] Stepanov S I and Petrov M P 1985 *Opt. Commun.* **53** 292-5
[7] Kamshilin A A, Tuovinen H, Prokofiev V V and Jaaskelainen T 1994 *Opt.Commun.* **109** 312-7

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Design of a Lenslet Array Based Free-Space Optical Backplane Demonstrator

B. Robertson, G. C. Boisset, H. S. Hinton, Y. S. Liu, N. H. Kim, M. R. Otazo, D. Pavlasek, D. V. Plant, D. Rolston.*

Department of Electrical Engineering, McGill University, 3480 University street, Montréal, Québec, Canada H3A 2A7 (Tel 514-398-4470).
*Now with University of Colorado at Boulder

W. M. Robertson

National Research Council, Institute for Information Technology, Montreal Road, Ottawa, Ontario, Canada K1A ORG (Tel 613-991-6839).

Abstract

The optical and optomechanical design of a representative portion of a lenslet array based free-space photonic backplane is described. Issues relating to the optical design, optomechanical layout and alignment tolerances will be discussed.

Introduction

The following paper will consider the optical and optomechanical design of a unidirectional lenslet array based PCB-to-PCB interconnection link. The system uses FET-SEED modulators and receivers to transmit data between two circuit boards and multi-level diffractive lenslets to interconnect the device planes. Differential logic is used to minimize the sensitivity of the system to optical power supply variations, thus each data channel requires two optical signal beams to encode the information. The design of the FET-SEED arrays and the performance of the backplane demonstrator are described elsewhere in these proceedings[1]. The aim of this work is to develop the technology required to implement a reconfigurable terabit capacity photonic backplane [2].

Optical Interconnect Design

Integrating a free-space photonic backplane into a commercial electronics environment will be a challenging task. Current electronic backplane standards typically demand a board-to-board separation of the order of one inch. Any competitive photonic backplane must therefore be compact enough to fit within this space. The optics must be designed such that it is possible to remove and insert circuit boards while maintaining precise registration between the optical signal beams and the transmitter and receiver arrays. In addition, alignment should be unaffected by temperature variations and external vibrations. This places exacting demands on the optical and optomechanical design of the backplane.

Figure 1 shows the optical layout of the photonic backplane demonstration circuit. Light from a Ti-Sapphire laser operating at 850nm enters the circuit via a single mode polarization preserving fiber. The light is collimated and passes though a non-separable binary-phase grating which generates an array of optical beams at the input to the optical interconnection assembly. The beams are collimated by lenslet array L_1, reflected by the polarizing beam splitter (PBS) and focused by lenslet array L_2 onto the array of FET-SEED quantum well modulators (figure 2). The intensity modulated signal beams pass back through the PBS (the polarization of the light having been rotated by a quarter waveplate) and are finally focused by lenslet array L_3 onto the array of receivers. The layout of a single pixel is shown in figure 3. Due to the close proximity of the window pairs, each lenslet must simultaneously carry two signal beams. It was therefore decided to use a telecentric optical relay to ensure a high optical

throughput, simplify the optical design and maximize the tolerance of the interconnect to misalignment and fabrication errors.

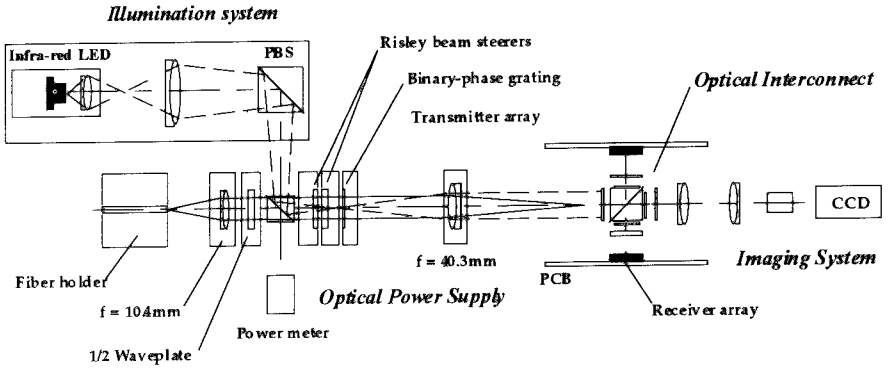


Figure 1 - Optical layout of PCB-to-PCB interconnection link.

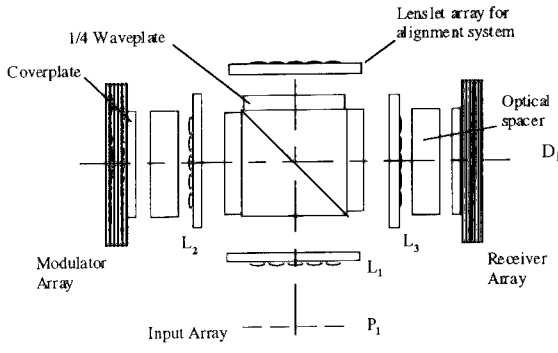


Figure 2 - Optical interconnect layout.

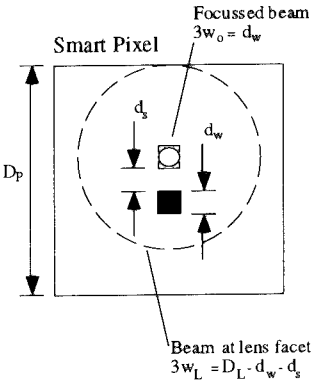


Figure 3 - Pixel window layout.

The optimum lenslet array geometry may be calculated by using a gaussian beam analysis to find the dependence of the window size, d_w, on channel density, C_D (channels per cm^2). The analysis presented here assumes the use of square packed lenslet arrays with dimensions of $D_L = D_P$. In addition, the restrictions $w_L = 1/3\ [D_L - (d_s + d_w)]$ and $3w_o = d_w$ have been imposed, where w_L is the gaussian ($1/e^2$) beam radius at the lenslet array plane and w_o is the beam radius of the focused beam. These limits ensure that minimum clipping of the beam occurs as it propagates through the interconnect [3]. Figure 4 shows the variation of d_w with C_D as a function of window spacing for a relay with focal length of f = 6.5mm (to obtain a PCB-to-PCB pitch close to current backplane standards) and an operating wavelength of 850nm. Higher channel densities can be obtained using a multi-channel (or super pixel) lenslet array configuration in which the windows of four adjacent pixels are clustered together and the optical signals relayed through a single lenslet facet having dimensions of $D_L = 2\ D_P$. The smart pixel layout is illustrated in figure 5. Using a similar gaussian beam analysis it can be shown that a multi-channel relay with a channel density of 1000 channels/cm^2 may be implemented using a lenslet array with focal length of f = 6.5mm, d_s = 20μm, an operating wavelength of 850nm and a window size of 34μm.

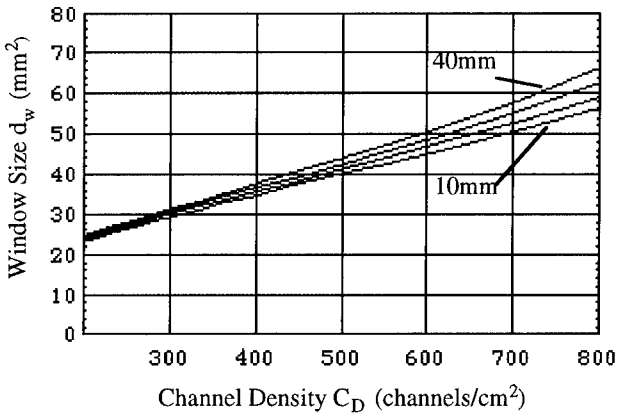


Figure 4 - Variation of d_w with C_D: f = 6.5mm, λ = 850nm, d_s = 10, 20, 30 and 40μm.

Optical Demonstrator Design

The FET-SEED arrays employed in this prototype backplane demonstrator had a pixel separation of 200μm, a window size of 25μm, a window spacing of 50μm and an operating wavelength of 850nm [1]. The optics were designed to connect a 2x2 array of channels spaced by 600μm with a device plane-to-device plane separation of just over 33mm. Multi-level diffractive lenslet arrays fabricated at Heriot-Watt University were used to implement the interconnect [4]. These had eight phase levels, a focal length of 6.5mm and a measured throughput of 90%. Optical simulations of the set-up described in figure 1 predict a theoretical gaussian $1/e^2$ beam radius of 9.98μm and a geometric spot size of 2.59μm at the receiver plane.

Optomechanics and Alignment

One of the disadvantages of a lenslet array based optical interconnect is its sensitivity to misalignment. An analysis of the systems optical alignment tolerances was carried out to determine the effect of a specific positional error on optical throughput (Table 1) . The results of these calculations were used to design an optomechanical set-up capable of meeting the required alignment tolerances. The illumination and imaging systems shown in figure 1 were also constructed to allow the receiver and transmitter arrays to be viewed during assembly.

Meeting the necessary alignment tolerances and space constraints required the development of custom optomechanics. The optical components were mounted in steel cylinders, prealigned and positioned on a magnesium baseplate using a magnetic slot channel arrangement [5]. The magnets were mounted on top of steel bars to increase the restraining force. This force can be varied by altering the thickness of the bars. A chamfered groove structure was chosen to reduce wear and improve alignment tolerances (figure 6).

The PCBs were attached to commercial 5-axis kinematic translation stages bolted to the baseplate. These were adjusted to get the required registration between the two device planes. Alignment of the input beams with respect to the modulator array was provided by a Risley beam steerer pair. The holders for these elements employed ball bearing mounts to reduce wear and improve rotational alignment accuracy.

Acknowledgments

The authors would like to thank M. R. Taghizadeh and the diffractive optics group at Heriot-Watt University, Scotland for fabricating the lenslet arrays. This work was supported by the

Canadian Institute for Telecommunications Research and the McGill BNR-NT/NSERC Industrial Chair in Photonics Systems. In addition, DVP acknowledges support from NSERC (#OPG0155159), FCAR (#NC-1415), and the McGill University Graduate Faculty.

References

[1] D.V. Plant et. al., "A FET-SEED Smart Pixel Based Optical Backplane Demonstrator", International Conference on Optical Computing, paper ThA3 (1994).

[2] H. S. Hinton, Canadian Institute for Telecommunications Research, Research Program 1993-94, pp. 143-156, 1993.

[3] F. B. McCormick, F. A. P. Tooley, T. J. Cloonan, J. M. Sasian, H. S. Hinton, K. O. Mersereau and A. Y. Feldblum, "Optical Interconnections Using Microlens Arrays", Optical and Quantum Electronics, **24**, 4, April 1992, pp. 265-477.

[4] M. R.Taghizadeh and J. Turunen, "Synthetic Diffractive Elements for Optical Interconnection", Optical Computing and Processing, **2**, 221-242 (1992).

[5] F. B. McCormick, F. A. P. Tooley, J. L. Brubaker, J. M. Sasian, T. J. Cloonan, A. L. Lentine, R. L. Morrison, R. J. Crisci, S. L. Walker, S. J. Hinterlong, and M. J. Herron, "Optical Applied Science and Engineering", SPIE Proc. **1533**, 91 (1991).

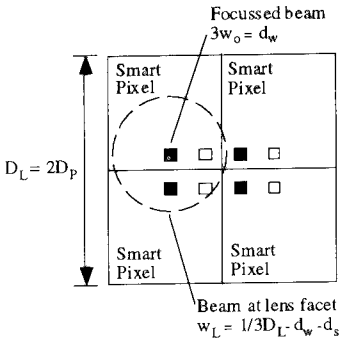


Figure 5 - Multi-channel interconnect.

System parameter	
Focal length of bulk lenses	±0.25%
Wavelength	±1nm
Δx, Δy of input array	±5μm
Δx, Δy of lenslet arrays	±10μm
Rotation of lenslet arrays	±1°
Translation of receiver array	±50μm
Δx, Δy of lenslet arrays	±1μm
Rotation of device planes	±0.25°
Tilt of modulator array	±0.25°

Table 1 - Alignment tolerances.

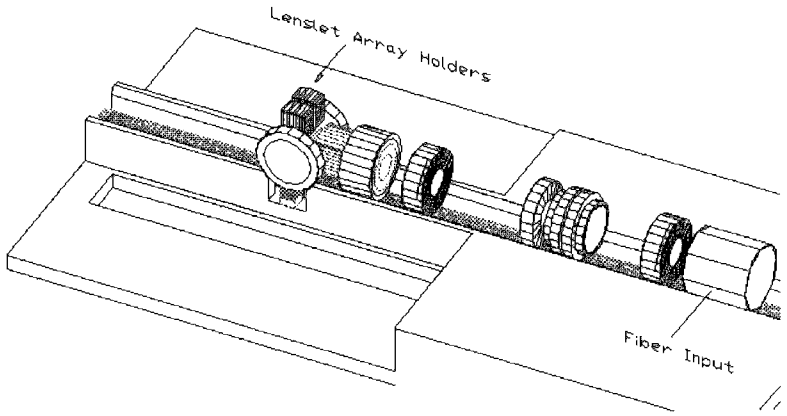


Figure 6 - Optomechanical layout of baseplate.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

An Approach to Active Alignment of Free-Space Optical Interconnects

*G.C. Boisset, B. Robertson, and H. S. Hinton**
McGill University, Department of Electrical Engineering, 3480 University Street,
Montréal, Québec, Canada H3A 2A7, email: boisset@photonics.ee.mcgill.ca
*Now with University of Colorado at Boulder

Abstract

An active alignment system is demonstrated. The error between the actual and desired spot location on a quadrant detector is used to compute the angular displacement required for two Risley Beam Steerers to centre the spot. Theoretical design and experimental results are given.

Introduction

Free-space optical interconnects hold the promise of alleviating communication bottlenecks in future connection-intensive electronic systems. A separable interconnect offering the massive connectivity of free-space optics will greatly increase the data throughput between multi-chip-modules and/or printed circuit boards (PCBs) on an optical backplane.

Alignment is a key issue facing free-space optical systems. For a system to be of practical use, it must be operate continuously for long periods of time in harsh industrial conditions.

At least two approaches exist to overcome alignment drift:

1) Design an extremely rigid system which will not drift over time. This can be performed by removing as many degrees of freedom as possible[1] and by prealigning.
2) Implement active alignment, a process in which system parameters such as throughput or error in spot position are monitored and fed back to a controller which realigns the system by altering the state of the optics. Such feedback loops exist in CD players[2].

For an optical backplane system with manually removable PCBs, active alignment will likely be a necessity. This paper describes an active alignment demonstrator for a free-space optical link. The basic control loop for the system is shown in Figure 1 and is explained below.

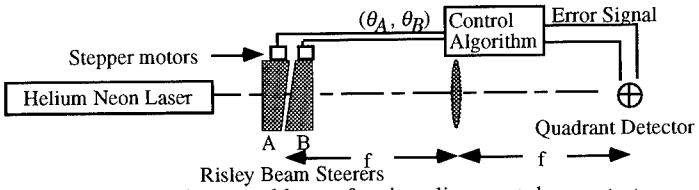


Figure 1: Basic control loop of active alignment demonstrator

Risley Beam Steerers (RBSs) were chosen as the optical components to move in an active alignment experiment because they have already been used for x-y alignment in free space-optical systems [1] and have proven to be a simple and cheap way of aligning free-space interconnects. Furthermore, components requiring rotational, as opposed to lateral, displacement can easily fit into a slotted plate, barrel or other integrated optomechanical setup. In addition, angular motion is easier to control than rectilinear motion when compensating for vibrational effects [3].

Definitions and assumptions

Let

1) An RBS be a nearly cylindrical optical component with a wedge in it, and whose purpose is to steer beams of light by imparting an angular displacement to them, as shown in Fig.2 and 3.
2) The two RBSs in the system be named RBS A and RBS B.
3) β_A and β_B be the wedge angle of RBS A and RBS B respectively.

4) θ_A and θ_B be respectively the angular rotation of RBS A and B about the z-axis (Fig.3).
5) (x_1,y_1) be the Cartesian representation of an arbitrary displacement imparted to a spot by the RBSs, and let d_1 be the magnitude of the displacement (x_1,y_1):

$$d_1 = \sqrt{x_1^2 + y_1^2} \tag{1}$$

6) θ_1 be a modified form of arctan (y_1/x_1), such that $0 \le \theta_1 < 2\pi$.
7) (x_u,y_u) be the position of the spot centre on the quadrant detector (QD) had it not been moved by the RBS pair.

Effect of two RBSs in series on the trajectory of a beam

From first order theory, the angular deviation, δ, imparted by an RBS on a beam is $\delta=(n-1)\beta$, where n is the index of refraction of the RBS (Fig. 2). As such, at the focal plane of a lens with focal length f, the lateral displacement, r, to the spot imparted by the RBS is:

$$r_A = f\tan\delta_A = f\tan((n-1)\beta_A) \tag{2a}$$
$$r_B = f\tan\delta_B = f\tan((n-1)\beta_B) \tag{2b}$$

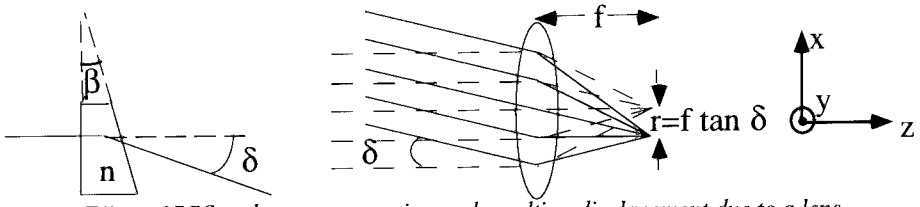


Figure 2 Effect of RBS on beam propagation and resulting displacement due to a lens.

Rotating the wedge about the z axis will cause the spot in the focal plane to move along the circumference of a circle of radius r in the x-y plane at the lens' focal plane. If RBS A and B are laid out in series along the beam path, the angular displacements imparted by the two RBSs will add. When RBS A and B are rotated about the z-axis, the spot will move on the periphery of a first circle of radius r_A whose centre is located on the periphery of a second circle of radius r_B, as shown in Figure 3.

Given the above explanation and definitions, it is possible to outline a centering algorithm. The following is an algorithm for determining the angular rotational displacement, (θ_A, θ_B), of the two RBSs required to centre on the QD a spot which is currently misaligned by $(\Delta x_e,\Delta y_e)$:

1) Determine (x_u,y_u), where (θ'_A, θ'_B) are the previous values of (θ_A, θ_B).

$$\begin{aligned} x_u &= \Delta x_e - \left(r_A\cos\theta'_A + r_B\cos\theta'_B\right) \\ y_u &= \Delta y_e - \left(r_A\sin\theta'_A + r_B\sin\theta'_B\right) \end{aligned} \tag{3}$$

2) Determine (θ_A, θ_B) necessary for the RBSs to impart a displacement of $(-x_u,-y_u)$. The Cartesian to Risley (CtoR transform) maps $(-x_u,-y_u)$ to (θ_A, θ_B):

$$d_1 = \sqrt{x_u^2 + y_u^2} \tag{4}$$

$$\theta_1 = \arctan(-y_u / -x_u) \text{ where } 0 \le \theta_1 < 2\pi. \tag{5}$$

$$\theta_a = \arccos\left(\frac{r_B^2 - r_A^2 - d_1^2}{-2r_A d_1}\right) \tag{6a}$$

$$\theta_b = \arccos\left(\frac{r_A^2 - r_B^2 - d_1^2}{-2r_B d_1}\right) \quad (6b)$$

The new value of (θ_A, θ_B) is thus:

$$\theta_A = \theta_1 + \theta_a \quad (7a)$$

$$\theta_B = \theta_1 - \theta_b \quad (7b)$$

(θ_A, θ_B) is now the desired input pair of angles to be fed to the stepper motor system controlling the RBSs. For the cases in which more than one solution is possible (equal wedge angles), the algorithm chooses the solution consistent with the geometry of Figure 3.

It is interesting to note the restrictions on d_1: 1) $d_1 < r_A + r_B$ and 2) $d_1 > |r_A - r_B|$.

The first condition simply implies that the total misalignment distance, d_1, must be less than the total radius of action of the RBSs. The second condition is more problematic and implies that there is a region, a 'blind area', to which the RBSs cannot centre a spot.

The blind area does not mean that the mismatched RBSs will never be able to centre a spot. A blind area means that the RBSs will never be able to impart a net displacement of zero μm to the spot. The corollary of this statement is: only if the spot is more than $|r_A - r_B|$ away from the centre before RBS insertion can it be centered after RBS insertion. In the future, matched wedges obtained by cutting one RBS in two along its diameter will be investigated.

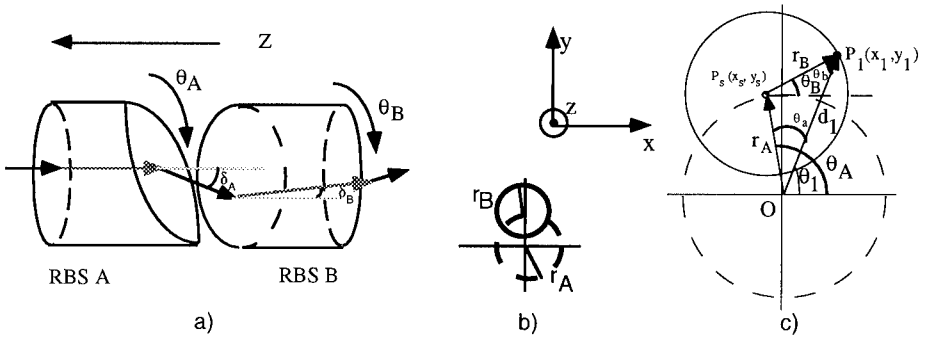


Figure 3 a): Effect on beam of two RBSs in series . Z axis is coaxial with RBSs. b) locus of spot movement on the focal plane, looking down the z-axis. c) expansion of b) with mathematical constructions used in derivation of transforms.

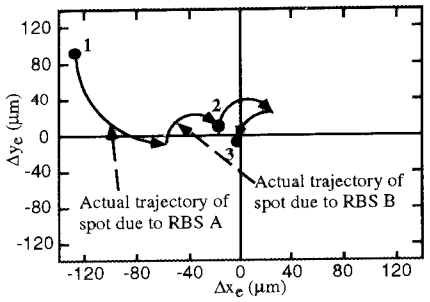
Initial conditions

Step 1 of the algorithm assumes that the previous values of (θ_A, θ_B) are known. However, the algorithm is only valid for the "steady-state": in order to begin, an initialization routine ('hunting') must determine the initial angular displacement of the RBSs. Given an unknown initial angular displacement (θ'_A,θ'_B) of the RBSs and an arbitrary ($\Delta x'_e$,$\Delta y'_e$), hunting involves rotating both RBSs until the spot's Δx_e coordinate is at a maximum (maximum deflection to the right). By definition, this angular position corresponds to (θ_A,θ_B)=(0,0).

Demonstrator and experimental results

The physical values used in the demonstrator were as follows: f = 100 mm, β_A = 0.208° and β_B = 0.278°. The beam was defocused on the QD such that the spot size was $3w \sim 585$ μm (99% power) in order to obtain a greater range for error detection. The lens and a custom holder for the RBSs were installed on a Spindler & Hoyer microbench setup built around stepper motors and a gear reduction system for rotating the RBSs. The QD was a Hamamatsu S4349. With a 10 volt applied reverse bias voltage, the generated photocurrents produced voltage swings across 330 kΩ resistors of 0-10 V which were sampled by a National Instruments Lab-NB™ I/O board on a Macintosh; the algorithm was encoded on LabView™ running on the Macintosh. The voltage pulses for turning the motors were sent out by the Lab-NB™. A UCN 5804 chip between the board and motors provided power gain for the motors.

Fig. 5a shows a typical run. The spot centre started at position 1, (Δx_e,Δy_e)= (-127,92) μm. After one iteration, the spot ended at position 2, (Δx_e,Δy_e)= (-17,12) μm. The error was likely due to laser fluctuations, mechanical backlash, and measurement errors occurring when the spot is far from the QD centre. A second iteration brought the spot centre to position 3, (Δx_e,Δy_e)= (-3.8,-6) μm. Fig 5b and 5c show a spot before and after a typical steady-state centering run. Table 1 outlines the results obtained from eight consecutive centerings performed by the demonstrator. In each case, the spot was aligned to within ~20 μm in one iteration and to within 10 μm in two iterations. The measurement accuracy of the system near the origin was ± 5 μm for each coordinate (Δx_e,Δy_e), and mechanical backlash caused an uncertainty of ± 2 μm in the spot position.



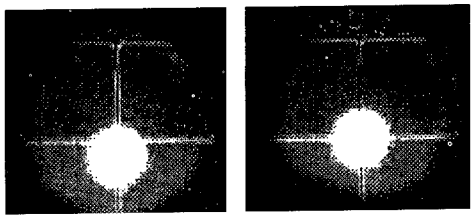


Figure 5a): centering operation. b) and c): spot before and after a centering run. Initially, the spot was manually misaligned to position (-213,-7) μm. After one iteration, the spot was at (-8,20) μm. (Image is inverted and reflected by imaging system)

Initial (Δx_e,Δy_e)	Initial distance to centre	Final (Δx_e,Δy_e)	Final distance to centre	iterations
(-127,92)	157	(-3.8,-6)	7.1	2
(109,5)	109.1	(-0.8,-2)	2.15	2
(-53,20)	56.64	(1.1,-6.1)	6.2	1
(106,-55)	119.4	(7.2,-1.4)	7.3	2
(-117,-89)	147	(5,-2.9)	5.8	2
(-85,19)	87.1	(-2,-5)	5.4	1
(90,103)	136	(7.1,6.8)	9.8	1
(89,76)	117	(-1.3,3.9)	4.1	3

Table I: Result of consecutive centering operations. All distances and positions in microns.

Conclusion and future directions

An active alignment system using Risley Beam Steerers to centre a spot on a quadrant detector was successfully designed, built and implemented. Future research will involve scaling the system to arrays of interconnects, for example by using higher orders of spot arrays to illuminate quadrant detectors on the periphery of a device array.

Acknowledgment

This work was supported by the Canadian Institute for Telecommunications Research and the McGill NT/BNR-NSERC Industrial Chair in Photonic Systems.

References:

[1] F. B. McCormick, T. J. Cloonan, A. L. lentine, J. M. Sasian, R. L. Morrison, M. G. Beckman, S. L. Walker, M. J. Wojcik, S. J. Hinterlong, R. J. Crisci, R. A. Novotny, and H. S. Hinton, "Five-stage free-space optical switching network with field-effect transistor self-electro-optic-effect-device smart-pixel arrays", *Appl. Opt* **33**, 1601 (1994).

[2] W. H. Lee, "Optical technology for compact disk pickups", *Lasers & Optronics*, **6**, No. 9, 1987.

[3] Y. Netzer, "Line-of-Sight steering and stabilization", *Optical Engineering*, **21**, 96 (1982).

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A Novel Programmable Liquid Crystal Lens Device for Adaptive Optical Interconnect and Beamforming Applications

Nabeel A. Riza and Michael C. DeJule

General Electric Corporate Research and Development Center
P. O. Box 8, KWB 617, Schenectady, New York 12301, USA

Abstract. A nematic liquid crystal three terminal device for optical beam steering is demonstrated using a thin film resistor network on the substrate layer forming a near continuous index perturbation adaptive lens. A 12 cm minimum focal length is demonstrated for a 1 mm square device.

1. Introduction

The use of nematic liquid crystals (NLCs) for making variable focal length lenses was first described as early as 1979, where S. Sato showed how NLC cells shaped like a plano-convex lens or a plano-concave lens could be electrically controlled to form adaptive lenses [1]. In 1984, S. T. Kowel et. al used a simple parallel geometry electrode structure (each electrode has a different applied voltage) in a uniform thickness NLC cell to demonstrate the optical focusing effect [2-3]. Later, in 1988, the same group showed optical beam translation with a similar electrode array geometry device, and described the effects of fringing fields of the electrodes on the NLC index perturbation [4]. Although, using a grid of independently driven parallel electrodes does produce an adequate refractive index perturbation across the NLC cell area to cause optical beam focusing, a useful amount of light is lost via the non-ideal (ideally, quadratic index pertubation with distance is required) near step-wise index variation caused by the discrete nature of the electrodes and their applied voltages. Therefore, it would be beneficial if the index variation induced by the electrode structure could made more gradual and smooth, thus improving the optical efficiency of the NLC lens. Also note that a useful NLC lens with this parallel electrode geometry would typically require > 100 electrodes, implying a 100 independent electronic drivers. Thus the whole purpose of using a compact, lightweight, low cost programmable lens would be defeated by the cost, weight, and power consumption of the electronics used to control the lens. Recently, we have described how the index perturbation between the electrode pairs can be made smooth, and how the need for independent electronic

drivers can be eliminated thus making a simple 3 terminal control device with one electronic driver [5-6]. This paper highlights these novel NLC device concepts.

2. Device Design and Experiment

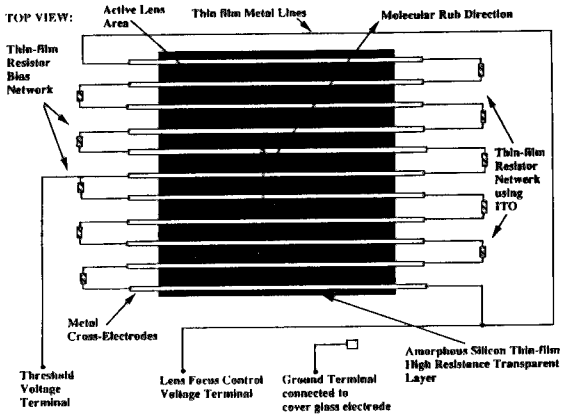


Fig.1 The novel thin-film resistor network biased nematic liquid crystal lens design (top view).

Fig.1 shows the top view of the novel thin-film resistor biased NLC cylindrical lens. An indium tin oxide (ITO) thin-film resistive network of specific resistor values generates a quadratically varying voltage gradient across the symmetrical (about the lens center) parallel electrode structure, with the resistors in series with the molybdenum metal electrodes. By applying a constant amplitude ac (e.g., 1 KHz square wave) signal to the top and bottom terminals (or electrodes) of the NLC device (see Fig.1), with the center electrode biased near the NLC threshold value (near 1 V) for molecular activation, a piece-wise quadratically varying index perturbation is generated by the electrode structure using a single driver or voltage level. By varying the voltage level, the focal length of this lens can be easily changed. Because there is a gap between the parallel electrodes, the NLC molecules in this region do not cause the required quadratic index pertubation behavior. We have somewhat alleviated this problem by adding a thin-film of very high resistance (e.g., 10 Mega-ohms/square mm) amorphous silicon to the lens area before depositing the much lower resistance (e.g., 0.4 ohms/square mm) parallel electrode structure. Thus, the amorphous silicon acts as a resistor between the conductor electrodes, generating a linear voltage (index) gradient between the electrode gaps. This behavior results in a smoother approximation to the desired continuous quadratic index perturbation required for the lens effect, thus improving the lens efficiency by reducing the diffracted light that is lost when the lens approximation occurs in discrete steps.

Fig.2 shows the experimental results from our 1 mm square NLC device. The figures on the left show a color-coded (or grey-scale) approximately quadratic left to right

(symmetrical at center of lens) optical phase/index perturbation shown by the device for two different applied voltage levels (or focal lengths) when viewed with white light using the device between crossed polarizers. The top photo on the right shows collimated 633 nm He-Ne light passing through the electrically turned-off NLC device.

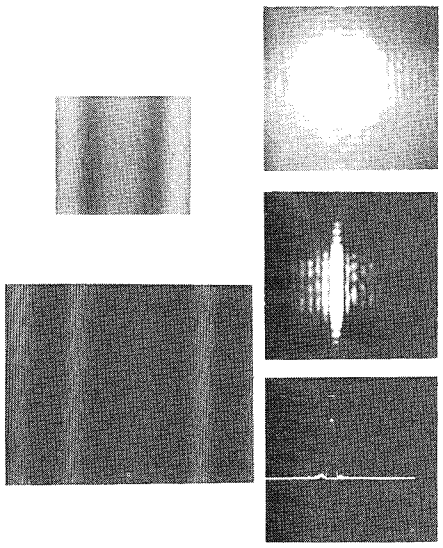


Fig.2 shows the demonstration of the 1 mm square NLC device fabricated at GE-CRD.

The diffraction effects observed are due to the light opaque square perimeter metal mask surrounding the active area of the device. The middle right photo shows the focusing effect when the device is turned on. Both pictures are taken at the NLC minimum focal length of 12 cm using a CCD. The focal length was variable from this minimum focal length to infinity by changing the amplitude of the single driver from 4 V to 1 V (the NLC threshold). The bottom photo on the right shows the focussed main beam and its sidelobes in temporal form as the output CCD video signal observed on an oscilloscope. Note that the sidelobes are much lower than the main beam, a desired feature of any good lens. In order to observe these relatively low sidelobes, we increase the laser beam power incident on the device which results in the main lobe saturating the CCD but producing a much brighter image of the previously low-level sidelobes (see middle photo).

3. Applications

Key applications for this adaptive NLC device include optical beam spoiling for laser communications applications, misalignment correction in optical recording/reading systems and beam deflection/alignment applications in optical computing and interconnection architectures. Figures 3 and 4 show two such programmable lens applications. It is important

to note that these lens devices, when appropriately manufactured, can be low cost, low electrical power devices, with useful apertures available in a variety of sizes because of the large area amorphous silicon LCD technology.

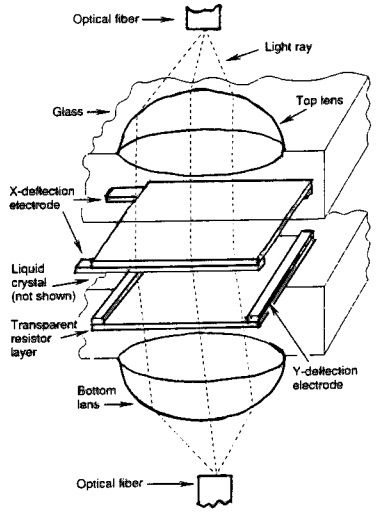


Fig.3 NLC lens used to correct for fiber misalignment errors.

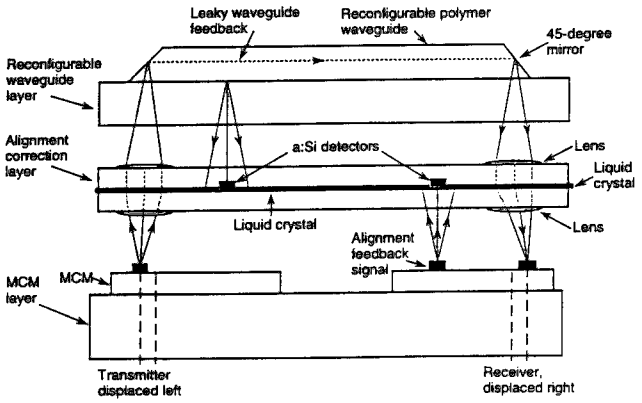


Fig.4 An adaptive NLC lens optical interconnect application for multichip modules.

References

[1] S. Sato 1979 *Japanese J. of Appl. Phys.* **18** 9 1679-1684

[2] S. T. Kowel, D. S. Cleverly, and P. G. Kornreich 1984 *Appl. Opt.* **23** 2 278-289

[3] S. T. Kowel, P. G. Kornreich, and A. Nouhi 1984 *Appl. Opt.* **23** 16 2774-2777

[4] P. F. Brinkley, S. T. Kowel, and C. Chu 1988 *Appl. Opt.* **27** 21 .4578-4586

[5] N. A. Riza and M. C. DeJule 1994 *General Electric Tech. Info. Series* 94CRD076

[6] N. A. Riza and M. C. DeJule 1994 *Opt. Lett.* 1994 **19** 14 1013-1015

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Integration of Fiber Holder and Microlens Arrays by deep Proton Irradiation

Maria Kufner, Stefan Kufner, Pierre Chavel

Institut d'Optique, CNRS, BP. 147, 91403 Orsay, France
e-mail: maria.kufner@iota.u-psud.fr, Tel. +33-1-69416870, Fax. +33-1-69413192,

Abstract. A monolithic integration of fiber holder arrays and microlens arrays well aligned to each other can be achieved by deep proton irradiation.

1. Introduction

For almost all fiber applications the combination of fibers and microlenses is of evident relevance. A microlens can interact with a fiber by coupling light from a source into a fiber and by collimating light coming out of a fiber. In both cases the longitudial and the lateral displacement between fiber and lens has to meet strictly some tolerances especially if single mode fibers are required. Also of interest is the alignment between arrays of fibers and arrays of microlenses. Problems arise from the fact that convential methods of fabricating a fiber holder and a microlens are farely distinct. Since several years many different fabrication methods for microlenses have been published [1,2,3]. Due to the lithographic design methods it is easy to fabricate arrays of microlenses instead of only a single microlens. This is not the case for fiber holders. If several optical fibers have to be aligned one line of fibers can be arranged in V-grooves [4,5,6]. A more recent publication tries to integrate a 4 by 4 array of single mode optical fibers in one substrate using excimer laser ablation [7]. But even here there is still a need for a microlens array well adapted to the spacing and the tolerances of the fiber holder array.

A fabrication method for both fiber holder arrays and microlens arrays monolithically integrated on one substrate is the technique of deep proton irradiation [8,9]. This is a tool for the fabrication of microlens arrays but as well for the fabrication of hole arrays which are suitable for fiber holding. A combination of these two kinds of arrays allows an assembly of fibers with lenses on top. Tolerances in the fiber diameter can be reduced by the possibility of selfcentering of the fibers in the holder.

2. Fabrication Method

The fabrication method is based on the fact that the irradiation of PMMA with protons changes the molecular weight. Thus those areas irradiated more than a critical dose can be dissolved in a special developer [8]. Furthermore the reduction of molecular weight increases the swelling potential of the material which means that monomer vapor can diffuse only in the domains with sufficiently low molecular weight and produces a significant swelling [9].

For the fabrication of fiber holder arrays and microlens arrays a PMMA substrate is irradiated through a mask with an array of circular apertures corresponding to the size and

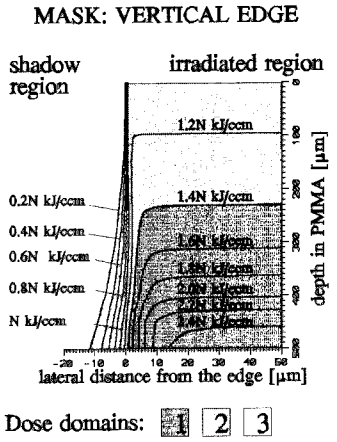


Fig. 1 Dose distribution in PMMA under an ideal edge

pitch of the fiber. Important in this context is the distribution of molecular weight caused by the proton irradiation. The simulation [10] in fig. 1 shows lines of equal dose deposition in a PMMA sample of 500 μm thickness (which is in this range equivalent to the distribution of molecular weight). Dose deposition is determined by the energy dependance of the absorption coefficient and by scattering. An ideal edge is assumed irradiated homogeneously with N * 10^{13} particles per cm^2 with a proton energy of 7 MeV.

The graph shows that at the bottom a significant dose deposition can be obtained in the geometrical shadow region of an edge which is caused by the lateral straggling of the protons. Furthermore it can be seen that the dose increases downwards which means that the maximum dose is deposited at the bottom of the structure. The choice of the dose parameter N it determines the level where in the substrate the development processes stops. In fig. 1 three areas are labelled which are of interest with respect to applications for fiber optics [11].

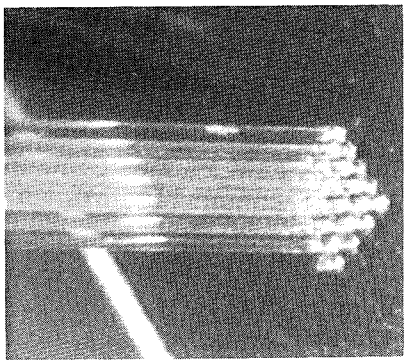


Fig. 2 Fiber holder array fixing 4x5 single mode fibers with a diameter of 125 μm

For the fabrication of fiber holders the dose is choosen that way that the developer can dissolve domains 1 and 2. This results in a fan-shaped hole through the substrate. A fiber can be plugged into this hole and afterwards the remaining domain 3 can be swollen in a styrene vapor athmosphere thus centering the fiber. Fig. 2 shows a photograph of a 4x5 single mode fiber array fabricated by this method. The fiber diameter is 125 μm.

For the monolithic integration of fiber holders and lenses the dose given to the substrate has to be choosen less than before so that only domain 1 is developed. The resulting structure now is a blind hole where a fiber can be plugged in. A subsequent diffusion of styrene vapor will swell domain 2 and 3. The swelling of domain 2 forms a surface lens on top of the substrate well aligned to the fiber and as above domain 3 is swollen for centering and fixing the fiber in the blind hole. Fig. 3 shows the depth of domain 1 as a function of the dose in a PMMA substrate with a thickness of 500 μm.

Fig. 4 is a photograph of a 4x5 array of blind holes irradiated with $15.5*10^{13}$ Ions/cm^2. The depth of the holes is 160 μm. The hole is slightly fan shaped with a bottom diameter of 140 μm and a top diameter of 130 μm.

Fig. 5 is side view of a plugged fiber with a monolithically integrated surface microlens on top. A fiber with a diameter of 140 μm is plugged into a 160 μm deep blind hole. The substrate thickness is 500 μm.

3. Summary

The method of deep proton irradiation allows the fabrication of fan shaped fiber holders which can compensate small tolerances in the fiber diameters by swelling the holder. Arrays of fiber holders and microlenses can be fabricated together in the same PMMA substrate. This kind of monolithic integration ensures that the microlenses are exactly on top of their corresponding fiber holders, therefore no further alignment is necessary.

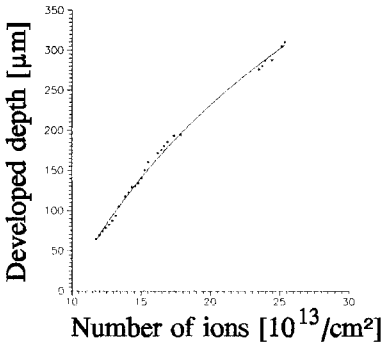


Fig. 3 Measured depth of development as a function of the dose

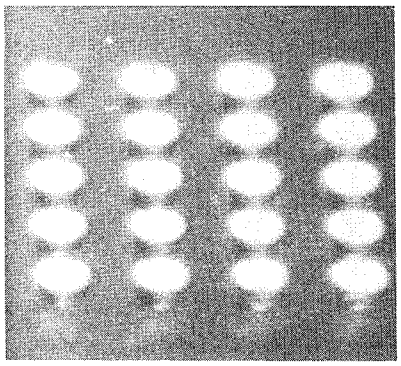


Fig. 4 4x5 array of blind holes in PMMA; depth: 160 μm, diameter: 140 μm (bottom), 130 μm (top)

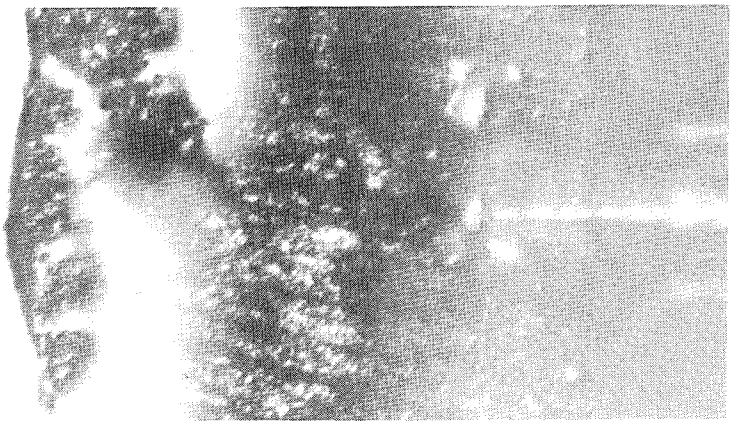


Fig. 5 Side view of a monolithically integrated fiber lens combination; parameters are as in fig. 4

4. Acknowledgment

This work has been supported by CNRS (Centre National de la Recherche Scientifique), Ultimatech and HCM (Human Capital and Mobility) projects.

5. References

[1] N. F. Borelli, et al., Appl. Opt. 24, 2520 (1985).
[2] M. C. Hutley, J. of Modern Optics 37 (2), 253 (1990).
[3] K. Iga, et al., Appl. Optics 21, 3456 (1982).
[4] M. S. Cohen, et al., Photon. Tech. Lett. 3, 985 (1991).
[5] C. A. Armieuto, et al., Electron. Lett. 27, 1109 (1991).
[6] A. Sasaki, et al., Photon. Tech. Lett. 4, 908 (1993).
[7] G. M. Proudly, et al., Opt. Engeneering 33 (2), 627 (1994).
[8] S. Kufner, et al., Pure and Appl. Opt. 2, 111 (1993).
[9] M. Kufner, et al., Pure and Appl. Opt 2, 9 (1993).
[10] J. F. Ziegler, et al., The stopping of Ions in Solids, (Pergamon Press, New York, 1985).
[11] M. Kufner et al., accepted by Optics Letters

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Packaging tolerance constraints in the optimisation of Fresnel computer generated holograms for optical interconnections.

KRESS Bernard♣, SCHILLINGER Marc♦, MEYRUEIS Patrick♣

♣ Laboratoire des Systèmes Photoniques, ENSPS/ULP
Boulevard Sébastien Brandt, F-67400 ILLKIRCH, France.
Tel: (33) 88 65 51 61, Fax: (33) 88 65 51 60

♦ Institut d'Optique, Bât. 503 - Centre Scientifique d'Orsay
F-91400 ORSAY, France.

Abstract. We address the needs to developp new design algorithms and optimisation techniques for Computer Generated Holograms (CGH) that have to work efficiently within an optoelectronic packaged system. First, we develop practical solutions to increase sensibly the performances of a modified Gerchberg-Saxton (G-S) iterative algorithm. We then analyse the various effects of the different packaging tolerance constraints on the CGH performances, and finally propose some iterative optimisation techniques to overcome these drawbacks.

1. Introduction.

The design and optimisation process of Computer Generated Holograms (CGHs) for free-space optical interconnections (FSIO) has been widely investigated [1],[2],[3]. Recently [4], we have demonstrated the potentials of a modified G-S algorithm, and compared its performances to several analytical generation methods to design Fresnel CGHs for FSOI. On the other hand, considerable work has been done on packaging and tolerance analysis in opto-electronic packaged systems [5],[6]. In this paper we investigate the possibility to design and optimise Fresnel CGHs that accomodate some of the tolerances constraints that occur in packaged systems.

2. The modified G-S iterative algorithm.

The G-S algorithm is a powerful and fast algorithm to design and optimise Fresnel CGHs for FSIO, when multiple fan-out and focussing properties are desired at the same time in the same element [4]. However, much improvement can still be done in the uniformity of the reconstruction by implementing different solutions into the basic G-S loop.

We have chosen a 16 phase levels element, 250µm square, composed of 128 by 128 kinoform encoded square cells, a focal length f=800µm and a fan-out of 11 reproducing the letters 'OC'. For the source, we use a λ=830nm laser diode with an elliptic beam waist.

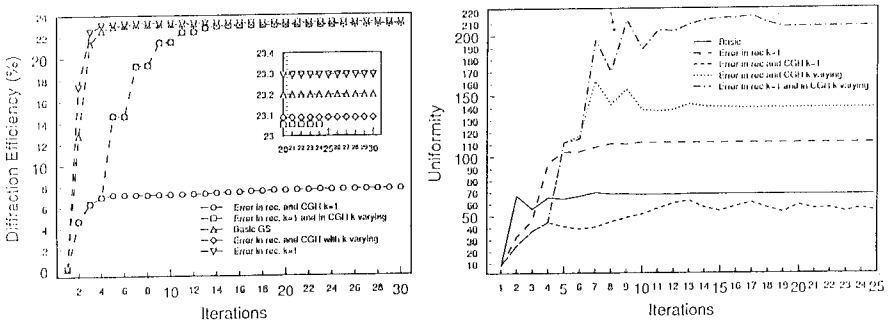


Fig.1: Diffraction efficiency and uniformity of reconstruction at focal plane, when applying several error reinjection methods into the G-S algorithm.

The results in diffraction efficiency (after approx. 10 loops) are not affected by the changes in the algorithm, but the uniformity has been much increased, especially when the error to be reinjected is scaled by a factor ***k*** which depends on the convergence rate (see ***Fig.1***). Also, it shows that the reinjection in the reconstruction plane is not necessary, whereas the reinjection in the CGH plane makes all the difference.

3. Effects of packaging tolerances on the CGH performances

For a CGH to work efficiently within a packaged system, it has to meet the standards of tolerances of the system in which it has been packaged. These include:

- Fabrication tolerances in both Diffractive Optics Plane (DOP) and Processing Plane (PP): systematic errors (etch depth errors, lateral misalignements), random errors (waviness and roughness of substrate) and encoding errors (square cells for kinoform encoding).
- Alignment tolerances between the DOP and PP (use of spacers, flip-chip bonding,...).
- Operating tolerances (spectral bandwidth, operating mode and beam waist of source).
- Local thermal expansion coefficients of DOP, spacers, PP, and thermal effects on source (spectral shift).
- Architecture tolerances (placement of processing elements), dimension of active surface on detectors, position of dots within the PP.

We have numerically simulated the effects of these tolerances by considering the diffraction efficiency, uniformity, shift and defocalisation of the spots for four different packaging tolerance constraints.
Fig.2 reports the results for several packaging tolerance constraints.

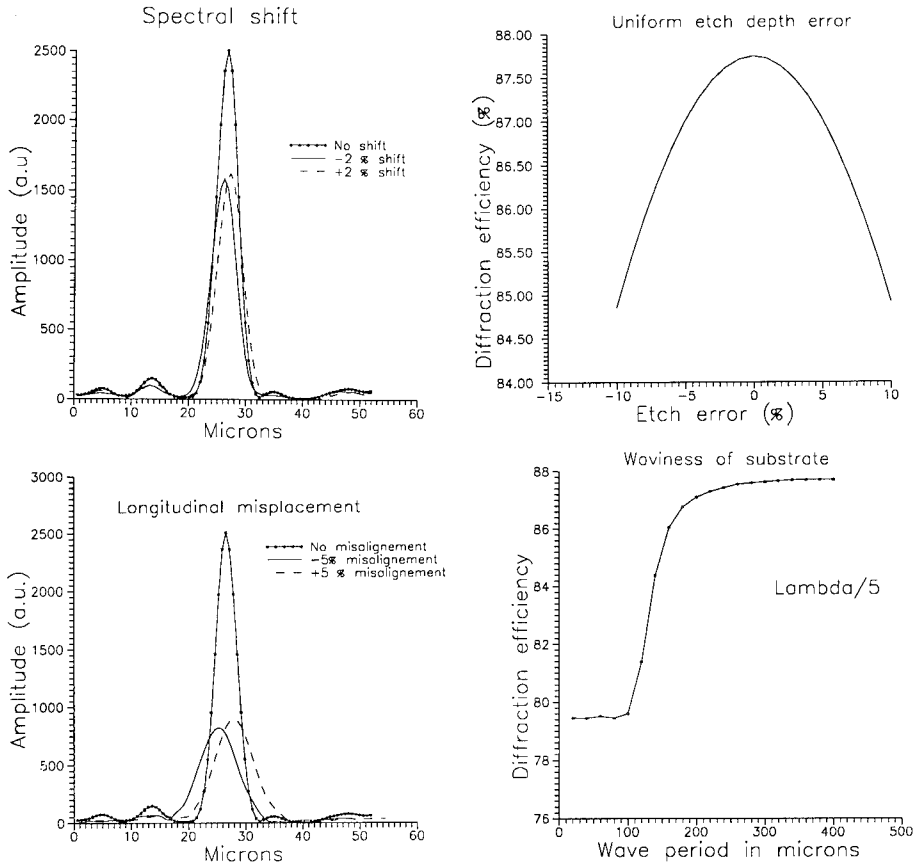


Fig.2: Effects of spectral shift of source, uniform etch depth errors, longitudinal misplacement and waviness of substrate on CGH performances.

4. Fresnel CGH designs algorithms including packaging tolerance constraints

The packaging tolerances investigated in 3 have in common the variation of the wavelength λ (spectral shift and indirectly etch depth errors and waviness), and/or focal length f (longitudinal misalignement). Also, it has been shown that the thermal expansion coefficient relates the change in f to the change in temperature [7]. Besides, the kernel functions we use do not see the variation of λ or f distinctly, but only the variation of the factor λ*f (e.g. fast Fresnel transform). So most of the tolerances described here can then be related to longitudinal misplacement (e.g. variation of f). If we can design a CGH with extended depth of focus, we can accomodate these various tolerances. We have therefore developed 3 different algorithms based on the G-S loop.

The first one bounces back and forth between the CGH, the focal plane and two planes each side of the focal plane, where the constraints are the waists of the defocussed beams decreased by a factor α (thus increasing the depth of focus).

The second one is based on a G-S optimised Fourier CGH which performs the reconstruction in the far field: this reconstruction is brought into the near field by modulating the Fourier CGH by the following phase profile.

$$\Phi = \frac{\pi.(x^2 + y^2)}{\lambda.\left[f_0 + f_{depth}.\cos\left(arctg(\frac{y}{x}) \right) \right]}$$

(a spherical lens with a focal length varying $f_0 - f_{depth}$ to $f_0 + f_{depth}$).

The third method uses an extended kernel function instead of the Fresnel transform in the G-S loop (both quadratic phases in the expression of the Fresnel transform are replaced by the above analytical expression). To compare their performances to the straight forward Fresnel G-S algorithm, we consider the amount of light falling into 11, 10.0μm square detectors in the PP plane, and the beam waists and lateral shifts of the 11 spots, when the longitudinal displacement varies from -10% to +10%. ***Fig.3*** reports these variations.

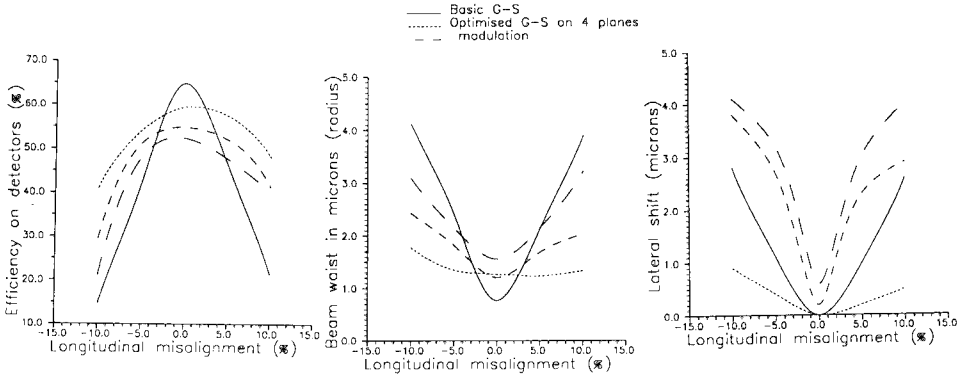


Fig.3: Variation of efficiency, beam waist and lateral shift, for the different algorithms, when longitudinal misalignements occur.

5. Conclusion.

We have discussed the needs to design and optimise Fresnel CGHs for free-space optical interconnections by considering closely several packaging tolerance constraints. We analysed the effects of these tolerances on the CGH performances, and proposed different algorithms to design multiple-fan-out Fresnel CGHs that will accomodate these tolerances.

References

[1] **S.H.Lee** 1993 *Critical review on diffractive and miniaturized optics, pp.290-301.*
[2] **J. Stack, M.Feldman** 1992 *Applied Optics, 31, pp.4839-4846.*
[3] **H.Bartelt, J.Horner** 1985 *Applied Optics, 24, pp.2894-2896.*
[4] **B. Kress, S.H.Lee** 1993 *OSA Topical Meeting on Optical Computing, pp.22-25.*
[5] **J.Jahns, Y.H.Lee** 1992 *IEEE Phot. tech. Lett. Vol 4, pp.1362-1372.*
[6] **Rick L.Morrison** 1993 *OSA Topical Meeting on Optical Computing, pp.250-253.*
[7] **T.H.Jamieson** 1981 *Optical Engineering, 20 pp.156-160.*

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Multilevel diffractive phase elements with trapezoidal shaped pulses.

P. Blair, H. Lüepken[1], M. R. Taghizadeh and F. Wyrowski[2]

Department of Physics, Heriot-Watt University, Edinburgh, Scotland.
[1]Department of Physics, Essen University, Essen, Germany.
[2]Department of Diffractive Optics, Berlin Inst. of Optics, Germany.
Email - paul@phy.hw.ac.uk

Abstract: The implementation of a general design concept for the design of diffractive phase elements, with a trapezoidal pulse structure, for array illumination is described. Results are presented for on-axis two-dimensional array illuminators designed using this method. The final designs are within 5% of the calculated upper bound of the diffraction efficiency for the given signal and signal non-uniformity is below 1%.

1. Introduction

Design theory for diffractive optical elements (DOEs) is a rapidly expanding area of research, especially within the area concerned with the design of DOEs that operate within the paraxial domain. In particular the needs of digital optics and material processing lead to a requirement for diffractive phase elements (DPE) that split an incoming light beam into an array of prescribed light spots, i.e., array illuminators.

The design problem for array illuminators is expressed in general terms as a desire to find a surface-relief phase profile $\theta(\bar{u})$ (where $\bar{u}=(u,v)$) that has a phase transmittance or reflectance $H(\bar{u})=\exp[i\theta(\bar{u})]$ which satisfies (within the paraxial domain)

$$\left|S(\bar{x})\right|^2=\left|\mathcal{F}[H(\bar{u})]\right|^2 \qquad \bar{x}\in\mathcal{W}, \tag{1}$$

where $\bar{x}=(x,y)$, $\mathcal{F}$ is the Fourier transform operator and $S(\bar{x})$ is the complex amplitude signal of prescribed intensity defined within the signal window $\mathcal{W}$. For the case of a general array illuminator the prescribed signal is given by

$$S(\bar{x})=\sum_{m,n} s_{m,n}\delta(x-m\Delta x)\delta(y-n\Delta y), \tag{2}$$

where $s_{m,n}$ are the desired complex amplitudes of the signal, $\delta(x)$ is the dirac-delta function and Δx and Δy are the spacing of the orders in the signal plane.

Typically the diffraction efficiency of the signal is also specified such that

$$\eta(Z)\geq\eta_{\min}, \tag{3}$$

where $\eta(Z)$is the diffraction efficiency of the final DPE of Z phase levels and $\eta_{\min}$ is the minimum diffraction efficiency required for the application. The dependence of the diffraction efficiency upon Z is a product of the fabrication techniques commonly employed

to realise DPEs. The most widely used is multi-mask microlithography. This requires the phase profile $\theta(\bar{u})$ to be quantised to a discrete number of phase levels (ideally 2^N, where N is an integer). Thus diffraction efficiency becomes a function of Z. Information about the minimum value of Z necessary to satisfy Eq. (3) would facilitate the fabrication process and reduce intrinsic fabrication errors. It is clear that the greater the number of masking steps, the larger the error becomes [1].

A design concept embedded within the theoretical framework of diffraction elements that solves the design problem above has been described in Ref. [2]. Given here is an implementation of that design concept for multilevel array illuminators. Particular emphasis is made upon the advantages of a parameterisation that leads to trapezoidal phase pulses.

2. Optimisation of the phase of $S(\bar{x})$.

Equation (1) is only concerned with the magnitude of the complex amplitude signal $S(\bar{x})$; this leaves the phase of $S(\bar{x})$ as a free parameter. This can be used to provide information about the upper bound of the diffraction efficiency, $\eta_l(Z)$, of quantised DPEs and suggest an initial starting point for the DPE design [2]. The diffraction efficiency upper bound is given by [3]

$$\eta_l(Z)=\frac{<|H(\bar{u})|\cos[\Delta\theta(\bar{u},Z)]>^2}{<|H(\bar{u})|^2>}, \tag{4}$$

where $\Delta\theta(\bar{u},Z)$ is the minimum phase change needed to map $\theta(\bar{u})$ onto an allowed phase level and $< >$ indicates an averaged value. By optimising the phase of $S(\bar{x})$ we minimise $\Delta\theta(\bar{u},Z)$ and $\eta_l(Z)$ approaches a maximum. In this way we gather information that relates diffraction efficiency to phase levels for a given signal and can ensure that Eq. (3) will hold *before* continuing with the DPE design.

Once the optimised complex signal $S'(\bar{x})$ is found (all optimised quantities will be indicated with a dash) a projection of the Fourier spectrum of $S'(\bar{x})$ is made onto the allowed phase levels,

$$H'(\bar{u})=\exp\left[i\left(\theta'(\bar{u})+\Delta\theta(\bar{u},Z)\right)\right]. \tag{5}$$

This provides a set of phases for a sampled quantised phase hologram. The parameterisation of this data set is dealt with in the next section.

To find $\eta_l(Z)$ and $H'(\bar{u})$, i.e. to implement Eqs. (4) and (5), we use the iterative Fourier transform algorithm (IFTA) [4]. This is a computationally efficient non-linear optimisation technique that employs the projection method. A fast Fourier transform is used to project a sampled complex data set between signal and hologram planes. In these planes the data set is subject to a number of constraints stemming from the design problem stated above. In the signal plane, while the phase is a free parameter, the magnitude of the complex signal is maintained at the correct level. In the hologram plane the constraints ensure that the requirements for a quantised phase only element holds. The optimal signal phase is found and $\eta_l(Z)$ determined by alternating between signal and hologram planes for a fixed number of iterations (typically 100). A final projection provides $H'(\bar{u})$. By running the algorithm with different random starting values for the signal phase we ensure that stagnation does not occur.

3 Trapezoidal parameterisation

If we assume that the profile is smooth between sample points (which is the case if a large enough number are used) then the sample points are transition points of a complicated phase topology. At present no design concepts exist to describe a hologram of arbitrary shaped pulses in closed form. A parameterisation does exist, however, to approximate the topology using trapezoidal pulses confined to strips [5]. If we define this parameterisation as the operator $\mathcal{P}$ then the phase profile becomes ${}^{T}\theta'(\overline{u}) = \mathcal{P}[\theta'(\overline{u})]$.

It is useful in the final stages of the DPE design to define a parameter to describe how close the DPE signal approaches the ideal. Here we use the reconstruction error, ΔR, defined in Ref. [4]. ΔR is a measure of the maximum relative deviation of the DPE signal intensity from the average power. For most applications $\Delta R < 0.1$, although some optical computing applications are more demanding and require $\Delta R < 0.01$.

Using the closed form power spectrum for a DPE with trapezoidal features, ΔR can be minimised with a suitable non-linear optimisation algorithm. The choice of algorithm is simplified by noting the hologram is "pre-optimised" for diffraction efficiency and the trapezoidal approximation should be close to the ideal solution in hologram space. ΔR is highly sensitive to movement of transition points and an optimisation routine that can take advantage of a near-optimal solution is desirable. Two approaches that fit the criterion have been implemented; Brent's method and a greedy search (GS). In both cases the cost function to be optimised, M, is defined in Ref. [4]. Brent's method [6] minimises ΔR by fitting changes in M from transition point movement to a parabola. By consecutively finding the minimum point of the parabola (and thus the corresponding minimising transition point) for each transition point in the DPE, ΔR approaches zero. This method will only work for a cost function that behaves as a parabola around the global minimum; that is the case here. The GS algorithm is perhaps the oldest of all non-linear optimisation methods and is generally considered to be too computational intensive for effective use. However, because of the almost optimal problem presented this is not the case and the ease of implementation of the GS method becomes an advantage. The GS algorithm works by moving successively through the transition points with a steadily decreasing trial change to test if this decreases M.

At this point in the optimisation a further design constraint is added to the design problem. As the DPE is designed within the paraxial domain care must be taken to ensure that the feature size of the trapezoidal pulses does not become too small. This would both invalidate the paraxial approximation and increase ΔR by the introduction of high spatial frequencies in ${}^{T}\theta'(\overline{u})$.

4. Experimental results

A number of DPE array illuminators have been optimised for a variety of fanout sizes for both $Z = 4$ and $Z = 8$; a selection of results are presented in Table 1. For the even fanout gratings the even-orders missing (EOM) symmetry [6] was used to further simplify optimisation. An important point to note is the slight decrease in $\eta_l(Z)$ after the final optimisation step. This is a result of both the parameterisation and the optimisation step to reduce ΔR. However, this reduction is less than 5% and can be taken into account in the early stages of the DPE design. In all cases ΔR is within the limit required by digital optics. Table 1 also provides a measure of the optimisation time required for the *final* step, which is

Table 1. Array illuminators designed using the described design concept.

Fanout Size	Z	Sample Size	$\eta_i(Z)$	$\eta'(Z)$	ΔR / %	$\eta_i(Z) - \eta'(Z)$	t (s)
3 x 3	4	32 32	0.787	0.774	0.219	0.013	7
	8	32 32	0.892	0.881	0.250	0.011	13
5 x 5	4	64 64	0.791	0.774	0.400	0.017	136
	8	64 64	0.885	0.868	0.500	0.010	197
7 x 7	4	64 64	0.787	0.760	0.050	0.027	416
	8	64 64	0.878	0.867	0.004	0.011	912
9 x 9	4	64 64	0.782	0.761	0.004	0.021	2377
	8	64 64	0.879	0.861	0.496	0.018	6352
11 x 11	4	128 128	0.801	0.750	0.049	0.051	5399
	8	128 128	0.892	0.865	0.665	0.027	7485
4x4(EOM)	4	32 32	0.823	0.801	0.589	0.022	142
	8	64 64	0.912	0.910	0.589	0.002	142
8x8(EOM)	8	64 64	0.873	0.857	0.190	0.020	-

generally greater than the IFTA phase, especially for high fanout, but still remains low enough to make the design method practical (optimisation took place on a Sun SparcStation 10/51).

5. Conclusion

An implementation has been described of a general design concept which uses a trapezoidal parameterisation to design two-dimensional DPE array illuminators. The design concept provides information about the diffraction efficiency upper bound that can be obtained for a given signal as a function of allowed phase levels. The final DPEs have diffraction efficiencies which fall within 5% of the upper bound. In the final stages a trapezoidal parameterisation has been used to maintain the phase topology obtained from optimising the signal phase. This parameterisation also provides a sufficient number of transition points to reduce signal non-uniformity to a level demanded to by today's DOE applications.

Acknowledgements. This joint research program has been supported by NATO under a collaborative research grant. The research at Heriot-Watt was also supported by SERC under the Scottish Collaborative Research Initiative in Optoelectronics (SCIOS).

References

[1] Miller J M, Taghizadeh M R, Turunen J, Ross N, Noponen E and Vasara A 1993 *Appl. Opt.* **32** 2519-2525

[2] Lüepken H and Wyrowski F 1994 *Appl. Opt.* (Submitted)

[3] Wyrowski F 1992 *Opt. Comm.* **92** 119-126

[4] Wyrowski F and Bryngdahl O 1988 *Opt. Soc. Am. A* **5** 1058-1065

[5] Vasara A, Taghizadeh M R, Turunen J, Westerholm J, Noponen E, Ichikawa H, Miller J M, Jaakkola T and Kuisma S 1992 *Appl. Opt.* **31** 3320-3336

[6] Press W H (Ed) 1992 *Numerical Recipes in Fortran* (Cambridge University Press) 395-399

[7] Morrison R 1992 *J. Opt. Soc. Am. A* **9** 464-471

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Demonstration and discussion of an interlaced fan-out interconnect

N. Collings and C. Berger

Institute of Microtechnology, Rue A.-L. Breguet 2, University of Neuchâtel, CH-2000 NEUCHATEL, Switzerland.

Abstract
2D fully interconnected processing systems where the input array is dilute can use a fan-out interconnect where the extra channels are accommodated within the repeat spacing of the input array.

1. Introduction

A commonly used 2D fan-out (fan-out **A**), where the input array is replicated at neighboring spatial locations which are located at distances of at least the array size from the original array, is illustrated in Fig.1.

Fig.1. Conventional 2D array fan-out.

A complementary fan-out is illustrated in Fig. 2 (fan-out **B**). It is complementary in the sense that, whereas **A** is useful when the subsequent fan-in is to an array of larger size repeat spacing than the input, **B** is useful when the output array is of smaller repeat spacing. This can be useful in cascaded interconnect systems, for example.

Fig.2. Interlaced 2D array fan-out.

A further advantage of **B** is that it allows lower resolution fan-out gratings to be used than **A**. This in turn eases the fabrication tolerances of the grating. The reduced fan-out angle also allows more compact systems to be envisaged.

2. Demonstration of interlaced fan-out

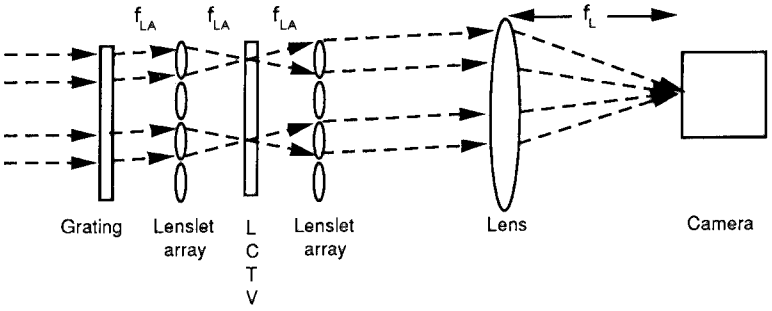
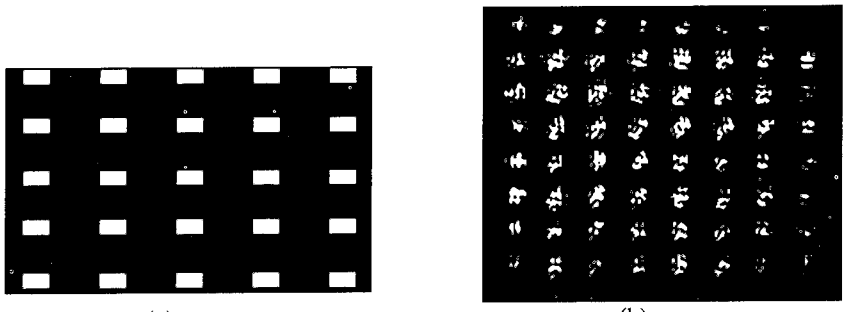


Fig. 3. Component layout of interlaced fan-out (polarisers omitted); f_{LA} is the focal length of each lenslet array, and f_L is the focal length of the focussing lens.

The layout of the optical system is shown in figure 3. Collimated light of 6 mm diameter from a HeNe laser beam is used in this pilot experiment, and the apertures of the beamlets (of which two are shown in figure 3) are defined by the apertures of the lenslets. In a real application of this interconnect, the beamlets would be defined by the source/modulator array and associated optics. The collimated beam is incident on two crossed 8 x 1 Dammann gratings of 690 μm and 567 μm periods respectively, which were fabricated at the Paul Scherrer Institute, Zürich (PSIZ). A lenslet array of 25 mm focal length and 1.15 x 1.4 mm repeat spacing (also fabricated at PSIZ [1]) acts as an array of Fourier Transform lenses to form multiple fan-outs in the focal plane of the lenslets. In the focal plane, a liquid crystal television (LCTV) (Seiko-Epson VPJ-2000) is placed. The grating periods and the focal length of the lenslets are chosen such that an 8 x 8 array of spots is formed by each lenslet on the 56 μm x 46 μm repeat spacing of the pixel modulators of the LCTV. The repeat spacing of the lenslets is chosen such that 25 pixel repeat spacings separate each array of 8 x 8 (Figure 4a). A lenslet array of the same specifications follows the LCTV, thus forming a telescope of unity magnification.



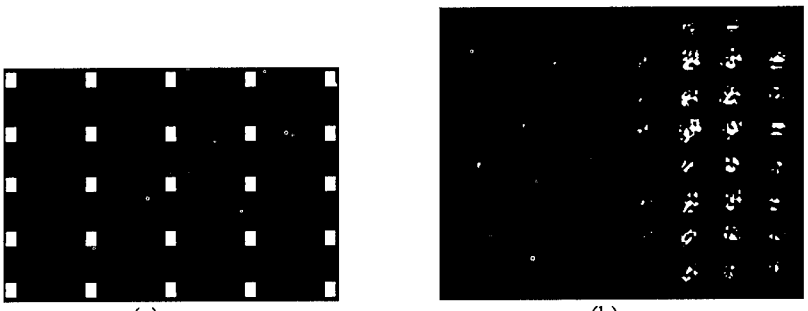
(a) (b)

Fig. 4. Layout of illuminated regions on the LCTV (a), and rear focal plane of lens (b).

A 100 mm achromat fans-in the parallel beams from each 8 x 8 array onto a spot on the camera. Therefore, spot (2,1) in figure 4b has received light from all the pixels (2,1) in the interconnection matrix in figure 4a. In the present experiment, the distance between the second lenslet array and the lens is 120 mm. We believe that this distance is noncritical.

3. Results

In the rear focal plane of the lens (output plane) there is just a single 8 x 8 array as would be formed by the combination of Dammann gratings and 100 mm focal length lens (Fig. 4b). When the number of columns of open apertures in the LCTV is reduced from 8 to 3 (Fig. 5a), a 8 x 3 array results in the output plane (Fig. 5b). The faint appearance of the 5th column in figure 5b is due to the limitations in the LCTV address electronics. Although the 5th column is shut off, electronic crosstalk from the 6th column gives the pixels a finite transmission.



(a) (b)

Fig.5. Layout of 8 x 3 apertures on the LCTV (a), and rear focal plane of lens (b).

4. Discussion

There are a number of uses for this type of fan-out. In the present case, it was developed in the design procedure for an optoelectronic multilayer Perceptron [2]. The intermediate processing plane (PP) in this Perceptron has the same physical dimensions as the interconnection plane. The two design options are either to demagnify the output of the PP and use a conventional fan-out (Fig. 1) or to maintain the size of the output of the PP and use an interlaced fan-out (Fig. 2). The latter eases the fabrication tolerances on the grating and allows larger spectral bandwidth illumination sources, due to the reduced fan-out angle requirement. Moreover, since both the lenslets and lens work at small field angles, no special lens design features are required. We expect improved spot sizes in the output plane in a optimised system because the fan-in is incoherent across the full aperture of the lens.

A further potential application area arises when the Hadamard product of a matrix and its transpose is required, eg in Delta rule and backward error propagation learning in neural networks [3]. The matrix which results from a conventional fan-out of the letter L is represented in figure 6a.

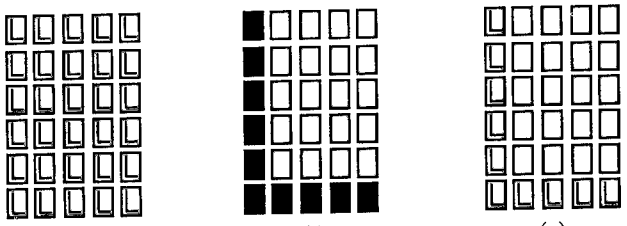


Fig. 6. The Hadamard product of (a) and (b) forms a Delta rule weight update matrix (c).

The matrix which results from an interlaced fan-out is represented in figure 6b. This matrix (A_{klij}) can be obtained by transposing the high frequency row and column indices of the matrix in figure 6a (A_{ijkl}). Holographic techniques have also been developed for performing this transposition [4]. The point-by-point or Hadamard product of these two matrices (Fig. 6c) is the matrix which is required for updating the interconnection matrix in Delta rule learning. Similar outer products between the neural output matrices and the error matrices are required for weight correction in backward error propagation learning. Also, outer products are used to form the weight matrix in high-order networks [5].

The hardware for an optically updatable interconnection matrix is being developed which uses Delta rule learning [6]. The optical system around this learning chip will have to generate matrices corresponding to Figs. 6a and 6b. Therefore, the interlaced interconnect could be useful in this context.

Finally, it should be noted that the latest generation of optical array sources/modulators are sufficiently dilute to warrant consideration of an interlaced interconnect. For example, vertical cavity surface emitting lasers have large spacings (up to 250 μm) in 8 x 8 array format [7]. Smart SLMs have large spacing between the modulators (up to 400μm) for electronic processing at each pixel. Moreover, the desire to build compact systems with these devices demands an interconnection technology which eases the fabrication requirements on the components. Interlaced fan-out does this for the gratings, but demands high quality lenslet arrays as a complement. It is hoped that the present emphasis on fabricating lenslet arrays succeeds in producing high quality components.

References

[1] Kunz R E and Rossi M 1993 *Opt. Comm.* **97** 6-10
[2] Collings N 1994 *Workshop on optical pattern recognition* eds. P. Réfrégier and B. Javidi (SPIE Press) 314-333
[3] Rumelhart D E and McClelland J L 1986 *Parallel Distributed Processing* (MIT Press, Cambridge, Mass.)
[4] Sirat G Y, Maruani A D, Chevallier R C 1989 *Appl. Opt.* **28** 1429-1435
[5] Lee Giles C and Maxwell T 1987 *Appl. Opt.* **26** 4972-4978
[6] Yu T C B et al 1994 *Summer Topical Meeting Digest on Smart Pixels* (IEEE LEOS) 6-7
[7] Data sheet on LA-1-850 Lase-Array (Photonics Research Inc.)

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 British Crown Copyright

Investigations into the use of lenslet arrays in optical signal processing and computing.

M F Lewis and R A Wilson

DRA Malvern, St. Andrews Road, Malvern, Worcs. WRI4 3PS, UK

Abstract. This paper describes several novel components and Fourier Transform - related techniques for the generation of very large uniform arrays of spots. One technique also forms a valuable assessment tool for arrays of components such as lenslets.

1. Introduction to array generators

The generation of large arrays of optical spots is of great importance for many optical signal processing and optical computing applications such as the matrix vector multiplier.[1] Several different array generation techniques exist such as lenslet arrays[2] and Dammann gratings.[3] However the generation of large arrays (~ 1000*1000) of uniform spots still proves to be a challenging problem. For example in the case of the lenslets being used in the usual manner (figure 1a) the output array is sensitive to defects and the nonuniform intensity profile of the illuminating beam. In this paper the authors describe several new methods using either lenslet arrays or a structure referred to as the *chirp Dammann grating* which are potentially capable of producing large uniform array of spots.

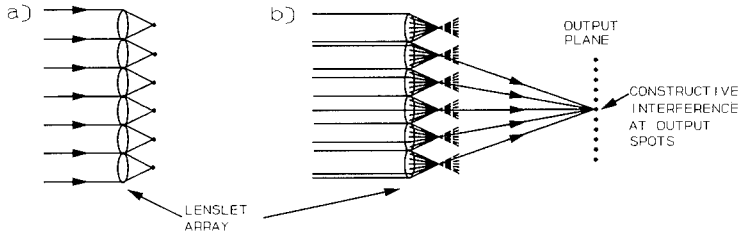


Figure 1 Spot array generation using lenslets in a) the direct mode and b) a self-imaging mode.

1.1 Self-imaging/Talbot effect for defect correction

Lenslet arrays illuminated with laser light exhibit a "self-imaging" effect (figure 1b) related to that recently reported by Heaton et al in integrated optic waveguides,[4] and to the

classical Talbot effect.[5] The self-imaging effect can be used to produce replicas of the lenslets' focal plane with different multiplicity ie number of spots and spacing.[6] An additional advantage over the direct method of use is that local defects can be "healed". Figure 2 illustrates this for a lenslet array with one lens blacked out; Figure 2a shows the focal plane with the spot missing and 2b shows a self image plane with the spot restored.

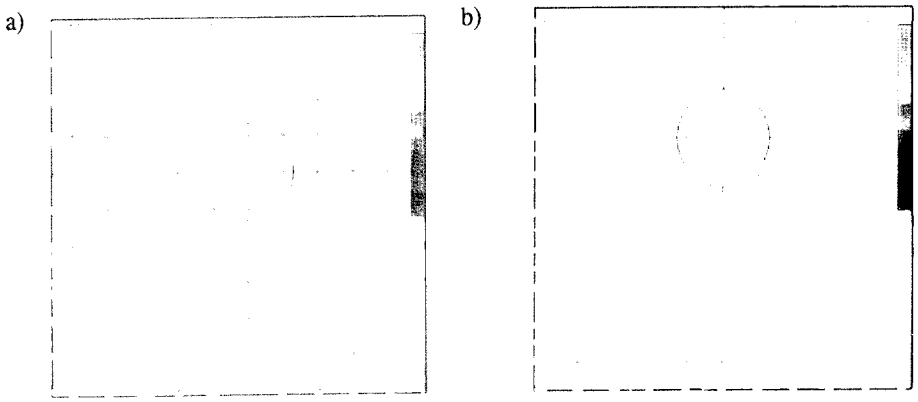


Figure 2 Intensity plots of a) the focal plane of a lenslet array with missing focal spot and b) a self-image plane with "self-healed" spot. The "healing" arises because each output spot derives from many input spots as shown in Figure 1b.

1.2 Chirp Dammann gratings

An alternative approach to the generation of uniform spot arrays is the use of Dammann gratings.[3] To date the design of such gratings has been non-analytical and confined to arrays of modest size.[7] Here a systematic approach to the initial design of such gratings can be demonstrated for the generation of *large* arrays - the result is a structure called the *chirp Dammann grating*, since it resembles a surface acoustic wave (SAW) *chirp* filter[8,9] This structure is also closely related to a one-sided classical zone-plate and phase-reversal diffractive lens. Figure 3a shows the calculated Fourier Transform (FT) of such a device, displaying regions of uniform intensity either side of the absent zero order. Note that the profile is not perfectly flat as it suffers from Gibbs phenomenon. Computer simulations show that this can be dramatically smoothed by apodising the grating. The FT of an *array* of such chirp gratings is an array of spots under the envelope; Figure 3b shows a small section of the FT of a 1x6 array of such chirp gratings.

1.3 Lenslets in the Fourier Transform mode

This line of reasoning has brought us full-circle back to the use of (eg refractive) lenslet arrays for spot array generation, not, however, used simply, or in a Talbot self-image

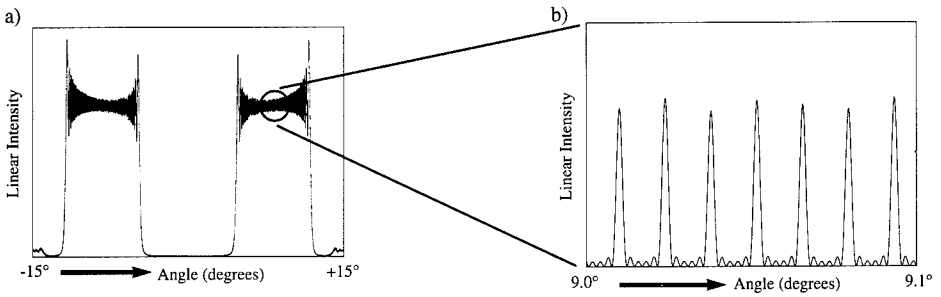


Figure 3 (a) Calculated FT of a single 1D binary phase chirp grating showing two regions of uniform intensity either side of the absent central order. (b) Detail of the calculated FT of a 1x6 array of chirp gratings.

plane, but as a replacement for the Dammann grating, ie in an FT mode, figure 4.[10] The use of lenslet arrays in this manner retains the defect self-healing property discussed above, and is potentially suitable for the generation of very large arrays, say 1000*1000, or greater. This is a consequence of the large space-bandwidth product (SBP) of the individual lenses. In fact the use of lenslets and the chirp Dammann grating contrasts with other array generator methods since they operate best for large numbers of spots, and perfectly in the limit of infinite arrays. However for finite arrays the use of computers will be necessary to optimise performance, as in the case of SAW.[8,9] A different use of this approach is as a non-mechanical means of assessing the average quality of the lenslets since the FT of an array is dependent on the profile of the individual lenses.

Calculations show that a spherical lens of square aperture provides a good approximation to a square array of uniformly intense output spots. This has been verified by measuring the FT of a convex lens masked with a square aperture, Figure 5. This shows a quasi-uniform envelope, accompanied by Gibbs phenomenon. As in the case of the chirp Dammann grating, the use of lenslet arrays results in an array of spots under this envelope, the number of which is determined by the SBP of the lenslets and not by the total number of lenses illuminated. Similarly apodisation results in more uniform spot intensities. Furthermore, and in contrast to most other techniques devised to date, this lenslet array approach does not suffer the problem of an actual or incipient zero-order (dc) spot of high/variable intensity due to manufacturing imperfections.

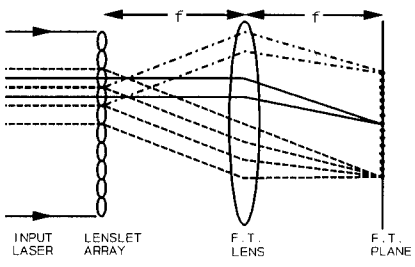


Figure 4 Schematic of large array generation using lenslets in a FT mode

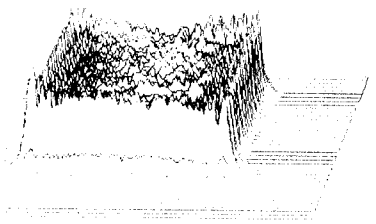


Figure 5 Measured FT (intensity) of a single lens of square aperture under laser illumination.

2. Conclusions and discussion

It is clear that the use of the FT mode and other associated modes has many advantages over other techniques for the generation of large spot arrays. For example the self-imaging phenomenon demonstrates the ability to perform local defect correction, as well as the capability to produce a multiplication in the number of spots. This has obvious attractions when considering the possible introduction of defects and dirt during lenslet manufacture and use. An assessment of the average shape and quality of lenslets in an array can be performed using the FT of the whole array. This may provide a quick and inexpensive means of assessing an array without the need to inspect individual lenslets.

The potential capability of generating *large* arrays ($>>100*100$) of spots of nominally equal intensity from *small* numbers of elements ie lenslet or chirp Dammann gratings has important implications for the manufacture of array generators. The chirp Dammann grating approach employing binary phase gratings is inherently prone to the dc spot problem as a consequence of variability in the manufacturing process. However this problem, along with interference from unwanted higher grating orders, can be alleviated by designs which isolate the first grating orders spatially to form a block of four output arrays. Such a structure may itself be a beneficial feature of the subsequent SLM, as it provides better access to the electronic addressing signals, the performance of current SLM's often being limited by addressing/multiplexing difficulties. Techniques exist to reformat such an array into one larger fully-filled array if so desired.

Acknowledgements

The authors would like to thank Mike Hutley at NPL for useful discussions on lenslets.

References

[1] Goodman J W, Dias A R, and Woody L M 1979 *SPIE* **190** 485
[2] Daly D, Stevens R F and Hutley M C 1990 *Meas. Sci. Technl.* **1** 759
[3] Dammann H and Gortler K 1971 *Opt. Comm.* **3** 312
[4] Heaton J M, Jenkins R M, Wight D R, Parker J T, Birbeck J C H, and Hilton K P 1992 *Appl. Phys. Lett.* **61 15** 1754
[5] Talbot H F 1836 *Philos. Mag.* **9** 401
[6] Bonet E, Andres P, Barreiro J C, and Pons A 1994 *Opt. Comm.* **106** 39
[7] Jahns J, Downs M M, Prise M E, Streibl N, Walker S J 1989 *Opt. Eng.* **28** 12 1267
[8] Nudd G R 1972 *Elect. Lett.* **8** 339
[9] Newton C O 1976 *Wave Electronics* **1** 387
[10] Jahns J, Streibl N, Walker S J 1989 *Proc of the SPIE* **1052** 198

© British Crown Copyright 1994/DRA
Published with the permission of the Controller of Her Britannic Majesty's Stationery Office.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical implementation of crossover an optical interconnect network using a Dammann grating with 65×65 spot arrays

Fengguang Luo, Mingcui Cao, Linqiang Huang, Hongpu Li, Jun Ai, Jun Xu and Zhaiguang Li

The National Lab. of Laser Technology, Huazhong University of Science and Technology, Wuhan 430074, P. R. China

Abstract. A Dammann grating for 65×65 spot arrays has been optimized using a simulated annealing-tempering-generalized reduced gradient (SAT-GRG) optimized algorithm and fabricated. A crossover optical interconnect network of 64×64 pixels has been performed with a Dammann grating.

1. Introduction

A Dammann grating is a binary phase grating that can be used to generate equal intensity spot arrays.[1] Recently, interest in Dammann gratings has been stimulated by the development of optical computing. In optical computing systems, the sizes of arrays (N) as high as 64×64, or more, are required. Due to complexity of optimization in the case of large spot arrays, the application of Dammann gratings will be limited.[2] In order to optimize Dammann gratings, we present an efficient optimized algorithm: simulated annealing-tempering-generalized reduced gradient algorithm (SAT-GRG). Dammann grating with 65×65 spot arrays is fabricated and used in a crossover optical interconnect network.

2. Optimization and fabrication of the Dammann grating

The principle of Dammann gratings is based on Fraunhofer diffraction theory. Dammann gratings are illuminated by plane wave from a laser source. The output pattern appears in the back focal plane of a Fourier converging lens. The amplitude of the diffraction pattern is given mathematically by the Fourier transform of the complex amplitude transmittance of the grating $G(x,y)$, where $G(x,y)$ is a binary periodic function. Since the output pattern is supposed to be separable in the x and y directions, so the complex amplitude transmittance in the x direction $g(x)$ can be given by:

$$g(x) = \sum_{-\infty}^{+\infty} A_m exp(2\pi imx) \tag{1}$$

A_m represent the amplitudes of the diffraction orders in the Fourier plane, A_m can be expressed as:

$$A_0 = 4\sum_{n=1}^{N}(-1)^{(n+1)}x_n + (-1)^{(N+1)} \tag{2}$$

$$A_m = 2/(\pi m)\sum_{n=1}^{N}(-1)^{(n+1)}sin(2\pi m x_n) \qquad m \neq 0 \tag{3}$$

the diffraction intensity of m-th order Im can be written as:

$$Im = |A_m|^2 \tag{4}$$

For the optimization, we define an objective function:

$$Merit = \sum_{m=-N}^{N}[Im(x)-I]^2 \tag{5}$$

$$I = \eta/(2N+1), \qquad \eta = \sum_{m=-N}^{N} Im(x)$$

where η is the diffraction efficiency of the Dammann grating, Im(x) is the diffraction intensity of the m-th order, I is the mean value of the intensity. In order to obtain equal intensity light spots, the intensities Im of all diffraction peaks must be equal:

$$I_0(x) = I_{\pm 1}(x) = \cdots\cdots = I_{\pm N}(x) \tag{6}$$

Considering errors produced in the fabricating process, the objective function has been rectified. Assuming the fabricating error factor is ω, the phase error is $\delta=\pi\omega/2$, then the objective function after rectifing can be writen as:

$$Merit = \sum_{m=1}^{N}[2 \times (Im - Im')^2] + (I - I_0)^2 \tag{7}$$

I_0 and Im′ are the 0-th order and higher order diffraction intensity respectively.

The simulated annealing algorithm (SA) is one of optimization, but its disadvantage is that it requires very long operating times for calculating large spot arrays. In order to reduce the calculation time and enhance the calculation efficiency, we add a tempering process to the SA algorithm, the SAT algorithm avoids too much time spent at the local minima of the merit function. Furthermore, on the base of the SAT algorithm we add a generalized reduced gradient algorithm (GRG) with much higher precision. With the SAT-GRG algorithm, the structure of a 65×65 spot array of Dammann grating has non-uniformity of 0.4% and diffraction efficiency of 82%, as shown in Fig. 1.

In order to fabricate Dammann gratings for 65×65 spot arrays, the VLSI technique is used in our experiment. By depositing a Si_3N_4 film on a quartz glass substrate using a PECVD system to achieve the desired thickness, then using a photolithographic mask and reactive ion etching, the phase pattern structure of the Dammann grating is transfered onto the Si_3N_4 film. The phase etching depth is given by:

$$d = \lambda/2(n-n_0) \tag{8}$$

where n is the index of Si_3N_4 film and n_0 is the index of air. The photograph of the Dammann grating is shown in Fig. 2.

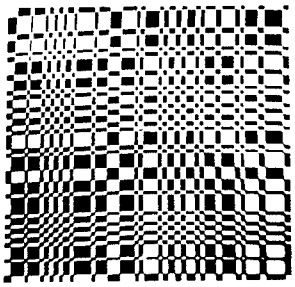


Fig. 1. The structure of 65×65 spot array Dammann grating

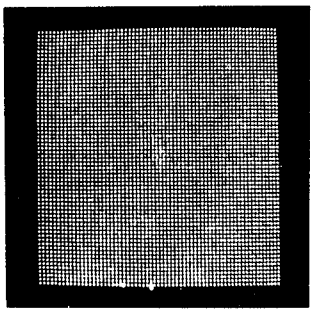


Fig. 2. Photograph of 65×65 spot array

3. Optical implemention of a crossover network with 64×64 pixels arrays

The crossover interconnection network is one of most important interconnection networks[3]. A crossover network with 2N channels has $n=\log_2 N$ stages of interconnection links. The 2N channels can be classified in two groups, one group with N channels corresponds to a straight connection, the other N channels accomplish cross permutation. The crossover optical interconnection network can be implemented in relatively simple, low loss and economical free-space optical hardwares. It can be used in an opto-electronic hybrid parallel computer system to connect large numbers of electronic processors. This system has the advantages of both the high intelligence of electronics and the massive parallelism of optics, which could enhance the speed of current computers.

The optical setup of one stage of the optical crossover network in our experi-

ment is shown in Fig. 3. The beam from the laser is projected onto the Dammann grating(D) through a combination lens(L_1). In order to demonstrate exchange in the output pattern, a mask with character "E" is placed on the focal plane of lens(L_2), where an input pattern of character "E" of 64×64 pixels is produced. The beam-splitter (BS) splits the input pattern into two different paths. The path E-P_2-CCD with a mirror in the P_2 plane implements the straight connection and the path E-P_1-CCD implements the cross connection. The output pattern of the $i=0$ stage for cross connection can be achieved by placing a reflecting 90^0 prism in the P_1 plane. For implemention of the output pattern of the i-th stage cross connection, we can place a prism grating of period 2^i in the plane of P_1. Using six such setups, we can complete the crossover network interconnection function for 64×64 spot arrays.

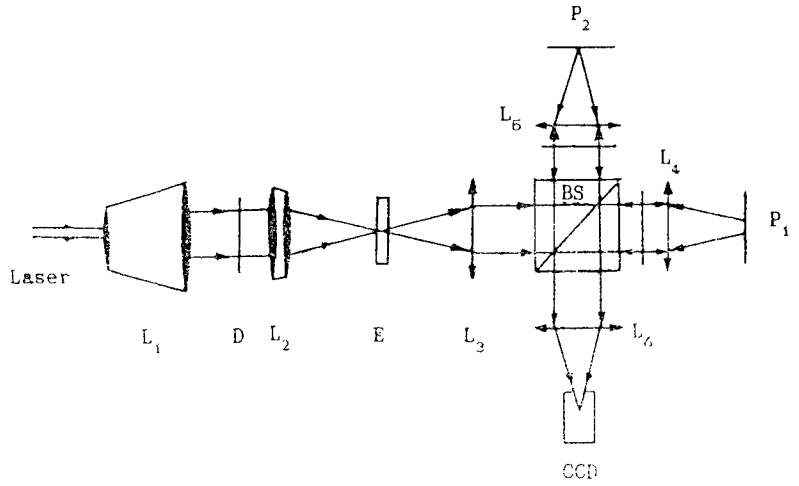


Fig. 3. Optical setup of one stage of crossover network

4. Conclusion

A Dammann grating for generating a 65×65 spot array has been optimized using the SAT -GRG algorithm and fabricated by thin film deposition, photolithography and reactive ion etching. A crossover optical interconnection network for 64×64 spot arrays with a Dammann grating has been performed. This experimental setup shows the advantages of equal optical path length and no light energy loss.

References

[1] Dammann H and Gortler K 1970 *Commun.* **3** 312-5
[2] Jahns J 1989 *Opt. Eng.* **28** 1267-71
[3] Jahns J and Murdocca M J 1988 *Appl. Opt.* **27** 3155-60

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Demonstration of a 3D integrated refractive microsystem

Jörg Moisel and Karl-Heinz Brenner

Physikalisches Institut der Universität Erlangen-Nürnberg, Lehrstuhl für Angewandte Optik, Staudtstraße 7/B2, 91058 Erlangen, Germany [1]

Abstract. A 3D integrated microsystem which uses only refractive components (planar microlenses and microprisms) is demonstrated. The system performs the overlay of two dataplanes. The output plane is $400\mu m$ x $400\mu m$ and contains 8 x 8 data channels.

1. Introduction

3D integrated microsystems promise a higher packaging density and complexity for optical interconnects and functional elements compared to 2D OEICs (eg. Brenner 1991, McCormick 1993, Özaktas/Goodman 1994). However, the additional degrees of freedom demand more efforts for aligning and mounting the system. Two different approaches can be distinguished: the use of diffractive components has the advantage that different functions (eg. focussing and aperture division) can be easily combined in one element and only one fabrication technique is needed. On the other hand, diffractive elements suffer from a high sensitivity to wavelength variations and the fabrication technique generally requires facilities with submicron alignment accuracy. Refractive elements are much more insensitive to wavelength variations and can be produced in many cases by very simple means which are standard in a optical workshop, but usually different techniques are used for the production of different elements and these elements have to be combined in order to build a system. So far, mostly diffractive systems were demonstrated (eg. Brenner/Sauer 1988, Jahns 1990, Acklin/Jahns 1994). We present a very simple refractive microsystem and demonstrate that the required technologies are available and manageable.

2. Design of the system

The function of the system under investigation is to interlace two dataplanes (Fig. 1). Possible applications of this functionality are symbolic substitution algorithms or interconnection networks (eg. Brenner *et al* 1986, Miller *et al* 1992). Depending on type, number and relative position of the components that are to be used, there are different possibilities for the optical realisation (Fig. 2). Approach a) performs the

[1] Tel.: +49 9131 858378, Fax.: +49 9131 13508, e-mail: jm@ao.physik.uni-erlangen.de

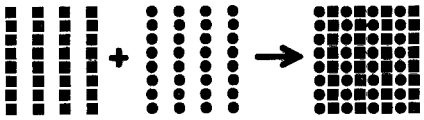


Figure 1. Overlay of two dataplanes

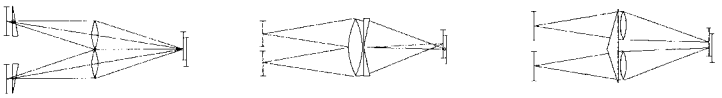


Figure 2. Three possibilities for the optical realisation

overlay without prisms but requires prisms in the input plane for high light efficiency and homogenous illumination. Approach b) uses aperture divison. It is a feasible approach in the macrooptic case but is not scalabel to microoptic dimensions out of two reasons: First, the space-bandwidth-product SBWP scales with the square of the lens diameter, resulting in typical values of order 10^3 in the microoptic domain. Therefore it is not favourable to reduce the SBWP further due to the aperture division. The second problem is that the prism edge may be rounded due to the fabrication process. If the edge runs across the center of the lens, it will reduce the image quality by stray light. Approach c) avoids both problems by shifting the prism edge between two lenses. This also results in lateral alignment tolerance between lenses and prisms because the critical design parameter is the prism angle only. Approach c) and a) are equivalent in terms of SBWP and numbers of optical elements but not in terms of manufacturing simplicity. Using approach c), the prisms and lenses can be aligned and fixed resulting in one single element whereas in case a) the prisms must be aligned with respect to the lenses and to the input plane. In the subsequent experiment, approach c) was chosen.

3. Fabrication

3.1. Lenses

The lenses are planar gradient index microlenses. These lenses were fabricated by the Na-Ag ion exchange in glass which has been reported elsewhere (eg. Brenner *et al* 1992). This process allows the fabrication of lenses with f#=5 and less than a quarter wavelength of spherical aberration. In the experiment, lenses with diameter $D = 200\mu m$, focal length $f = 1100\mu m$ (in air) and lens pitch $p = 400\mu m$ are used.

3.2. Prisms

The prisms were made by thermal moulding and casting (Fig 4.). First, a glass master of the prism is made with standard techniques. The glass master is then heated, embossed in a PMMA substrate and cooled down (a). Repeating the process generates a regular

array of identical negativ prisms (b). The negativ array is now used as master for the production of the final prism array. For this purpose, it is filled with UV-curing glue and covered with the planar microlens array (c). Both arrays are aligned under a microscope and exposed with UV. After curing of the glue, the negative is separated from the system (d). Figure 4e) shows an interferometric measurement of a section of a negativ prism.

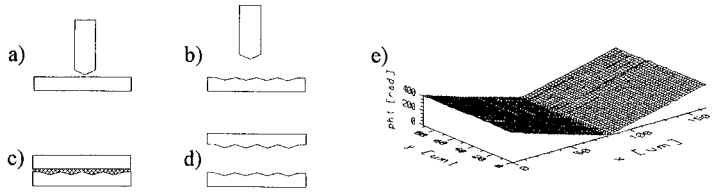


Figure 3. Fabrication of the prisms by thermal moulding and casting

4. Experimental setup

A chromium mask which contains circular and rectangular holes is used as input of the system. The mask is illuminated by two LEDs. Each LED is demagnified by a microscope objective and imaged in the plane of the microlenses. By this we can switch both channels independently. The output plane is magnified by a second microscope objective and imaged on a CCD-camera. The camera is connected to a frame grabber in order to provide facilities for quantitative measurements. The LEDs can be replaced by a halogen lamp and colour filters

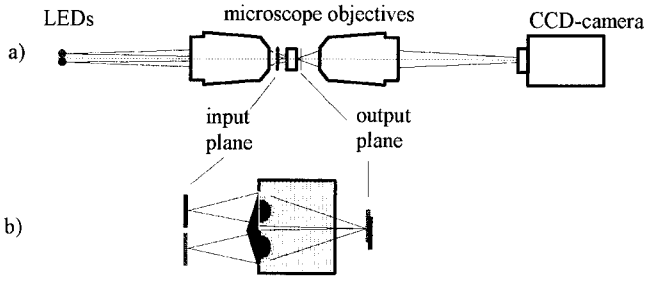


Figure 4. Experimental setup: a) overall setup, b) microsystem.

5. Experimental results

Fig. 5 shows camera images of both channels switched independently and together which demonstrates the required functionality. Table 1 gives the related data, for the S/N-ratio a detector area (40μm x 40μm) and detector separation (10μm) equal to the mask was assumed.

Table 1. Experimental data

size of output plane	400μm x 400μm
# of data channels	8 x 8
size of rectangles/circles	40μm x 40μm / 40μm diam.
separation of data channels	10μm
S/N-ratio	17:1 (450nm), 11:1 (633nm)
light efficiency (excl. Fresnel losses)	89%
peak/channel homogeneity	7%/4%

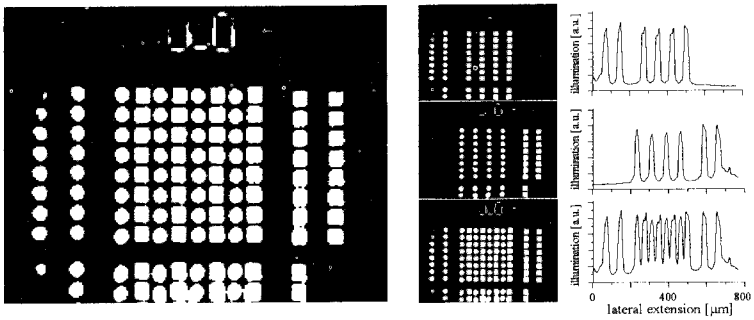


Figure 5. Output of the microsystem. Enlarged CCD image of the overlaid data planes and CCD images of independently switched channels together with related linescans.

6. Acknowledgements

Parts of this work are funded by the german ministry of research and technology (BMFT) under grant TK 0585/5.

References

Acklin B and Jahns J 1994 *Appl. Opt.* **33** 1391–7

Brenner K.-H., Huang A and Streibl N 1986 *Appl. Opt.* **25** 3054–60

Brenner K.-H. and Sauer F 1988 *Appl. Opt.* **27** 4251–4

Brenner K.-H. 1991 *Proc. Conf. Optical Computing, Salt Lake City* pp 25–28

Breuner K.-H., Moisel J, Singer W, Sinzinger S, Spick T and Testorf M 1992 *Proc. ICO Topical Meeting on Optical Computing, Minsk* pp 234–42

Jahns J 1990 *Appl. Opt.* **29** 1998

McCormick F 1993 *Proc. Int. Conf. Optics, Budapest* pp 352–6

Miller A, Cheng L, Sawchuk A 1992 *Proc. Conf. Optical Computing, Minsk* pp 129–139

Özaktas H and Goodman J 1994 *Appl. Opt.* **33** 2968–87

Inst. Phys. Conf. Ser. No 139: Part II .
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Planar integration of free space microoptical systems with refractive elements

Werner Eckert[1], Karl-Heinz Brenner and Christoph Passon

Universität Erlangen–Nürnberg, Physikalisches Institut, Angewandte Optik, Staudtstr. 7/B2, 91058 Erlangen, Germany

Abstract. Architectural concepts for the planar microintegration of digital optical data processing systems are presented. From these concepts a microoptical system consisting of stacked layers of refractive microcomponents is derived. The optical system is folded using off–axis imaging. The feasability of this configuration is analysed theoretically and prooved by experimental results.

1. Introduction

Microintegrated systems for digital optical information processing on 2D data planes impose a series of requirements on the optical realization. Light efficiency is favourable to achieve low bit–error rates with the available optical power. For constructing complex systems the design has to be modular. It should be possible to cascade an arbitrary number of processing stages. Only a few technologies should be necessary for the production of a microintegrated system.

2. Fourier stages

Optical Fourier transform stages are necessary in free space optical systems to perform spaceinvariant operations. Single lens $2f$ Fourier stages (fig. 1) exhibit certain disadvantages when used in microoptical setups consisting of arrays of identical components. A lens diameter of two times the pupil diameter is necessary to avoid vignetting. In an integrated system, the $2f$ configuration requires an additional spacer in order to place the lens in the center. From the fabrication point of view it would be more desirable to place the lenses at the substrate's surface. A more general two–lens configuration is shown in fig. 2, which is identical to the $2f$ system when b becomes zero. This setup is chosen symmetrically, because the space–bandwidth product reaches a maximum, as was shown in [1]. Assuming symmetric configurations, the number of resolvable pixels

[1] E-mail: we@pi5sun1.physik.uni-erlangen.de

is maximal when $b = f$ and $a = 0$ [4]. This represents the 'light pipe' configuration. In practical implementations, a small distance a is often required for technological reasons. With microoptical systems, space-bandwidth is a critical parameter. The distance b should therefore be chosen close to f to obtain maximal resolution.

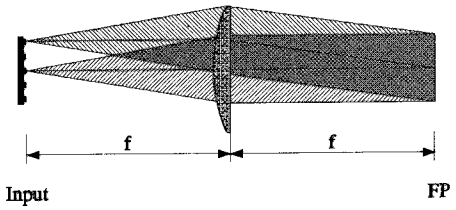


Figure 1. $2f$ system

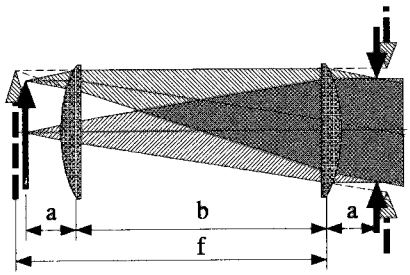


Figure 2. Two–lens Fourier setup

3. Pupil division

Digital optical systems often require the generation of multiple copies of a data– plane and the superposition of different dataplanes. For these split and join operations pupil division is most advantageous since the join operation can be performed almost lossless for an arbitrary number of joins. Pupil division may be performed either by splitting the aperture of a single lens or by using a lens for each channel in the pupil plane. In microoptical setups, multi–lens pupil division is preferable, because the full space-bandwidth product of the light pipe is available in all channels. Additionally, multi–lens pupil division can also be used for space invariant nearest neighbor interconnections.

4. Layer architecture

A transfer of these considerations to a micro–integrated system is shown in fig. 3. This layer architecture [2] consists of two modified lightpipes with multi–lens pupil division. Precise alignment can be achieved if each layer is produced with lithographic precision. The light pipes can be realized on one substrate with microlens arrays on both sides. Mask layers are introduced on additional spacers. The prism layers, realized by embossing, perform space invariant split and join operations. Since for thermal and practical reasons, active devices should be located at the system surfaces. In most cases active devices on one surface are sufficient. Cascading of several stages is then realized by placing a mirror layer on the top system surface.

4.1. Folded systems

The most compact setup is acheived when the optical system is folded. The lenses have to be used in an off–axis way in this case [3]. Only a single layer of microlenses, a layer of active components and a layer of reflective components for space invariant (shift, split and join) operations in the Fourier domain is required. Structured mirrors in the plane

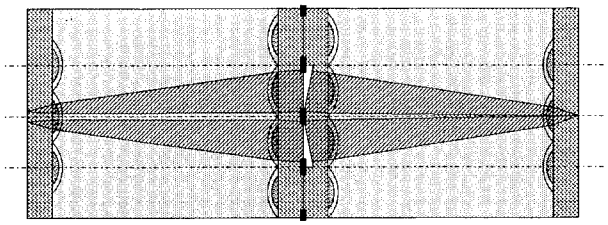


Figure 3. Layer system with lenses, prisms and masks

of the active components may be used for space variant operations. These mirrors can be produced together with the electical connectors for the active components on top of the microlens array. The active components may be precisely aligned to these structures via flip–chip bonding. The tilted mirrors in the Fourier planes are produced by embossing in polymer materials and thin film coating.

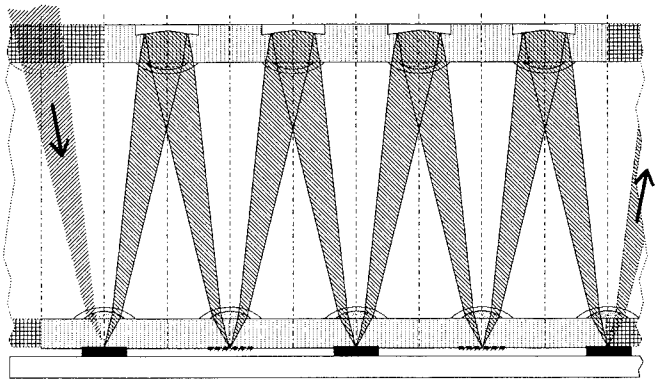


Figure 4. Microsystem with a single lens layer

4.2. Off–axis imaging

Multi–lens pupil division and folded optical systems require off–axis imaging. The resulting aberrations strongly depend on the type of lenses used and the location of the field stops in the Fourier planes and the lens shape. To get maximum space–bandwidth product the stops should be placed close to the lenses (Section 1). The off–axis aberrations Coma and astigmatism can be avoided by placing the stop in the center of a spherical refractive index distribution. For aspheric distributions or flat diffractive lenses full correction is acheivable for a single field position with astigmatic lenses. Curvature of field causes mainly a rotation of the plane of best focus around the center of the image plane of 2θ if the lights incident angle is θ. The field size is limited by the resulting defocus. The defocus inside the field must be smaller than the focal depth of the imaging system. Distorsion is very critical in digital optical systems. It does not occur in this symmetrical systems when astigmatism and coma is corrected.

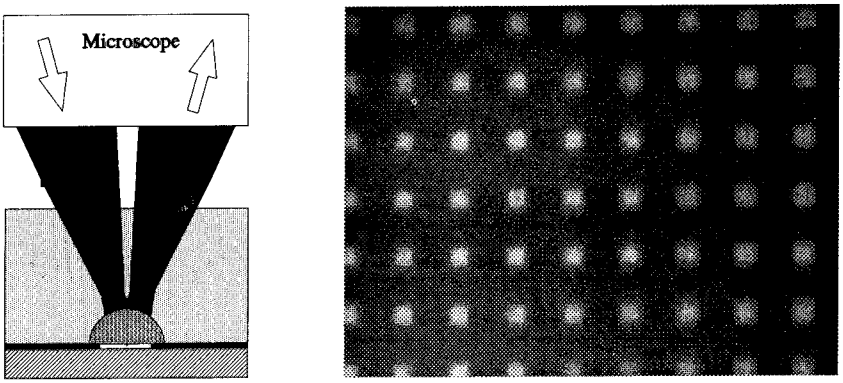


Figure 5. Off-axis imaging with a spherical microlens

5. Experimental verification

The off-axis imaging setup is shown in fig. 5. It consists of a ion exchanged microlens, with spherical refractive index distribution ($f = 2.4mm$ in glass), an aperture stop ($D = 200\mu m$) and a mirror. A mask is projected in the front focal plane of the microlens. This object is imaged by the microlens back into the same plane. This image, observed with a microscope is shown in fig. 5. The image with dimensions of $400\mu m x 300\mu m$ is free of coma and astigmatism as expected from the theoretical analysis.

6. Conclusion

From the architectural requirements for digital optical processing and the techical data of existing microcomponents a layer architecture was derived. A single layer of microlenses with active components and structured mirrors on one side and tilted mirrors on the other side, is sufficient, to perform all space variant and space invariant operations required for optical processing systems. The considerations on aberration correction were proven by an imaging experiment.

Acknowledgments

The authors want to thank J. Bähr, J. Moisel and M. Testorf for many fruitful discussions and suggestions. This work was partially founded by the German ministry of research and technology (BMFT) under grant BV 225/9.

References

[1] Lohmann A 1991 *Opt. Comm.* **86** 365-370

[2] Iga K, Kokubun Y and Oikawa M Fundamentals of microoptics Academic Press Tokio 1984

[3] Jahns J, Acklin B 1993 *Opt. Lett.* **18** 1594-1596

[4] Brenner K-H, Eckert W and Passon C 1994 *Opt. Laser Technol.* accepted

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Direct laser writing of planar Fresnel elements for optical interconnects

M T Gale, M Rossi and R E Kunz

Paul Scherrer Institute, Badenerstrasse 569, CH-8048 Zurich, Switzerland

G L Bona

IBM Research Center, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland

Abstract. Phase-matched Fresnel elements have been fabricated for laser array to fibre-cable coupling, fan-out and beam shaping operations. The continuous-relief planar microoptical elements are produced by direct laser writing in photoresist and copied by replication techniques.

1. Introduction

Parallel optical links are attracting increasing interest in the fields of data communications and optical computing. Typical applications include lens arrays for light source to fibre array coupling and fan-out elements for parallel, high-speed, high-capacity interconnects for optical processors and data links. This paper describes progress in the realisation of Phase-Matched Fresnel Elements (PMFEs) for such applications. These computer-generated Diffractive Optical Elements (DOEs) are designed with segment profiles and phase steps optimised to maintain proper phase relationships at the design wavelength [1]. By suitable choice of phase step and segment dimensions, they can be designed to combine the advantages of geometrical and diffractive optical components and their optical characteristics can be considered as resulting from a combination of refractive and diffractive behaviour.

PMFEs are fabricated as surface-relief microstructures by direct laser writing in photoresist and can be reproduced by replication techniques. They can be produced with very high numerical aperture, arbitrary aperture shape and as large area, close-packed arrays. Examples of laser diode to fibre imaging lens arrays with high numerical aperture and of focusing Nx1 fan-out elements are described.

2. Fabrication by direct laser writing

The fabrication of phase-matched Fresnel elements with continuous-relief surface profiles represents a challenging area of modern optical fabrication technology, in particular for high aperture lenslets with segment sizes in the order of micrometers. Laser writing (Fig. 1) offers a highly flexible approach for the fabrication of such DOEs.

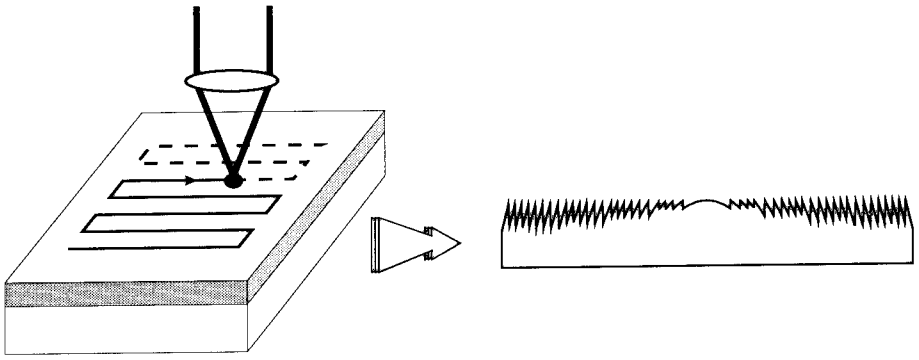


Fig. 1. Continuous-relief phase-matched Fresnel elements are fabricated by direct writing in photoresist. Segment profiles and phase steps are optimised to maintain proper phase relationships at the design wavelength.

Technology for direct laser writing in photoresist has been developed over a number of years at the Paul Scherrer Institute in Zurich. The system, described in more detail elsewhere [2], uses a HeCd laser to expose a photoresist coated substrate which is raster scanned under the focused beam using a high precision xy-stage. The surface-relief microstructure is defined in a data file which can be generated or computed by a variety of standard or custom design procedures. This data is used to control the exposure of a resist-coated substrate which is raster scanned under the focused HeCd laser beam. The laser beam intensity is synchronously controlled (256-level gray scale) to write a fully programmable, 2-dimensional exposure pattern. Development of the resist then results in a 3-dimensional continuous-relief microstructure with the desired surface profile. Continuous-relief PMFE microstructures, typically with minimum feature size of 5 μm, up to 5 μm depth and up to about 50 x 50 mm^2 in size, are routinely fabricated with this system.

PMFE microstructures can be characterised by the maximum relief depth and the minimum segment size. In laser writing technology, the maximum depth is given by the resist layer thickness, typically ~ 5 μm. The minimum segment size at the perimeter of the lenslet area is determined by the numerical aperture (NA) of the lenslet, together with the phase (height) step at the segment boundary. The minimum segment size can be maximised by choosing the largest phase step possible within the limits of the resist thickness - a 0.5 NA lens with 8π phase step requires a resist thickness of at least 4.4 μm. On the other hand, the performance of structures with high phase steps is more sensitive to errors in the relief profile. This leads to microstructures of the type shown in Fig. 1, in which the phase step is increased by 2π each time the segment size approaches the limiting value as the radius increases. Lenslets with high NA (~>0.5) have a dominantly diffractive behaviour and function best with monochromatic (laser) or narrow band (LED) illumination.

Resist surface-relief microstructures are electroformed to a Ni shim and reproduced by replication in plastic or epoxy materials. Replication technologies which are commercially available and currently under investigation include hot embossing, injection moulding and uv-replication techniques.

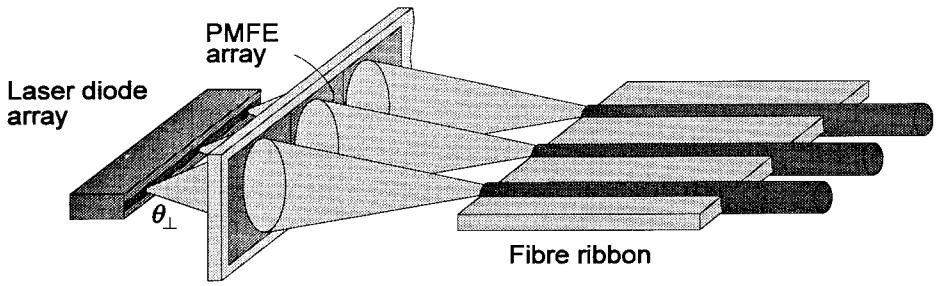


Fig. 2. PMFE array for the coupling of a laser diode array into a fibre ribbon cable.

3. Fresnel lens arrays for laser to fibre coupling

High numerical aperture, planar microstructures and flexible design in shape and array formats are attractive features for the application of PMFEs in the coupling of laser diode arrays to fibre ribbon cable. An example (see Fig. 2) is a PMFE array which has been designed for coupling an array of laser diodes ($\lambda = 831$ nm) into a ribbon of 12 fibres. The laser diodes each emit an astigmatic beam with divergence angles $\theta_\perp = 8°$ and $\theta_{//} = 28°$ (FWHM) in the horizontal and vertical planes. The MT-connector compatible ribbon cable consisted of 12 multimode fibres with 0.21 NA and 50 µm core diameter. The PMFE fabricated achieved a measured efficiency of 60% with an object side NA of 0.5 [3].

Typical laser diodes with an elliptical near field and longitudinal astigmatism produce an elliptically shaped image at the fibre plane. Anamorphic PMFE arrays have also been designed and fabricated for beam-shaping and circularising the image to obtain a better match to the circular fibre geometry. Fig. 3 shows the measured irradiance distribution at the fibre plane for such an anamorphic PMFE. The $1/e^2$ radius of the circularised image is about 16 µm.

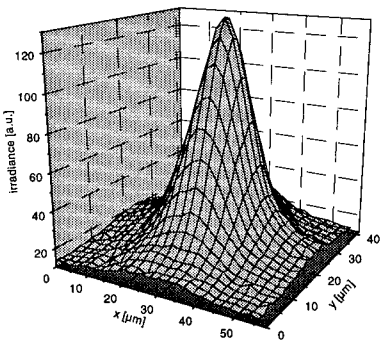


Fig. 3. Measured irradiance of a circularised image at the fibre plane.

4. Spatially interlaced lenses for fan-out applications

Novel PMFE structures have been designed to perform focusing fan-out functions (see Fig. 4.) in a single planar microoptical element [4]. The desired fan-out function is implemented by combining different PMFEs (e.g. one for each interconnection channel) in a special area sharing arrangement. Each PMFE is divided into a subarray structure, leading to an array of focused diffraction orders centred around the focal point of the basic PMFE. The period of the different subarray structures is chosen such that the diffraction orders coincide with and are coherently superimposed upon the desired image points. Simple and fast procedures have been developed for optimising such fan-out PMFEs, based upon a low number of well-defined physical-optical parameters. The concept can easily be extended for realising 2D fan-out elements.

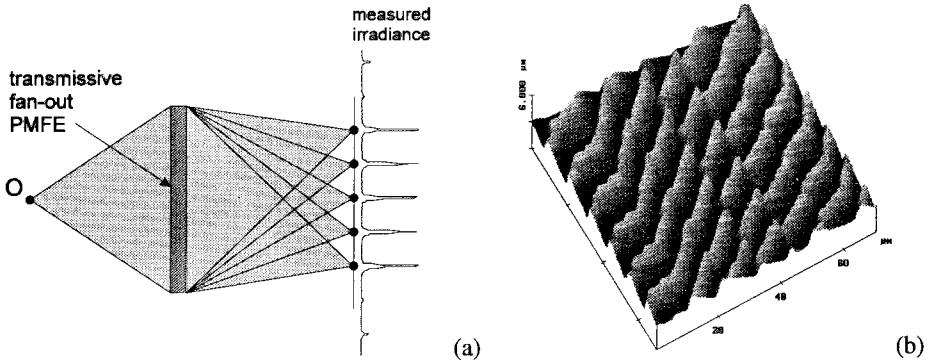


Fig. 4. (a) Focusing fan-out function using interlaced PMFEs and measured irradiance.
(b) AFM image of an interlaced PMFE microrelief (70 μm x 70 μm).

A major advantage of the PMFE fan-out approach is the large tolerance with respect to fabrication errors. In a series of experiments in which depth scaling errors over a range of ±20% were introduced, PMFE fan-outs showed a uniformity error of < 3%. This is much superior to the performance of conventional surface relief fan-out elements for which a uniformity reduction of ~10% typically results from a depth scaling error of only a few percent. PMFEs with a 5x1 fan-out have been fabricated and evaluated. Figure 4b shows an Atomic Force Microscope (AFM) image of the surface of such a fan-out PMFE (~1 μm relief amplitude) with the subarray structure clearly visible. The measured uniformity error was about 2%.

5. Conclusions

The direct laser writing of planar Fresnel elements has been developed for the flexible and reproducible fabrication of microoptical elements. Novel optimal imaging properties are obtained using Phase Matched Fresnel Elements (PMFEs) in which the segment profiles and phase steps are optimised for the given wavelength and imaging configuration. Examples have been given of PMFE arrays for the coupling of diode laser arrays to fibre ribbons, including anamorphic elements for circularising the image irradiance, and of PMFEs implementing fan-out functions by spatially interlacing different elements. Many other novel imaging and beam shaping functions can be realised.

Acknowledgements

Illuminating discussions with Prof. R. Dändliker and the experimental assistance of C. Appassito, J. Pedersen, H. Schütz and R. Stutz and are gratefully acknowledged. This work was supported in part by the Swiss Priority Program OPTIQUE.

References

[1] R.E. Kunz and M. Rossi., "Phase-matched Fresnel elements", Opt. Comm. **97**, 6-10 (1993).
[2] M.T. Gale, M. Rossi and H. Schütz, "Fabrication of continuous-relief microoptical elements by direct laser beam writing in photoresist," Proc. SPIE, **2045** (1994).
[3] M. Rossi, R.E. Kunz and G.L. Bona, "Phase-matched Fresnel elements of the optical interconnects", OSA Technical Digest Series, Vol. 11, 321-324 (1994).
[4] M. Rossi, G.L. Bona and R.E. Kunz, "Focusing fan-out elements based on spatially interlaced phase-matched Fresnel elements", accepted for publication in Opt. Comm. (1994).

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 British Crown Copyright

Strategies for realising large Computer Generated Holograms from the calculation of sub-holograms.

R.I. Young[1], A.E. MacGregor[2] & C.W. Slinger[1].

[1] DRA. St Andrews Road. Malvern. Worcs. WR14 3PS.
[2] Dept EEEP. Northumbria University. Ellison Place. Newcastle upon Tyne. NE1 8ST.

Abstract. Several schemes that decrease the calculation time of a computer generated hologram through utilising sub-holograms are presented. The applicability of these schemes to differing encoding schemes and design criteria is discussed.

1. Introduction

The performance of a CGH is restricted by it's limited information capacity. This limit arises either from the difficulty in calculating the CGH or from the resolution of the fabrication device used to realise them. For example calculating the hologram using the directed binary search, DBS[1], algorithm is computationally very intensive due to the number of points which have to be tested during the search.

For CGH too large to be conveniently directly determined an alternative design strategy is to interlace several small CGH together. This poster presents a new algorithm involving the partitioning of the CGH into sub-holograms and then calculating these sub-holograms.

2. Interlacing methodologies

Two patching methodologies have appeared in the literature. The first approach is to repeat the basic hologram pattern periodically, which turns the initial calculated structure into a sub hologram and increases the SBWP of the final hologram without adding any extra sampling data to it. This technique is trivial in that it requires no more calculation and results in a punctuated replay with reduced speckle noise[2].

The second technique is the Iterative Interlaced Technique, IIT, of Ersoy et al.[3] which iteratively determines the sub-holograms prior to combination. By manipulating the encoding errors of the sub-holograms the algorithm improves the fidelity of the large hologram's replay. The IIT is applicable to all CGH encoding schemes and offers higher fidelity replays and calculation times than by realising a single large hologram.

Using an interlacing algorithm allows the realisation of holograms faster than from a single large hologram. This approach is faster because there are fewer possible pixel permutations to compare in each sub hologram to evaluate it's structure and there are less points to Fourier Transform during the calculation process.

The general applicability of the HT methodology is limited due to restrictions placed upon it by the calculation. For example since it uses a complex amplitude based cost function (to allow constructive interference between the sub-holograms at the replay plane) no phase freedom[4] is available. Furthermore there is also a reduction in the available amplitude freedom[3,4] to avoid the overlapping of shifted images. The patched CGH will have a low efficiency because the resultant replay is the average of all the sub-holograms replays.

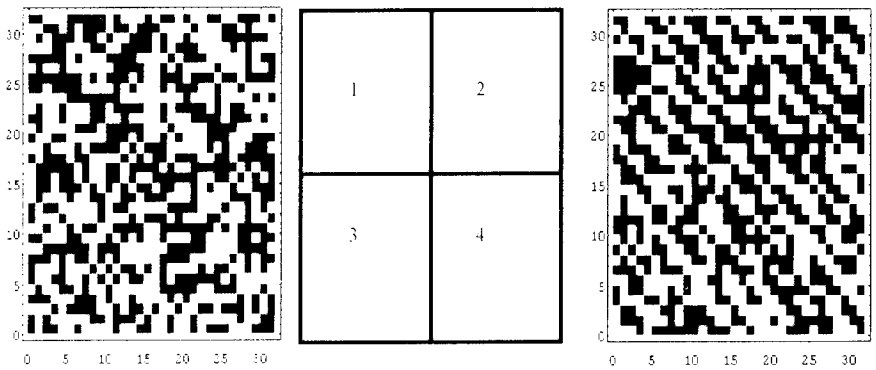
3. Partitioning methodologies

If the effects of all the sub-holograms were considered simultaneously an intensity based cost function could be employed with more amplitude freedom available to design the CGH. This is the basis for the partitioning algorithm, here demonstrated with a modified DBS algorithm. We refer to the partitioning version of DBS as PDBS.

The accelerated convergence of PDBS in comparison to DBS is due to the minimisation of pixel interaction in the encoding process as explained in the discussion section. The FFT carried out in this design process is for the full CGH.

3.1. PDBS described

PDBS starts by determining an initial structure for the entire hologram. The hologram is then partitioned, for the demonstration calculations a 32*32 pixel CGH is divided into four 16*16 pixel sub-holograms. The structure of each sub-hologram is then tested using a normal DBS methodology. Several partitioning approaches have been tested, two are related here, the difference between them being that in one each sub hologram is fully determined before the next sub hologram is optimised, PDBS, whilst in the other a *single* DBS iteration is conducted for each sub-hologram in turn, PDBS2.



(1) Determine the initial structure for the CGH

(2) Partition this CGH into 4 sub-holograms & undertake DBS upon each

(3) To give the final CGH

Figure 1. Outline for PDBS

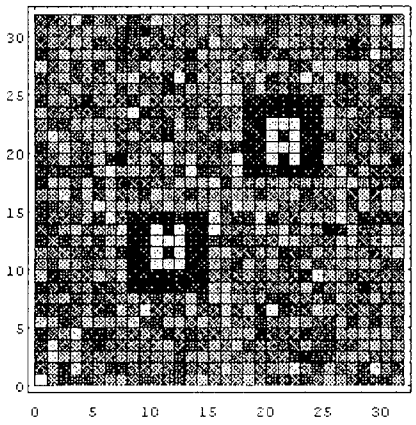
3.2 Experiments

To compare the performance of PDBS to DBS several CGHs were made to replay an intensity based target. This was undertaken using random pixel configuration starts and also optimum starts as found using Wyrowski's IFTA algorithm[5]. The performance of the hologram is measured in terms of the replay error.

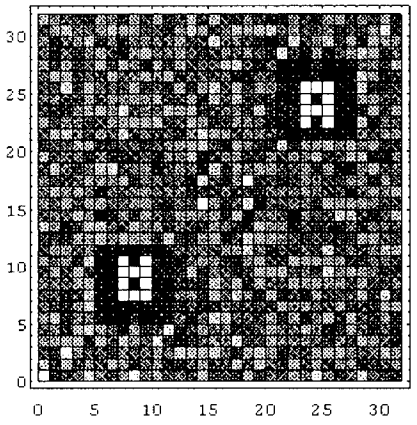
3.3 Results

Table 1. Calculation times and resultant replay errors for CGHs designed using different versions of the DBS algorithm.

Algorithm	Start	Run time /s	Replay error
DBS	Random	5418	0.000739
PDBS	Random	2555	0.003200
PDBS2	Random	4522	0.000807
DBS	Optimum	5332	0.000467
PDBS	Optimum	4008	0.001250
PDBS2	Optimum	5010	0.000573



Replay from a PDBS CGH with four sub holograms.



Replay for a 32*32 DBS CGH from a good start.

Figure 2. Computer simulations of the replays of some of the calculated CGHs

4 Discussion

An explanation of why the partition searches curtail more quickly than for a normal DBS for the same number of pixels is as follows. DBS works to reduce the encoding errors introduced into the replay through using a binary hologram[4] by examining the interaction between pixels. The smaller the region of pixel interaction considered the larger the replay error of the final hologram. Thus limiting the amount of pixel

interaction limits the ability for DBS to find a good solution as there are fewer possible pixel configurations to consider. That is the PDBS algorithms have less solution space to search and a reduced solution space takes less time to search giving the modified DBS algorithm the appearance of an accelerated convergence.

For all of these partitioning methodologies the resulting CGH is also a solution for normal DBS (i.e. a normal DBS starting from a pixel configuration determined by a partitioning algorithm will not flip any more pixels). The PDBS results appear to be poor local minima because the PDBS fidelities are much lower in comparison to the fidelities found from DBS for holograms of the same size.

Whilst the partitioning methodology is applicable to all the major encoding schemes (i.e. DBS, error diffusion and the IFTA) it is most appropriate for DBS which has the longest calculation time. A possible application for PDBS is in an optical learning system where a SLM is used as the CGH and the Fourier transform is undertaken optically. In this regime the CGH is generally designed using DBS, applying PDBS would allow quicker realisations of the CGH.

5 Summary

The partitioning approach is a method for accelerating the design time for CGH made by the DBS encoding scheme. This acceleration is achieved by the use of sub-holograms to minimise the pixel interaction, the trade off for reducing calculation time is an increase in replay error. It is left to the user to decide upon the balance between calculation speed and replay error. Unlike earlier patching schemes the partitioning approach allows the sub-holograms calculation from an intensity cost function.

References

[1] Jennison BK. Allebach JP Sweeney DW 1989 *Optical Engineering* **28**(6) 629-637.
[2] Lee WH CGH: techniques and applications 1978 *Progress in Optics* **16** North Holland.
[3] Ersoy OK & Zhuang Z 1992 *Applied Optics* **32**(17) 6894-6901.
[4] R.I. Young Design strategies for encoding Computer Generated Holograms 1994 Ph.D. Thesis.
[5] Peter T. Wyrowski F & Bryngdahl O 1993 *J Mod Optics* **40**(4) 591-600.

Acknowledgments

This work was conducted at the University of Northumbria at Newcastle upon Tyne and funded by the DRA (formerly Royal Signal and Radar Establishment, Saint Andrews Rd, Malvern, Worcs. WR14 3PS) under contract D/ER1/9/4/2050/010/RSRE.

British Crown Copyright 1994/DRA. Published with the permission of the Controller of Her Britannic Majesty's Stationery Office.

Inst. Phys. Conf. Ser. No 139: Part II
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Three dimensional distribution of light generated by a diffractive element

Rafael Piestun†, Boris Spektor‡and Joseph Shamir†

† Department of Electrical Engineering
Technion - Israel Institute of Technology, Haifa 32000, Israel

‡Institute of Automation and Electrometry
Russian Academy of Science, Siberian branch, Novosibirsk

Abstract. A method to design diffractive elements to generate 3D light distributions is presented. Low resolution binary amplitude modulated diffractive elements are designed to generate "nondiffracting beams" and arrays of them in on-axis configuration.

Fourier and Fresnel computer generated holography are usually applied to generate a desired field intensity distribution over a specified transversal plane and sometimes on a limited depth of field.

In a recent letter [1] a method to control the wavefront propagation over long distances in the Fresnel and Fourier regions was presented. Low resolution diffractive elements including phase and amplitude were designed for generation of the so called "nondiffracting beams". In this paper we applied that method to generate three dimensional distributions of light with a binary amplitude diffractive element in on-axis configurations. Preliminary experimental results are also presented.

Applications of the technique include display of information, spatial distribution of energy, optoelectronic interconnections, precision measurements and alignment.

1. Procedure

The problem under consideration is: 'Given a known WF incident on a thin low-resolution, amplitude modulated diffractive element, design this element to obtain a desired intensity distribution within a given three-dimensional domain.' Fundamental physical limitations do not allow a solution for any arbitrary distribution. However, if a solution does not exist, it still may be valuable to derive the closest solution to the constraints.

The general procedure is as follows (see Fig. 1)

1. Define a region D behind the diffractive element and the desired intensity distribution $I(\xi, \eta, z)$ with appropriate tolerances.

2. Transform $I(\xi, \eta, z)$ into constraints on sufficiently close transversal planes within this region.

3. Impose constraints on the diffractive element (space-bandwidth, amplitude only modulation).

4. Solve the optimization problem: 'Find an input distribution which satisfies the constraints in (2) and (3) or a reasonable approximation according to certain measure'.

In reference [1] we calculate the maximal distance between consecutive transversal planes that assures that longitudinally continuous light distributions are achieved.

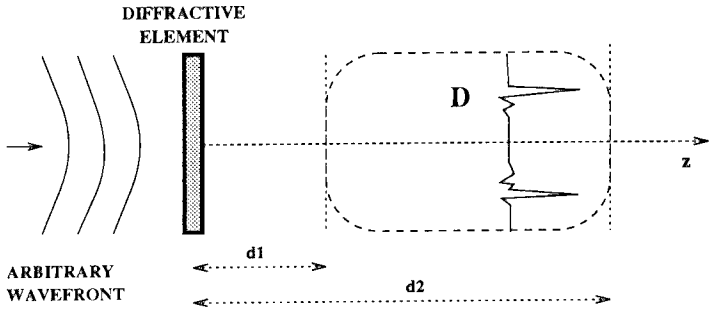


Figure 1. The basic scheme for generation of 3D light distributions

The algorithm of projections onto convex sets (POCS) adapts to the described method, leading to good designs after a few iterations.

The complex amplitude over a plane at a distance z can be described by the Fresnel transform **FrT**. The constraints prescribed on this plane can be transformed into constraints on the field at $z = 0^+$ and represented as C_z. There is a projection operator $\mathcal{P}_z$ associated with each constraint C_z.

Given a function f describing the input field, we find its projection onto the constraint C_z by calculating the field at the distance z, projecting this function onto the constraints of that plane and finally performing the inverse **FrT** to obtain the corresponding field at the input:

$$g = \mathcal{P}_z f = \mathbf{FrT}^{-1}\left\{\bar{\mathcal{P}}_z \mathbf{FrT}\,(f)\right\} \tag{1}$$

where $\bar{\mathcal{P}}_z$ represents the projection operator on the plane at the distance z.

If we take n planes we have a set of constraints $C = C_0, C_1 ... C_n$, where C_0 is the constraint at $z = 0$. In the case of amplitude modulated elements, this constraint can be expressed as follows

$$\mathcal{P}_0\{f(x,y)\} = \begin{cases} \mathbf{Re}[f] & \text{if } \mathbf{Re}[f] \geq 0 \\ 0 & \text{if } \mathbf{Re}[f] < 0 \end{cases} \tag{2}$$

In the problem under consideration, a function f is searched, which lies simultaneously in all sets C_i $(i = 0, 1 ... n)$,

$$f \in C = \bigcap_{i=1}^{n} C_i \tag{3}$$

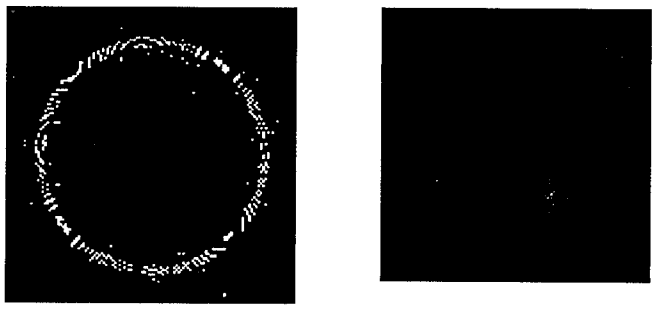


Figure 2. a) Binary amplitude diffractive element (128 × 128 cells, 100 × 100μm) for generating a "nondiffracting beam". b) Experimental result of the intensity distribution on a transversal plane at $z = 53$cm. The peak remains parallel to the optical axis at a distance 1.2 mm, along 10 cm.

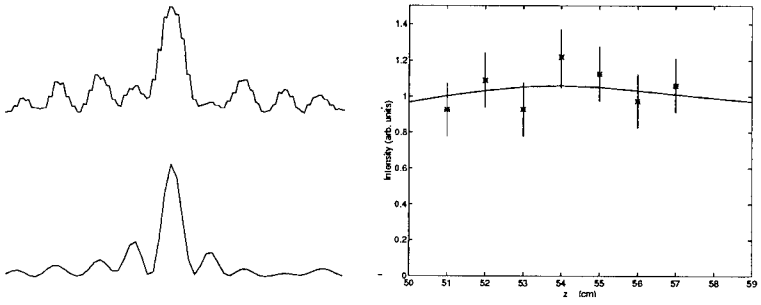


Figure 3. a) Cross section of the intensity peak at $z = 50$cm. b) Simulation result at the same distance. The peak width is similar to the diffraction limit. c) Simulation (continuous line) and experimental measurements of the peak intensity along the selected longitudinal region.

The projection onto C can be obtained by a recursive process using the composite operator $\mathcal{T} = \mathcal{P}_n \mathcal{P}_{n-1} ... \mathcal{P}_1$:

$$f_m = \mathcal{T}^m f_0 = \mathcal{T} f_{m-1} \qquad m = 0, 1, 2... \tag{4}$$

where f_0 is an arbitrary initial function.

We applied the method to design low resolution elements having 128 × 128 binary cells which work in on-axis configuration. This is attractive because there is a reduction in design and fabrication efforts and real time devices are available.

In a first experiment a "nondiffracting beam" was generated. We seeked a light distribution that presents a constant intensity peak along ten centimeters in the longitudinal direction z. The peak should be situated out of the optical axis and remain parallel to it in a dark region around the axis (see Fig. 2).

The input WF is a plane wave and the binary cells are 100μm squared. The domain D was chosen between $d_1 = 50$ cm and $d_2 = 60$ cm for $\lambda = 630$ nm. An amplitude only

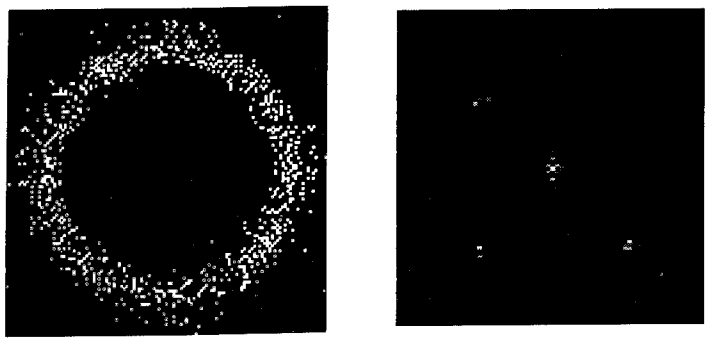


Figure 4. a) Binary amplitude diffractive element for generating a non-symmetrical array of "nondiffracting beams". b) Experimental result of the intensity distribution at $z = 52$cm.

mask was first generated and then direct binarization with an appropiate threshold was performed. The design obtained after 20 iterations is shown in Fig. 2(a). Observe that this binary mask has a reduced information content. The element was produced on chromium coated plates by photolithography.

The light distribution at different transversal planes was measured with a photodiode camera. Fig.2(b) shows a typical transversal distribution at $z = 53$cm. In Fig. 3a. we show an experimental cross section of the peak region together with the result of the simulations. The peak width is under 25μm, comparable to the diffraction limit.

The graph in Fig. 3(b) shows the experimental constancy of the peak intensity in the selected area compared to the simulations.

The second example is an array of five "nondiffracting beams" instead of only one. The binary diffractive element for a system of five peaks propagating for 5 cm is shown in Fig. 4(a). The transversal intensity at $z = 52$ cm is shown in Fig. 4(b). The distribution starts at $z = 50$cm and the peak width is 25μm for $\lambda = 630$ nm.

In conclusion, preliminary experiments indicate that the proposed approach leads to efficient solutions for various problems in accordance with the design constraints and the simulations. As examples, "nondiffracting beams" and "nondiffracting beam" arrays were demonstrated. The procedure adapts the design to the available space bandwidth and to the imposed constraints. As the method is not restricted to incident plane waves, the design of diffractive elements can be adapted to any WF such as beams coming directly from laser sources.

Acknowledgments

This work was performed within the Technion Advanced Opto-Electronics Center established by the American Technion Society (ATS) - New York.

References

[1] R. Piestun, J. Shamir, "Control of wave-front propagation with diffractive elements", *Opt. Lett.* **19**, 771-773, 1994.

III. ANALOGUE OPTICAL PROCESSING

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Theoretical results on accuracy limitations in analog optical processors

D A Timuçin, J F Walkup and T F Krile

Department of Electrical Engineering, Texas Tech University, Lubbock, Texas 79409-3102, USA

Abstract. A complete statistical analysis and modeling of a generic three-plane optical processor is described. The device models used are detailed and the output signal statistics are determined for a number of interesting special cases. The concepts of accuracy and precision are then defined and the problem of accuracy enhancement is formulated in detection- and estimation-theoretic contexts.

1. Introduction

The most important and prominent features of optical processors are their speed and parallelism potentials. However, they also suffer from some fundamental drawbacks, the most critical being low computational accuracy, that limit their application to real-world problems. In view of the intimate (inverse) relationship between the highest achievable accuracy and the maximum attainable speed and degree of parallelism, it is understood that a mastery of accuracy limitations and enhancement potentials is crucial for the full exploitation of the strengths and capabilities of optical processors.

The noise sources that limit the accuracy of an optical processor fall into two groups. On one hand, we have the system noise sources, such as diffraction, crosstalk, and background radiation, that arise from the processor architecture and operation. On the other hand, we have the device noise sources that introduce inevitable inaccuracies associated with the physical representation and interpretation of the information-bearing signals in the system. Of particular significance here are the noise in the source field intensities due to photon, excitation, and emission fluctuations, the randomness in the spatial light modulator (SLM) intensity transmittance due to transmission and polarization fluctuations, and the noise in the detectors due to nonunity quantum efficiencies, gain fluctuations, shot noise, dark current, and thermal noise. Thus, given a statistical description of the system operation as well as the statistical models for the component devices, one can utilize tools from statistical optics to establish the statistics of the system output signal. Optimal detection- and estimation-theoretic techniques can then be applied to these statistics in an effort to improve the accuracy of the processor.

In this paper, we first review our past work on statistical analysis and modeling of optical processors, establishing the overall system output statistics for a number of interesting device combinations. We then provide a quantitative definition of accuracy for analog optical

processors, first from a detection-theory perspective and then from an estimation-theory perspective. With the availability of the output signal probability density function (PDF), the highest achievable accuracy for each special case can then be determined.

2. Statistical analysis and modeling

The starting point for our efforts was a rigorous statistical analysis of a generic three-plane optical processor whose architecture is the backbone of a wide class of systems including optical correlators, optical interconnects, and optical linear algebra processors [1]. We established the statistics of the output signal for this processor without committing ourselves to a specific set of sources, modulators, and detectors. Specifically, we found the conditional characteristic function of the detector output voltage to have the form

$$\boldsymbol{M}_{V|R,\mathrm{P}}(\omega) = \exp\left(\int_0^{\infty} p_Q(q) \int_0^{T_p} [\alpha r(t) + \rho] \left\{ \exp[j\omega e q f(t-\tau)] - 1 \right\} d\tau \, dq - \frac{1}{2}\sigma_{V_r}^2 \omega^2 \right) \quad (1)$$

where e is the electronic charge, q is the random gain in the photodiode with the PDF $p_Q(q)$, $f(t)$ is the photon-to-voltage impulse response of the detection and post-processing electronics, $r(t)$ is the stochastic rate process due to the incident field, ρ is the random dark excitation rate, and $\sigma_{V_r}^2$ is the variance of the Gaussian zero-mean thermal noise voltage.

We then proceeded to insert statistical device models into our general equations, obtaining system output statistics for various combinations of sources, modulators, and detectors [2]. Specifically, we considered classical sources such as laser and light-emitting diodes, and popular detectors such as *p–i–n* and avalanche diodes. Since statistical models for most SLMs are currently unavailable, we considered an ideal static device as well as one with hypothetical random transmittance. Finally, near- and far-field free-space propagation were assumed between processor planes. Table 1 summarizes the various statistical and deterministic device models used [3–7].

The overall system modeling was performed in stages of increasing sophistication [2]. We started with an idealized system where the only realistic components were the sources, thus reproducing well-known results. We then moved on to consider realistic detectors in combination with realistic sources. The next step was to bring in the randomness due to the use of a realistic SLM. Finally, the most general form of the processor was reached by including background, crosstalk, and diffraction effects.

For each case, we first calculated the statistics (means, mutual coherence functions, and PDFs) of the field incident on the detectors in terms of those of the sources and the SLM in the processor. These intensity statistics were then used to derive the statistics of the rate process $r(t)$. Finally, the output signal statistics were obtained by removing the conditioning on $r(t)$. In the ideal detector case, we obtained the well-known negative-binomial and multi-fold Laguerre PDFs for thermal and laser sources, respectively [4], while for realistic detector cases, we were able to invoke the Gaussian approximation obtained by the Maclaurin expansion of the complex exponential in Eq. (1) [8]. The results for the realistic SLM case were then readily obtained from these expressions by simple averaging.

3. Accuracy limitations and enhancement

Computational accuracy of a processor can best be specified in terms of the signal resolution it affords at its output while simultaneously satisfying an average probability-of-error criterion.

Table 1. Description of statistical and deterministic device models.

COMPONENT	DEVICE	MODEL
Sources	Laser diodes	Phase-diffused coherent and narrow-band chaotic superposition as the output of a nonlinear van der Pol oscillator in the quasi-linear regime
	Light-emitting diodes	Circular complex Gaussian random process due to independent, spatially uniform, temporally Poisson, homogeneously-broadened emissions
Spatial light modulator	Static screen	Constant transmittance due to lack of spatio-temporal material fluctuations
	Random screen	Circular complex Gaussian random process due to transmittance and/or polarization fluctuations in the medium
Detectors	Ideal photon counter	Unity gain, rectangular impulse response, no dark or thermal noise
	p–i–n diodes	Constant dark excitation rate, unity gain, exponential (op-amp) impulse response, Gaussian thermal noise
	Avalanche diodes	Constant dark excitation rate, Yule–Furry random gain, exponential (op-amp) impulse response, Gaussian thermal noise
Propagation structures	Near-field free space	Dirac-delta point-spread functions due to geometrical-optics regime propagation
	Far-field free space	Quadratic-phase point-spread functions due to Fresnel-regime propagation

Quantitatively, this signal resolution can be defined in terms of the maximum number of identifiable signal levels L within the signal dynamic range $[V_{min}, V_{max}]$ or, equivalently, the maximum number of bits n, where $n = \log_2 L$. It is trivial, yet crucial, to recognize that, in a fixed dynamic range, accuracy (i.e., number of levels or bits) and precision (i.e., probability of error per level or bit error rate), both of which should be specified for a meaningful expression of processor performance, exhibit an inverse relationship.

The determination of the maximum number of resolvable signal levels involves the optimal quantization of the signal dynamic range. This problem can be solved either iteratively by the Lloyd–Max algorithm or recursively by adaptive dynamic programming [9]. In the case of signal-independent noise, the solution is trivial since uniform quantization proves to be the optimal scheme, resulting in uniformly distributed signal levels with equal-

size decision regions centered at these levels and decision thresholds placed halfway between consecutive levels. In our case, the problem is complicated by the fact that the noise at the processor output is signal-dependent, calling for a more sophisticated quantization scheme.

Formally, for a given maximum tolerable average probability of error per signal level, denoted P_e, and for equal *a priori* probabilities $P(v_i) = 1/L$, $i = 1, 2, ..., L$, the maximum attainable accuracy can be found by solving for the maximum value of L in the equation

$$\frac{1}{L}\sum_{i=1}^{L-1}\left[\sum_{j=1}^{i}\int_{z_i}^{z_{i+1}} p_V(v\,|\,v_j)\,\mathrm{d}v + \sum_{j=i+1}^{L}\int_{z_{i-1}}^{z_i} p_V(v\,|\,v_j)\,\mathrm{d}v\right] = P_e, \quad p_V(z_i\,|\,v_i) = p_V(z_i\,|\,v_{i+1}) \tag{2}$$

where the choice of signal levels v_i, $i = 1, 2, ..., L$, and decision thresholds z_i, $i = 0, 1, ..., L$, subject to the constraints $V_{min} = z_0 \le v_1 \le z_1 \le v_2 \cdots \le v_{L-1} \le z_{L-1} \le v_L \le z_L = V_{max}$, comprise the optimal quantization scheme [10].

Alternatively, we can formulate the problem in the context of parameter estimation theory by ascribing an *a priori* parameter PDF $p(m_s)$ to the signal we wish to estimate. Upon the observation of k samples of the output signal, creating the sample vector $\underline{v}$ with a joint PDF $p(\underline{v}\,|\,m_s)$, we can form the *a posteriori* parameter PDF

$$p(m_s\,|\,\underline{v}) = \frac{p(\underline{v}\,|\,m_s)\,p(m_s)}{\int p(\underline{v}\,|\,m_s)\,p(m_s)\,\mathrm{d}m_s} \tag{3}$$

which we ideally expect and desire to satisfy $p(m_s\,|\,\underline{v}) \to \delta(m_s - \hat{m}_s)$ as $k \to \infty$. Here $\hat{m}_s$ is the true value of the signal [10]. In this approach, accuracy can be quantified by the Cramér–Rao lower bound on the variance of the estimate [11], which offers a tradeoff between accuracy and speed.

In both approaches, the difficulties due to the signal dependence of noise at the output can be alleviated by the use of suitable normalizing transforms which remove from the noise power the dependence on the signal mean [10, 12].

References

[1] Timuçin D A *et al.* 1994 *J. Opt. Soc. Am. A* **11** 560–571
[2] Timuçin D A *et al.* 1994 *J. Opt. Soc. Am. A* (to be submitted)
[3] McIntyre R J 1972 *IEEE Trans. Electron. Devices* **19** 703–713
[4] Saleh B E A 1978 *Photoelectron Statistics* (Berlin: Springer-Verlag)
[5] Teich M C *et al.* 1984 *J. Opt. Soc. Am. B* **1** 366–389
[6] Goodman J W 1985 *Statistical Optics* (New York: Wiley)
[7] Saleh B E A and Teich M C *Fundamentals of Photonics* (New York: Wiley)
[8] Papoulis A 1991 *Probability, Random Variables, and Stochastic Processes* 3rd ed. (New York: McGraw-Hill)
[9] Trushkin A V 1993 *IEEE Trans. Inform. Theory* **39** 1180–1194
[10] Timuçin D A *et al.* 1994 *J. Opt. Soc. Am. A* (to be submitted)
[11] Van Trees H L 1968 *Detection, Estimation, and Modulation Theory* Part I (New York: Wiley)
[12] Prucnal P R and Saleh B E A 1981 *Opt. Lett.* **6** 316–318

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Space-variant filtering in fractional Fourier domains

Haldun M. Ozaktas and Billur Barshan
Bilkent University, Electrical Engineering, 06533 Bilkent, Ankara, TURKEY
David Mendlovic
Tel-Aviv University, Electrical Engineering, 69978 Tel-Aviv, ISRAEL
Hakan Urey
Georgia Institute of Technology, Electrical Engineering, Atlanta, GA 30332, U.S.A.

Abstract

Signals with significant overlap in both the space and frequency domains may have little or no overlap in a fractional Fourier domain. Spatial filtering in these domains may allow us to eliminate distortion components which cannot be eliminated in the ordinary Fourier domain.

1 Introduction

Space-invariant filtering may be performed by multiplying the Fourier transform of the input signal by the Fourier transform of the impulse response. Recently we have discussed how various space-variant operations can be performed by multiplying with a filter function in a fractional Fourier domain [1]. These operations can be realized optically, because the fractional Fourier transform can be realized optically. One approach is based on the use of quadratic graded index media [2, 3, 4], whereas another is based on the use of bulk lenses [5]. The graded index approach is closely connected to the definition of the fractional Fourier transform in terms of its spectral decomposition, whereas the bulk implementation is closely connected to its definition in terms of its linear transform kernel [1].

The many mathematical properties of the fractional Fourier transform, its relation to the Wigner space-frequency distribution, wavelet transforms, and chirp basis expansions, its applications to signal processing, and issues relating to its optical implementation are discussed in the references. Due to limited space, we will here content ourselves with the presentation of two examples of how space-variant filtering can be achieved by applying simple binary masks in fractional Fourier domains. Among the many things we cannot mention, of particular interest is correlation in fractional Fourier domains and its application to pattern recognition.

2 Definition of the fractional Fourier transform

The ath order fractional Fourier transform of a function $f(\cdot)$ is denoted by $\mathcal{F}^a[f](x)$ and may be defined as [1, 6]:

$$(\mathcal{F}^a[f])(x) = \int_{-\infty}^{\infty} \frac{e^{i(\pi\hat{\phi}/4-\phi/2)}}{|\sin\phi|^{1/2}} \exp[i\pi(x^2\cot\phi - 2xx'\csc\phi + x'^2\cot\phi)]\, f(x')\, dx',$$

where $\phi = a\pi/2$ and $\hat{\phi} = \mathrm{sgn}(\sin\phi)$. Some of its properties are: i.) linearity; ii.) $\mathcal{F}^0$ and $\mathcal{F}^4$ correspond to the identity operation; iii.) $\mathcal{F}^1$ corresponds to the conventional Fourier transform; iv.) $\mathcal{F}^{a_1}\mathcal{F}^{a_2} = \mathcal{F}^{a_1+a_2}$.

One of the most important properties states that performing the ath fractional Fourier transform operation corresponds to rotating the Wigner distribution by an angle $\phi = a(\pi/2)$

in the clockwise direction. We are unable to discuss the Wigner distribution here, although it is necessary for a complete understanding of the filtering examples discussed below. The reader is encouraged to consult [1] and the references given there. Roughly speaking, the Wigner distribution of a function $f(\cdot)$, denoted by $W_f(x,\nu)$, can be interpreted as a function that indicates the distribution of the signal energy over space x and frequency ν. Defining the rotation operator $\mathbf{R}_\phi$ for two-dimensional functions, corresponding to a counterclockwise rotation by ϕ, the property mentioned above can be expressed as $W_{\mathcal{F}^a[f]}(x,\nu) = \mathbf{R}_{-\phi} W_f(x,\nu)$. Another version of this property [7] is $\mathcal{R}_\phi[W_f(x,\nu)] = |\mathcal{F}^a[f]|^2$, where the operator $\mathcal{R}_\phi$ is the Radon transform evaluated at the angle ϕ. The Radon transform of a two-dimensional function is its projection on an axis making angle ϕ with the x axis.

Applications other than that discussed in this paper may be found in the references.

3 The fractional Fourier transform in optics

3.1 Optical implementation of the fractional Fourier transform

Analog optical implementations of the fractional Fourier transform have already been presented. In [1, 2, 3, 4] we discussed the fractional Fourier transforming property of quadratic graded index media. Lohmann suggested two systems consisting of thin lenses separated by free-space [5]. That the two approaches were equivalent and represented the fractional Fourier transform as defined above was demonstrated in [8].

The fact that the fractional Fourier transform can be realized optically means that the filtering examples discussed below can also be realized optically. Experimental results may be found in [9].

3.2 The fractional Fourier transform as a tool for analyzing optical systems

In [10, 11] we show that there exists a fractional Fourier transform relation between the (appropriately scaled) optical amplitude distributions on two spherical reference surfaces with given radii and separation. It is possible to determine the order and scale parameters associated with this fractional transform given the radii and separation of the surfaces. Alternatively, given the desired order and scale parameters, it is possible to determine the necessary radii and separation.

This result provides an alternative statement of the law of propagation and allows us to pose the fractional Fourier transform as a tool for analyzing and describing a rather general class of optical systems.

One of the central results of diffraction theory is that the far-field diffraction pattern is the Fourier transform of the diffracting object. It is possible to generalize this result by showing that the field patterns at closer distances are the fractional Fourier transforms of the diffracting object [10].

More generally, in an optical system involving many lenses separated by arbitrary distances, it is possible to show that the amplitude distribution is continuously fractional Fourier transformed as it propagates through the system. The order $a(z)$ of the fractional transform observed at the distance z along the optical axis is a continuous monotonic increasing function. As light propagates, its distribution evolves through fractional transforms of increasing orders. Wherever the order of the transform $a(z)$ is equal to $4j+1$ for any integer j, we observe the Fourier transform of the input. Wherever the order is equal to $4j+2$, we observe an inverted image, etc. [10, 11]

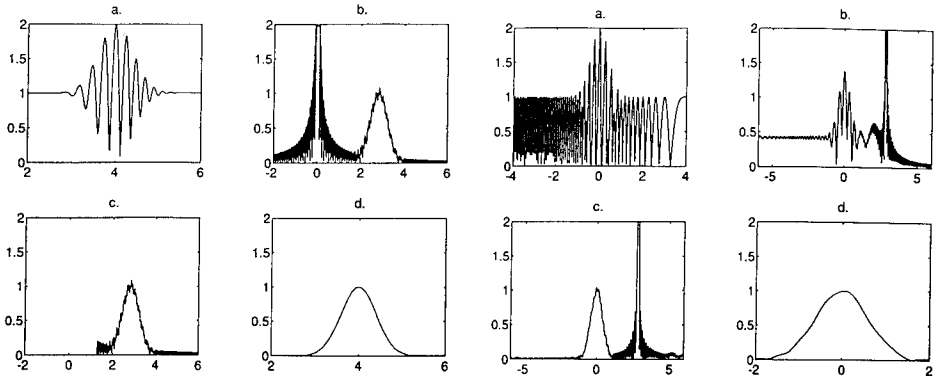


Figure 1: Example 1 (left) and example 2 (right)

4 Filtering examples

Consider the signal $\exp[-\pi(x-4)^2]$ distorted additively by $\exp(-i\pi x^2)\mathrm{rect}(x/16)$. The magnitude of their sum is displayed in part a., on the left hand side of the figure. These signals overlap in the frequency domain as well. In part b., we show their $a = 0.5$th fractional Fourier transform. We observe that the signals are separated in this domain. The chirp distortion is transformed into a peaked function which does not exhibit significant overlap with the signal transform, so that it can be blocked out by a simple mask (part c.). Inverse transforming to the original domain, we obtain the desired signal nearly perfectly cleansed of the chirp distortion (part d.).

Now we consider a slightly more involved example in which the distorting signal is also real. The signal $\exp(-\pi x^2)$ is distorted additively by $\cos[2\pi(x^2/2-4x)]\mathrm{rect}(x/8)$, as shown in part a., on the right hand side of the figure. The $a = 0.5$th transform is shown in part b. One of the complex exponential chirp components of the cosine chirp has been separated in this domain and can be masked away, but the other still distorts the transform of the Gaussian. After masking out the separated chirp component (not shown), we take the $a = -1$st transform (which is just an inverse Fourier transform) to arrive at the $a = -0.5$th domain (part c.). Here the other chirp component is separated and can be blocked out by another simple mask. Finally, we take the 0.5th transform to come back to our home domain (part d.), where we have recovered our Gaussian signal, with a small error.

The examples above have been limited to chirp distortions which are particularly easy to separate in a fractional Fourier domain (just as pure harmonic distortion is particularly easy to separate in the ordinary Fourier domain). However, it is possible to filter out more general types of distortion as well. In some cases this may require several consecutive filtering operations in several fractional domains of different order [1]. There is nothing special about our choice of Gaussian signals other than the fact that they allow easy analytical manipulation. Also, there is nothing special about the 0.5th domain. It just turns out that this is the domain of choice for the examples considered above.

In the above examples we have demonstrated that the method works, but did not discuss what led us to transform to a particular domain and what gave us the confidence that doing so will get rid of the distortion. This becomes very transparent once one understands the relationship between the fractional Fourier transform and the Wigner distribution. This relationship,

as well as the general philosophy behind such filtering operations is discussed in [1].

5 Conclusion

What we know as the space and spatial frequency domains are merely special cases of fractional domains. These domains are indexed by the parameter a. The representation of a signal in the ath domain is the ath fractional Fourier transform of its representation in the $a = 0$th domain, which we define to be the space domain. The representation in the $a = 1$st domain is the conventional Fourier transform. If we set up a two-dimensional space, called the Wigner space, such that one axis (x) corresponds to the $a = 0$th domain (the conventional space domain) and the other (ν) to the $a = 1$st domain (the conventional spatial frequency domain), then the ath domain corresponds to an axis making an angle $\phi = a\pi/2$ with the x axis.

A desired signal and noise may overlap in both conventional space and frequency domains, but not in a particular fractional domain. Even when this is not the case, spatial filtering in a few fractional domains in cascade may enable the elimination of noise quite conveniently. It is possible to implement these operations optically, as well as with a fast digital algorithm.

It is a pleasure to acknowledge the contributions of A. W. Lohmann of the University of Erlangen-Nürnberg in the form of many discussions and suggestions.

We acknowledge the support of the NATO Scientific Affairs Division within the framework of the Science for Stability Program.

References

[1] H.M. Ozaktas, B. Barshan, D. Mendlovic and L. Onural, *J. Opt. Soc. Am.* A, 11:547-559, 1994

[2] H.M. Ozaktas and D. Mendlovic, *Opt. Commun.*, 101:163-169, 1993.

[3] D. Mendlovic and H.M. Ozaktas, *J. Opt. Soc. Am.* A, 10:1875-1881, 1993.

[4] H.M. Ozaktas and D. Mendlovic, *J. Opt. Soc. Am.* A, 10:2522-2531, 1993.

[5] A.W. Lohmann, *J. Opt. Soc. Am.* A, 10:2181-2186, 1993.

[6] A.C. McBride and F.H. Kerr, *IMA J. Appl. Math.*, 39:159-175, 1987.

[7] A.W. Lohmann and B.H. Soffer, *Technical Digest of the 1993 Annual Meeting of the OSA*, p.109, 1993.

[8] D. Mendlovic, H.M. Ozaktas and A.W. Lohmann, "Graded index fibres, Wigner distribution functions and the fractional Fourier transforms", to be published in *Appl. Opt.*.

[9] R.G. Dorsch, A.W. Lohmann, Y. Bitran, D. Mendlovic and H.M. Ozaktas, "Chirp filtering in the fractional Fourier domain: experimental results", to be published in *Appl. Opt.*.

[10] H.M. Ozaktas and D. Mendlovic, "The fractional Fourier transform as a tool for analysing beam propagation and spherical mirror resonators", to be published in *Opt. Lett.*.

[11] H.M. Ozaktas and D. Mendlovic, "Fractional Fourier Optics", submitted to *J. Opt. Soc. Am.* A.

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 British Crown Copyright

A novel form of incoherent optical correlator

M F Lewis

DRA Malvern, St. Andrews Road, Malvern, Worcs. WRI4 3PS, UK

Abstract. Much work on digital signal processing, DSP, and optical signal processing, OSP, is aimed at pattern recognition. Here we describe a simple and rugged but promising hybrid digital/incoherent optical approach to this problem.

1. Introduction

The problem of identifying the presence of a given pattern in a scene is superficially similar to that of identifying a given radio or radar signal in a noisy environment. The optimum solution to the latter problem was devised in WW2 and takes the form of either a *correlator* or a *matched filter*, appropriately modified if the background noise does not have a white spectrum. For this reason the first and many subsequent approaches to pattern recognition have been based on these components.[1-4] However in reality there are numerous differences between these superficially similar problems. For example, the electronic problem is one-dimensional (time) while in pattern recognition we need to process a two-dimensional projection of a three-dimensional object, often with no a priori knowledge of its scale or orientation. In addition it turns out that the background spectrum is never white. Much of the research effort over the past thirty years has sought to overcome such complications of the pattern recognition problem. The work described here is a revisit of an earlier incoherent optical correlator, but with the incorporation of new features to improve its performance in various respects. In particular we describe a hybrid OSP/DSP processor incorporating spectral whitening which exploits each technology to advantage, and a multi-channel version of this device for use with scrolling input data/scenes.

2. Summary of new work on incoherent correlators

A simple and elegant incoherent correlator is described in various standard textbooks.[2,3] It operates by a shadow casting process based on geometrical optics, as shown schematically in Figure 1. As it stands this setup is non-ideal because all objects correlate with a given reference to some extent, the output deriving from optical intensity which is positive definite. This difficulty arises because the (spatial) spectrum of the scenery is not "white".

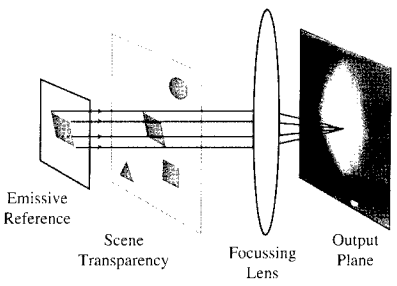


Figure 1. Schematic incoherent optical correlator

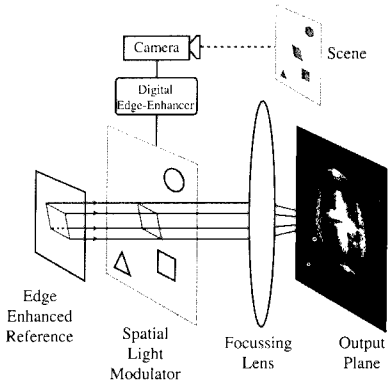


Figure 2. Modified incoherent optical correlator incorporating edge-enhancement

In particular, because light intensity is positive definite the spectrum is always peaked at zero spatial frequency. According to the earlier analysis of temporal signals, the ideal processor should "prewhiten" this spectrum prior to matched filtering. Computer simulations show that the consequence of prewhitening is to require scenes and references which are essentially *real edge-enhanced versions* of the originals. This is agreeable, as it means that the original apparatus can be employed; Figure 2 shows a version similar to the one we have used in which the input scene is edge-enhanced in the time domain at video rates using DSP chips. It should be emphasised that true edge enhancement requires bipolar operation, requiring duplication of part of the processor. To date we have not implemented this for reasons of simplicity, and because the use of a single-channel allows identification of dark signals on a light background as well as light signals on a dark background.

Various other points emerge from the computer simulation, eg we have identified the importance of *corners* as well as edges (for objects which possess corners) and have found that if the target possesses some unique feature then one should emphasise this feature in the correlation process.

The scheme shown in Figure 2 has been built and demonstrated, but with the scene and reference transposed, the former employing a Panasonic LC TV input and the latter being a fixed edge-enhanced transparency. Using artificial scenery this has enabled us to identify a given target in cluttered and noisy backgrounds. An important issue in the design is the width of the edges to be used in practice, since mathematically these are vanishingly small. We have chosen to implement the edges of the central component with quite wide slits (of order 0.5mm in our setup) as this conveys various benefits on the system:

(a) It allows most light through to the output stage.

(b) It gives the correlation process some tolerance to scale and orientation, an important

issue in pattern recognition. When wider edges are used, the effect is similar to that obtained using a synthetic discriminant function (SDF) filter[4].

(c) It minimises the deleterious effect of optical diffraction through narrow slits, which degrades the correlation peak.[2]

An agreeable feature of edge-enhancement is that it causes a sharpening of the correlation peak, enabling the centre of the object to be located precisely. This is evident from a comparison of the output planes of Figures 1 and 2. A similar phenomenon occurs in certain coherent correlators which deliberately or accidentally incorporate a degree of spectral prewhitening, such as those employing "phase-only matched filters". However, such systems are very dependent on a close match between reference and scene. A useful compromise is to use wider slits/sources to implement the edges, as discussed above. A further advantageous feature of *digital* edge-enhancement is that it can eliminate the effects of non-uniform illumination in the scene, eg a target in the shade of a tree can be made to appear just as bright as a fully illuminated target if it exhibits the same pattern of edges.

Incoherent correlators are greatly superior to their coherent counterparts in respect of size, cost and ruggedness and can be realised in a variety of optical configurations. A potent implementation might employ an LC SLM to input the edge-enhanced scene at video rates, and an array of surface emitting lasers, which can be modulated at GHz rates[5], to input multiple references in very rapid succession, eg >100 per frametime, to cover further variations in scale and orientation. The incoherent correlator possesses a number of additional advantages over coherent optical correlators. Two of the more important of these are:

(1) Unlike the coherent correlator, the incoherent correlator can readily employ references designed to mimic the *colour content* of a known target. (Colour could also be used to implement the bipolar processor mentioned earlier if so required).

(2) Optical correlators exhibit invariance to object location *in two dimensions*. However there are many circumstances in which it is only necessary to have invariance in one dimension, eg when dealing with scrolling text, when items translate along a production line, or when visible, ir, or radar images are gathered from a moving aircraft. In such circumstances the incoherent correlator can exploit the second dimension to good effect. In the laboratory we have demonstrated this by employing the second dimension to seek three different objects, namely the letters, N, O and A. The apparatus used to achieve this is shown schematically in Figure 3. The principal differences from Figures 1 and 2 are (a) illumination of the scene with a narrow "line" source and cylindrical lens, the output of which diverges horizontally but is constrained vertically, and (b) optics which separates the outputs of the three channels in the output plane; a segmented lens was used for simplicity. The operation of this structure is illustrated in Figure 4. The central photograph shows the three separate correlation peaks arising if the letters N,O,A are input to the correct channels. The other four photographs show the lower-level cross-correlations that occur if the input pattern is stepped up/down by 1/2 steps. The camera exposure was the same for all five photographs.

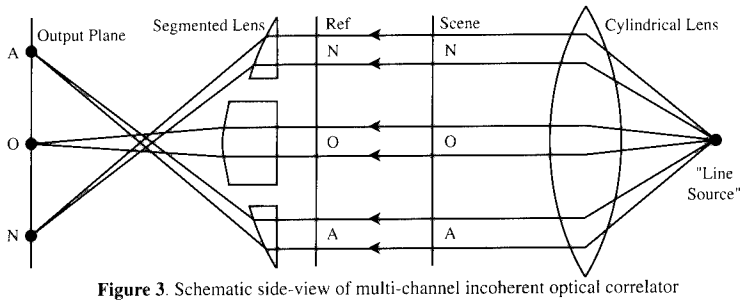


Figure 3. Schematic side-view of multi-channel incoherent optical correlator for use with scrolling input scenes

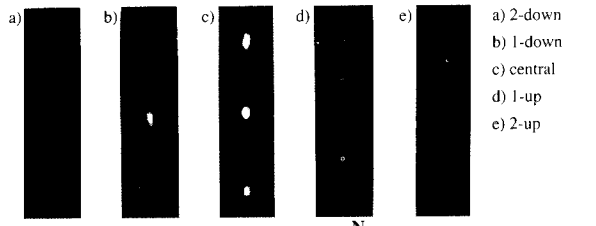


Figure 4. Multichannel correlator on output using $\begin{matrix} N \\ O \\ A \end{matrix}$ for both scene and references

3. Conclusion

This paper describes various modifications to a simple and rugged incoherent correlator. The incorporation of edge-enhancement leads to greatly improved inter-class discrimination, while a further modification, the multi-channel correlator, is particularly appropriate to the processing of scrolling data. Acceptable geometries for this type of correlator are discussed in another paper at this conference.[6]

Acknowledgment

The author acknowledges helpful discussions with Rebecca Wilson and Rupert Young. The work described here was begun in collaboration with Chris West, and the real-time edge enhancement was provided by Dave Barrett.

References

[1] Vander Lugt A 1964 IEEE Trans *IT-10* 139-145
[2] Goodman J W 1968 *Introduction to Fourier Optics* McGraw-Hill
[3] Hecht E and Zajac A 1974 *Optics* Ch 11 Addison-Wesley
[4] Casasent D 1984 *Appl Opt* **23** 1620-1627
[5] Tauber D et al 1993 *Appl Phys Lett* **62** 325-327
[6] Laude V, Chavel P, & Refregier, Ph, this Proceedings.

© British Crown Copyright 1994/DRA
Published with the permission of the Controller of Her Britannic Majesty's Stationery Office.

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Experimental implementation of a joint transform correlator providing rotation invariance

Laurent Bigué, Michel Fracès* and Pierre Ambs

Ecole Supérieure des Sciences Appliquées pour l'Ingénieur - Mulhouse (ESSAIM), Université de Haute Alsace, 12 rue des Frères Lumière, F-68 093 Mulhouse Cédex France

* ONERA-CERT, Département d'Optique, 2 avenue Edouard Belin, BP 4025, F-31 055 Toulouse Cédex France

Abstract. We optically experiment a joint transform correlator (JTC) where the input image is replaced by a pseudo-inverse synthetic discriminant function (SDF) filter. This latter allows us to ensure a practical invariance over a 10° range. Since our input amplitude SLM cannot display other kinds of SDFs, we propose to post-process them with a DBS-based method and provide simulations.

1. Introduction

Optical correlation, and especially joint transform correlation [1], has been known as an efficient alternative to its electronical counterpart since the 1960's and compact implementations of joint transform correlators (JTCs) have recently been presented. But correlation is inherently rotation- and scale- variant and a traditional correlation operation cannot perform pattern recognition in most practical cases. In order to overcome this limitation, Javidi [2] has proposed and simulated a JTC using synthetic discriminant function (SDF) filters. Juvells *et al.* [3] have even provided optical experimental results, but have not tested the correlator training ability to recognize intermediate rotated or scaled views. In this paper, we propose such experimental results. But our input system proves unable to input all types of SDF filters because it cannot display negative values. Instead of replacing our input device with a more sophisticated one allowing us to control both amplitude and phase [4], we propose to postprocess our filters with a Direct Binary Search (DBS)-based method. This method is of general purpose and we show that it allows us to display negative-valued images.

2. Basic Theory

The basic principle of joint transform correlation [1] is now widely known, and we shall not present it. The need of ensuring distortion invariances appears clearly: on our correlator, at the CERT Optical Science Department, an in-plane rotation of one degree or a ± 5% scale change causes a dramatic energy loss of more than 3 dB. To achieve an invariance for such a

phenomenon, we decided to use the SDF technique [5], a general method to achieve any invariance. In this paper, we decided to limit ourselves to in-plane rotations: in this case, the SDF technique consists in replacing the reference image with several rotated views. The correlation peak height will be equal, provided the input view belongs to the training set. The point is to test intermediate-rotated input views, between those of the training set.

The simplest SDF filter, a linear combination of views from the training set, is called pseudo-inverse SDF (PI-SDF). Other types of filters have then been derived from it: the minimum variance SDF [6] (MV-SDF) which is coloured noise resistant, the minimum average correlation energy [7] SDF (MACE-SDF) which sharpens the correlation peak and the optimal trade-off SDF [8] (OT-SDF) which combines the previous two qualities.

3. Preliminary simulations

We tested the four kinds of SDFs mentioned above, in order to verify their training ability. For each type of filter, we calculated four filters, each of them containing 6 views of F-4 or F-18 planes, with a various angular step between two consecutive views: 0.5°, 1°, 2° or 4°.

The simulations give the same results as those in the literature for the 4f architecture: the PI-SDF training ability is average: there is a marked difference between the images belonging to the training set and images of the same airplane, but not in the training set. The discrimination between the plane of the training set and another lifelike plane is unambiguous (the ratio of the correlation peak heights is over 4). The MACE-SDF only recognizes the views it contains and so proves useless, and the MV-SDF produces a very poor discrimination, sometimes no discrimination at all. As expected, the OT-SDF produces intermediate results, and is worth testing experimentally.

4. Experiments

Such experiments had already been reported [9], but never with a joint transform correlator. The experimental setup was implemented on the CERT Optical Science Department demonstrator. The input SLM of the correlator, a GEC-Marconi liquid crystal light valve (LCLV), perform an amplitude modulation [10].

We tried to test the four kinds of filters previously simulated. Unfortunately, MACE-

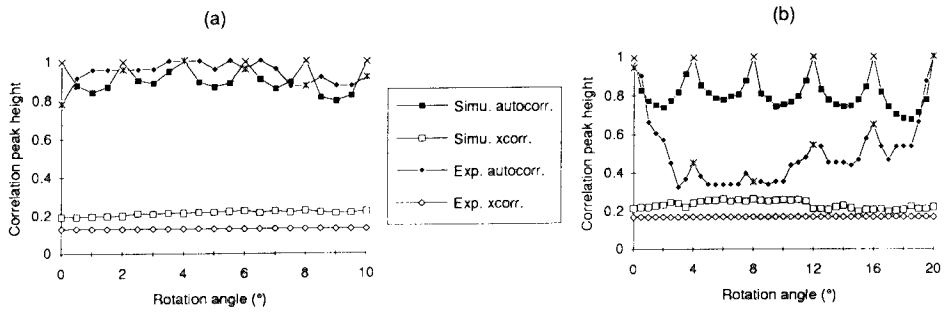


Fig. 1.Simulated and experimental correlation peak heights for PI-SDFs whose views differ from 2° (a) or 4° (b). In all figures (1(a), 1(b) and 2), crosses denote autocorrelation when the input view belongs to the training set and 'xcorr' curves are relating to cross- correlations.

SDFs, MV-SDFs and so OT-SDFs proved unusable: they are non-positive, so produce a non-zero background on our amplitude SLM and especially their dynamic range is high. The input setup cannot display both a large dynamic range and tiny details.

We experimented 6-view PI-SDFs, with respective angular steps of 0.5°, 1°, 2° and 4° between two consecutive planes. For a 0.5, 1 or 2° difference, the output response is quite uniform over the angular range: +/- 7 % (see fig. 1(a) for the 2° step). The training ability is good: we cannot even find any difference between the images belonging to the training set and images outside the training set. The discrimination is also good (the ratio of the peak heights is at least 1:5). For a 4° difference, the results deteriorate dramatically (fig. 1(b)) and are no longer in good agreement with the simulations.

5. Simulations

To overcome the incorrect display of non-positive images, there are two solutions: to post-process the images in order to fit the SLM coding domain, or to replace the input system with a more sophisticated one [4]. We chose the first way, and decided to make the filters fit our input device at best. Juvells *et al.* [3] have recently proposed to code the filter with a detour phase technique (derived from Lee's method [11] of generating holograms by computer), which allows displaying the whole filter information with only positive values. Such methods have been extensively studied in the field of holography, and one of the best proved to be Direct Binary Search (DBS) [12,13].

This iterative method was originally dedicated to binary holograms, Legeard *et al.* [14] have extended it to grey level holograms: a random image -the "hologram"- is first generated, and its FFT -"the reconstruction"- is then computed. An error function is calculated by comparing this FFT and the FFT of the original image. The grey level of each pixel is modified pixel after pixel and a new error is computed each time between the new "reconstruction" and the FFT of the original image. If the new error is smaller than the previous one, then the modification of the pixel is maintained; if not it is cancelled. This iterative process goes on until no modifications are required.

We proposed to code our OT-SDF filters with this method, with a number of grey levels ranging from 2 to 256. The filters behaviour has been simulated and results normalized in relation to the post-processed filter energy are depicted fig. 2. The difference between views belonging, or not, to the training set is less marked after post-processing. The discrimination ability is maintained. The simulations also show that a reduced number of grey levels is

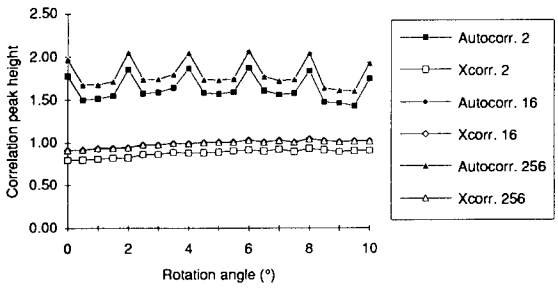


Fig. 2. Simulated correlation peak heights for OT-SDFs whose views differ from 2°, and which have been post-processed with DBS. 2,16 or 256 indicate the number of grey levels in the filter.

sufficient, with hardly any difference -invisible on a graphics- between the 16-level filter and the 256-level one. This allows us to content ourselves with filters comprising a reduced number of grey levels, computationally less expensive.

There is a huge advantage of such a method over a detour phase one: the available space-bandwidth product is not magnified. But on the other hand, the computational cost is high, since this method is iterative. Furthermore, as in the holographic case, we can hope that the results are better, as far as error is concerned [13].

6. Conclusion

We have implemented a joint transform correlator using synthetic discriminant function filters. The experiments proved that the MACE-SDF and the MV-SDF -and so the OT-SDF- cannot be displayed correctly on an amplitude SLM. But the PI-SDF which is correctly displayed, works. We showed that for an in-plane rotation - provided the angular step between the views it contains is not too high - this filter can lead to an unambiguous discrimination, for any object orientation over the filter angular range: it showed a certain training ability and especially, that it could be optically implemented, without too much degradation, into JTC architectures whose compact examples have become popular. Further simulations concerning OT filters showed that provided the filters were post-processed to fit SLM constraints (in our case an amplitude device), they could be implemented into JTC architecture without losing their inherent qualities.

Acknowledgements

The authors thank Luc Legeard, from the Université de Haute Alsace - ESSAIM, Mulhouse, for his computing of DBS filters and his enlightening discussions.

References

[1] Weaver CS, Goodman JW 1966 *Appl. Opt.* **5** 1248-9.
[2] Javidi B 1989 *Appl. Opt.* **28** 2490-3.
[3] Juvells I, Carnicer A, Vallmitjana S, Campos J 1994 *J. Optics* **25** 33-40.
[4] Bigué L, Fracès M, Ambs P 1994 *Opto 94* (Paris: ESI Publications), 430-3.
[5] Caulfield HJ, Maloney WT 1969 *Appl. Opt.* **8** 2354-6.
[6] Vijaya Kumar BVK 1986 *J. Opt. Soc. Am. A* **3** 1579-84.
[7] Mahalanobis A, Vijaya Kumar BVK, Casasent D 1987 *Appl. Opt.* **26** 3633-40.
[8] Réfrégier P 1991 *Opt. Lett.* **16** 829-31.
[9] Nicholson MG, Gibbons GG, Cooper IR, Petts CR 1987 *Progress in Holography* (SPIE), 54-9.
[10] Bigué L, Fracès M, Ambs P 1993 *Opto 93* (Paris: ESI Publications), 201-6.
[11] Lee W-H 1970 *Appl. Opt.* **9** 639-43.
[12] Seldowitz MA, Allebach JP, Sweeney DW 1987 *Appl. Opt.* **26** 2788-98.
[13] Legeard L, Ambs P 1992 *Opto 92* (Paris: ESI Publications), 149-58.
[14] Legeard L, Ambs P, Marchand PJ, Esener S 1993 *Fourth International Conference on Holographic Systems, Components and Applications* (Neuchatel: IEE), 222-7.

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Associative memory for rotation pre-processing and projection invariant pattern recognition .

Guy Lebreton

GESSY, Université de Toulon, B.P.132, 83957 La Garde cedex, France

Abstract. For invariant pattern recognition, an original technique using a simple associative optical memory is presented; it detects the orientation of any object before identification. Then the implementation of another associative optical memory for pattern recognition invariant to scale and projections is discussed.

1. Introduction

To increase the computing efficiency with moderate data flow (Video rate typical), the original idea of the hybrid neural network architecture (modified Hopfield type) is to utilise a non-linear optical amplifier between two cascaded optical correlators, the first one giving the weights (inner product with central correlations), the second yielding the weighted memory output for electronic thresholding and feed-back to the optical processor. In the implementation proposed at OC'92 [1], the high capacity programmable optical memory was a photothermoplastic plate, and the non-linear amplifier was a BGO crystal with phase conjugation (Fig.1). It was shown that a high non-linearity (of the order 4 at least) maintains the dominant influence of the $\Gamma(0)$ peak in an extended window required for shift invariance [2]).

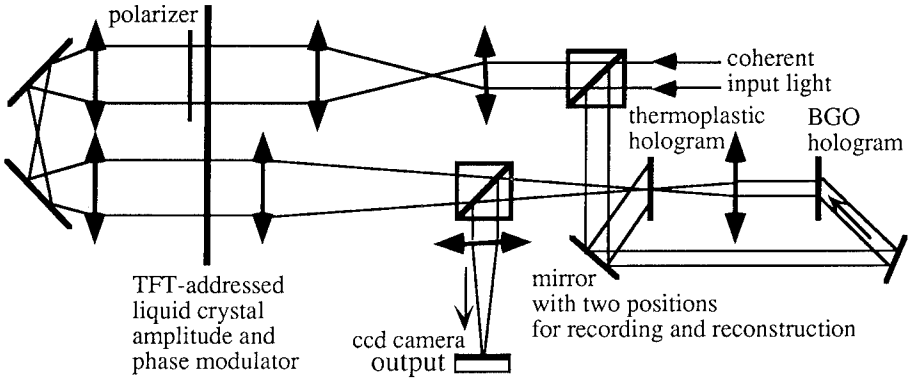


Figure 1 : double-correlator optical associative memory with complex input

Instead of performing the standard outer product, computing the weighting coefficients as the sum of central correlation peaks between the input image f and the memory models f_m, the output of the first-stage correlator keeps the correlation window, to provide for shift invariance. The second-stage correlator convolves it with the memories for the auto-associative reconstruction; the blurr which would result from this low-pass filtering is compensated for by the highly non-linear detection (power 4 instead of the power 2 used in ref.2) :

$$h(x,y) = \sum_m \left| \Gamma_{f_m f}(\zeta,\eta) \right|^4 * f_m(\zeta,\eta)$$

To obtain this high non-linearity, the photorefractive crystal should be used twice, but this requires too much optical input power. The new version proposed will replace the crystal with a non-linear optically addressed ferroelectric liquid crystal.

The progresses presented here do not concern the non-linear optical layer, but the algorithms of the process, and should be also of interest for electronic computing. The goal was to utilise the above double-correlator optical associative memory for two successive operations : first, find the orientation of an unknown input object and rotate it as a pre-processing stage; then, identify the object with projection and scale invariant detection. The first operation has been successfully completed and will be fully described. The second operation is now under investigation and actual projects for its implementation will be discussed.

2. Rotation pre-processing before projection-invariant recognition.

The filter previously reported [1] for finding the orientation of an input from a set of 4 aircrafts was the one (the left one on Fig.2), that had a detectable cross-correlation with all other objects of this small data base. The memory contained replicas of this model, rotated from ten degrees. The results were satisfactory on contour objects, but were restricted to one class of correlated objects and required edge detection pre-processing. Then a simplified model object was usedas shown in Fig.2 below.

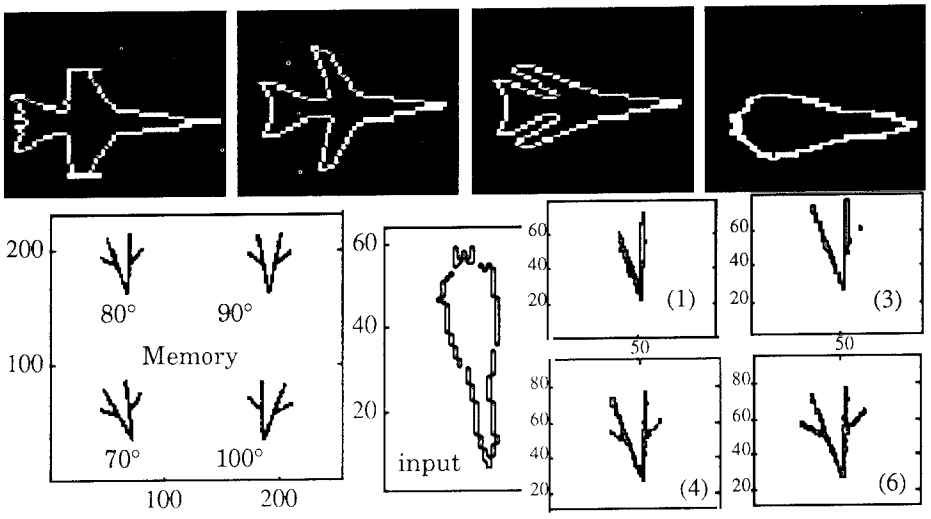
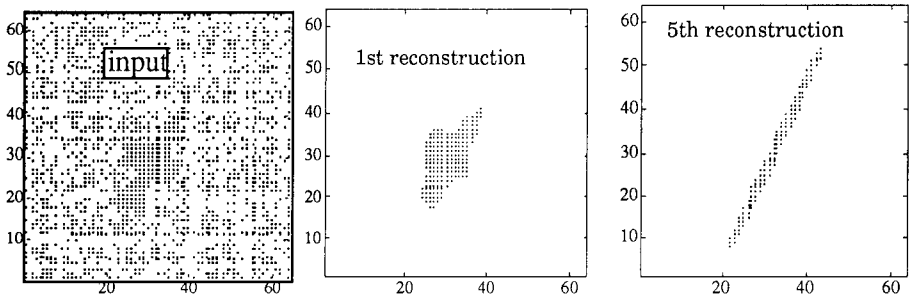


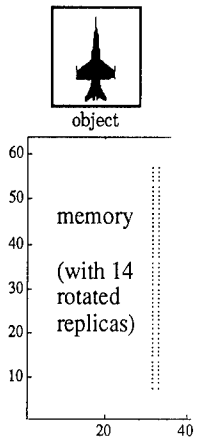
Figure 2 : data base of 4 aircrafts and iterative reconstruction of a model

The model oriented closer to the direction of the input aircraft is reconstructed after a few iterations, but it appears that the simplified wings are not useful. After their suppression, it was found that the two lines of the remaining angle in the model could induce a conflicting competition to select the orientation. The final model has been reduced to one thick line (2x50 pixels) and the pre-processing of the input for edge detection was suppressed. Then one obtained excellent results for any type of input object, quite independent of distortions, as illustrated on Fig.3.



iterations →

thresh / angle	0	1	2	3	4	5	6	7	8	9
0	64	45	40	34	44	33	50	28	50	
-55	66	24	16	12	9	6	6	5	6	
-45	66	24	16	11	8	5	5	5	5	
-35	67	26	18	13	10	6	6	5	6	
-25	65	26	20	15	12	8	8	7	8	
-15	65	30	24	18	14	8	8	7	8	
-5	66	36	31	23	20	13	13	11	13	
5	65	38	39	33	25	18	18	16	18	
15	69	37	42	44	43	32	32	27	32	
25	81	49	62	80	94	90	95	85	99	
35	73	45	48	48	40	27	27	23	28	
45	69	33	34	29	23	15	15	13	15	
55	64	30	27	21	16	12	12	10	12	



Correlation peaks (weight coefficients)

Fig.3 : detection of the orientation of an unknown object

The selectivity, tested with ellipses as input, reaches a 1 dot difference in the discrete image. For a direction exactly between two nearest recorded angles of the filter, one obtains a stable output with identical values for the nearest angles, if all rotated replicas have exactly the same weight. To provide for this condition with the discrete 64x64 image, a longer model line was used to compute each rotated version and then truncated to a uniform weight of 100 pixels.

An important improvement for the convergence speed has been to program an adaptive thresholding, to increase the gain when the competition is low between the output modes. In a standard Hopfield model, the threshold plane is the correlation plane, which cannot be accessed here in the double-correlator architecture. Thus one utilises the Parseval Theorem, saying that the correlation peak amplitude equals the energy of the signal. Our threshold criterion is then the sum over the output plane. The adaptive threshold % shown on Fig.3 was adjusted at each iteration to keep only in the new reconstructed image1.1 times the number of pixels of the model filter ±10%.

3. Projection and scale invariant pattern recognition.

After this pre-processing, a full invariance in pattern recognition can be achieved using the same associative memory with a scale and projection invariant filter. The interest of using the associative memory architecture instead of a simple correlation is to enhance the signal to noise ratio and the tolerance to distortions. First a Log-Log single harmonic filter (2-D Mellin transform) was tested: it is invariant to projections on both axes, but it would require a Mellin transform on the real-time input as in previous implementations [3]. A new filter has been synthesised recently, by doing first a Mellin-radial single harmonic filter f_R^N, invariant to scale, and then a 1-D Log Transform f_{LR}^N with the same harmonic N [4] :

$$f_R^N(r,\theta) = r^{i2\pi N-1}.\int_{r_0}^{R} f(r,\theta).r^{-i2\pi N-1}dr$$

$$f_{LR}^N(x,y) = |x|^{i2\pi N-1/2}.\int f_R^N(x,y).x^{i2\pi N-1/2}dx$$

First experiments in correlation show an insufficient discrimination capability for inserting this filter in the associative memory. The rejection of another object g might be forced with the synthetic discriminant function approach :

$f_{LR}(x,y) = |x|^{i2\pi N-1/2}\left[a.f_R^N(x,y) + b.g_R^N(x,y)\right]$, with complex numbers a and b solutions of $\begin{bmatrix} \Gamma_{ff_{LR}}(0,0) & \Gamma_{gf_{LR}} \\ \Gamma_{fg_{LR}} & \Gamma_{gg_{LR}} \end{bmatrix}.\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

But the experience on the rotation filter shows that a partial correlation from contributing neighbours in the memory is necessary to enable energy transfers between the intermediate iterations, so we should allow 1/2 instead of 0 for the cross-correlation. This solution is under investigation, but it is expected that the b coefficient might be sensitive to distortions of the input. Another solution currently being tested is using only the Mellin-radial single harmonic filter for scale invariance, which has a good selectivity, and leaving the associative memory to provide for some distortion invariance. But the Mellin-radial harmonic will never retrieve hidden information, such as specific top-view, profile and front-view contents: this information will have to be introduced as a hidden layer in the neural network, as illustrated for neural network classifiers [5].

4. Conclusion.

Previously, the memory content for orientation detection had to be selective for one class of object, with limited tolerance to distortions, and worked only on contours with low optical efficiency. Now, excellent results have been obtained for any type of filled objects, using as memory a simple line with rotated replicas.

For the projection invariance, much work is still to be done, but the harmonic decomposition approach is expected to avoid much of the redundancy of a synthetic discriminant filter by providing at least for scale invariance.

References

[1] Lebreton G, Mendlovic D et al. 1993 *SPIE Proc.* **1806** 305-312
[2] Psaltis D, Hong J 1987 *Opt. Eng.* **26** 10-15
[3] Mendlovic D et al. 1989 *Appl. Opt.* **28** 4982-4986.
[4] Mendlovic D, Lebreton G et al. 1994 *to appear in Appl. Opt.*
[5] Casasent D 1992 *Opt. Eng.* **31** 971-978

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Adaptive pattern recognition with rotation, scale and shift invariance

Ezra Silvera, Tuvia Kotzer and Joseph Shamir

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel.

Abstract. Mutually orthogonal pattern distortions are handled by an adaptive optical recognition system. A double channel system is presented which implements pattern recognition with rotation, scale and shift invariance. The overall recognition process is performed efficiently and can be executed in real time.

1. Introduction

The recognition process is based on a two stage operation: An object independent determination of one distortion parameter (the scale, in the example presented here) and then the recognition is completed by a shift and rotation invariant optical correlator which is adapted to the measured parameter (in our example the input is scaled up or down, according to the mesured scale). Thus, complete invariance to three distortion parameters is achieved by the combination of two channels. See also Fig. 1.

2. Scale measurement procedure

The overall recognition process is performed in two stages where the scale measurement is performed first. This procedure dictates several properties: The scale measurement must be rotation and shift invariant since the object's orientation and position are not known *a priori*. It must be independent of the object as long as the object belongs to the class of objects to be recognized. The whole process must be easily implemented optically and, finally, it must be fast enough for efficient operation of the complete system.

Let $f(r,\theta)$ be the input function and $F(\rho,\phi)$ its Fourier transform (in polar coordinates). When the scale of the input is extended by a factor a, the energy, using Parseval's theorem, becomes:

$$E_a = \int_0^{2\pi}\int_0^{\infty} \left|a^2 F(a\rho,\phi)\right|^2 \rho d\rho d\phi = a^2 E_0, \tag{1}$$

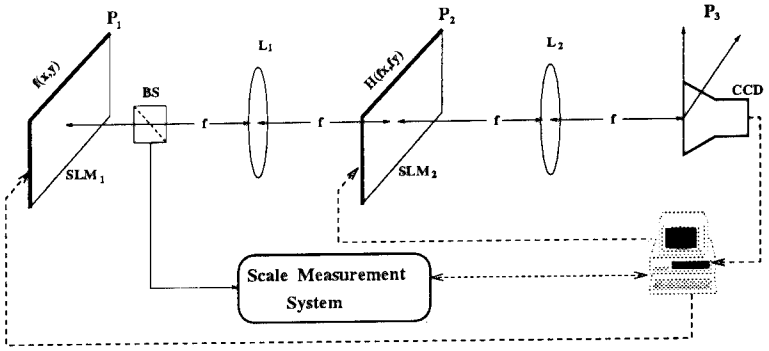


Figure 1. Schematic configuration of the overall recognition system.

where, E_0 is the energy of the original function $f(r,\theta)$. Let $E_a(\hat{\rho})$ be the energy of the object in the radial spatial frequency interval $0 < \rho < \hat{\rho}$:

$$E_a(\hat{\rho}) = \int_0^{2\pi}\int_0^{\hat{\rho}} \left|a^2 F(a\rho,\phi)\right|^2 \rho d\rho d\phi = a^2 \int_0^{2\pi}\int_0^{a\hat{\rho}} |F(\rho,\phi)|^2 \rho d\rho d\phi \tag{2}$$

We define the ratio between $E_a(\hat{\rho})$ and the total energy E_a by:

$$T(a,\hat{\rho}) = \frac{E_a(\hat{\rho})}{E_a} = \frac{\int_0^{2\pi}\int_0^{a\hat{\rho}} |F(\rho,\phi)|^2 \rho d\rho d\phi}{E_0} \tag{3}$$

Clearly, $T(a,\hat{\rho})$ depends on $\hat{\rho}$ only through the integration limit. Hence, by requiring the value of T to assume a fixed (reference) value T_0 **independent** of the scale, the scale factor, a, can now be easily extracted from the simple relation:

$$a = a_0 \frac{\hat{\rho}_0}{\hat{\rho}} \tag{4}$$

where a_0 is the reference scale and $\hat{\rho}_0$ the corresponding radial frequency.

One of the most important parameters we must take into consideration is the operation speed. Therefore, in the laboratory system, we used an approach based on the binary search algorithm, see [1]. In this way for an input object of $N \times N$ pixels the scale is found after no more than $log_2(N)$ steps.

3. Filter Design

We use a two-stage recognition process in which the input object is adapted (re-scaled) before the correlation is made. Thus, theoretically, in the correlation channel we have to use a rotation and shift invariant filter. However, due to the errors in the scale measurement procedure, there is a scale mismatch between the filter and the object. Conventional filters (and especially rotation and shift invariant filters), cannot operate properly with such a scale mismatch. This is the main reason for developing a new filter which is invariant to a limited scale range as well as remaining fully rotation and shift invariant. The scale-invariance range was chosen according to the scale measurement errors, so that those errors are fully compensated by the correlation channel.

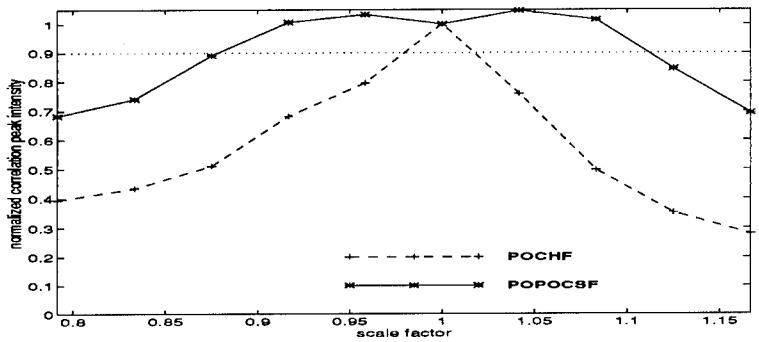


Figure 2. The POPOCSF sensitivity to scale in comparison with that of the POCHF. The intensities are normalized so that both peaks are equal.

The filter was designed using the POCS (Projection Onto Convex Sets) algorithm and will be called the POCS filter (POCSF). The requirements that are explicitly used in the design procedure are: high correlation peak intensity, good discrimination ability (in this case detect the letter 'F' while rejecting the rest), full rotation invariance, the filter being a passive element, and finally, limited scale invariance. This algorithm and its implementation for our purposes are discussed in detail in [2] and [1], respectively, and hence will not be repeated here.

The laboratory implementation of the optical correlator employs SLM's, both in the input and in the Fourier planes. This architecture dictates several requirements from the filter, including high light efficiency, high correlation peak and the ability to record the filter on the SLM which has a limited number of pixels. A phase-only version of the designed filter (POPOCSF) is a good candidate for implementation. The invariant scale range ($[a_1, a_2]$) was chosen to be $\pm 8\%$.This range fully compensates for the inaccuracies of the scale estimation.

Fig. 2 clearly demonstrates that for the scale range required, due to the limited accuracy of the scale measurement, the POPOCSF is indeed invariant and even for a scale range of $\pm 10\%$ the peak remains above 95% of its maximum value. The phase-only CHF (POCHF), as expected, is very sensitive to scale and, obviously, cannot be used in the proposed system. The laboratory correlation system was based on a 4-f optical correlator in which both filter and input were written on SLM's. Several experiments were performed in order to investigate the filter characteristics, including discrimination and sensitivity towards rotation and scale changes of the input function. As it can be seen from Fig. 3. the letter 'F' was recognized while the letter 'P' was rejected with a rejection ratio of, approximately, 50%.

The experimental performance of the complete system was tested and excellent results were obtained, as expected from the results presented in previous sections.

4. Conclusions

A double channel optical pattern recognition system was introduced, implementing distortion invariant operation. The basic principle is quite general and can be extended to

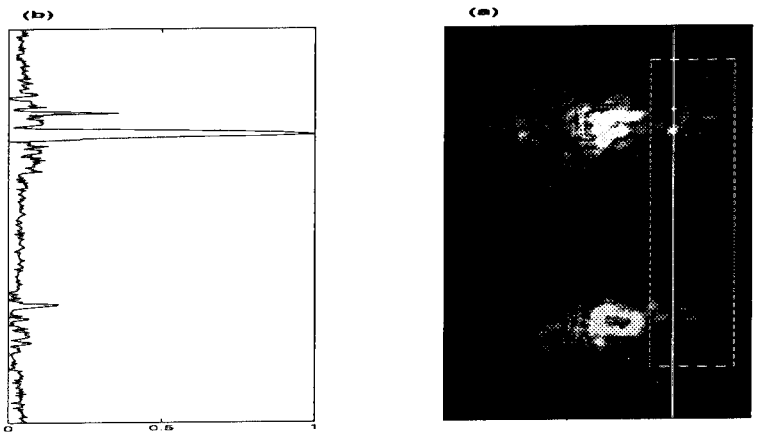


Figure 3. Correlation response of the binary POPOCSF for the letters 'F' and 'P'. **(a)** The output plane. **(b)** Cross section of the correlation peaks .

more than two channels implementing measurements on more than two distortion parameters. To keep this work within reasonable limits, the discussion here was restricted to the problem of three parameter distortion invariant pattern recognition (scale, shift and rotation). Laboratory experiments using alphanumeric characters demonstrated excellent performance in agreement with computer simulations.

Because the overall recognition process is performed in two stages, the main consideration is the operation speed of every channel and the time it takes to adapt the correlator. The advantages and efficiency of the two-stage procedure are judged primarily by the speed criterion. Special emphasis was put on this subject by using real-time components such as SLM's and CCD's, both in the scale measurement channel and in the correlation channel. This lead to a very rapid scale measurement procedure (see section 2) while the correlation channel, using a static *a priori* designed filter produces the correlation results in the speed of light. Thus, practically, the correlator may be used in real time.

The real time adaptive correlator enables us to use the system for many related applications, such as iterative filter design, pattern classification etc.

Acknowledgement: This work was performed within the Technion Advanced Opto-Electronics Center established by the American Technion Society (ATS), New York.

References

[1] E. Silvera, T. Kotzer and J. Shamir, "Adaptive pattern recognition with rotation, scale and shift invariance", Submitted for publication to *Appl. Opt.*

[2] A. Levi and H. Stark, "Signal restoration from phase by projections onto convex sets," *J. Opt. Soc. Am.*, **73**, 810–822, (1983).

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Application of optical multiple-correlation to recognition of road signs: the ability of multiple-correlation

K Matsuoka, M Taniguchi and Y Mokuno

Osaka National Research Institute, 1-8-31 Midorigaoka, Ikeda, Osaka 563, JAPAN

Abstract. We apply correlation filters to the location problem of multiple objects in a real scene and discuss the ability of optical correlators. Firstly, we design a correlation filter to reduce false signals caused by background and noise in an input scene. Then we introduce a color processing for locating multiple road signs. The performances of these filtering are verified by the computer simulations.

1. Introduction

Optical correlators are one of most powerful techniques for locating multiple objects and they are expected as a real-time processor for a real world scene. Processing for a real world scene requires robustness against object distortions, noise, and background of an input scene. Especially, robustness against background is very important because many unknown images will appear in the background of real scenes. These unknown images are too many to use as training images in the filter design process.

Many correlation filters have been developed to achieve distortion invariant processing (e.g. [1]-[4]). However, they are mainly designed to achieve the invariance for the object shape distortion and the robustness against the background has not investigated well.

In this paper, we apply multiple correlation filters to a location problem for multiple objects, that are road signs here, in a real scene and discuss their abilities of locating and robust processing.

2. Design of robust filter for unknown background

A final result of locating is obtained by thresholding the correlation pattern of an input image and a correlation filter. Therefore, it is needed to discriminate the true signals from the false signals caused by noise, background, and sidelobes around a correlation peak by an appropriate threshold level. Some correlation filters have been proposed (e.g. [1]-[4]) which can suppress the sidelobes or reduce the ill effects of noise.

To get an optimum filter improving the discrimination, we examined the abilities of well-known filters to reduce the false signals. We designed 5 filters, the minimum noise and correlation energy (MINACE) filters [3] with different noise variance, from the training images in Fig.1 and evaluated the performance of them by applying to a real scene. The designed MINACE filters are to detect the #0 and #1 patterns of Fig.1.

Table 1 shows the characteristics of filters; the noise variance σ^2 used at the filter

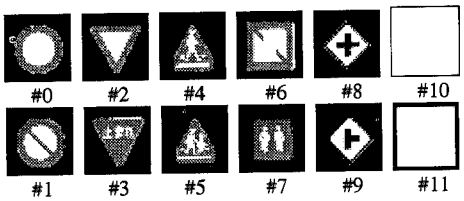


Fig.1 Training set images for the evaluation of correlation filter.

Table 1. Evaluation result of MINACE filter by changing noise variance σ^2.

σ^2	filter energy $\|f\|^2$	PSR		maximum value for true pattern		maximum value for false pattern	
		true	false	inside	outside	inside	outside
SDF filter	10.86	0.45	0.15	1.0	0.10	0.31	0.39
σ^2=50	10.90	0.57	0.17	1.0	0.09	0.29	0.35
σ^2=10	11.44	0.96	0.21	1.0	0.06	0.19	0.20
σ^2=1	15.72	0.96	0.16	1.0	0.02	0.06	0.07
σ^2=0	34.23	0.94	0.08	1.0	0.02	0.03	0.03

synthesis on the first column, the energy of correlation filter (the sum of the squares of the pixel values of filter) on the second column, the average ratios of the peak intensity to sidelobe intensity (PSR) for the true signal and for the false signal on the third and forth columns, respectively, and maximum pixel values inside the peak area and outside the peak area on the fifth to eighth columns. The peak area in this case is the central 10x10 pixel region.

A filter with lower filter energy can more reduce the ill effects of noise and a filter with larger PSR for the true pattern generates smaller sidelobes. However, any of designed filters does not have both of a low filter energy and a large PSR together.

Figure 2 shows the correlation results by applying the designed MINACE filters to the test scene in Fig.2 (a). The brightness of these results is reversed. Each threshold level for the discrimination was determined to be the largest value under the condition in which two objects in the test scene can be detected at least. In this experiment, Fig.2 (d) shows the best result to reduce false signals, while other results have many false signals as shown in Fig.2 (b),

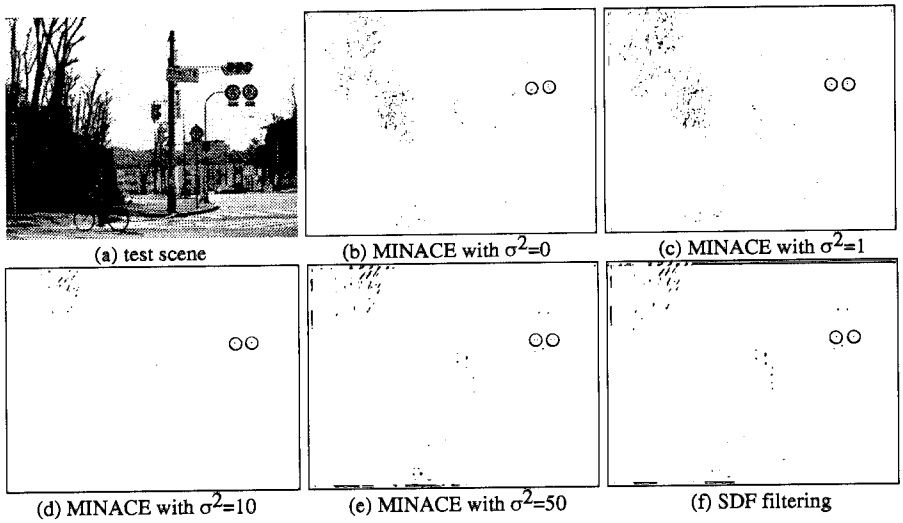


Fig. 2 Results of MINACE filtering and SDF filtering on the test scene (a).

(c), (e) and (f) due to the large filter energy and the poor PSR.

These results denote that a correlation filter can be designed to reduce false signals caused by background and a filter with a small filter energy and a large PSR for true pattern is preferable. It seems difficult to eliminate all false signals by using a single correlation filter.

3. Color image processing for locating road signs

Single correlation filtering described in the previous section can not eliminate false signals perfectly. To reduce the false signals, we introduce multiple correlation to perform a color image processing as shown in Fig.3.

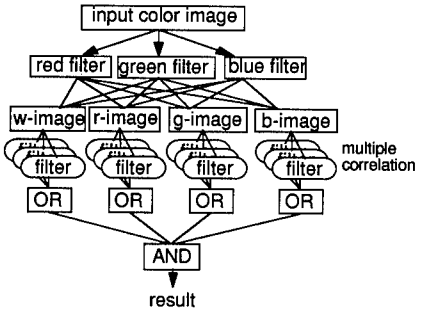


Fig. 3 Procedure for color image processing.

In this processing, an input image separates to color component images (red, green, blue and gray components), and correlation filters are applied to individual color components. The final result is obtained by applying AND operation to all component results; i.e. locating signals appear only when the objects are detected at the same position in all component images. To improve the scale invariance, each component processing is performed by multiple correlation, in which different size of objects can be detected by different filters.

We applied this color processing to the problem of locating red road signs in a real scene. In this application, we used only two color components; i.e. gray component and red component, because all objects have mainly gray and red components. Each correlation filter is synthesized with 33 training images and 9 types of correlation filers are provided for detecting various sizes of object.

Figure 4 shows the example of correlation peak values obtained by 9 correlation filters and objects with various sizes. The horizontal axis denotes the object size and the vertical axis denotes the correlation peak value. It is demonstrated in Fig.4 that almost all the size of object can be detected by one of these filters by an appropriate threshold level.

Figure 5 shows the computer simulation results by applying this color processing to three different test scenes shown in Fig. 5 (a), (b) and (c); each of which has three objects,

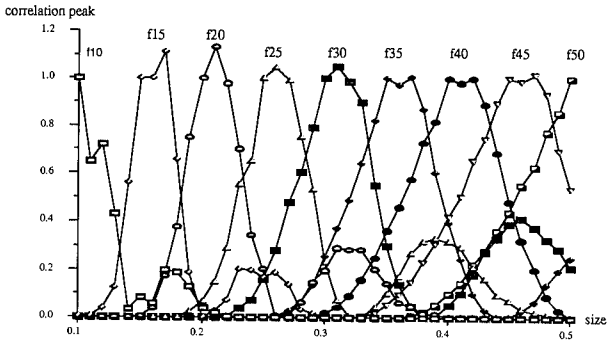
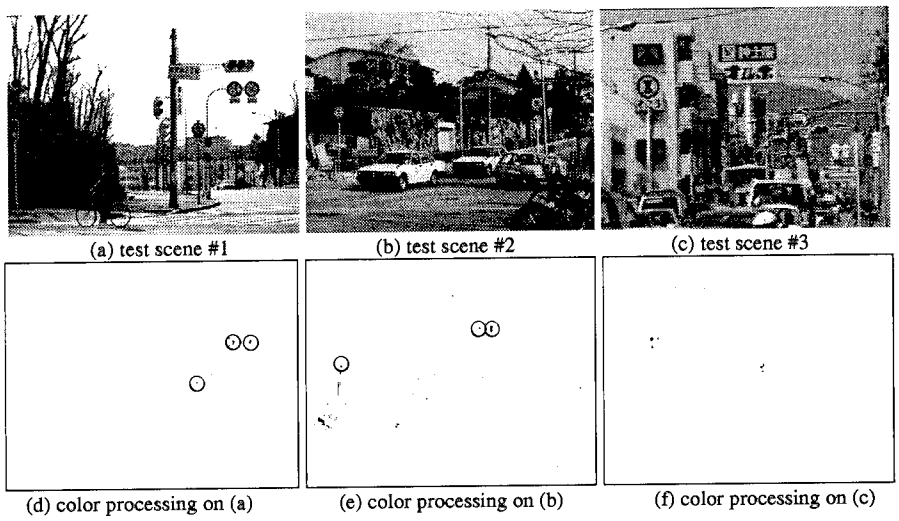


Fig. 4 Correlation peak values obtained by the SDF filters (f10, f15, ..., f50) and different size objects.



(a) test scene #1 (b) test scene #2 (c) test scene #3

(d) color processing on (a) (e) color processing on (b) (f) color processing on (c)

Fig. 5 Results by applying color image processing to the three test scenes.

four objects, and two objects, respectively. These test scenes have different image complexity. Figure 5 (d), (e) and (f) show the processed results for test scenes (a), (b) and (c), respectively. Perfect result is obtained in Fig.5 (d), i.e. three objects can be perfectly located without false signals while single correlation filtering could not eliminate the false signals as described in the previous section. In other results in Fig.5 (e) and (f), only few false signals appear though perfect locating cannot be performed. These results denote the great ability of multiple correlation to achieve robust processing against the background.

4. Conclusion

To investigate the ability of optical correlators, we applied the multiple correlation to the problem of locating multiple road signs in a real scene. By the computer simulations, we showed that a well-designed correlation filter can reduce false signals caused by background of an input image, but it is difficult to eliminate all false signals by a single correlation filter. Then we introduced a multiple correlation to perform a color image processing. The computer simulation results of color processing denoted a multiple correlation has a great ability to suppress false signals compared with the single correlation. However, this color processing cannot perfectly succeed to locate road signs in a real scene. More robust processing will be needed for the real world scene recognition.

References

[1] Casasent D 1984 Appl. Opt. 23 1620-1627
[2] Mahalanobis A, Kumar B V K V and Casasent D 1987 Appl. Opt. 26 3633-3640
[3] Ravichandran G and Casasent D 1992 Appl. Opt. 31 1823-1833
[4] Casasent D, Ravichandran G and Bollapragada S 1991 Appl. Opt. 30 5176-5181

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Parametric design of correlation filters for multiple-object discrimination

M. Taniguchi and K. Matsuoka

Osaka National Research Institute, 1-8-31 Midorigaoka, Ikeda 563, Osaka, JAPAN

Abstract. We propose a new algorithm for the design of multiple-object discriminant correlation filters. This technique makes it possible to control the characteristics of the filter. By considering the properties of recording media in the filter synthesis, improvements in filter design are possible. Results of computer simulations are demonstrated.

1. Introduction

The technique of using matched spatial filters for pattern classification or pattern recognition has been well studied. Many methods have been proposed to use the filters for detection of patterns in the presence of noise and distortion, or for multiple-object classification (e.g., [1]). However, as the functions of the filter increase, the more information that has to be recorded on it. Therefore, if recording media have insufficient gray level, the filter cannot be recorded correctly. Because of this inaccuracy, the expected correlation response is not obtained.

We studied the case of recording correlation filters on discrete type recording media, such as discrete detour phase computer-generated holograms (CGHs). In this type of CGHs, the Fourier coefficients of filter function are encoded discretely. Therefore, the amplitude and phase components for each pixel of the filter are quantized. We studied the effects of quantization in the correlation peak value for the cases of two popular types of multiple-object discriminant filter, namely, synthetic discriminant function (SDF) filter [1] and minimum average correlation energy (MACE) filter [2]. It was shown that the effects are not negligible.

In this paper, we propose a new type of spatial filter for pattern classification. The proposed filter is calculated by using a simulated annealing (SA) algorithm [3] with considering the properties of a recording medium. The filter is convenient to implement on a recording medium which has poor gray level. In addition, the characteristics of the filter, such as correlation peak value at the origin, sidelobe level and sensitivity for noise and distortion, can be controlled parametrically by the method. Results of computer simulations showed that the filter had the expected correlation response.

2. Effects of quantization on correlation

In this section we show the effects of quantization on the correlation results. We studied the case of recording filters as the discrete detour phase CGHs, which are the modified version of Lohmann type CGHs [4] and are suitable for recording. The CGHs consist of small regions or cells corresponding to each Fourier coefficient. Each cell consists of more than one subcell.

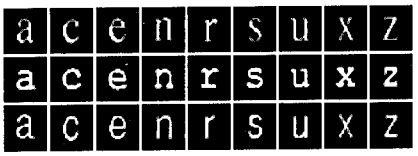


Fig. 1. Training patterns.

(a1) (a2) (c1) (e1)

Fig. 2. Test patterns.

Table 1. Correlation responses of the filters

input pattern	SDF	MACE	qSDF	qMACE
a1	1.00	1.00	1.19	0.70
a2	1.00	1.00	1.08	0.73
c1	0.00	0.00	0.20	0.01
e1	0.00	0.00	0.11	0.00

In each cell, the amplitude component of the coefficient is controlled by the number of open subcells and the phase one is determined by the positions of open subcells. If each cell consists of 4 x 4 subcells, the amplitude and phase components are quantized to 5 and 4, respectively.

To estimate the effects of quantization on the correlation results, we calculated the SDF filter and MACE filter. Figure 1 shows the training patterns used for filter synthesis. Each pattern was decimated to 32 x 32 pixel resolution. The filters were designed to distinguish the character "a" from the others. Therefore, we specified the correlation peak values of 1.0 and 0.0 for patterns of "a" and the others, respectively. We also calculated the quantized versions of the filters encoded with 5 amplitude and 4 phase levels. These filters shall be called qSDF and qMACE, respectively. Figure 2 shows the test patterns used for the estimation. The patterns are included in the training patterns. Table 1 lists the values of intensity at the origin of correlation plane. The correlation values deviate by 20 percent (SDF) or 30 percent (MACE) from which specified values. These effects detract from the advantage of SDF or MACE filter that the correlation values are specifiable.

3. Filter synthesis

We applied a simulated annealing (SA) algorithm to the construction of correlation filters recorded as discrete detour phase CGHs. When we apply the algorithm to an optimization problem, an energy function to be minimized has to be determined. In this paper, we introduce three target values as the degree of the characteristics of the filter, namely, correlation peak value at the origin, sidelobe level and sensitivity for noise and distortion. Thus we use the weighted sum of the three target values as the energy function.

The correlation peak values are evaluated by the difference between the actual correlation value and the specified one. The target value E_1 is defined as follows:

$$E_1 = \sum_{i=1}^{N} \left(\left| \iint T_i^*(X, Y) \cdot F(X, Y)\, dXdY \right|^2 - c_i \right)^2, \qquad (1)$$

where N is the number of training patterns, $T_i(X, Y)$ $(i = 1, 2, ..., N)$ is the Fourier transform of training pattern, $F(X, Y)$ is the filter function to be obtained, c_i is the specified correlation value for the training pattern and * denotes the complex conjugate.

The sidelobe level is estimated by the energy of correlation plane. The target value E_2 is given by,

$$E_2 = \sum_{i=1}^{N} \iint |T_i^*(X, Y) \cdot F(X, Y)|^2 \, dXdY \quad . \tag{2}$$

The sensitivity for noise and distortion depends on the transmissivity of the filter in the high frequency area. Therefore, we define the target value E_3 as

$$E_3 = \sum_{i=1}^{N} \iint (X^2+Y^2) \cdot |F(X, Y)|^2 \, dXdY \quad . \tag{3}$$

Therefore, the energy function E to be minimized is defined as

$$E = \lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 \quad , \tag{4}$$

where λ_1, λ_2 and λ_3 are positive constants. The filter function which minimizes the energy function is obtained by the SA algorithm. The characteristics of the filter depend on the parameters. In other words, the properties of the filter can be controlled parametrically.

4. Results of computer simulations

We designed many filters by the SA algorithm using the different parameters. Figure 3 shows the typical correlation planes for (a) qSDF filter, (b) qMACE filter, (c) filter designed by SA with parameters (λ_1, $\lambda_2 > 0$, $\lambda_3 = 0$), (d) same as (c) with parameters (λ_1, $\lambda_3 > 0$, $\lambda_2 = 0$), (e) same as (c) with parameters (λ_1, λ_2, $\lambda_3 > 0$). These filters designed by SA shall be called SA1, SA2 and SA3, respectively. The input patterns are as those shown in Fig. 2.

We also tested the correlation response for other patterns which are not included in the training patterns. The patterns used for the test are in Fig 4. All of the patterns are distorted or noisy versions of the pattern (a1) of Fig. 2. Table 2 and 3 list the correlation peak values at the origin and the values of maximum sidelobe for the five filters, respectively. The results for filters designed by SA shows that the sidelobe level and the sensitivity for noise or distortion vary depending on the parameters. Filter SA1 shows sharp correlation peaks, filter SA2 has tolerance for degraded patterns and filter SA3 has characteristics between SA1 and SA2.

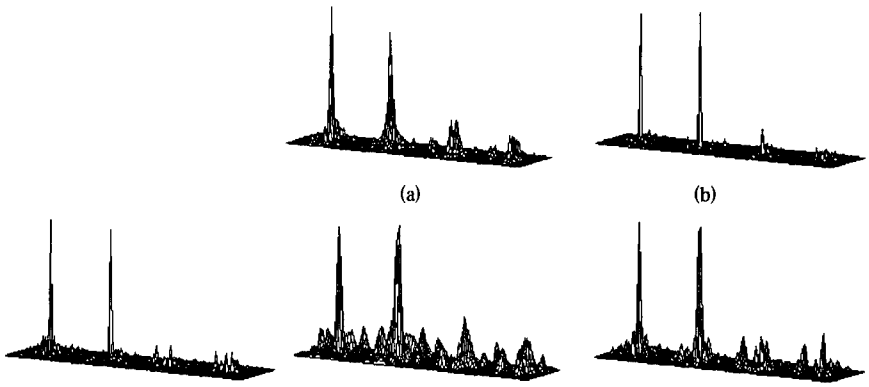


Fig. 3 Typical correlation planes for (a) qSDF, (b) qMACE, (c) SA1, (d) SA2, (e) SA3.

(a) (b) (c) (d)

Fig. 4. Test patterns not included in the training patterns.

Table 2. Correlation peak values for patterns not included in the training patterns

input pattern	qSDF	qMACE	SA1	SA2	SA3
(a)	0.64	0.23	0.02	1.14	0.90
(b)	0.47	0.20	0.01	0.70	0.41
(c)	0.26	0.11	0.18	0.44	0.40
(d)	0.79	0.41	0.66	0.86	0.89

Table 3. Maximum sidelobe for patterns not included in the training patterns

input pattern	qSDF	qMACE	SA1	SA2	SA3
(a)	0.21	0.10	0.06	0.64	0.21
(b)	0.19	0.21	0.08	0.55	0.32
(c)	0.15	0.12	0.07	0.31	0.41
(d)	0.15	0.21	0.06	0.41	0.31

Sidelobe is defined as the intensity in the correlation plane excluding a 5 x 5 region around the origin.

5. Summary

We have proposed a new technique for the design of multiple-object discriminant correlation filters. In the filter synthesis, improvements in filter design are possible by considering the properties of recording media. In addition, the characteristics of the filters can be controlled parametrically.

References

[1] Casasent D 1984 *Appl. Opt.* **23** 1620-1627
[2] Mahalanobis A, Vijaya Kumar B V K and Casasent D 1987 *Appl. Opt.* **26** 3633-3640
[3] Aarts E and Korst J 1988 *Simulated Annealing and Boltzmann Machines* (John Wiley & Sons)
[4] Brown B R. and Lohmann A W 1966 *Appl. Opt.* **5** 967-969

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

An Analog Retina for Edge Detection

Chunyan WANG and Francis DEVOS

Institut d'Electronique Fondamentale, Bat 220, Université Paris Sud
91405 Orsay Cedex FRANCE
Tel. 33-1-69416574

Abstract We present an optoelectronic analog test retina fabricated in a standard CMOS process. The retina is to be reproduced into a pixel array. In each pixel there is a photodiode and an elementary operator consisting of 21 transistors. As the operator is a "programmable" analog cell, it is possible to recombine the processing sequences in the circuit in order to diversify its functions and to achieve a spatial and/or temporal differentiation. By testing the prototype, the performance for the acquisition and edge detection of the circuit is evaluated.

1. Introduction

The retina is a part of a system of classification of road signs with an optical correlation [1]. The system consists of two matrices, each of which has 64*128 cells. Controlled by a FLC micro mirror device[2], the first matrix displays the images sequentially from an "edge library". The second one performs four functions: acquiring the signal of an image carried by the incident light; detecting the edges of the image; displaying the edge image and eventually outputing the processed data. The circuit presented here is an elementary test retina built for the evaluation of the second matrix.

2. Analysis of the problem

Our goal is to integrate a matrix of 8192 cells in a reasonable size (≈ 0.5 cm2). Therefore, the functions of a cell should be simple enough to be achieved with fewer than thirty transistors. To fulfil this task within its limitations we have to choose :

1. The implementation via an analog circuit, as there is no question of a digital one.
2. A compromise between the simplicity of the circuit and the processing time.

The algorithm Sobel is one of the most common methods for differential edge detection, but because of the non-feasability of a large number of the interconnections (due to its nature of an 8 pixel neighbourhood) we didn't choose it. The idea of implementing the detection method is originated from Roberts [3], which, by means of the convolution, is only a four pixel neighbourhood. The following are the two modifications of the Roberts method:

1/ To get the detection efficiency comparable to that of Sobel, we modified Roberts by introducing two supplementary differentiators into the template impulse response array (Fig.1). The comparison of the algorithms, Roberts, modified Roberts and Sobel is made using the Khoros simulation. The results show that with the modified Roberts method the compromise between the performance and the complexity can be reached (Fig.2).

NW	NE
SW	SE

1	0
0	-1

H_{R1}

0	1
-1	0

H_{R2}

1	-1
0	0

H_{S1}

1	0
-1	0

H_{S2}

Fig.1 H_{R1}, H_{R2} : classic Roberts differentiators H_{S1}, H_{S2}: supplementary differentiators

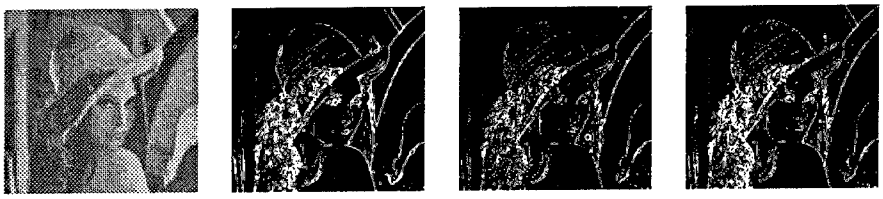
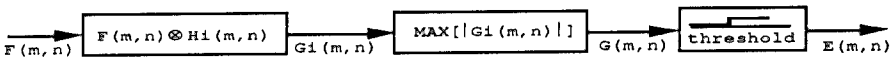


Fig.2 (a) (b) (c) (d)
(a) Original image (b) Edge map of Sobel (c) Edge map of Roberts
(d) Edge map attained by using the four Robert operators. Some diagonal lines missing on Fig.2-c show up here.



F(m,n) : original image Gi(m,n) : gradient in the i th direction
Hi(m,n): spatial differentiator G (m,n) : gradient after the compass E (m,n) : edge map

Fig.3 General procedure for Roberts detection

2/ The other modification is in the procedure shown in Fig.3, done in order to reduce the varieties and the amount of the physical operators to be implanted. If the variation of the incident light between two samples is negligible, it is possible to simplify the circuit by extending the processing sequence. In short, the two calculations in the procedure shown in Fig.3 can be eliminated by means of two replacements :

1/ eliminating "absolute value": Gij = | Fi -Fj | = Max {(Fi -Fj), (Fj -Fi)}

2/ eliminating "max value": E =Sign (Gmax - thr.)= OR{Sign (Gij - thr.), ≠ differentiators}

[y = sign (x) =1 case x≥0 ; otherwise, y = 0]

Thus, the general procedure is transformed into the iterated simple sub-procedures (Fig.4-b) which are composed of only three operations : accumulation/subtraction, comparison and memorisation of logic "1"(OR in terms of time).

We can make such an arrangement to complete a calculation for each pixel:

1/ The execution of the sub-procedure can be iterated successively with the eight differentiators (Fig.4-a) which choose the combination of the input data.

2/ Each execution can be started by the signal acquisition. The elimination of the input memory is performed at the cost of the iterative signal acquisitions.

3. Design and implantation of the circuit

3.1. Circuit design principle

The procedure of Fig.4-b makes it possible to produce a cellular system (fig.5) with a single sequential SIMD control. Functionally, each pixel consists of :

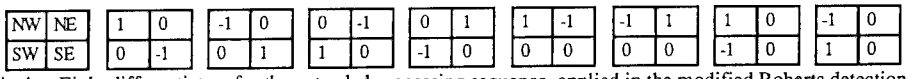


Fig.4-a Eight differentiators for the extended processing sequence, applied in the modified Roberts detection.

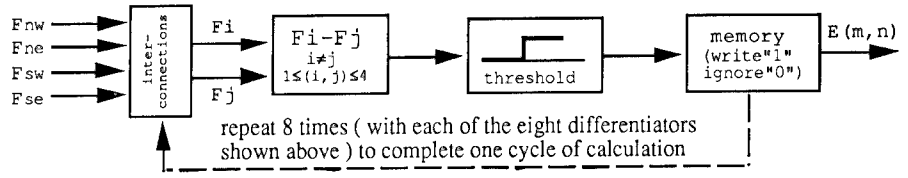


Fig.4-b Simplified procedure of the modified Roberts detection for the "parallel and squential" implantation.

1/ an elementary operator which executes the three basic operations; 2/ a photodiode for the optical data acquisition; 3/ the interconnections which make each photodiode connectable at anytime to only one of its four neighbouring elementary operators according to the address code; 4/ a control logic which programmes the function of the whole system (operative code and address code). This control can carry out the real programmable algorithms for several spatio-temporal convolutions.

3.2. *Detail of the elementary operator and evaluation of the performance*

The diagram of the operator and its four interconnection switches (K1-4) are shown in Fig.6.

The left part is for the acquisition and subtraction. The input is the photocurrent produced alternatively by one of the four neighbouring photodiodes. With a double integration (first additive then subtractive) via the sole current mirror, the difference between the incident intensities is transformed into Vc which is the voltage on the capacity "C" (Fig.7). The bias of each photodiode is stablized by the transistor "Tbias" in order to get the relatively linear transformation of Δintensity-voltage.

In Fig.6 (right side) a "quasi-static" binary memory cell has three functions : 1/ comparing "Vc" with the threshold (its value depends on the transistor sizes here); 2/ writing and memorizing only logic "1", which happens when the result of the comparison is "true" (one point of edge is detected); 3/ serving as a shifter. The data on each pixel can be outputed by successive shifting through the entire neighbourhood.

The function of a matrix with four elementary operators and nine photodiodes (Fig.5) was simulated with a Hspice simulator in Cadence environment before the implantation.

3.3. *Implantation of the prototype*

A prototype circuit is processed with CMOS technology : ES2 1.5 μm (dual layer metal and single poly). This prototype (Fig.8) contains an elementary operator (as in Fig.6, surface size: 50 * 92 μm2), four photodiodes (each size : 30 * 30μm2) and a capacity of 1pf (size : 100*100 μm2) which will not exceed 0.35pf in the final version owing to the total circuit size.

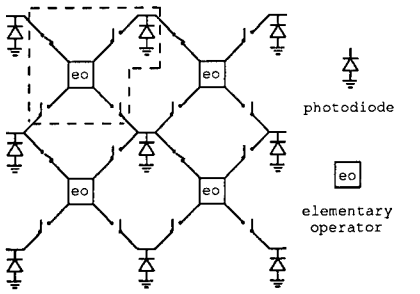


Fig.5 Pixel array

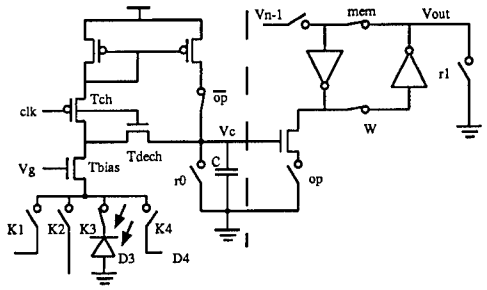


Fig.6 Diagram of an elementary operator

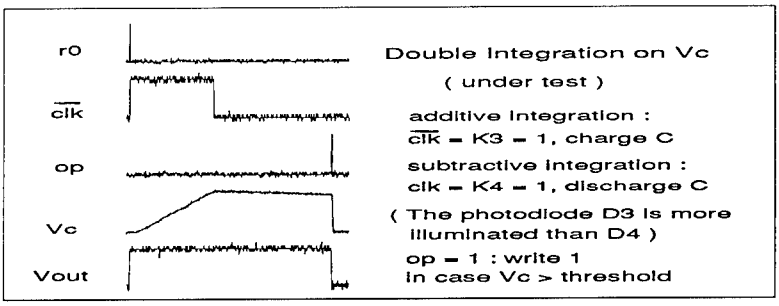


Fig.7 Waveform shows the subtraction of the photocurrents achieved by the double integration on "Vc".

4. Test results of the prototype, conclusion and perspectives

Having tested the prototype we then verified the global linearity of the analog operator (acquisition and spatial or temporal differentiation).The results (Fig.9) show that when the sample rate for the incident image is within the limits of 10 kHz, the non-linearity percentage was in the lower 10 per decade range.

During the tests we were confronted with two problems : the problem of cutting off the photocurrent by the interconnection switches (due to the photo-absorption of the light incident upon the substrate); the strong sensitivity of the current mirror to the parametric dispersion in weak current range. These two points were considered in a later version of the circuit. In that version, in order to lower its potential enough to turn the switch off, a control signal is first applied to a transistor follower whose source is connected to a tiny additional photodiode. For the balance compensation of the current mirror we inserted a capacity between the two grids of the mirror and adjusted the charges by Tunnel Effect [4] to counter-balance the potentials on the two grids. Consequently we obtained the "reflected" current which followed closely the "original" one, the photocurrent. Therefore the two coefficients (one for the additive integration and the other for the subtractive one) coincide.

The work we presented here gives an example of the computing implementation that adapts to an analog cellular circuit. Its key points : 1/ preserve the density of the cellular computing; 2/ exploit the variety, density and nature of the spatial distribution — the utilizable mechanism; 3/ verify the resolution and global linearity of the completed operators; 4/ take into account the spatial dispersion for eventual compensation; 5/ preserve the minimal resources of configuration (switches, operation codes, parameters, etc) for the minimal adaptability of the control sequences.

Acknowledgements — We thank our colleagues R.Reynaud, S.Bouaziz and F.Lavainne for valuable discussions and help on the circuit tests, with particular thanks to C.Arcile for providing the CAD tools.

References

[1] L.Guibert, G.Keryer, P.Pellat and J.L.de Bougrenet de la Tocnaye,*Compact Optical Correlator for Roadsign Recognition*

[2] R.Moignard,E.Daniel, P.Cambon and J.L.de Bougrenet de la Tocnaye 8th Workshop on Optical Computing 1992 *Design of a Silicon VLSI/FLC Smart Light Valve for Parallel Optical Information Processing*

[3] William K.Pratt Wiley Interscience 1991 *Digital Image Precessing*

[4] M.Zhang, J.F.Pone and F.Devos GCIS Grenoble Fevrier 1993 *Etude d'une Cellule de Mémorisation Analogique non-volatile par EFFET Tunnel en Technologie Standard CMOS 2,4um*

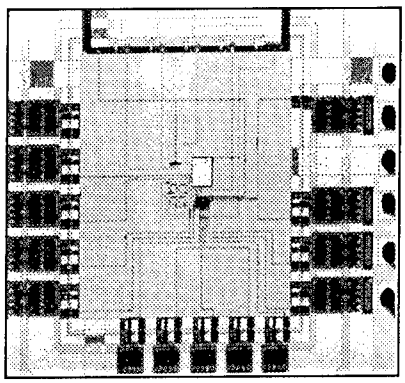


Fig.8 Photograph of the prototype

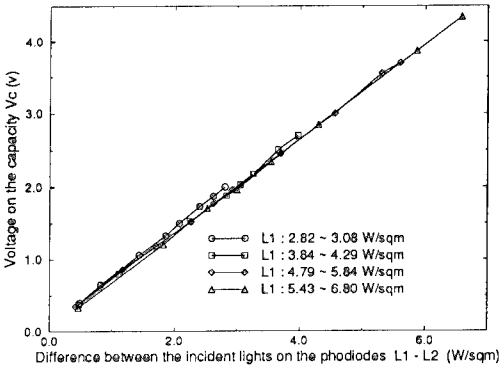


Fig.9 Characteristic of the Vc/ΔL transformation

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optoelectronic implementation of a phase-retrieval Vander Lugt correlator.

Santiago Vallmitjana, Arturo Carnicer, Estela Martín-Badosa, Ignacio Juvells.

Universitat de Barcelona, Laboratori d'Optica, Departament de Física Aplicada i Electrònica. Diagonal 647, E08028 Barcelona, Spain.

Abstract. An implementation of a Vander-Lugt correlator, which operates with a single spatial light modulator is proposed. Optical phase-retrieval manipulation, based on the symmetrization of the input scene is required. Theoretical analysis, simulations and some experimental results are presented.

1. Introduction.

In recent years, some authors have proposed the same compact correlation architecture [1-3] for different purposes. These setups are based on the use of a spatial light modulator (SLM) to introduce and display information, a Fourier lens system, a video-camera (CCD) to register light distributions and a computer that controls the whole system (Figure 1).

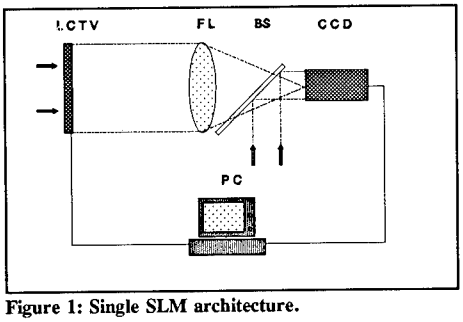


Figure 1: Single SLM architecture.

If the setup performs as a joint transform correlator, both the scene and the reference are displayed on the SLM and the CCD in the Fourier plane registers the joint interferences between them. This power spectrum is processed by a computer and displayed again on the same SLM. A second optical Fourier transform is obtained and the correlation is recorded by the CCD. Rosen et al. have proposed the use of the same architecture to implement Vander-Lugt correlators [4], in which the scene is shown on the modulator and its power spectrum is recorded by the CCD. To recover the phase, they register the joint interferences between the object and a plane wave, and in a second step, the reference is handled in the same way. Finally, the correlation between the scene and the target is obtained.

In this article we present an alternative to the original proposal that simplifies the experimental procedure and improves the method's capabilities.

2. Theoretical analysis.

The idea of retrieving the phase lost in the recording process is based on the symmetrization of the images used, as in the following expression: f(x-a, y)+f(-x-a,-y), where 2a is the separation between the original image and its symmetrical. Figure 2a show the symmetrized scene (three satellites). The images should be placed as close as possible, without superposition, in order to take advantage of the spatial bandwidth of the modulator. When there is interference between the light distribution and a plane wave in the focal plane, the following intensity is obtained:

$$| \, | F(u,v) | \exp(i\phi(u,v))\exp(-i2\pi au/\lambda f) + | F(u,v) | \exp(-i\phi(u,v))\exp(i2\pi au/\lambda f) + C \, |^2$$

$$= | \, 2 \, | F(u,v) | \cos[\phi(u,v)-2\pi au/\lambda f] + C \, |^2 \qquad (1)$$

where f is the focal of the Fourier lens system.

By using a square-root input response CCD and if the value of the amplitude plane wave C is high enough, the equation (1) becomes,

$$2 \, | F(u,v) | \cos[\phi(u,v)-2\pi au/\lambda f] + C \qquad (2)$$

The constant value C, can easily be subtracted by computing. The reference g(x,y) can be processed in a similar way: g(x+a,y)+ g(-x+a, -y). Figure 2b show the geometry used in this case. In the Fourier plane, the light distribution can be written as

$$| \, | G(u,v) | \exp(i\gamma(u,v))\exp(i2\pi au/\lambda f) + | G(u,v) | \exp(-i\gamma(u,v))\exp(-i2\pi au/\lambda f) + C \, |^2$$

$$= | \, 2 \, | G(u,v) | \cos[\gamma(u,v)+2\pi au/\lambda f] + C \, |^2 \qquad (3)$$

and finally considering the response of the CCD, equation (3) can be obtained in the following terms:

$$2 \, | G(u,v) | \cos[\gamma(u,v)+2\pi au/\lambda f] + C \qquad (4)$$

The product between $| F(u,v) | \cos[\phi(u,v)-2\pi au/\lambda f]$ and $| G(u,v) | \cos[\gamma(u,v)+2\pi au/\lambda f]$ is performed in the computer:

$$| F(u,v) | \, | G(u,v) | [\exp(i\phi(u,v))\exp(-i2\pi au/\lambda f) + \exp(-i\phi(u,v))\exp(i2\pi au/\lambda f)]$$
$$[\exp(i\gamma(u,v))\exp(i2\pi au/\lambda f) + \exp(-i\gamma(u,v))\exp(-i2\pi au/\lambda f)] = \qquad (5)$$

$$| F(u,v) | \, | G(u,v) | \, [2\cos(\phi(u,v)+\gamma(u,v)) +$$
$$\exp(i\phi(u,v)-i\gamma(u,v)-4\pi iau/\lambda f) + \exp-(i\phi(u,v)-i\gamma(u,v)-4\pi iau/\lambda f)] \qquad (6)$$

This result is displayed in the SLM and in the focal plane we obtain the cross-correlation between f(x,y) and g(x,y) at a distance $\pm 2a$ from the origin. As a conclusion, the function $| G(u,v) | \cos[\gamma(u,v)+2\pi au/\lambda f]$ can be considered as a classical matched filter (CMF). If the amplitude of this filter is modified properly, several spatial correlation filters can be designed.

3. Simulated results

Several simulated results have been carried out in order to show the procedure described. Figure 3a presents the cross-correlation between the scene of Figure 2a and the reference of Figure 2b. As the filter has not been processed, it corresponds to the CMF.

A second set of simulations has been carried out taking into account a simplification method obtained by binarizing the distributions of equation (6). The binarization is made considering the zero value as the threshold. As a result, Figure 3b shows a representation of the cross-correlation between the scene and the satellite in Figure 2b, obtained with this method.

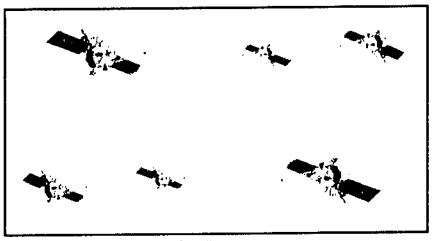


Figure 2a.- Symmetrized scene.

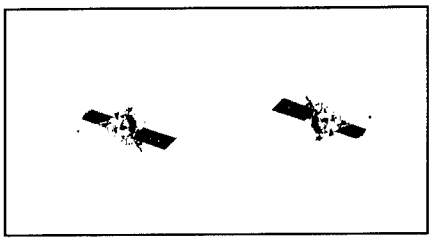


Figure 2b.- Symmetrized reference.

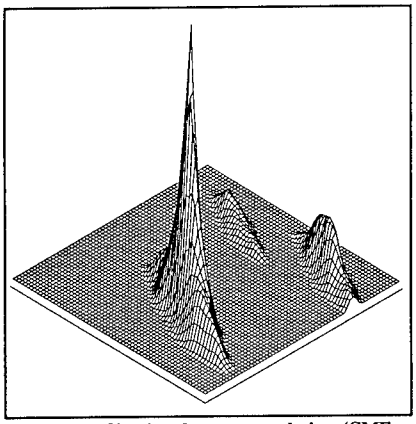


Figure 3a.- Simulated cross-correlation (CMF filter).

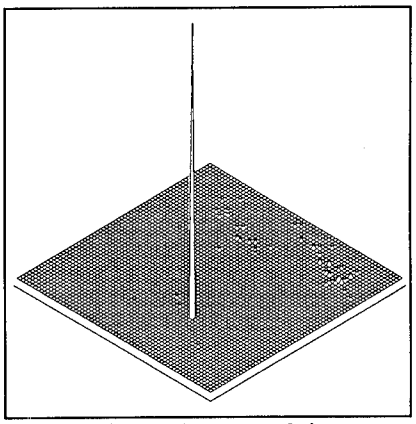


Figure 3b.- Simulated cross-correlation (binarization method).

4. Experimental results.

A single SLM correlator which operates with a low cost liquid crystal television (LCTV) has been implemented [5]. The LCTV used was removed from an Epson 100PS videoprojector [6]. The symmetrized scene is displayed in the LCTV. A CCD videocamera with a square root input response was connected to an 8-bit digitizer board. The LCTV is illuminated and the light distribution in the Fourier plane of the lens system interferes with a plane wave. This intensity is registered by the CCD and stored in the computer memory. In a second stage, the power spectrum of the symmetrized reference is stored in the same

way. Then, the product between these distributions is computed and the result is displayed in the LCTV. Finally, the correlation is obtained in the focal plane of the Fourier lenses.

The experiments have been carried out with this correlator using the matched filter and binarizing the distribution in the LCTV. Figure 4a show the cross-correlation zone using the CMF and Figure 4b the detection obtained with the binary method.

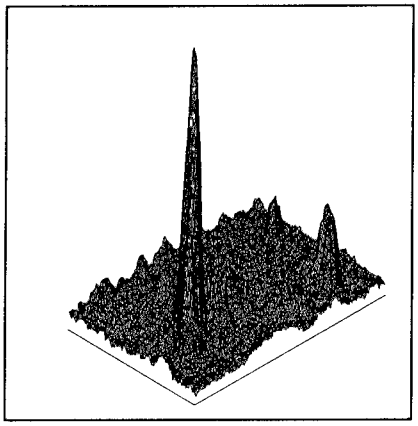


Figure 4a.- Optical correlation (CMF filter).

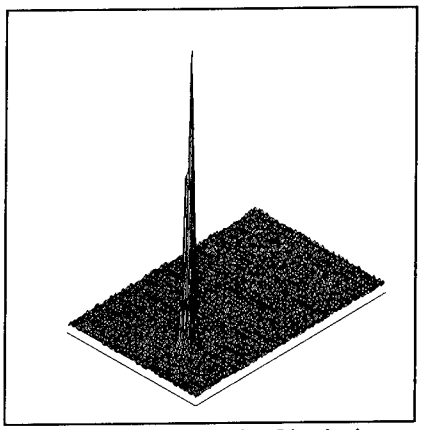


Figure 4b.- Optical correlation (binarization method),

5. Conclusions

The use of single SLM architecture allows the simultaneous implementation of Vander-Lugt and joint transform correlator with similar recognition capabilities. The possibility of incrementing the discrimination using the non-linear properties of binarization in Fourier plane are shown.

Acknowledgements.- This paper has been supported in part by the Spanish CICYT (Comisión interministerial de Ciencia y Tecnología) under contract ROB91-0554. E. Martín-Badosa gratefully acknowledges a grant of the Generalitat de Catalunya.

6. References

[1] B. Javidi and J. L. Horner, "Single spatial light modulator joint tranform correlator", 1989, Appl. Opt **28**, 1027-1032.

[2] T. Kotzer, J. Rosen and J. Shamir, "Phase extraction pattern recognition", 1992, Appl. Opt. **31**, 1126-1137.

[3] T. Kotzer, J. Rosen and J. Shamir, "Multiple-object input phase extraction correlation", 1993, Appl. Opt. **32**, 1919-1932.

[4] J. Rosen, T. Kotzen and J. Shamir, "Optical implementation of phase extraction pattern recognition", 1991, Opt. Commun., **83**, 10-14.

[5] A. Carnicer, E. Martín-Badosa, I. Juvells and S. Vallmitjana, "Spatial envelop-free nonlinear joint transform correlation", 1994, Opt. Commun (submitted).

[6] J. C. Kirsch, D. A. Gregory, M. W. Thie and B. K. Jones, "Modulation characteristics of the Epson liquid crystal television", 1992, Opt. Eng. **31**, 963-969.

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 British Crown Copyright

Performance of an acousto-optic joint transform correlator

C D Reeve

SECEE, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK

A W Houghton

EE Department, Royal Naval Engineering College, Manadon, Plymouth, PL5 3AQ, UK

F D Bryant

Department of Physics and Astronomy, Western Kentucky University, Bowling Green, KY 42101-3576, USA

Abstract. This paper describes a joint transform correlator, whose input is provided by two acousto-optic cells. The performance of the correlator is analysed and modelled and results from a practical system are presented. It is shown that the system is capable of detecting and direction finding spread spectrum signals which are more than 10 dB below the noise level.

1. Introduction

When attempting to detect a spread spectrum signal of unknown waveform and bandwidth, which may be 10 dB or more below the noise level, a single receiver is generally unable to form the matched filter for the transmitted waveform. Since a matched filter is essentially a correlator, an alternative approach is to cross-correlate the signals from two spatially separated receivers assuming that the noise in the two receivers is uncorrelated. The computation rates required to perform the correlations digitally are high for signals which may have bandwidths of the order of 100 MHz. Acousto-optical correlators offer a possible solution since they are able to perform analogue cross-correlation in real time. Three basic architectures are available. Space-integrating correlators using two AO cells require the input to one of the cells to be time reversed [1]; this can be done [2] but the process is inefficient. Time-integrating correlators [1] do not require time reversal of signals but they suffer from unwanted bias terms which tend to saturate the detector array and limit integration time. The proposed joint transform correlator (JTC), which combines space-integrating and time-integrating attributes, suffers from none of these drawbacks. No time reversal is required and unwanted terms are easily separated from the required correlation.

2. Correlator architecture.

A schematic diagram of the system is shown in Figure 1. It is a standard joint transform correlator [3] in which the inputs are provided by the first order diffracted light from two identical AO cells. The lenses and spatial filters required to extract the first order light from the cells have been omitted for clarity. The input function for a joint transform correlator, with two

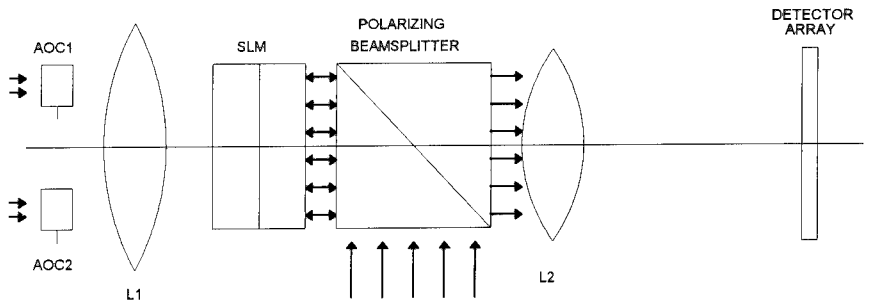


Figure 1. The acousto-optic joint transform correlator - schematic

objects separated by a distance 2a, is of the form

$$o(x) = f(x-a) + g(x+a) \tag{1}$$

It is straightforward to show that if the Fourier transforms of f(x) and g(x) are written as $|F(u)|\exp\{j\varphi(u)\}$ and $|G(u)|\exp\{j\phi(u)\}$ respectively then the response of the square law detector of the spatial light modulator (SLM) will be

$$|O(u)|^2 = |F(u)|^2 + |G(u)|^2 + |F(u)|\cdot|G(u)|\cos\{2ua + \varphi(u) - \phi(u)\} \tag{2}$$

The final term in equation (2) represents a set of sinusoidal fringes of spatial frequency 2a, amplitude modulated by the amplitude of the cross spectral density (CSD) of f(x) and g(x) and phase modulated by the phase of the CSD. In the acousto-optic JTC the input functions are of the form f(x-a-Vt) and g(x+a-Vt), where V is the acoustic velocity in the cells. A rigorous analysis of the acousto-optic system will not be presented due to space limitations. However it is clear that, since the acoustic velocity in the two cells is the same, the two functions will maintain their separation as they propagate across the cells. The spatial frequency of the fringe pattern will therefore remain constant. Consider first the case where the cells are illuminated only for the period that the two functions are fully within the cells. The light intensity at the detector will be as given in equation (2) because the extra phase term introduced by the term Vt will cancel out. This places quite a severe restriction on the length of signal that can be processed; the maximum length of an AO cell window is around 50 μs. Consider now the case where the signals are longer than the cell window and the cells are constantly illuminated. The separation of the signals will still be maintained at 2a plus a term due to any (constant) time shift between the signals so the spatial frequency of the fringes will remain constant. The light intensity at the detector at any instant will be of the form of equation (2) where F(u) and G(u) now represent the transforms of the signals within the cells at that instant. As the signals propagate through the cells a composite intensity pattern will be built up on the SLM by time integrating on the detector.

3. Computer simulations.

Simulations have been carried out using Mathcad to investigate the performance of the system. Figure 2(a) shows a snapshot of the joint transform obtained at a particular instant as a chirp of length 100 μs and bandwidth 1 MHz, centred on 15 MHz, passes through a pair of Bragg cells

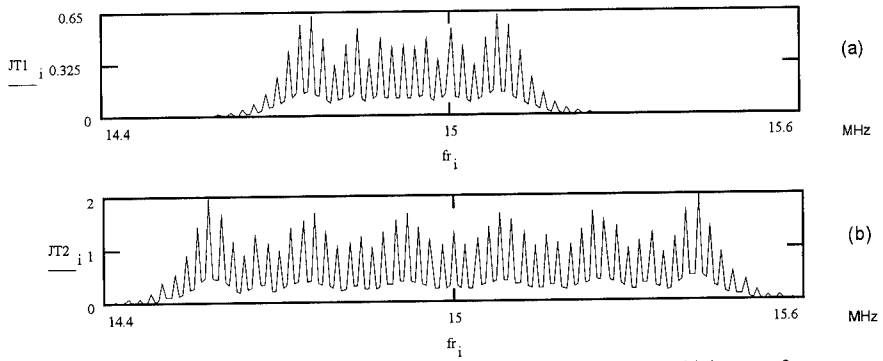


Figure 2. Simulations: (a) a snapshot joint transform and (b) an integrated joint transform.

whose time aperture is 50 μs. Figure 2(b) shows the result of integrating the joint transform over the 150 μs interval between the chirp first entering the cells and finally leaving them. The spatial frequency of the fringes remains constant and would require a spatial light modulator with an aperture of 50 mm and a resolution of 30 line-pairs/mm. Such devices are available. Figure 3 shows the results of simulations carried out using a chirp of length 100 μs and bandwidth 10 MHz, centred on 15 MHz, buried in noise of bandwidth 30 MHz with a signal to noise ratio (SNR) of -10 dB. In Figure 3(a) the relative time shift (τ) between the chirps in the two cells was zero; in Figure 3(b) τ = 500 ns. The plots show the square modulus of the Fourier transform of the joint transform, which is the output that would be expected from the detector array. The output values have been normalised to a peak value of unity. The time shift is clearly visible and the output SNR, calculated as the peak to mean ratio of the output function, is around 14 dB. This is a reasonable result; the time-bandwidth product of the chirp is 1000 giving a signal processing gain of 33 dB. The missing 9 dB can be attributed to signal-noise and particularly noise-noise correlation that would not occur in a matched filter [4].

4. Experimental results

The system has been tested in principle using a single Bragg cell through which two identical chirps were passed. The cells used had a time aperture of 50 μs and a bandwidth of 30 MHz

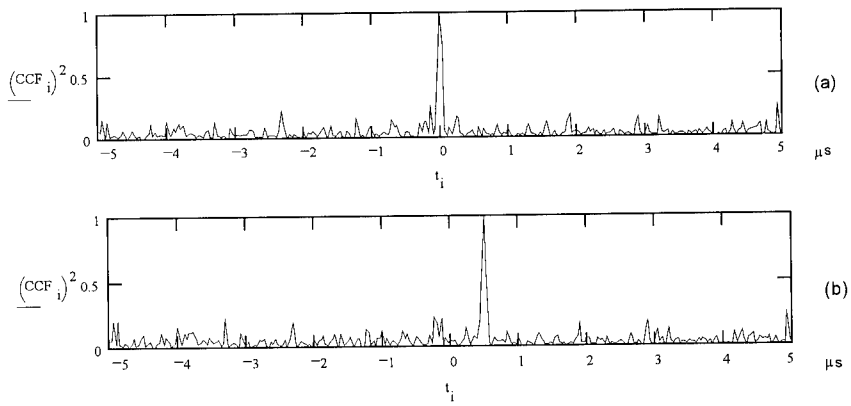


Figure 3. Simulated array output with input SNR = -10 dB for (a) τ = 0, (b) τ = 500 ns

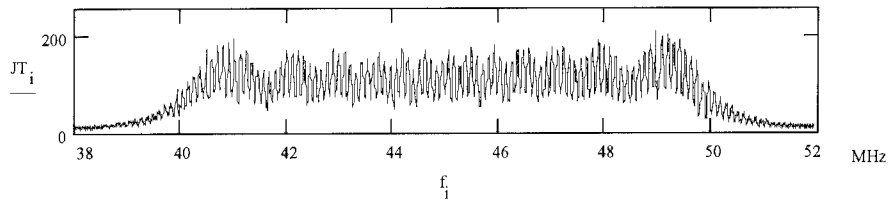


Figure 4. Experimental joint transform

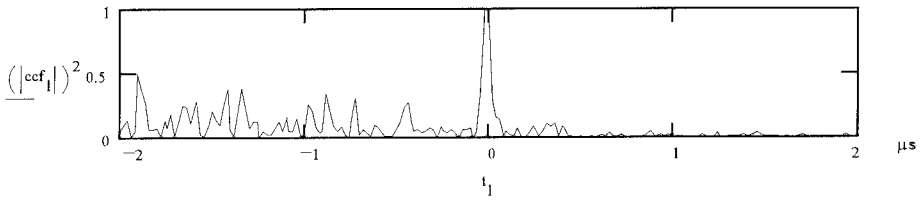


Figure 5. Experimental correlation output

centred on 45 MHz. The chirps used were of length 5 μs and bandwidth 10 MHz centred on 45 MHz. Figure 4 shows the joint transform obtained when the spacing between the two chirps within the cell was 5 μs. Figure 5 shows the correlation output obtained when the above chirps were buried in noise with noise bandwidth 30 MHz and a SNR of around -5 dB. This result was obtained by digitising the joint transform recorded by a 2048 element detector array and plotting the square modulus of its Fourier transform. The apparent asymmetry in the correlation function is due to the reduced response of the array at higher spatial frequencies. The output SNR ratio is around 8 dB, which again is a reasonable figure. The time bandwidth product of the chirp was 50 (signal processing gain of 20 dB) and the input SNR was -5 dB.

5. Conclusions

It has been shown that, in principle, the acousto-optic joint transform correlator described is capable of detecting spread spectrum signals buried in noise by correlating the signals received at two spatially separated receivers. The signal processing gain of the system is worse than a matched filter by up to 9 dB. This is to be expected since a matched filter is, by definition, the system which gives the greatest signal to noise enhancement. Analysis of this system shows that its signal processing gain will be at least 3 dB worse than a matched filter [4], the degradation being due to extra signal-noise and noise-noise correlation terms. The system also has potential for direction finding since the time of arrival at the two receivers, and hence the separation of the signals in the Bragg cells, depends on the angle of arrival of the signal. Bearing resolution will depend upon the separation of the receivers and the width of the compressed correlation pulse.

References

[1] Vanderlugt A 1992 *Optical Signal Processing* (Wiley)
[2] Caputi W J 1971 *IEEE Transactions* **AES-7**, 269-278
[3] Weaver C S and Goodman J W 1966 *Appl. Opt.* **5**, 1248-1249
[4] Houghton A W and Reeve C D 1994 *RNEC Research Report* **94014**

Copyright © HMSO, London, 1994

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 British Crown Copyright

A zeroth order non-heterodyning space integrating acousto-optic correlator

A W Houghton

EE Department, Royal Naval Engineering College, Manadon, Plymouth, PL5 3AQ, UK

C D Reeve

SECEE, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK

Abstract. Both a theoretical analysis and practical results are presented for a non-heterodyning acousto-optic space integrating correlator that uses the zeroth diffraction order to produce a true correlation function containing both amplitude and phase information.

1. Introduction

The structure of the zeroth order non heterodyning space integrating correlator is as shown in figure 1.

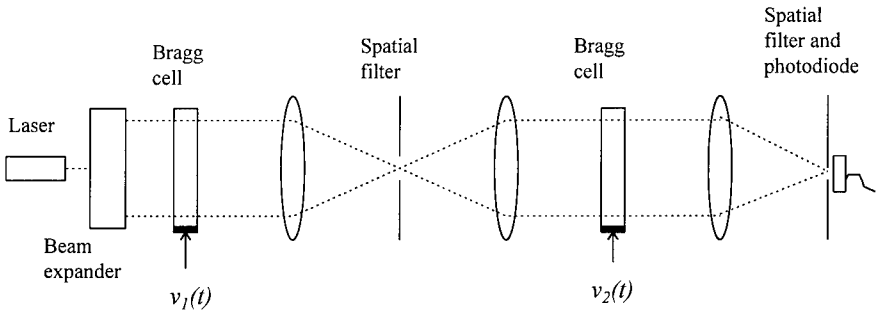


Figure 1. Physical structure of zeroth order space integrating acousto-optic correlator.

The basic architecture is similar to many correlators previously described in the literature [e.g. 1,2,3], but there are two important differences. The spatial filters select all or part of the zeroth diffraction order. Most correlators use the first order, sometimes treating the zeroth order as if it contained no information. In reality, the zeroth order contains as much information as all the other orders put together, but it can be difficult to access due to large optical power incident on a photodetector causing saturation or excessive shot noise. However, it can be shown that the zeroth diffraction order has sub-orders that can be used very effectively. The second important difference lies in the way the signals are introduced into the acousto-optic cells. A common technique is to insert the signals as double sideband

suppressed carrier modulation of a carrier at the centre frequency of the acousto-optic cell. The modulation used here is double sideband large carrier. The difference is crucial.

2. Information in the zeroth order.

Using the co-ordinate frame such that the x axis lies in the direction of propagation of the acoustic wave with its origin at the transducer and the z axis lies in the direction of propagation of the incident optical plane wave, and assuming Raman-Nath diffraction the optical wave at the output of the acousto-optic cell can be described by [3]

$$u(x,t) = \mathrm{Re}\left\{\exp\left[j(\omega_o t - kz)\right]\exp\left[-j\beta v\left(t-\tfrac{x}{V}\right)\right]\right\}\mathrm{rect}\left(\tfrac{x}{W}\right) \quad (1)$$

where ω_o and k are the angular frequency and wave vector of the input optical signal, $v(t)$ is the electrical input signal to the cell, V is the acoustic wave velocity, $\mathrm{rect}(x/W)$ is the windowing function due to the acousto-optic cell of aperture width W, and β is a modulation index which is a function of input signal power and cell parameters. $v(t)$ will usually comprise a carrier, at the centre frequency of the acousto-optic cell, modulated by a signal of interest. We will only consider amplitude modulation here, but two types of amplitude modulation are possible and the difference is crucial. The zeroth order correlator makes use of information in the zeroth diffraction order. This information is not available if we use double sideband suppressed carrier modulation but it is available if we use double sideband large carrier modulation. The electronic input signal to the acousto-optic cell then takes the form

$$v(t) = \left(1 + m\,a(t)\right)\sin\left(\omega_c t\right) \quad (2)$$

where $a(t)$ is the input signal of interest, m the amplitude modulation index, and ω_c the angular frequency of the carrier. The optical wave at the output of the acousto-optic cell becomes

$$u(x,t) = \mathrm{Re}\left\{\exp\left[j(\omega_o t - kz)\right]\exp\left[-j\beta\left(1 + m\,a\left(t-\tfrac{x}{V}\right)\right)\sin\left(\omega_c t - K_c x\right)\right]\right\}\mathrm{rect}\left(\tfrac{x}{W}\right) \quad (3)$$

where K_c is the acoustic wave vector at the carrier frequency. Expanding out the second exponential term in this expression and including second order terms enables us to obtain an expression for the plane wave component, ${}^{o}u(x,t)$, that gives the zeroth diffraction order. Most papers on acousto-optic correlation [e.g. 1,3] include only first order terms in this expansion since this is sufficient to describe the first diffraction order usually used. If we perform this expansion and gather terms up to second order then we obtain

$${}^{o}u(x,t) = \left\{1 - \tfrac{\beta^2}{4}\left[1 + m\,a\left(t-\tfrac{x}{V}\right)\right]^2\right\}\cos\left(\omega_c t - kz\right)\mathrm{rect}\left(\tfrac{x}{W}\right) \quad (4)$$

If we consider a sinusoidal component of $a(t)$ of the form $\cos(\omega_m t)$, then the component of ${}^{o}u(x,t)$ due to such a component is

$$\left\{1 - \tfrac{\beta^2}{4}\left[1 + m\cos\left(\omega_m t - K_m x\right)\right]^2\right\}\cos\left(\omega_c t - kz\right)\mathrm{rect}\left(\tfrac{x}{W}\right) \quad (5)$$

Further expansion of this expression reveals [4] that the zeroth diffraction order can be broken down into diffraction sub-orders 0,n , where n = 0, ±1, ±2. The 0, ±2 sub-orders are very small but the 0, ±1 sub-orders are of significant amplitude and, when focused into the Fourier transform plane of a lens, comprise the Fourier transform of *a(t)*, convolved with the Fourier transform of the window function rect*(x/W)*. Figure 3 shows a prediction of the form of the 0,0 and 0, ±1 diffraction sub-orders, generated using Mathcad, for an input signal *a(t)* which is a linear chirp pulse sweeping from 8 to 12 MHz in 20 μs, with m = 0.3 and β = 0.5 and an acousto-optic cell with a time window of 50 μs.

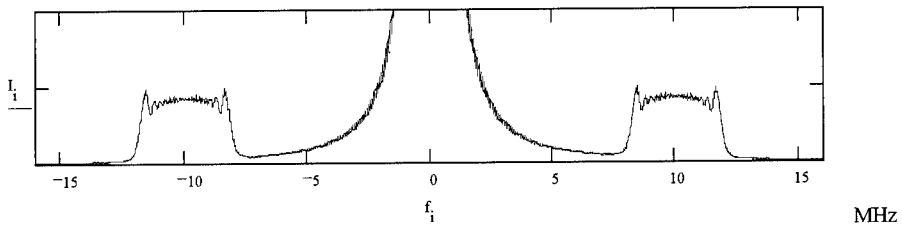


Figure 2. Predicted intensity distribution of the zeroth diffraction order, showing 0,0 and 0,±1 sub-orders

3. Analysis of the zeroth order correlator

In the zeroth order space integrating correlator, as shown in figure 1, the signal input, $v_1(t)$, to the first cell is time reversed so that the signal actually delivered to the transducer is $v_1(T-t)$, where T is the time window of the cell. The travelling wave in the acousto-optic cell can then be described as proportional to $v_1(x-Vt)$. If the second signal is not time reversed but the sense of the x axis is reversed while maintaining its origin at the transducer of the second cell, then the travelling wave in the second cell can be described as proportional to $v_2(x+Vt)$. The zeroth order plane wave component of the output from the first cell is given by equation 4 which can be rewritten as

$$^{o}u_1(x,t) = \left\{1-\tfrac{\beta^2}{4}\left[1+m\,a_1(x-Vt)\right]^2\right\}\cos(kz-\omega_c t)\,\mathrm{rect}\left(\tfrac{x}{W}\right) \qquad (6)$$

The first spatial filter passes the whole of the zeroth diffraction order and so, if the lenses are large enough to treat their apertures as infinite, this component is imaged onto the second cell where reversal of the sense of the x axis enables us to use equation 6 unchanged to describe the optical input to the second cell. The zeroth order plane wave component at the output of the second cell is then

$$^{o}u_2(x,t) = \left\{1-\tfrac{\beta^2}{4}\left[1+m\,a_1(x-Vt)\right]^2\right\}\left\{1-\tfrac{\beta^2}{4}\left[1+m\,a_1(x-Vt)\right]^2\right\}\cos(kz-\omega_c t)\,\mathrm{rect}\left(\tfrac{x}{W}\right) \qquad (7)$$

Expanding out this expression produces a large number of terms [5]. The full expansion is not given here due to the limitations of space. The terms of interest are the first order terms taking the form

$$\left\{\tfrac{\beta^4 m}{8}-\tfrac{\beta^2 m}{2}\right\}\left\{a_1(x-Vt)+a_2(x+Vt)\right\}\cos(kz-\omega_c t)\,\mathrm{rect}\left(\tfrac{x}{W}\right) \qquad (8)$$

Provided that $a_1(t)$ and $a_2(t)$ occupy bandwidths of less than one octave and that their lowest

frequency components are sufficiently high, the final spatial filter can select out one of the 0,±1 diffraction sub-orders containing these first order terms. If a photodiode placed behind this spatial filter is of sufficient area to receive the whole of this diffraction sub-order then the spatially integrated intensity of the light incident on the photodiode will contain slowly varying functions plus the desired cross correlation function in the form

$$\frac{\beta^4}{16}\int_{-\infty}^{+\infty} a_1(x-Vt)\,a_2(x+Vt)\,\mathrm{rect}\left(\tfrac{x}{W}\right)dx \qquad (9)$$

Because of the windowing term the correlation function can only be formed if the functions $a_1(t)$ and $a_2(t)$ are shorter than the time window of the cells so that the overlap between them will lie within this time window at all times. Subject to the restrictions on the signal bandwidths and temporal lengths, this relatively simple correlator produces the complete correlation function, compressed in time by a factor of two, giving both amplitude and phase information. Figure 3 shows the autocorrelation function of a linear chirp pulse sweeping from 4.25 to 5.75 MHz in 15 μs obtained using the correlator illustrated in figure 1.

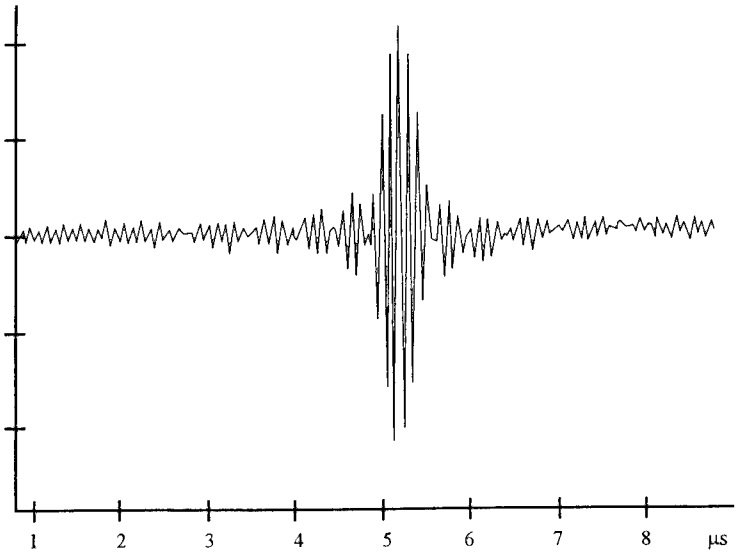


Figure 3. Autocorrelation function of a linear chirp pulse produced by the correlator shown in figure 1.

References

[1] Lee J N and VanderLugt A 1989 *Proc. IEEE* 77 1528-1557
[2] Reeve C D and Wombwell J F 1989 *Proc. IEE F* **136** 185-190
[3] Rhodes W T 1981 *Proc. IEEE* **69** 65-79
[4] Houghton A W and Reeve C D 1992 *Royal Naval Engineering College Research Report* 92033
[5] Houghton A W and Reeve C D 1993 *Royal Naval Engineering College Research Report* 93026

Copyright © HMSO, London, 1994

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Edge enhancement in photorefractive Joint Transform Correlators

Olivier Daniel, Jean-Michel C-Jonathan and Gérald Roosen
Institut d'Optique Théorique et Appliquée - Unité Associée au CNRS n° 14
Centre Scientifique d'Orsay, Batiment 503,
BP. 147, 91403, ORSAY CEDEX -FRANCE

Abstract. We give an analytical expression for the space charge field induced at steady state when a photorefractive crystal, used in the diffusion regime, is illuminated by a non uniform interference pattern. Its consequences on the response of a photorefractive Joint Transform Correlator, including the effect of the reading beam are predicted and experimentally observed. The observed edge enhancement is explained.

1. Introduction

Joint Transform optical Correlators (JTC) are commonly implemented using liquid crystal Spatial Light Modulators for the input of the images to be compared. CCD cameras allow a quadratic detection in the Fourier plane and a linear JTC is thus obtained [1]. The system is quite different when a photorefractive crystal is used.

The purpose of this paper is to demonstrate the specific non linearity of the photorefractive response and show how it affects the output of the correlator.

2. The photorefractive response

The usual Kukhtarev's solution of the band transport model for the photorefractive effect fails when the spatial distribution of illumination varies rapidly on the crystal or is two dimensional. It also fails to describe the case where the modulation depth of the interference pattern locally becomes close to unity. This is what happens in the experimental conditions of a photorefractive Joint Transform Correlator. To predict the steady state photorefractive response in these experimental conditions, we use a method [2], which takes into account the highly non uniform, two dimensional, non periodic illumination.

With a single additional assumption, it provides an analytical expression for the steady state space charge field and an efficient model in the specific experimental conditions of our photorefractive JTC where the use of thin crystals and the low diffraction efficiency do not require beam propagation.

2.1. The photorefractive space charge field

At steady state, with no external field applied to the crystal, the induced space charge field is ruled by 3 equations. Most important, the charge transport equation describes the equilibrium, which is reached between drift and diffusion of the charge carriers. It is completed by the rate equation and the Poisson's equation [3]. With the approximations, $N_D^+(\vec{r}) \ll N_D$ and $n(\vec{r}) \ll N_A$ (low illumination) (where N_D is the volume density of photorefractive sites, N_D^+ that of the ionized ones and N_A equals N_D^+ in the dark), one finds as in [2]:

$$\vec{E}(\vec{r}) - \frac{k_B T}{e} \frac{\vec{\nabla}\left(\vec{\nabla}.\vec{E}(\vec{r})\right)}{\left(k_0^2 \frac{k_B T}{e} + \vec{\nabla}.\vec{E}(\vec{r})\right)} = -\frac{k_B T}{e} \vec{\nabla} \ln\left(I(\vec{r}) + I_d\right) \tag{1}$$

where $\vec{E}(r)$ is the space charge field and $k_0 = \sqrt{\frac{e^2 N_A}{\varepsilon_{DC} k_B T}}$. $I(\vec{r})$ is the illumination of the crystal and I_d is phenomenologically added to describe the effects of dark conductivity and illumination by the reading beam.

Additionally, we assume that the variation of the induced space charge field on a distance equal to the Debye screening length $2\pi/k_0$ is small compared to the maximum possible value $\frac{1}{2} k_0 \frac{k_B T}{e}$ of the space charge field induced by diffusion. This approximation is valid in a wide range of experimental conditions. In the case of a sinusoidal space charge field, it would only fail when the modulation depth gets close to unity and the grating wave vector k close to k_0. We thus assume that $m << \frac{k_0}{k}$ instead of $m << 1$ as in the Kukhtarev's solution. In BSO where $k_0 \approx 1.5\,10^5\, cm^{-1}$, this is verified even at high modulation for the values of k corresponding to our JTC. The derivative in the denominator of (1) may then be neglected and one gets:

$$\vec{E}(\vec{r}) - \frac{1}{k_0^2} \vec{\nabla}\left(\vec{\nabla}.\vec{E}(\vec{r})\right) = -\frac{k_B T}{e} \vec{\nabla} \ln\left(I(\vec{r}) + I_d\right) \tag{2}$$

By the Fourier transforming back and forth, the space charge field is found to be:

$$\vec{E}(\vec{r}) = -\frac{k_B T}{e} \left\{\vec{\nabla} \ln[I(\vec{r}) + I_d]\right\} \otimes \left\{\frac{k_0}{2} \exp(-k_0 |\vec{r}|)\right\} \tag{3}$$

where $\otimes$ denotes the convolution.

The space charge field is obtained as a vector. Expression (3) of the induced space charge field, is valid for any distribution of illumination. It discloses the non linearity of the photorefractive response, including how details smaller than the screening length are smoothed down by the convolution. The logarithm compresses the dynamics, which is very useful in the Fourier plane. The derivative then performs edge enhancement in the plane of the photorefractive crystal. The effect of this non linear processing in the Fourier plane of a JTC is described in the following paragraph.

3. The photorefractive JTC

3.1. Predicted behavior

We now show the consequences of relation (3) in the photorefractive JTC. $O(\vec{r}')$ and $S(\vec{r}')$ are centered at $(-X_0', -Y_0')$ and (X_0', Y_0') in the input plane. These coordinates and the Fourier transformer define the carrier spatial frequency $k = 4\pi\sqrt{X_0'^2 + Y_0'^2}/\lambda f$. The intensity distribution on the BSO crystal is: $I(\vec{r}) = I_0(\vec{r})\left(1 + m(\vec{r})\cos(\vec{k}.\vec{r} + \delta)\right)$

with $I_0(\vec{r}) = \left(I_d + |o(\vec{r})|^2 + |s(\vec{r})|^2\right)$, $m(\vec{r}) = \frac{2|o(\vec{r})|.|s(\vec{r})|}{I_d + |o(\vec{r})|^2 + |s(\vec{r})|^2}$, and $\delta = \arg[o(\vec{r}) s^*(\vec{r})]$.

Replacing $I(\vec{r})$ into (3), one straightforwardly finds an expression involving ratios of trigonometric functions whose Fourier expansions are tabulated [4]. Assuming that the frequencies in $m(\vec{r})$ are small compared to the carrier frequency k, they may be used to

obtain a new expression for the space charge field. These Fourier expansions introduce the reduced modulation $p(\vec{r}) = \frac{1}{m(\vec{r})}\left[\sqrt{1-m^2(\vec{r})}-1\right]$:

$$\vec{E}(\vec{r}) = -\frac{k_B T}{e}\left\{\begin{array}{l}\vec{\nabla}\ln\left(\frac{I_0(\vec{r})}{1+p^2(\vec{r})}\right)+2\vec{k}\sum_{n=1}^{+\infty}p^n(\vec{r})\sin\left[n(\vec{k}.\vec{r}+\delta)\right] \\ -2\frac{1}{p(\vec{r})}\vec{\nabla}p(\vec{r})\sum_{n=1}^{+\infty}p^n(\vec{r})\cos\left[n(\vec{k}.\vec{r}+\delta)\right]\end{array}\right\}\otimes\left[\frac{k_0}{2}\exp(-k_0|\vec{r}|)\right] \quad (4)$$

The first term between braces gives rise to the center peak of the correlation plane. The other ones generate the usual correlation peak and its symmetrical counterpart. Higher harmonics are caused by the non linearity of the photorefractive response. The effect of high modulation is described by $p(\vec{r})$ as previously found in [5]. The last term is the correction brought to the Kukhtarev's model. The first order of diffraction responsible for the correlation is thus:

$$\vec{E}(\vec{r}) = -2\frac{k_B T}{e}\left\{\vec{k}\,p(\vec{r})\sin(\vec{k}.\vec{r}) - \vec{\nabla}p(\vec{r})\cos(\vec{k}.\vec{r})\right\}\otimes\left[\frac{k_0}{2}\exp(-k_0|\vec{r}|)\right] \quad (5)$$

and may be extracted. The assumption that frequencies in $m(\vec{r})$ are small compared to the carrier frequency k, guaranties that overlapping of orders is avoided. The low diffraction efficiency in the diffusion regime (<< 1%) and the use of a thin crystal thus do not allow multiple diffraction that would couple back higher order diffracted terms.

The role of I_d appears here. When the two objects are identical, $m(\vec{r}) = \frac{2|o(\vec{r})|^2}{I_d + 2|o(\vec{r})|^2}$. If I_d is neglected, $m(\vec{r})$ and therefore $p(\vec{r})$ are identical to unity. The cosine in equation (5) vanishes and the sine one is of constant amplitude. The modulated space charge field contains no information about the object. A delta function is observed in the correlation plane at the position (X_0', Y_0') of the object. It is precisely localized but the "correlation peak" bears no information about its shape.

Introducing I_d broadens that peak by bringing back frequencies into $p(\vec{r})$. It reveals the second term, i.e. the space derivative of $p(\vec{r})$.

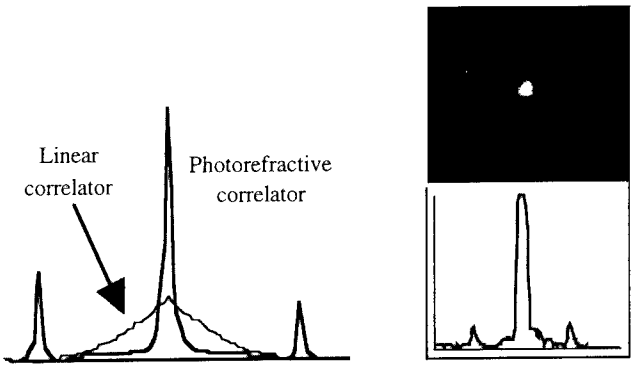
3.2. Experimental set-up and results

The geometry of our experimental set-up is directly derived from that described by Rajbenbach & al [6]. The object O to be recognized and the scene S where it is expected are introduced in the input plane using a TFT nematic liquid crystal spatial light modulator illuminated by the collimated beam of a diode pumped doubled YAG laser. Their Fourier transforms o and s are formed by the same lens. The resulting distribution of intensity is recorded as a modulation of refractive index by a BSO crystal. It is observed using the 670 nm beam of a laser diode incident from the back at an angle adjusted to fulfill the Bragg condition. At that wavelength, the sensitivity of the BSO crystal is usually considered as negligible compared to that at the writing wavelength (532 nm), so that reading and writing may take place simultaneously. The Fourier transform of the amplitude distribution diffracted by the photorefractive grating is observed in real time in the correlation plane P_C, using a CCD camera.

Figure 1 gives experimental evidence for that prediction. One half of the correlation plane is represented in the case of the autocorrelation of a disk. It shows (a) a simulation of a linear response correlator and of a photorefractive JTC from which the central order of diffraction is removed by polarizers, (b) a scan through an experimental "correlation peak" of our photorefractive JTC.

The bright central peak in b) would be the only correlation signal in the absence of the reading beam. In the real conditions, where the sensitivity to the red is only 1/7 of the

sensitivity of the green and as described by the model, this peak widens and is surrounded by a weaker ring.



a) Prediction b) Observation

Figure 1: Autocorrelation of a disk

6. Conclusion

The model, valid at steady state in the diffusion regime, provides original results. We can now give an accurate description of the effects induced by a non uniform illumination of the crystal and predict the response of a photorefractive JTC using a BSO crystal in the Fourier plane. It shows that the photorefractive response is strongly non linear. This non linearity is controlled by the dark conductivity and uniform illumination by the reading beam. Opportunities of controlling it are then open. They are now being studied.

These results are part of a contribution to the NAOPIA II Project, funded by the Commission of the European Community under the ESPRIT Program for Research and Development.

References

[1] B. Javidi and C.J. Kuo, Appl. Opt., Vol. 27, n°4, pp.663-665, (1985).
[2] M. Cronin-Golomb, Optics Communications 89 , pp. 276-282, (1992).
[3] N.V. Kukhtarev, U.B. Markov, S.G. Odulov, M.S. Soskin and V.L. Vinetskii, Ferroelectrics 22 (1979) 949.
[4] I.S. Gradshteyn and I.M. Ryzhik, Table of integrals, series, and products, Corrected and enlarged edition,Academic Press, Inc, p40, 1.447.
[5] M. G. Moharam, T. K. Gaylord, R. Magnusson and L. Young, J. Appl. Phys. 50 (1979) 5642.
[6] H. Rajbenbach, S. Bann, P. Réfrégier, P. Joffre, J.-P. Huignard, H.-S. Buchkremer, A.S. Jensen, E. Rasmussen, K.H. Brenner, G. Lohman, Appl. Opt., Vol. 31, n°26, pp.5666-5674, (1992).

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical correlators for the analysis of images with the computer processing of the results

A.V. Korolkov and Y.S. Mosyakin

Scientific Research Institute "KVANT", 15, 4-th Likhachevsky Lane, 125438 Moscow, Russia

Abstract. This article concerns the experimental construction and application of a precision optical correlator for analysis of large format images with the computer processing of output fields for fast and precise determination of correlation maxima coordinates.

1. Introduction

For the structural analysis of highly informative images, for example, astro- and aerial photography, we effectively use optical correlators [1]. To provide a very high speed of processing of information the correlator should include an automatic image analyser, which has a link to the computer. But for detecting and analysis of correlation images, a special unit based on a 2-D optical bistable element and two linear photosensors can be used.

In this paper we describe the experimental realisation of an optical correlator with the computer processing of output images and with the analyser on the basis of an optical bistable element.

2. Automated correlator with analyser based on a linear scanning photosensor unit

For the solution of the task of precise determination of coordinates of correlation maxima for the analysis of large format images (180*180 mm) we have developed a precision coherent optical correlator according to the scheme of Vander Lugt, based on an optical Fourier processor with aperture 500 mm and resolution to 100 lin/mm, holographic registrator type HRS 1001 with the working area 40*40 mm and resolution to 1500 lin/mm, scanning photoreceiving units Eikonix 1000 (Kodak Co.) with the resolution of elements 4096*4096, connected with the IBM PC AT-computer.

The output image of the correlation function formed by the Fourier processor, is read by the scanning unit, digitised to 256 levels and put into the computer for the digital processing. Realised so, algorithms permit one to automatically determine the location of correlation peaks, by increasing the operator threshold, independently of their dimensions and form and with regard for unevenness of the background, on which images are optically formed. Algorithms are optimised according to fast activity and permit analysis of one correlation field with information capacity 4096*4096*8 bits in a period of five minutes and writing results into the file and the possibility of depicting the map of location of detected correlation maxima and fragments of read image in pseudocolours on the monitor of the computer.

The Fourier optics for our correlator was produced by Vavilov State Optics Institute (S.-Petersburg). The experimental method of conducting investigations was directed to receive maximum sensitivity in locating correlation maxima in images, which results in optical processing of large format highly informative images. The usage of the described automated correlator permits one to significantly shorten the time of analysis of images (less than 200 sec by Compaq 386) and ensure high precision and reliability of received results. The programs were written on Borland C++.

3. Correlator with analyser on the basis of OBE

The speed of the correlator is limited by the fame frequency of the communication link to the computer. But in some tasks it is enough to establish the presence or absence of one or more correlation maxima in the output image, received on a free working correlator. In this case the output analyser may be built on the basis of a two-dimensional optical bistable element (OBE) in the mode of a threshold element, behind which we place a collecting lens and point (one-element) photoreceiver. By exceeding threshold in any point of the two-dimensional OBE there happens its commutation into transparent state, which increases photoelectric current in the chain of photoreceiving element. A similar analyser, equipped with light dividing prism, two crossed cylinder lenses and linear photodetectors, connected with computer, may be used for the determination of correlating maximum coordinates in the output image of the optical processor. Thereby because of threshold processing with the help of the OBE the amount of information at the output is substantially reduced. The coordinates of the maximum pixel exceeding the tuning threshold of the OBE are easily determined by the cell coordinates in the linear detectors on which light falls. The result of processing is given to the computer, to which all units are connected; this computer ensures their synchronisation and makes decisions on recognition. Arising from simultaneous presence of some correlating maxima, uncertainty in their coordinates is easily removed by additional testing on the computer. The described analyser can ensure high speed, because it interrogates only 2N photoreceiving cells, but not N*N as in the case of using a two-dimensional matrix.

We used 2-D thin film optical bistable devices for the experimental analysis of coherent images, formed by an optical correlator at very high speed. The unit which makes this analysis consists of one 2-D optical bistable device, two multi-element linear photosensors and one cubic beamsplitter [2]. The 2-D optical bistable devices had 38 mm diameter and were made in the Division of Optical Problem of Informatic (Academy of Sciences of Belarus). This OBE has less than 4% spatial nonlinearity of its threshold. The switching time is 200 ms and resolution is 30 lin/mm. This unit provided a good signal/noise ratio and we have today sensitivity to discover correlation maxima only to 4 times less.

References

[1] Edited by H. Stark, 1982, *Applications of Optical Fourier Transforms* (Academic Press).

[2] Korolkov, A.V., 1992, Fast correlator on the basis of optical bistable elements, *Technical Digest ICO Topical Meeting on Optical Computing '92*, paper 29E36.

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical Coordinate Transformations: A New Architecture

Ronglong Tian, Masahide Itoh, and Toyohiko Yatagai

Institute of Applied Physics, University of Tsukuba, 1-1-1 Ten'nohdai, Tsukuba, Ibaraki 305, Japan

Abstract. A new architecture for optical coordinate transformations with two phase-only filters is proposed, which is described by the Wigner distribution function. A precise theory and experimental results are presented.

1. Introduction

Many optical coordinate transformation method have been proposed.[1-8] Bryngdahl first proposed the use of computer-generated holograms in coordinate transform.[1] Holograms for scale and rotation invariant transforms have been designed.[2,3,6] Bartelt and Case proposed to use a multifacet hologram to the point-to-point mapping.[5] Most of the methods studied are based on the coherent optical Fourier transformation and paraxial approximation in the coordinate transformation. Davidson et al. proposed to determine the grating vector of a Fourier transform hologram.[7] In these methods, however, it is difficult to retain the resolution of the input images.

We propose here an alternative non-Fourier transform approach, in which two phase-only filters are used and the input plane and the output plane of the system are conjugate each other.

2. Phase-only filter design

Consider an optical system, in which two phase-only filters are used and the input plane and the output plane are conjugate each other, as shown in Fig.1. Two phase-only filters $\exp[\mathrm{j}\Phi_1(\boldsymbol{r})]$ and $\exp[\mathrm{j}\Phi_2(\boldsymbol{r})]$ are positioned at a distance of z from the input plane and at the Fourier transform plane of a lens L_1, respectively. The Fourier transform of the first filter is obtained in the second filter plane. The final output is the Fourier transform of the transmitted light from the second filter. In terms of paraxial approximation, the optical system is generally described by the following double Wigner distribution function:[8]

$$K(\boldsymbol{r}_o,\boldsymbol{a}_o,\boldsymbol{r}_i,\boldsymbol{a}_i)=\frac{1}{\lambda^2}\iint h\left(\boldsymbol{r}_o+\frac{\boldsymbol{r}_o'}{2},\boldsymbol{r}_i+\frac{\boldsymbol{r}_i'}{2}\right)h^*\left(\boldsymbol{r}_o+\frac{\boldsymbol{r}_o'}{2},\boldsymbol{r}_i+\frac{\boldsymbol{r}_i'}{2}\right)\exp\left[-\mathrm{j}k\left(\boldsymbol{a}_o\boldsymbol{r}_o'-\boldsymbol{a}_i\boldsymbol{r}_i'\right)\right]\mathrm{d}\boldsymbol{r}_o'\,\mathrm{d}\boldsymbol{r}_i', \quad (1)$$

where $h(\boldsymbol{r}_i,\boldsymbol{r}_o)$ denotes the point spread function of the system. Equation (1) gives a

relationship between heights $\boldsymbol{r}$ and angles $\boldsymbol{a}$ in the input and the output planes. In the particular case of Fig. 1, the optical system is represented by

$$K(\boldsymbol{r}_o,\boldsymbol{a}_o,\boldsymbol{r}_i,\boldsymbol{a}_i)=\iiint\iiint\iiint\delta(\boldsymbol{r}_1-\boldsymbol{r}_i-z\boldsymbol{a}_i)\delta(\boldsymbol{a}_1-\boldsymbol{a}_i)\delta(\boldsymbol{r}_1'-\boldsymbol{r}_1)\delta\left[\boldsymbol{a}_1'-\boldsymbol{a}_1-\frac{1}{k}\frac{\mathrm{d}\Phi_1(\boldsymbol{r}_1)}{\mathrm{d}\boldsymbol{r}_1}\right]\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{a}_1$$
$$\times\delta(\boldsymbol{r}_2-f\boldsymbol{a}_1')\delta\left(\boldsymbol{a}_2+\frac{\boldsymbol{r}_1'}{f}\right)\mathrm{d}\boldsymbol{r}_1'\mathrm{d}\boldsymbol{a}_1'\delta(\boldsymbol{r}_2'-\boldsymbol{r}_2)\delta\left[\boldsymbol{a}_2'-\boldsymbol{a}_2-\frac{1}{k}\frac{\mathrm{d}\Phi_2(\boldsymbol{r}_2)}{\mathrm{d}\boldsymbol{r}_2}\right]\mathrm{d}\boldsymbol{r}_2\mathrm{d}\boldsymbol{a}_2$$
$$\times\delta(\boldsymbol{r}_o-f\boldsymbol{a}_2')\delta\left(\boldsymbol{a}_o+\frac{\boldsymbol{r}_2'}{f}\right)\mathrm{d}\boldsymbol{r}_2'\mathrm{d}\boldsymbol{a}_2'. \tag{2}$$

Suppose the phase of the second filter is written by the following form:

$$\exp[\Phi_2(\boldsymbol{r}_2)]=\exp\left[\left(-\frac{\mathrm{j}k}{f}\right)\int S\left(\frac{\boldsymbol{r}_2}{f}\right)\mathrm{d}\boldsymbol{r}_2\right]. \tag{3}$$

We have the double Wigner distribution function written by:

$$K(\boldsymbol{r}_2',\boldsymbol{a}_2',\boldsymbol{r}_2,\boldsymbol{a}_2)=\delta(\boldsymbol{r}_2'-\boldsymbol{r}_2)\delta\left[\boldsymbol{a}_2'-\boldsymbol{a}_2-\frac{1}{k}\frac{\mathrm{d}\Phi_2(\boldsymbol{r}_2)}{\mathrm{d}\boldsymbol{r}_2}\right]=\delta(\boldsymbol{r}_2'-\boldsymbol{r}_2)\delta\left[\boldsymbol{a}_2'-\boldsymbol{a}_2-\frac{1}{f}S\left(\frac{\boldsymbol{r}_2}{f}\right)\right]. \tag{4}$$

We generally consider the following coordinate transformation:

$$\boldsymbol{r}_o=g(\boldsymbol{r}_i). \tag{5}$$

Then we have final phase distributions of filters as follows:

$$\Phi_1(x,y)=k\left\{\int R_x(x_i,y_i)\mathrm{d}x_i+\int\left[R_y(x_i,y_i)-\frac{\partial}{\partial y_i}\int R_x(x_i,y_i)\mathrm{d}x_i\right]\mathrm{d}y_i\right\}, \tag{6}$$

$$\Phi_2(x,y)=-\frac{k}{f}\left\{\int g_x\left[x_i\left(\frac{x}{f},\frac{y}{f}\right),y_i\left(\frac{x}{f},\frac{y}{f}\right)-x_i\left(\frac{x}{f},\frac{y}{f}\right)\right]\mathrm{d}x+\int\left\{g_y\left[x_i\left(\frac{x}{f},\frac{y}{f}\right),y_i\left(\frac{x}{f},\frac{y}{f}\right)-x_i\left(\frac{x}{f},\frac{y}{f}\right)\right]\right.\right.$$
$$\left.\left.-\frac{\partial}{\partial y}\int g_x\left[x_i\left(\frac{x}{f},\frac{y}{f}\right),y_i\left(\frac{x}{f},\frac{y}{f}\right)-x_i\left(\frac{x}{f},\frac{y}{f}\right)\right]\mathrm{d}x\right\}\mathrm{d}y\right\}. \tag{7}$$

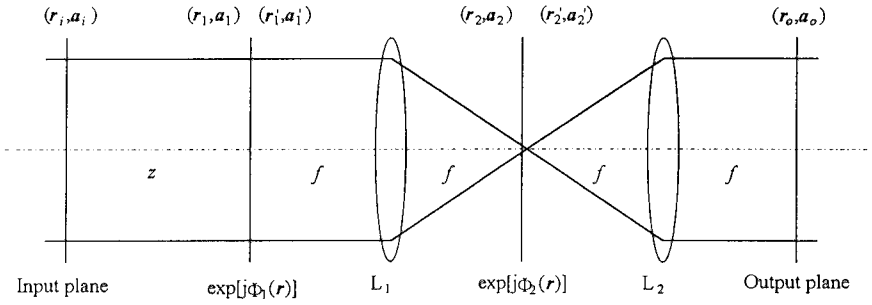


Fig. 1 Optical system for coordinate transform.

As a typical example, we consider the following coordinate transformation:

$$g_x(x_i, y_i) = \eta \ln(x_i) \tag{8}$$

$$g_y(x_i, y_i) = \xi y_i \tag{9}$$

The phase distributions of the filters are given by

$$\Phi_1(x,y) = -\frac{k}{z}\left[\frac{x^2}{2} - \frac{x^3}{6\eta} + qx + \left(1 - \frac{1}{\xi}\right)y^2\right] + c \tag{10}$$

$$\begin{aligned}\Phi_2(x,y) = -\frac{k}{f}\Bigg\{\eta\Bigg[&\left(x + \frac{qf}{z}\right)\ln\left(\sqrt{1 + \frac{2zx}{\eta f} + \frac{2q}{\eta}} - 1\right) \\ &+ x\ln(|\eta|) - \frac{\eta f}{4z}\left(\sqrt{1 + \frac{2zx}{\eta f} + \frac{2q}{\eta}} + 1\right)^2 - x \\ &+ \frac{\eta f}{3z}\left(\sqrt{1 + \frac{2zx}{\eta f} + \frac{2q}{\eta}}\right)^3\Bigg] + \frac{qf}{z}\ln\left(\sqrt{1 + \frac{2zx}{\eta f} + \frac{2q}{\eta}} - 1\right) - \frac{\xi z}{2f}y^2\Bigg\} + c'\end{aligned} \tag{11}$$

3. Experiment

We have designed the phase-only filters of Eqs. (10) and (11). The fabrication of phase-only filters is technically difficult, and so we have made binary filters with the computer-generated hologram technique, as shown in Fig.2. Figure 3 shows an input pattern example (a) of the grating whose pitches are exponentially changed. Because of the logarithmic coordinate transformation system, we have the output (b) of an equi-spaced grating image.

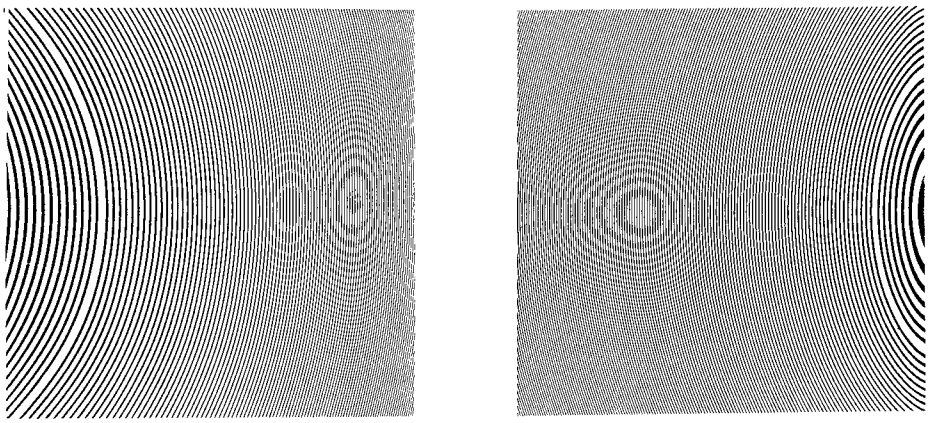
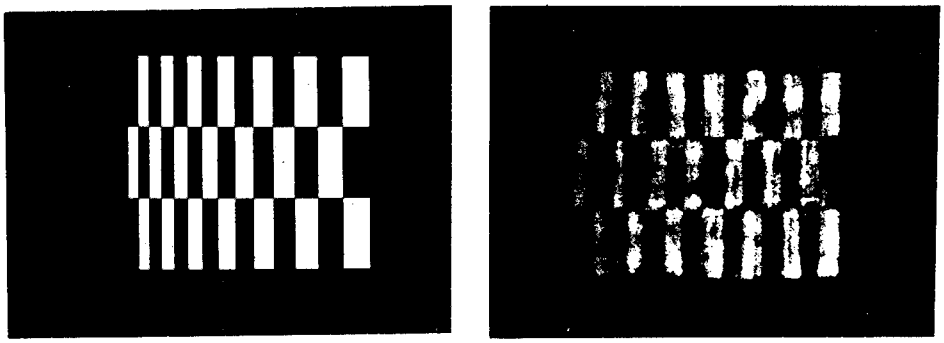


Fig. 2 Binary computer-generated holograms for Eqs. (10) and (11).



(a) (b)

Fig. 3 Logarithmic coordinate transformation. (a) Input image, (b) output image.

4. Conclusion

We have proposed a new architecture of the optical coordinate transformation using a pair of phase-only filters and derived the analytical solution of the filters. Optical experiments with computer-generated filters verifies the analytical solution for logarithmic coordinate transformation. Features of the architecture are discussed in comparison with Bryngdahl Fourier transform method.

References

[1] O. Bryngdahl, "Geometrical transformation in optics," J. Opt. Soc. Amer., 64, 1092-1099 (1974).
[2] D. Casacent and C. Szczutkowski, "Optical Mellin transforms using computer-generated holograms," Opt. Commun., 19, 217-222 (1976).
[3] Y. Saito, S. Komatsu and H. Ohzu, "Scale and rotation invariant real time optical correlator using computer-generated hologram, Opt. Commun., 47, 8-11 (1983).
[4] J. Cederquist and A. M. Tai, "Computer-generated holograms for geometrical transformations," Appl. Opt., 23, 3099-3104 (1984).
[5] H. Bartelt and S. K. Case, "Coordinate transformations via multifacet holographical optical elements," Opt. Engr., 22, 497-500 (1983).
[6] H. Szu, Holographic coordinate transformations and optical computing," Proc. SPIE 634, 480-484 (1986).
[7] N. Davidson, A. A. Friesem and E. Hasman, "Optical coordinate transformations," Appl. Opt., 31, 1067-1073 (1992).
[8] M. J. Bastiaans, " The Wigner distribution function and Hamilton's characteristics of a geometrical-optical system," Opt. Commun., 30, 321-326 (1979).

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Breaking Symmetry in the Binary Phase Only Matched Filter

T. D. Wilkinson and R. J. Mears
Cambridge University Engineering Dept,
Trumpington St, Cambridge CB2 1PZ
Tel + 44 223 332600
email tdw@eng.cam.ac.uk

Abstract

A new means of breaking symmetry in binary phase only matched filters is reported. A randomly pixellated diffractive element is combined with a dynamic binary filter displayed on a ferroelectric liquid crystal spatial light modulator to produce a pseudo four-level phase matched filter. The four phase levels produced are sufficient to break inversion symmetry without penalising the peak to noise ratio of the correlator. A simulated annealing algorithm is used to generate dynamic section of the filter.

The Binary Phase Only Matched Filter (BPOMF) has been well documented since Spatial Light Modulators (SLMs) capable of binary phase modulation were developed[1]. A less publicised feature of the BPOMF is the rotational symmetry due to the binary phase levels of 0 and π. The Hermitian impulse response of the BPOMF means that the filter cannot discriminate between the reference and the same reference rotated by 180°. For most pattern recognition tasks, the inherent rotational invariance is an advantage, but for some applications such as distinguishing between the characters d and p, the property can lead to incorrect decisions. In an application such as road sign recognition, triangular signs have very different implications for each orientation. An asymmetric BPOMF is thus desirable. Symmetry can be broken by increasing the number of phase levels from two to four. Previous work has shown that two binary SLMs in a Mach Zehnder interferometer geometry[2] can be used to display the filter, but this technique is expensive and complicates the optical design. Recently, a pseudo four level phase system has been used to break the symmetry in computer generated holograms displayed on a single binary phase SLM[3]. In this letter we show that such an approach can be used with great advantage in the BPOMF. With filters displayed on a single SLM we demonstrate that it is possible to break the BPOMF symmetry without penalising the peak to noise ratio.

The correlator is a classical $4f$ Vander Lugt BPOMF. Immediately following the filter SLM, (in binary phase mode) is a Pixellated Diffractive Element (PDE) which is used to create a pseudo four level system. The PDE has the same pixel pitch as the filter SLM and is aligned pixel to pixel with the SLM. Each pixel on the random phase array is set to either 0 or $\pi/2$

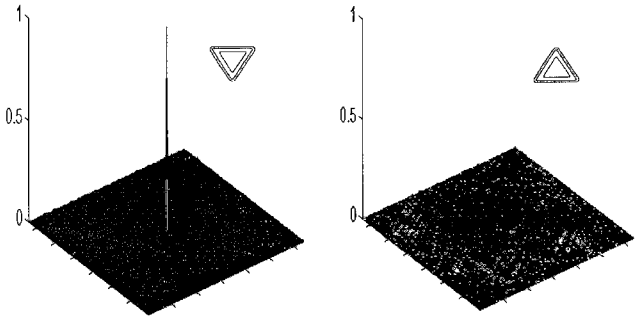


Figure 1: Simulation of the pseudo four phase level system applied to a triangular road sign

and is fixed. Each pixel on the SLM can be set to 0 or π and is dynamic. The combination of the array and the SLM means that phase levels from either $[0 \quad \pi]$ or $[\pi/2 \quad 3\pi/2]$ can be selected depending on the pattern on the SLM. Ideally, the phase array would be physically etched onto the glass surface of the SLM, but for initial demonstration, we have constructed a photoresist pattern on a $\lambda/10$ optical flat to generate the $\pi/2$ phase step.

The BPOMF was generated by a simulated annealing algorithm[4, 5] which was adapted to include the random phase array in the calculation. An ideal correlation (delta function) was used as the target for the algorithm and a plane of zeros used as the target for the rejection criteria. An 87x56 pixel triangular warning sign sampled from the U.K. Highway Code was used as the input to be recognised and the rotated sign used for rejection. Figure 1 shows a simulation of the pseudo four phase filter with a $\pi/2$ phase step and a triangular road sign input in both orientations. Although the PDE constrains the design of the matched filter, there is sufficient phase freedom in the system for the filter to optimise very efficiently. The suppression of the symmetric order for an ideal phase step of $\pi/2$ was 16dB and there was no change in the peak to noise ratio, 27.5dB, when compared to a symmetric BPOMF generated by the same algorithm. For a phase step of more than $\pi/2$, the suppression was reduced, but is still sufficient for good discrimination. The PDE constructed had a delay of 2.09rad, the symmetric suppression was reduced to 5.1dB but the peak to noise ratio remained at 27.5dB.

A chrome on glass mask, with 128x128 pixels (220μm pitch matching the filter SLM), was made containing a totally binary random pattern. The mask was then used to to transfer the pattern onto a layer of photoresist spun onto an optical flat. The flat was then developed leaving the raised pattern in photoresist. The exact thickness of the photoresist determines the retardation of the step. A step size of 247nm was required for $\pi/2$ at $\lambda = 633$nm. The thinnest photoresist layer achieved was 318nm using Shipley S1400-17 spun at 4000rpm for 40 seconds, which creates a phase delay of 2.09rad at 633nm which is sufficient to demonstrate the asymmetric properties.

The experimental layout is a modified $4f$ correlator shown in Figure 2. The SLMs used were

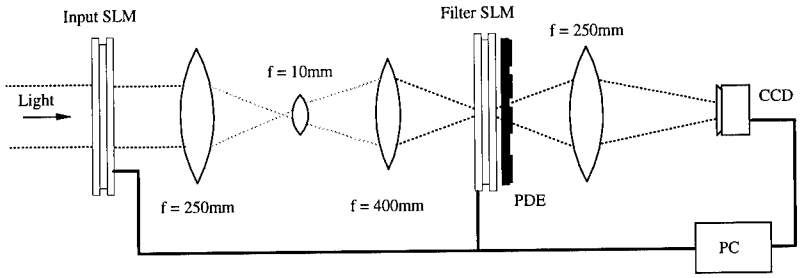


Figure 2: Experimental layout of the modified $4f$ BPOMF

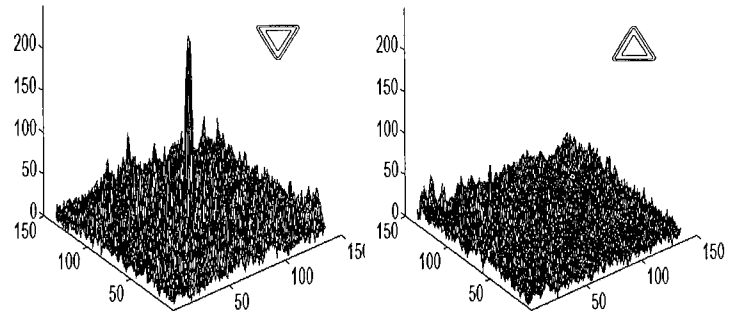


Figure 3: Experimental BPOMF correlation plane for the correct and mirrored triangular roadsign

128x128 pixel ferroelectric liquid crystal devices from THORN EMI CRL. The pixel pitch was 220μm for both the input and filter SLM. The three lens system in Figure 2 was used to compress the effective length of the transform lens, whilst keeping the wavefront distortion to below one wavelength, The PDE was mounted as close to the SLM analyser as possible to minimise the effects of diffraction through the pixels. Figure 3 shows the results of the experiment. The symmetry suppression was limited by the noise in the output plane to 4.8dB which is sufficient for good discrimination between the two rotated images. The noise was mainly due to nonuniformities in the photoresist layer of the PDE and diffraction across the gap between the analyser and the PDE.

We have presented a new means of suppressing the symmetric correlation peaks in the BPOMF without requiring an extra SLM or penalising the peak to noise ratio in the correlation plane. Performance would be further improved by constructing an SLM with the random $\pi/2$ phase pattern accurately etched onto the glass surface of the SLM cell. Furthermore, as

the BPOMF is generated by simulated annealing, it is possible to train it for applications which require other invariant properties whilst suppressing the symmetry.

The authors would like to thank Dr J. Brocklehurst and M. Birch of CRL for providing the SLMs and Dr D.C. O'Brien for many useful discussions. The work was partially supported by ESPRIT project number 7050, HICOPOS.

References

[1] D. Psaltis, E.G. Paek, and S.S. Venkatesh. Optical image correlation with a binary spatial light modulator. *Optical Eng*, 23(6):698–704, November 1984.

[2] F.M. Dickey and B.D. Hansche. Quad-phase correlation filters for pattern recognition. *Applied Optics*, 28(9):1611–1613, May 1989.

[3] T.D. Wilkinson, D.C. O'Brien, and R.J. Mears. Dynamic asymmetric binary holograms using a ferroelectric liquid crystal spatial light modulator. *Optics Comms*, 109:222–226, July 1994.

[4] M.P. Dames, R.J. Dowling, P. McKee, and D. Wood. Efficient optical elements to generate intensity weighted spot arrays: design and fabrication. *Applied Optics*, 30(19):2685–2691, July 1991.

[5] T.D. Wilkinson, D.C. O'Brien, and R.J. Mears. Scale invariant binary phase-only matched filter using a ferroelectric liquid crystal spatial light modulator. *Applied Optics*, 33(20):4452–4453, July 1994.

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Universal computation and optical correlators

M. Montes-Usategui, I. Juvells, J. Campos(*) and J. R. de F. Moneo

Universitat de Barcelona, Departament de Física Aplicada i Electrònica, Laboratori d'Òptica. 08028 Barcelona, Spain.
(*) Universitat Autònoma de Barcelona. Departament de Física, Laboratori d'Òptica. 08193 Bellaterra, Barcelona, Spain

Abstract. We show the possibility of implementing a cellular automaton called Life by means of an optical correlator. Since this cellular automaton is an universal computer there is no inherent limitation in the information processing capabilities of correlators.

1. Introduction.

Optical correlators are probably the most popular and well-studied devices used in optical pattern recognition. They are applied to a wide variety of problems including those requiring multiple invariances. However, there is no guarantee that a particular problem can be solved by means of an optical correlator, and therefore it is a legitimate question whether they can be considered as general purpose classifier devices.

The main difficulty of single filter correlations is the limited number of images that can be stored in the filter thus limiting the complexity of the problems that can be solved. These limitations have been analyzed in Ref.1. A possible way to overcome this drawback is the use of multichannel correlators which enable us to obtain arbitrarily complex processings [1,2].

In this communication we show that it is possible, at least in theory, to use iterative techniques as an alternative to multichannel schemes to solve the inherent limitations of single channel correlators. We propose an iterative process involving a single-channel correlation processed by a simple nonlinear function. This process enables us to optically implement a cellular automaton, called Life, with powerful capabilities to process information.

2. Cellular automata.

Cellular automata are usually described as a regular infinite array of cells that can take one of several allowed states. Automata evolve in discrete time steps (generations) governed by deterministic rules. The state of a given cell in the generation k+1 is a function of its previous state (that at generation k) as well as of the states of its neighbors (also at generation k). The updating of the cells is performed synchronously, that is all at the same time. Some of these cellular machines exhibit a behavior complex enough to support universal computation [3]. They are called class-IV automata and one of the most studied is the game of Life [4]. The possibility of being implemented by correlation leads to the conclusion that a correlator is as powerful, in its information processing capabilities, as the most powerful machines known (or for those who accept the Church-Turing thesis as the most powerful machines that can exist), namely universal computers.

3. Optical implementation of Life.

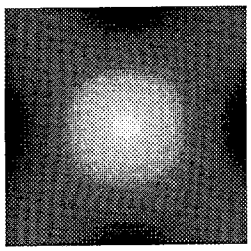
The cells of Life can be in one of the two possible states, dead or alive. The transitions between these two states are controlled by the states of its eight neighboring cells following these basic rules:

- A cell that is alive at time k continues living at time k+1 if and only if it has 2 or 3 living neighbors.
- A cell that is dead at time k becomes alive at time k+1 if it has exactly three living neighbors.

From these rules it is evident that the state of any cell is determined by the total sum (assuming one for live cells and zero for the dead ones) of the values of its neighbors, and does not depend on their individual values. This property is called totalistic and enables a simple optical implementation. By correlating an input Life pattern with the filter represented in Fig.1 and processing the output intensity distribution by the function f(x) = rect (x/10 -1) we obtain the next generation.

1	1	1
1	0.5	1
1	1	1

a)



b)

Figure 1. a) Filter in object space. b) Filter in Fourier space.

As an example of the computing capabilities of the automaton we show the construction of an AND gate. The correlation was simulated by using the Fast Fourier Transform (FFT) algorithm over an array of 128x128 cells. The input pattern is that shown in Fig. 2.a where the two inputs are labelled by the letters A and B. The signals are carried by patterns formed by five cells called gliders (labelled with ones) which move one position diagonally each four generations. The presence or absence of a glider represents the bit 1 or 0 respectively. The structure at the bottom left corner is a glider gun, a periodic pattern that emits a glider after 30 generations. With the proper timing and alignment the collision (indicated by V) of two gliders produces a vanishing reaction which annihilates both patterns, as shown in Fig. 2.b. The logic operation is based on these collisions.

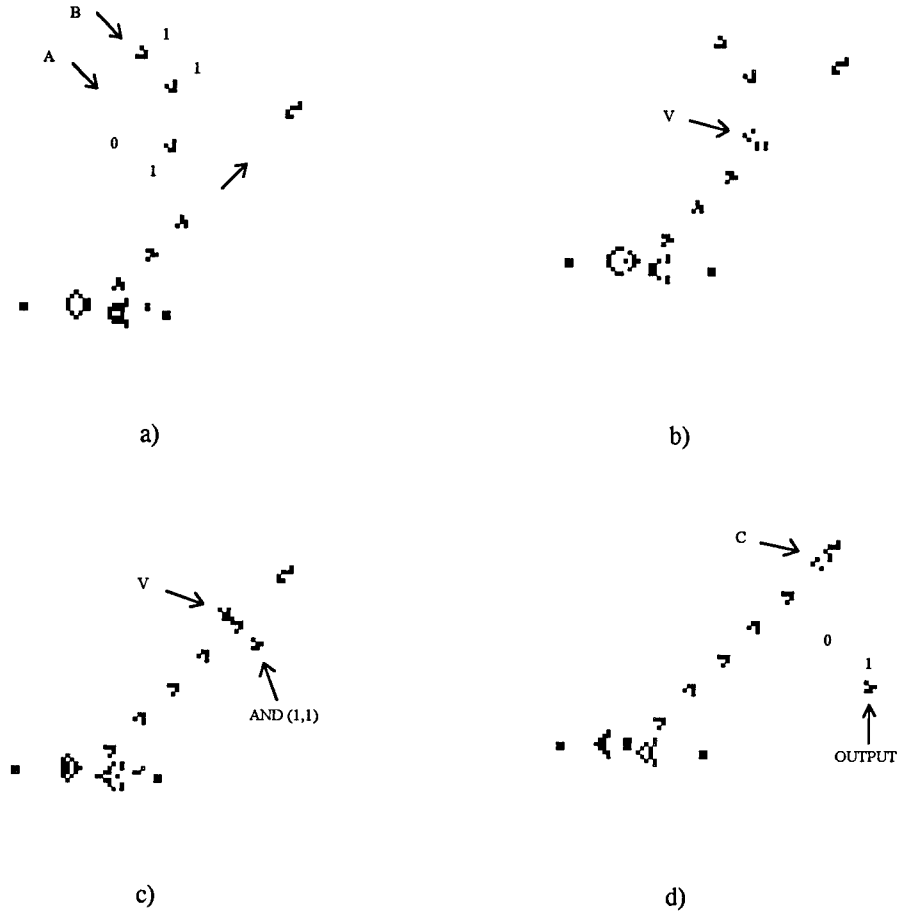


Figure 2. Life AND gate. a) The desired operations are AND (A=1,B=1) and AND (A=0,B=1). b) The first glider of channel A opens a hole in the stream. c) The first glider of channel B pass while the solitary second glider of channel B is destroyed. d) The eater eliminates the residual gliders

When there is a glider in both channels (A and B) the collision of one of them opens a hole in the stream produced by the glider gun, which is profited by the second to pass through (Fig. 2.c). When there is only one glider it is eliminated by the stream (Fig. 2.c) and therefore by detecting the presence or absence of gliders at the position indicated in Fig. 2.d at predetermined times we obtain the output of an AND gate. The rest of the stream produced by the glider gun is eliminated by the eater (the collision C shown in Fig. 2.d) to avoid interferences with other structures.

As can be observed the repeated application of the simple rules that govern the evolution of Life patterns produces a very rich behaviour; the automaton permits the construction of travelling structures (the gliders) to send signals, these gliders can be eliminated by colision with other gliders or with stationary patterns (the eater), etc. Based mainly on these patterns we can perform a complete imitation of an universal computer, that is we can build different logic gates and interconnect them to build memory elements for example [4]. As a consequence, the iterative correlation process by which we have implemented Life, can be used to solve arbitrarily complex mathematical problems such as those involved in pattern recognition. However, it is not necessary to codify pattern recognition problems as Life structures to solve them. We have used the game only to show that the proposed architecture is free of the a priori restrictions of single-filter correlations. This is the usual way to show the equivalence of computational power between different architectures of computing devices. Therefore it is necessary to find the practical form to take profit of this framework by designing specific processes that permit to solve a particular pattern recognition problem.

Acknowledgements.

This work was supported in part by CICYT projects ROB91-0554 and TAP93-0667-C03-01.

4. References.

1. J. Campos, M. Montes-Usategui, I. Juvells and M.J. Yzuel. "On the necessity of multiple filters in optical pattern recognition," in Euro-American Workshop on Optical Pattern Recognition, B. Javidi and Ph. Réfrégier, Eds. SPIE Monograph **PM12**, 137-166 (1994).
2. F. M. Dickey, M. L. Yee, K. T. Stalker. "Correlation, functional analysis, and optical pattern recognition," in Euro-American Workshop on Optical Pattern Recognition, B. Javidi and Ph. Réfrégier, Eds. SPIE Monograph **PM12**, 3-25 (1994).
3. S. Wolfram. "Universality and Complexity in Cellular Automata". Physica D **10**, 1-35 (1984).
4. E. R. Berlekamp, J. H. Conway, and R. K. Guy. *Winning Ways for Your Mathematical Plays*. Academic Press. New York (1982).

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A filter design to improve the recognition of defocused images

J. Sallent, M. Montes-Usategui, J. Campos (*), S. Bosch

Universitat de Barcelona, Departament de Física Aplicada i Electrònica, Laboratori d'Òptica. 08028 Barcelona, Spain.
(*) Universitat Autònoma de Barcelona. Departament de Física, Laboratori d'Òptica. 08193 Bellaterra, Barcelona, Spain

Abstract. The use of Minimum Average Correlation Energy (MACE) and Minimum Noise and Correlation Energy (MINACE) filters for the recognition of defocused images is studied. We propose to adapt the filter simultaneously to the original and to the blurred images to improve the recognition process. In this way the value of the correlation peak is stable for a wide range of blurring in the input image.

The input image in an optical correlator may be degraded by different sources of blurring. A common degradation is defocus. Usually the defocusing will correspond to an image acquisition process which is essentially incoherent, but, for optical processing purposes, the blurred images will always be the input of a coherently illuminated correlator.

In the recognition process of real images we can distinguish two different cases. In a first situation the blurring of the whole scene containing different images is fixed, so simply the ratio of the corresponding correlation peaks may be a good criterion for a final thresholding. A second case corresponds to different blurrings for the different images. Here, to apply a threshold it is necessary that the peak value does not change with defocusing.

We will consider only the second case, therefore we will be interested in the maximum stability of the peaks and in keeping the difference between the correlations of the target and the rejected images. We need differences high enough to allow a thresholding in the correlation plane to classify the objects.

The theoretical performances of several correlation filters like the Classical Matched Filter, Phase Only Filter and Inverse Filter have been previously studied [1]. Moreover, their implementation in convergent correlators [2] or in a joint transform architecture [3] has been recently studied.

One possibility to improve the performance of the correlation method is the use of Minimum Correlation designs such as the MACE and MINACE filters[4]. With these designs it is possible to control the correlation at the origin with different input objects (the training set). This allows to include the blurred images as part of the training set and therefore to obtain defocusing invariant filters. We will show that this is a useful technique to improve the performance of a MACE or MINACE filter for a wide range of defocused scenes.

Figure 1. Test objects used in the experiments.

Figure 1 shows the original (sharp) object to be detected together with the two blurred versions. We show also the two objects to be rejected. The butterflies are 54x64 pixels in size, zero-padded to a total of 128x128 pixels. The Point Spread Functions of the blurring are approximated by circles of a diameter d=1 (non defocused), 4.16 and 20.8 pixels.

Figures 2 and 3 represent the values of the correlation intensity peaks (obtained with the different filters) as a function of the defocusing. In both figures, the lines with * correspond to a filter designed to give a central correlation value 1 with the non-defocused target and 0 with the two non-targets. Although the discrimination capability is good, we see that the correlation value decreses rapidly as the defocus increases.

Since the MACE and MINACE filter design methods allow multiple target correlation conditions, we can use a defocused object as one of the targets and we can impose the central correlation value with the non defocused and with the defocused target to be the same. The lines with 'o' correspond to filters designed to give the same central peak for the non-defocused and maximum defocused targets (diameter = 20.8 pixels). Note the good stability for intermediate degradations. Lines with 'x' are for non-defocused and slight defocusing (diameter = 4.16 pixels). In this case the performances also decay as blurring increases. In all cases the MACE filters are less resistant than MINACE ones.

This lack of resistance to defocusing in the MACE design when no blurred images are included in the training set, is due to the high values of the filter components corresponding to frequencies of low energy in the training images. Since it can be assumed with wide generality that these low energy components are those of high frequency, the MACE filter will be very sensitive to even small changes in the high frequency components of the input

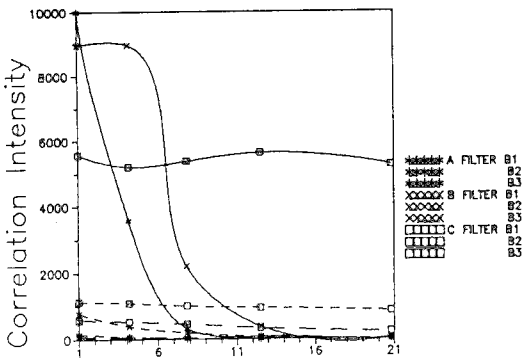


Figure 2.- MACE Filter

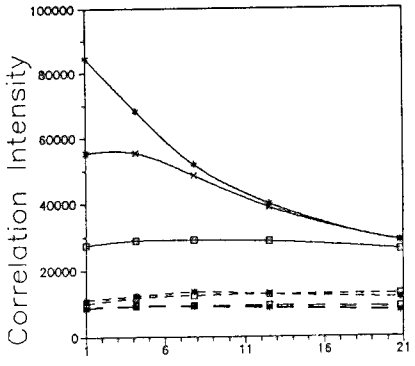
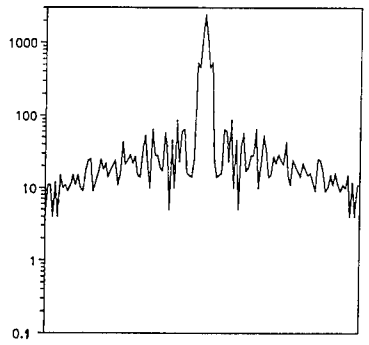


Figure 3.- MINACE Filter

Correlation intensity peaks versus defocusing obtained with Mace and Minace filters. Target is butterfly B1. Cross-correlation with the other butterflies are also given.



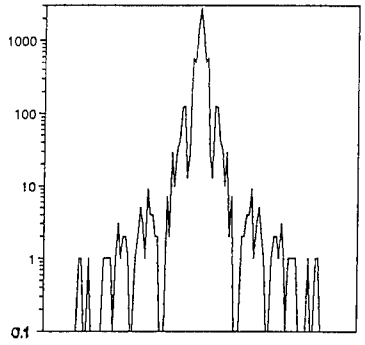


Figure 4 Figure 5

Cross-section of B1 power spectrum, 4) sharp image and 5) defocused image (d=20.8).

images. On the other hand, as we can see in figures 4 and 5, the defocusing of an image mainly affects to its high frequency components (the effect can be approximated to a low-pass filter) being the error largely amplified and thus producing significative variations in the correlation peaks.

The MINACE filter, although specifically designed for noise resistance, shows better performances due to the limitation imposed to the maximum value of a filter component corresponding to a low-energy frequency. As expected, the autocorrelation values of MINACE filters are higher than the same autocorrelations corresponding to MACE filters for the same training set.

Figures 6 shows 3D views of different correlations in the case of MACE filters. In figures a), and b) the filter is designed using the non-blurred objects. In c) and d) the blurred target (PSF of 20.8 pixels) is also included. These correlations are obtained with non blurred (figures 6a, and 6c) and with blurred butterflies (figures 6b, and 6d) B1, B2 and B3 of 20.8 pixels in each picture. Notice the stability of the peaks when the blurred target is included in the training set

In figure 7 we see the peak to correlation energy (PCE) of the different MACE and MINACE filters. For MACE filters, the PCE is quite constant within the range considered for their design, decreasing rapidly outside. For MINACE filters, the behaviour is more stable with respect to defocusing.

Figure 8 shows the Signal-to-Noise Ratio for all the filters. In all cases the SNR is stable for the defocusing range used in the design. Outside this range the SNR degrades exponentially. As expected because of the design, the SNR with MINACE filters is higher than that of MACE filters.

In conclusion, we can use MACE or MINACE filters designed with blurred images in the training set in order to provide resistance to defocusing in our identification system even in the most restrictive case of different defocus for each captured image. We have found that it is enough to include the original and the maximum degraded images to obtain good

stability in all the range. If we use MINACE filters we add noise tolerance, higher stability and we obtain more light efficiency.

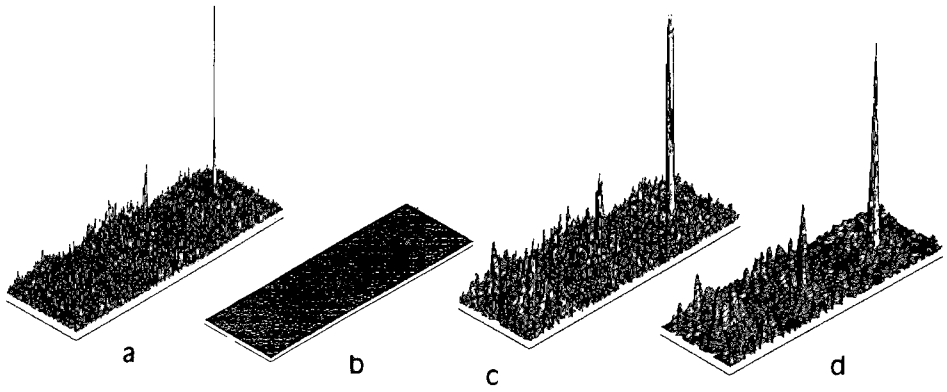


Figure 6.- Correlation intensities with the three butterflies. A MACE filter is designed to give the following responses: a), b) B1=1, B2=0, B3=0; c), d) B1=1, B2=0, B3=0, B1=1(defocused with d=20.8). a), c) sharp scene, b), d) scene defocused with d=20.8.

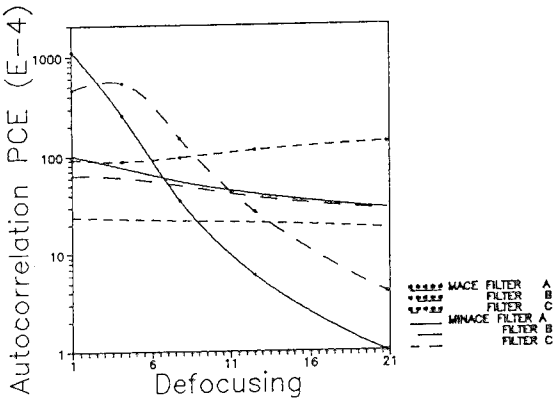


Figure 7

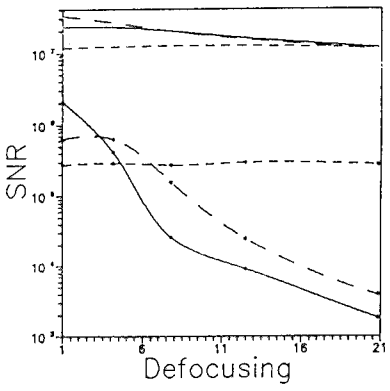


Figure 8

Lines with (*) correspond to MACE filters. Continuous lines are for the case when only the sharp target is used. Long dashed lines use also the d=4.16 blurred target, and short dashed lines correspond to using the sharp and the most defocused target (d=20.8). The same convention applies to MINACE lines.

Acknowledgements This work was partly supported by the CICYT project ROB91-0554.

References

[1] J. Campos, S. Bosch and J. Sallent, 1991 Opt. Comm., **82**, 370-379.
[2] J. Campos, S. Bosch, J. Sallent and A. Berzal, 1994 J. Optics, **25**, 25-31.
[3] B.K. Karch, M.A. Karin and D.L. Flannery, 1993 Opt. Eng., **32**, 2709-2719.
[4] G. Ravichandran and D. Casasent, 1992 Appl. Opt., **31**, 1823-1833, (1992).

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical Haar wavelet transform using a time-division-multiplexed shadow casting system

Ju-Seog Jang† and Shin Il Jeong

Department of Telematics Engineering, National Fisheries University of Pusan, 599-1 Daeyun-Dong, Nam-Gu, Pusan 608-737, Korea
† Current Address: Computation and Neural Systems, MS 116-81, California Institute of Technology, Pasadena, CA 91125, USA

Abstract. An optical system for the Haar wavelet transform of binary images based on a time-division-multiplexing technique is described. We show that a simple parallel image differentiator can handle the bipolar wavelet function efficiently in the system.

1. Introduction

Recently, various optical architectures of the wavelet transform (WT) have been investigated to get fast transform results by taking advantage of parallelism of optics [1]. Following Yang *et al.* [2], we performed experiments on the optical Haar WT using a shadow casting system. To handle bipolar nature of the wavelet functions, Yang *et al.* used a polarization encoding technique in their shadow casting system with a coherent light source. Through a properly rotated analyzer, the phase difference between the lights from the positive and negative parts of the wavelet function becomes π, by which the amplitude subtraction is obtained in their experiment. However, we could not obtain good amplitude subtraction results because the sum of two coherent beams with a phase difference π still interfered constructively at higher-order positions in the output focal-plane.

To represent bipolar images in incoherent systems, area encoding, wavelength multiplexing, and time division multiplexing (TDM) techniques have been commonly used [3]-[5]. In these schemes, a separate subtraction process is required after the two outputs of the different signs are obtained. Here we adopt the TDM technique and show that a simple parallel image differentiator can be used efficiently for the subtraction.

2. The Haar WT based on TDM technique

The Haar WT of a 2-D image $s(x,y)$ is mathematically described by

$$W(p,q,a,b) = \frac{1}{\sqrt{ab}} \int\int s(x,y) h\left(\frac{x-p}{a}, \frac{y-q}{b}\right) dxdy \quad (1)$$

where $h(x,y)$ is the mother Haar wavelet function, a and b are dilation factors, and p and q are shifting factors. Let us denote the positive and negative parts of the wavelet function $h^+(x,y)$ and $h^-(x,y)$, respectively. That is

$$\begin{aligned} h^+(x,y) &= \text{rect}(x-0.5, y-0.5) + \text{rect}(x+0.5, y+0.5) \\ h^-(x,y) &= -\text{rect}(x+0.5, y-0.5) - \text{rect}(x-0.5, y+0.5). \end{aligned} \quad (2)$$

Thus $h = h^+ + h^-$.

The Haar WT based on TDM technique means that h^+ and $|h^-|$ are presented alternately during $2jt_o \le t < (2j+1)t_o$ and $(2j+1)t_o \le t < (2j+2)t_o$, respectively. $j = 0, 1, 2, \cdots$, and t_o is the flipping time interval of the positive and negative wavelet patterns. Thus the TDM version of Eq. (1) can be written as

$$\begin{aligned} W_{TDM}(p,q,a,b,t) &= W^+(p,q,a,b) \sum_j [u(t-2jt_o) - u(t-(2j+1)t_o)] \\ &+ W^-(p,q,a,b) \sum_j [u(t-(2j+1)t_o) - u(t-(2j+2)t_o)] \end{aligned} \quad (3)$$

where $W^{\pm}(p,q,a,b) = (1/\sqrt{ab}) \int\int s(x,y)|h^{\pm}((x-p)/a, (y-q)/b)|dxdy$ and $u(t)$ is the Heaviside step function. Note that W_{TDM} is always unipolar if $s(x,y)$ is a unipolar function.

To obtain the Haar WT result, $W(p,q,a,b)$, the subtraction $W^+ - W^-$ should be executed. We can get this subtraction by differentiating W_{TDM} of Eq. (3) in time domain, *i.e.*,

$$\frac{\partial W_{TDM}}{\partial t} = W^+(p,q,a,b)\delta(t) + \sum_{k=1}^{\infty} (-1)^k [W^+(p,q,a,b) - W^-(p,q,a,b)] \delta(t - kt_o) \quad (4)$$

where $\delta(t)$ is the Dirac delta function. Practically, $k = 1, 2$ is enough. Eq. (4) implies we can get the magnitude of $W^+ - W^-$ by detecting the pulse energy at every output position (p,q) after $t > 0$. Note that there is no ideal differentiator and thus ideal delta function. A practical differentiator is the high pass filter. In this case, the pulse amplitude is proportional to $|W^+ - W^-|$ in Eq. (4). The sign of subtraction result can be also obtained if the pulse polarity is detected.

3. Optical Haar WT system based on TDM technique

Our system that implements Eq. (3) and (4) is shown in Figure 1. The first part that calculates W_{TDM} is a shadow casting system whose structure is similar to that in [2]. The pattern for h^+ and $|h^-|$, which are flipping alternately, are represented by the light transmittance of a spatial light modulator (SLM1). A sequence of discrete shifting factors (p_m, q_n), where m and n are integers, is represented by the positions of the point sources. The dilation factors a and b can be modified by moving the position of the flipping

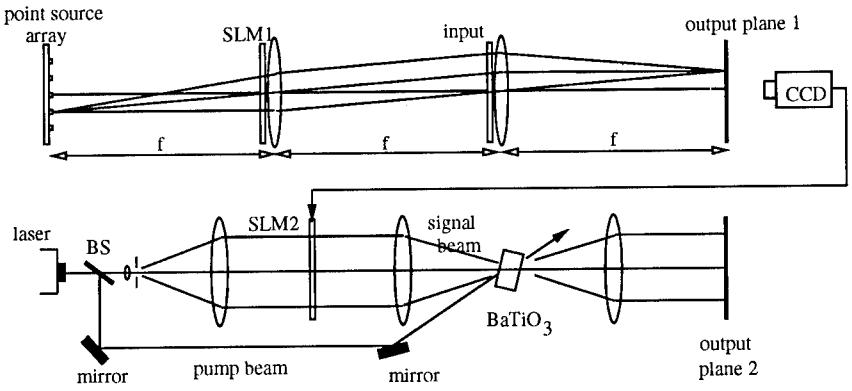


Figure 1: The optical Haar WT system based on TDM technique.

wavelet pattern (SLM1) along the optical axis. This movement enlarges or shrinks the shadow size of the wavelet pattern according to the moving directions. Thus for a given size of the wavelet pattern, a and b, $W_{TDM}(p_m, q_n, t)$ is detected and fed to SLM2 in the image differentiator or novelty filter (NF) [6],[7].

Either two-beam coupling or four-wave mixing NF can be used for the image differentiator. The two-beam coupling NF shown in the lower part of Figure 1 is simple and works with not only phase-modulated but also intensity-modulated images. The light intensity of the pump beam is much higher than that of the image-bearing signal beam. In the photorefractive crystal ($BaTiO_3$), holographic gratings are generated by these two beams. The crystal is oriented so that the signal beam is strongly deamplified. Physically, this takes place because the diffracted beam by the hologram from the pump beam into the the direction of the signal beam interferes destructively with the directly passing signal beam. In other words, the diffracted beam and the signal beam are nearly equal field intensity and π out of phase with each other in the steady state [6],[7]. Thus they tend to cancel out.

Suppose some portion of the input signal beam, for example, $W^+(p_m, q_n)$ changes suddenly into $W^-(p_m, q_n)$ at $t = (2j+1)t_o$. The hologram in the crystal cannot follow this change instantly because of the photorefractive time constant. Therefore, only the amount of beam intensity that is not cancelled by the diffracted beam from the pump beam, $|W^+(p_m, q_n) - W^-(p_m, q_n)|$, appears in the image plane of the output signal beam at that instant. Since there is no further variation in the signal beam until next t_o duration, new gratings are formed after a time lapse of about a photorefractive time constant. The destructive interference is regained. An example of this circumstance is depicted in Figure 2(a). The time interval, t_o, should be greater than the photorefractive time constant. The way of our system works is depicted in Figure 2(b) when the input unipolar function $s(x, y)$ is a rectangular image.

4. Discussion and Conclusion

Both the two-beam coupling and four-wave mixing NF's detect only the magnitude of $|W(p_m, q_n)|$ like other optical wavelet transform systems do [1], because they produce only unipolar light pulses whenever W_{TDM} changes suddenly. To detect both the sign

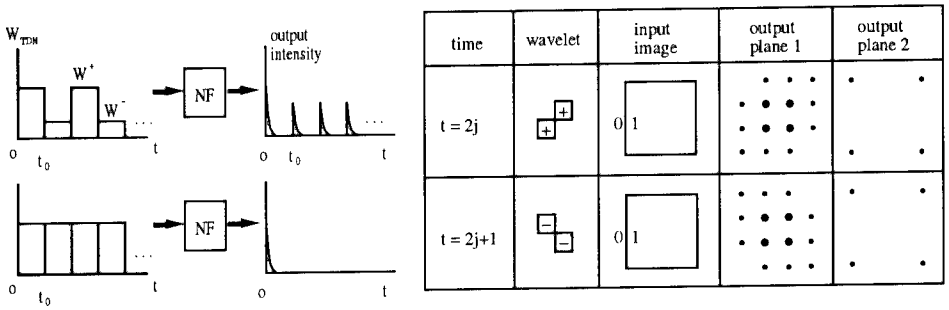


Figure 2: An example of the way our system works (gedanken experiment). (a) The action of two-beam coupling differentiator at an output position (p_m, q_n); (b) The Haar wavelet transform result for an input of unipolar rectangular image.

and magnitude of W, a beam fanning differentiator may be used, which produces positive light pulses only when W_{TDM} increases [6],[8]. Therefore, if the pulses are generated at $t = kt_o$ of even (odd) k, the sign of W is positive (negative). In this case the photorefractive crystal should be placed near the image plane of W_{TDM} so that the beam fanning can occur independently for every beam spot of (p_m, q_n). In general, the speed of beam fanning depends on the input beam intensity. Thus one disadvantage of this scheme will be that t_o should be long enough for the beam fanning to take place sufficiently at possible weak beam spots.

The system in Figure 1 is not a fully parallel processor, because of the CCD detector and SLM2, although their speed is a real-time video rate. They can be replaced with a parallel photodetector-SLM module, for example, a liquid crystal light valve.

The time needed to obtain $|W(p_m, q_n)|$ with this optical system is t_o, roughly speaking, the photorefractive time constant that is determined by the strong pump beam intensity. It is noteworthy that this time constant is independent of the size of the input image $s(x, y)$ and the number of point sources in the array (p_m, q_n).

In conclusion, we explained an optical shadow casting system for the Haar wavelet transform based on TDM technique. We showed that a simple image differentiator using a photorefractive crystal can handle the bipolar wavelet function. Thus a serial subtraction process related with the bipolar nature of the wavelet function can be eliminated efficiently.

References

[1] See, for example, 1992 *Opt. Eng.* **31** 1823-1898

[2] Yang X, Szu H, Sheng Y, and Caulfield H J 1992 *Opt. Eng.* **31** 1846-1851

[3] Trabka E A and Roetling P G 1964 *J. Opt. Soc. Am.* **54** 1242-1248

[4] Psaltis D, Casasent D, and Carlotto M 1979 *Opt. Lett.* **4** 348-350

[5] Yang X, Lu T, and Yu F T S 1990 *Appl. Opt.* **29** 5223-5225

[6] Cronin-Golomb M, Biernacki A, Lin C, and Kong H 1987 *Opt. Lett.* **12** 1929-1931

[7] Anderson D Z and Feinberg J 1989 *IEEE J. Quantum Electron.* **25** 635-647

[8] Cronin-Golomb M and Yariv A 1985 *J. Appl. Phys.* **57** 4906-4910

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Fractional Fourier transforms, imaging systems and correlators

L. M. Bernardo, F. J. Marinho and O. D. D. Soares
CETO, Centro de Ciências e Tecnologias Ópticas, Faculdade de Ciências, Universidade do Porto, 4000 Porto, Portugal.

Abstract. Fractional Fourier transforms of integer, fractional and complex degrees can be implemented by lenses. Optical units that realize such basic transformations are shown to be useful in the implementation of more complex optical systems such as imaging systems and correlators.

1. Introduction

Recently, Lohmann [1] has shown that fractional Fourier transforms (FRFTs), defined on the basis of Wigner functions, can be optically implemented by lenses, after generalizing the usual single-lens and two-lenses geometries used for the implementation of the conventional Fourier transform. The Lohmann's analytical definition is based on collimated illumination. A more general definition for divergent and convergent illumination can be written for both single-lens and two-lenses geometry configurations [2].

2. Implementation of FRFTs with lenses

Figure 1(a), indicates the geometry for the single-lens configuration of FRFTs. The variables z_1 and f affect the value of the degree p of FRFT, $F^{(p)}$, through the expressions:

$$sin\phi = \frac{1}{f}\sqrt{z_1(2f - z_1)} \; ; \qquad cos\phi = \frac{f - z_1}{f} \tag{1}$$

where $\phi = p\pi/2$. The angle ϕ and therefore p can be complex numbers. In this case, the sinusoidal functions can be expressed in terms of sinusoidal and hyperbolic functions of real arguments. The scale factor of the transform, f_ρ, depends on z_1, f and the point source distance ρ to the object plane either for divergent ($\rho>0$) or convergent ($\rho<0$) illumination, through the expression:

$$f_\rho \equiv \frac{\rho f}{\rho + z_1 - f} \tag{2}$$

The distance z_2 of the plane, where $F^{(p)}$ is formed, to the lens depends on ρ, f and z_1, according to the expression:

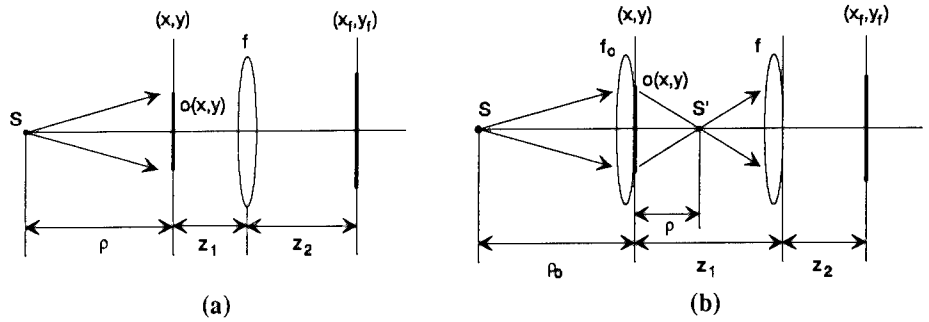


Figure 1. General geometry for (a) the single-lens configuration, (b) the two-lens configuration, for $\rho<0$; o(x,y) is the object transparency, (x,y) the object plane and (x_f,y_f) the FRFT plane.

$$z_2 = \frac{\rho + f}{\rho + z_1 - f} z_1 \tag{3}$$

If the illumination is not collimated, the scale and the position of $F^{(p)}$ change in general, as compared to the case of collimated illumination. However, as it is well known, this is not the case for $F^{(1)}$, the perfect conventional Fourier transform, where only the position changes.

Figure 1(b) presents the general geometry of the two-lenses configuration for $\rho<0$. This case can be reduced to the case of a single-lens geometry with convergent illumination. The first lens of focal distance f_o determines the value of ρ; the second lens of focal distance f realizes the FRFT. The equations (1), (2) and (3) are also valid for the two-lenses configuration, where $\rho=-\rho_o f_o/(\rho_o+f_o)$. For collimated illumination, ρ_o is infinite, $\rho=-f_o$, and then

$$z_2 = \frac{f - f_o}{z_1 - (f + f_o)} z_1 \tag{4}$$

In the two-lenses configuration, if both lenses have the same focal length f, then $z_2=0$. The focal length f_o and ρ_o determines the value of ρ, and thus the scale factor f_ρ and z_2.

The most general expression of $F^{(p)}$ [2] for both configurations is:

$$F^{(p)}_{x_f,y_f}\{o(x,y)\} = \iint o(x,y)\, e^{i\frac{k\cos\phi}{2f_\rho \sin^2\phi}\left[(x^2+x_f^2)+(y^2+y_f^2)\right]} e^{-i\frac{k}{f_\rho \sin^2\phi}(xx_f+yy_f)}\, dx\, dy \tag{5}$$

where ϕ and f_ρ are defined by equation (1) and equation (2), respectively.

The inverse FRFT, $F^{(-p)}$, can be obtained from equation (5) after substituting there i by $-i$. The implementation of inverse FRFT is realized also by the two configurations, substituting

the positive lenses by negative lenses and interchanging the object and the transformation planes [3]. The object and the transform distributions are in those cases virtual. According to equation (1), the FRFTs may have a complex degree for certain values of z_1.

For a collimated illumination, the geometries in both configurations take the form represented in figure 2(a), where $z_2=z_1$, and in figure 2(b), where $z_2=0$; the general expression for $F^{(p)}$ (equation (5)) takes different forms in the two configurations [3].

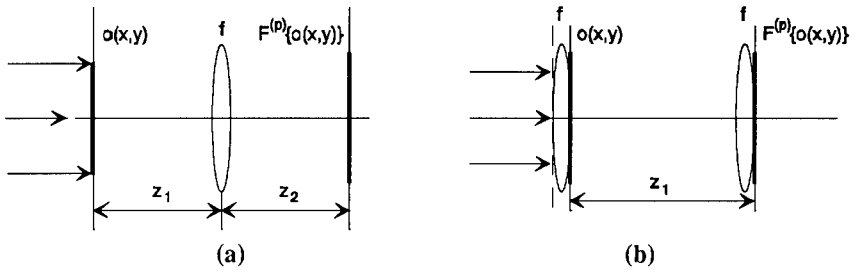


Figure 2: (a) single-lens configuration, where $z_2=z_1$, (b) two-lenses configuration, where $z_2=0$.

An important parameter, associated to any $F^{(p)}$, is the so called standard focal length, f_1 [1]. The expressions of f_1, for the single-lens configuration and for the two-lenses configuration, are respectively:

$$f_1 = f\sin\phi \;;\quad f_1 = f\frac{1-\cos\phi}{\sin\phi} \tag{6}$$

Transforms $F^{(p)}$ with the same parameter $|f_1|$ belong to the same family, meaning that, when they are implemented and cascaded in tandem, they satisfy the properties of the FRFTs expressed by the relationship:

$$F^{(\alpha)}\left\{F^{(\beta)}\{f(x,y)\}\right\} = F^{(\beta)}\left\{F^{(\alpha)}\{f(x,y)\}\right\} = F^{(\alpha+\beta)}\{f(x,y)\} \tag{7}$$

The meaning of the standard focal length f_1 becomes clear. It represents the scale factor of $F^{(1)}$, resulting from the application, in tandem, of other FRFTs belonging to the same family.

3. Generalized afocal imaging systems and correlators

Perfect imaging optical systems or correlators, without additional phase curvatures at the output or $F^{(1)}$ planes, can be realized by cascading basic units (see figure 2) which implement fractional Fourier transforms of the same family. Particularly, $F^{(\alpha+\beta)}$ may

represent a conventional Fourier transform of an object distribution if $\alpha+\beta=2n+1$ (n=0,1,2...), an inverted image if $\alpha+\beta=4n+2$, or an erect image if $\alpha+\beta=4n$.

Figure 3 shows two examples of imaging systems based on fractional FRFTs of the same family. One complex unit of the single-lens configuration, followed by another complex unit of two-lens configuration are used in figure 3(a); three real units of the two-lenses configuration are used in figure 3(b). In the last case, it happens that the perfect conventional Fourier transform $F^{(1)}$ is formed in plane π. This can be understood by symmetry considerations. Therefore, this imaging system is also a correlator with a scale factor $f_1=f/2$.

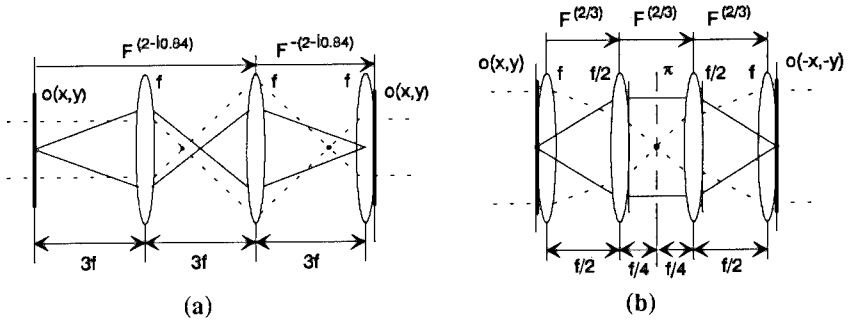


Figure 3: (a) perfect imaging system with two units of complex degrees, (b) perfect imaging system and perfect correlator; o(x,y) at the input plane is the object transmission function.

4. Conclusion

It has been shown that it is possible to implement imaging systems and correlators based on FRFT units constituting generalized afocal systems. This concept may simplify the design of complex systems.

Acknowledgements

This work was partially supported by the European Union Human Capital and Mobility (HC&M) project, "Technologies for hybrid-integrated optoelectronic systems" contract no. CHRX-CT93-0215 (DG DSCS).

References

[1] Lohman A 1993 *J. Opt. Soc. Am. A* **10** 2181-2186.

[2] Bernardo L and Soares O 1994 *J. Opt. Soc. Am. A* (to be published).

[3] Bernardo L and Soares O 1994 *Opt. Comm.* (to be published).

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Programmable 2-D optical wavelet transform in the space domain

Dong-Xue Wang, Ju-Wei Tai, *Yan-Xin Zhang and *Yu Ma

Institute of Automation, Academia Sinica, P O Box 2728, Beijing, PRC,
*Institute of Modern Optics, Nankai University, Tianjin, PRC

Abstract. In the paper, an optical wavelet transform system with a lenslet array as well as its implementation in the space domain is presented. Several optical experimental results have been demonstrated which are in keeping with computer simulations.

1. Introduction

In recent years, much attention has been paid to the wavelet transform (WT) for it shows great potential in transient signal processing [1-4]. WT is an efficient way to represent a signal (image) in both time (space) and frequency domains. On the other hand, the WT is a time consuming process, and high operation speed computational tools are needed to implement it. The optical implementation of the WT, taking advantage of high speed, parallel processing, is desirable. Some optical wavelet transform (OWT) methods in which WT is considered as a bank of filters and some results that have been obtained by the Fourier transform method through spatial filtering in the frequency domain have been reported. OWT in the space domain has also been proposed [5-6]. In this paper, a programmable OWT in which a lenslet array is employed, and the optical Haar wavelet forms is realised in the space domain for binary images, is reported.

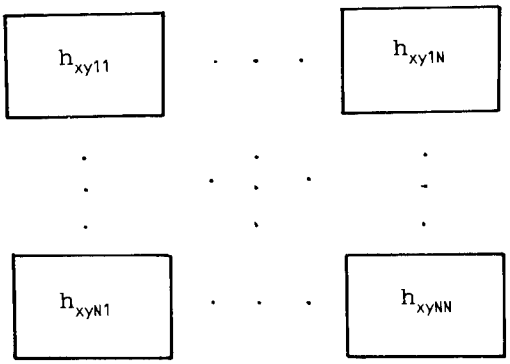


Fig. 1 Space arrangement of wavelets

2. 2-D Discrete wavelet transform representation

There are two forms for Wavelet Transform: the continuous and discrete form. In the case of a 2-D discrete wavelet transform (DWT), the WT expansion coefficients W_{ijkl} of the image s(x,y) can be formed as follows:

$$W_{ijkl} = \sum_x \sum_y s(x,y) h\left(\frac{x - x_i}{a_j}, \frac{y - y_k}{a_l}\right) \tag{1}$$

The omission of the normalisation factor in the formula does not effect the WT operation [7].

In the 2-D DWT, the coefficients W_{ijkl} form a 4-D matrix, with the discrete translation factor (x_i,y_k) and the dilation factor (a_j,a_l). The WT coefficients with a given dilation factor can be represented by a 2-D matrix indexed with a translation factor (x_j,y_k). In our experiment, the results of WT with NxN discrete translation factors and a selected dilation factor are obtained in parallel on a 2-D output plane. The coordinates of the output plane correspond to the discrete translation factor (x_i,y_k). The WT for different dilation factors can be implemented step by step, and each step corresponds to a WT under a fixed dilation factor. Because the programmable OWT system is under the control of a PC, dilation factors can be easily selected.

3. Programmable OWT system

For the fixed dilation factor (a_j,a_l), the wavelets are represented by h_{xyik}:

$$h_{xyik} = h\left(\frac{x - x_i}{a_j}, \frac{y - y_k}{a_i}\right) \tag{2}$$

It can also be considered as a 4-D matrix, and the wavelet matrix (WM) can be arranged with a 2-D array (indexed by i,k) of 2-D submatrices (indexed by x,y). These wavelets are displayed on a high resolution video monitor addressed by a CRT monitor according to the above space arrangement (Fig. 1).

The experiment uses 8x8 discrete wavelets, and each wavelet consists of 8x8 discrete values which can be calculated from Eq. (2). The indexes of the translation factor are from (1,1) to (8,8) with one fixed dilation factor.

In the experiment, a lenslet array consisting of 8x8 lenses is to be used to implement the optical summation of Eq. (1). Each lenslet of the array is of 8.2 mm diameter and 385 mm focal length. Each lens images a specific wavelet submatrix onto the input image which is an SLM with 8x8 binary pixels. The images of the wavelet submatrices can be added to the input image, and the transmitted beam from the SLM is collected onto the output plane to form a 8x8 pixel array which represents 8x8 products of the 2-D wavelet submatrices with an input image (Fig. 2 and Fig. 3). Each of the pixels indexed by the translation factor (x_i,y_k) on the output plane corresponds to a WT of the input image under one wavelet indexed by the same translation factor with a specific dilation factor.

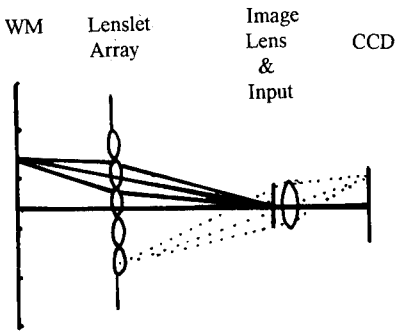


Fig. 2 Optical design of the OWT system

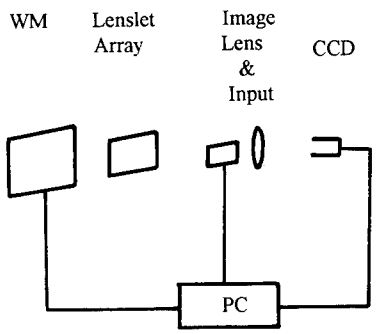


Fig. 3 Schematic diagram of the OWT system

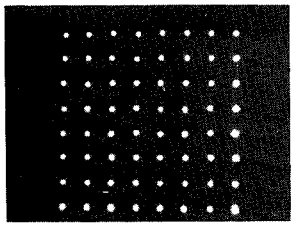


Fig. 4a The wavelet mask of OWT

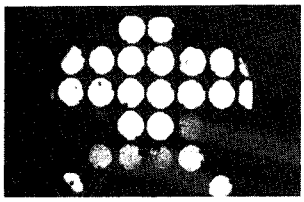


Fig. 4b OWT of the input image "Da"

The wavelets have both positive and negative values, there are several common ways to process the negative value optically [2]. In the experiment we choose two steps: first, all the positive values of the wavelets are displayed and the OWT results for the input for positive wavelets are available, then the negative wavelets are processed in the same way. Finally, the above results detected by a CCD are subtracted with a computer.

4. Optical Haar Wavelet Transform

In the Haar Wavelet Transform, a 2-D Haar mother wavelet can be described as follows:

$$h(x,y)=\text{rect}(x-0.5,y-0.5)+\text{rect}(x+0.5,y+0.5)-\text{rect}(x+0.5,y-0.5)-\text{rect}(x-0.5,y+0.5) \quad (3)$$

We choose the Haar wavelet: $a_j=a_l=1, x_i, y_k=1,2,...8$. The input binary image s(x,y) also consists of 8x8 discrete pixels. In this case, substituting Eq. (3) into Eq. (1), we can obtain the equation: $W_{ijkl}=s(i,k)$; so the WT coefficients with the dilation factor can be denoted by w_{ik}; therefore, these coefficients of selected dilation factors can form a 2-D matrix. So the WT of an image has the same form as the input image for the dilation factor: $a_j=a_l=1$. The optical result is in keeping with computer simulation. Although the WT of the image under the dilation factor has no obvious sense, the results can be easily tested through computer simulation, and provide a pragmatic analysis of confirmation of the optical results, especially for a processed image with a few pixels.

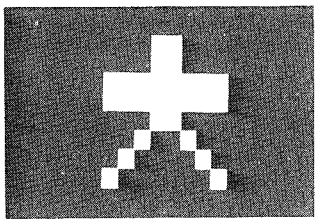


Fig. 5 The input image "Da"

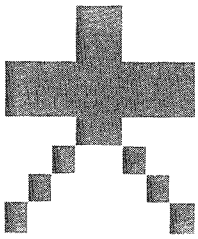


Fig. 6 Computer simulation result of WT

The Chinese character "Da" is an input binary image for the experiment, and the character makes up 8x8 dots. The discrete binary image, Haar wavelets with the dilation factor (a_j=1,a_i=1) have no negative values, and the result can be obtained from Eq. (1). The wavelet mask displayed on a video monitor is shown in Fri. 4a, and the OWT of the input image "Da" is shown in Fig. 4b. This result is in keeping with computer simulations (Fig. 6).

5. Conclusion

We proposed an architecture of an OWT in the space domain. The optical Haar wavelet transform is implemented with the method. However, we need a great deal of engineering techniques to design a lenslet array consisting of a large number of lenses to implement a DWT with fine dilation and translation factors for images made up of many pixels. Moreover, we also need an SLM with a large dynamic range of grey levels to represent wavelets with multiple values.

Acknowledgement: The work was supported by the National Natural Science Foundation of China.

References

[1] Szu H, Sheng Y and Chen J *Appl. Opt.* **31** 3267 (1992)

[2] Sheng Y, Lu T, Roberge D and Caulfield H J *Opt. Eng.* **31** 1859 (1992)

[3] Zhang Y and Li Y *Opt. Lett.* **17** 21 (1992)

[4] Freszy E, Pouligny B, Argoul S and Arneodo A *Phys. Rev. Lett.* **64** 745 (1990)

[5] Yang X, Szu H, Sheng Y and Caulfield H J *Opt. Eng.* **31** 1846 (1992)

[6] Wang D-X, Tai J-W and Zhang Y-X *Appl. Opt.* **33** Aug 10 (1994)

[7] Burns T J, Fielding H, Rogers S K, Pinski S D and Ruck D W *Opt. Eng.* **31** 1852 (1992)

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Low Resolution Fresnel Encoded Lenses Applied to Pattern Recognition

E. Carcolé, I. Juvells

Universitat de Barcelona, Departament de Física Aplicada i Electrònica, Laboratori d'Òptica, 08028 Barcelona

J. Campos

Universitat Autònoma de Barcelona, Departament de Física, 08193 Bellaterra, Spain.

Abstract: The frequency response of Low Resolution Fresnel Encoded Lenses is calculated in terms of all parameters that characterizes these lenses and possible applications to pattern recognition are discussed.

1. PSF of Low Resolution Fresnel Encoded Lenses

When a Fresnel Lens with focal length f for a wavelength λ is encoded in a pixelated low resolution device, with a center to center pixel distance given by Δx, Δy, infinite new focal regions appear at the coordinates (kX,lY) where $X = \lambda f/\Delta x$ and $Y = \lambda f/\Delta y$ where k,l are arbitrary integers[1]. A low resolution fresnel encoded lens (LRFEL) seems an array of lenses with a size given by XY. Then, if the device has NxM pixels with a rectangular pupil of dimensions $L_x = N\Delta x$ and $L_y = M\Delta y$, the number of lenses appearing in the device is given by $W_x = L_x/X$ and $W_y = L_y/Y$. In Fig.1 we can see a LRFEL with $W_x = W_y = 3$.

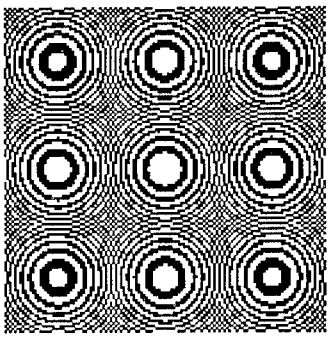


Fig.1. LRFEL with $W_x = W_y = 3$.

The amplitude distribution at a k,l focus for a plane wave illumination is given by (Eq. 20 of Ref.[1]):

$$U_{k,l}(x_i,y_i) = \frac{1}{i\lambda f}\frac{L_x L_y}{\Delta x \Delta y}\left[sinc\left(\frac{L_x X}{\lambda f},\frac{L_y Y}{\lambda f}\right)\exp\left[i\frac{2\pi}{\lambda f}(x.kX+y.lY)\right]\right] * rect\left(\frac{x}{\Delta x'},\frac{y}{\Delta y'}\right) \quad (1)$$

The sinc function comes from the Fourier transform of the rectangular pupil of the device and the rect function defines the pixel of dimensions $\Delta x'\Delta y'$. Another useful parameters are $c_x=\Delta x'/\Delta x$ and $c_y=\Delta y'/\Delta y$.

2. Coherent transfer function for a (k,l) focus

In order to calculate the frequency response we must perform the Fourier transform of Eq.1.Taking as a natural unity of length the pixel size we obtain:

$$H_{k,l}(g_x,g_y) = \frac{\lambda f c_x c_y}{i} rect\left(\frac{g_x-k}{W_x},\frac{g_y-l}{W_y}\right) sinc(c_x.g_x,c_y.g_y) \quad (2)$$

For a (k,l) focus, the rect function is shifted k in the g_x direction and l in the g_y direction. In a general case (any shape of the pupil or the pixel) it will happen the same, we will have the integral of the Fourier transform of the pixel function multiplied by the shifted pupil function. Now, let us study some different cases:

First we consider (k=l=0). In this case the cut-off frequency is determined by the rect function at the frequency (in pixels^{-1}) $g_{c,x}= W_x/2$, $g_{c,y}= W_y/2$. The first zero for $W_x > 2/c_x$ or $W_y > 2/c_y$ of H is given by $g_{0,x}=1/c_x$ $g_{0,y}=1/c_y$. If $W_x < 2/c_x$ or $W_y < 2/c_y$ then $g_{0,x} = g_{c,x}$ or $g_{0,y} = g_{c,y}$. For smaller values of the focal length H take negative values and inversions of contrast may happen. For $c_x->0$ and $c_y->0$ we obtain H corresponding to the infinite resolution case (with a diffraction efficiency tending also to zero). Then the limitations imposed for the low resolution condition is a function of c_x, the ratio between the pixel size and the center to center pixel distance.

We can now make general considerations for different k,l. When $c_xW_x/2$, $c_yW_y/2 >> 1$ then the rect function is large compared with the sinc function and his effect will not be appreciable for $k<W_x/2-1/c_x$ and $1<W_y/2-1/c_y$. In that conditions each focal region verifying the inequalities will replicate near the same information. This situation is sketched in Fig 2 for $W_x=W_y=40$ and $c_x=c_y=1$. Note that the inequalities can be verified for the particular case of a short focal lens with $c_x,c_y\sim1$, or a large focal length with $c_x,c_y << 1$. When $c_xW_x/2$, $c_yW_y/2 << 1$, then the rect function is thinner than the sinc function,

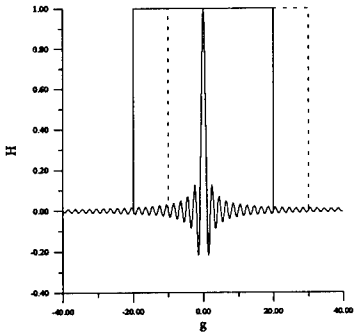


Fig.2 H for W_x = 40 and c_x = 1.

for k,l>0 H has the aspect of Fig. 2 The H for each k,l corresponds only to a little window of frequencies. In such case each focus will concentrate information corresponding to a certain range of frequencies. If W_x, W_y = 1 then the windows are the ones sketched in Fig 3 In this case each focus replicates different information but all information is replicated.

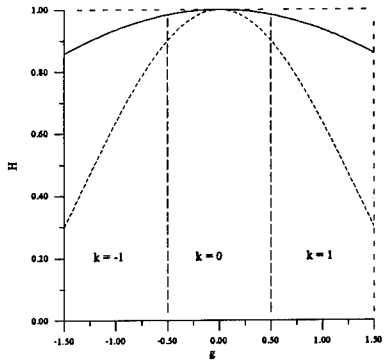


Fig.3 H for W_x= 40 and c_x= 0.3 (Dashed) and 0.2 (Solid)

3. Application to pattern recognition

We have seen that in certain conditions each focus can replicate different information of the object. The central focus can replicate the information corresponding to a window of low frequencies and the others ones can replicate windows corresponding to high frequencies. This will permit to compare separately the information corresponding to high and low frequencies and different kind of filters can be used in order to process appropriately this information.

We have also seen that for short focal length encoded each focus replicates near the same information. This opens the door to apply different types of filters for the same image in order to improve the recognition process.

Acknowledgments

This work has been financed by CYCIT (Comisión Interministerial de Ciencia y Tecnología). Project TAP93-0667-C03-01 and ROB91-0554

Bibliography

[1] E. Carcolé, J. Campos, S. Bosch, "Diffraction Theory of Fresnel Lenses encoded in low resolution devices" Appl. Opt. (In Press).

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Region of support to improve discrimination of Phase Only Filter

E. Ahouzi, J. Campos and M.J. Yzuel.

Physics Department, Autonomous University of Barcelona 08193 Bellaterra, Spain
e-mail: ifop1 at cc.uab.es.

Abstract We propose a Phase only filter with maximum discrimination capability. A zero modulation state is introduced to block certain spatial frequencies to equalize the phase difference histogram. The discrimination capability is increased significantly compared with that of the POF.

1. Introduction

The performance of different spatial filters with high discrimination capability (DC) has been investigated in recent years. Several types of filters have been proposed to increase this parameter (DC): Phase Only Filter (POF)[1], Optimal Filter[2] and other techniques [3]. The POF has a better discrimination than the classical matched filter and is more sensitive to noise. Hence, in [4] B.V.K.V. Kumar and Z. Bahri introduced the notion of optimal support function for a Phase Only Filter and they defined the Optimal Phase Only Filter (OPOF) in the sense of maximizing the signal to noise ratio SNR.. The support function indicates which pixels in the filter have magnitudes equal to 1 and which pixels have zero magnitudes. Other support functions to optimize different criteria or multicriteria have been proposed [5]. In all of these designs the support functions are calculated by taking into account the amplitude distribution. The phase distribution plays a crucial role in the discrimination capability and, in consequence, we propose a support function to optimize the DC based on the phase information.

In this contribution we propose a method that optimizes the discrimination capability by using a POF with a support function. The procedure is based on the modification of the phase difference histogram by blocking some frequencies. We investigate the DC and we show the improvement in DC obtained with this method. Simulated and experimental results are presented.

2. Theoretical analysis

We denote by t(x,y) and d(x,y) the target and nontarget respectively, when the target t(x,y) is present at the origin, and the nontarget d(x,y) is placed in the scene at coordinates that the maximum of the cross-correlation is at the origin. Let $T(u,v) = |T(u,v)| \exp[i\phi_t(u,v)]$ and $D(u,v) = |D(u,v)| \exp[i\phi_d(u,v)]$ be their Fourier transforms. When the POF is used the auto and the cross-correlation functions in Fourier domain are given by:

$$\begin{aligned} C_t(u,v) &= |T(u,v)| \\ C_d(u,v) &= |D(u,v)| \exp[i\Delta\phi(u,v)] \end{aligned} \qquad (1)$$

where $\Delta\phi(u,v) = \phi_d(u,v) - \phi_t(u,v)$ is the phase difference. The discrimination capability, DC, is a parameter that measures the ability of the correlator to discriminate between two very similar input objects, and it is defined as:

$$DC = 1 - \frac{\left| \sum_{u,v} | D(u,v) | \exp [i\Delta\phi(u,v)] \right|^2}{\left| \sum_{u,v} |T(u,v)| \right|^2} \qquad (2)$$

Let us divide the 2π phase interval in N steps of width $\delta\phi = 2\pi/N$. Let S_j be the set of pixels (u,v) for which the phase difference $\Delta\phi$ is in the interval $\Delta\phi_{j-1} < \Delta\phi \leq \Delta\phi_j$, where $\Delta\phi_j = \mathbf{j}.\delta\phi$. Let us define the weighted phase difference histogram as:

$$P_j = \sum_{(u,v)\in S_j} |D(u,v)| \qquad (3)$$

We approximate the phase difference of those pixels which belong to S_j by $\Delta\phi_j$, then, the DC can be written as:

$$DC = 1 - \frac{\left| \sum_{j=1}^{N/2} [\, P_j - P_{N/2+j} \,] \, e^{i\Delta\phi_j} \right|^2}{\left| \sum_{j=1}^{N/2} T_j \right|^2} \qquad (4)$$

The DC is optimized if we equalize $P_j = P_{N/2+j}$, for $\mathbf{j} = 1...N/2$. This can be made by blocking some frequencies. After the blocking process the phase histogram P_j is equal to $P_{N/2+j}$, and, in concequence the DC given by equation (4) is equal to 1. In practice it is not neccesary to make $P_j - P_{N/2+j} = 0$. It is enough that this difference be $|P_j - P_{N/2+j}| < \epsilon$, then $DC < N.\epsilon^2/2.E_{nbe}$. where E_{nbe} is the nonblocked energy. The DC is improved for small value of ϵ.

3. Computer simulation and experimental results

We invesigate the comparative performance of conventional POF and optimized POF in terms of discrimination capability between two similar objects . To illustrate the improvements of the discrimination with the proposed method, we perform numerical expriments using alphanumeric characters. Figure 1 shows the input scene, the target is the letter **F**, and the ob ect to be re ected is the letter **E** (letter **F** is contained in letter **E**). In figure 2a and 2b, we show the original and the modified weighted phase difference histograms, with N equal to 100. The corresponding correlations are shown in figure 3a and 3b. By comparing these figures it is shown that the algorithm proposed in this paper, enhances clearly the discrimination capability of the POF.

In the optical experiment we used the scene of figure 1. The scene and the filter were obtained with the commercial polygraphic machine Linotronic 630 with a resolution of 3251 dpi. The filters are codified by Burckhardt's method [6] in a size of 9X9 mm.. The light source is a 15 mW linearly polarized He-Ne Laser.

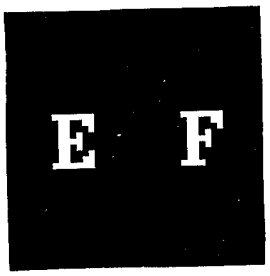


Fig.1. Input scene.

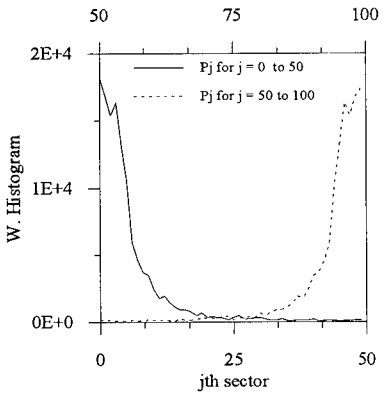


Fig. 2a. The original Weighted Phase Difference Histogram.

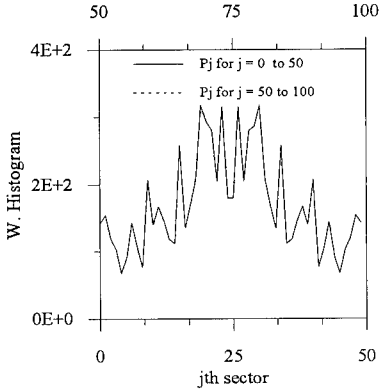


Fig. 2b. The modified Weighted Phase Difference Histogram.

The optical correlation is recorded by a Pulnix CCD detector array and stored in a computer by means of a 8-bit frame grabber. The experimental results with the conventional POF and with the optimized POF to improve DC, are illustrated in Fig. 4. Fig. (4a) shows a 3D plot of the correlation plane when the conventional POF is used. The DC is about 0.12. Fig. (4b) represents a 3D plot of optical correlation with the optimized POF for ϵ=300. The simulation and optical results are given in the table 1 for differente values of ϵ.

Table 1: Discrimination capability (DC) of optimized POF for different values of ϵ.

	POF	ϵ= 300	ϵ= 400	ϵ= 600	ϵ= 1000
Simulation	0.12	0.91	0.88	0.81	0.71
Optical experiment	0.14	0.89	0.90	0.87	0.74

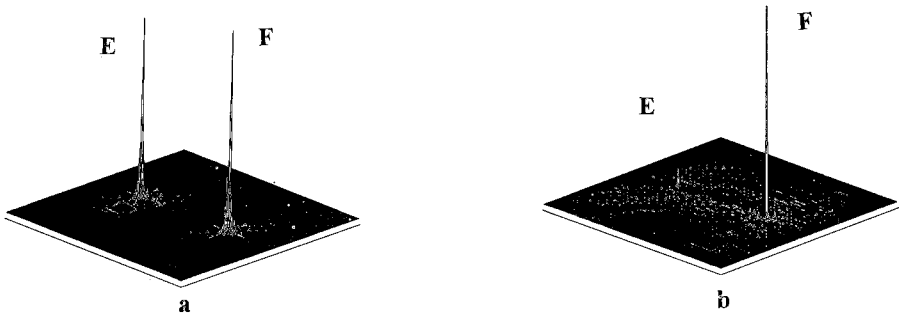


Fig. 3. Correlation planes with computer simulation. a) With conventional POF. b) With optimized POF according with DC.

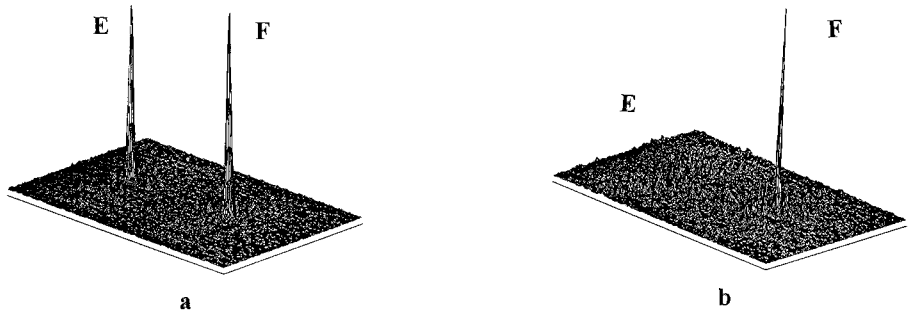


Fig. 4. Correlation planes (experimental results). a) With conventional POF. b) With optimized POF according with DC.

4. Summary.

A method to improve the DC of POF is given. It consist in blocking some frequencies. The DC increases when ϵ descreases. The experimental results are in accordance with the computer simulated ones. The sharpness of the autocorrelation peaks is improved with optimized POF.

Acknowledgments

This work has been financed by CICYT(Comision Interministerial de Ciencia y Tecnologia). Project: TAP 93-0667-C03-01. One of us E. Ahouzi wants to tank the Spanish Ministry of Education and Science for a grant.

References

[1] Horner J L and Gianino P D 1984 Appl. Opt., **23** 812-818
[2] Campos J, Turon F, Yaroslavsky L P and Yzuel M J 1993 Int J. Opt. Computing **2**, 341-365
[3] Zhou H, Zao F,Yu F T S, Chao T H 1993 Opt. Eng. **32** 2720-2725
[4] Kumar K V and Bahri Z 1989 Appl. Opt. **28** 250-257
[5] Réfrégier Ph, Vijaya Kumar B V K and Hendrix C 1992 J. Opt. Soc. Am. **A9** 2118-2125
[6] Campos J, Turon F, Yzuel M J and Vallmitjana S 1990 Appl. Opt. **29** 5232-5234

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Modified Filter Synthetic Discriminant Functions for Pattern Recognition

R.K. Wang, David Budgett, and C.R. Chatwin
Laser and Optical Systems Engineering Group
Glasgow University, Glasgow G12 8QQ, UK

Abstract: The correlation performance of SDF's is improved using a filter modulator to enhance higher frequencies from individual training images. A second operator containing all device dependencies allows a hardware specific optimal filter to be derived.

1 Introduction

Although filter synthetic discriminant functions (fSDFs)[1,2] are capable of producing distortion torelant filters, they suffer from poor dynamic range when implemented on spatial light modulators (SLMs) capable of encoding multi-level values. This is due to most of the energy being concentrated in a very narrow low frequency band leaving the mid-band area with energy levels below the smallest modulation level on an SLM. In order to solve this problem, Wang and Chatwin[3,4] investigated the utility of modulating individual training set images and linearly combining them to produce an SDF which itself was then modulated; this synthesised filter is called the modified fSDF (MfSDF). The MfSDF presented here synthesizes the SDF from the linear combination of a set of training images which are already filter modulated so that the SDF constructed is dominated by the higher, not the lower, frequency components of individual training set images. The filter-encoding constraint is then applied to the SDF. This scheme enables multi-grey level SLM capabilities to be fully utilised. The performance of this scheme is compared with that possible using a binary SLM.

2 MfSDF Filter and its Design

The MfSDF extends the fSDF filter design[1,2] to make it more flexible. It begins with a set of centered training set images $t_n(x,y)$, $n = 0, 1, ..., k$, spanning the desired distortion invariant feature range. The modulation operator, $\mathcal{N}$, is first imposed on the individual training image set to establish the pre-processed training set images $t'_n(x,y)$:

$$t'_n(x,y) = \mathcal{F}^{-1}\mathcal{N}\mathcal{F}[t_n(x,y)] \tag{1}$$

where $\mathcal{F}$ is the Fourier transform operator. The pre-processed training image set is used to construct the MfSDF, $s'(x,y)$, for a selected filter modulation, $\mathcal{M}$. The desired peak correlation response of $s'(x,y)$ is a constant, c_n, for each training image $t_n(x,y)$:

$$\int\int t_n(x,y)s'^*(x,y)dxdy = t_n(x,y) \odot s'(x,y) = c_n \tag{2}$$

where the integral is taken over the area of the input field. The function $s'(x,y)$ includes the filter modulation, $\mathcal{M}$, through the equation:

$$s'(x,y) = \mathcal{F}^{-1}\mathcal{M}\mathcal{F}[s(x,y)] \tag{3}$$

The purpose of the MfSDF procedure is to determine the function $s(x,y)$ which solves Eq(2) for a specified modulation function, $\mathcal{M}$. The function $s(x,y)$ is chosen to be a linear combination of the pre-processed training set images. Thus, a general MfSDF synthesis equation is derived as:

$$t_n(x,y) \odot \mathcal{F}^{-1}\mathcal{M}\mathcal{F}\{\sum_{n=0}^{k} a_n \mathcal{F}^{-1}\mathcal{N}\mathcal{F}[t_n(x,y)]\} = c_n \tag{4}$$

Fig1 gives a simple flow chart of this MfSDF filter design procedure.

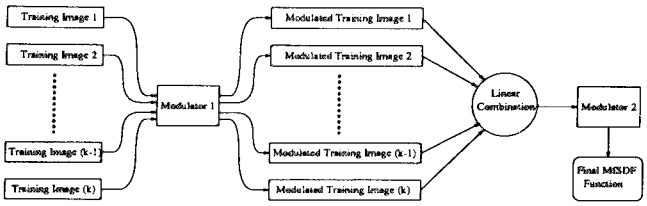


Fig1. Flow chart of the MfSDF design procedure

The modulation operators, $\mathcal{M}$ and $\mathcal{N}$, can be specified to take on any desired form. Whether or not the MfSDF produces equal correlation peaks (ECP) for all the training set images depends on the choice of the modulation operators $\mathcal{M}$ and $\mathcal{N}$. Thus, Eq(4) is frequently a system of nonlinear equations which may be solved using an iterative procedure[5] based on the Newton-Raphson algorithm.

3 Simulation and Results

The SDF function $s(x,y)$ was constructed using POF modulation of the individual training set images. The modulation operator $\mathcal{M}$ was specified to give phase and amplitude modulation within the limitations of the LCTV, as this may be implemented in real time on the commercially available Epson LCTV SLMs. Consistent with the Epson SLM performance, the number, N, of discrete levels of phase and amplitude (MLAP) information was chosen to be 16. The binary phase only filter (BPOF) modulation in consistent with using a binary SLM which are commercially available; hence, the MLAP scheme is compared with that possible using a binary SLM.

The following filter performances are considered: distortion invariant range, discrimination sensitivity between in-class and out-of-class images, and ability to accommodate the input image noise. The training set images used in the simulations consist of in-plane rotated images of the in-class Bradley APC vehicle and out-of-class Abram MI tank. Each image is centered and normalised to unit energy.

Filters designed to be invariant to in-plane rotation, over distortion ranges up to 180°, are constructed using training images of the in-class APC vehicle and out-of-class MI tank separated by a rotation increment of 5°. The correlation peaks are specified to give a constant value of c for the in-class training set images and zero for the out-of-class training set images at the center position of output plane. For example, the MfSDF designed for invariance to in-plane rotation over a distortion range of 45° is constructed

from ten in-class and ten out-of-class training set images. After construction, filters are correlated with images, spanning their entire design range of 45°, at 1° intervals.

The PSR_w data, which is usually used to evaluate the distortion invariant range of the filter, for the MLAP/MfSDF and BPOF/MfSDF with distortion ranges up to 130° are displayed by Fig2. From Fig2, it can be seen that the value of PSR_w drops below 1 for a distortion range of approximately $65° \sim 70°$ for the BPOF/MfSDF; whereas, for the MLAP/MfSDF it happens at $120° \sim 125°$, almost twice that of the BPOF/MfSDF.

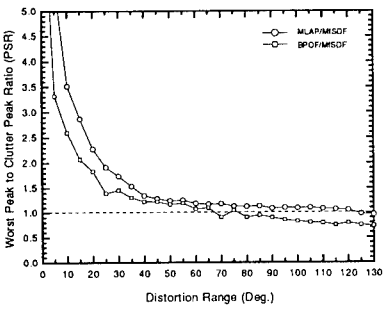


Fig2. The worst correlation peak response to secondary peak response ratio (PSR_w) over the distortion ranges from 0° to 130°.

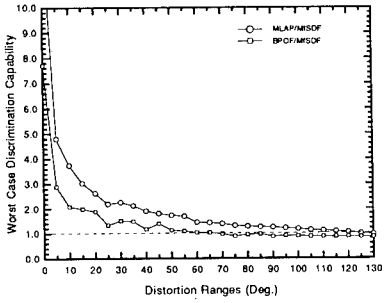


Fig3. The worst discrimination capability (DC_w) of filters between in-class and out-of-class images over the distortion ranges up to 130°.

Fig3 shows the worst-case discrimination values DC_w of MLAP/MfSDF and BPOF/-MfSDF measured at any angle within each distortion range up to 130°. It can be seen from Fig3 that, for the MLAP/MfSDF, the value of DC_w drops below one for a distortion range of $120° \sim 125°$; hence, the MLAP/MfSDF can be designed for an invariant distortion range of at least 120° whilst still maintaining 100% discrimination against the out-of-class targets with training images spaced at 5°. Whereas, for the BPOF/MfSDF, the DC_w value drops below one at a 60° distortion range, just half the value achieved by the MLAP/MfSDF method. It should be noted that although the BPOF/MfSDF can achieve the invariant distortion range of 65° from the conclusion of the previous paragraph, it does not guarantee a 100% discrimination capability between the in-class and out-of-class images; thus its effective invariant distortion range is 60°. This is an issue frequently not addressed by other authors.

The filter noise resistance performance is examined by using in-plane rotated images of the noise corrupted in-class Bradley APC vehicle and out-of-class Abram MI tank; the ratio of image energy to noise energy is equal to 0.5 which means that the images are severely corrupted by noise.

Fig4a shows the in-class and out-of-class peak correlation responses, using the noise corrupted in-class and out-of-class images as the input images, for a MLAP/MfSDF constructed from the noise free in-plane rotated training images with a distortion range from 0° to 40°. The training images used are 5° apart. A similar gragh for the BPOF/MfSDF is shown in Fig4b. From Fig4, it is clear that the MLAP/MfSDF filter is invariant to the distortion range of 40° whilst delivering a superior discrimination capability (the value of $DC > 1.6$ everywhere) between the in-class and out-of-class images in this noise corrupted case; whereas, the BPOF/MfSDF does not give complete invariant distortion over this range as several secondary correlation peaks and out-of-class correlation peaks exceed the lowest in-class correlation peaks.

The PSR_w data for the MLAP/MfSDF and BPOF/MfSDF, with distortion ranges up to 60°, are displayed by Fig5. The resulting data obtained are similar to those in Fig2 and Fig3 except that the input images used are the noise corrupted images. From

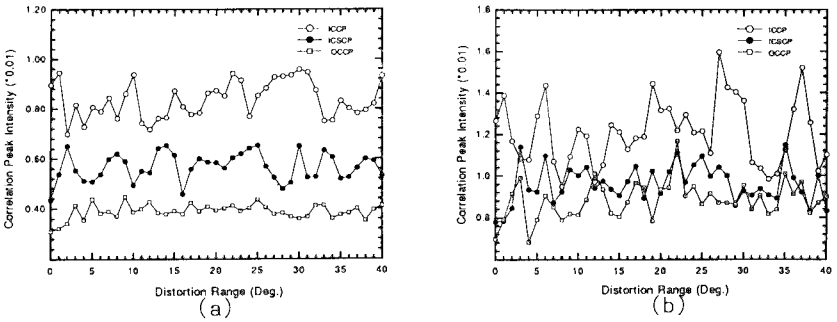


Fig4. In-class Correlation peak (ICCP), secondary peak (ICSCP) and out-of-class correlation peak (OCCP) responses with a distortion range from 0^o to 40^o; (a) MLAP/MfSDF and (b) BPOF/MfSDF.

Fig5, for the noise corrupted image inputs, the MLAP/MfSDF still delivers distortion invariance, with a 100% discrimination capabiltiy between the in-class and out-of-class images, up to at least 45^o; whereas, the BPOF/MfSDF achieves less than 15^o.

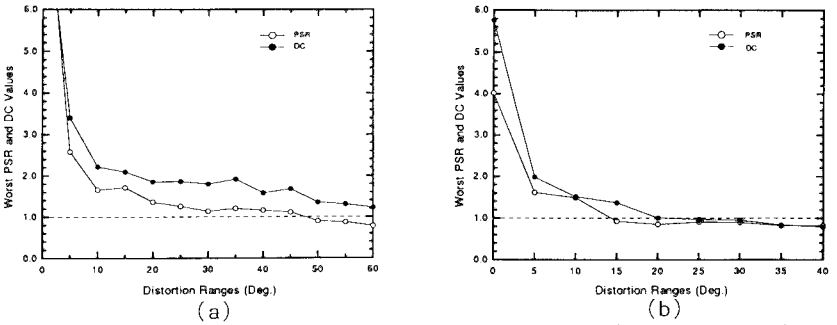


Fig5. The worst correlation peak response to secondary peak response ratio (PSR_w) and the worst discrimination capability (DC_w) of filter between in-class and out-of-class images over the distortion ranges from 0^o to 60^o: (a) MLAP/MfSDF and (b) BPOF/MfSDF.

4 Conclusions

Using 16 levels of modulation, a modulator operator has been formulated to represent the SEIKO-Epson SLM. Phase only modulation on the individual training images was selected to allow an initial investigation of the MfSDF design, it is not the optimal choice. The results show that the MLAP/MfSDF filter offers much improved correlator system performance with a greater allowable image distortion range whilst maintaining 100% discrimination between in-class and out-of-class images; furthermore, it shows an improved ability to accommodate input image noise when compared to the MfSDF with a binary phase only constraint.

References

1. D. Jared and D. Ennis 1989 *Appl. Opt.* **28** 232-9
2. D. Jared 1989 *Appl. Opt.* **28** 4335-39
3. R.K. Wang, C.R. Chatwin and M.Y. Huang 1994 *Appl. Opt.* **33** in press
4. R.K. Wang, and C.R. Chatwin *Appl. Opt.* **34** in press
5. G.F. Schils and D.W. Sweeney 1986 *J. Opt. Soc. Am.* **A3** 1433-42

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Multi-distance, Multi-Projection, Parallel Projection Method

Tuvia Kotzer† Nir Cohen† Joseph Shamir† and Yair Censor‡

† Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel.

‡ Department of Mathematics and Computer Science
University of Haifa
Mt. Carmel, Haifa 31905, Israel

Abstract. A parallel projection method is introduced which solves shortcomings of conventional projection methods and alleviates many of their limitations. This process can be efficiently employed for image processing, pattern recognition, optimization and many other signal processing tasks.

1. Background and Development

With the increasing demands and complexity of signal processing systems, renewed interest emerged in the set theoretic formulations [1] applied to optimization problems such as signal synthesis in pattern recognition, constrained deconvolution, image restoration etc. The task is usually: "Given some *a priori*, corrupted, signal, h^o, which is known to have, supposedly, satisfied N constraint sets, try to restore it to a signal which will satisfy the constraint sets". A similar framework appears in design problems.

Projection methods have been suggested for this task. A projection of an arbitrary element, h, onto a closed convex set C_i is that unique element in the set ,h', closest to h, where "close" is measured by some distance function d_i. Namely,

$$P_{C_i}^{d_i}(h) = h' \quad \text{if and only if } \inf_{h_1 \in C_i} d_i(h_1, h) = d_i(h', h); \quad h' \in C_i.$$

Given N sets C_i one usually constructs operators of the form $T^{k+1} = \sum_{i=1}^{N} w_k(i) P_{C_i}^{d_i}$ where $w_k(i)$ is some, iteration dependent, averaging function. This leads to a sequence $h^{k+1} := T^{k+1}(h^k)$. One strives to show that the sequence generated converges to a solution satisfying all constraint sets, i.e., that $\lim_{k\to\infty} h^k \in C_0 := \bigcap_{i=1}^{N} C_i$. The results known to date, basically, amount to the following: if $C_o \neq \emptyset$ **and** d_i are equal **for all** sets involved **and** all sets C_i are convex then, under some general conditions, the sequence $\{h^o, h^1, \cdots\}$ converges (weakly) to an element in C_o.

However, what happens if the constraints are inconsistent, i.e., $C_o = \emptyset$ (which is a frequent occurrence due to detection error, slight mis-characterization of the set, too optimistics requests etc.)-see [2]? Is there convergence at all and if so to what? Also, how will one project onto a constraint set when the constraint is given indirectly/implicitly, as some pre-image of another explicit set? This may be transformed in many instances to an easier projection onto an explicit set but then the distance function is modified [3], what then? Finally, in many image recovery problems more than 2 constraints exist, where not all are convex:- can a projection based algorithm be employed (classically we are limited to two sets [4])?

Fortunately, all of the above problems may be solved by use of a special parallel projection method, based on [5]. It amounts to performing projections of the current estimate onto all N sets involved and then taking a special average of these projections.

2. Algorithm

The special parallel projection algorithm assumes the following form:
Initialization: $h^o(x)$ is an arbitrary function.
Iterative step: given the function $h^k(x)$, calculate,

$$v_i^{k+1}(x) \;:=\; P_{C_i}^{d_i}(h^k), \quad \text{for all } i = 1,\ 2, \cdots N, \tag{1a}$$

and then, set

$$h^{k+1}(x) \;=\; \mathcal{F}^{-1}\left\{ \frac{\sum_{i=1}^N \beta_i W_i(u) \mathcal{F}\{v_i^{k+1}\}}{\sum_{i=1}^N \beta_i W_i(u)} \right\}, \tag{1b}$$

where W_i are, frequency dependent, weighting functions in accordance with (2), see [2, 6, 7]. $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier Transform and its inverse, respectively, $\{\beta_i\}_{i=1}^N$ are user determined positive real numbers and upper case letters denote the Fourier transform of the lower case functions, e.g., $V_i(u) := \mathcal{F}\{v_i(x)\}$.

3. Characteristics of the algorithm

The algorithm has the following, desirable, characteristics:

1. Even if the sets are nonconvex or C_o is empty, the appropriate Summed (squared) distance error (SDE) decreases along the iterates, for any, finite, number of closed constraint sets, where the *appropriate* SDE, $\hat{J}$, is defined by

$$\hat{J}(h) = \sum_{i=1}^N \beta_i d_i(h, P_{C_i}^{d_i}(h)), \quad \text{and} \quad d_i^2(h_1, h_2) = \int W_i(u) |H_1(u) - H_2(u)|^2 du. \tag{2}$$

   We note that d_i are just weighted norm, with weight W_i.

2. If C_o is nonempty, the algorithm converges to C_o. Even if C_o is empty, the algorithm converges under general conditions [2] (provided all sets C_i are closed and convex). Moreover, it converges to one of the best solutions possible. Namely, one minimizing its distances (squared) from all sets involved, i.e., to an element minimizing globally $\hat{J}$.

3. Projections onto the individual sets may be performed w.r.t **different** distance functions. Thus, many implicit constraints may be converted to explicit constraints with modified distance functions, which are handled efficiently by this method [2, 6].

4. Examples

The present algorithm has already been employed successfully for pattern recognition purposes [6]. As an illustrative example to demonstrate the power of the algorithm we employ it to image restoration with inconsistent constraints. We asssume a linear degradation model of the form $g = f * h + n$ where f is the original image, h is the blurring function, n denotes the observation noise and g is the recorded, blurred and noisy observation. The task is to restore the original image f from its degraded observation g, where we assume that h is given as well as the statistics of the noise.

In our restoration example we use the following constraints:

1. The estimated image convolved with the blurring function yields a distribution which is *close* to the observed distribution, g, both in the spatial domain and in the frequency domain.

2. The estimated image distribution should be *close* to the Wiener solution which is known to be optimal in certain circumstances (in the MSE sense).

3. The estimated image should be real and positive with finite support, say $[a, b], \quad a < b$.

Any image, f, which satisfies all of the above four constraints is considered a solution. For full details and discussion see [7].

We consider a tank (of size 70×70 pixels) blurred by a blurring function 3×3 with additive white Gaussian noise. Our initial estimate (f_o), for initialization of our algorithm, is a constant background, space limited to $[4, 67]$ in both dimensions. Shown in Fig. 1 are the intensities of the original image, blurred and noisy image, the Wiener solution (f_w) and the restoration after 10 iterations, respectively. The improvement of our iterative approach over the Wiener result is quite evident. To quantify the results the mean square error (MSE) was calculated for reconstruction by the present algorithm and the Wiener solution. After 10 iterations with our algorithm the MSE was less than that of the Wiener solution by a factor of 79. This is, seemingly, surprising as the Wiener solution is the best result obtained by a linear filter, where the blurring mechanism is linear, and optimality is judged by the MSE criterion. Nevertheless, our restoration algorithm yields better results since the Wiener filtering approach does not exploit all the information (constraints) available.

Acknowledgements: This work was performed within the Technion Advanced Opto-Electronics Center established by the American Technion Society (ATS), New York. The work of Yair Censor was partially supported by NIH grant HL-28438 at the Medical Image Processing Group (MIPG), Dept. of Radiology, Hospital of the University of Pennyslvania, Philadelphia, PA, USA.

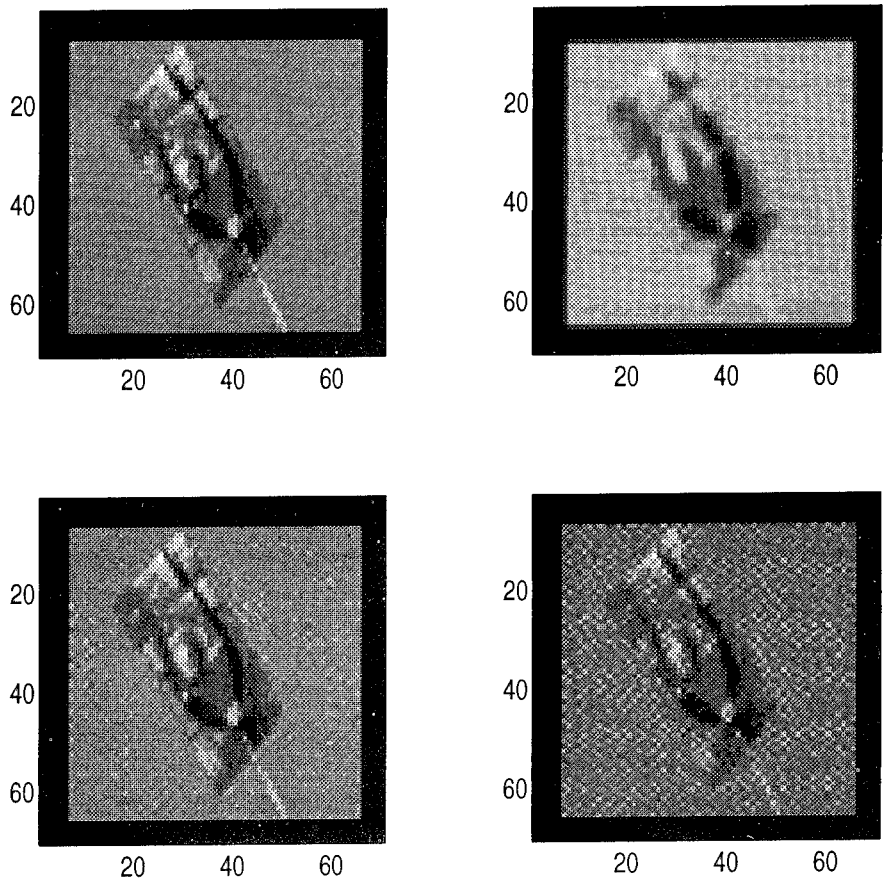


Figure 1. The restoration results. Clockwise from upper left: The original image, the degraded blurred and noisy image, the Wiener solution and the result of our Algoriothm after 10 iterations.

References

[1] P. L. Combettes, *Proc. IEEE* **81**, p. 182 (1993).

[2] T. Kotzer, N. Cohen and J. Shamir, EE Pub. No. 920, Technion- I.I.T., August 1994.

[3] T. Kotzer, N. Cohen and J. Shamir, EE Pub. No. 900, Technion- I.I.T., November 1993.

[4] A. Levi and H. Stark, *JOSA A* **1**, p. 932 (1984).

[5] Y. Censor and T. Elfving, "A multiprojection algorithm using Bregnman projections in a product space", Technical Report, Dept. of Math. & Comp. Sci., University of Haifa, Israel, December 1993, and accepted for publication to *Numerical Algorithms.*

[6] T. Kotzer, N. Cohen, J. Shamir and Y. Censor, EE Pub. No. 908, Technion- I.I.T., August 1994.

[7] T. Kotzer, N. Cohen and J. Shamir, EE Pub. No. 919, Technion- I.I.T., June 1994.

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Fourier optics and holography of temporal signals

Yu T Mazurenko, A G Spiro, S E Putilin and A G Beliaev

S I Vavilov State Optical Institute, 199034, St Petersburg, Russia

Abstract. We review the methods and possible applications of recording, reconstruction and transformation of optical time signals, while using interference, diffraction and nonlinear interaction of waves obtained by spatial spectral decomposition. The experimental results on time-to-space conversion of nanosecond signals using spectral nonlinear optics are considered.

Consider spatial spectral decomposition of a short optical pulse. The diagram of the optical spectral instrument used for this purpose is shown in the fig. 1. In this diagram G is the dispersive optical element that carries out the angular spectral decomposition of light. In our example G is the transmitting diffraction grating. The dispersive element is placed in the front focal plane F of the lens L. The spectrum plane of the instrument S coincides with the back focal plane of the lens L.

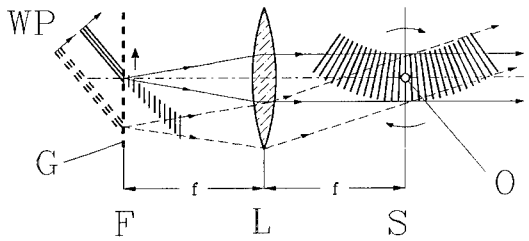


Fig. 1.

Let a short wave packet WP be incident on the diffraction grating G. The radiation of the wave packet scatters from the rules of the grating. We may consider that the moving wave packet gives rise to a small source of monochromatic radiation that moves over the plane of the grating. The radiation of a small monochromatic source moving along the front focal plane is converted by the lens L into the parallel beam. This beam rotates around the back focal point O of the lens. Irrespective of the beam rotation the velocity of wave propagation is the same at any point of the beam cross section. This is possible only if the beam is bent as is shown in the figure. The wavelength of the bent beam depends on the transverse coordinate and this just means the spectral decomposition of an

incident light pulse. Such rotating beam is the useful visual representation of a spectral decomposition wave (SDW) of a short light pulse.

The spatial amplitude-and-phase distribution in SDW corresponds to the amplitude-and-phase distribution in the Fourier image of the pulse. So the SDW contains the complete set of the spatially separated Fourier components of the initial temporal signal. The length of the SDW is equal to the inverse of the frequency resolution of the used spectral device and should be much more than the length of the initial wave packet.

Let us direct onto the entrance of the spectral device two short pulses - a reference pulse (having simple shape) and a signal pulse. Even if these pulses are separated in the time scale, their SDWs may be superimposed. Therefore two SDWs may interfere. The result of recording the interference pattern can be called a spectral hologram. It can be shown that the spectral hologram contains the total record of the signal pulse [1].

To reconstruct a signal pulse we use the double spectral device. The left-hand part of this device is the same as the instrument depicted in fig. 1, and its the right-hand part may be obtained by the mirror reflection of the left-hand part from the spectral plane S. The left-hand part forms the SDW of the incident time signal on the spectrum plane while the right-hand part executes the inverse operation of composing of frequency component. As a result we again obtain temporal signal at the exit of the double spectral device. We put the developed spectral hologram in the spectrum plane, and we direct the reference pulse onto the entrance of the device. The SDW of a reference pulse strikes the hologram and diffracts from the hologram inhomogeneities. As a result we obtain a new wave. It may be shown that this wave contains the replica of the SDW of a signal pulse. After spectral composing this radiation by the right-hand part of the device, we obtain the replica of the signal pulse stored on the hologram. Besides, we may obtain its time-reversed replica.

The reconstruction of optical temporal signals by diffraction of spectral decomposition waves was proposed in [2]. These results were further developed and reviewed [3,4]. The first experiments were done in [5,6]. Considered area may be called spectral holography.

There is a close analogy between the transformations of time-varying signals by filtering their SDWs and the methods of normal Fourier optics. Using this analogy it is possible to develop Fourier optics for temporal signals from the well-known principles of Fourier optics for steady state radiation. First it concerns with the Fourier transform holography that is the most powerful method of the Fourier optics [7].

The natural consequence of the analogy of spectral holography with the normal holography is the methods of dynamic spectral holography and spectral nonlinear optics [8]. SDWs of several light pulses may be superimposed in a common optical nonlinear medium. This makes possible the dynamic interaction of SDWs. As the result of interaction a new wave may generate. It can be shown that this newly generated wave usually has the properties of a SDW of a certain wave packet. This new wave packet can be obtained in the real form with the use of a spectral device operating on backward waves. The spectrum of a newly generated pulse may be considered as the product of spectra that participate in the interaction. Therefore the newly generated temporal signal is actually the result of fast convolution and/or correlation of the initial temporal signals.

In addition, interactions of SDWs with the monochromatic inhomogeneous waves are possible. We may consider the monochromatic wave as produced by the optical Fourier transform of some spatial signal (coherent 1-D image). In this case we obtain convolution or correlation, in which both temporal and spatial signals participate. The result of such transform may be a spatial as well as a temporal signal. In other words, we may mix

spatial and temporal signals in real time and obtain new spatial or temporal signal. This allows to carry out dynamic space-to-time or time-to-space transformations of signals [8,4,5].

The upper bond for the duration of optical pulses that can be processed by the methods of spectral holography and spectral nonlinear optics is imposed by the frequency resolution of a spectral device. While using a biconfocal Fabry-Perot interferometer [9] this bond may be as great as hundreds of nanoseconds.

Here we examine in more detail dynamic time-to-space transformation using sum-frequency generation of two SDWs.

Consider optical device described in the fig. 2. The left-hand part of this device is the optical spectral instrument such as depicted in fig. 1. The SDWs of the incident pulses are formed near the spectrum plane S. This plane is also the entrance plane of the optical Fourier processor, which is the right-hand part of the fig. 2 device. The lens L of the right-hand part of the device transforms spatial distribution of monochromatic field on the S plane into its Fourier image on the F plane.

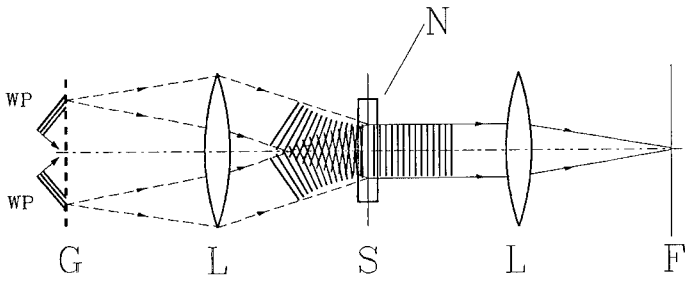


Fig. 2.

Let two wave packets (optical pulses) WP enter the device of fig. 2. One of the pulses is the reference one and the other is the signal one. The SDWs of two pulses are superimposed in the spectral plane S. The spectral device of is so arranged that the directions of the spectral dispersions for two considered pulses are opposite. This means that the "blue" end of one spectrum on the S plane corresponds to the "red" end of another spectrum. This also means that SDWs of the two pulses rotate in the opposite directions. In our example this is realized by projecting two pulses on the diffraction grating G from the different directions. Due to this, two light sources move over the grating with the same speed but in the opposite directions, giving rise to the two parallel beams near the S plane that rotate with the same speed but in the opposite directions.

We place a plate of the quadratic nonlinear medium N in the vicinity of the spectral plane S and examine the interaction of two SDWs in this medium. We may divide the nonlinear medium into elementary volumes, determined in the transverse direction by the spatial resolution of optical system and in the longitudinal direction by the depth of field. Radiation interaction in any of these small volumes may be thought of as the interactions of plane monochromatic waves. Due to the interaction of two SDWs in the quadratic medium we obtain sum-frequency generation in every elementary volume. Since the changes of frequency for two SDWs along the transverse co-ordinate have the opposite directions, we obtain new radiation whose frequency does not depend on the transverse co-ordinate, i.e., we obtain monochromatic wave. The spatial distribution of the amplitude and phase for the newly generated monochromatic radiation over the S plane is essentially the same as the corresponding distribution in the SDW of the signal

wave packet. So we the newly generated monochromatic wave is the analog of the Fourier image of the signal pulse.

It is not difficult to see that the amplitude-and-phase distribution of monochromatic radiation in the F plane repeats the temporal amplitude-and- phase structure of the signal pulse. So we obtain dynamic transformation of the temporal signal into the spatial signal. The time required for this transformation is limited only by the light velocity and the length of the optical device.

Consider some results of our experiments. The source of ultrashort light pulses was the Nd-YAG-laser with mode synchronization emitting ultrashort pulses with the duration of about 240 psec and wavelength 1.064 mm. These pulses served as the reference ones. Signal pulses with complicated shape were formed from the initial laser pulses using multiple reflections with time delays. To obtain spectral resolution sufficient for the operating with nanosecond pulses we used Fabry-Perot etalon with the side entrance as the dispersive element [4].

The example of dynamic time-to-space transformation is presented in the fig. 3. Here *J(t)* is the shape of the initial temporal signal and *J(x)* is the spatial distribution of monochromatic radiation obtained after the described transformations.

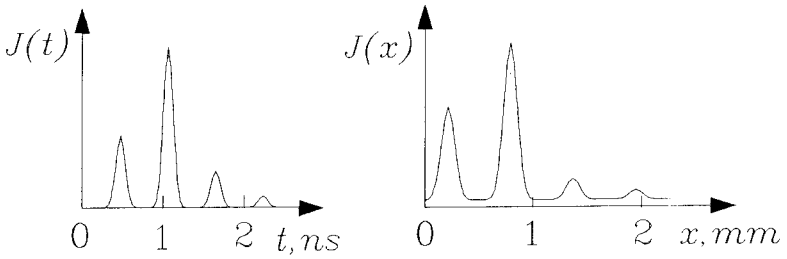


Fig. 3.

The research described in this publication was made possible using the grant of Russian Foundation for Fundamental Researches (Project code 93-02-14826) and NATO Collaborative Research Grant (Project code HTECH.CRG 931141).

References

[1] Froehly Cl, Colombeau B and Vampouille M 1983 *Progress in Optics* **20** 63-153
[2] Mazurenko Yu T 1984 *Opt. Spectr.* **56** 357
[3] Mazurenko Yu T 1990 *Applied Physics* **B50** 101-114
[4] Mazurenko Yu T 1992 *Optical Engineering* **31** 739-749
[5] Mazurenko Yu T et al 1991 *Opt. Spectr.* **71** 385-388
[6] Weiner A M et al 1992 *Optics Letters* **17** 224-226
[7] Mazurenko Yu T 1984 *Opt. Spectr.* **57** 1-6
[8] Mazurenko Yu T 1985 *Opt. Spectr.* **59** 35-37
[9] Mazurenko Yu T 1990 *Opt. Spectr.* **69** 147-148

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

"Spatial brightness amplification" for increasing signal–to–noise ratio in 2D optical information processing

G V Sinitsyn, S P Apanasevich, F V Karpushko and A V Lyakhnovich

Division for Optical Problems of Information Technologies,
Academy of Sciences of Belarus,
P.O.Box 1, 220072 Minsk, Republic of Belarus [1]

Abstract. The concept of "spatial amplification" based on transverse effects in optical bistability is proposed. The possibility is experimentally shown to detect with the help of a "spatial amplifier" weak optical signals of a power that is 10^7–10^8 times less than the power of noise.

1. Introduction

In a number of algorithms for optical information processing one needs to detect the presence or absence of information at least in one pixel of a wide–aperture matrix. Since the "0"–level signal power in optical systems is not equal to zero, the power of the background signal coming from the whole matrix is many times higher than the power of the information signal in any individual pixel. Therefore, recording of a weak signal is a serious problem and a challenge to experiment.

In the present contribution we propose the concept for increasing the signal–to–noise ratio in 2D optical information processing systems, which is based on application of transverse effects in optical bistability (OB). This concept allows to develop a technique for identification of the states of separate information pixels in wide–aperture 2D systems for information processing.

2. Principle of operation

The basic idea of the concept proposed is to provide favourable conditions for the emergence and propagation of switching waves in optical bistable layers [1,2] if the switch–on threshold has been exceeded at least in one local area. This leads to a many–fold

[1] E-mail: dopit%bas02.basnet.minsk.by@demos.su

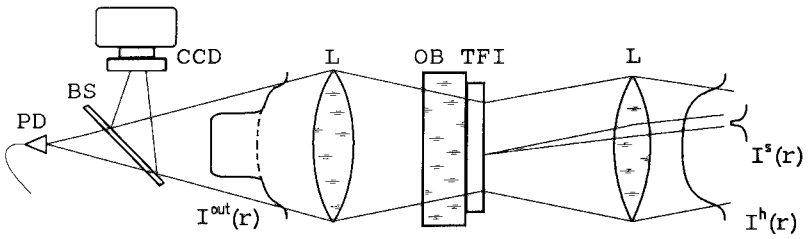


Fig. 1. Experimental setup. **OB TFI** – optically bistable thin-film interferometer; **L** – lenses; **BS** – beam splitter; **PD** – photodiode; **CCD** – CCD or TV camera; $I^S(r)$, $I^h(r)$ and $I^{out}(r)$ – intensity profiles of signal, holding and output beams respectively; dashed line shows output profile when $I^S(r)$="0".

increase of output light power up to the highest value which corresponds to total switching of the whole bistable plate.

Using an OB-matrix formed on an optically uniform nonlinear layer one can switch on not only a specified pixel but its nearest surrounding as well, and thus the resulting output signal which is registered and analyzed can be increased many-fold. The integral output beam is further directed either to photodetectors, when an optoelectronic information processing system is employed, or to an optical threshold device, when an all–optical computing system is considered.

3. Experimental

The technique proposed has been tested by modeling a "spatial brightness amplifier" experimentally using a thin-film bistable ZnS–interferometer (TFI) [3]. The experimental layout is shown schematically in Fig. 1. The samples used were prepared under special conditions which provided high uniformity of optical parameters over the whole optical aperture of about 36 *mm*.

Inhomogeneity of the nonlinear optical characteristics of the TFI was examined in the bistable mode of operation. Switching thresholds of bistable pixels were measured at seven equidistant points along the sample's diameter in at least two perpendicular directions. A set of bistable loops for different points with corresponding switch–on and switch off thresholds registered by oscilloscope is presented in Fig. 2. One can see that the relative spread of measured threshold powers does not exceed $\sim$ 10%.

Spatial resolution of our 2D bistable device is determined as the highest packing density of simultaneously and independently operating channels and it has been measured following the technique described in [4]. As a natural resolution criterion one can consider the minimal distance between nearest bistable pixels when the information

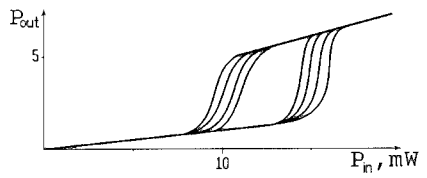


Fig. 2. Spread in threshold powers of bistable TFI along its diameter; oscilloscope drawing.

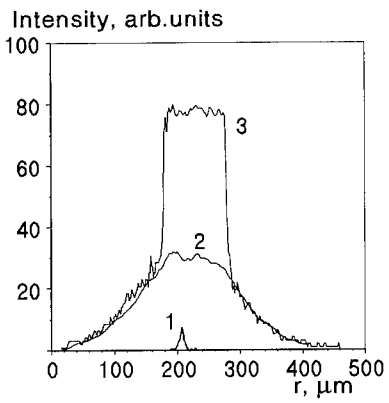


Fig. 3. Intensity profiles of beams: (1) – signal $I^S(r)$, (2) – output $I^{out}(r)$ when $I^S(r)$ ="0", (3) – output $I^{out}(r)$ when $I^S(r)$ ="1"

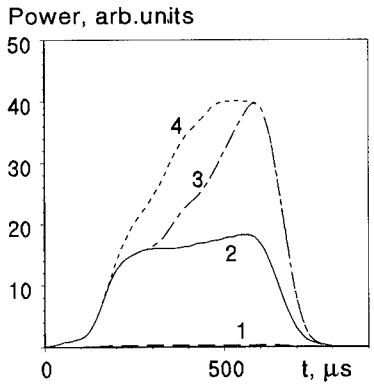


Fig. 4. Kinetics of integral output power of the beams: (1) – signal, (2) – output when $I^S(r)$ ="0", (3) and (4) – output when $I^S(r)$ ="1"

state of any pixel is not influenced by the states of its 4 neighbours, provided all of them are controlled by holding intensities within bistable region. For individual pixels with a size of about 10 μm the measured spatial resolution is 350...400 $lines/cm$, that corresponds to a total number of (1.2...1.6)·10^6 pixels over the whole aperture of TFI.

While modelling the "brightness amplifier" experimentally we formed in the plane of the bistable interferometer a certain distribution of incident light intensity that consisted of a relatively wide illuminated area with approximately uniform intensity distribution and a sharply focused gaussian beam with diameter of about $10\mu m$. The size of the wide area was limited by the maximum power of holding laser beam. In our experiment the output beam of an Ar–laser with power of 500...600 mW could illuminate an area of diameter $\sim$ 180 μm. The second beam modelled an individual pixel, the information state of which is to be detected. In order to form the above distribution of input light the laser beam was split into two beams. One of them (with lower intensity) was directly focused by the lens onto the surface of the TFI. The other (with higher intensity) was first widened by an additional lens. The power of each beam was controlled independently.

4. Results and discussion

The main experimental results are presented in Figs. 3–4.

Fig. 3 shows the intensity profile of the signal beam (1), the output distribution of background intensity on the whole illuminated area (2), and the output intensity profile (3) which indicates the presence of signal somewhere within the switched area. On Fig. 4 time dependencies are plotted for the power of the signal beam to be detected (1) and for the power of output light beam when there is no signal (2), and while detecting a local signal with maximum intensity that is 10% (3) and 20% (4) of the background intensity.

For the OB device operating as a "spatial brightness amplifier" the gain factor K can be defined as the ratio of integrated output power increase in the channel of the holding beam (or special probe beam), resulting from the emergence of ON-state in one of the pixels, to the integrated power of the latter.

$$K \simeq \frac{SI^hT_\uparrow - SI^hT_\downarrow}{\delta SI^ST_\downarrow} = \frac{S}{\delta S}\frac{I^h}{I^S}\left(\frac{T_\uparrow}{T_\downarrow} - 1\right),$$

where δS and I^S - are the square of cross-section and intensity of the signal light beam which switches the pixel being detected, I^h - intensity of holding beam (with a uniform distribution), S - the square of switched-on area, $T_\uparrow$ and $T_\downarrow$ - transmission of bistable device in switched-on (upper) and switched-off (lower) states. In our experiment $\delta S \simeq 75\mu m^2$, $S \simeq 2.7 \cdot 10^4 \mu m^2$, $I^S/I^h \simeq 0.1$, $T_\uparrow \simeq 60\%$, $T_\downarrow \simeq 25\%$. Thus, experimentally the obtained gain factor of our "spatial brightness amplifier" was about $5{\cdot}10^3$ and was limited by the output power of the laser available.

The above expression allows to estimate also the gain factor that can be reached when the holding beam totally covers the whole operating area. The diameter of the illuminated spot is then $36mm$, whereas the diameter of the signal beam and the intensity ratio are assumed to be $10\mu m$ and $I^S/I^h \simeq 0.1$, as in the experiment. The estimated limit of the gain factor reaches the value $K \simeq 1.8 \cdot 10^8$.

The gain factor of a "spatial amplifier" can be also increased at the expense of increasing the precision of intensity control (then a lower value of the I^S/I^h ratio can be reached). But this requires much better stabilization of the holding beam, both in its power and spatial uniformity, than that reached in our experiment.

The concept of "spatial brightness amplification" has been used also for analysis of output light fields of the coherent correlation processor.

Acknowledgments

This work has been supported by the Foundation for Fundamental Research, Republic of Belarus.

References

[1] Rozanov N N, Semenov V E and Khodova G V 1982 *Sov. J. Quant. Electr. (USA)* **12** 193–7; transl. of: 1982 *Kvant. Elektronika (USSR)* **9** 354–60

[2] Apanasevich S P and Lyakhnovich A V 1988 *Phys. Stat. Solidi (b)* **150** 507–12;
Sinitsyn G V, Apanasevich S P, Karpushko F V and Lyakhnovich A V 1988 *Bull. Acad. Sci. USSR Phys. Ser. (USA)* **52** 135–8; transl. of: 1988 *Izv. Akad. Nauk SSSR Ser. Fiz. (USSR)* **52** 369–73

[3] Apanasevich S P, Karpushko F V and Sinitsyn G V 1988 *Laser Optics of Condensed Matter, Proc. 3 USA–USSR Symposium* ed J L Birman *et al* (New York: Plenum Press) pp 475–80

[4] Sinitsyn G V, Apanasevich S P, Lyakhnovich A V, Karpushko F V 1993 *Optical Computing, Proc. SPIE* vol 1806 ed A M Goncharenko *et al* (Washington: SPIE — The International Society for Optical Engineering) pp 559–65

Inst. Phys. Conf. Ser. No 139: Part III
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical heterodyning noise-protected with stochastic resonance

M I Dykman

Department of Physics, Stanford University, Stanford, CA 94305, USA

G P Golubev, I Kh Kaufman, D G Luchinsky, E A Zhukov

VNIIMS, Andreevskaya nab. 2, Moscow, 117965, Russia

P V E McClintock, N D Stein, N G Stocks[1]

School of Physics and Materials, Lancaster University, Lancaster, LA1 4YB, UK

Abstract. A new form of heterodyning in an all-optical bistable system, related to stochastic resonance, is reported. It is shown theoretically and experimentally that the heterodyne signal and signal-to-noise ratio can be enhanced by the addition of noise to the input signal.

1. Introduction

Prospective applications of optical bistability (OB) in digital optical computing and signal processing are related to the development of miniaturised devices with reduced threshold power [1]. This implies that noise is one of the important limiting factors (see, for example, [1]) and highlights the problem of controlling the signal-to-noise ratio (SNR) in OB devices. It is thus of particular interest to investigate in OB systems phenomena related to stochastic resonance SR [2] in which the signal amplitude and SNR can be *increased* rather then suppressed by adding noise to the system.

Most studies of SR to date (including those for OB systems [3-5]) have related to signals of low frequency [6], i.e less than the reciprocal relaxation time of the system. It was shown recently [7], however, that in the practically important case where two *high*-frequency signals are mixed nonlinearly, SR can be used to *strongly* increase the amplitude of the resulting heterodyne signal and also to enhace its SNR.

In this paper we extend the idea to obtain a new form of optical heterodyning noise-protected with stochastic resonance. In Sec.2 a theory of noise-enhanced optical heterodyning for the case of noise on the intensity of the input signal is outlined (the case for noise on the reference beam will be considered elsewhere). In Sec. 3 experimental results demonstrating an increase of the heterodyne signal and SNR with increasing noise intensity are presented. Sec.4 contains concluding remarks.

[1]Present address: Department of Engineering, University of Warwick, Coventry, CV4 7AL, UK.

2. Theory

The system considered is a bistable interferometer with intensity dependent phase gain ϕ driven, in addition to the resonant signal beam, by a nonresonant reference beam. The intensity I_T of the transmitted resonant radiation is related to the input intensity $I_{\rm in}$ via a periodic transmission coefficient $N(\phi)$. The phase gain is assumed to depend linearly both on the intensity of the resonant intracavity radiation and on the intensity $I_{\rm ref}$ of the reference beam.

Assuming that the intensity $I_{\rm in}$ of the resonant radiation has a modulated high-frequency component and zero-mean gaussian noise $\Delta I(t)$, whereas $I_{\rm ref}$ has an *unmodulated* high-frequency component, the dynamics of the system can be described by a Debye relaxation equation for the phase gain ϕ

$$\dot{\phi} + \frac{1}{\tau}\Delta\phi = (\bar{I}_{\rm in} + A_{\rm in}(t)\cos(\omega_0 t + \psi(t)) + \Delta I(t))\, M(\phi) + \bar{I}_{\rm ref} + A_{\rm ref}\cos\omega_0 t, \quad (1)$$

Here $\Delta\phi = \phi - \phi_0$, where ϕ_0 is the phase in the dark, $\bar{I}_{\rm ref}$ and $\bar{I}_{\rm in}$ are constant components in the intensities of the two laser beams, $A_{\rm ref}$ is the amplitude of the reference signal, and $A_{\rm in}(t)$ and $\psi(t)$ are the modulated amplitude and phase of the input signal. The periodic functions $M(\phi)$ and $N(\phi)$ characterize the nonlinear response of the system and have well-known forms for simple models (cf. [1]).

If the characteristic frequencies of modulation $\Omega \sim \dot{\psi}$ and $\dot{A}_{\rm in}/A_{\rm in}$ are low compared to τ^{-1}, and $\tau^{-1} \ll \omega_0$, the motion of the system consists of fast oscillations at frequency ω_0 superimposed on a slower (averaged over $\delta t = 2\pi/\omega_0$) motion which can be described as overdamped Brownian motion, with the coordinate $x^{(sl)}(t)$ (cf. [7,8]), in a bistable potential $U(x^{(sl)})$

$$\begin{aligned} &\dot{x}^{(sl)} + U'(x^{(sl)}) = A\left[\partial F/\partial x^{(sl)}\right]\sin\psi(t) + \Delta I(t), \quad A = A_{\rm ref}A_{\rm in}(t)/2\omega_0 \\ &x = \textstyle\int_0^{\phi(x)} d\phi M^{-1}(\phi); \quad U(x) = -\bar{I}_{\rm in}x + \int_0^{\phi(x)} d\phi M^{-2}(\phi)\left(\frac{1}{\tau}(\phi-\phi_0) - \bar{I}_{\rm ref}\right). \end{aligned} \quad (2)$$

Here $F(x) \equiv M^{-1}(\phi(x))$. For small modulation amplitude $A_{\rm in}/\omega_0, A_{\rm ref}/\omega_0 \ll 1$, heterodyning can be characterized by the amplitude of the low-frequency signal $x^{(sl)}(t)$ for sinusoidal modulation, $A_{\rm in} = {\rm const}$, $\dot{\psi}(t) = \Omega$, and standard linear response theory can be applied to the analysis of Eqs. (1)-(2). In doing so one has to find the susceptibility $\chi(\Omega)$ for the averaged intensity of transmitted radiation

$$\langle I_T(t)\rangle = \sum_{n=1,2} w_n \bar{I}_{Tn} + {\rm Im}\left[\chi(\Omega) A \exp(-i\Omega t)\right],$$

where w_n are the stationary populations of the states. For low noise intensity D the susceptibility is given by the sum of partial contributions from vibrations about the stable states $\bar{I}_{T_n} = \bar{I}_{\rm in}N(\phi(x_n))$, weighted by w_n, and from the interwell transitions

$$\chi(\Omega) = \sum_{n=1,2} w_n \left(N'(\phi_n) + M'(\phi_n)\frac{\partial \bar{I}_{Tn}}{\partial \bar{I}_{\rm ref}}\right) + \quad (3)$$

$$\frac{w_1 w_2}{D}[\bar{I}_{T1} - \bar{I}_{T2}][M^{-1}(\phi_1) - M^{-1}(\phi_2)]\frac{W_{12} + W_{21}}{W_{12} + W_{21} - i\Omega}.$$

Here, W_{nm} is the probability of the transition $n \to m$ between the stable states for $A_{\rm in} = A_{\rm ref} = 0$. The dependence of W_{12}, W_{21} on the noise intensity D is of the activation type, and therefore the amplitude of heterodyne signal ($\propto |\chi(\Omega)|$) may increase with increasing noise intensity. The SNR for the heterodyning can be characterized by the ratio R of the low-frequency signal in the intensity of the transmitted radiation, given by $\frac{1}{4}A^2|\chi(\Omega)|^2$, to the value of the power spectrum $Q^{(0)}(\omega)$ at the same frequency for $A = 0$ which can be evaluated in a similar way to $\chi(\Omega)$. For $\omega \ll U''(\phi_n)$ it has the form

$$Q^{(0)}(\omega) = \sum_{n=1,2} w_n \frac{D}{\pi}\left[N^2(\phi_n) + M^2(\phi_n)\left(\frac{\partial \bar{I}_{Tn}}{\partial \bar{I}_{\rm ref}}\right)^2 + N(\phi_n)M(\phi_n)\frac{\partial \bar{I}_{Tn}}{\partial \bar{I}_{\rm ref}}\right] + \qquad (4)$$

$$\frac{w_1 w_2}{\pi}[\bar{I}_{T1} - \bar{I}_{T2}]^2 \frac{W_{12} + W_{21}}{(W_{12} + W_{21})^2 + \omega^2}.$$

It follows immediately from (3), (4) that, for $\Omega \ll \tau_r^{-1}$, not only the heterodyne amplitude, but also R can indeed increase with increasing noise intensity in the range of $\bar{I}_{\rm in}, \bar{I}_{\rm ref}$ for which $w_1 \approx w_2$.

3. Experiment

These predictions were first tested quantitatively by electronic analogue simulation [7] for the case of overdamped motion in the symmetric bistable Duffing potential, for $F(x) \equiv x$. *Noise-enhanced heterodyning* was thereby demonstrated for both low- and high-frequency noise, in good agreement with theory.

We now report an optical experiment based on a double-cavity membrane system (DCMS) [8] was investigated, consisting of a thin $\sim 1\mu$m GaSe semiconductor film separated from a dielectric mirror by a metal diaphragm $\sim 500\mu$m in diameter.

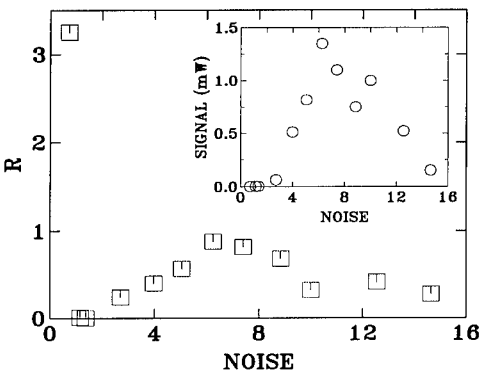


Fig.1 The signal-to-noise ratio R, also as a function of noise intensity (arbitrary units). The inset shows magnitude of the heterodyne signal for $\omega_0 = 2.1$kHz, $\Omega = 3.92$ Hz as a function of noise intensity .

The incident beam from an argon laser, of wavelength 514.5 nm, propagating along the normal to the mirror provides an input signal. An additional beam of wavelength 488 nm from the argon laser, inclined with respect to the DCMS axis, provides a reference signal. The intensities of the laser beams are modulated in time by two electro-optic modulators, to which periodic signals and noise are applied. Optical bistability arises because of thermoelastic bending of the membrane caused by the main (514.5 nm) laser beam. The phase gain of the air-resonator is linear in bending and thus follows adiabatically the thermal relaxation of the film. Heating of the DCMS by the 488 nm reference signal is directly proportional to its intensity.

The measured relaxation time τ_r of the DCMS was $\sim$2 ms. Thus to met the condition $\Omega \ll \tau_r^{-1} \ll \omega_0$ the reference signal (488 nm) was modulated periodically at a frequency

$\omega_0 = 2.1$kHz,while the input signal (514.5 nm) was modulated at $\omega_0 + \Omega$, with $\Omega = 3.92$ Hz and in addition by noise with a cutoff frequency of 5 kHz. A heterodyne signal at frequency $\Omega = 3.92$ Hz was detected in the transmitted light intensity I_T at wavelength 514.5 nm. With noise-induced fluctuationaltransitions occurring between two stable states of the DCMS, a strong heterodyne signal appeared, superimposed on the zero-frequency Lorenzian peak in the spectral density of fluctuations of I_T [8].

Strong enhancements of both the heterodyne signal (by a factor of $\sim$1000) and of the signal-to-noise ratio R were observed with increasing noise intensity, as shown in Fig.1. The dependence of R on the noise intensity is of the characteristic reversed-N shape familiar from earlier studies of SR in bistable systems [6] and consistent with the theory outlined above. The enhancement of the SNR occurs within a restricted range of noise intensities, as expected, and the ratio between the value of R at the minimum to that at the local maximum (i.e., the maximum noise-induced "amplification" of the signal-to-noise ratio) is $\sim$10.

4. Conclusion

In conclusion, we have predicted, and observed experimentally in an all-optical system, the new and potentially useful phenomenon of noise-enhanced optical heterodyning. Similar effects can be mediated by the thermal noise in the device. A fuller discussion and a detailed quantitative comparison with the theory will be given elsewhere.

One of us (DGL) would like to acknowledge the warm hospitality of Lancaster University. The work was supported by the Science and Engineering Research Council, by the European Community, by the Royal Society (London), and by the Ministry of Education and Science of Russia.

References

1 Gibbs H 1985 *Optical Bistability: Controlling Light with Light* (New York: Academic).

2 Benzi R, Sutera A and Vulpiani A 1982 *J. Phys. A* **14** L453-7
Nicolis C 1982 *Tellus* **34**, 1-9
Benzi R, Parisi G, Sutera A and Vulpiani A 1982 *Tellus* **34** 10-6

3 Special issue 1993 of *J. Stat. Phys.* **70**, ns.1/2 and references therein

4 McNamara B, Wiesenfeld K and Roy R 1988 *Phys. Rev. Lett.* **25** 2626-9

5 Dykman M I, Golubev G P, Luchinsky D G, Velikovich A L and Tsuprikov S V 1991 *JETP Lett.* **53** 193-7

6 Gross J, Apanasevich S, Jung P, Issler H, Burak D and Klingshirn C 1994 *Phys. Rev. A* **49** 2199-202 and references therein

7 Dykman M I, Golubev G P, Luchinsky D G, McClintock P V E, Stein N D and Stocks N G 1994 *Phys. Rev. E* **49** 1935-42

8 Dykman M I, Golubev G P, Luchinsky D G, Velikovich A L and Tsuprikov S V 1991 *Phys. Rev. A* **44** 2439-49

IV. NEURAL NETWORKS

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical Implementations of a Stochastic Neural System

WA Crossland‡, TJ Hall††, JS Shawe-Taylor† and M van Daalen†

‡Department of Engineering, Cambridge University, Cambridge, CB2 1P

††Department of Electronic and Electrical Engineering, King's College, University of London, Strand, London WC2R 2LS

†Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX

Abstract. The paper addresses the design of an opto-electronic implementation of a stochastic bit-stream neural system which operates by manipulating digital bit streams to create emergent activation functions using extremely simple logic.

1. Introduction

This paper investigates the possible strategies for implementing a bit stream stochastic neural network (BSN) design [6] making use of the latest opto-electronic technologies [3, 4]. The network design operates by manipulating digital bit streams to create an emergent activation function using extremely simple logic. The neural design and implementation strategy will potentially yield a number of impressive benefits.

- The optical connections have a very high bandwidth which can overcome the I/O bottleneck in VLSI implementations of the stochastic neural design [5].
- The stochastic approach introduces real values through a precisely controlled probabilistic technique, which makes possible a complete and exact mathematical description and simulation of the network functionality.
- In contrast to analogue implementations, digital networks can be combined without introducing further uncertainties in the accuracy of the computation. Hence implementations can be scaled up without major modifications.
- Both the speed and the digital nature of the basic operations will mean that effective on-chip learning may be incorporated into the design, by exploiting the stochastic properties of the network operation.

2. Bit stream neural design

A standard neural design involves a network of neurons each calculating the weighted sum of its inputs and then applying a sigmoid-like function to the result. Hence, one of the fundamental problems inherent in a massively parallel implementation of a neural architecture is how to multiply together the real inputs and their corresponding weights without resorting to cumbersome bit-parallel digital circuitry. The analogue solution to this problem incurs a number of difficulties, including relatively low resolution, cross-chip variations in component performance, and the general problems in constructing large-scale reliable systems.

The stochastic bit-stream approach combines the benefits of digital circuitry and analog simplicity. In this approach real values are represented by stochastic bit streams. Such a bit stream is a sequence of 0's and 1's with a fixed probability p of a 1 occurring. They are used to represent the real value p in the interval $[0, 1]$ for unsigned values, or in an alternative representation the value $(2p - 1)$ in the interval $[-1, 1]$ for signed values [2]. Each stochastic input bit to a neuron is weighted by either ANDing, when operating on unsigned values, or XORing, when operating on signed values, with a corresponding weight bit. One bit of each of the individual weighted input streams is summed and compared with a threshold value to determine one bit of the output stream. The interaction of the probability distributions of the sum and the thresholds creates a sigmoid-like functionality, which can be precisely described [5]. The design of a single neuron is therefore extremely simple, making it possible to map a large network onto an established implementation technology.

The functionality of the neuron has been demonstrated through various simulations. A handwritten digit recognition task has been chosen to test the learning and generalization. A clean set of 500 16×16 pixel samples chosen from the NIST database were used. The fully connected feed-forward network has 256 input neurons, 12 hidden neurons and 10 output neurons. It has been successfully trained on 400 input patterns. This was achieved with a network structure of comparable complexity to a relatively compact sigmoid network for the same problem. The error (the vector distance between outputs) on the training sample was less than 0.1 for each example, though this was larger than that achieved with the sigmoid network. The generalization error for the network was slightly larger than for the sigmoid network, which appears to contradict the conjecture that the functionality of a BSN network is more restricted.

Simulations of BSN networks have also demonstrated the application of the Mean Field Annealing algorithm to the problem of graph bisection. The quality of solution and number of iterations is comparable with standard neurons. Massively parallel BSN implementations could process large graphs in real time.

3. The Two Optical Architectures

In this paper we will present two opto-electronic implementations of the neural architecture. They are illustrated in figure 1. One design is fully spatially multiplexed (maximally parallel), whilst the other introduces the extra dimemsion of time multiplexing. A neuron in a time multiplexed system receives the first bit from each connection in turn and only delivers an output bit when all have been processed. This considerably simplifies the design of the neurons, but at the expense of their operational speed. It also simplifies the beamlet optics for a given number of neurons when compared to the fully spatially multiplexed design.

3.1. Overview of optical components

Both of the optical implementations require image replication and beam handling optics. Image replication has been demonstrated [4] using computer generated holograms in conjunction with bulk lens systems to produce output arrays consisting of up to 1000×1000 beamlets. Beam detection is achieved by the use of silicon/GaAs photodetectors, and modulation is possible through the use of various opto-electronic devices. Modulation is most easily achieved with liquid crystal spatial light modulators, but only relatively slow speeds in the region of 10 MHz may be realised. High speed modulation can be achieved

by the use of devices such as LED based photothyristors [1] featuring switch off times less than 10ns, or arrays of individually addressable surface emitting semiconductor lasers [3] operating at speeds of up to 1 GHz. The choice of output modulator will depend on the architecture used. Those with a high level of internal multiplexing will have relatively low output bit rates, these may be achievable with SLMs. Non-multiplexed designs with high output bit rates ($\approx$ 100MHz) will require the use of fast output modulators.

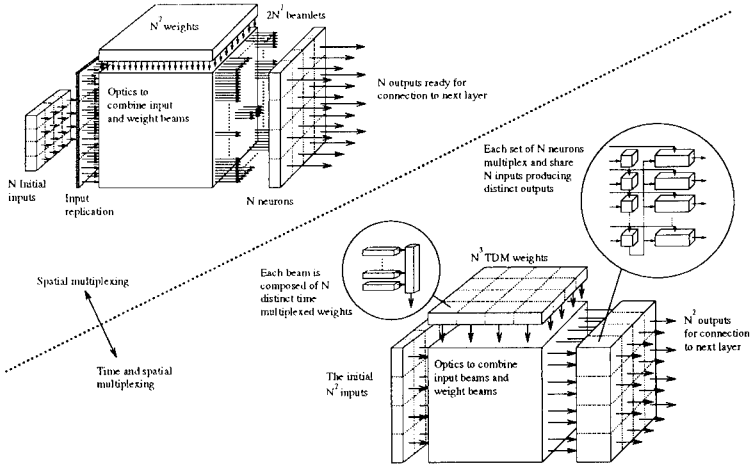


Figure 1. *The spatial/time multiplexing scale of implementations*

3.2. Detail of spatial architecture

In the fully spatial implementation each neuron receives all of its inputs values and their corresponding weights in parallel. On a given operational cycle a neuron will receive one bit from each weight and input connection. These input/weight pairs are first multiplied, the choice of AND or XOR will depend on the actual stochastic representation being used. The resulting weighted input bits must now be accumulated, in parallel, prior to thresholding and the generation of a stochastic output bit. Accumulation may be achieved in several ways.

The first approach is to focus the n weighted bits onto a single analogue detector that produces an analogue output proportional to the number of 1's. Current analogue detectors of this type can resolve values for n of 64 whilst operating at speeds around 10MHz. A second purely digital approach would involve the use of fast adders incorporating carry save and look-ahead techniques. This approach is not ideal as relatively large amounts of circuitry would be required, which in turn will impact performance. The final approach is based around a novel population detector that directly exploits the stochastic nature of its inputs and produces an appropriately thresholded stochastic output. At the time of writing the details of this device cannot be disclosed pending a patent application.

3.3. Detail of time multiplexing architecture

In this approach a group of n input bit streams is transmitted to each of n neurons that together form a layer. Such a fully connected layer will require a total of n^2 distinct weight bit streams. The example illustrated in figure 1 shows n such layers operating in

parallel. Here we effectively have n^2 inputs, n^3 weights and n^2 outputs. Each neuron has two physical connections designed to accept time multiplexed input and weight bit streams. Full connectivity within a layer is achieved by connecting together all of the physical input connections of each of the constituent neurons, which is then supplied with the time multiplexed input bits. The corresponding weights are separately time multiplexed and supplied directly to each neuron. These two multiplexing processes are depicted in the circled areas in figure 1. Once the neurons have completed an operational cycle, i.e. they have received n input/weight bit pairs and have performed the appropriate computations, an output bit is generated and can be passed to the next layer. The critical components in this approach are the shift registers used to multiplex both the input bits and the weight bits. These subsystems must clock at n times the bit rate of the input and output bit streams. Typical clock speeds will be of the order of 100MHz. Each neuron is composed of an AND or XOR gate to perform the input/weight multiplication and a simple counter to compute the sum and activation, see [6] for details.

4. Conclusions

One of the attractions of the stochastic design is its adaptability to hardware constraints. This means that there is a great deal of flexibility in the choice of hardware configurations. At the cheapest end of the spectrum, simple FPGAs can be used to provide prototyping systems to verify design methodologies for larger systems. At the other extreme, this paper has indicated how the latest opto-electronic techniques can be utilised to give very high throughput speeds in a massively parallel mode.

As an example of the flexibility, we mention that a neuron can easily generate a bitstream for the derivative of its output with respect to a particular input. This opens the way to implementing a parallel network to calculate the backpropagation of error through the main network and hence perform gradient descent learning on chip.

The other very important property of the design is its modular capability. Since the functionality can be precisely described, there is no sense in which uncertainty or error accumulates through layers of the network. This contrasts with analogue solutions, where there is a need to train out errors in larger networks. Hence, for the bit stream design there is no obstacle to concatenating several subnetworks into a larger system and backpropagating through one network in order to train another. Recent work by Werbos [7] suggest that these more complex structures may be the key to realising the full potential of Neural Networks in Neurocontrol applications.

References

[1] P. Heremans *et al*, 'Image transcription between arrays of N-p-n-P optoelectronic switches', IEEE Transactions on Electron Devices, 39(10), 1992, pp. 2248-2253

[2] P. Jeavons *et al*, 'Generating Binary Sequences for Stochastic Computing', to appear in IEEE Transactions on Information Theory

[3] J. L. Jewell *et al*, 'Vertical-cavity surface-emitting lasers: design growth, fabrication, characterization', IEEE J. Quantum Electronics, 27, 1991, pp. 1332-1346

[4] A.G. Kirm *et al*, 'Experimental implementation of an optoelectronic matrix-matrix multiplier which incorporates holographic multiple imaging', Optical Computing and Processing, 2, 92, 293–304

[5] J. Shawe-Taylor *et al*, 'Probabilistic Bit Stream Neural Chip: Theory', Connection Science, 3, 91, 317–328

[6] M. van Daalen *et al*, 'A Stochastic Neural Architecture that Exploits Dynamically reconfigurable FPGAs', IEEE Workshop on FPGAs for Custom Computing Machines, pp.202–211, April 93

[7] P. J. Werbos, 'Neurocontrol: where it is going and why it is crucial', Artificial Neural Networks 2, Proceedings of 92 ICANN Conference, 1992, pp.61–68

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

An opto-electronic neural network with on-line learning

D. Pignon, T.J. Hall, L.Q. Xu, K.B. Russ*, S.R. Cherry*, N.C. Roberts**, P. Prewett*** and P. Hallowell***

Department of Physics, King's College London, U.K.
*Pilkington Group Research, U.K.
**British Aerospace, Sowerby Research Centre, U.K.
***Rutherford Appleton Laboratories, U.K.

Abstract: We present an opto-electronic hardware system that performs matrix-matrix multiplication using an electrically addressable spatial light modulator as the switching divice and computer generated holograms for connection routing. This system is well suited for the implementation of a quantised neural network. The operation of two demonstrator systems is described as well as a new algorithm for on-line learning with quantised weights and hard thresholding based on function smoothing and stocastic approximation. Finally, some results of test runs are presented.

1. Introduction

The system described here implements a three-layer perceptron. Restrictions imposed by the binary nature of the Spatial Light Modulator (SLM) cause the quantisation of weights available and force hard thresholding neurons. Thus standard gradient descent algorithms such as Back-Propagation of Error (BPE) do not work and other solutions to learning had to be found.

The first is to run a computer simulation of the hardware using continuous variables in conjunction with the hardware itself.

The second method was the use of the Stocastic Approximation with function Smoothing (SAS) algorithm [1], futher developed at King's College specifically for this problem [2]. The SAS algorithm allows direct on-line learning.

2. Demonstration systems

Both demonstrators perform the same matrix-matrix multiplication operation,

$$C = W \cdot A$$

where A is the input pattern to layer n of the network, W is the weight matrix, C is the unthresholded output and · represents the inner product operation. The control computer then applies the thresholding function F completing the pass from layer n of the network to layer n+1,

$$O_{n+1} = F(W_n \cdot A_n)$$

where O_{n+1} is the output from layer n+1. Next we set $A_{n+1} = O_{n+1}$ and repeat the process through the same two layer hardware to obtain O_{n+2}. In this way time multiplexing is used to construct a network of as many layers as desired, in our case three.

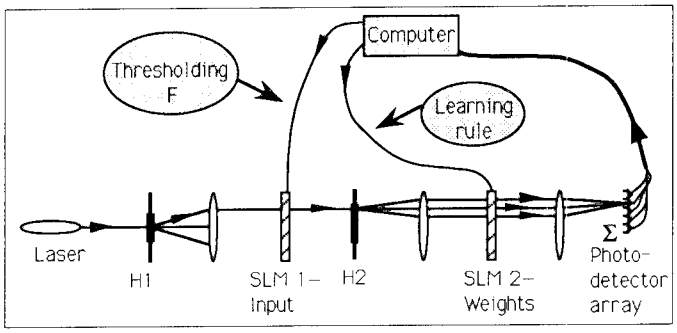


Figure 1: Transmissive demonstrator. H1 & H2 are computer generated holograms.

2.1 Transmissive system

The first demonstrator to be built was the transmissive version (figure 1) [4][5]. An expanded laser beam is fanned out into 8×8 spots by hologram H1 which then proceed on to SLM1 where the input pattern A is encoded as a binary intensity matrix. Hologram H2 further multiply images the input matrix 5×5 times onto SLM2 where each pixel of the input matrix is given a weighting factor for each 'neuron'; the 25 neurons being represented by the 25 blocks on SLM2. The summation for each neuron is performed optically at the 5×5 detector array.

The control computer applies the thresholding function F and modifies the weight matrix according to the learning rule in use. Operation is very slow due to the use of the serial computer and the slow SLM addressing time.

The computer generated holograms [3] are phase elements etched into quartz and have proved very accurate. In order to encode the weights at SLM2, H2 actually produces a 2×2 sub-fanout of each pixel of the input matrix of differing intensities. A 'maxi-pixel' at SLM2 (consisting of 4 true SLM pixels) then selects different combinations of these beamlets to encode a weight with 16 levels.

2.2 Reflective system

The reflective demonstrator (figure 2) is the same in function as the transmissive one except folded in on itself by making hologram H2 reflective. This way the hardware requirements and system size are reduced.

This is possible because the input matrix is displayed on the central area of the SLM whereupon H2 reflects and multiply images onto the surrounding area on the return leg. However, this arrangement now becomes much more sensitive to the polarisation states produced by the SLM and we find that it is much more difficult to get a good contrast ratio at the output due to the SLM producing elliptically polarised light. In addition, reflections, which were not a problem in the transmissive system, are now mixed in with the desired signal and have proved very troublesome.

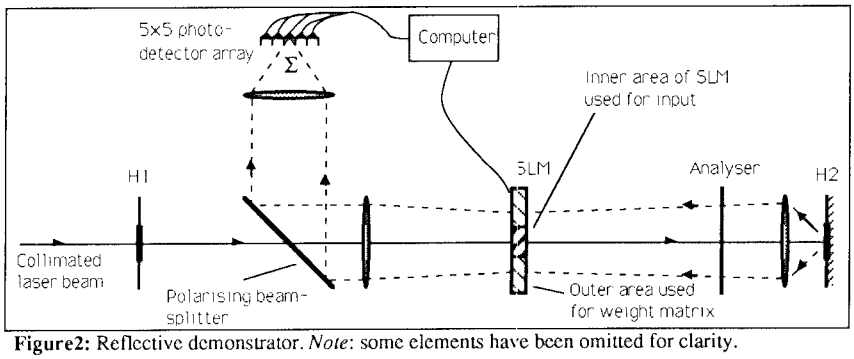


Figure2: Reflective demonstrator. *Note*: some elements have been omitted for clarity.

2.3 Neural network operation

Both demonstrators are run as a 64-24-9 perceptron, i.e. we have run tests with 9 memories of 64 bits each. A standard network requires both positive and negative weights yet an optical signal can only be positive. We circumvent this problem by defining a total intensity level as the unchanged input and calculating the negative weights for each neuron (that has positive weight connections) thus:

Negative=Total -Positive.

3. On-line learning methods

Quantised weights and hard thresholds produce a sampled discontinuous and hence non-differentiable error (cost) function in weight space which make it unsuitable for minimisation using standard back propagation of error. Some way round this has to be found in order to proceed.

3.1. Back propagation of error using a software simulation

If a continuous weight, smooth threshold simulation of the network is run in software, it can be used to predict the approximate hidden and output layer states using standard BPE with a quantisation stage at the end. The hardware weights may then be adjusted according to a simple local rule to produce the required states [5]. But this is 'cheating'!

3.2 Stochastic approximation and function smoothing

When the sampled discontinuous cost function is convolved with a Gaussian probability function G(P,β) it is transformed into a smooth differentiable form on which a gradient descent method may be used to find the minimum (figure 3). In addition to actually allowing a minimisation to take place, the gradient descent algorithm can avoid local minima if the smoothing coefficient β is started off at a suitably high value.

This smoothes the cost function to such an extent that all the local minima are ironed out and the lowest point of the function will be in the vicinity of the global minimum (top trace in figure 3). As β is reduced during the search, the smoothed function gradually becomes less smooth and approaches the original cost function while the solution settles in the true global minimum (botom trace in figure 3).

Stochastic approximation implements a technique to estimate the gradient of the cost function using only a few random samples rather than an exhaustive calculation. The expectation of the search trajectory, however, corresponds to a direct search. This reduces the number of samples of the function required, and thus the number of passes through the hardware per learning iteration, without compromising accuracy.

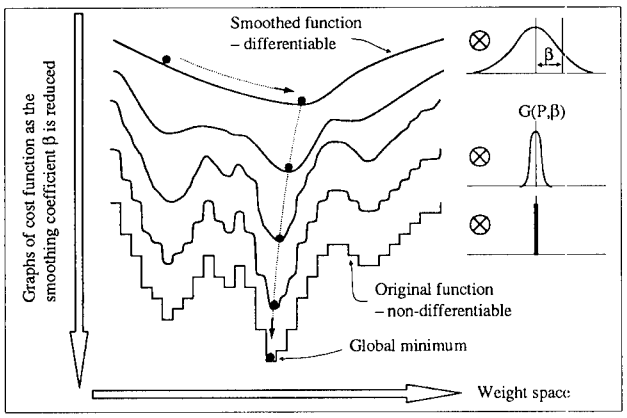


Figure 4: Convolution with a smoothing function G turns the cost function into a differentiable form. By starting off with a high β value and then gradually decreasing it, a gradient descent search will not get stuck in a local minimum.

4. Experimental results

The net was trained to recognise 9 patterns consisting of pixellated letters of the alphabet with various levels of noise added (to improve generalisation). Testing was then carried out with the same memories with added noise of up to 30%.

The software simulation method produced correct classification of near 100% with up to 10% input noise, dropping to 60% with 30% noise.

The SAS algorithm proved difficult to implement due to the slow hardware cycle. The drift of the components (due to temperature variations) was faster than the learning time so the hardware characteristics changed too fast for the learning procedure to adapt to them! Nevertheless, limited success was achieved with the error reduced to 33%, and software simulations of the hardware worked perfectly.

5. Conclusion

The application of a quantised optical matrix-matrix multiplier system to a neural network implementation has been demonstrated. A novel learning scheme has been used which gets over the problem of quantised weights and hard thresholds imposed by the binary nature of the hardware. In the course of learning, the network also learns the imperfections in the hardware enabling a high degree of robustness.

The hardware is at present too slow for on-line learning but it has been demonstrated that it does work and with improvements in speed the SAS algorithm would work well.

Computer generated holograms have been used for beam routing and intensity coding and have proved very reliable. Other advances in technology such as smart-pixel directly addressable SLMs and an SAS adapted to 'on-chip' learning should allow much faster wholly parallel systems to be developed. This, together with miniaturisation along the lines of the reflective system with the use of microlens arrays, laser diode arrays, smaller and faster detector arrays, and smaller-pixel SLMs might produce systems for real time direct image processing applications.

Grey-level input can be implemented in the same way that the weights are coded on the SLM 'maxi-pixels', allowing a greater information content in the input pattern and relaxing the hard-threshold constraint.

References

[1] Styblinski M A and Tang T S, "Experiments in nonconvex optimisation: stochastic approximation with function smoothing and simulated annealing," *Neural Networks, Vol. 3, 467-483,* 1990.

[2] Xu L Q and Hall T J, "From coarse to fine: A novel way to train neural networks," *Proceedings of ICANN' 94, 4th Annual Conference of ENNS,* Naples, Italy, May 1994.

[3] Kirk A G, "Diffractive Optical Elements within Parallel Optoelectronic Systems", *PhD thesis 1992,* King's College London.

[4] Hall T J, Xu L Q, Jamieson S, Dainty P, Russ K B, Cherry S R, Roberts N C and Walton A L, "Free-space optical networks using diffractive and refractive micro-optics," *4th International Conference on Holographic Systems, Components and Applications,* University of Neuchatel, Switzerland, September 1993.

[5] Russ K B and Cherry S R, "HIPPS Joint Demonstrator," *DTI LINK 'HIPPS' project final report (Ref. HIP/TR/BAE/NR931222), pages 111-121,* December 1993.

Acknowledgements

The work of D Pignon is supported by an EPSRC CASE award co-sponsored by GEC-Marconi.

The systems described in this summary were developed as part of the "Highly Interconnected Parallel Processing Systems" project under the DTI LINK scheme (Ref. HIP/TR/BAE/NR931222); co-ordinator N. C. Roberts.

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical Lateral Inhibition Networks using Self-Linearised SEED's

Paul Horan

Hitachi Dublin Laboratory, O'Reilly Institute, Trinity College, Dublin 2, Ireland.
Email <paul.horan@hdl.ie>

Abstract. A family of lateral inhibition architectures which use the self-linearised SEED effect to implement optical subtraction are described, and their operation demonstrated in simulation.

1. Introduction.

Lateral inhibition networks, where nodes within a processing layer inhibit one another, form a very important class of networks. They are not amenable to processing on conventional serial machines, since they require some type of iterative self-consistent solution, thus making parallel implementation very attractive. The dense, recurrent nature of the interconnection is ideally suited to an optical approach, and the finite range of the interconnection means there is no intrinsic limit on scaling up the system.

The essence of any inhibition network is that the activity level of an individual node must decrease in response to increasing input from neighbouring nodes *i.e.* subtraction, which is difficult to do optically. One approach has been to use an optically coupled *pnpn* light emitter and an *npn* phototransistor.[1] However, the Self Linearised SEED effect, observed in a serial photodiode-SEED modulator combination offers an alternative method of optical subtraction.[2,3] Feedback provided by the common current causes the modulator reflectivity to decrease as the light falling on the photodiode increases. The small range of modulation can be improved by incorporating the SEED device in a resonant Fabry-Perot optical cavity.[4] This combination is made more attractive when we consider that the SEED is, itself, a *pin* diode, allowing the possibility that the detector and modulator can be made together in the same integrated process. Operating the diode within an optical cavity configuration will require that the diode to be biased, thus solving problems with quantum well diode voltage independence. [3] However, the requirement that the modulator and detector photocurrents are similar implies that the output is a small fraction of the input, thus limiting cascaded operation. Having

established that we have a good optical subtraction technology, we will go on to investigate a family of architectures based on inhibition.

2. Simple lateral inhibition.

2.1 Single stage inhibition

A simple lateral inhibition optical system is shown in figure 1a in a one-dimensional format, although this can be extended to a two dimensional computing surface. Each node consists of a detector-modulator pair, and receives an external input which is incident on the modulator. A portion of the reflected signal is redirected to the detectors of neighbouring nodes in some manner, which defines the lateral interconnection function $g_{(x)}$. The node does not inhibit itself. The optical element is not specified and may be reflective, refractive, diffractive, or holographic. Thus, the reflectivity of a given modulator is dependent on the total inhibitory signal received from neighbours, and decreases linearly in response to it. A point to note is that the internal state of the node, which we identify with the modulator reflectivity, is determined purely by the neighbours, and is not dependent on the input to that node, although the final output is. This single pass geometry effectively results in a convolution of the input with the lateral interconnection function. Figure 1b shows the results of a typical simulation with flat lateral inhibition $g_{(x)}$ over ±3 nodes. A periodic boundary condition is imposed to avoid edge effects. This network performs a simple edge-enhancement function. The response of the system reflects the essentially linear nature of the device. While still linear, the state of the node can be made positively dependent on the input, as well as inhibited by neighbouring nodes, by using a differential pair of detectors, following the example of Miller.[3]

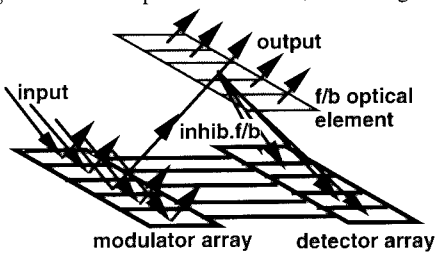


Figure 1a Simple lateral inhibition system

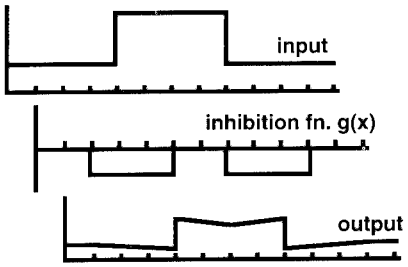


Figure 1b Input, lateral inhibition function $g_{(x)}$ and the output.

2.2 Single stage inhibition with input subtraction

From figure 1b it can be seen that the result of a single stage inhibition is basically a reduced scale version of the original input with some small perturbations representing the processing. This suggests that if a suitably scaled version of the original input is subtracted from the output, then the processing can be enhanced. Since this involves subtraction a second self-linearised

modulator detector pair can be used as shown in figure 2. A fraction of the initial input is taken off and directed, on a one to one basis, to the detectors of a second detector/modulator array. The output of the lateral inhibition processing is sent to the modulators of the "subtraction" array, effectively subtracting the input and giving an enhanced, if small, output. This could form the input to a further processing stage, such as the winner-take-all layer described in the next section.

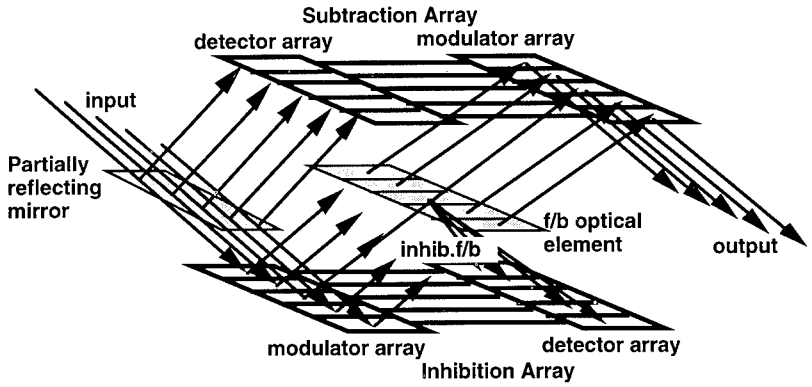


Figure 2 Simple lateral inhibition with input subtraction.

3. Lateral inhibition with feedback

By introducing a nonlinear, recurrent feedback of the output to the input winner-take-all (W-T-A) behaviour can arise. This can be done by using a second modulator array for input, as shown in figure 3a. The modulator array is uniformly illuminated and the external inputs presented. The modulated beams are then directed to the laterally connected layer as before. However, if the output is monitored and used to control the input modulator array a recurrent loop is established. All that is required is that the input is presented for a sufficient time for the feedback loop to be established, thereafter the input is removed and the system converges to a stable state. Again, this is a one-dimensional example of a possibly two-dimensional surface.

Cheng & Wan have shown that there must be gain, however small, in the feedback loop; if not, all signals will decay to zero.[5] More importantly, from an optical point of view, they have examined the situation of limited lateral inhibition. Global interconnection, leading to a single winning node, is not feasible in an optical system, as finite optical power can only be distributed to a limited neighbourhood. They have shown that in a finite interconnection network local winners will arise, one per neighbourhood as defined by the range of interconnection.[5] A point not stressed by Cheng & Wan is that each cell must have a winner, even if the input in that area is perfectly uniform. Small spurious variations arise during the iteration process, which become amplified, eventually giving a "winner". Figure 3b shows an example of an input and the result that the simulation converged upon. As can be seen, local

maxima are located, but a spurious peak also occurs (around node 22). Such a system could form the heart of a self-organising network, when prefixed by a Hebbian network.[6]

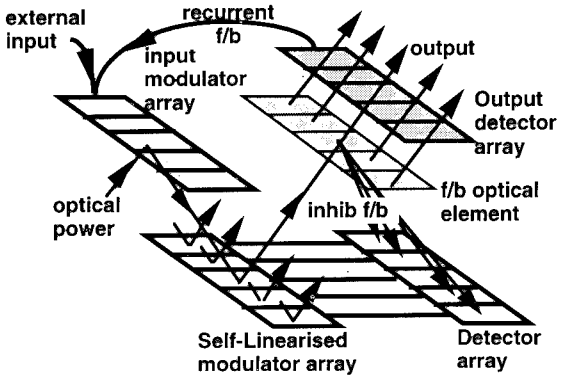


Figure 3a Lateral inhibition with feedback

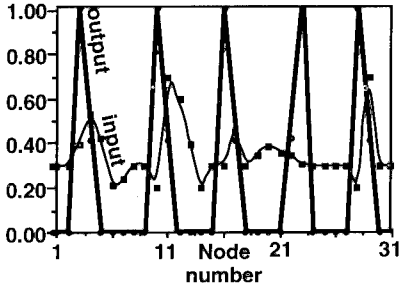


Figure 3b Simulation of local winner-take-all behaviour.

The problem of spurious peaks arising from positive gain in the feedback could be ameliorated by introducing a nonlinear feedback function. For example, if a sigmoidal gain function were used, signals less than the point of inflection value would experience loss and so be damped, only signals larger than the threshold set by the inflection value would grow.

A further nonlinearity is introduced in shunting networks where the lateral inhibition signal is modulated (multiplied) by the value of the node. [7] In the context of this work an optical implementation can be envisaged whereby the inhibition optical signal is reflected from the node modulator before detection. This would necessitate a more complex optical arrangement. It should be noted that this would not correspond directly to the conventional shunting network as, in this case, the inhibition signal is being multiplied by the modulator reflectivity rather than the single state variable generally used.

References.

[1] Nakano Y, Ikeda M, Kawakami W and Kitayama K, 1991 *Appl. Phys. Lett.* **58** 1698-1700

[2] Miller D, Chemla D, Damen T, Wood T, Burrus C, Gossard A, and Wiegmann W 1985 *IEEE J. Quant. Electron* . **QE-21** 1462-1475

[3] Miller D 1993 *IEEE J . Quant. Electron* . **29** 678-698

[4] Shoop B, Pezeshki B, Goodman J and Harris J 1992 *Opt . Lett* . **17** 58-60

[5] Cheng Y and Wan Z 1992 *Proc. of the Intl. Joint Conf. on Neural Networks, Baltimore, Maryland, 7-11 June 1992* (IEEE 92 CH 33114-6) p III-456 - III-460

[6] Kohonen T 1990 *Proc. IEEE* **78** 1464-1480

[7] Grossberg S 1988 *Neural Nets* . **1** 17-61

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical spatial/frequency filtering and adaptive neural networks based on fractional Fourier transforms

Soo-Young Lee

Korea Advanced Institute of Science and Technology
Department of Electrical Engineering
373-1 Kusong-dong, Yusong-gu, Taejon 305-701, Korea (South)

Abstract: Based on fractional Fourier transforms a new optical architecture is analyzed for filtering both in spatial and frequency domains, and is further extended to adaptive neural networks. Both error back-propagation and genetic algorithm are applicable for the training of the neural networks. The developed single-layer neural network is suitable for large-scale optical implementations, and may be utilized as a feature extractor or classifier. Complicated multi-layer neural networks may incorporate several of these layers in serial and/or parallel architectures.

1. Introduction

Recently Fourier transforms of fractional order had been developed.[1,2] Unlike the ordinary Fourier transform, i.e. the fractional Fourier transform with order 1, the fractional Fourier transforms extract features which combine both spatial and frequency characteristics of the original images. One may easily expect that the fractional Fourier transformed image is something between the original image and full Fourier transformed image. Also simple optical architectures with only single lens or double lenses with proper spacing were reported.[2]

It is well known that a Vander Lugt correlator can perform 2-D correlation and shift-invariant filtering based on Fourier transformed frequency information. On the other hand neural networks has been widely utilized for position-dependent classification and associative memory.[3] The shift-invariant correlation may be understood as characteristics of feed-forward neural networks with Toeplitz interconnection matrix, which shares appropriate interconnection weights. However this weight-sharing greatly reduces network complexity to result in limited storage capacity. Instead several neural network models incorporate *local* shared-interconnections to provide *local* shift-invariance with good classification performance.[4-6]

In this paper we extend the Vander Lugt correlator to incorporate both the shift-invariant (frequency) and position-dependent filtering based on 2

fractional Fourier transforms, and develope analogy between this architecture and neural networks. By virtue of the simple optical architectures for the fractional Fourier transforms, the neural networks is easy to implement in large-scale.

2. Optical filters based on fractional Fourier transforms

We substitute the ordinary Fourier transforms in the *4-f* Vander Lugt filter into fractional Fourier transforms with orders p_1 and p_2. Following Ref.2, we adopt an explicit integral formula for the fractional Fourier transform, i.e.

$$U_p(k) = F^p(u(x)) = \int u(x)\ \exp\left[\frac{i\pi}{\lambda f_o \tan\phi}(x^2+k^2)\right]\ \exp\left[-\frac{2\pi i}{\lambda f_o \sin\phi}xk\right]\ dx\ , \qquad (1)$$

and simple one thin-lens architecture with lens focal length $f = f_o/\sin\phi$ and distances $Z = f_o \tan(\phi/2)$, where $\phi = p\pi/2$, λ is the wavelength, and f_o is an arbitrary constant. The infinite domain of the integral is understood and suppressed throughout this paper. Within this approximation the input-output mapping relationship of the p_1-p_2 filter is now derived as

$$v(z) = \int U_{p_1}(k)H(k)\ \exp\left[\frac{i\pi}{\lambda f_o \tan\phi_2}(z^2+k^2)\right]\ \exp\left[-\frac{2\pi i}{\lambda f_o \sin\phi_2}zk\right]\ dk\ ,$$

$$= \int u(x)\ h(t)\ \exp\left[\frac{i\pi}{\lambda f_o}\left(\frac{x^2-t^2}{\tan\phi_1}+\frac{z^2-t^2}{\tan\phi_2}\right)\right]\ dx\ , \qquad (2)$$

where $u(x)$ and $v(z)$ are input and output, respectively, and $h(.)$ is the fractional Fourier transform with order $p_f = 2\phi_f/\pi$ of the filter $H(k)$ located between the two fractional Fourier transforms. The ϕ_f and t satisfy

$$\frac{1}{\tan\phi_1}+\frac{1}{\tan\phi_2}=\frac{1}{\tan\phi_f} \quad \text{and} \quad \frac{x}{\sin\phi_1}+\frac{z}{\sin\phi_2}=\frac{t}{\sin\phi_f} \qquad (3)$$

with $\phi_1 = p_1\pi/2$ and $\phi_2 = p_2\pi/2$.[8]

It is worth noticing to look at several special cases. If both p_1 and p_2 are 1 for the Vander Lugt filter, ϕ_f becomes $\pi/2$ and $t = x+z$. Now Eq.(2) becomes correlation integral. When p_1 and p_2 satisfy $p_1+p_2=2$ ($\phi_1+\phi_2=\pi$), ϕ_f again becomes $\pi/2$ and one obtains

$$v_{\phi_1}(z) = \int u(x)\ h\left(\frac{x+z}{\sin\phi_1}\right)\ \exp\left[\frac{i\pi}{\lambda f_o}\frac{x^2-z^2}{\tan\phi_1}\right]\ dx\ , \qquad (4)$$

where the $h(.)$ is now ordinary Fourier transform of the filter $H(k)$. The subscript ϕ_1 emphasizes that $v(z)$ is function of ϕ_1. Compared with the Vander Lugt filter, Eq.(4) now has scaling factor ($\sin\phi_1$) in the filter function. This scaling factor and exponential modulation term destroy shift-invariance, and provide position-dependent classificastion with slight local shift invariance. Feature extraction and/or classification of characters, images, and speech signals may be useful applications of this optical filter. In

addtion to the filter in Eq.(4), it may be possible to obtain ambiguity function and Wigner distribution function.[8]

3. Neural Networks based on Fractional Fourier Transforms

In Fig.1 the neural network analogy of the fractional Fourier transform and the spatial/frequency filters are shown, where the u_n and v_l are input and output, respectively, and U_m and V_m are corresponding fractional Fourier and inverse transforms. The fractional Fourier transform operations are now substituted by 2 complex synaptic weights, $W_{mn}^{(1)}$ and $W_{lm}^{(2)}$, and linear summation at U_m and v_l. For classification problems additional Sigmoid function $S(.)$ may be used at the output. The H_m's are filter transmissions between the two fractional Fourier transforms and adaptively trainable.

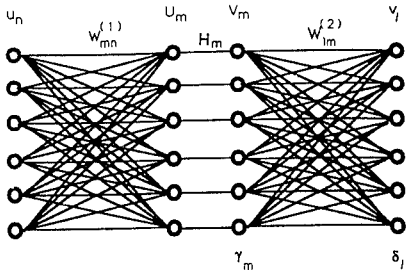


Fig.1 Neural network analogy

In mathematical notations one obtains

$$U_m = \sum_n W_{mn}^{(1)} u_n , \quad V_m = H_m U_m , \quad v_l = \sum_m W_{lm}^{(2)} V_m , \quad y_l = S(|v_l|^2), \tag{5}$$

where from Eq.(1) the fixed global interconnections become

$$W_{mn}^{(1)} = \exp\left[\frac{i\pi}{\lambda f_o \tan\phi_1}(m^2\Delta k^2 + n^2\Delta x^2) - \frac{i2\pi}{\lambda f_o \sin\phi_1}\Delta k\Delta x\ mn\right], \tag{6}$$

$$W_{lm}^{(2)} = \exp\left[\frac{i\pi}{\lambda f_o \tan\phi_2}(m^2\Delta k^2 + l^2\Delta z^2) - \frac{i2\pi}{\lambda f_o \sin\phi_2}\Delta k\Delta z\ lm\right]. \tag{7}$$

The Δx, Δz, and Δk are distances between pixels, and designed to be equal. It is worth noting that both the $W_{mn}^{(1)}$ and $W_{lm}^{(2)}$ are symmetric.

This architecture with fixed global synapses and adaptive local control gains is similar to the TAG (Training by Adaptive Gain) model [9], and the popular error back-propagation learning algorithm is still applicable. By defining a total output error as

$$E = \frac{1}{4}\sum_s \sum_l (t_l^s - y_l^s)^2 \tag{8}$$

for all stored input-output pairs $(\boldsymbol{u}^s, \boldsymbol{t}^s)$ one may train the neural networks by error back-propagation [8] or genetic algorithm. Not only the filter function H_m's but also the order of fractional Fourier transform ϕ_1 may be adapted. For the latter case with error back-propagation learning algorithm the gradient may be calculated as

$$\frac{\partial E}{\partial \phi_1} = \sum_{m,n} \frac{\partial E}{\partial W_{mn}^{(1)}} \frac{\partial W_{mn}^{(1)}}{\partial \phi_1} + \sum_{l,m} \frac{\partial E}{\partial W_{lm}^{(2)}} \frac{\partial W_{lm}^{(2)}}{\partial \phi_1}. \tag{9}$$

The adaptation of the fractional order is essential for proper feature extraction in multi-layer classifier network, where the first layer does feature extraction and subsequent layers do classification.

The single-layer neural networks in Fig.1 may be implemented by 2 thin lenses and one spatial light modulator (SLM), which is much simpler than existing optical architectures utlizing multifacet holograms, lenslet arrays with SLM, or volume holograms. Error back-propagation also requires one thin lens. Multi-layer neural networks consist of several these modules. In this case the first layer may emphasize local features by assigning higher order ($p_1 \approx 1$) for the fractional Fourier transform, while the orders may become smaller for position-dependent classification at the latter layers. The whole networks may be adapted by error back-propagation or genetic algorithm.

One problem might arise at practical optical implementations. Even for symmetric real function $h(x)$, its fractional Fourier transform $H(k)$ becomes complex function. Performance of phase-only filter may need to be investigated in this case.

For some cases including multiple feature extractions, the ambiguity function and Wigner distribution function cases, one would like to have fractional Fourier transforms with many different orders simultaneously. These can be done by lenslet arrays instead of single lens. To use same distance Z, the focal length of each small lens in the lenslets should be $f=f_o/\sin\phi = Z/(\tan(\phi/2)\sin\phi)$.

4. Conclusion

In this paper we present a new optical architecture for both spatial and frequency filtering and also an adaptive neural networks module based on fractional Fourier transforms. Complicated multi-layer neural networks may consist of several modules in serial and/or parallel architectures. With this simple architecture large-scale optical implementation of adaptive neural networks becomes feasible.

Acknowledgements: The author would like to express his thanks to Prof. Adolf Lohmann for introducing him to the subject of fractional Fourier transforms. The basic idea on this paper was initiated during his sabbatical leave at the University of Erlangen-Nürnberg as a Humboldt Research Fellow hosted by Prof. Gerd Häusler.

References

[1] Mendlovic D and Ozaktas H M 1993 J. Opt. Soc. Am. A **10** 1875-1881
[2] Lohmann A W 1993 J. Opt. Soc. Am. A **10** 2181 2186
[3] Lippmann R P 1987 IEEE ASAP Magazine **3** 4-22
[4] Fukushima K 1988 Neural Networks **1** 119-130
[5] Waibel A, Hanazawa T, Hinton G, Shikano K, and Lang K 1989 IEEE Trans. ASSP **37** 328-339
[6] Le Cun Y, Boser B, Denker J S, Henderson D, Howard R E, Hubbard W, and Jackel L D 1989 Neural Computation **1** 541-551
[7] Lee S-Y and Marcuvitz N 1986 IEEE Trans. Ant. Prop. **AP-34** 613-625
[8] Lee S-Y and Szu H 1994 Opt. Eng. 33 2326-2330
[9] Lee H J, Lee S Y, Shin S Y, and Koh B Y 1991 Neural Computation **3** 135 143

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

All-optical self-organizing map applied to character recognition

K.Heggarty, J.Duvillier, E.Carpio Pérez and J.-L. de Bougrenet de la Tocnaye.

Département Optique, ENST de Bretagne, BP 832, 29285 Brest, France.
Tel +33 98001205, Fax +33 98001025, Email heggarty@gosc.enst-bretagne.fr

Abstract. We present character recognition results obtained with an self-organizing map neural network. The optical experimental system is built around two Ferroelectric Liquid Crystal Bistable Optically Addressed Spatial Light Modulators (FLC-BOASLMs) in a resonator configuration.

1. Network Model

The network we implement is based on Kohonen's self-organizing map [1] where the spatial neighbourhood of the active neurons is taken into account during the learning process to produce a topological organization of the activity: similar inputs produce similar map-layer activities. Two major modifications are made to the original Kohonen model to adapt it to our optical implementation.

The first concerns the neural decision where we replace the Winner-Take-All (WTA) decision by a global neural-map layer thresholding. This is because the WTA is a non-local nonlinearity which is difficult to perform with our BOASLMs. The WTA function can be realized with silicon-backplane SLMs [2] but our computer simulations (section 3) show that the simpler global thresholding can nevertheless lead to a system which learns and classifies as required, albeit with an expected reduction in capacity.

The second modification concerns the weights renormalization, again made difficult by the use of BOASLMs. The BOASLMs do, however, permit switching in both directions: to opaque and transmissive states. We use this capacity to reinforce active connections and weaken inactive ones, a procedure which, when correctly balanced, prevents weight saturation. This technique has also been confirmed in simulations.

With these modifications for a pixelized, binary, unipolar input pattern, X_{ij} $(i,j = 1 \cdots n)$, an array of weight maps $\{W\}_{ij}$, a map-layer neural activity, $\{Y\}$, and a map-layer neural firing, $\{Y'\}$, the network algorithm becomes:

$$\{Y\} = \sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij} \{W\}_{ij} \qquad \{Y'\} = f(\{Y\}) \tag{1}$$

$$\{W\}_{ij}(t+1) = \{W\}_{ij}(t) + \mu X_{ij}\{Y'\} - \nu X^c_{ij}\{Y'\} \tag{2}$$

where $\{\}$ indicates a quasi-continuous quantity (the BOASLMs are non-pixelized devices), X^c is the complement of X, μ, and ν are learning rate coefficients and f is a combined thresholding neighbourhood operator.

2. Experimental system

The experimental system, described in detail in [3], is centred around two BOASLMs. If such a modulator is illuminated with a given pattern and an electrical pulse is applied to the electrodes, the pattern can be binarized and stored on the modulator for subsequent rereading. One modulator is used in this way to perform the thresholding of the neural activity. Although the BOASLMs are intrinsically bistable, by operating them near the threshold voltage and using spatial integration techniques they can be made to show a grey-level behaviour [4]. A second modulator is used in this manner to store and update the synaptic weights. The sensitivity of the modulators can be varied by varying the height and duration of the electrical pulse. We use these parameters to change the threshold level and the learning rate.

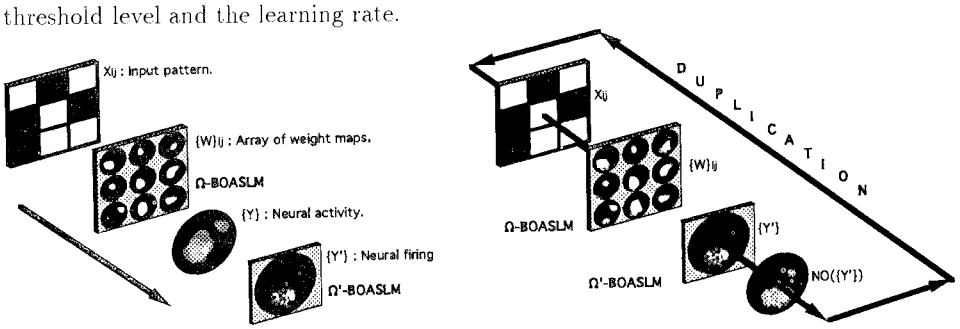


Figure 1. (a) Creation of the neural activity. (b) Weight updating

The basic operation of the system is shown above. The input pattern is presented to the system with an electrically addressed SLM built by CRL. The SLM is a 128x128 pixel FLC device, we group the pixels 12x12 to form one input pixel. The input pattern is imaged onto the array of weight maps (first BOASLM) thus activating the corresponding weight maps. An optical crossbar, consisting of a lenslet array and a collimating lens, then images and superimposes the activated weight maps onto the second (neural map) BOASLM thus realizing the required sum (eq.1a). The resulting neural activity is binarized and stored on the second BOASLM. The neighbourhood function is introduced by varying the height and duration of the voltage pulse sent to the BOASLMs to dilate the neural firing pattern.

This firing is read with a plane wave, duplicated with a Dammann grating and fed back, through the input image, to the weight map BOASLM where it is added to the active weight maps and subtracted from the inactive ones. The effect is to reinforce the connections between active input neurons and active zones in the neural map, while at the same time weakening non-productive connections. In this way a Hebbian-type learning rule is implemented.

The whole system is synchronized and the electrical pulse parameters varied with a personal computer which otherwise plays no part in the implementation. All neural calculations, thresholding, weight updating etc. are performed in parallel in optics.

3. Computer simulations

To confirm the feasibility of the implementation we have performed some computer simulations of the algorithm and the optical system. The simulations take some of the limitations of the optical system into account, in particular instead of weight renormalization they use the reinforcing/weakening procedure described above and a simple global threshold instead of the WTA decision. In addition the neighbourhood function is a simple dilatation. Three input classes were used: 8x8 pixel images of the characters F,I,L with three examples per class.

Figures 2 and 3 show some results of the simulations. In figure 2 from left to right are the input image, the weight map after learning, the map-layer activity (before thresholding) and the map-layer firing (after thresholding and dilatation). Figure 3 shows the inputs and corresponding firings for the other input images.

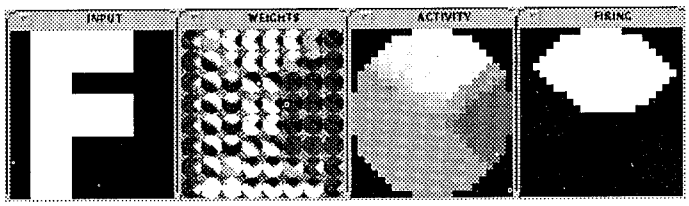


Figure 2. Results of computer simulations

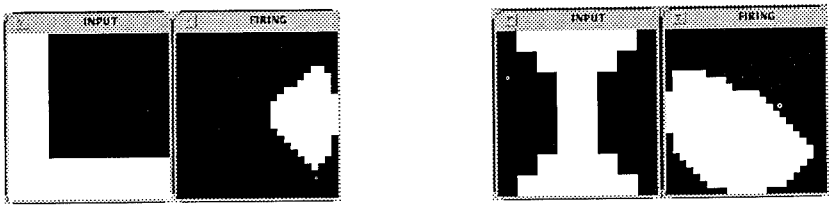


Figure 3. Simulated map-layer firings for different inputs.

We can see that the neural-map layer activity has been spatially organized by the learning process; different zones of the map being active for the different input classes. The learning also generalizes giving essentially the same output for each example in a class.

4. Experimental results

The first step in the experimental verification consisted in initializing the array of weight maps with an post-learning array obtained from the simulations and using it to check

the capacity of the experimental system to recognize the different inputs correctly. The initialization was performed by sending the array to the EASLM (using its full resolution) and imaging it onto the weight map BOASLM. The activities produced on the neural-map BOASLM for the different inputs classes appear in figure 4, from left to right the firings for inputs F,I and L. We can see a clear correspondance with the results of the simulations (figure 3).

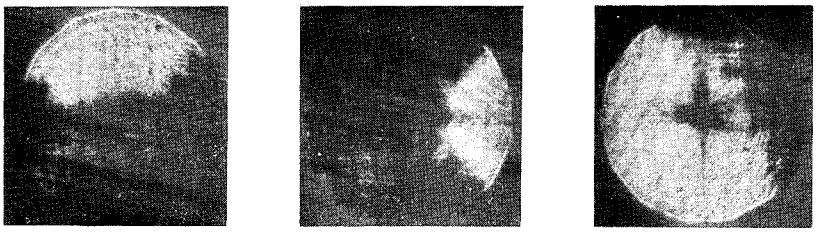


Figure 4. Experimental map-layer activities for different inputs.

The second more difficult step in our experiments consists in the implementation of all optical learning. Our first results in this direction appear in figure 5 where we see the initial and final weight maps and the activities for two different input images. The weight map has effectively been modified and differing activities produced.

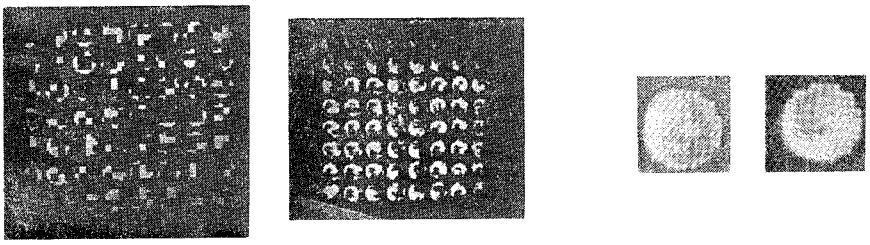


Figure 5. Initial map Final map Activities.

5. Conclusion

We have demonstrated the capacity of our system to recognize input characters and shown evidence of a learning capability. Improvements could be made by using a more complex decision stage, (for example with a smart OASLM) which appears to be the major limitation at present.

References

[1] T.Kohonen, 1990, IEEE Proceedings, **78** 1464-1480.

[2] K.Wagner and T.M.Slagle, 1993, Appl. Opt., **32** 1408-1435.

[3] J.Duvillier, M.Killinger, K.Heggarty, K.Yao and J.L. de Bougrenet de la Tocnaye, 1994, Appl. Opt., **33**, 258-266.

[4] M.Killinger, J.L. de Bougrenet and P.Cambon, 1991, Ferroelectrics **122**, 89-99.

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Evaluation of a Fast Access Phase Encoded Photorefractive Memory

Clara Alves, Lionel Gérard, Gilles Pauliat, Gérald Roosen

Institut d'Optique Théorique et Appliquée
Unité associée 14 au Centre National de la Recherche Scientifique
Bât. 503, BP 147, 91403 Orsay Cedex - France

Abstract
We experimentally analyze the origin of cross-talk in a photorefractive memory in which images are multiplexed with the deterministic phase encoding technique. We demonstrate a simple method to efficiently reduce most of this cross-talk.

1. Introduction

All the methods used to multiplex holograms in photorefractive crystals, as angular, frequency or phase encoding multiplexings, suffer from one common source of cross-talk due to the energy diffracted from non-Bragg-matched gratings. This source of noise is not very restrictive and allows large storage capacities [1]. Nevertheless, the phase encoding multiplexing scheme presents other sources of cross-talk, originating from setup imperfections, that may severely limit the storage capacity.

In this technique, each image is recorded by interfering the image beam with N overlapping reference beams. The set of relative phases between the reference beams is the image code. One may construct N sets of orthogonal phase codes so that N images can be stored in the material. The reading-out of a stored image is realized by illuminating the crystal with all N reference beams with the set of phases used to record that image. We use binary $(0,\pi)$ orthogonal phase codes (Hadamard codes). Theoretically, noiseless reconstructions appear because the reconstruction of undesired images interferes destructively. Nevertheless, if the phase shifts slightly differ from $(0,\pi)$ and/or if the diffracted reference beam amplitudes are not the same then the sets of phase codes are not perfectly orthogonal. That induces noise. We previously [2] derived the expressions for the amplitude signal S_p of the p^{th} image, and for the amplitude noise B_p on this image due to cross-talk with other recorded images. They are proportional to :

$$S_p \;\alpha\; A_p \sum_{n=1}^{N} u_{np} \qquad \text{and} \qquad B_p \;\alpha\; \sum_{m=1,m\neq p}^{M} A_m \sum_{n=1}^{N} u_{nm} e^{j(\varphi_n^p - \varphi_n^m)} \tag{1}$$

φ_n^m is the phase of the n^{th} reference beam during the recording of the m^{th} image. u_{nm} is the diffracted amplitude from the n^{th} reference beam onto the grating corresponding to the m^{th} image. A_m is the amplitude of the m^{th} binary recorded image, $A_m = (0,1)$ so that $S_p \;\alpha\; (0,N)$.

One sees on this formula that noise B_p may appear either if $u_{nm} \neq u_{n'm}$ or if $\varphi_n^p - \varphi_n^m \neq (0, \pi)$.

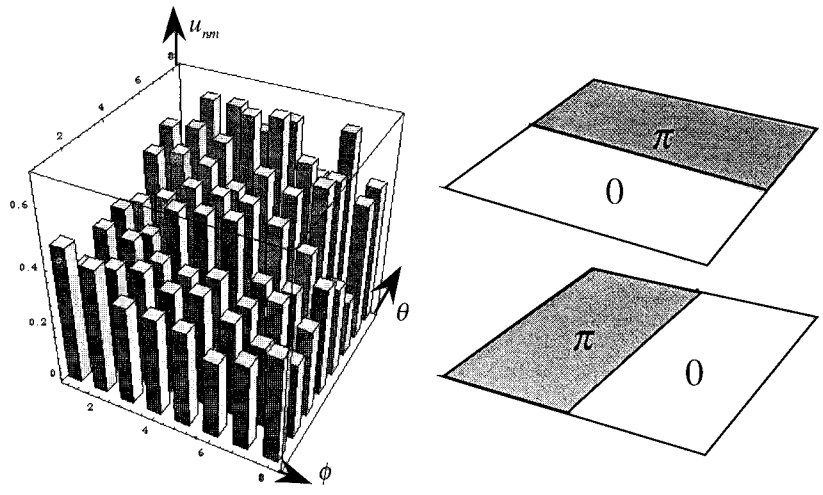
Hereafter, we experimentally investigate the influence of such imperfections on noise. A technique to decrease this cross-talk will be demonstrated.

2. Experimental arrangement

Our set-up was described elsewhere [3]. With 64 phase coded reference beams we can store 64 images in a $BaTiO_3$ sample. A dynamic 8x8 phase modulator allows a fast retrieval of all images. In that set-up the uniformity of u_{nm} and the accuracy of the phase modulator insure a large signal to noise ratio. We measured it to be, in intensity, $SNR_{\text{int}} > 140$.

In order to investigate the origin of noise in that kind of multiplexing, we intentionally misaligned our set-up. First a disturbance of the pixellated phase modulator produces a phase shift of $\pi + \varepsilon$ (instead of π), with ε a systematic phase error. A random phase error with a zero mean value was indeed proven to negligibly contribute to the cross-talk. Second, we introduced an amplitude noise ($u_{nm} \neq u_{n'm}$) by rotating the photorefractive crystal and working in an angular range where the diffraction efficiency greatly varies with the incident angle of the reference beam.

To measure these diffracted amplitude modulations, we modified the phase modulator and used it as an amplitude modulator. Each reference beam could be switched on while all other ones were off. We recorded N holograms (one for each reference beam) of an entirely white image. We then measured the diffracted signal when reading out the hologram with each reference beam. We thus determined $(u_{nm})^2$. A representation of these measurements is plotted in figure 1.a. The reference beams are arranged in a square matrix. The θ direction is in the incident plane, the ϕ is the perpendicular direction. We see that we have a random noise, $\sigma_u \approx 25\%$, plus slow variations in the θ and ϕ directions.



a) b)

Figure 1 : a) is a 3D plot of the 64 u_{nm} values versus the two angles θ and ϕ; b) represents two noisy codes.

These fluctuations come from: first the fan-out grating, which transforms the unique reference beam into 64 reference beams (the intensity variation between the reference beams is lower than $\sigma \approx 4\%$); second the phase modulator induces a fluctuation lower than 2%; third, the main error is due to the variation of angles between the different reference beams and the image beam in the crystal, it explains the dependence in the θ direction.

3. Phase error

Because the phase noise is a systematic noise, we can easily derive analytical expressions from formulae (1) in case $u_{nm} = u_{n'm}$. We find the same expression for the cross talk in all images except for the image recorded with the uniform phase code $(\varphi_1^1, \varphi_2^1, ..., \varphi_N^1) = (0,0,...,0)$. The image recorded with the uniform code is the only one to contribute to the noise at the first order in ε. From formulae (1), we find:

$$B_1 \ \alpha \ \frac{N}{2}(1-\cos\varepsilon + j\sin\varepsilon)\sum_{m=2}^{M} A_m \qquad (2)$$

and for $p \neq 1$, $\quad B_p \ \alpha \ \frac{N}{2}\left[(1-\cos\varepsilon)(\sum_{m=2, m\neq p}^{M} A_m) + (1-\cos\varepsilon - j\sin\varepsilon)A_1\right]$ (3)

4. Computer simulations and amplitude noise

Because of the random distribution of u_{nm}, it is impossible to derive an analytical formula for that cross-talk. Therefore, we computed values for formulae (1).

First of all, we checked that a phase only error (with $\varepsilon << 1$) produces uniform cross-talks between all images, except for the one recorded by the uniform code. These results are in agreement with formulae (2) and (3). Then we introduced in the simulation all the measured values for u_{nm}, the noise increases. We found it to be now different for each image. The noisiest image (largest B_p for $p \neq 1$) depends on the spatial distribution of the fluctuation on u_{nm}. For instance with the distribution shown in figure 1.a, the noisiest images are the ones recorded with the uniform code and with the codes shown in figure 1.b).

We experimentally checked that point. We inserted a neutral density filter on the right half of the reference beam matrix to partly compensate for the dependence of u_{nm} with θ. Consequently we measured a reduction of the noise on the image recorded with the first code shown in figure 1b. We can say that the appearance of noise in some images corresponds to the similitude of the code geometry with the geometry of variations of the u_{nm} (see figure 1).

5. Technique for cross-talk reduction

We propose two step technique to reduce cross-talk between images. They are derived from the observation of formulae (2,3).

The first idea is to set the amplitude of the first image $A_1 = 0$. The computer simulation showed this is efficient to reduce cross-talk even in case $u_{nm} \neq u_{n'm}$. We experimentally recorded first a set of images with the first image being white, $A_1 = 1$, and second a set of images with this first image being black, $A_1 = 0$. For instance, for $\varepsilon = 2.10^{-2} rad$, we measured an intensity signal to noise ratio divided by a factor 2.

The second idea to decrease the cross-talk is to discard the term $\sum_{m=2}^{M} A_m$ in formulae (2,3) while keeping $A_1 = 0$. The addition of a supplementary uniform phase modulator on the image arm was previously proposed [2]. A $\psi = 0$ or π phase shift is randomly superposed to the image amplitude during the recording process. On the average $\sum_{m=2}^{M} A_m e^{j\psi_m}$ becomes close to zero. The computer simulations showed that this technique is also efficient when we have both $\varepsilon \neq 0$ and $u_{nm} \neq u_{n'm}$. We experimentally verified that technique by inserting a uniform phase modulator on the reference arm. For instance for $\varepsilon \approx 3\,10^{-2} rad$ and without that technique we found an intensity signal to noise ratio (averaged on all images except the noisiest) larger than 100; for the image recorded with the uniform code we found 3 only; and for the images recorded with the codes shown in figure 1.b we got 7. By applying that technique we measured an intensity signal to noise ratio for all images larger than or about 150. There is no more noise on particular images.

6. Conclusion

We recorded 64 images in our set-up with low cross-talk between images. The sources of cross-talk have been increased by misaligning the set-up in order to make the analysis easier. We demonstrated a simple and very efficient solution to reduce that cross-talk and therefore to enhance the memory capacity.

Acknowledgment: This work has been conducted in the frame of the Parallel Optical Processors And Memories (POPAM) project funded by Commission of the European Community under the ESPRIT Program for Research and Development.

References

[1] K. Curtis and D. Psaltis, J. Opt. Soc. Am. A 10, 2547 (1993).
[2] C. Denz, G. Pauliat and G. Roosen and T. Tshudi, Appl. Opt. 31, 5700 (1992).
[3] C. Alves, G. Pauliat, G. Roosen, accepted for publication in Opt. Lett. (1994).

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Photorefractive Implementation of Neural "Self-Correcting" Learning Algorithms

Tigran Galstyan, Gilles Pauliat, André Villing, Gérald Roosen

Institut d'Optique Théorique et Appliquée, Unité de Recherche Associée 14 au Centre National de la Recherche Scientifique (CNRS), Centre Scientifique d'Orsay Bât. 503, B.P. 147, 91403 Orsay Cedex, France

Abstract. Some "self correcting" learning laws are compared with photorefractive dynamics. Possible optical implementations of corresponding neural basic units with photorefractive crystals are proposed and the key characteristics, required for self organizing network realization, are experimentally demonstrated. An optoelectronic network architecture based on presented above units and on a self-imaging feedback system is discussed.

1. Introduction

Neural self-correcting learning algorithms allow feature extraction in a stochastic set of input signals and selectively self-adapt to provide maximal activation for signals exhibiting the same features [1]. The learning is conducted by comparing the input signal with the weight vector in which are stored the common features extracted by the algorithm during the learning of previously processed signals. The mismatch is considered as an effective error. The learning process of the current input then results in a correction of this error by updating the weight vector components proportionally to that mismatch. So, the algorithm itself finds the classification code from the real time inputs. Digital simulations are mostly used to get such kind of algorithms and only the simplest versions of algorithms are implemented because of the complexity of the others.

However, the fundamental dynamics of hologram creation in some nonlinear optical materials appears to be perfectly fitted to implement these algorithms [2,3]. In this work we show, theoretically and experimentally, that different learning algorithms can be implemented by taking advantage of the kinetics of photorefractive holograms [2].

2. Theoretical model

Consider a neural unit performing weighted summation on the components of each input vector $\vec{S} = (S_1,\ldots,S_n)$ with its weight vector $\vec{W} = (W_1,\ldots,W_n)$. The neuron output η is a function f of the scalar product $\vec{W}.\vec{S}$. During the training period, a set of input vectors $(\vec{S}^1,\ldots,\vec{S}^m)$ is presented to the neuron and $\vec{W}$ is adapted as a function of η. Some of the most interesting self-adaptation laws are described by the vectorial equation [1]:

$$d\vec{W}/dt = a\eta^q \vec{S} - b\eta^p \vec{W} \qquad (1)$$

where t is the time, a and b are positive constants, parameters p and q are positive integers defining the learning law. The potentialities and performances of the neuron considerably differ according to the values of p and q, $p = 0,1,2$, $q = 0,1$ [1]. For instance, this adaptation law makes $\vec{W}$ become proportional to the "moving average of $\vec{S}$" for $p = q = 0$ (it does not show any interesting self organizing characteristics) [1]. On the contrary, the choice of parameters as $p = 1, q = 1$ or $p = 2, q = 1$ produces interesting adaptive behaviors. These

adaptation laws show such key characteristics as *Competitive learning* and *Feature extraction*. Namely, $\vec{W}$ converges towards one of learned vectors when, say, two orthogonal vectors were presented at the input (see computer simulation results at fig.1, e.g., for $p = 1, q = 1$, f is linear, and a=b=1). It converges towards the vector which represents the principal characteristics of the group of winning images when different orthogonal groups (composed by correlated images) are presented to learn. This is $\vec{W} = (0.5, 0, 0, 0, 0, 1, 1, 1, 0.5, 0)$ or $\vec{W} = (0, 0.5, 1, 1, 1, 0, 0, 0, 0, 0.5)$ for the above learning law and the inputs:

$$\vec{S}^1 = (0,1,1,1,1,0,0,0,0,0); \vec{S}^3 = (0,0,0,0,0,1,1,1,1,0) \qquad (2)$$
$$\vec{S}^2 = (0,0,1,1,1,0,0,0,0,1); \vec{S}^4 = (1,0,0,0,0,1,1,1,0,0) \ .$$

A traditional representation of an adaptive neuron and its holographic analog are sketched in fig. 2.a,b. During the recording, a reference beam, whose electric field amplitude is A_R, interferes with n signal beams whose amplitudes are $(A_{S1},\ldots,A_{Sn})$. n holograms with index modulations δn_i are thus recorded in the crystal.

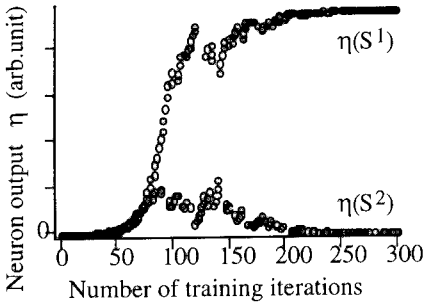
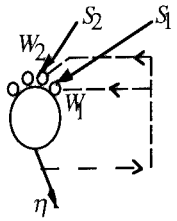


Fig.1



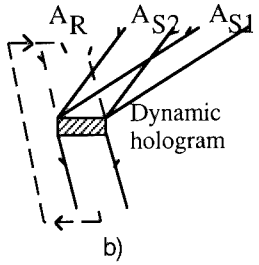


Fig.2

During each readout (the reference beam is closed), all corresponding diffracted beams add coherently and, assuming a small diffraction efficiency, the diffracted amplitude A_d is proportional to the weighted summation: $\sigma = \sum_{i=1}^{n} \delta n_i . A_{Si} = \sum_{i=1}^{n} \delta n_i . \|A_{Si}\|$. Thus, if $(A_{S1},\ldots,A_{Sn})$ are made proportional to the neuron input vector components $(S_1,\ldots,S_n)$, A_d will represent the neuron output η where the index modulations δn_i are proportional to the components of the weight vector. In order to implement an adaptive neuron, these index modulations must be adapted according to the value of η by using a feedback loop (see eq.1).

For the geometry depicted in fig. 2.b, n space-charge electric fields E_i and corresponding holographic δn_i gratings are recorded in the photorefractive material [4]. Assuming a diffusion process, the kinetics of these E_i is governed by the following n equations:

$$\begin{cases} \dfrac{dE_i}{dt} = \dfrac{I_t}{\tau_{0,i}} \left(\dfrac{2\, A_{Si}^{*} . A_R}{I_t} E_{sc,i} - E_i \right) \\ \text{with} \quad I_t = A_R . A_R^{*} + \sum_{j=1}^{n} A_{Sj} . A_{Sj}^{*} \end{cases} \qquad (3)$$

$E_{sc,i}$ is the maximum steady state space charge electric field that can be induced in the material. $\tau_{0,i}$ is the photorefractive time constant for unit illumination. Both parameters $E_{sc,i}$ and $\tau_{0,i}$ depend on the grating wave vector induced by the interferences between A_{Si} and A_R. Nevertheless, we assume that the whole set of signal beams lies in an angle small enough for the variations of these parameters to be negligible. Therefore, effective electrooptic coefficients will not depend on the signal beam orientation, i.e. of parameter i, and equation (3) can be re-written by changing E_i into δn_i.

Similarly to computer simulations, we update the neuron weights by a discrete time method:

i) For each input vector $\vec{S}$ the diffracted amplitude (with the reference beam switched off) provides the neuron output η.
ii) The weights are updated by re-enforcing the gratings during a time period Δt with the reference beam switched on. Both amplitude $\|A_R\|$ of the reference beam and time period Δt are adjusted according to the neuron output and the desired adaptive law. Changes ΔE_i are:

$$\Delta E_i = \frac{2\,\|A_{Si}\|\cdot\|A_R\|}{\tau_0} E_{sc}\,\Delta t - \frac{I_t}{\tau_0} E_i\,\Delta t. \qquad (4)$$

From the comparison of equations (1) and (4) we derive the following adaptation rules:
a) Case $p = q = 0$: This trivial case is achieved by setting to a constant value both $\|A_R\|$ and Δt whatever the output result η.
b) Case $p = 1, q = 1$: This case is easily implemented by setting $\|A_R\|$ to a constant value but by making the recording time Δt proportional to η. There is now a total correspondence between the updating law and the kinetics of the photorefractive effect.
c) Case $p = 2, q = 1$: To implement this case, we adjust the different beam intensities so that the reference beam intensity is much larger than the sum of the signal beam intensities. Now, Δt is a constant but $\|A_R\|$ is made proportional to η.

3. Experimental realization

The scheme of the experimental implementation of a photorefractive neural basic unit is presented in fig.3 [2]. The initial plane wave from a CW Ar laser at 514 nm is split into the reference R and the signal S arms by beam splitter BS. The signal beam is modulated by a liquid crystal television screen, SLM, (model XV100ZM from Sharp) and then exploded in a multitude of spherical beamlets by two crossed grids of cylindrical microlenses ML. The image focal plane of these microlenses is set in the object focal plane of lens L. The photorefractive crystal ($BaTiO_3$) is set in the image focal plane of L. Thus each beamlet from a microlens is transformed into a plane wave so that all beams are superposed in the crystal. The set up is driven by a personal computer PC. It is used to impress the input vectors (images) on the SLM, to digitize the signal of diffraction detected by the photodiode and to control the optical shutter SH disposed on the reference arm.
For the case $p = q = 0$, both the recording time period and amplitude of reference beam are made constant and thus independent of output η. Two orthogonal vectors $\vec{S}^1 = (1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0)$ and $\vec{S}^2 = (0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1)$ are randomly presented on the input SLM. As expected for this learning law, no selective learning (selective adaptation) exists during experiment. We obtain a temporarily averaged hologram recording.

The recording time period is then made proportional to η (i.e. to the diffracted amplitude or to diffracted intensity according to the chosen f) to implement the case $p = q = 1$

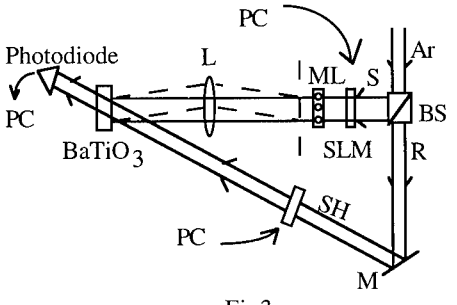


Fig3.

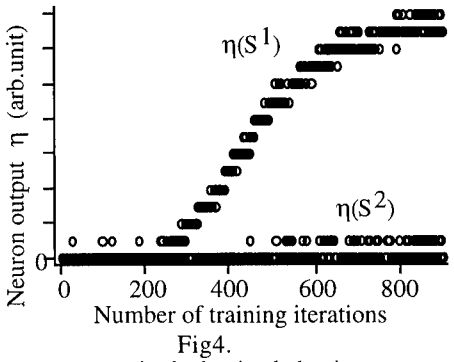


Fig4.

Conversely to the previous case, now the neuron presents required adaptive behaviors.
a) *Competitive learning*: The dynamic behavior for the same input vectors can be seen in fig.4

where the output of neuron is presented versus the number of processing steps. f is linear. There is very weak initial diffusion (with approximately equal intensity) towards the photodiode (neuron output) for both input vectors. After some period of initial processing the randomly accumulated fluctuations break the symmetry and one of these vectors, $\vec{S}^1$, becomes the winner. With further processing the weight vector formation (and respectively the output) of neuron reaches its saturation level. We receive the following picture of normalized weight components trying to readout the neuron weights with separate vector elements :

$$\vec{W}^1 = (0.7, 0.2, 0.8, 0.1, 0.7, 0.2, 0.8, 0.1, 1.0, 0.1, 0.9, 0.1, 0.8, 0.1, 0.7, 0.2)$$

The weight vector is thus oriented parallel to input vector $\vec{S}^1$. This winning vector is not defined *a priori* and the learning results into the winning of one of the vectors with equal probability if the same conditions are provided. That is, the scheme demonstrates "flip-flop"-like behavior.

b) *Feature extraction*: Two orthogonal groups of inputs, each one composed by two correlated vectors, given by equation (2) are presented to the neuron in a random sequence. After 800 iterations have been performed, the weight components are read out. We get the following distribution of components:

$$\vec{W} = (0.53, 0.23, 0.18, 0.22, 0.22, 1.0, 0.93, 0.93, 0.56, 0.13)$$

The system has recognized and selectively adapted a common feature in the two vectors $\vec{S}^3$, $\vec{S}^4$ (see the computer simulation results).

4. Discussion

The proposed implementation for a photorefractive neuron is able to perform quite interesting data processing. We are currently working on the implementation of a self organizing neural network via juxtaposition of a large number of identical neural units in different locations of the same photorefractive crystal. This is a self-imaging optoelectronic feedback system created by changing the photodiode (see fig.3) for a CCD camera (at the image plan of the $BaTiO_3$) and by replacing the optical shutter by another spatial light modulator (which is imaged on the $BaTiO_3$).Very large capacities should thus be obtainable. Namely, the maximal dimensionality of input n here can be estimated as the possible maximal number of angularly multiplexed images. Optical photorefractive memories with as many as 5000 angularly multiplexed images have already been experimentally demonstrated [5]. Diffraction limit for a crystal with 1x1x0.5 cm^3 allows for the juxtaposition of more then 100x100 neural units. An electronic lateral feedback (to control neural neighbourhood relations) and presented above fundamental photorefractive kinetics provide together the global network organization algorithm (e.g., Kohonen Map). It would permit to implement the higher order neural adaptation laws (cases q=p=1 and q=1; p=2) which are usually strongly simplified in computer simulations, limiting thus network learning capacities and requiring additional external intervention [1]. This first experimental demonstration encourages us to look for other nonlinear optical materials as well as for suitable holographic characteristics (dynamics, dark memory, geometrical dependence etc.) to implement different neural computing algorithms.

Acknowledgements
Dr. Tigran Galstyan acknowledges the support of Direction des Recherches Études et Techniques (DRET).

References

[1] Kohonen T 1989 Self-Organization and Associative Memory (3rd ed., Berlin, Heidelberg, Germany, Springer-Verlag)
[2] Galstyan T, Pauliat G, Villing A, Roosen G 1994 Opt.Comm. **109** 35-42.
[3] Wagner K, Slagle T M 1993 Appl. Optics **32** 1408.
[4] Günter P and Huignard J P eds 1989 Photorefractive Materials and Their Applications I-II (Springer-Verlag, Vol. 61-62, Berlin, Heidelberg, Germany).
[5] Mok F H 1993 Optics Letters **18** 915.

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22-25 August 1994
© 1995 British Crown Copyright

Dynamic Effects in Volume Interconnects with Feedback

Chris Slinger and Pat Conway

DRA Malvern, St Andrew's Road, Great Malvern, Worcestershire, WR14 3PS, UK.
slinger@signal.dra.hmg.gb

Abstract. An attempt is made to exploit the inherently rich physics of volume holographic interconnects. The effects of crosstalk and nonlinearities are considered. When combined with a novel feedback configuration, complex, self-organising behaviour is seen to emerge.

1. Introduction.

Volume holographic interconnects [1] represent one way of applying optics in connectionist approaches to signal processing. They combine the usual advantages of optical interconnects, together with high storage capacities ($>10^9$ weighted connections cm^{-3}) and offer the possibility of weight update during training [2].

However, when used in the 'conventional' way, several factors constrain their performance. These include effects such as diffracted amplitudes being a nonlinear function of grating strength; cross gratings and multiple grating interactions giving crosstalk; the recording of modifying of gratings invariably changing the strengths of other gratings in the hologram and nonlinearities introduced by the recording process itself. All the foregoing can limit the use of volume interconnects, even when the training process takes account of these factors [3]. An alternative approach is to attempt to try and use these phenomena to good advantage - in effect, to exploit more fully the rich physics of this dynamic, multiple grating system. This may enhance the interconnects capabilities and indeed, may enable it to perform useful information processing tasks in its own right. To investigate this idea, it is necessary to study the regimes of behaviour of the interconnect in various configurations. In the following, one such arrangement is described and investigated. Some conjectures on the utility of the modified architecture are then briefly discussed.

2. Architecture.

A schematic of a modified architecture is shown in Figure 1. A single mode fibre has been added to the conventional volume interconnect arrangement. This fibre feeds light from one of the outputs, back to input 1, via a beamsplitter of reflectivity R. Superficially, this arrangement has similarities to the ring resonator systems studied by Ikeda, Moloney, Anderson, Firth and others [e.g. 4], but here the resonator path occupies a channel of a dynamic, multiple grating interconnect, rather than passing through a nonlinear dielectric slab. Additionally, the fibre is single mode to remove transverse effects in the resonator.

An input pattern (e.g. an image or the output of some array of processing nodes in a connectionist architecture) is presented to the system as a spatially coherent, complex amplitude distribution. The wave from each element of this input array interferes with those from other elements of the array to form a set of diffraction gratings in the dynamic holographic material. The waves from the input array, in addition to being responsible for the formation of the gratings, are diffracted by them, before they propagate to the output

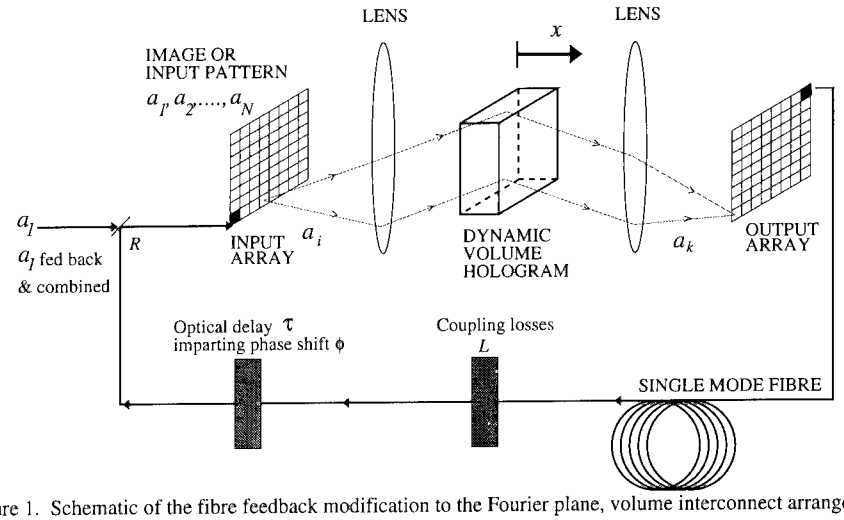


Figure 1. Schematic of the fibre feedback modification to the Fourier plane, volume interconnect arrangement. Note the dynamic holographic material (e.g. a liquid crystal cell or a photorefractive crystal) and the absence of a training array.

array (perhaps for subsequent post processing). For a more details of the conventional volume architecture and the grating formation see [2].

3. Theory.

After some approximations [2], the governing equations describing the propagation of light through the interconnect can be written as :

$$\left\{\frac{n}{c}\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right\} a_m + \frac{j\,\gamma}{\cos\theta_m} \sum_{\substack{n=1 \\ m \neq n}}^{N} G_{mn}\, a_n = 0 \qquad (1)$$

$$\frac{\partial\, G_{mn}}{\partial\, t} = -A\, G_{mn} + B\, a_m\, a_n^* \qquad (2)$$

where n is the bulk refractive index of the hologram material, c is the velocity of light, θ_m is the angle of propagation of wave m, normalised amplitude a_m. G_{mn} is the mnth grating strength. γ, A and B are material constants. Equations (1) are coupled wave equations describing the diffraction of the waves as they propagate through the hologram: the amplitude of each wave is coupled to the other diffracting waves amplitudes through the appropriate diffraction gratings. Note that cross gratings and multiple diffraction paths are included. Equations (2) describe the grating formation in the hologram - for the case shown, a photorefractive type characteristic is assumed. This type of response can also approximate that of other materials e.g. liquid crystals. The second term on the right of (2) shows that the rate of growth of the mnth grating is related to the strength of the appropriate interference pattern, whilst the first is a decay term. The boundary conditions are: $a_m(x=0, t) = a_{m0}$ $(m \neq 1)$ and $a_1(x=0, t) = a_{10}\,(1\text{-}R) + R\,(1\text{-}L)\, a_1(x=d, t=t\text{-}\tau)\, \exp(j\phi)$, where L is the loss, τ is the time delay in the feedback loop and ϕ is the phase shift associated with propagation through it.

The time delayed feedback, combined with the nonlinearities and time dependent nature of multiple gratings in the hologram, can endow the arrangement with very rich and varied behaviour. In the following subsections, this is looked at more closely, starting with a stability analysis - a standard procedure when studying nonlinear dynamical systems.

3.1. Stability analysis.

The size of the time delay τ can play an important role in the behaviour of the arrangement, but a stability analysis is difficult due to the complexity of the system. Some progress can be made, however, if it is assumed that the gratings do not vary significantly through the depth of the hologram, i.e. $G_{mn}(0) \equiv G_{mn}(x)$ for $(0 \leq x \leq d)$. The problem then reduces to studying the properties of the infinite dimensional, delay differential equation :

$$\frac{d\,\boldsymbol{G}(t)}{d\,t} = f\,[\boldsymbol{G}(t),\, \boldsymbol{G}(t-\tau)] \qquad (3)$$

where $\boldsymbol{G}$ is the matrix of the grating strengths. The stability can be determined by perturbing about the system's steady state solution $\boldsymbol{G}_{ss}$ and studying the way in which the perturbations grow. This is equivalent to analysing the roots ω of

$$\det\,[\mathbf{J}(\boldsymbol{G}_{ss},\, \boldsymbol{G}_{ss}) + \mathbf{J}_{\tau}(\boldsymbol{G}_{ss},\, \boldsymbol{G}_{ss})\,\exp(-\omega\tau) - \omega\,\mathbf{I}\,] = 0 \qquad (4)$$

where $\mathbf{I}$ is the identity matrix and the Jacobians $J_{ij} = \{\partial f_i[\boldsymbol{G}_i(t), \boldsymbol{G}_i(t-\tau)]/\partial G_j(t)\}_{|G(t)=G(t-\tau)=Gss}$ and $J_{\tau ij} = \{\partial f_i[\boldsymbol{G}_i(t), \boldsymbol{G}_i(t-\tau)]/\partial G_j(t-\tau)\}_{|G(t)=G(t-\tau)=Gss}$ [5]. Roots having real parts greater than zero correspond to unstable behaviour.

In general, the main results are that firstly, the number of distinct roots (and therefore the complexity) of the system increases as the time delay increases, and secondly, the system stability can change as the time delay is varied. These observations agree qualitatively with Farmer's analysis of another delayed nonlinear system, the Mackay-Glass equation [6]. In our case, the effects of delay start to become important at $\tau.A > 0{\cdot}01$. For most photorefractives, time delays produced by fibres are therefore too short to significantly alter the interconnect dynamics, but systems incorporating rotating photorefractives have demonstrated suitably long, coherent delays, together with gain [7]. It may thus be possible to vary the dynamics of the system by control of the delay time τ.

Further stability analyses in terms of variation of other system parameters have also been carried out. We have some new analytic results for multiple grating systems which allow such studies to be made *without* making the assumption that the gratings do not vary with hologram depth. Space precludes presentation of these results here, but system behaviour is found to vary strongly with, for example, the ratio of the complex photorefractive material coefficients B to A.

3.2. Bifurcation behaviour.

Several bifurcation analyses have been undertaken, illustrating how the dynamics of the arrangement evolve with variation in parameters. Figure 2 shows a bifurcation diagram resulting from the variation of the feedback loss L. Suitable adjustment of this easily varied parameter should therefore provide an excellent method to "tune" the system dynamics.

4. Discussion.

We have made some preliminary experimental observations using barium titanate crystals in multiple grating, feedback configurations. Initial results do suggest that complex dynamics can occur. Such modes of operation are being increasingly viewed as advantageous in pattern processing. For example, the ability of an associative memory to change its dynamics, during a search task, has been demonstrated to give superior classification performance [8]. Additionally, complex dynamics in a classifier can also give it advantages in noisy environments. The volume interconnect system described here also has direct equivalence to many features of the coupled map lattice paradigm [9]. Looked at in a

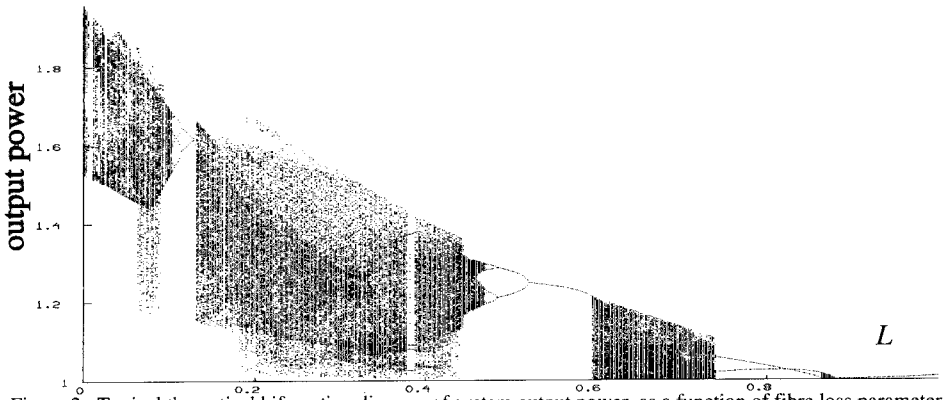


Figure 2. Typical theoretical bifurcation diagram of system output power, as a function of fibre loss parameter L. Fixed points (e.g. at L=1, 0.55), limit cycles, period doubling routes to chaos, and chaotic behaviour with 'gaps' where periodic motion returns, can all be seen. N=2, γd=50, $a_{10}=a_{20}$=1.0, $\phi=\pi/2=\angle B/A$, $|B/A|$=0·327.

slightly different way, the iterative processing, generated by the feedback, endows the system with properties not obtained with layered networks of comparable size.

Self organisation in pattern processing can be useful as a preprocessing stage of a pattern recognition system (e.g. as in Kohonen's self organising feature maps [10]). The interconnect described here also has its own self organising features. In particular, it is able to associate input patterns with different attractors at its output. Inputs that differ slightly (within the corresponding basin of attraction) can give rise to the same attractor. In this way, the system performs classification and clustering without external supervision.

5. Conclusion.

A system has been formulated that attempts to exploit, more fully, the complex physics available in the volume interconnect. This includes the use of many processes conventionally regarded as detrimental to interconnect behaviour. By suitable control of accessible parameters (e.g. L), it may be possible to "tune" the behaviour of the arrangement. When combined with the merits of the conventional volume interconnect, unusual, and possibly beneficial, modes of behaviour may result.

References.

[1] Psaltis D, Brady D, Gu X G, Lin S, 1990 "Holography in artificial neural networks" *Nature* **343** 325-330.
[2] Paek E G, Wullert J R, Patel J S, 1989 "Holographic implementation of a learning machine based on a multicategory perceptron algorithm" *Optics Letters* **14** 1303-1305.
[3] Slinger C W, 1991 "Weighted volume interconnects for adaptive networks" *Optical Computing and Processing* **1** 219-232.
[4] D'Alessandro G, 1992 "Spatiotemporal dynamics of a unidirectional ring oscillator with photorefractive gain" *Physical Review A* **46** 2791-2802.
[5] Epstein I R, Luo Y, 1991 "Differential delay equations in chemical kinetics. Nonlinear models : the cross-shaped phase diagram and the Oregonator" *Jrnl. Chemical Physics* **95** 244-254.
[6] Farmer J D, 1982 "Chaotic attractors of an infinite dimensional dynamical system" *Physica* **4D** 366-393.
[7] Zhou G, Anderson D Z, 1993 "Photorefractive delay line for the visualization and processing of time dependent signals" *Optics Letters* **18** 167-169.
[8] Nara S, Davis P, Totsuji H, 1993 "Memory search using complex dynamics in a recurrent neural network model" *Neural Networks* **6** 963-973.
[9] Kaneko K (ed), 1993 *Theory and applications of coupled map lattices* (Nonlinear Science Series: Wiley).
[10] Kohonen T, 1989 *Self-organization and associative memory* (Series in Information Sciences: Springer).

© British Crown Copyright 1994/DRA

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Holographic Associative Memory Using Learning Pattern Method

Ho Hyung Suh, Chong Hoon Kwak, Gyungock Kim, and El-Hang Lee

Research Department, Electronics and Telecommunications Research Institute
P.O. Box 106, Yusongku, Taejon, Korea

Abstract. We have developed a new holographic associative memory (HAM) which can be implemented easily using learning patterns derived from an adaptive learning rule. The principle of learning pattern method (LPM) and simulation results of the present HAM are presented.

1. Introduction

Although the optical implementation of an associative memory such as the Hopfield model can be done easily with vector-matrix multiplications[1] or holographic systems[2],[3], the performance may be limited due to the practical difficulties of having pseudo-orthogonality in memory patterns and the existence of undesirable stable states. The LPM utilizes the high performance of adaptive learning by using the simple outer-product learning implementation of HAM.

2. Principle of LPM

Let b^m be the m-th memory pattern to be stored, then the error function E is defined as

$$E = \frac{1}{2}\sum_{m}^{M}\sum_{i,j}^{N}(o_{ij}^{m} - t_{ij}^{m})^2, \tag{1}$$

where $o^m$and $t^m(=b^m)$ is the actual output state and its target state(i.e. memory pattern), respectively.

$$o_{ij}^{m} = f\left[\sum_{k,l}^{N} W_{ijkl} b_{kl}^{m}\right], \tag{2}$$

We have adopted gradient-descent algorithm to get the optimal interconnection weight W.

$$\Delta W_{ijkl} = -\eta \frac{\partial E}{\partial W_{ijkl}}$$
$$= -\eta \frac{\partial E}{\partial o_{ij}^m} \cdot \frac{\partial o_{ij}^m}{\partial W_{ijkl}} \tag{3}$$
$$= -\eta \left(o_{ij}^m - t_{ij}^m \right) b_{kl}^m$$

$$W_{ijkl} = 0, \text{ if } W_{ijkl} \le 0. \tag{4}$$

where, η is a constant of learning rate. The above equation is used so as to obtain the interconnection matrix of discrete and positive elements with appropriate value of η. After optimization the the interconnection matrix W can be expressed as the liner combination of binary matrices T^s as follows:

$$W = \sum_{s=1}^{S} a_s T^s = \sum_{s=1}^{S} \alpha_s \mathbf{u}^s (\mathbf{v}^s)^T. \tag{5}$$

where α_S is the constant coefficient and the matrix T^s is the outer product between learning patterns u^S and v^S with binary elements (1,0).

3. Holographic Implementation of LPM

The three two-dimensional binary patterns L, P, and M are selected as the memory patterns as shown in Fig.1, and they can be written in the form of 25(5×5)-bits vectors as follows:

(L) : b^1 = (1000010000100001000011111),
(P) : b^2 = (1111010001111101000010000),
(M): b^3 = (1000111011101011000110001).

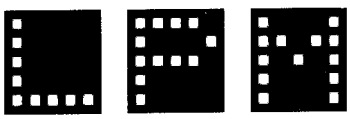


Fig. 1. Three two-dimensional memory patterns

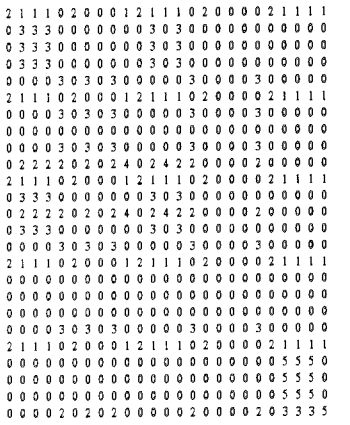


Fig. 2. The interconnection matrix W obtained from the adaptive learning rule.

Figure 2 illustrate the 25×25 interconnection matrix obtained by using the learning algorithm in Sec. 2. We can derive ten pairs of learning pattern pairs (v^s,u^s) and their coefficients α_s out of the matrix W after simple algebra (Fig.3). As shown in Fig. 4, the interference patterns between the collimated beam passing through the pattern v^s and the scattered beam from the pattern u^s and ground glass construct holographically the outer-product between v^s and u^s. [3], [4] The coefficients α_s controls the exposure time. The experimental setup for the retrieval of stored informations is shown in Fig. 5.

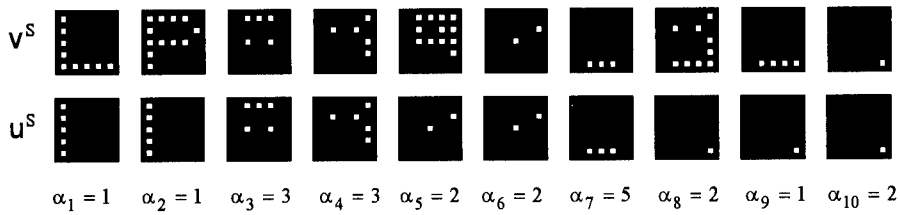


Fig. 3. Ten learning pattern pairs (v^s,u^s) and their coefficients α_s (s=1,2,...,10).

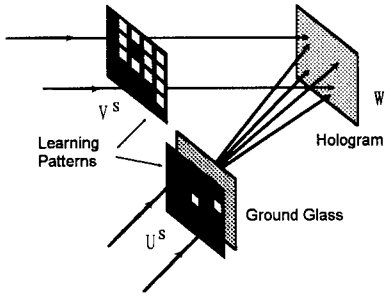


Fig. 4. Schematic diagram of recording the outer-product of learning patterns holographically.

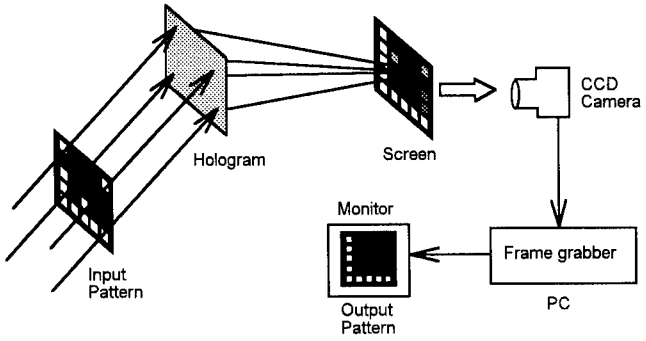


Fig. 5. Setup of the retrieval of the memory pattern for the input pattern.

The energy function U is defined as follows:

$$U = -\frac{1}{2}\sum_{ij}^{N}\sum_{kl}^{N} W_{ijkl} b_{ij}^{s} b_{kl}^{s} + \lambda (N_0)^2 . \tag{6}$$

where N_0 is the number 1's of input b^s and λ is a constant. Since the elements of the matrix are positive, the input-dependent threshold level should be used. It is important to use appropriate value of λ, because the stable state of energy and the input-dependent threshold level depend on it. Figure 6 shows the simulation results of the LPM for the various values of λ. The recognition probability is defined as the number of correct recognition per total number of the inputs. The simulations were repeated over 300 randomly generated inputs of each Hamming distance $(0 \sim 15)$, and the results are averaged for the three memory vectors L, P, and M. We could get best results in the range of $0.6 \le \lambda \le 0.7$. In conclusion, all memory patterns can be stored in stable state and the recognition capability was improved considerably by using the learning patterns instead of direct using the memory patterns.

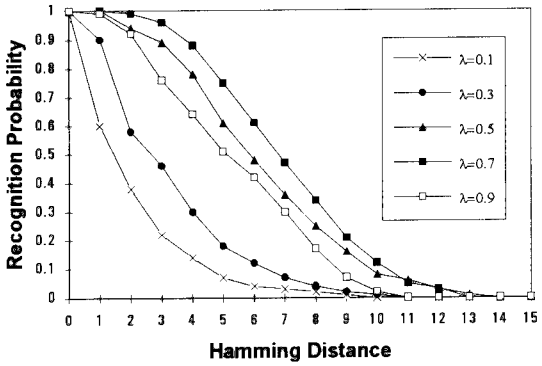


Fig. 6. Recognition probability versus Hamming distance of input patterns with various constant scale factors λ.

References

[1]. Farhat N H, Psaltis D, Prata A and Paek E 1985 Appl. Opt. **24** 1469-75

[2]. Jang J S, Jung S W, Lee S Y and Shin S Y 1988 Opt. Lett. **13** 248-50

[3]. Suh H H and Lee S S 1992 Appl. Opt. **31** 199-204

[4]. Suh H H, Kwak C H and Lee S S, 1993 KOGAKU (Jpn. J. of Opt.) **22**, 721-6

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical associative memory with bipolar edge-enhanced learning

Xu-Ming Wang, Trevor J. Hall and Jian Wang

Department of Physics, King's College London, Strand, London WC2R 2LS, England

Abstract

An optical associative memory with bipolar edge-enhanced feature learning by using a ferroelectric liquid crystal spatial light modulator and a barium titanate crystal is presented. A high discrimination ability is achieved. Experimental result and computer simulation are given.

1. Introduction

In an associative memory, the recalled output is the weighted summation over all the stored patterns. The weight is determined by the similarity of the input stimulus and the stored pattern. To achieve a better association we increase the wanted weight and suppress the unwanted ones. We can find the same objective in optical correlation pattern recognition where we seek better discrimination abilities for a correlator. Therefore techniques used in pattern recognition are applicable to the optical implementation of associative memories. The phase-only filter (POF) has been known for its optimal discrimination. Investigations revealed that the magnitude squared of the impulse response of the optical correlation system using a POF is the edge outline of the original reference and the histogram of the impulse response consists of both positive and negative values; the positive part lies predominantly inside the outline while the negative part lies predominantly outside the outline. Therefore we can produce ternarily valued edge-enhanced versions of patterns to be stored by the associative memory. In the edge-enhanced version of a pattern, pixels are assigned values of 1 and -1, respectively, if their positions in the original pattern are adjacent to the outline of the object and are inside and outside the object, respectively, while the other pixels are given the value of 0. The bipolar edge-enhanced pattern has equally distributed positive and negative pixels, which meets the equal-distribution requirement for an associative memory.

We set up an optical associative memory which uses a barium-titanate photorefractive crystal (PRC) as the memory recording medium, a ferroelectric Spatial Light Modulator (SLM) as the programmable input device and a charge coupled device (CCD) camera as the output detector. The use of a PRC enables the interconnect hologram to be recorded in real-time, which also introduces reconfigurability into the system. The use of the SLM permits a programmable system.

2. Holographic associative memory

Suppose that there are M patterns f_i, $i = 1, 2, \ldots, M$, with binary values (0, 1), to be stored in a holographic associative memory. There exists an input-output pair of patterns, f_i' and f_i'', used in hologram recording for the storage of pattern f_i. When the learning process is finished and the associative recall is in progress with a stimulus, s, being input into the system, the expected weighted output, E, satisfies

$$E = \sum_{i=1}^{M} \{s \cdot f_i'\} f_i'', \tag{1}$$

where $\{s \cdot f_i'\}$, known as the weight of f_i'' in the output, represents the inner product of patterns s and f_i'. A threshold on the expected output gives the actual output of the system expressed by

$$O = T\left[\sum_{i=1}^{M} \{s \cdot f_i'\} f_i''\right], \tag{2}$$

where $T[\cdot]$ denotes a sigmoid-like threshold function. If an original pattern is correctly memorized, the thresholded output must be that pattern when it is input into the system. To achieve a best performance, we should have the unthresholded output be as similar to the correct stored as possible. One of the methods is to design the patterns f_i' and f_i'', by othogonalization preprocessing for example [1], so that when one of the stored pattern is used as the stimulus the unthresholded output is the stored pattern itself. Another method is to leave the function, f_i'', to be the original pattern, f_i, and to design f_i' only so that the weight for the correct stored memory is much higher than all the other weights. The bipolar edge enhanced learning belongs to the later.

3. Learning principle of the optical system

The optical system performing the bipolar edge-enhanced feature learning is schematically shown in Fig. 1. A laser beam is enlarged and focused

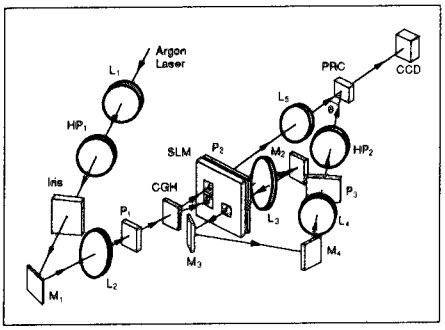


Figure 1: Optical system for the learning; HP_1 and HP_2 are halfwave plates, L_1 – L_5 are lenses, M_1 – M_4 represent mirrors, and P_1, P_2, and P_3 denote polarizers.

onto the SLM by lenses L_1 and L_2. A computer generated hologram (CGH), which fans out the incident beam into 33×33 equal-intensity beams, is used to pixelate the patterns displayed on the SLM and to ensure that the high contrast ratio of the SLM is not compromised by the scattering by interpixel dead space. The other purpose of using a CGH is to introduce a random phase function. The original pattern and its bipolar edge-enhanced version are simultaneously displayed on a binary SLM with the edge enhanced and the original corresponding to the input and output in the input-output pair, respectively.

The ternary modulation, as illustrated in Fig. 2, is realized in the following way: Two pictures, la-

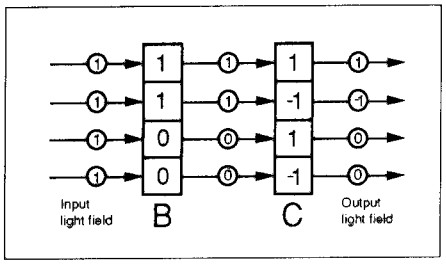


Figure 2: Ternary modulation of a binary SLM; a number in a box denotes the value written on the SLM, and a number in a circle is the value represented by the light field.

belled B and C, respectively, are simultaneously displayed on the SLM. The fan-out beams of the CGH illuminate B. Beams passing through B will be reflected by mirror M_2 onto pattern C. The system is so arranged that B is coincidedly imaged pixel-wise onto C. Polarizers P_1, P_2 and P_3 are so placed that light passing through B (and A) is binarily amplitude modulated while light passing back through C is binarily phase modulated.

For each of the to-be-stored patterns, an original pattern, $f_i(x, y)$, is displayed in binary amplitude form on the SLM at position A while two patterns are placed at positions B and C, respectively, to form a ternary edge-enhanced version of that pattern, $f_i'(x, y)$. Light passing through A goes via the upper path of the system with the CGH and pattern A being imaged onto the PRC and CCD, respectively, by lens L_5. So we have the field behind the SLM be expressed by

$$\hat{u}_i(x, y) = f_i(x, y)\,\mathcal{H}(x, y), \tag{3}$$

where (x, y) denotes the coordinates system of the SLM plane,

$$\mathcal{H}(x, y) = \sum_{m=-16}^{16} \sum_{n=-16}^{16} \exp\left[i\phi(m, n)\right] \delta(x - md, y - nd) \tag{4}$$

denotes the diffraction of the CGH, d is the period of the SLM pixels, and coefficient constants have been dropped for convenience. It can be shown that the light distribution on the PRC is

$$\hat{u}_i'(\alpha, \beta) = \exp\left[-\frac{ik}{2d_4}\left(\alpha^2 + \beta^2\right)\right] \left[\mathcal{F}_i\left(\frac{k(d_1 + d_2)}{d_1 d_3}\alpha, \frac{k(d_1 + d_2)}{d_1 d_3}\beta\right) * h\left(-\frac{d_1 + d_2}{d_3}\alpha, -\frac{d_1 + d_2}{d_3}\beta\right)\right], \tag{5}$$

where (α, β) is the coordinate system in the PRC plane. $\mathcal{F}_i(p, q)$ is the Fourier transform of $f_i(x, y)$, which satisfies

$$\mathcal{F}_i(p, q) = \int f_i(x, y) \exp\left[-i(px + qy)\right] dx\, dy, \tag{6}$$

$h(u, v)$ is the transmittance of the CGH, that is

$$\mathcal{H}(x, y) = \int h(u, v) \exp\left[-\frac{ik}{d_1}(xu + yv)\right] du\, dv, \tag{7}$$

$*$ denotes convolution, and k is the wave vector of the light source.

Light passing through both B and C is directed by the lower path of the system with C being imaged onto the PRC. The light distribution on the crystal for this path is written as

$$e_i(\alpha, \beta) = \exp\left[\frac{ik(d_5 + d_6)}{d_6^2}(\alpha^2 + \beta^2)\right] f_i'\left(-\frac{d_5}{d_6}\alpha, -\frac{d_5}{d_6}\beta\right) \mathcal{H}\left(-\frac{d_5}{d_6}\alpha, -\frac{d_5}{d_6}\beta\right). \tag{8}$$

The light from both paths overlaps and interferes there. Photorefractive effect creates a volume phase hologram inside the PRC. We have the diffraction term been expressed as

$$t_i(\alpha,\beta) = \hat{e}_i^*(\alpha,\beta)\,\hat{u}_i'(\alpha,\beta), \tag{9}$$

where $*$ in superscript means complex conjugation. Here we ignore the crystal thickness and the angle, θ, made by the two paths for simplicity. But this ignorance does not effect the result for the associative recall [1].

Patterns are memorized by being input into the SLM in turn [1, 2]. When all the learning has been fulfilled, we can modify the above equation to obtain the final diffraction expression

$$t(\alpha,\beta) = \sum_{i=1}^{M} t_i(\alpha,\beta). \tag{10}$$

Further learning is permitted whenever it is needed. But the learning of one of the new memories will cause the previously stored memories' being gradually forgotten. We have to balance the tradeoff to get an optimal learning procedure, that is to justify the photorecording time length for each of the memories according to their position in the learning sequences.

4. Associative recall

Associative recall is carried on in the system with the upper path being blocked. Besides feeding an edge-enhanced pattern, we can also use an original pattern or its incomplete version as the input stimulus. Feeding an input stimulus, $s(x,y)$, into the SLM at place B or C, we have the light distribution just in front of the crystal as

$$\hat{s}(\alpha,\beta) = \exp\left[\frac{ik(d_5+d_6)}{d_6^2}(\alpha^2+\beta^2)\right] s(-\frac{d_5}{d_6}\alpha, -\frac{d_5}{d_6}\beta)\,\mathcal{H}(-\frac{d_5}{d_6}\alpha, -\frac{d_5}{d_6}\beta). \tag{11}$$

Passing through the crystal, effected by the recorded volume hologram, the light field just behind the crystal is given by

$$\hat{s}'(\alpha,\beta) = \hat{s}(\alpha,\beta)\,t(\alpha,\beta) \tag{12}$$

The intensity distribution on the output plane, i.e. the CCD plane, is simplified and given by

$$I(\xi,\eta) = \left|\sum_{i=1}^{M}\{s\cdot f_i'\}\,f_i(-\omega\xi,-\omega\eta)\right|^2, \tag{13}$$

where

$$\omega = \frac{d_1 d_3}{(d_1+d_2)d_4}. \tag{14}$$

This is similar to Eq. (1) except for the modulus squared. The intensity distribution at the output plane is the weighted summation of all the stored patterns. The weights of the weighted output are determined by the inner products of the input stimulus and the corresponding edge-enhanced patterns. If s is an original pattern, f_k, or a portion of it, the inner product of f_k and f_k' is relatively much higher than those between f_k and other memories. Hence high discrimination amongst the stored patterns is achieved.

A threshold on $I(\xi,\eta)$ results in

$$T[I(\xi,\eta)] = f_k(-\omega\xi,-\omega\eta). \tag{15}$$

If the input stimulus is an incomplete version of a memorized pattern, this thresholded output may not be an exact replica of the corresponding memory. A feedback process is needed to readdress the SLM with this thresholded output as the new input stimulus. The feedback enables the system to update its output, in which each loop yields a more similar output to that stored, until a stable output is achieved.

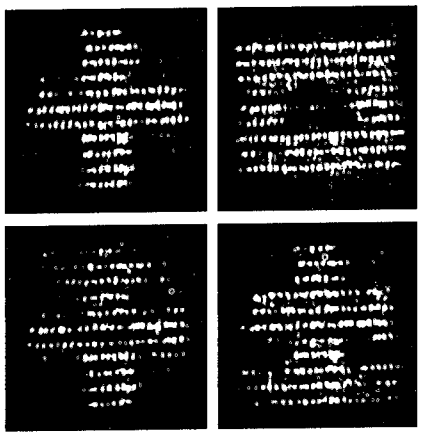
5. Experimental demonstration and computer simulation

5.1 Optical demonstration

Five 13 × 13-pixel Chinese characters, four as memory patterns and one as false input stimulus, were chosen for optical demonstration. The edge-enhancement was carried out in two steps. At the first step, the 0-valued pixels adjacent to a 1-valued pixel were set to -1 while the others were kept unchanged and secondly, the 1-valued pixels with no -1 valued near neighbors were set to 0 while the others were unchanged.

We calculated the weights for both the original bipolar Hopfield model (BH) [2, 3], where both the input stimulus, s, and the original pattern, f_i, take bipolar form, and our bipolar edge enhanced model (EE). An average improvement of discrimination from 73% (BH) to 80% (EE) was achieved. Discrimination is defined as the average ratio of difference in the weights for the expected memory and an unexpected memory to the weight of the expected memory. This system is equivalent to a high-order nonlinear one with an order of 1.3.

Figure 3(a) shows the recalled output where a stored pattern is used as the recall stimulus. Output threshold level is determined by the input energy. The output for a nonstored stimulus is shown in Fig. 3(b). There is a dramatic drop in the output energy, normalized to the input energy, when a nonstored pattern is input into the system. This implies that the system has a strong discrimination among the stored and nonstored patterns. It also suggests that this phenomenon can be used as a measure to decide whether the input is a memory of the system.



(a)

(b)

Figure 3: Recalled output when the input stimulus is (a) a stored pattern or (b) a non-stored pattern.

Partial information addressability was also investigated, where the partial information was the right half of the original pattern. The results indicate that the system can perform the partial association except for a bit of decrease of the output signal to noise ratio. When input stimuli are the edge-enhanced versions, the output is better than that in the first two cases. The edge enhanced versions of the patterns can be treated as phase coded address, so the system is addressable by both address and content information.

5.2 Computer simulation

We also performed a computer simulation to evaluate the retrieval error of the proposed learning algorithm. Most of the performance evaluation of a neural network is statistically carried out by computer simulation where the stored patterns are randomly generated. But a random pattern might be meaningless in practical applications, where patterns involved are obtained from the real world. So the storage capacity is lower than the simulation based on random patterns.

To do the statistical evaluation on the retrieval error of the proposed learning algorithm, we choose the frequently used Chinese character set, which consists of 3755 characters in 16×16 bitmap format. We simulated the relationship between the retrieval error and the number of the stored patterns, as shown in Fig. 4, averaged over

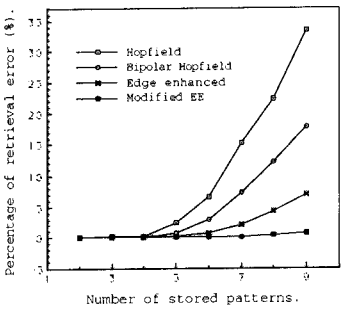


Figure 4: Retrieval error via number of stored patterns in computer simulation.

200 independent simulations where stored patterns were selected randomly from the 3755 character library. We calculated the cases for the Hopfield model (HM), its bipolar counterpart, *i.e.* the BH, and EE. Computer simulations show that the retrieval error for EE is much smaller than those of HM and BH. When the stored pattern number is 9, the error for the three are 33.5% (HM), 17.8% (BH) and 6.99% (EE), respectively. We also proposed a modified EE algorithm (ME), where the edge-enhancement is so carried out that for a 1-valued pixel its value will be reset to be the number of its 0-valued nearest neighbour pixels, and for a 0-valued pixel its value will be reset to the negative of its 1-valued nearest neighbour pixels. With the same simulation as above we find the corresponding retrieval error is now 0.732%.

The work of Dr. Xu-Ming Wang was supported by K. C. Wong Education Foundation on a Royal Society K. C. Wong Fellowship. The work of Mr. Jian Wang is supported by a K. C. Wong Scholarship.

References

[1] X.-M. Wang, G.-G. Mu, and Y.-X. Zhang, "Optical associative memory using an orthogonalized hologram," Opt. Lett. **16**, 100-102 (1991).

[2] X.-M. Wang and G.-G. Mu, "Holographic associative memory with bipolar features," Proc. Soc. Photo-Opt. Instrum. Eng. **1558**, 518-528 (1991).

[3] X.-M. Wang and G.-G. Mu, "Optical neural network with bipolar neural states," Appl. Opt. **31**, 4712-4719 (1992).

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

SINGLE RAIL OPTICAL ASSOCIATIVE MEMORY

Alain Bergeron[1,2], Henri H. Arsenault[2], Denis Gingras[1]
bergeron@ino.qc.ca, arseno@phy.ulaval.ca, gingras@ino.qc.ca
[1]National Optics Institute/Institut national d'optique
369 Franquet, Sainte-Foy,
Québec, Canada, G1P 4N8
[2]COPL, Université Laval,Sainte-Foy
Québec, Canada, G1K 7P4

INTRODUCTION

Associative memories were the first optical neural network architecture to be implemented[1,2]. Previous optical implementations have some drawbacks such as the need for an interruption of the light beam path between the different layers of the neural network due mainly by the lack of suitable optical thresholding units. As a consequence, most of the implementations so far used detectors coupled with light sources[3]. This solution however may cause data bottlenecks while requiring additionnal components.

Paek and Psaltis have also proposed an all-optical associative memory based on a double correlator architecture[4]. However, it lacked translation invariance due to the use of pinholes to suppress sidelobes of the correlation peaks. It also required the use of additional components, such as a powerful laser and expensive optical components, i.e., an image intensifier and a liquid crystal light valve, that makes this approach impractical.

A feed-forward associative memory implementation based on the double correlator architecture is proposed. In order to allow translation invariance, all the pixels of the input plane have to be analyzed. This may be acccomplished by means of an optoelectronic thresholder[5]. This module also allows an uninterrupted propagation path so only one HeNe laser source with a relatively low power of 35 mW is required to operate the associative memory. The holograms required to perform both the correlation and the association are computer-generated holograms with a relatively small number of degrees of freedom. In order to obtain good image quality and sufficient diffraction efficiency, the associative layer's CGHs are computed using the global iterative coding method[6].

ASSOCIATIVE MEMORY ARCHITECTURE

This associative memory is based on the double correlator architecture. Our implementation is shown in Fig. 1. An input image is presented to the first correlator. The choice of filter is crucial as it directly determines the signal that will be transmitted to the second layer of the neural net. The second layer simply performs the recall of the information stored in the second memory. The first filter requires a good discrimination and a sharp peak of correlation to approximate a delta function. The phase-only filter satisfies those two requirements ans is used in the first layer of the neural network.

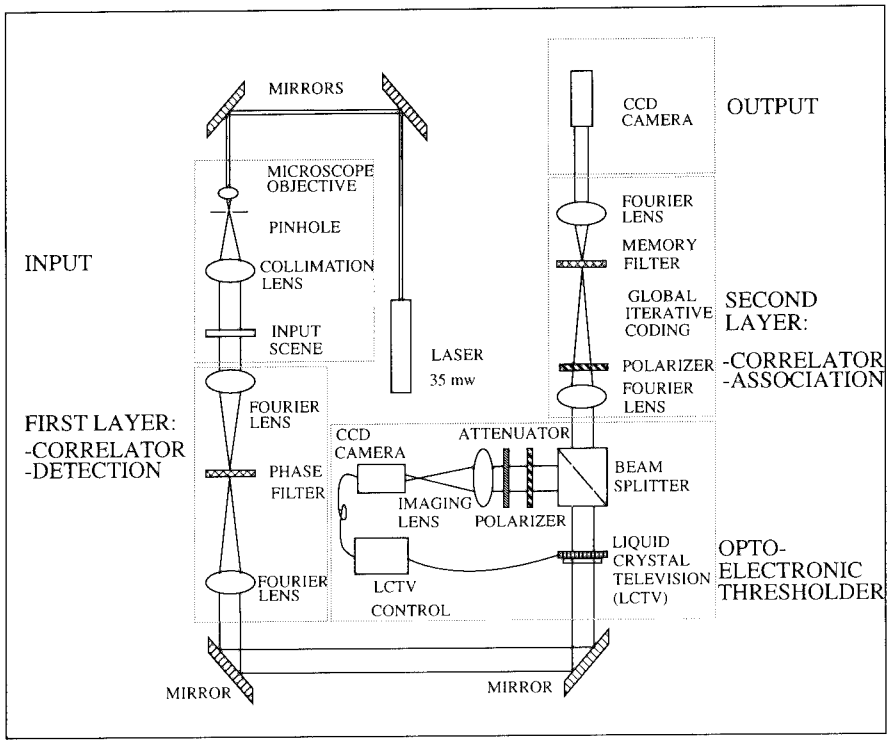


FIGURE 1. Optical memory architecture.

An optoelectronic module was inserted between the two layers to perform the thresholding operation. This module taps the output of the first layer and modulates the output of an electrically addressed spatial light modulator in order to attenuate small signal intensities while keeping higher signal values unchanged. This kind of thresholder

allows the light path to be uninterrupted so that the photons arriving at the system output come directly from the input laser without seeing any electronic relay. Part of the photons are detected by the tapping section of the threshold module[5].

It should be noted that the threshold is applied to each point of the output plane of the first layer over a spatial extent given by the resolution of the modulator. Such an operation allows to track the input object even as its location changes. The thresholder maintains the translation invariance property exhibited by the first correlator. The thresholder not only eliminates the cross correlations but also helps to sharpen the detection peak by attenuating the sidelobes with intensities lower than the threshold value. The delta function is more closely approximated after the thresholding operation.

After this process, the recall is carried out by the second layer. In order to obtain a good reconstruction quality, a CGH is calculated with the help of the global iterative coding[6] (GIC). This coding allows the fast computation of the CGH and a good reconstruction quality. Furthermore, the use of the GIC allows to detect, in the first diffraction order, more than eight times the energy detected with conventional coding.

EXPERIMENTAL RESULTS

An optical associative memory corresponding to the set-up of Fig. 1 was built. The filters were recorded on high-resolution photographic film with the help of a laser writer. The input images were binary images of 256X256 pixels representing standard security symbols. The first filter had dimensions of 256X256 pixels, the dimensions of the second filter were only of 128X128 pixels. This small number of degree of freedom required the use of a coding scheme scheme with good performance on small supports, such as the GIC. A single 35 mW He-Ne laser was used to perform the experiments.

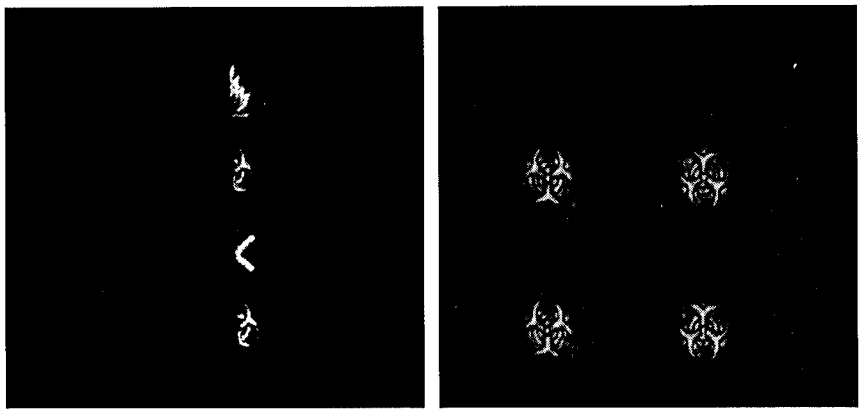


FIGURE 2. Occluded input by 50 % (left) and corresponding output (right).

The system behavior when a degraded object is presented at the input is illustrated in the Fig. 2. The objects to be detected were the two biological hazard symbols, whereas the fire and wheat symbols have to be discarded. In this experiment, 50% of the input image was occluded. The system recognized the object, and the association was correct. The amount of degradation allowed by the system strictly depends on the choice of the filter in the first layer. A filter showing good discrimination capabilities will allow the neural network to be robust. Note that only one cycle had to be performed to obtain the output.

CONCLUSION

A continuous optical associative memory was implemented. Experimental results show a good behavior of the system in the presence of unwanted objects and of degraded targets. The neural network can be the basis of a system allowing an input scene to be cleaned from unwanted noise or objects while the desired occluded objects could be reconstructed. The architecture is versatil because it can easily be modified by changing the filters in the first and in the second layers of the neural network.

ACKNOWLEDGEMENT

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Fonds pour la formation des chercheurs et l'aide à la recherche (FCAR) program of Québec and the JSTF program of the Canadian Ministry of External Affairs.

REFERENCES

1-N. H. Farhat, D. Psaltis, A. Prata and E. Paek, "Optical Implementation of the Hopfiel Model", Applied Optics, vol. 24, no. 10, pp. 1469-1475, 1985.

2-B. Kosko, "Adaptive Bidirectional Associative Memories", Applied Optics, vol. 26 no. 23, pp. 4947-4960, 1987.

3-S. Jin-Yuan, Z. Yan-Xin, "Optical Pattern Recognition System Based on a Winner-Take-All Model of a Neural Network", Optical Engineering, vol. 32, No. 5, pp. 1053-1056, May 1993.

4-E. Paek and D. Psaltis, "Optical Associative Memory Using Fourier Transform Holograms", Optical Engineering, vol. 26 No. 5, pp. 428-433, May 1987.

5-A. Bergeron, H. H. Arsenault, E. Eustache, D. Gingras, "Optoelectronic Module for Winner-Take-All Operations in Optical Neural Network", to be published in Applied Optics.

6-A. Bergeron, H. H. Arsenault, J. Gauvin, D. Gingras, "Computer-Generated Holograms Improved by a Global Iterative Coding", Optical Engineering, vol. 32, no. 9, pp. 2216-2226.

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

High Efficiency Photorefractive Storage for Multimode Phase Conjugate Resonators

C. C. Chang and D. R. Selviah

Department of Electronic and Electrical Engineering
University College London, Torrington Place, London WC1E 7JE, UK
Tel: +44-(0)71-3877050 Fax: +44-(0)71-3874350
E-mail: dselviah@eleceng.ucl.ac.uk

Abstract. Two phase conjugate mirrors, used during recording, improve the diffraction efficiencies of angularly multiplexed holographic gratings. These were used to demonstrate a phase conjugate resonator with 15 modes using non-degenerate multi-beam induced phase conjugation.

1. Introduction

There has been considerable interest in the use of resonators incorporating photorefractive holographic memories with phase conjugate mirrors to give optical feedback for use in optical associative memories[1-3]. Difficulty has been experienced in realising such resonators in practice due to the poor diffraction efficiencies of the holograms. In this paper we propose and demonstrate a novel technique in which two phase conjugate mirrors are used during the recording to store holograms with high diffraction efficiencies suitable for use in a resonator. We go on to demonstrate such a phase conjugate resonator with 15 modes.

Consider the recording of a single grating in a thick (~0.7cm) photorefractive crystal. During recording, the refractive index modulation, Δn, is determined by $\Delta n = \Delta n_{sat}(1-e^{-t/\tau_r})$[4], where Δn_{sat} is the saturation value of refractive index modulation, and τ_r is the recording time constant. The refractive index modulation magnitude decays exponentially with depth into the crystal since the writing beams are absorbed progressively with depth of penetration. In our technique we generate two phase conjugate beams from the remnants of the two writing beams succeeding to completely penetrate through the crystal. The beams are generated using two phase conjugate crystals. One is in the "Cat"[5] configuration and phase conjugates the object beam. The other crystal is "induced"[6,7] to enable it to phase conjugate complex reference beams which may have components propagating over a wide range of angles. In this way both the reference beam and object beam (no matter how complex) are returned to retrace their original paths but in the opposite direction with the same phase as they had on their incoming journeys. So the returning beams add constructively to the original writing beams. In addition, the grating that the phase conjugate beams create is exactly in phase with the previous grating, and therefore the latter gets enhanced and sustained. However, the amplitude of the returning conjugate beams decays exponentially from the back to the front of the crystal in the opposite sense to the initial writing beams. This offsets, to a some extent, the original non-uniform depth profile of refractive index modulation, resulting in better utilisation of the crystal volume and a higher diffraction efficiency.

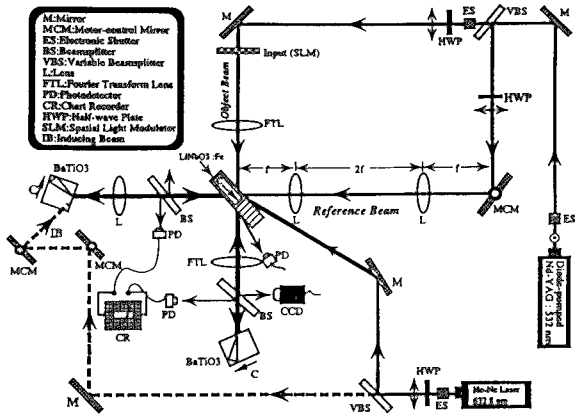


Fig. 1 Schematic showing experimental layout for high efficiency photorefractive storage.

2. Experiment and Results

In our experiments (Fig. 1) we used an ADLAS 400mW diode pumped, frequency doubled Nd:YAG laser. Both reference and object beams were plane and collimated The total exposure intensity was 56mW/cm^2 on the crystal surface and both beams had an equal intensity and diameter. The size of the region illuminated was a circle of diameter 1.5mm. Each beam was angled at θ=45^o to the normal and was extraordinarily polarised (in the plane of the figure). The photorefractive crystal, grown by Shanghai Institute of Ceramics, at the Chinese Academy of Science, was Y-cut lithium niobate with 0.05% weight of iron doping resulting in a 34±0.2% transmissivity. Lenses were used to focus the wide beams emerging from the iron doped lithium niobate into the small barium titanate crystals (5mm cubes) used as the phase conjugate mirrors. The reflectivities of the self-pumped crystals with incident intensity of 28 mW/cm^2 was 30% and 35%. The first crystal was grown in China and supplied by Photox Optical Systems, UK and the second by Hughes in USA. The diffraction efficiency was measured during the recording by using a red probe beam from a He-Ne laser (λ=632.8nm).

The experimental results (Fig. 2) show a plot of the diffraction efficiency as a function of time. Curve 1 was obtained experimentally without phase conjugate mirrors (PCMs) while curve 2 was obtained with phase conjugate mirrors. Both curves have similar forms showing initial growth and relaxation back to a lower more stable value with time. The curves grow at the same rate for the first 15 seconds and, thereafter, the curve using PCMs, grows at a much faster rate than that without PCMs and attains a higher efficiency of 22.1% after 220 seconds while that without PCMs has a maximum of 17.9% after 179 seconds. This represents an improvement in diffraction efficiency of 4.2%. Both curves tend to a diffraction efficiency of about 16.5% at longer exposure times of 9 mins so to achieve the highest efficiencies the exposure time must be set correctly. The curves do not appear to fit well to a model based on the usual inverse exponential saturation, but simply for the sake of comparison, we have calculated the writing time constants assuming the maximum refractive index modulation is achieved in each case at the maximum diffraction efficiency using the same method as that used in reference 8 but assuming that it is a function of the total writing power. This gives a writing time constant of 30.4s for the curve using PCMs and 42.2s for the curve without. This experiment was repeated several times giving similar improvements in diffraction efficiency. The increased rate of growth of diffraction efficiency also observed in each case may be due to the increased illumination arising from use of the PCMs. This is beneficial as multiple multiplexed recordings can be made in a shorter time. The increased maximum diffraction efficiency may be due to the better refractive index modulation uniformity with depth.

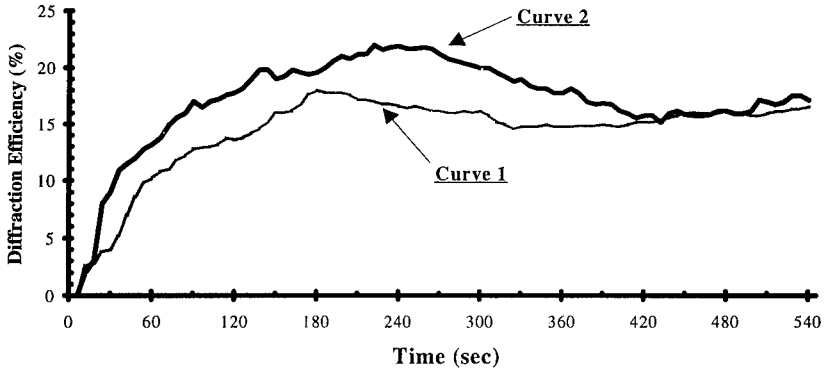


Fig. 2 The growth of holographic grating diffraction efficiency with time, curve 1 uses conventional recording, curve 2 uses the novel phase conjugate mirror recording configuration.

Fig 3 shows the experimental layout of our multimode phase conjugate resonator which incorporates a set of angularly multiplexed gratings stored by the aforementioned technique, within the cavity. In order to store several holographic gratings it is necessary to use angled reference beams over a large angular range since there must be a minimum angular change between holographic recordings in order to reduce crosstalk. Consider the case of 15 beams with 2.5mm diameter beam widths, as an example. The power entering the first PCM (shown on the left of figure 3) immediately after recording and on initial playback was 0.3mW. After phase conjugation the power was 20μW giving a 11.4% reflectivity (neglecting Fresnel losses) which is less than the 35% measured earlier using a single beam. The holographic gratings were not fixed and so were gradually erased by the reading beam. This can be clearly seen in figure 4 which shows the optical power phase conjugated by the first PCM as measured after the 50% beam splitter. Zero time corresponds to when all of the 15 green beams illuminated the phase conjugate barium titanate crystal. After one minute a red (He-Ne) inducing beam was switched on to illuminate the crystal with a power of 30μW at an angle of 60 degrees, a distance of 1mm from the corner of the crystal. The curve clearly shows that the phase conjugated power suddenly increases from zero to 20 μW in 2.5 mins. At this time the inducing beam was turned off and the subsequent exponential decay is due to the erasure of the holographic gratings in the lithium niobate. After 25 mins when the power had decayed to 4 μW it suddenly dropped to 0.3 μW. This indicates the thresholding effect of the PCM. Figure 5 shows the 15 simultaneously phase conjugated beams. We observed each eigenmode of the system as a spot of light. It was confirmed that the system was resonating by chopping all of the beams in front of each PCM in turn. If the spot was extinguished when each set of beams were blocked then this confirms that signals are circulating around the system. The CCD camera showed a large spot of light with detail multimode structure.

3. Conclusion

In conclusion, this paper introduces a novel recording technique, in which two crystals of barium titanate are used as phase conjugate mirrors during recording to enhance the resulting holographic diffraction efficiency. We demonstrated a phase conjugate resonator having fifteen modes incorporating an angularly multiplexed set of gratings recorded using our novel technique. Finally, we demonstrated phase conjugation of fifteen beams by means of an inducing beam of another wavelength (non-degenerate multi-beam induced phase conjugation).

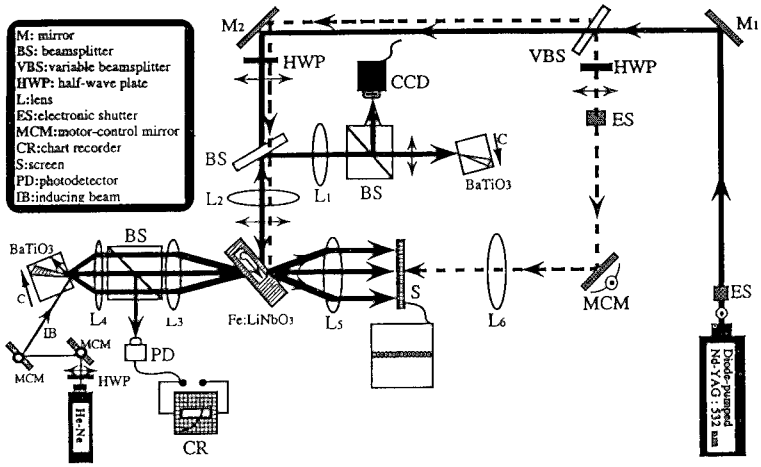


Fig. 3 Multimode phase conjugate resonator.

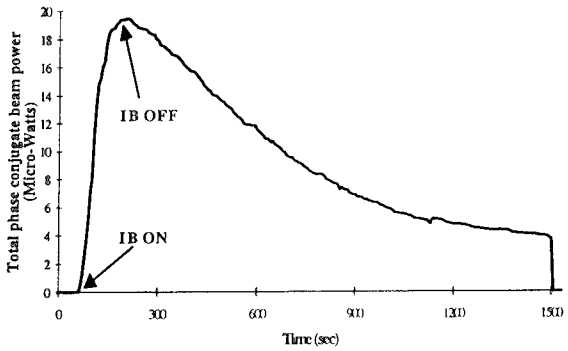


Fig. 4 Total optical power phase conjugated by the first phase conjugate mirror rises abruptly when the inducing beam (IB) is switched on and then decays exponentially.

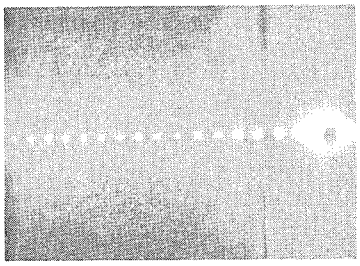


Fig. 5 Fifteen modes simultaneously supported in the phase conjugate resonator as observed on screen S in figure 3.

Acknowledgement

The authors wish to thank Laser Lines Ltd for supplying the laser and the National Science Council of Taiwan, ROC for financial support.

References

1. B. H. Soffer, G. Dunning, Y. Owechko, and E. Marom 1986 Optics Letters **11** 118-120.
2. A. Yariv, and S-K Kwong 1986 Optics Letters **11** 186-188.
3. A. Yariv, and S-K Kwong, and K. Kymura 1986 Appl. Phys. Lett **48** 1114.
4. E. S. Maniloff and K. M. Jonhson 1991 J. Appl. Phys. **70** 486-491.
5. J. Feinberg 1982 Optics Letters **7** 486-488.
6. H. F. Yau, Y. P. Tong, C. C. Chang and N. C. Wang 1992 Optics Comm. **98** 457-462.
7. D. R. Selviah and C. C. Chang 1993 OSA International Topical Meeting on Photorefractive Materials, Effects and Devices, Kiev PD6.1-PD6.4.
8. S. Tao, D. R. Selviah and J. E. Midwinter 1993 Optics Letters **18** 912-914.

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical learning neural network with reversal input superposing technique

Masahiko Mori, Masaru Kondo[1], Yoshio Hayasaki[2], Ichiroh Tohyama[1] and Toyohiko Yatagai[1]

Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba, 305 Japan
[1]University of Tsukuba, 1-1 Tennodai, Tsukuba, 305 Japan
[2]The Institute of Physical and Chemical Research, 19-1399, Koeji, Nagamachi, Aoba-ku, Sendai, 980 Japan

Abstract. Reversal input superposing technique is applied to an optical learning neural network. Optical neural networks introduced the technique are not necessary to use negative weights and subtraction, and inherently constructed all optical systems.

1. Introduction

In general neural network models, the synaptic weights and the weighted sum are the finite real values. It means that some of them have minus values. In the other hand, to implement optical neural networks, usually the output of neurons and the values of weights are represented by optical intensities, transmittances or refrectivities. It is hard to realize minus values under these representations. Therefor, in some optical neural network systems, the twice weights, separated plus and minus weights are used. After each weighted sum operation, these values are subtracted by electronic circuits[1,2]. In this approach, twice pixels of weight matrix and electronic subtract circuits are required and the system speed is limited by these circuits. In another case, the weights are added with a bias and the thresholds of each neuron are changed according to the total values of inputs. In this case special hardwares to realize controlling the each threshold are needed, and an all optical system is impossible.

We proposed the reversal input superposing technique (RIST) to avoid above problems[3,4]. In this paper, we report an optical learning neural network with RIST. The optoelectronic neural system based on RIST is implemented and learning capability is realized by using a Pockels readout optical modulator (PROM) device[5].

2. Principle of RIST

Consider general discrete neural network models consisting of M neurons, which receive the same input signals $X = \{x_1,...,x_i,...,x_N\}$ from N input neurons, and emit respective output signals $V = \{v_1,...,v_j,...,v_M\}$. Let w_{ji} be the synaptic weight from the i-th input neuron to j-th output neuron. The j-th weighted sum of input signals u_j and the output v_j are written as

$$u_j = \Sigma_j w_{ji} x_i - h_j \qquad (1)$$

and

$$v_j = f(u_j), \qquad (2)$$

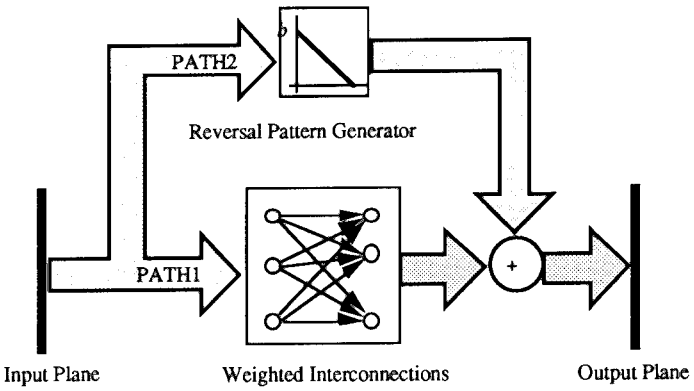


Fig.1 Schematic block diagram of RIST in a neural network.

where h_j is the j-th value of a set of the thresholding values $H = \{h_1,...,h_j,...,h_M\}$ and **f** is a nonlinear output function. Introducing a constant α, Eqs.(1) and (2) are rewritten as

$$u'_j = \Sigma_i(w_{ji} + \alpha)x_i + \alpha\Sigma_i(1 - x_i) \qquad (3)$$

and

$$v_j = f'(u'_j), \qquad (4)$$

where $f'(y) = f(y-N\alpha)$. The thresholding values $H = \{h_j\}$ are neglected because they can be treated as same as one of the synaptic weights. The constant α is a bias of the weights w_{ji} and is determined to $\alpha > -\min(w_{ji})$, where $\min(w_{ji})$ denotes the minimum values of a set of the weights $W = \{w_{ji}\}$. Hence, the biased synaptic weights $\{w_{ji}+\alpha\}$ of the first term in Eq.(3) are necessary to be positive. The input x_i is ranged from 0 to 1 in many neuron models. In these cases, $\{1-x_i\}$ of second term in Eq.(3) are the reversal values of the inputs x_i in the range of 0 to 1, and positive. Under these conditions all terms of Eq.(3) are all positive and easily realized by optical techniques. **f'** is the nonlinear function with the thresholding value which is slid the constant $N\alpha$ from **f** in Eq.(2).

Figure 1 shows a schematic block diagram realizing Eq.(3). The lower path (PATH1) of the diagram performs a weighted sum operation which is the first term of Eq.(3) with input signals. The upper path (PATH2) makes reversal inputs of the second term. The calculated results on the output plane are applied to the devices which perform the nonlinear function with the slid thresholding value.

3. Optical Implementation and Experiments

We apply polarization encoding method in order to get reversal inputs. The polarization encoded patterns and two orthogonal analyzers are easily made the input and reversal patterns simultaneously.

The experimental system of an optoelectronic neural network with RIST is shown in figure 2. He-Ne laser light (633nm) is used at the recalling process. Input signals of input

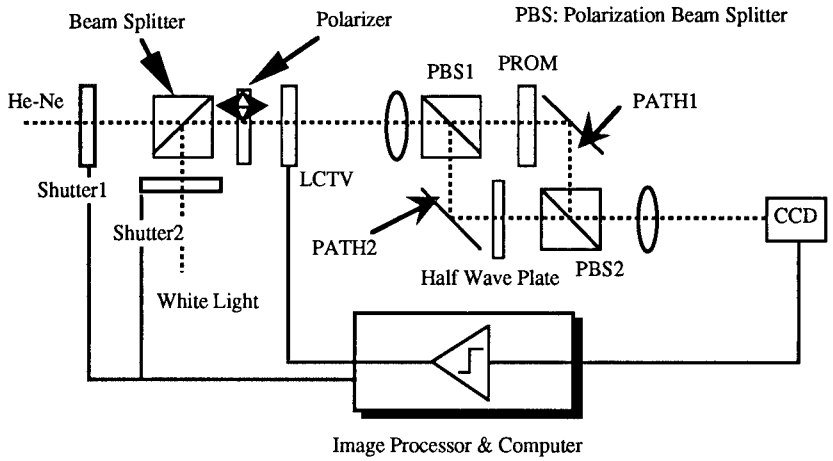


Fig.2 Experimental setup of the optical learning neural network with RIST.

neurons are polarization encoded by a liquid crystal TV (LCTV) panel which has not an analyzer. After the polarization beam splitter (PBS1), input patterns pass a PROM device which displays the weights and weighted sum operations are performed (PATH1). On the other hand, reversal patterns pass another arm (PATH2) and superposed with the calculated results at PATH1. The superposed results are detected by a CCD camera. The balance of these signals are adjusted by controlling the angel of a half wave plate in PATH2. The nonlinear function with a slid threshold value are calculated in a computer.

To achieve the learning capability on the system, a PROM device is located in PATH1. The PROM is written or erased by white light according to the values which are displayed on the LCTV. Therefor it is not necessary to align the weight matrix on the PROM and input signals on the LCTV. The modification values of the weights are calculated from the teaching signals and the recalling results by a computer and displayed on the

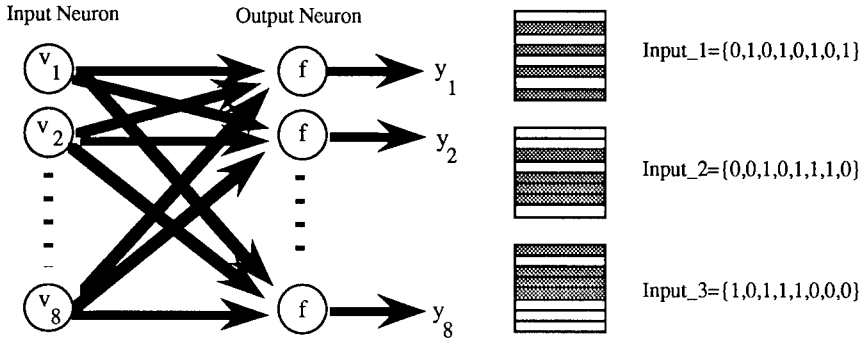


Fig.3 (a) 8-8 neuron network structure and (b) learning patterns (8 neuron, 3 patterns).

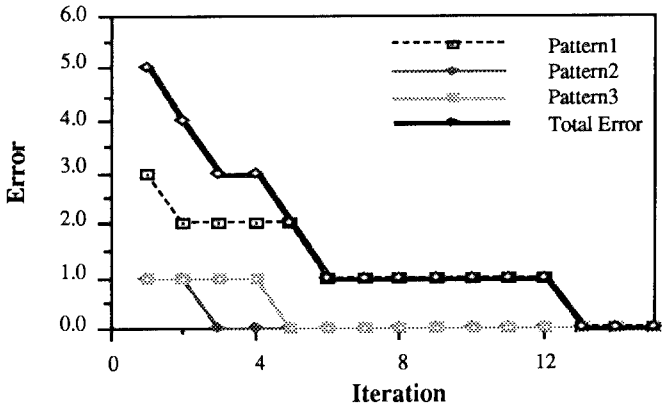


Fig.4 Learning curves of each stored pattern and the total error.

LCTV through an image processor, and are written and/or erased on the PROM.

On the experimental system, a single layer structure with 8 input and 8 output neurons is realized (Fig.3(a)). 3 input patterns of 8 bits are stored on the system according to the orthogonal learning rule (Fig.3(b)). The weights on the PROM start from random values nearby the constant α. Here, we determine α with the half value of the dynamic range of the device. The learning curves are shown in figure 4. By 13 iterations, the weights went to correct values and the error went down to 0.

4. Conclusion

Reversal input superposing technique is applied to an optical learning neural network. Optical neural networks introduced the technique are not necessary to use negative weights and subtraction. The learning capability was applied to the optoelectronic system using a PROM device and verified at the performance with an 8-8 neurons network.

The authors thank S.Ishihara, H.Yajima and T.Hidaka of Electrotechnical Laboratory for useful discussions and encouragements. Thanks are also due to Y.Osugi of NGK Insulators, Ltd. for providing the PROM.

References

[1] Ohta J, Tai S, Oita M, Kuroda K, Kyuma K and Hamanaka K 1989 Opt. Lett. **14** 844-6
[2] Kasama N, Hayasaki Y, Yatagai T, Mori M and Ishihara S 1990 Jpn. J. Appl. Phys. **29** L1565-8
[3] Hayasaki Y, Tohyama I, Yatagai T, Mori M and Ishihara S 1993 Tech. Digest OC'93 (Palm Springs, USA) 140-3
[4] Hayasaki Y, Tohyama I, Yatagai T, Mori M and Ishihara S 1994 Appl. Opt. **33** 1477-84
[5] Osugi Y, Abe H, Honda A, Hamajima A and Toyoda S 1990 Conf. Record of OC'90 (Kobe, Japan) 21-2

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Feature Extraction Preprocessing and Optical Neural Network

Ichiroh Tohyama, Masaru Kondoh, Toyohiko Yatagai,

Institute of Applied Physics, University of Tsukuba, 1-1 Tennoudai, Tsukuba, 305, Japan

Masahiko Mori,

Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba, 305, Japan

Yoshio Hayasaki

The Institute of Physical and Chemical Research, 19-1339 Koeiji, Nagamachi, Aoba-ku,Sendai, 980 Japan

Abstract. A new architecture combined with an optical Fourier transformational processor and a three-layer optical neural network proposed. This hybrid system can reduce a number of neural elements.

1. Introduction

Optical neural network computing is of interest in terms of massively parallel computing. Many optical neural network systems based on the matrix-vector multiplication architecture,[1] however, are limited in a scale of parallelism, because of spatial resolution of optical systems and spatial light modulators. Besides, when the bipolar synaptic weights are represented by optical intensity, they are divided into positive unipolar weights and negative unipolar weights[2],[3]. This means that twice number of the synaptic weights are needed. To solve these problems, as an alternative method, we introduce optical feature extraction techniques. The feature extraction can map the data onto a new feature space. The data on the feature space are subjected to input data for an optical neural system. If we can design a suitable mapping system, the ability , the function, and the size of the optical neural system could be reduced.

We propose here a new type of neural computing with a preprocessor for feature extraction to reduce the number of neural units. A variety of optical feature extraction techniques have been discussed, which include analog[4] and digital methods[5]. The Fourier transform, the Hough transform, the polar coordinate transform, the scale-invariant transform and so on are typical examples of analog methods. In this paper, we use a Fourier Transform optical system as a preprocessor, because of its data compression and shift-invariant properties.

2. Feature Extraction and Hybrid Computing

Figure 1 shows a concept of a hybrid optical neural network using the Fourier transform which is one of feature extraction preprocessing. We note that a Fourier transform of a real object is shift invariant and point-symmetrical. A arbitrary upper half region of the Fourier transformed image is connected with positive synaptic weights, the other region negative synaptic weights. This procedure reduces half the number of weights to display on a spatial modulator.

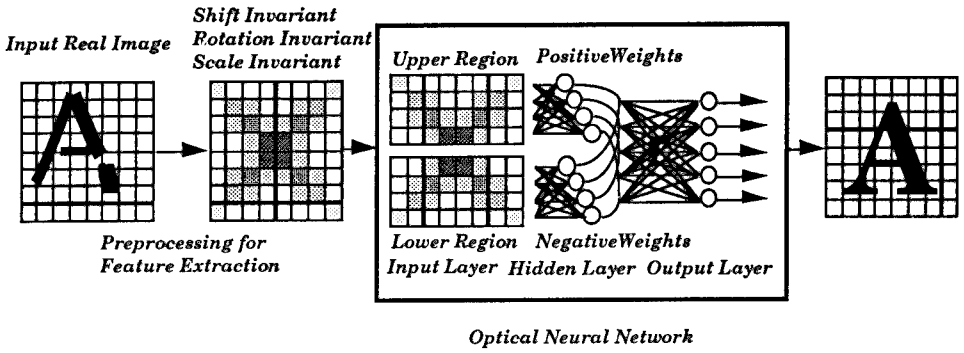


Fig. 1. Concept of a hybrid optical neural computing.

3. Computer Simulation

Computer simulation for the feature extraction preprocessing and the neural network with the back-propagation learning algorithm is performed. We consider a three-layers neural network consisting of 64x64 input layer neurons, 64 hidden layer neurons, and 5 output neurons, to recognize 26 alphabet characters of 64x64 pixels.

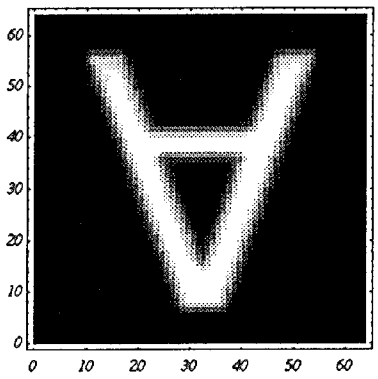


Fig. 2 Input character example "A".

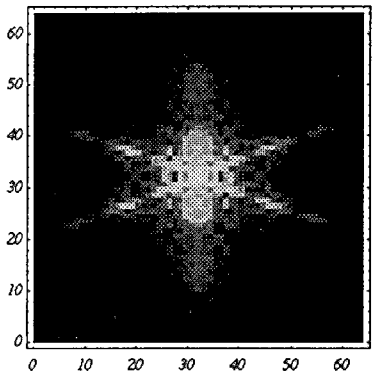


Fig. 3 FT of the character "A".

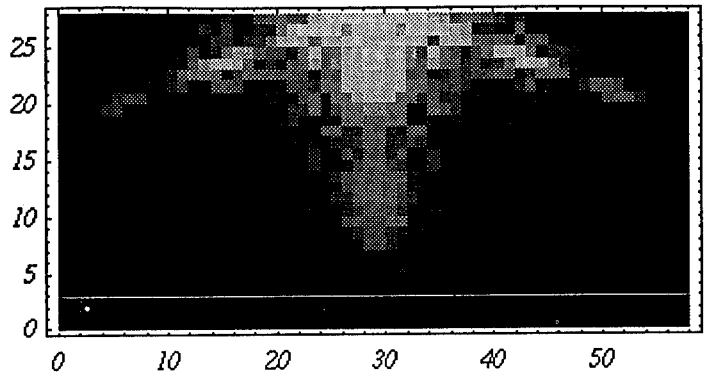


Fig. 4 An arbitrary region of transformed patters .

Figure 2 shows an example of inputs of 26 alphabet characters. These characters are represented by a gray scale of 8bit. At first, we calculate Fourier transforms of input characters. Figure 3 shows a Fourier transformation of the character of Fig. 2. Transformed patterns are normalized, and they are also represented by a gray scale of 8bit. Because of the point symmetry of Fourier transformed images, an arbitrary half part of transformed patters as shown in Fig. 4 are used for feature extraction.

Figure 5 shows the change of the squared sum error of 26 characters vs iteration times of learning in the cases when inputs for the neural network are;

(1) Original character images, (Fig.2)
(2) Fourier transformed images, (Fig.3)
(3) Arbitrary half of Fourier transformed images. (Fig.4)

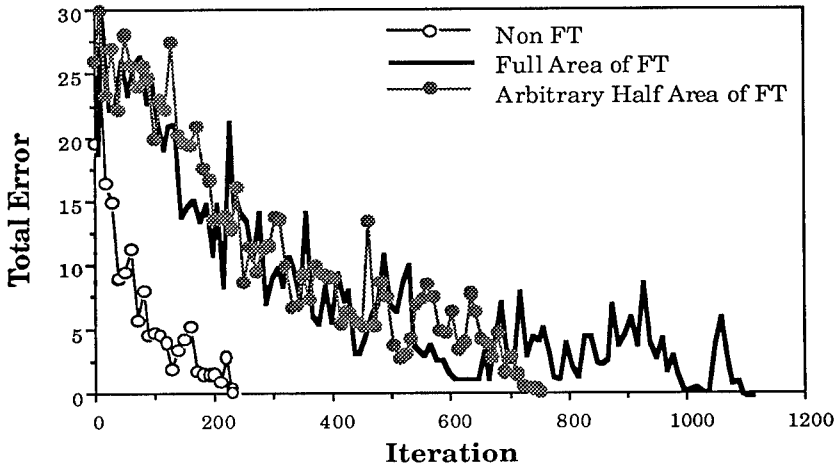


Fig. 5 Change of the squared-sum error.

The conventional optical neural networks are not able to recognize position shifted input patterns. Because of the shift invariance in the Fourier Transform, we have estimated the recognition ability for non-shifted and for shifted input images using the values of weights and offsets which are obtained when the learning is completed. Figure 6 shows the error rates of recognition for each input pattern with and without random shift within 4 pixels. The results indicate reduction of the error rates due to shift invariance of the Fourier transform.

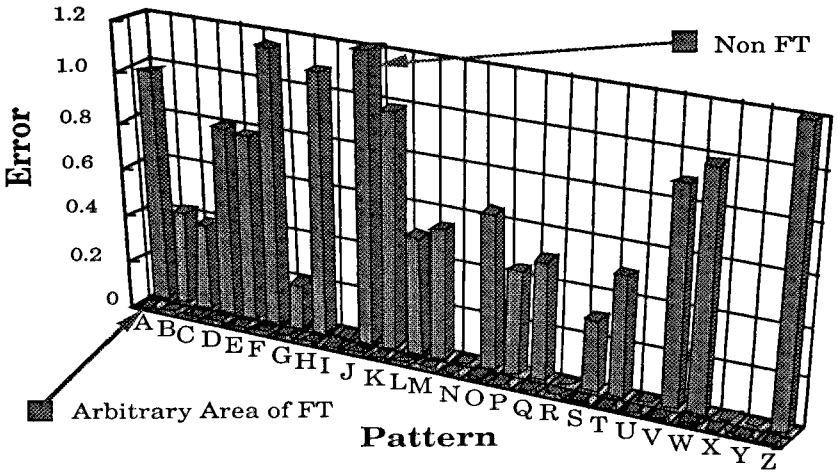


Fig. 6 Error rates for the shifted each input.

4. Conclusion

We propose a new type of neural computing with a preprocessor for feature extraction. The Fourier transform of a real object, which is one of feature extraction preprocessing, is shift invariant and point-symmetrical. The feature extraction preprocessing we use here has two advantages; reduction of the number of neuron units and shift-invariance for input patterns.

Acknowledgment

The authors thank S. Ishihara, H. Yajima and T. Hidaka of Electrotechnical Laboratory for their useful discussions and encouragements.

References

[1]N. Farhat, D. Psaltis, A. Prata and E. Peak, Appl. Opt., **24** (1985) 1469.

[2]N. Kasama, Y. Hayasaki, T. Yatagai, M. Mori and S . Ishihara, Jpn. J. Appl. Phys., **29** (1990) 1565.

[3]I. Tohyama, Y. Hayasaki, T. Yatagai, M.Mori, and S. Ishihara, Proc. SPIE **1806** (1992) 271-278.

[4]K. Sasaki, D. Casasent and S. Natarajan, Proc. SPIE, **1384** (1990) 228-233

[5]T. Yatagai, Opt. Lett., **11**(1987) 270-272

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical Digit Recognition based on Kohonen Maps

M. Barge, R. Chevallier, E. Curatu† and A. Maruani

Département Images, ENST, 46 rue Barrault, 75634 PARIS Cedex 13, France
Tel : +33-1-45817320, Fax : +33-1-45813794, email : barge@ima.enst.fr

† Polytechnical Institute of Bucharest, Applied Optics Laboratory, Spl Independentei No 313, R-77206 Bucharest, Romania

Abstract. We present an optical implementation of the recognition stage of a Kohonen-like neural network based on a supervised learning algorithm. The set up includes Diffractive Optical Elements to both store and address the weights.

1. Introduction

Our work involves the study of optical implementation of neural networks. We describe and show the feasibility of the optical implementation of the recognition stage of a neural network based on Kohonen feature maps. Indeed, this stage can be understood as a multichannel correlation and thus is well suited to an optical implementation. The learning stage is done numerically. The set up uses two binary phase holograms : a 16x16 microlens array to achieve the replication of the Fourier transform of the inputs and a Computer Generated Hologram (CGH) that contains the weights of the computed Kohonen map.

2. Learning Algorithm

We use a supervised learning algorithm based on Kohonen self-organizing feature maps for pattern recognition. While in the classical Kohonen algorithm the map is self-organized by the inputs according to their probability density function, pattern recognition applications need some supervision to map similar stimuli in terms of their context and not in terms of a known vector distance metric. We adopt here an idea previously used for semantic map organization [1] adapted to pattern recognition [2]. The basic motivation is to organize the map by the patterns and their association targets simultaneously (*eg* in our case the class which the input pattern belongs to). A by-product of this process is that the class-labelling of neurons on the map emerges during the learning

phase, while in the classical Kohonen maps the labelling is achieved after the learning stage.

We present in figure 1 the feature map used for our experiments. It shows a 8x8 neuron map ; for each neuron, a 16x16 pixel image shows the weights between the input layer and the neuron itself, the intensity of each weight being proportionnal to its darkness. This representation exhibits several properties of the model :

- Each neuron is specialized ; hence, it is possible to "recognize" patterns on the map.
- Inside each class, the great differences between input patterns induce the fact that several neurons belong to this class while the small variations are taken into account by the fact that the weights are calculated in such a way that the patterns on the map appear as "blurred" images.
- There is a topology on the map since all the neurons belonging to one class are close to each other.
- One can observe the clustering of the neurons on the map according to the classes.

The learning database includes 750 patterns of an actual French postal code database which was already segmented. Each pattern is a 16x16 binary pixel image.

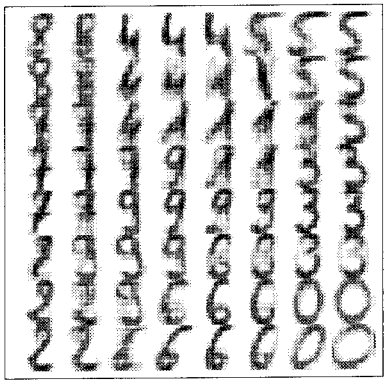


Figure 1. 8x8 feature map used for our experiments (see the clustering of the map according to the classes).

3. Numerical Studies

The recognition stage and its application for handwritten digit recognition have been numerically implemented. More precisely, two points relevant to an optical implementation were studied :

- while the selection of a neuron, for classical Kohonen feature maps, is made by minimizing a distance between the input and the weights, it is easier for an optical version to maximize an inner product. Those two conditions are strictly equivalent under the assumption that the weights and the inputs used for the learning stage are normalized.

- another point of matter is to study the influence on the results during the recognition stage when using a binary hologram instead of an actual weights map.

The studies were done using a 1024x1024 binary CGH that was computed using mainly an error diffusion algorithm. They show for a 50 pattern test database (5 patterns for each class) the validity of the computation of the CGH and the feasibility of the optical implementation.

4. Experimental Work

For our set up, the size of the inputs is 16x16 pixels and the Kohonen map has 8x8 neurons so that 256x64 weights must be coded using a 1024x1024 pixel off-axis hologram. For the classification stage that requires 64x10 interconnections a micro-computer is used. This set up is a slightly improved version of a previously-used one for the implementation of an associative memory [3][4]. It is specially designed to achieve a multiplexing of connection holograms involving a special coding scheme for the weights matrix that we called Frequency Multiplexed Raster (FMR). This scheme allows one to encode a 4D matrix on a plane. Moreover, for an optical implementation, it leads to a standard spatial filtering set up except that the inputs and outputs must be sampled in a special manner. The holographic filter contains the FMR encoded synaptic array.

An holographic element (microlens array or Dammann gratings for instance) enables the achievement of the sampling scheme of the inputs. The first Fourier lens of the set up achieves the Fourier Transform (FT) of the so-sampled input and allows one to replicate the FT of each input pattern. This component and the connection hologram are binary phase elements processed by current means : contact lithography using e-beam designed masks and reactive ion etching of a silica substrate. This technological part of our work was done by the laboratories of the CNET-Bagneux, France.

The output image diffracted by the connection hologram is detected using a CCD camera. The image measures the activation of the neurons for a given input. Since each activation is the inner product of the input and one neuron (whose coordinates are its weights since it is fully connected to the input), the result can be seen as the correlation peak of the input and each of the patterns shown on figure 1. Hence, this optical implementation can also be understood as a multichannel correlator except that the filter is computed in such a way that the results for all the channels are obtained using a simple Fourier lens.

For the classical Kohonen algorithm, a special kind of non-linearity called "Winner-Takes-All" is used which means that only the most active neuron is retained. In our case, the difference between neuron activation is first emphasized by quadratic detection. The decision is then made by computer and suitable software.

We present in figure 2 two experimental results for two different 0's. The most important activations of the neurons are in the "right" part of the map (see figure 1 for the references) showing that these results are in good agreement with what was expected.

5. Conclusion

As a conclusion, we present results for three steps of an optical implementation of a Kohonen-like neural network : software implementation for the learning algorithm, numerical simulations for the optical recognition step and experimental results. The results

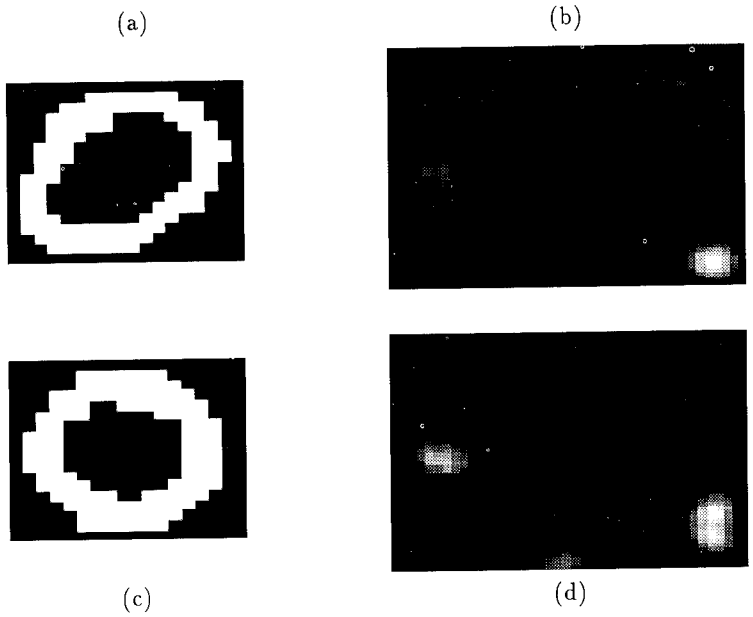


Figure 2. Experimental results (b) and (d) for two different inputs (a) and (c). The activations are to be compared with the feature map of the figure 1.

we get with actual postal code patterns show the validity of our encoding scheme and prove the possibility of achieving any kind of interconnections between two planes.

We are still working on the experimental part of this work in order to improve, in terms of dynamic range, the quality of the storage of the data for the connection hologram.

6. Acknowledgements

We thank J.L Oudar, J.F. Palmier and S. Sainson from the CNET-Bagneux for the technological realization of the components.

References

[1] S. Midenet and A. Grumbach, 1990, *Proceedings of INNC -Paris*, 773-776

[2] Y. Idan and R. Chevallier, 1991, *Proceedings of IJCNN'91-Singapore*, 2576-2581

[3] R. Chevallier, 1990, *International Journal of Optical Computing*,**1**, 71-87

[4] K. Heggarty and R. Chevallier, 1992, *Optics Communications*, **88**, 91-95

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Overcoming noise in a fast neural network

R. P. Webb

BT Laboratories, Martlesham Heath, Ipswich IP5 7RE, UK.
Email: webb_r_p@bt-web.bt.co.uk

Abstract. Noise will limit the size and speed of optoelectronic neural networks. The pattern recognition performance of an experimental network was measured against signal-to-noise ratio. Accuracy was improved by redundancy in the network configuration.

1. Introduction

The computing power that will be demanded by many future applications, for example the control of future packet-based telecommunications switches or searches through massive databases, is likely to grow beyond the capacity of conventional methods. Some of these tasks are amenable to neural processing and could therefore be accelerated by parallel operation on appropriate hardware. Optoelectronic networks could further expand the scale of parallel systems, but they must also be fast. Otherwise the extra parallelism will just compensate for lost speed, rather than extending performance beyond the bounds electronics.

In high-speed optical systems receiver noise becomes dominant; hence this paper first examines in general how noise limits network size and then how it affects performance in a particular instance. The effects of introducing redundancy into the network configuration have been modelled and confirmed experimentally. The example problem was binary pattern recognition, chosen both for ease of testing and because it resembles some simple operations on packet headers.

2. Noise limit to network size

Consider a network with fan-out from a modulator array to a weight matrix and fan-in to a detector array, with the total optical power evenly divided between all the connections. The detection circuits must be able to resolve the smallest significant change in the value of one weight, so the maximum number of connections that can be supported is:

$$MN = S\Gamma/hP_{res}$$

where M and N are the numbers of modulators and detectors, S is the total optical power available, Γ is the optical efficiency, h is the number weight levels and P_{res} is the smallest

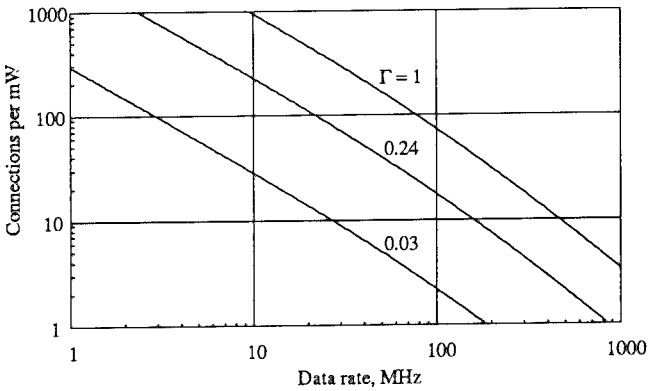


Fig. 1. Network size / data rate with 100 weight levels. Γ is overall optical efficiency.

resolvable optical power. P_{res} was calculated for a state-of-the-art communications receiver assuming an error rate of 10^{-3}, but with a large area photodiode (7 pF) and a high mean level of illumination ($100 \times P_{res}$). $\Gamma = \eta x_d w_d$, where x_d and w_d are the usable transmittance ranges of the modulators and mask respectively and η is the net efficiency of the other components. In the experimental system, $\eta = 0.3$, $x_d = 0.12$, and $w_d = 0.8$, giving $\Gamma = 0.03$. Network size/unit power was plotted against data rate for this value of Γ assuming 100 weight levels (Fig. 1). Two more curves were plotted, for ideal modulators or a laser array ($x_d = 1$, hence $\Gamma = 0.24$) and for an ideal system ($\Gamma = 1$). The results indicate that a large network ($MN >$ 1000, say) could run at speeds > 100 MHz powered by a semiconductor laser source, provided that an overall optical efficiency of 0.24 or better were attained.

3. Pattern recognition theory

A single neuron is sufficient for binary pattern recognition if its connection weight to each expected one in the input pattern is +1, and to each expected zero is -1. But with large patterns the threshold would be high and the weight tolerances tight, so in the experimental system margins were improved by dividing the target into sub-patterns with a neuron to recognise each. Some redundancy was introduced by having two overlapping sets of three sub-patterns and arranging a second-layer neuron to respond whenever more than four of the six sub-patterns were detected.

By assuming that the receiver noise at its input is gaussian, the probability that a neuron will be *on* may be calculated (Fig. 2). Each first-layer neuron has an initial probability of 0.5 which tends asymptotically to 1 or 0 as the signal-to-noise ratio is increased, according to whether or not the correct sub-pattern is present. The *on* probability of the output neuron first falls below 0.5 even with the target pattern before beginning to increase with signal-to-noise ratio. This is because the output neuron is susceptible both to errors induced by the noise at its own input and to errors occurring in the first-layer. For signal-to-noise ratios above 1.6, however, the majority vote system makes the output more accurate than the individual first-layer neurons. The output probability was also calculated for patterns close to the correct target. Patterns differing by just one bit are more likely to be misclassified than is the target, but discrimination could be improved with a greater degree of redundancy (more sub-pattern overlap).

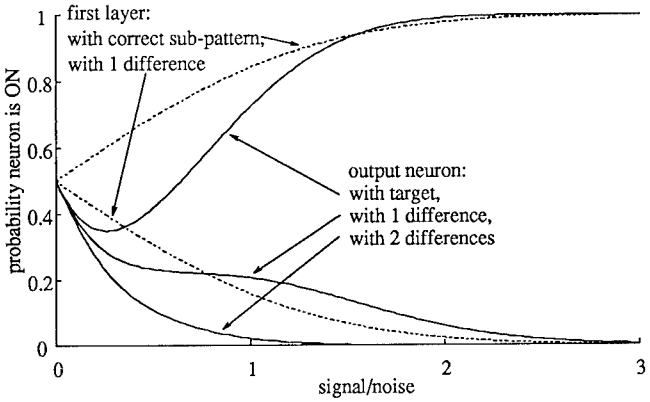


Fig. 2. ON probability / signal-to-noise for first-layer and output neurons.

4. Pattern recognition with an optoelectronic network

The experimental system [1] used to test these results employed an array of high-speed multi-quantum-well modulators [2] as inputs (Fig. 3). They were illuminated by a single 1.5 μm semiconductor laser using a computer-generated hologram [3] to divide the beam. A second hologram provided fan-out to the weight mask. Weights to recognise a target pattern were calculated and encoded on photographic film. Fan-in was achieved by imaging the weight mask onto a detector array feeding comparators. Bipolar weight values were mapped to unipolar mask transmittances using a shared threshold. The mapping scheme also cancelled offsets caused by the limited contrast of the modulators and mask and required only a small increase in network size. Because the neuron thresholds were part of the weight mask rather than voltages applied to the comparators, the laser power could be adjusted to control the signal-to-noise ratio while leaving the relative values of all neuron inputs unchanged.

A stream of input patterns including the target was presented to the network at 8.4 Mword/s. The frequencies of correct recognitions of the target and the six sub-patterns were measured as laser power was increased (Fig. 4). As predicted, the performance of the whole network exceeded the average of its first-layer neurons for input powers over 2 mW. Then

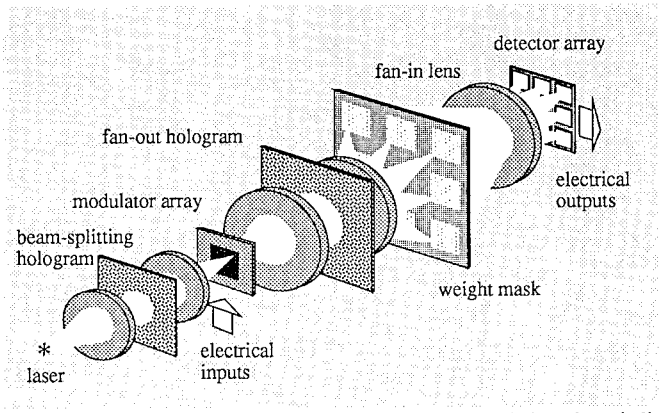


Fig. 3. Experimental system. The beams corresponding to one input only have been indicated.

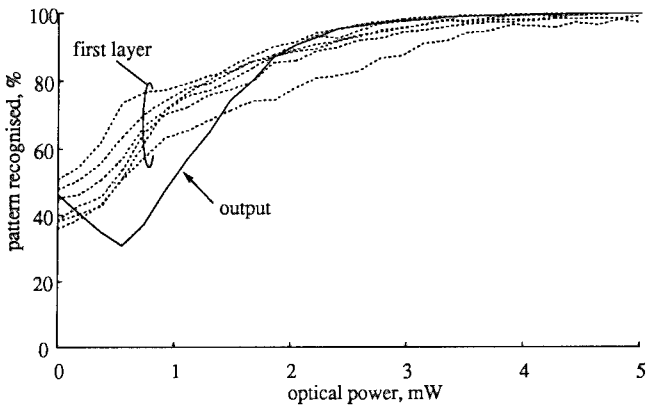


Fig. 4. Measured output and first-layer recognition frequencies / laser power.

responses to other patterns close to the target were recorded (Fig. 5). False recognition frequencies were also broadly as predicted, only a few patterns caused significant numbers of false recognitions at full power, probably because of inaccuracies in weight fabrication.

5. Conclusions

Networks large and fast enough to challenge electronic systems will operate close to the noise limit if powered by semiconductor sources, as they must be to allow compact integrated systems. However, this work demonstrates that redundancy in the network can reduce the frequency of noise-induced errors.

References

[1] Webb R P 1993 *International Journal of Neural Systems* **4** 435-444
[2] Rejman-Greene M A Z, Scott E G and McGoldrick E 1988 *Electronics Letters* **24** 1583-1584
[3] Dames M P, Dowling R J, McKee P and Wood D 1991 *Applied Optics* **30** 2685-2691

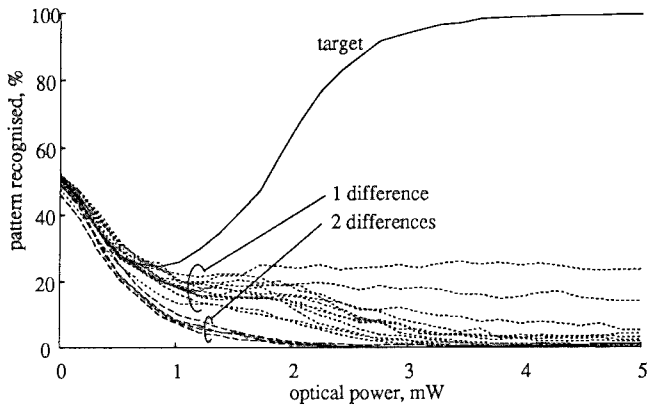


Fig. 5. Measured recognition frequencies with target and other patterns.

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical self-routing complex neural network using mixed negabinary numbers

Liren Liu, Guoqiang Li, Lan Shao, and Yaozu Yin

Information Optics Laboratory, Shanghai Institute of Optics and Fine Mechanics, Academia Sinica, P.O.Box 800-211, Shanghai 201800, P. R. China

Abstract. A general neural unit with only a complex-valued matrix depicting the integrated synaptic connection and lateral interaction is proposed. Various connections can be self-routed by the routing code contained in the input. Based on the mixed negabinary system, an inner-product algorithm is developed to calculate the complex matrix-vector multiplication without carries, signs, and decimal points. An incoherent optical correlator with spatial coding is suggested to execute the algorithm in parallel.

1. Introduction

The biological neurons receive input connections from different sources: external input signals through the incoming synaptic connections, and lateral interaction within the same area[1]. And various neural network model were developed[2]. Recently, a general neural unit was suggested as a building block [3], which is a cascade of a synaptic connection layer and a lateral interaction layer.

In this paper, we suggest that the combined stimuli states of neuron are represented by complex numbers: the synaptic connections are indicated by real numbers; the lateral interaction by imaginary numbers; and the integrated effect by a pair of routing code. Thus, a general neural unit with only a complex-valued matrix is established to calculate the integrated connections. Thus, various connections can be self-programmed by the routing code contained in the inputs. To deal with the complex-valued and multi-valued matrix-vector multiplication, we develop a digital inner-product algorithm on the basis of the real and imaginary decomposition and the mixed negabinary system. It has the features as no carries, no signs, no decimal points, and simple pre- and post-processing. Correspondingly, an optical architecture of incoherent optical correlation with spatial digital coding is suggested, which can execute the algorithm in parallel.

In the previous optical neural networks[4,5], the analog and bipolar matrix-vector multiplications were usually utilized. In optical matrix processing, various digit systems and architectures were used to improve the accuracy[6-9], but they have the problems on carries, signs, decimal points, and pre- and post-processing.

2. Self-routing general neural unit

The complex-valued general neural unit is shown in Fig.1. The states of the input signals are denoted by a complex-valued vector, $[v_j(K_1+jK_2)]$. $v_j \in \{1,-1\}$ is the input stimuli and K_1, $K_2 \in \{1,0,-1\}$ assemble the routing code. The complex-valued matrix consists of the real and the imaginary part as $[T^R_{ij}]+j[T^I_{ij}]$. The real and imaginary parts of the resulting vector are

$$[v'^R_i] = K_1[v_i^R] - K_2[v_i^I] \quad and \quad [v'^I_i] = K_2[v_i^R] + K_1[v_i^I],$$
$$where \quad [v_i^R] = [T_{ij}^R][v_j] \quad and \quad [v_i^I] = [T_{ij}^I][v_j]. \qquad (1)$$

By using different routing code the resulted real and imaginary parts both have nine connection functions (see Table 1). Thus either one can be selected as the outputs. In the figure, the imaginary outputs are used as the terminals.

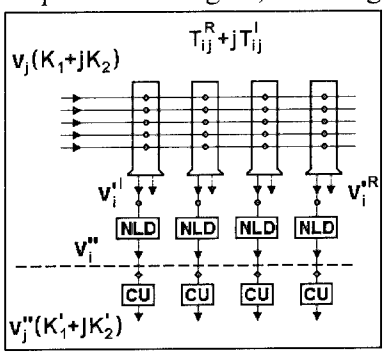


Figure 1

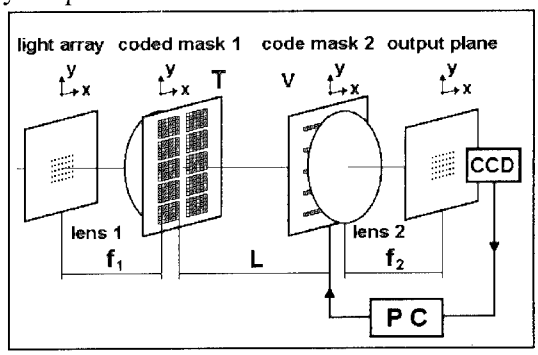


Figure 2

Table 1. Self-routed connection functions

K_1	K_2	$[v'^R_i]=K_1[v^R_i]-K_2[v^I_i]$		$[v'^I_i]=K_2[v^R_i]+K_1[v^I_i]$	
1	1	$[v^R_i]-[v^I_i]$	(1)	$[v^R_i]+[v^I_i]$	(10)
	0	$[v^R_i]$	(2)	$[v^I_i]$	(11)
	-1	$[v^R_i]+[v^I_i]$	(3)	$-[v^R_i]+[v^I_i]$	(12)
0	1	$-[v^I_i]$	(4)	$[v^R_i]$	(13)
	0	0	(5)	0	(14)
	-1	$[v^I_i]$	(6)	$-[v^R_i]$	(15)
-1	1	$-[v^R_i]-[v^I_i]$	(7)	$[v^R_i]-[v^I_i]$	(16)
	0	$-[v^R_i]$	(8)	$-[v^I_i]$	(17)
	-1	$-[v^R_i]+[v^I_i]$	(9)	$-[v^R_i]-[v^I_i]$	(18)

The results are then processed by nonlinearity devices (NLD)[2]. The coding units (CUs) are used to index the real signals with the routing code for the cascaded complex neural unit.

The complex-valued general neural unit itself pictures a variety of self-organized neural networks such as the Hamming network[2] and the Kohonen

network[1]. It can be simplified to the Hopfield associative memory[2] by a real matrix, or to the laterally interconnected neural network[1] or the MAXNET network[2] by a complex matrix with the real part unitary matrix. Furthermore, it is able to represent multi-layer networks by the cascade of units.

The synaptic matrix is denoted by $[T^{sy}_{ij}]$ (NxM, M≥N) and the lateral interaction matrix $[T^{la}_{ij}]$ (NxN). The learning rule is

$$[T^R_{ij}] = [T^{sy}_{ij}] \quad and \quad [T^I_{ij}] = ([T^{la}_{ik}] - [I])[T^{sy}_{kj}]. \tag{2}$$

3. Negabinary inner-product algorithm

The negabinary number system has the radix of -2[10]:

$$a = \sum_{n=-L}^{P} a_n(-2)^n \qquad (a_n \geq 0), \tag{4}$$

Where the digit a_n may be a positive fraction in its mixed form.

The addition of two real numbers is simply the additions of their digits. For the multiplication, we suggest a vector format to negabinary numbers. The multiplicand is denoted by the negabinary vector of **a** and the multiplier vector is written in a reversed and shifted negabinary vector as $\mathbf{b}_n$:

$$\begin{aligned} \boldsymbol{a} &= \{a_K, a_{K-1}, \cdots, a_1, a_0, a_{-1}, \cdots, a_{-L+1}, a_{-L}\}, \\ \boldsymbol{b_n} &= \{b_{n-K}, b_{n-K+1}, \cdots, b_{n-1}, b_n, b_{n+1}, \cdots, b_{n+L-1}, b_{n+L}\}. \end{aligned} \tag{5}$$

The nth digit of multiplication result is the inner-product $c_n = \mathbf{a} \bullet \mathbf{b}_n$.

A complex number can be decomposed into a real part and an imaginary part. The negabinary multiplicand vectors of the real and imaginary elements of the complex matrix are denoted as $\mathbf{T}^R(i,j)$ and $\mathbf{T}^I(i,j)$, and the negabinary multiplier vectors of the inputs of $K_1 v_j$ and $K_2 v_j$ as $v^1(j)$ and $v^2(j)$. Thus, the nth real digit and imaginary digit at the ith row of the resulting negabinary vector are

$$\begin{aligned} v_n^R(i) &= \sum_j \boldsymbol{T}^R(i,j) \cdot \boldsymbol{v_n}^1(j) + \frac{1}{2} \boldsymbol{T}^I(i,j) \cdot \boldsymbol{v_{n-1}}^2(j), \\ v_n^I(i) &= \sum_j \boldsymbol{T}^R(i,j) \cdot \boldsymbol{v_n}^2(j) + \boldsymbol{T}^I(i,j) \cdot \boldsymbol{v_n}^1(j). \end{aligned} \tag{6}$$

4. Incoherent optical correlation architecture

An incoherent optical correlator with spatial coding is used to implement the complex-valued and multi-valued matrix-vector multiplication in parallel[11], which results in a correlation between two coded masks similar to the representation by the negabinary inner-product algorithm. The conversion to negabinary numbers, the thresholding necessary for neural nonlinearities, and the iterative feedback are executed by a PC computer (see Fig.2).

In the experiment, we use a laterally inhibited Hopfield model. The input complex signal and the complex matrix are coded as Fig.3 and Fig.4. The predicted and experimental outputs are shown in Fig.5 and Fig.6.

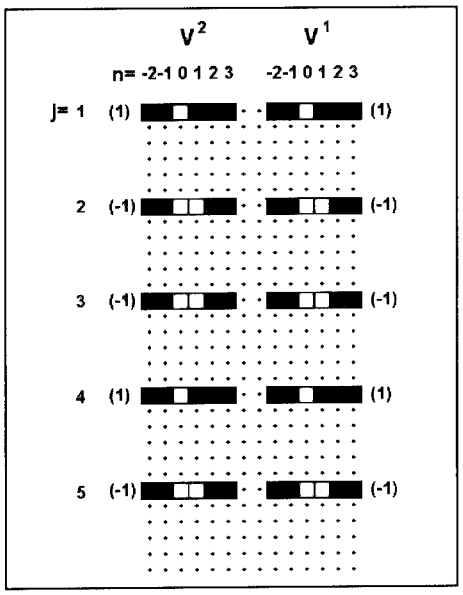


Figure 3

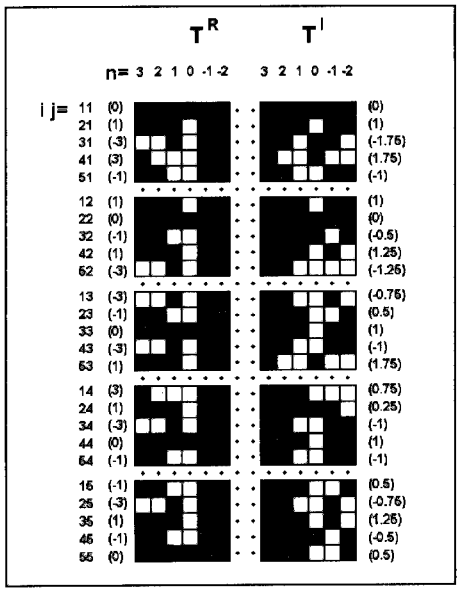


Figure 4

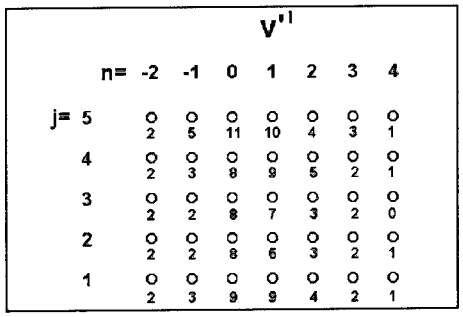
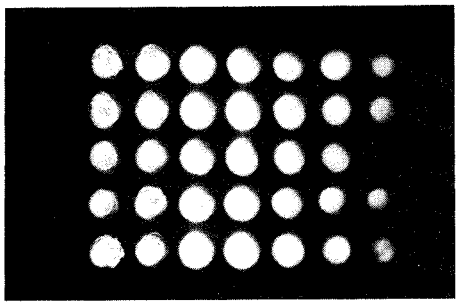


Figure 5



Acknowledgment

This work was supported by the National Natural Sciences Foundation of China.

References

[1] T.Kohonen, Appl. Opt. 1987, **26** (1987) 4910.
[2] R.P. Lippmann, IEEE ASSP Magazine April (1987) 4.
[3] I.Aleksander and H.B.Morton, Electron. Lett. **27** (1991) 1776.
[4] D.Psaltis and N.Farhat, Opt. Lett. **10** (1985) 98.
[5] C.C.Mao and K.M.Johnson, Appl. Opt. **32** (1993) 1290.
[6] R.P.Bocker, Opt. Eng. **23** (1984) 26.
[7] B.H.Soffer, Y.Owechko, E.Maron, and J.Grinberg, Appl. Opt. **25** (1986) 2295.
[8] D.Casasent and B.K.Taylor, Appl. Opt. **24** (1985) 1476.
[9] E.Baranoski and D.Casasent, Appl. Opt. **28** (1986) 5351.
[10] M.P.de Regt, Comput. Design, **6** (1967) 53.
[11] L.Liu, G.Li, and Y. Yin, Opt. Lett. (in press).

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Neural networks with unipolar weights and normalized thresholds

Jan Stanisław Bródka and Bohdan Macukow

Warsaw University of Technology, Institute of Mathematics, Plac Politechniki 1, 00-661 Warsaw, Poland

Abstract. This paper is concerned on problem of simulation of one neural network by the other one. Formal definitions of neural network and simulation are proposed. Two theorems are presented, the first one shows, that each neural network can be replaced by a neural network with non-negative connection weights only; the second one shows, that each neural network can be replaced by a neural network with the connection weights equal 0 or 1.

1. Introduction

Artificial neural net models or simply *neural networks* have been studied for many years. The models are based on our present understanding of biological nervous system and are composed of many nonlinear computational elements - neurons. These neurons are arranged in patterns (or layers) and connected via weights that are usually adapted during use to improve the net performance [1]. Neural networks are finding many areas of applications. Although they are particularly well-suited for applications related to associative recall such as content-addressable memories, neural networks can perform many other applications ranging from logic operations to the solution of optimization problems.

Simplified model of neuron was formulated in 1943 by McCulloch and Pitts [2]. In their model, neuron is a binary element, which has two stages: active (1) and non active (0). Connection weight may be positive (excitatory), negative (inhibitory) or zero (no connection). Neuron becomes active only when weighted sum of its inputs exceed a special value called threshold. Active neuron sends an output signal. This paper is based on McCulloch and Pitts model.

Although artificial neural networks (also called the neurocomputers) are much simpler than biological ones, they have some advantages over classical von Neumann computers. Great research effort is made to implement neural nets in hardware. Some problems arises with implementation in one network connections with various weights, especially with positive and negative weights [3,4].

2. Definitions

Neural network is the triple (V,E, P), where (V,E) is oriented graph with weights, V is the set of nodes (neurons), E is connection weight function, E: $V\times V \to \mathbf{R}$: (i.e. E(x,y) is the connection weight from neuron x to y), and P is threshold function, $P:V \to \mathbf{R}$, (i.e. P(x) is the threshold of node x).

The stage of the network is the function f, f: $V \to \{0,1\}$, f(x) is the stage of neuron x (1 - active, 0 - non-active). Set of all network stages is denoted by F (i.e. $F=\{f: V\to\{0,1\}\}$). The network changes its stage from f to f* according to *transformation rule T* (i.e. f*=T(f)), where T is defined by

$$T(f)(x)=\begin{cases}1 & \text{if } \sum_{y\in V} E(y,x)*f(y)\geq P(x)\\ 0 & \text{otherwise}\end{cases}$$

The expression $\sum_{y\in V} E(y,x)*f(y)$ is often called net(x).

This definition of the transformation rule corresponds to the synchronous Hopfield binary model [4].

The *action* of the net with the beginning state f_0 we will called infinite sequence of states $f_0, f_1, \ldots, f_n, \ldots$ where $f_{n+1}=T(f_n)$.

Network $S = (V,E,P)$ *is simulated* by the network $S^* = (V^*,E^*,P^*)$ when functions: $g: V \to V^*$ and $h: N \to N$, $h(0)=0$, (h - ascending) exist, that for all actions $(f_0,f_1,f_2,\ldots)$ of network S, actions $(f_0^*,f_1^*,f_2^*,\ldots)$ of network S^* satisfying the condition $(\forall x \in V)(\forall k \in N) f_k(x) = f^*_{h(k)}(g(x))$ exist.

Remarks:

1) Function g defines that neuron x of network S corresponds to neuron g(x) of network S^*.

2) Function h defines, that instant k in network S action corresponds to instant h(k) in network S^* action.

3) More intuitively we can say: network S is simulated by network S^* if the condition *neuron* x *is active in network* S *in the instant k when, and only when, neuron* g(x) *is active in network* S^* *in the instant* h(k) is satisfied for all neurons x and any instant k.

3. Networks with only non-negative connection weights

Theorem 1.

Any network S can be simulated by the network S^* with non-negative connection weights only.

Proof. Network S^* with the desired features will be constructed. At first we introduce simplified notation related to networks S and S^*.

Elements of network S are denoted as follows:

1) Set of nodes (neurons) of network S: $V = \{x_1,x_2,\ldots x_n\}$.

2) Thresholds of nodes of network S: p_i - threshold of node x_i, i = 1,2,...,n (i.e. $p_i = P(x_i)$).

3) Connections weights in network S: $w_{ji} = E(x_i,x_j)$ - connection weight from x_i to x_j (may be negative).

Elements of network S* are denoted as follows

1) Set of nodes (neurons) of network S^*: $V^* = \{x_1^p,x_1^d,x_2^p,x_2^d,\ldots,x_n^p,x_n^d\}$. Notice, that two nodes x_i^p and x_i^d in S^* replace one node x in S. Node x_i^d is called *the primary node* and x_i^d is called *the dual node*.

2) Thresholds of nodes of network S^*: p_i^p - threshold of node x_i^p, p_i^d - threshold of node x_i^d, i=1,2...,n.

3) Connection weights in network S^*: $w_{ij}^{pp} = E^*(x_i^p,x_i^p)$ - connection weight from x_i^p to x_i^p, $w_{ij}^{pd} = E^*(x_i^p,x_i^d)$ - connection weight from x_i^p to x_i^d, $w_{ij}^{dp} = E^*(x_i^d,x_i^p)$ - connection weight from x_i^d to x_i^p, $w_{ij}^{dd} = E^*(x_i^d,x_i^d)$ - connection weight from x_i^d to x_i^d.

Before the formal construction of the network S^* we describe intuition on which S^* is based. Node x_i^p is active in S^* exactly when x_i is active in S, and on the contrary, x_i^d is active in S^* when x_i is non-active in S. Then, in any instant, exactly one of nodes x_i^p, x_i^d is active in S^*, and activity of x_i^p in S^* corresponds to activity of x_i in S.

Now we describe formal construction of network S^*.

Each node of network S corresponds to two nodes of network S^* then each connection of network S corresponds to four connections of network S^*. Their weights are defined as follows:

If $w_{ij} \geq 0$, then $w_{ij}^{pp} = w_{ij}^{dd} = w_{ij}$, $w_{ij}^{pd} = w_{ij}^{dp} = 0$, and if $w_{ij} < 0$, then $w_{ij}^{pp} = w_{ij}^{dd} = 0$, $w_{ij}^{pd} = w_{ij}^{dp} = -w_{ij}$. Notice, that these weights are non-negative.

Definition of thresholds is more complicated.

Let $u_i = \sum_{j \in W_i^-} w_{ji}$, where $W_i^- = \{j: w_{ji} < 0\}$ i.e. W_i^- is the set of all nodes which connections with negative weight are coming from (to x_i), and u_i is the sum of these negative connection weights. $d_i = \sum_{j \in W_i^+} w_{ji}$, where $W_i^+ = \{j: w_{ji} \geq 0\}$ i.e. W_i^+ is the set of all nodes which connections with non-negative weight are coming from (to x_i), and d_i is the sum of these non-negative connection weights.

Threshold values are defined as follows: $p_i^p = p_i - u_i$ for primary nodes, $p_i^d = -p_i + d_i + \varepsilon_i$ for dual nodes, where ε_i is a positive number smaller than the minimal distance from summarized input at x_i which does not activate x_i, to threshold p_i (ε_i which satisfy this condition exists because of binary character of nodes).

This ε_i guarantee that if in original network $net(x_i) = p_i$ then in simulated network only node x_i^p becomes active, without ε_i both nodes x_i^p and x_i^d will be active.

Lemma 1.

Network S^* defined above simulates original network S for function g defined as $g(x_i) = x_i^p$ and function h defined as $h(k) = k$. Proof of this lemma is presented in [4].

Some thresholds in network S^* constructed in the Theorem 1 may be negative. It may occur even when in original network S all thresholds have been positive. Network S^* can be modified to have all threshold positive, by modifying connection weight between nodes x_i^p and x_i^d. More details are presented in [4].

Corollary 1.

Any network S is simulated by network S^* with only non-negative connection weights and only positive node thresholds.

4. Networks with weights equal 0 or 1 and normalized thresholds

Theorem 2.

Any network S with only non-negative connection weights and only positive node thresholds is simulated by network S^* with all connection weights equal 0 or 1 and all node thresholds positive and identical.

Proof. Network S^* with the desired features will be constructed.

This construction follows some steps.

1) At first notice that any network can be transformed to network with weights and thresholds represented by rational number.

2) Identical threshold of all nodes equals p can be reached by appropriate scaling of connection weights. New weights are defined by $w_{ji}^* = \frac{p}{p_i} * w_{ji}$.

3) After steps 1, 2 all weights and thresholds are non-negative rational numbers, then they can be transformed to natural number by multiply by appropriate constant.

Network obtained in the step 3) instead of initial network will be the base of following steps. For simplicity, threshold and weights obtained in step 3) will be denoted appropriately p and w_{ji}.

4) Any connection (with an integer weight w_{ji}) in network S is replaced by w_{ji} pairs of connections, each with weight equals 1, and by w_{ji} intermediate nodes. Set of intermediate nodes introduced to replace connection from x_j to x_i of origin network is denoted Y_{ji}. Set of all intermediate nodes is denoted Y.

5) All nodes, as well intermediate, should have threshold p, moreover individual impulse arrived from original network node should activate appropriate intermediate node.

Set $Z = \{z_1, z_2, \ldots, z_{p-1}\}$ of additional nodes is introduced to satisfy this condition. Elements of Z are always active and they are connected to each other. Connections from nodes' $z_1, z_2, \ldots, z_{p-1}$ arrive to every intermediate node $y \in Y$.

Generally, network S^* is composed of nodes corresponding to original network nodes X^*, intermediate nodes Y and additional nodes Z.

6) It should be described which way network S^* simulates network S, i.e. functions g, h and initial stage f_0^* should be defined.

Function g is defined as $g(x_i) = x_i^*$ and function h is defined as $h(k) = 2*k$, this means that network S^* acts *two times slower* than network S.

Initial stage f_0^* is defined as follows:

$(\forall i) f_0^*(x_i^*) = f_0(x_i)$; $\forall ij (\forall y \in Y_{ji}) f_0^*(y) = f_0(x_j)$; $(\forall z \in Z) f_0^*(z) = 1$.

Construction of network S^* guarantee that it simulates original network S.

Corollary 2.

Any network S is simulated by network S^* with all connection weights equal 0 or 1 and all node thresholds positive and identical.

Corollary 2 follows directly from Corollary 1 and Theorem 2.

5. Conclusion

First theorem has theoretical and practical meanings. It shows, that the network with inhibitory (negative) connections can always be replaced by network with only excitatory (positive) connections. This theorem also gives method of construction of simulating network, which has only twice more nodes and non-zero weighted connections than original net. Networks with only one kind of connections are often easier for hardware implementation, even if they have some more nodes. For instance they are simpler to optical implementation. Second theorem has rather theoretical meanings. It shows, that only one value of threshold and only one value of non-zero connection weight can be used to simulate any neural network. Practical application of this theorem may by in general difficult because transformed networks can have much more nodes than the original.

References

[1] Lippmann R P 1987 *IEEE ASSP Magazine* **3** 4 -22.
[2] McCulloch W S and Pitts W 1943 *Bull. of Math. Biophysics* **5** 115-133.
[3] Dickey F M and Delaurentis J M 1993 *Opt. Commun.* **101** 303-305.
[4] Bródka J S "The Equivalence of Neural Networks", (in press).
[5] Hopfield J J 1982 *Proc. Nat. Acad. of Sci. USA* **79** 2554 -2558.

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Experimental investigation of the performance of the two-layer neural network based on an optical vector-matrix multiplier

N N Evtihiev, R S Starikov, B N Onyky, V V Perepelitsa, I B Scherbakov

Dept. 37, Moscow State Engineering Physics Institute (Technical University),
Kashirskoe shosse 31, Moscow, 11509, Russia

Abstract. The paper presents the obtained results of learning of a two-layer neural network (TLNN), the results of tolerance to noise at the weight matrices and the results of hardware implementation. The imperfections of optics satisfy the requirements of the TLNN model.

1. Introduction

Different types of neural networks (NN) are under investigation now. Creating the working models of optical NNs is most useful and promising [1]. One of the most interesting is the TLNN, which can be easily made on the basis of a proposed optical matrix-vector multiplier (OVMM). Computing in TLNN is performed sequentially in OVMM, high performance speed is available due to the implementation of the high-speed high dynamic range multichannel multifrequency acoustooptic modulator (MAOM), laser array and CCD (charge-coupled device) array.

2. NN's mathematical models suitable for optoelectronic implementation

Due to the fact that the optoelectronic devices are the most effective for realization of the vector-matrix multiplication, it is very useful to put the main calculation load on them and to simplify the non-linear transformations as much as possible. That is why it is possible to assume that neural networks with the same transformations for all neurones (homogeneous NN) are the most suitable for the optoelectronic realizations.

From the viewpoint of hardware implementation the most promising architecture is the multilayer neural network. As it was proved [2] such networks may solve tasks of image recognition. The amount of interconnection in such networks is less than in completely-connected ones. This reduces the complexity of the hardware realization.

3. TLNN Hardware

Optical realization of the TLNN is an example of the effective application of MAOM. The main part of the calculation in this case is vector-matrix multiplication: A×B=C. The proposed OVMM architecture (fig. 1) allows one to perform this in one step.

The algorithm we propose is similar to DMAC (digital multiplication via analogue convolution) [3], but there is one significant difference concerned with the ability to use frequency coding in MAOM. Due to this fact OVMM allows one to multiply (convolve) 2D and 3D data flows and obtain the whole vector-matrix multiplication in one step with speeds about the frame rate of MAOM. In this case the vector As a 2D array is input into the laser array (LA). The elements b_{ij}^{s} of the 3D array (matrix B) are input into the MAOM: i is the channel number, j is the pulse position in the channel aperture at the exposure moment, s is the acoustic wave frequency number (f_s). When the MAOM aperture is filled the LA executes eight (according to the number of bits in each vector element) light pulses (expositions). Radiation from each laser uniformly illuminates the corresponding "window" of every channel. The optical system transfers the light diffracted in the MAOM onto a set of eight linear CCD arrays. Each linear CCD contains eight (according to the number of frequencies) points. The CCD operates in the shift-and-add mode. The shift velocity corresponds to the bit loading speed of the lasers. The charge is accumulated at a point of a linear CCD while the first exposition is shifted to the next point of this CCD for a second exposition. Each linear CCD represents a component of the output vector C in a mixed representation. The supporting electronic control system (CS) is responsible for converting the presentation of this data from mixed to hex format, the thresholding, the synchronization, etc.

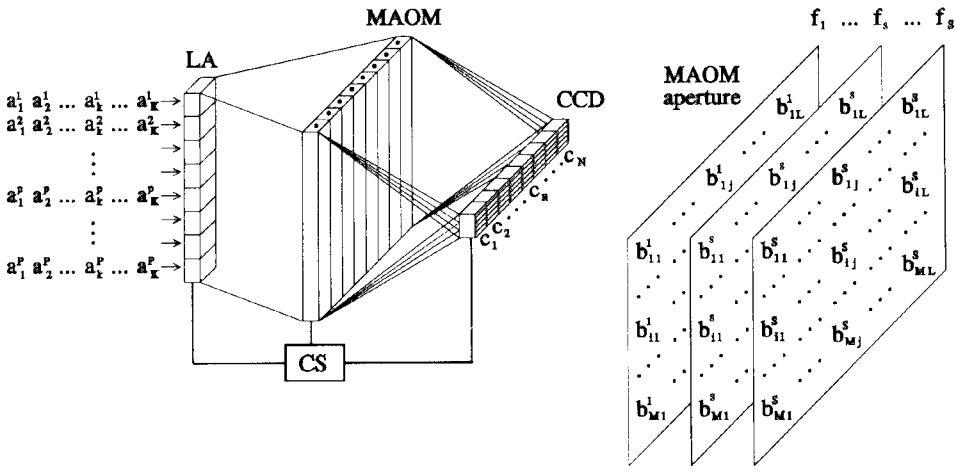


Fig. 1. Vector-matrix multiplier.

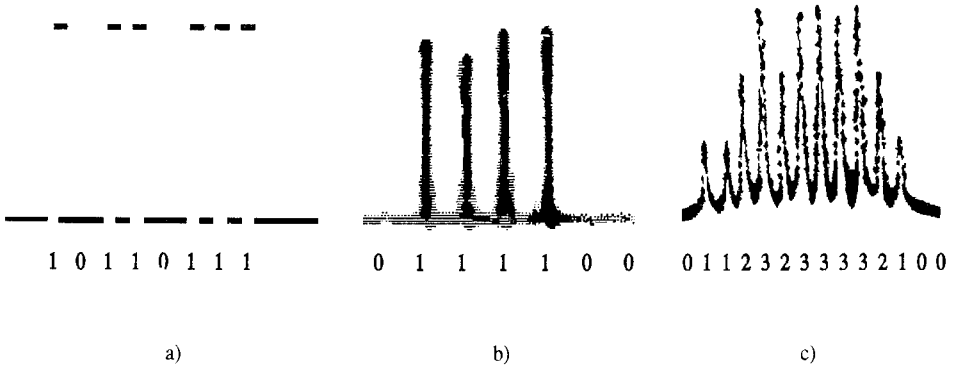


Fig. 2. Example of DMAC multiplication with implementation of multiplexing: a) lazer pulses, b) multiplexing signal in acoustooptic cell, c) CCD output signal in a mixed representation.

To check the possibility of performing DMAC in the described way, a breadboard model performing multiplication of 8 digit numbers was assembled. The rate of bit-by-bit entering was limited by the CCD array transfer frequency. The typical resulting signal of DMAC-multiplication with implementation of multiplexing in the MAOM being registered directly at the CCD output is presented in Fig. 2. Taking into account the utmost parameters of TeO_2 for MAOM as follows: 100 "windows", 8 frequencies, the period of laser pulse repetition is about 10 ns and the CCD array transfer frequency is about 120 MHz.

4. TLNN Noise tolerance

The TLNN simulation software program was designed. The NN was trained with gradient and stochastic algorithms.

In the hardware implementation of TLNN each parameter could differ from the calculated one. This could happen because the technology of optoelectronic NN fabrication does not allow the exact representation of the calculated meanings of parameters.

The aim of statistical experiments on noise was the definition of tolerance to deviation of meanings of the parameters from the calculated values. The criterion was the ability to recognize all the learned images.

As a result of modelling the men value of misrecognition and the mean data of unrecognized images are defined. The results of the experiments are the following. The TLNN retains the ability of correct recognition in two cases:

(i) The noise ratio (NR) of synaptic weights is less than 15%. The NR of all other parameters is less than 5%.

(ii) The NR of all the parameters is less than 10%.

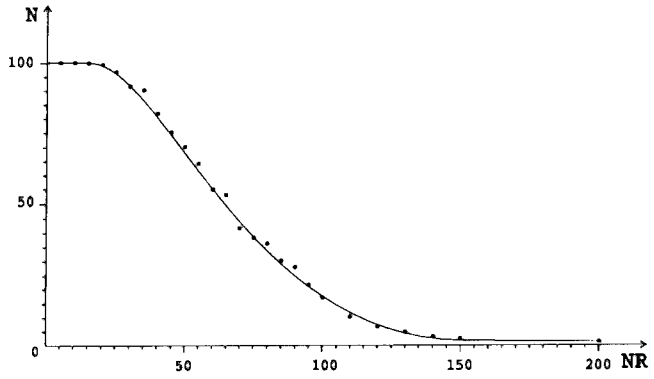


Fig. 3. Noise toleration of the TLNN.

So, it is proven that the technology of optoelectronic NN fabrication which allows to present all calculated parameters in hardware with accuracy better than 10% could be implemented for the hardware realization of TLNN.

Figure 3 presents the plot of noise tolerance of all the parameters of TLNN. It is important to note the linear decreasing of the number of correctly recognised images (N) against the NR of TLNN's parameters. While NR=60% the TLNN retains the ability of correct recognition of 50% of the images. For computer systems of previous generations another type of plot is inherent, namely avalanche dependence: 100% recognition at weak noise and 0% recognition beyond some noise level.

5. Conclusion

The described TLNN model is fast, versatile, stable and easy to implement. The computer experiment shows that such TLNN has significant noise resistance. These facts make the TLNN attractive for hardware realization.

References

[1] Barnard E and Casasent D 1988 *Proc. SPIE* **963** 537-45

[2] Muller B and Reinhardt J 1990 *Neural Networks* (Springer-Verlag)

[3] Beaudet P R et al 1986 *Appl. Opt.* **25** 18 3097

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Features of Application of Holographic Memories in PC Networks and Optical Neural Computers

V.N. Shahgedanov

Moscow Scientific & Research Institute of Instrument Engineering, Russia

Abstract. The characteristics of experimental prototypes of disc holographic memories of 1.5 Gb capacity and with 16 Mbs^{-1} data reading rate are given. The features of interfacing these devices with personal computers are discussed, as well as the question of the application of the data recorded in a holographic disc as a command matrix for a neural computer.

Optical memory systems have considerable advantages over the traditional memory systems, these include higher density of recording, higher rate of data reading and transmission, and storage of the information for long periods of time. Optical disc systems recording information in a form of holograms prove extremely promising for storage of the large volumes of information (exceeding 10 Gb). Usage of 1-D holograms is one way of implementing a multichannel recording device (Mikaelyan, 1981). Being a compromise between the 2-D holographic systems and optical bit systems, 1-D systems retain the advantages of both and fit well into the architecture of a computer.

Development of an element base for a 1-D memory system is less complicated in comparison with the element base of a 2-D system. At present, multichannel devices based on the 1-D holograms are being developed to be used as high speed data recorders which are employed in various systems of data processing and storage and as an external storage in computer networks. We suggest that a holographic disc may be used as a command matrix for neural computers. Characteristics of a memory system prototype based on 1-D holographic recording are presented in Table 1. For successful implementation of the holographic memory systems in PC networks and in the optical neural computers, high speed processing is required. This is why the major feature of the devices is implementation of specially designed lines of photodiodes based on the metal-resistor-semiconductor (MRS) structure. Due to a high conductivity of the resistive layer, the MRS-structure works under a constant voltage. Homogeneity of the process in the MRS-structure is achieved through the emerging negative feedback on the border of a semiconductor-resistor layer. Characteristics of the photodiodes are presented in Table 2.

Table 1. Characteristics of Memory System Prototypes

Technical Characteristics	HMS	Universal HMS	Optical Memory System APX-400
Disc capacity (one side), GB	1.5	0.4	0.4
Data transmission rate, $Mbsec^{-1}$	2.0	16.0	0.625
Seeking & reading time for data volume of 1Mb, sec	0.75	0.31	4.0
Probability of an error	10^{-12}	10^{-12}	10^{-12}

HMS - holographic memory system

Table 2. Photodiodes Characteristics

Wavelength, μm	0.63
Number of photosensitive elements	36
Noise equivalent power, nW	< 5
Speed, MHz	> 20
Multiplying factor	> 1000
Factor of connection between the elements	< 5%
Step between elements, μm	250
Size of the elements, μm	100 × 100

The analysis shows that the use of a direct current amplifier for a photodiode device at low levels of the detected optical signal; (which is typical of a holographic memory system) demands complicated methods of reconstruction of a constant component and control of comparator threshold. A simpler solution involves the installation of a low frequency filtering capacitor, i.e. the use of an alternating current amplifier. This, in turn, suggests a pre-set preparation of the recorded signal which is achieved through the use of positional codes which do not contain a constant component. Along with the certain advantages (such as, for example, simpler arithmetic), the positional codes possess a number of drawbacks. A disadvantage of a technical form is that a wide transmission band is required when transmitting coded signals. Indeed, long sequences of "0" and "1" codes are transmitted by a zero frequency signal, while alteration of "0" and "1" codes (width T_0) demands the upper transmission limit to be not less than $1.5\ T_0^{-1}$Hz. This is a strict requirement which results in various technical problems. A desire to avoid zero frequency has led to a number of codes where the lower frequency is shifted in time (Kovalenkov, 1989). These codes are either partially or completely compensated, i.e. the constant component is independent of codes. A particular example of a completely compensated code is the modified frequency modulated (MFM) Miller's code. This code is distinguished by a relatively simple technical realisation. As it consists of the three carrier frequencies, it does not require expansion of the band towards the high frequencies. Due to its qualities, Miller's code was widely accepted originally for magnetic recording devices, later in digital sound recorders, and at present in the optical recording discs and in the holographic recording discs in particular.

An algorithm of transformation of binary code into the MFM code is presented in Fig. 1a. Fig. 1b presents a scheme of a coder transforming an input code D_i into the MFM code. The synchronising frequency of D_i is F_0; an accompanying input frequency is $2F_0$. The positively rectifying differentiator forms the trigger input signal for a flip-flop unit which then forms the MFM signal. The widths of rectangular pulses are T_0, $1.5T_0$ and $2T_0$ (Fig. 1a). The maximum frequency does not exceed $0.5T_0^{-1}$, and the minimum frequency is $0.25T_0^{-1}$, i.e. the lower limit of the transmitting band is $0.25T_0^{-1}$Hz, and the upper limit is $0.75T_0^{-1}$Hz. When such an MFM signal is decoded by a decoder shown in Fig. 1c, the resulting code is the initial D_i shifted by $0.5T_0$.

References

Kovalenkov, L L, 1989, Mashinostroenie, 264.

Mikaelyan, A L, 1981, Radiotechnika, 36, 6.

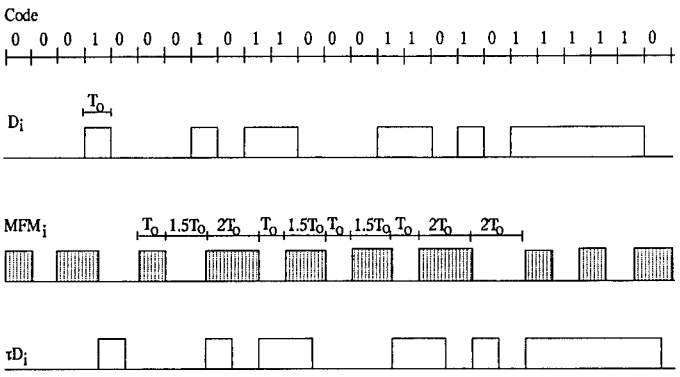


Fig. 1 a. Combined Time Diagram of "i" Channel of the MFM Coder and Decoder.

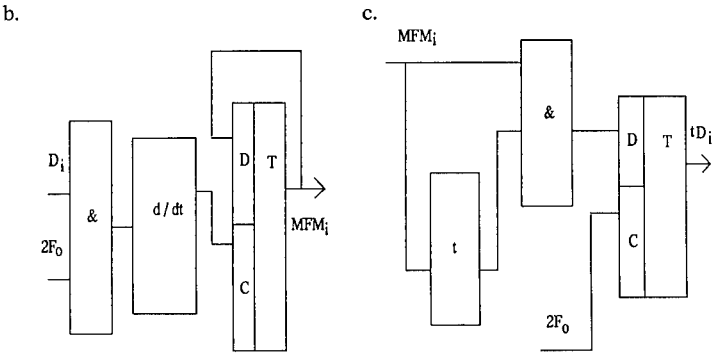


Fig. 1 b, c. Functional Diagram of "i" Channel of the MFM Coder (b) and MFM Decoder (c).

Inst. Phys. Conf. Ser. No 139: Part IV
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optoelectronic Fuzzy control of an inverted pendulum using light–emitting–diode arrays and position–sensing–devices

H Itoh, B Houssay and S Mukai

Optoelectronics Division, Electrotechnical Laboratory, Tsukuba, Ibaraki, 305 JAPAN

T Yamada and S Uekusa

School of Science and Technology, Meiji University, Higashi Mita, Kanagawa, 214 JAPAN

Abstract. For the first time real–time optoelectronic fuzzy control of an inverted pendulum is realized. The system uses arithmetic product–sum–gravity method. Antecedent and consequent operations using Gaussian–like membership functions and center–of–gravity operations are realized by combinations of LED array and a position–sensing–device.

1. Introduction

High–speed optoelectronic analog fuzzy inference technique has been proposed[1] using beam–scanning laser diodes (BSLDs)[2] and position–sensing–devices (PSDs). The inference process uses an arithmetic PRODUCT–SUM Gravity method with Gaussian membership functions and the controllability of the inference method is better than conventional MIN–MAX Gravity method with triangle membership functions[3]. The inference speed of the system will be more than several tens of MFLIPS (Mega Fuzzy Logical Inference Per Second) using high–speed operation of the beam scanning laser diodes. However, a real system control using the BSLDs has not been yet realized because of the present low resolution of their far–field pattern.

In this paper, configurations of optoelectronic antecedent and consequent fuzzy operation units using an light–emitting–diode (LED) array and a PSD are proposed. Furthermore, the first optoelectronic control system using the consequent unit is realized and its usefulness is demonstrated using an inverted pendulum.

2. Optoelectronic analog antecedent and consequent units

Figure 1 shows a schematic configuration of a fuzzy antecedent grade evaluation unit, which calculates the satisfaction of a fuzzy antecedent rule. In this figure, 1, 2 and 3 serially–wired LEDs are connected with the input in parallel. 1, 2 and 3LEDs start emission in turn with increasing the applied voltage. Because a PSD can detect a center–of–gravity of the radiation pattern, Gaussian–like membership function is realized by position and emitting characteristics of the LEDs. Figure 2 shows a schematic result of several membership functions. Peak

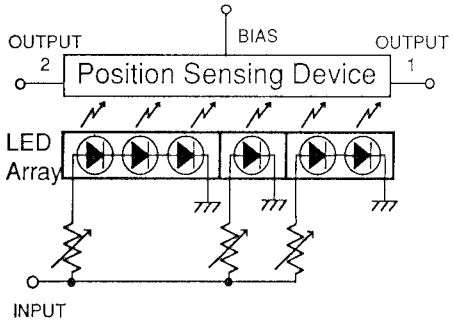


Fig.1 Optoelectronic antecedent grade evaluation unit using an LED array and a PSD

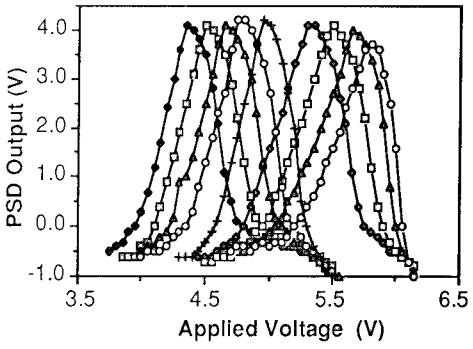


Fig.2 Examples of membership functions of an optoelectronic antecedent operation unit using an LED array and a PSD.

position of the function is controlled by value of the resistors and level-shift-diodes.

Figure 3 shows a schematic configuration of an optoelectronic fuzzy consequent operational unit. Unified output membership function by superposition of LED radiations and defuzzification of the radiation patterns are realized by an LED array and a position sensing device. Figure 4 shows an example of defuzzification of the unit. Unequal interval of each PSD output reflects the separation of each LED of the array.

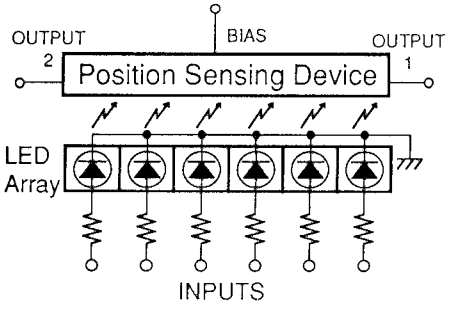


Fig.3 Optoelectronic consequent operation unit using an LED array and a PSD

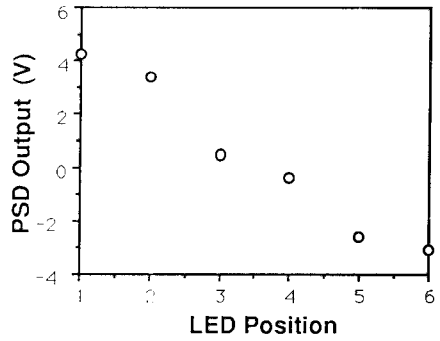


Fig.4 Output sample of a consequent operation unit using an LED array and a PSD.

3. Inverted pendulum control system

Figure 5 shows a schematic configuration of optoelectronic fuzzy inverted pendulum control system. A 50 cm-long free-rotating inverted pendulum is on a cart. The cart is on a 100 cm-long rail and is driven by a stepping motor. Input data for the fuzzy control are angle of the pendulum, which is measured by a rotary encoder, and angle velocity of the pendulum, which is calculated by the difference of measured angles. Consequent fuzzy inference processes are realized by the optoelectronic operation unit and antecedent fuzzy inference processes are simulated by a personal computer.

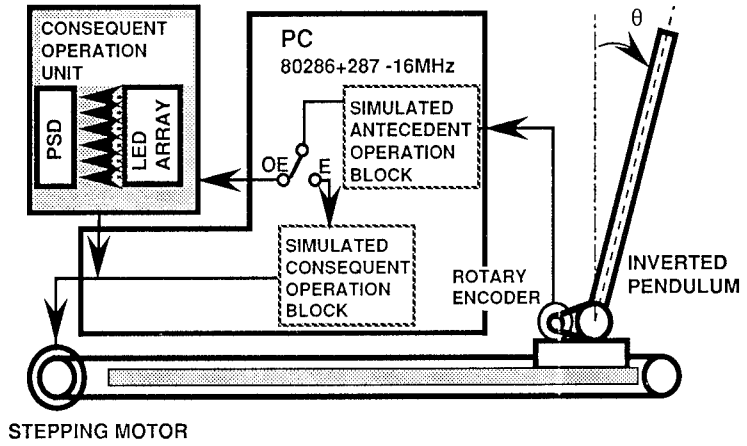


Fig.5 Inverted pendulum fuzzy control system with an optoelectronic/ electronic consequent operation unit

Figure 6 shows schematic experimental results of various experimental conditions. Angle of a controlled inverted pendulum after initial perturbation (40 cm/sec, 0.3 sec) is displayed as a function of time. Antecedent Fuzzy operations are numerically simulated by a Personal Computer. Cases of optoelectronic control "OE" and pure electronic fuzzy control "E" are shown in the figure. Control periods of "OE" and "E" cases are 0.015 sec/cycle and 0.08 sec/cycle, respectively. Difference of the cycle time are from introduction of optoelectronic processing. Control gain of "HIGH" in the figure 6 is 4.5 times than that of "LOW". The results show the superior controllability of the optoelectronic system.

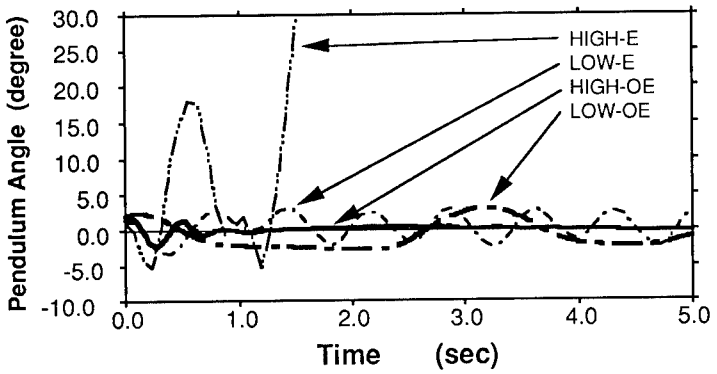


Fig.6 Traces of the angle of an inverted pendulum controlled by optoelectronic systems and electronic systems

4. Conclusion

Antecedent and consequent fuzzy operation units using a LED array and a position sensing devece were proposed and the characteristics were described. For the first time, real-time control of an inverted pendulum was realized based on an optoelectronic analog fuzzy inference system. Further better controllablity was demonstrated using an optoelectronic consequent Fuzzy inference control unit. This result comes from the operational speed of optoelectronic implementation is faster than that of numerical simulation.

The authors thank Yokoyama I for his technical support. Also the authors thank Mori M, Watanabe M, Hidaka T and Yajima H for fruitful discussions.

References

[1] Itoh H, Watanabe M, Mukai S and Yajima H 1993 *Proceedings of Topical Meeting on Optical Computing '93* (Palm Springs USA) OWD5 123–126

[2] Mukai S, Watanabe M, Itoh H, Yajima H, Hosoi Y and Uekusa S 1985 *Opt. and Quantum Electron.* **17** 431–434

[3] Itoh H, Yamada T, Houssay B, Mukai S and Uekusa S 1994 *Abstract of the Topical Meeting of the International Commission for Optics '94 –Frontiers in Information Optics–* (Kyoto Japan) 5–B2–19

V. ARRAY DEVICES

Sources, Smart-Pixels and Spatial Light Modulators

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Independently Addressable Vertical-Cavity Surface-Emitting Laser Diode Arrays

E. Zeeb, B. Möller, T. Hackbarth*, H. Leier*, K.J. Ebeling

University of Ulm, Department of Optoelectronics, D-89069 Ulm, Germany

*Daimler Benz Research Center, D-89013 Ulm, Germany

Abstract. The fabrication and testing of very uniform, planar proton-implanted independently addressable 10×10 vertical-cavity surface-emitting laser diode arrays is reported. Individual elements of 12 μm active diameter show threshold currents around 4 mA and emit up to about 300 μW single-mode. An electrical 3 dB modulation bandwidth of 4.6 GHz is determined. The potential total transmission rate of the array is several hundred Gbit/s.

1. Introduction

Independently addressable, two-dimensional laser diode arrays have many applications for example in optical scanners, display technology, optical interconnects, high capacity switching systems, or for the generation of high power coherent beams [1]. Using planar vertical-cavity surface-emitting lasers (VCSELs), such arrays with high packing densities are relatively easy to fabricate [2]. In addition VCSELs with small active diameters allow dynamic single-mode oscillation, very low threshold currents [4], and alignement tolerant and efficient coupling into single-mode fibers [3]. Here we report on 10×10 independently addressable vertical-cavity laser diode arrays with active diameters of 12 μm and an excellent homogeneity of the output characteristics over the whole array size. 3 dB modulation bandwidths of single-mode emitting array lasers are as high as 4.6 GHz at an optical output power of 340 μW.

2. Laser Structure and Device Fabrication

The laser structure under investigation was grown by molecular beam epitaxy. The one wavelength thick central region contains three 8 nm thick $In_{0.2}Ga_{0.8}As$ quantum wells which are embedded in GaAs barriers and $Al_{0.3}Ga_{0.7}As$ cladding layers for efficient carrier confinement. It is surrounded by two semiconductor Bragg reflectors. The upper mirror is p-doped and consists of 18 quarter-wavelength AlAs/GaAs pairs, the lower is n-doped and has 22.5 pairs. To reduce the resistivity of the reflectors we used two

intermediate $Al_{0.3}Ga_{0.7}As$ and $Al_{0.7}Ga_{0.3}As$ steps and enhanced doping concentration of $8 \cdot 10^{18}$ cm^{-3} near the heterointerfaces driven in backward direction. During growth of the upper AlGaAs cladding layer the rotation of the substrate was interrupted to produce a spatial thickness variation of the central lasing region. The thickness variation results in a spectral shift of the resonator mode across the wafer.

Lateral current confinement and insulation of the individual lasers is achieved by proton implantation with an energy of 300 keV and a dose of $5 \cdot 10^{14}$ cm^{-2}. The reflectivity of the upper Bragg reflector is improved with an additional Au layer on top. TiAu rings form the ohmic p-contacts. The 100 elements are connected to 25 bonding pads at either side of the 5×5 mm array chip by TiAu feeding lines which are deposited on an insulating SiO layer. A schematic of the vertical structure and a top view of a full chip is depicted in Fig. 1. Individual lasers have active diameters of 12 μm and the pitch size of the array is 250 μm.

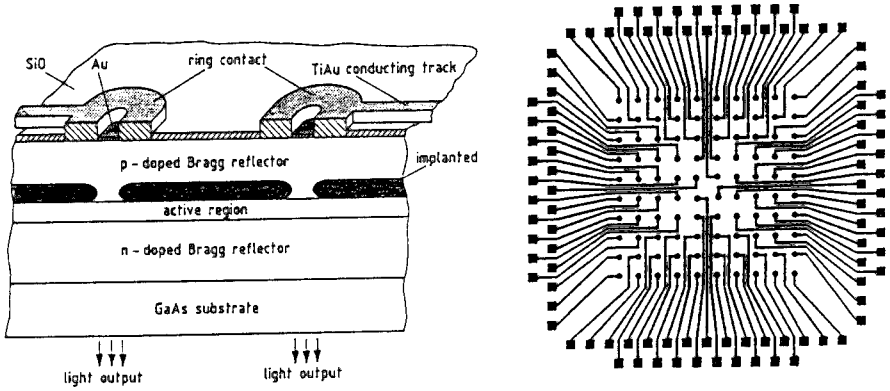


Fig. 1: Vertical structure and top view of an VCSEL array fabricated.

3. Output Characteristics

All processed arrays show a very good homogeneity of the output characteristics over the whole array size. Fig. 2 depicts a typical threshold current histogram of an array. Almost all lasers exhibit threshold currents between 4.1 mA and 4.7 mA with an average value of 4.4 mA. Each laser has a maximum single-mode output power in the order of 300 μW. Above this level higher order transverse modes appear in the emission. The maximum output power of about 500 μW is restricted by thermal roll-over due to still high threshold voltages.

Due to the wedge shaped central region the emission wavelength of individual lasers within the array slightly changes with position. Fig. 3 shows emission wavelengths at driving currents of 5 mA. Emission between 994.3 nm and 998.3 nm makes the array well suited for wavelength division multiplexing systems. Using different growth conditions the shift of the emission wavelength can be adjusted to vary up to 30 nm across the array [5, 6].

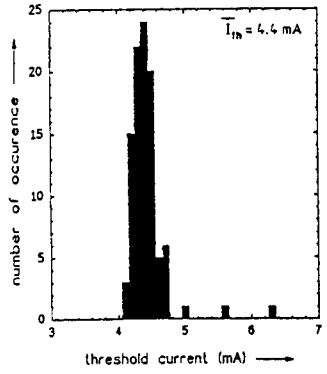


Fig. 2: Histogram of threshold currents.

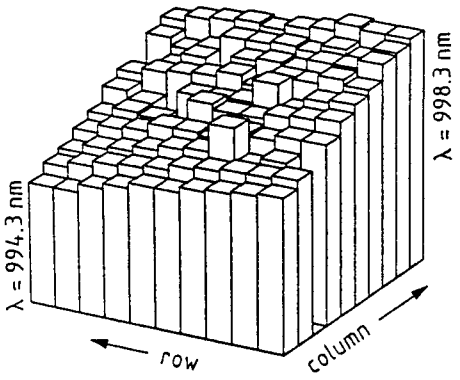


Fig. 3: Emission wavelength distribution.

Fig. 4 illustrates small-signal modulation responses of an individual lasing element for different driving currents above threshold together with corresponding light output characteristics. With increasing driving current the modulation bandwidth increases up to 6.2 mA, where the next order transverse mode starts to oscillate. A maximum electrical 3 dB modulation bandwidth of 4.6 GHz and a 5 GHz optical bandwidth are obtained for single-mode emission. Crosstalk to adjacent lasers both biased above threshold was measured to increase with increasing frequency and to remain below −15 dB at 4 GHz. Crosstalk is mainly caused by non-optimized microwave packaging. Operating all 100 elements in parallel bit rates of several hundred Gbit/s can be obtained.

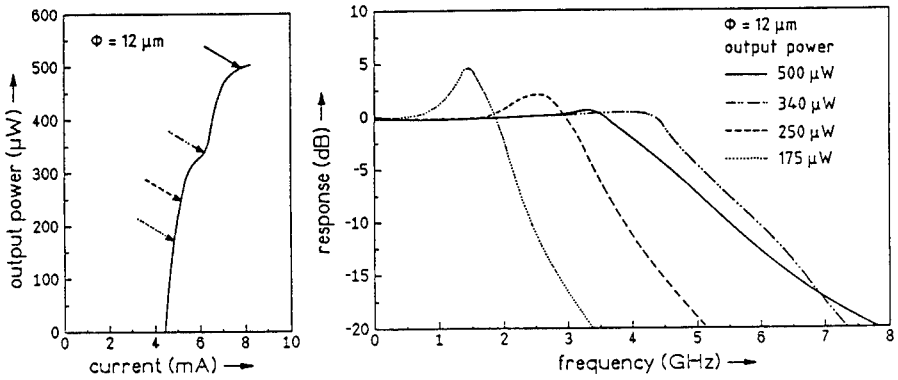


Fig. 4: Light output characteristics and electrical small-signal modulation response.

The output characteristics of the arrays can further be improved if laser structures with lower series resistances are used and when the arrays are solder bonded with a flip-chip bonding technique on insulating heat sinks. Tests with individual lasers already indicated that maximum output powers and modulation bandwidths can be enhanced if proper heat sinking is applied [7].

4. Conclusion

We have fabricated 10×10 independently addressable planar proton implanted laser diode arrays with 12 μm active diameters. The output characteristics of the lasers within the array are very uniform. The elements show an average threshold current of 4.4 mA, maximum output powers of about 500 μW and optical 3 dB modulation bandwidths of 5 GHz for single-mode emission. These results suggest that VCSEL arrays are ideally suited for data transmission of several 100 Gbit/s in wavelength division multiplexing systems or high capacity multi-channel optical interconnects.

5. Ackowledgments

The authors thank Dr. J. Wenger for help with the lithography masks and G. Strohm for bonding the devices. This work was supported by the German Ministry of Research and Technology (BMFT) and the Deutsche Forschungsgemeinschaft (DFG).

References

[1] Iga K 1990 *Proc. 16th European Conference on Optical Communication, Amsterdam, The Netherlands*, pp 895-931

[2] Vakhshoori D, Wynn J D, Zydzik G J and Leibenguth R E 1993 *Appl. Phys. Lett.* **62** pp 1718-1720

[3] Wipiejewski T, Panzlaff K, Zeeb E, Weigl B and Ebeling K J 1993 *Proc. 19th European Conference on Optical Communication, Montreux, Switzerland*, pp 333-336

[4] Hansing C C, Deng H, Huffaker D L, Deppe D G, Streetman B G and Sarathy J 1994 *IEEE Photon. Technol. Lett.* **6** pp 320-322

[5] Zeeb E, Möller B, Reiner G, Ries M, Hackbarth T and Ebeling K J 1994 *Proc. IEEE Semiconductor Laser Conference, Hawaii, USA*, W1.2

[6] Zeeb E, Reiner G, Ries M, Hackbarth T and Ebeling K J 1994 *Proc. 24rd European Solid State Device Research Conference, Edinburgh, Great Britain*

[7] Zeeb E, Hackbarth T and Ebeling K J 1993 *Proc. 23rd European Solid State Device Research Conference, Grenoble, France*, pp 873-875

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994

Nonuniformity Tolerance and Its Price for Different Modes of Operation of FET-SEED Smart Pixel Arrays

L. M. F. Chirovsky[a], A. L. Lentine[b], T. K. Woodward[c], G. Livescu[a], and G. D. Boyd[c]

AT&T Bell Laboratories, [a]Murray Hill, NJ 07974, [b]Naperville, IL 60566, [c]Holmdel, NJ 07733

Abstract. We quantify the nonuniformity tolerance of FET-SEED circuit arrays and show its dependence on optical output contrast, switching speed, available laser power, and the mode of operation. This analysis, utilizing recent experimental data, then shows that the preferred mode of operation may be either with differential optical signals or with single-ended optical signals, depending on the circumstances.

Although a particular kind of FET-SEED Smart Pixel circuit may have acceptable performance for a given application, one cannot assume immediately that all the Smart Pixels in a large array can exhibit the same performance in a system. Neither the circuits nor their concomitant optical beams can be identical even if designed to be so. There will be variations or nonuniformities due to many different causes. Thus the circuit and system designs should minimize the effects of possible variations, and failing that, be tolerant to the lingering effects of nonuniformities, so that all pixels operate properly under one global set of conditions. In recent years, as sources of nonuniformity were identified, iterative redesign has significantly reduced the size of variations which occur in FET-SEED circuits and the associated system hardware [1,2]. The use of "dynamic latching" greatly dissipates the effects of many of the nonuniformities not eliminated [3,4]. The remaining variations can be dealt with as if due solely to differences in the amount of optical power in the beams incident on the modulators in the FET-SEED circuits, hereafter called the "read beams". The principal issue is to guarantee that the two possible digital signals (logical "1" and "0") are correctly recognized as such, as data is optically transmitted from pixel to pixel, despite the fact that the read beam powers, P_r, are not all equal. Recall that data is optically transmitted by a FET-SEED circuit when a read beam impinges on an output modulator, that is either in a high reflectivity state, R_{HI}, for one logic state, or a low reflectivity state, R_{LO}, for the other logic state. The output signals are then either $R_{HI}P_r$ or $R_{LO}P_r$. Their ratio, $K_O \equiv R_{HI}/R_{LO}$, is often called the modulator output contrast.

A mode of FET-SEED circuit operation, where optical data encoding is done with transmitters which use only one modulator is called single-ended. If R_{LO} is not zero, and K_O is finite (as is usually the case for SEED modulators) then one can intuitively see that for single-ended operation, a nonuniformity problem can arise when $P_{ri} \neq P_{rj}$, where i,j denote the ith, jth pixel in an array. For instance, when $P_{ri} >> P_{rj}$, then either $R_{LO}P_{ri}$ becomes too large for one logic state or $R_{HI}P_{rj}$ becomes too small for the other logic state, as one tries to globally adjust all the P_r powers together. To date, this problem has usually been dealt with in systems using FET-SEED circuit arrays by operating in a mode where the data encoding is differential. Each transmitter has two modulators, configured to settle into complementary states, and receives two read beams, P_{ri1} and P_{ri2}. The two digital states are thus represented by complementary signal pairs ($R_{HI}P_{ri1}$, $R_{LO}P_{ri2}$) and ($R_{LO}P_{ri1}$, $R_{HI}P_{ri2}$). The optical receivers, already having dynamic latching for its benefits, are then designed to respond to the difference in power in a signal pair, so that the two logic states are registered by ($R_{HI}P_{ri1} - R_{LO}P_{ri2}$) or ($R_{LO}P_{ri1} - R_{HI}P_{ri2}$). If there is local uniformity (where $P_{ri1} \approx P_{ri2} \approx P_{ri}$, $P_{rj1} \approx P_{rj2} \approx P_{rj}$) and only global non-uniformity ($P_{ri} \neq P_{rj}$), then in differential operation the digital signals are $P_{ri}(\Delta R)$ or $P_{ri}(-\Delta R)$; $P_{rj}(\Delta R)$ or $P_{rj}(-\Delta R)$; etc., where $\Delta R = R_{HI} - R_{LO}$. Then, once all P_r are made large enough, all digital signals will be correctly recognized even if $P_{ri} >> P_{rj}$ or vice-versa. Thus differential operation is generally considered much more nonuniformity tolerant than single-ended operation and essentially dependent on ΔR rather than K_O.

Some of the assumptions underlying the above assertion do not hold steadfastly under close scrutiny.

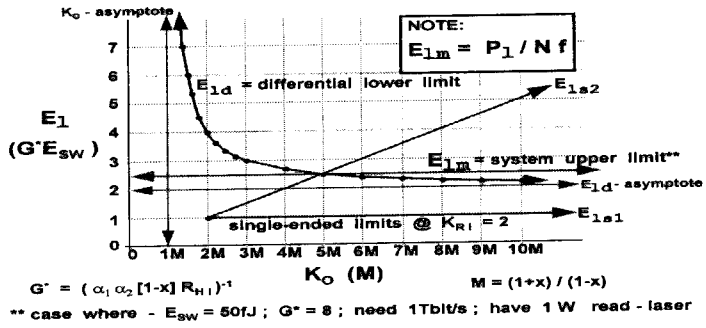


Figure 1. Plots of the limits on the launched laser energy, E_l, in units of ($G^* E_{sw}$) for differential (lower limit, E_{ld}) and for single-ended (lower limit, E_{ls1}; upper limit, E_{ls2}) modes of operation of FET-SEED Smart Pixel Arrays, versus signal contrast, K_O, in units of M, the nonuniformity tolerance margin. Superimposed in the figure, is a plot of E_{lm}, the maximum allowed launched laser energy for the particular system application case given in the figure. The position of E_{lm} can change with every application. In the region below E_{lm} and above E_{ld} or E_{ls1}, variations in the absolute value of E_l are tolerated above and beyond the nonuniformity variations across an array characterized by M.

The two principal ones are that there is local uniformity ($P_{ri1} \approx P_{ri2}$, etc.) and that the P_r's can be made as large as necessary. Recent measurements of beam powers in a real system [2,5] showed that, effectively, the local and global variations, as defined above, are about the same. Saturation and heating effects in the modulators as well as limited laser source power, for read beams, conspire to set an upper limit on the P_r's. The consequences of these facts can be quantified with the following analysis.

To deliver E_{sw} through a transmitter to a FET-SEED receiver which has a switching energy, E_{sw}, a source laser must launch an energy, E_l. These two energies have the relationship given by the equation:

$$E_l = GE_{sw}, \tag{1}$$

where $(G)^{-1}$ is the efficiency with which launched energy is converted to switching energy. $(G)^{-1}$ is reducible to the "Figure of Merit" (FM) discussed in [6]. For differential operation G_d is given by:

$$G_d = 2P_{ln}[\alpha_1\alpha_2(R_{HI}P_{li1} - R_{LO}P_{li2})]^{-1}, \tag{2}$$

where P_{ln} is the average laser launched power per beam, α_1 is the attenuation factor, due to system optics, between the laser and a transmitter's modulator, and α_2 is the attenuation factor, due to system optics, between transmitters and receivers. (Note that by these definitions $P_{ri} = \alpha_1 P_{li}$, etc. and $P_{ln} = P_l / N$, where P_l is the average laser power and N is the number of pixels). The required laser launched energy for differential operation, E_{ld}, is then given by:

$$E_{ld} = G_d E_{sw}. \tag{3}$$

Now let the beam powers vary such that over N pixels,

$$(1-x)P_{ln} \le P_{lia} \le (1+x)P_{ln}, \tag{4}$$

where x is a fraction ≤ 1, i = (1, 2, ... N) and a = 1,2.

Also define the nonuniformity tolerance margin, M, by

$$M = (1+x)/(1-x). \tag{5}$$

(Note that M is the ratio of the largest over the smallest beam power.) Then, for the worst case, which governs the array performance,

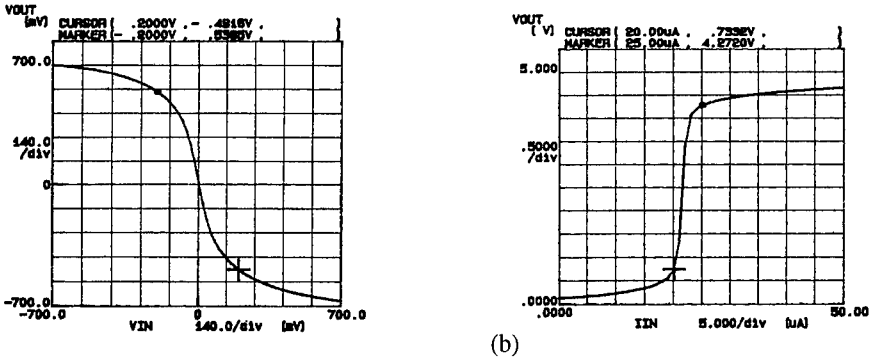
$$E_{ld} \ge 2G^*(1 - M/K_O)^{-1}E_{sw}, \tag{6}$$

where,

$$G^* \equiv [\alpha_1\alpha_2(1-x)R_{HI}]^{-1}. \tag{7}$$

Thus even for differential operation, the contrast ratio, K_O, is important, as was pointed out before in [7]. Error-free performance is not possible unless $K_O > M$, and then E_{ld} remains large until $K_O >> M$ (or at least $K_O \ge 3M$). In a system application, there will be an upper available limit on P_{ln} and a lowest acceptable limit on f, the data rate per pixel, setting an upper limit on the amount of laser energy launched per pixel, E_{lm}, given by:

$$E_{lm} = P_l/(Nf). \tag{8}$$



(a) (b)

Figure 2. a) This output voltage vs. input voltage plot for a MESFET inverter circuit, illustrates a sharp output voltage switching characteristic. The required input voltage swing contrast for proper digital output with this inverter is 1.8. b) This is a plot of the output voltage versus the input current for an experimental receiver circuit specifically designed for single-ended operation. The cursor and marker identify the threshold current, I_{th}, (essentially no switching for $I_{IN} < I_{th}$) and the switching current, I_{sw}, (essentially complete switching for $I_{IN} > I_{sw}$). Therefore for proper digital switching, here $K_{RI} = 1.25$.

Without a sufficiently large K_O, the allowable x and M become small and the nonuniformity tolerance of differential operation evaporates because one must have $E_{ld} \leq E_{lm}$. See Figure 1. A similar analysis shows that the methods of single-ended operation employed in many optical logic devices, using bias beams or reference beams, fare even worse. Is there an alternative?

The answer is possibly yes, in the form of single-ended operation with specially designed optoelectronic receiver circuits. These receivers must exhibit three important characteristics: 1) a sharp output switching threshold: 2) a quiescent state, for when no signal is applied, at a well defined distance from the threshold; and 3) a mechanism for returning the receiver to its quiescent state in the absence of light, which is based on a feedback principle or operates asynchronously with the signal.

Figure 2a depicts the input-output voltage characteristics of an electronic inverter illustrating at least one way of obtaining a sharp output switching threshold. If, for this inverter, one arbitrarily chooses V_{IN} at +0.7V as a starting point, then any logical "0" inputs can swing V_{IN} up to +0.2V (at the cursor) and produce essentially no output swing. Logical "1" inputs need to swing V_{IN} only beyond –0.2V (at the marker) to produce essentially complete output swing. Making the assumption that the input voltage swing, in a receiver, can be made proportional to optical input energy, one can then define the concept of required input contrast ratio K_{RI}, the minimum necessary input swing for switching divided by the maximum allowed input swing for no switching. The inverter of Fig. 2a has a K_{RI} of 1.8. Figure 2b shows the input-output response of the electronic portion of a receiver circuit designed specifically for single-ended operation. The K_{RI} here, defined as input switching current divided by the input threshold current, turns out to be 1.25. Figure 3 shows the results of an experiment where a FET-SEED receiver-transmitter, designed for differential operation, was actually operated by a novel single-ended mode of operation called Asynchronous Reset On Every Bit Input Contentionless Switching (AROEBICS), described briefly in the figure caption and in [8]. For the receiver in Fig. 3 and for complete single-ended receivers in general, there is a minimum necessary switching energy, E_{sw}, for bit 1 inputs and a maximum allowed energy, E_{th}, for bit 0 inputs (which produces essentially no switching).

Such receivers then have a required input contrast ratio, K_{RI}, obviously ≥ 1, such that

$$K_{RI} = E_{sw}/E_{th}. \quad (9)$$

Single-ended modes of operation have two constraints on E_l, namely, E_{ls1} and E_{ls2}, such that:

$$E_{ls1} \geq G_{s1}E_{sw} \quad \text{and} \quad E_{ls2} \leq G_{s2}E_{th}. \quad (10)$$

For the worst cases, which govern array performance,

$$G_{s1} = G^* \quad \text{and} \quad G_{s2} = (K_O/M)G^*. \quad (11)$$

Error free performance is not possible unless, (from eqs. 9, 10, 11)

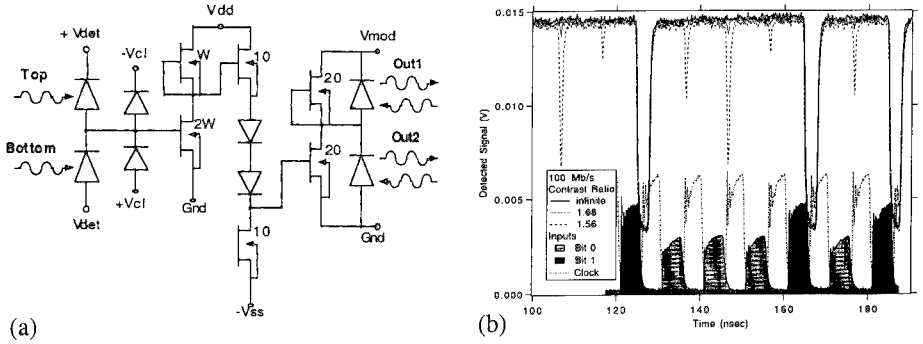


Figure 3. a) This is a schematic diagram of FET-SEED optical signal receiver-transmitter, designed for differential operation, but which can also be operated by the single-ended AROEBICS procedure. The input voltage value of the quiescent state is determined by the voltage $+V_{cl}$ on a forward-biased clamping diode. b) These are the results of an AROEBICS experiment performed with the circuit at a 100 Mbit/s data rate. The output is the reflection of a CW beam on the modulator "Out 2" and has a $K_O \approx 4$. "Return to Zero " (RZ) digital signal pulses arrive on the "Top" detector diode. Interleaved in time between the signal pulses, regularly clocked optical RZ pulses arrive on the bottom diode, to reset the receivers to the quiescent state, before each signal pulse. The K_{RI} is shown to be 1.68, so the allowed M can be as large as 2.38.

$$E_{ls2} \geq E_{ls1} \quad \text{i.e.} \quad K_O \geq MK_{RI}, \tag{12}$$

and $E_{ls1} \leq E_{lm}$. Plots of E_{ls1} and E_{ls2} are shown in Figure 1 for comparison with E_{ld}. Note that E_{ld} exceeds E_{ls1} by a factor of $2/(1\text{-}M/K_O)$. For receivers with $K_{RI} = 2$ and with a not very large K_O of 4, a very respectable $M = 2$ is allowed and $E_{ld} = 4E_{ls1}$. So, such a single-ended operation could be four times faster than the differential mode or alternatively could be done with four times less laser power. From a different perspective, one can also deduce that when a system application demands a high data rate with limited laser power, *the single-ended mode of operation can have greater nonuniformity tolerance than the differential mode of operation.* On the other hand, with a very relaxed data rate requirement, the differential mode will work with a lower K_O, by a factor of K_{RI}, than the single-ended mode would.

In summary, a simple method of analysis, which quantifies nonuniformity tolerance in FET-SEED circuit arrays by relating it to easily measurable parameters, leads to an understanding of how signal contrast, switching speed, available laser power, mode of operation and nonuniformity tolerance impact each other so that reasoned design choices can be made when such arrays are incorporated into systems.

[1] A. L. Lentine, R. A. Novotny, T. J. Cloonan, L. M. F. Chirovsky, L. A. D'Asaro, G. Livescu, S. Hui, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, K. G. Glogovsky, and T. K. Woodward, "4×4 arrays of FET-SEED embedded control 2×1 optoelectronic switching nodes with electrical fan-out" IEEE Photonics Tech. Lett., V.6, No. 9, (1994).

[2] F. B. McCormick "Design issues for free-space photonic switching demonstrations", paper WA3, OC'94.

[3] L. M. F. Chirovsky, "Dynamic Operation of Quantum Well SEED's", MRS Symp. Proc. V. 261, pp. 27-32 (1992).

[4] A. L. Lentine, L. M. F. Chirovsky, and T. K. Woodard, "Optical energy considerations for diode-clamped smart pixel optical receivers", IEEE JQE, V. 30, No. 5, pp. 1167-1171, (1994).

[5] F. A. P. Tooley and S. Wakelin, "Design of a Symmetric Self Electro-optic Effect Device Cellular-Logic Image Processor", Appl. Optics, V.32, No. 11, pp. 1850-1862 (1993).

[6] G. D. Boyd, L. M. F. Chirovsky, A. L. Lentine, and G. Livescu, "Wavelength optimization of quantum well modulators in smart pixels", to be published in Appl. Optics, (late 1994).

[7] A. L. Lentine and F. A. Tooley, "The relationships between speed and tolerances for self electro-optic effect devices", Appl. Optics, V.33, No. 8, pp. 1354-1367 (1994).

[8] L. M. F. Chirovsky, G. Livescu, G. D. Boyd, T. K. Woodward, R. A. Novotny, and A. L. Lentine, "Optoelectronic Circuits and Smart Pixels Which Use Asynhronous Reset on Every Bit for Input Contentionless Switching (AROEBICS)", Technical Digest, IEEE LEOS Topical Meeting on Smart Pixels, pp. 22-23 (1994).

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical, algorithmic and electronic considerations on the desirable "smartness" of optical processing pixels

M.P.Y. Desmulliez, B.S. Wherrett, J.F. Snowdon, J.A.B. Dines

Department of Physics, Heriot-Watt University, Riccarton EH14 4S, Scotland, U.K.

Abstract. The performance gains associated with optical computing schemes are analysed in terms of the complexity of individual processing nodes in order to optimise the efficiency of an optical processor. The sorting task is chosen as a practical example. Performance metrics are discussed for various component technologies and electronic node (smart-pixel) sizes.

1. Introduction

Advances in the integration of optoelectronic components with VLSI chips are now enabling arrays of electronic nodes to be interconnected optically through free space [1]. This hybrid technology offers performance gains in many computing tasks. Little has been done to quantify these gains or to compare the various options in specific component technologies or in electronic node size [2-4]. The "desirable smartness of a pixel" refers to the complexity of the individual nodes. Complex (very smart) pixels demand large chip area; hence only a few can be accommodated per chip and the 2-D parallelism offered by the optical interconnect may not be optimally exploited. Conversely, the optics may not be able to cope with vast numbers of simple pixels per chip, or the resulting high number of optically interconnected chips required to accomplish a given processing task may lead to inefficiencies.

We choose to consider the sorting task because, on algorithmic grounds, there is a reduction in computational time if implemented with space-variant non-local optical interconnects [5]. Sorting remains also one of the main tasks in computation [6]. Also, pixels of different complexities have already been designed and laid out to perform the sorting.

Three different classes of pixels have been studied. (i) The Symmetrical Self-Electrooptic-Effect Device (S-SEED) provides the least intelligent of the pixels in the sense that the logic functions of these p-i-n photodiodes are just NAND or NOR operations [7]. (ii) Any logic operation is a priori possible with the logic SEEDs (L-SEED) [8]. (iii) The last type of pixel integrates photoreceivers and modulators with VLSI compatible electronic circuitry. The integration is either monolithic (FET-SEED) [9] or performed by flip-chip bonding the optical transceiver chip onto a CMOS electronic chip [10]. The voltage gain of the first electronic stage provides both a substantial reduction in optical input switching energy and a thresholding unit. The pixel is designed to operate as a two-input, two-output exchange/bypass node [11]. Each of these technologies is considered to be operated at its optimum conditions.

In the optical domain, the limitations on the scaling possibilities originate from fundamental properties and technical constraints. For example, the aberration and diffraction effects of the imaging devices and the space-variant nature of the optical interconnections impose an upper limit on the space-bandwidth-product available and the field of view. Technical limitations encompass the micro-mechanical alignment and the stability of the optical system. The amount of source power per laser (taken to be 1 watt) restricts the power available per pixel, thereby limiting the switching rate of the nodes. In the electronic domain, the decrease of the yield with increasing chip size demands an area of no more than 1 cm^2, compatible with the optical field of view. In the thermal domain, it is assumed that 10 Wcm^{-2} of heat can be dissipated at each pixel array; this is achievable with conventional forced air cooling methods.

2. Effects of limited amount of laser source power, power dissipation and pixel complexity

The optimum smart pixel maximises the processing throughput rate while satisfying the constraints imposed by the limited heat removal capability and limited laser source power. The throughput rate of the sorter is defined as the number of 8-bit words that can be sorted per second and is expressed in millions of (byte) operations per second (Mops/s). On the sole assumption of limited laser power, the optimum throughput rate is a trade-off between the increase in number of processing elements (with increasing chip size) and the decrease in power available at each node (figure 1). The optimum throughput rate is reached for 36×36 S-SEED or L-SEED-based arrays. The dimensions of the (existing) devices are, respectively, $10\ \mu m \times 20\ \mu m$ and $55\ \mu m \times 55\ \mu m$. The optical windows are, respectively, $5\ \mu m \times 10\ \mu m$ and $5\ \mu m \times 5\ \mu m$ with a capacitance of $120aF\mu m^{-2}$. The contrast ratio of the output modulators is 4:1 if the device is operated at maximum absorption when unbiased and 6:1 otherwise. The optimum throughput rate of the 2×2 FET-SEED node of dimensions $560\ \mu m \times 280\ \mu m$ is 220 Mops/s and is obtained for an array size of 0.85 cm^2 whereas the optimum array size for 2×1 nodes is about 0.40 cm^2. The throughput rate for the 1 μm CMOS-SEED is limited by the maximum chip size and not by the power available. The CMOS pixel, of dimension $400\ \mu m \times 200\ \mu m$, is assumed to operate at a frequency of 100 MHz whereas the FET-SEED node has been successfully simulated at 310 MHz. At higher frequencies, more power is needed to convert the optical signals into electronic voltage swings. This explains the decrease of the throughput rate at large chip sizes for the FET-SEED-based array.

The power dissipated by the S-SEEDs and L-SEEDs originates from the photocurrent generated by the optical beams. The power-switching-time-product of these devices [7] suggests that high frequencies can be achieved if high optical power is available. The saturation of the absorption peak [12] at high power levels however limits the maximum frequency of operation if cascaded arrays of such devices are utilised. This is shown in figure 2 for the case of two S-SEED arrays. For this figure and the following ones, the power dissipated (left vertical axis) and the required laser source power (right vertical axis) per pixel are displayed as functions of the operating frequency of the devices. On each of the axes is also indicated by a dash the maximum powers allowed per device for different array sizes. The saturation irradiance of the SEED is assumed to be 1 $kWcm^{-2}$. A maximum frequency of 70 MHz is then achieved for 500 μW of laser power at each node, which corresponds to 350 μW of power dissipated per device. The limited laser source power is the fundamental

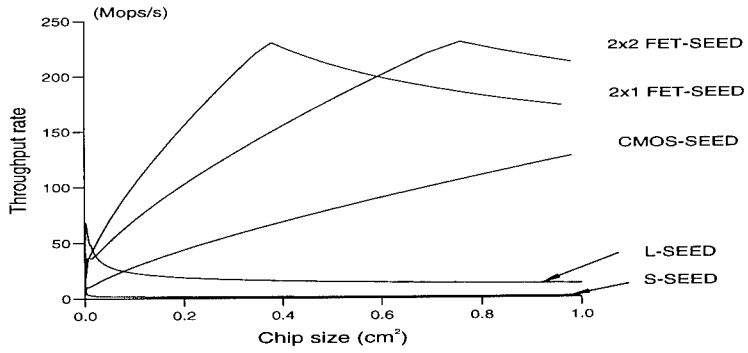


Figure 1

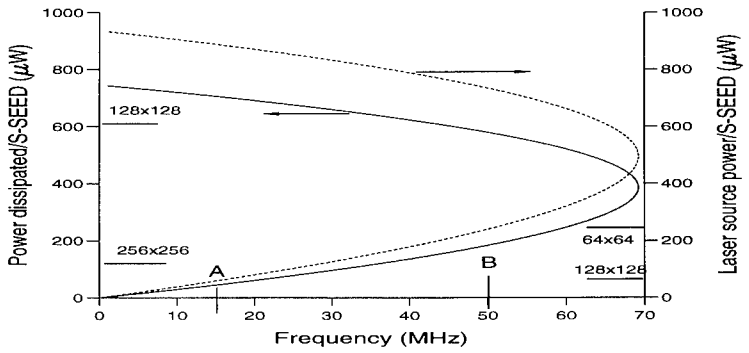


Figure 2

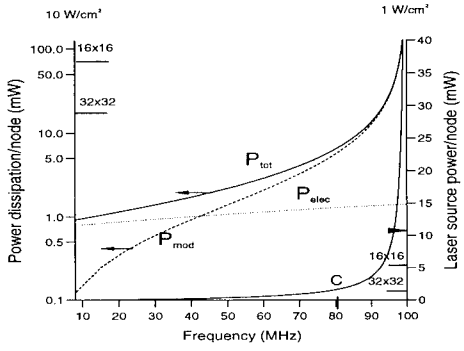


Figure 3

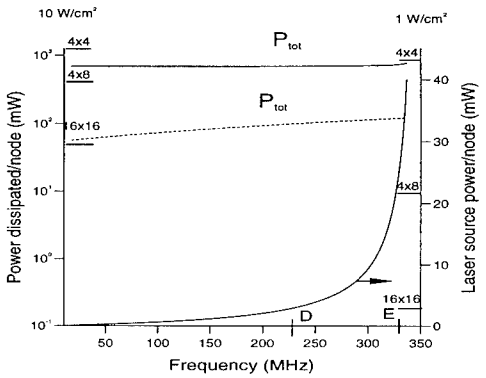


Figure 4

limit for the S-SEED; 11 Mops/s is achievable for a 128 × 128 array (A) whereas 12 Mops/s (B) is possible for a 64 × 64 array.

The total power dissipated for the CMOS and FET-SEED (P_{tot} in figures 3 and 4) comes from the heat generated by the output SEED modulators (P_{mod} in figure 3) and by the electronic circuitry (P_{elec} in figure 3). The total input capacitances of each pixel are taken to be 60fF. The frequency depends on the energy used to convert the optical signals into voltage swings. Pulse mode of operation of the lasers is hence preferred for high frequency in order to provide the minimum input energy in a short timescale [13]. For a 32 × 32 array of CMOS nodes (figure 3), the limited amount of laser source power restricts the throughput rate to 88 Mops/s (C) for a frequency of 80 MHz. The effect of the pixel complexity within the same MESFET logic family can be seen in figure 4. The power dissipated is essentially of electronic nature because of the buffered FET logic (BFL) family. The more complex node dissipates around 600 mW per node (solid line) whereas the simpler node dissipates only 53 mW (dashed line). This time, the heat removal capability dictates the maximum number of nodes possible on the chip. In the first case, a 4 × 4 array at 320 MHz produces a throughput rate of 38 Mops/s (D) whereas 100 Mops/s are possible for a 16 × 16 array at 230 MHz (E).

3. Conclusion

Optimum processing throughput rates of an optically interconnected bitonic sorter have been calculated for different pixel complexities and electronic node sizes. It has been shown that, in systems for which the electronic power dissipation of the pixels is negligible, the power of the laser source limits the throughput rate. High power, highly modulated laser sources are desirable in order to increase the throughput rate. The FET-SEED pixels, which dissipate a large amount of power must be kept simple in order to approach the performance of the CMOS-based node array. If Directly Connected FET Logic (DCFL) FET-SEED nodes are to appear, they will provide an improvement on CMOS-SEED devices for the same node complexity.

References

[1] E. Gluch et al., 1993, J. Mod. Opt., Vol. 40, **10**, 1857.
[2] F.E. Kiamilev et al., 1991, J. Light Tech., Vol. 9, **12**, 1674.
[3] D.T. Lu et al., 1992, Opt. Quant. Electron., **24**, S379.
[4] T.J. Cloonan, 1993, IEEE J. Quant. Electron., Vol. 29, **2**, 619.
[5] B.S. Wherrett et al., 1992, SPIE Proc. on Optical Computing, Vol. 1806, 333.
[6] D.E. Knuth, 1973, "The art of computer programming", Vol. 3, *Addison Wesley*, Reading.
[7] D.A.B. Miller et al., 1985, IEEE J. Quant. Electron., Vol. 21, **9**, 1462.
[8] A.L. Lentine et al., 1992, IEEE J. Quant. Electron., Vol. 28, **6**, 1539.
[9] L.A. D'Asaro et al., 1993, IEEE J. Quant. Electron., Vol. 29, **2**, 670.
[10] M.J. Goodwin et al., 1991, J. Light Tech., Vol. 9, **12**, 1639.
[11] C.W. Stirk and R.A. Athale, 1988, Appl. Opt., Vol. 27, **9**, 1271.
[12] A.M. Fox et al., 1991, IEEE J. Quant. Electron., Vol. 27, **10**, 2281.
[13] A.L. Lentine et al., 1994, IEEE LEOS Summer Top. Meet. Digest on Smart Pixels, 36.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Logic and layout design of optically interconnected self-routing exchange/bypass node arrays

M.P.Y. Desmulliez†[1],F.A.P. Tooley†,G.J. Crowder†,B.S. Wherrett†, P.W. Foulk‡, S. Ashcroft‡, M. Coy‡, R.A. Novotny§

† Department of Physics, Heriot-Watt University, Riccarton EH14 4AS, Scotland, U.K.

‡ Department of Computing and Electrical Engineering, Heriot-Watt University

§ AT&T Bell Labs, Naperville, Illinois, U.S.A.

Abstract. The design and implementation of an optically interconnected self-routing exchange/bypass node array is described. The nodes are laid out using silicon and gallium arsenide based technologies. 900 Mbs/s of throughput rate is achievable for the sorting of 1024 8 bit-words at an operating frequency of 100 MHz.

1. Introduction.

Sorting remains one of the most commonly performed tasks in any computation [1]. It is also a task for which the use of a space-variant non-local interconnect, such as the perfect shuffle, provides an increase in performance compared to a local interconnect implementation [2]. This increase arises from the reduction in the number of clock cycles needed to route the data to their respective addresses. When used in a system performing sorting, the perfect shuffle has the important property of being stage invariant. Only one set of hardware is thus required to implement the stages of a sorting network if a recirculating design is used. Recent advances in hybrid opto-electronics now allow the construction of optical demonstrators. This is most advantageous since it was shown that the throughput rate can be enhanced by two orders of magnitude if dedicated hardware in the form of smart pixels is provided to reduce the computational time of the sorting [2]. Self-routing exchange/bypass nodes have hence been designed and laid out in semi-custom 1 μm n-well double-metal CMOS technology and in AT&T FET-SEED monolithic integration [3]. Arrays of such processing units are to be included in a perfect-shuffle interconnected bitonic sorter [4].

[1] E-mail : marc@phy.hw.ac.uk

2. The bitonic sorting algorithm.

The network described here is based on Batcher's bitonic sorting algorithm [5]. It uses arrays of opto-electronic switching nodes, each of which compares, processes and outputs two optical digital streams of bits to the next array. The interconnection pattern between two successive processing steps can be implemented by the optical perfect shuffle (fig. 1). Each node has three functions : compare the two inputs and either interchange them or let the data pass straight through [6]. The network is self-routing in the sense that the control of the path is built into the logic. In our iterative system (fig.2) a sequence of pre-determined space-variant control images provides the optical signals that specify the functionality of each node at each cycle. If the control bit is set to zero, the output D (fig.1) is the maximum of the two inputs (A,B). If the control is set to the logic level 1, the maximum of the inputs is output to C. A second control bit prevents the node from comparing the data; the outputs (C,D) are then the images of respectively (A,B).

3. Logic and layout design of the switching node.

AT&T have developed a technology that integrates onto the same chip S-SEEDs and FET logic circuitry [3]. A 4x2 array of 2x2 switching nodes has been fabricated. The high value of the scale parameter, the availability of only depletion mode FETs and the complexity of the design have produced a node of area 560 μm by 280μm. The power density of the chip, estimated at 34Wcm^{-2} necessitates forced cooling capabilities.
Reduced node (smart pixel) area and reduced power demands are achieved using SEED technology for input optical detection/output modulation, combined with CMOS electronic circuitry. Solder bump flip-chip bonding will be used to interface the two technologies [7]. Figure 3 shows the silicon test node which is implemented in European Silicon Structures (ES2) 1.0 μm CMOS n-well double metal technology via Eurochip. The node, of area 370 μm by 230μm, encompasses two 8-bit long shift registers which store the two words at each processing step. A power density of 10 Wcm^{-2} at 100 MHz has been estimated which accounts for the optical losses of the system, the electrical power consumption, the power absorbed by the receivers and the modulation efficiency of the modulators.
The functionality of the node is described in figure 4. Two bit-serial words, A and B, are clocked through the node, most significant bit first. The first bit difference between the inputs is detected by the comparator and memorized by the following latch. A reset signal has set this latch in a neutral state before the application of the inputs. A control signal, P, is used to establish whether the minimum ot the maximum of the inputs is to be output at C (and the max/min at D). Alternatively a bypass signal is introduced in order to over-ride P and prevent data exchange. The output latch stage provides the state of the switch, which controls the multiplexer. The multiplexer then routes the input signals to the outputs C and D. The electrical feedback between the output stage and the comparator ensures that, having latched as a consequence of a bit-difference, the state of the latch is not altered by subsequent opposing bit-differences, until the end of the words. The remaining word bits are then clocked through the node until the application of the reset, P and/or bypass signals for the processing of the next words.

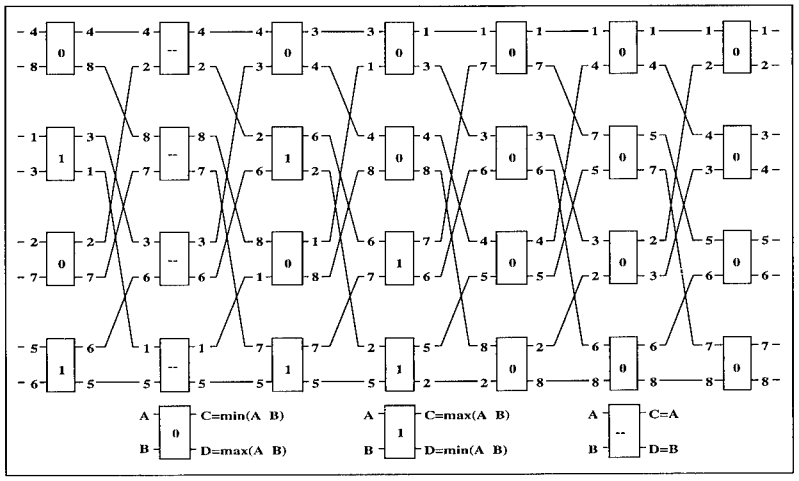


Figure 1 : Bitonic sorting with Perfect Shuffle interconnect

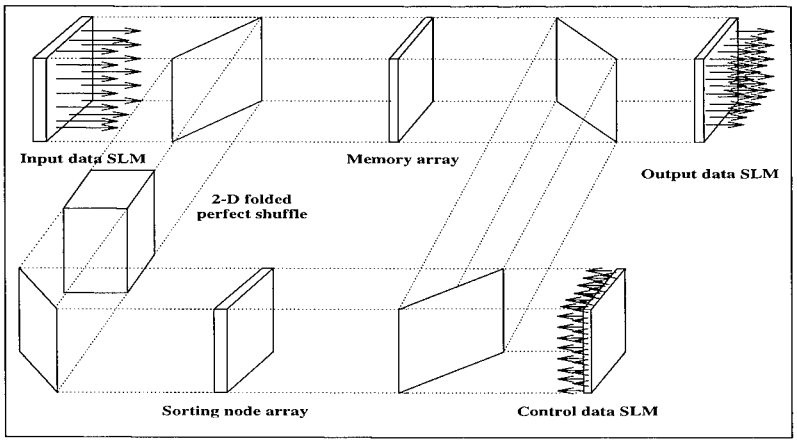


Figure 2 : Schematic implementation of the bitonic sorter

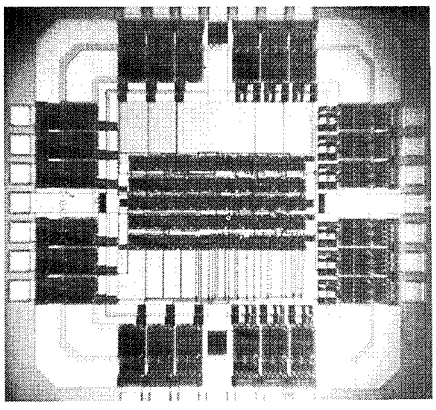


Figure 3 : CMOS test node

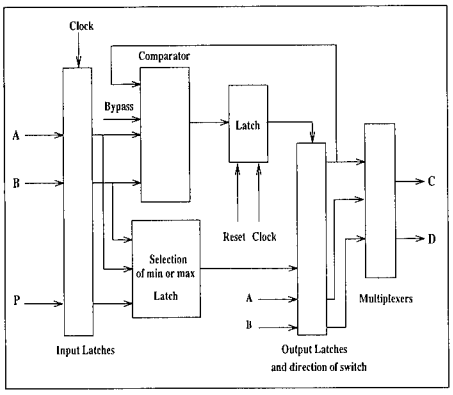


Figure 4 : Node functional schematic

4. Data bottleneck and laser power requirement.

The main components of the sorter are the input spatial light modulator (SLM), the memory array, the sorting node array and the perfect shuffle implementation (figure 2). The SLM is taken to be an electrically addressed detector/modulator array, based on SEED technology, and flip-chip bonded onto the CMOS electronic circuitry [7].
The CMOS circuitry is capable of a 100 MHz operating frequency, the sorting of 32 by 32 words of 8 bits therefore requires 9 μs. The set of words is to be presented as a 32 by 32 array for each of the eight bit-planes. The SLM must then output eight images, at 10 ns intervals, once every 9 μs. Each SLM pixel must be provided with at least 8 bits of memory in order to achieve this. 32 lines of electronic inputs, each working at 28 MHz are sufficient to provide the refresh rate demanded. This is within the capabilities of electronics and thus there is no data bottleneck at the input SLM. To ensure a 100 MHz operating frequency at the sorting node, the incident power levels must exceed aroung 80 μW per beam. Taking into account the optical losses within the system, the SLM read lasers then need to have a power greater than 1.4W. The low duty factor and use of CMOS drive circuitry means that the SLM would not require special heatsinking. A potential data bottleneck does occur however at the control SLM since this modulator array needs to be refreshed every clock cyle; solutions have been found to eliminate this bottleneck.

5. Conclusion.

Exchange/bypass switching nodes have been designed, fabricated and tested. They provide the first elements of a perfect shuffle interconnected bitonic sorter. The next generation of these chips will be flip-chip bonded onto S-SEED modulators and receivers. The bit output data rate of 32x32 words of 8 bits to be sorted at 100 MHz operating frequency is 900 Mbs/s. This corresponds to 10^{13} gate-Hertz or 10^{11} pin-Hertz.

References

[1] D.E. Knuth, 1973, "The art of computer programming", Vol. 3, *Addison Wesley*, Reading, MA.

[2] B.S. Wherrett, J.F. Snowdon, S. Bowman, A. Kashko, 1992, "Digital optical circuits for 2-D data processing", *SPIE Proceedings on Optical Computing*, Vol. 1806, pp. 333-346.

[3] A.L. Lentine, D.A.B. Miller, 1993, "Evolution of the SEED technology : bistable logic gates to optoelectronic smart pixels", *IEEE J. Quant. Electron.*, Vol. 29, **2** , pp. 655-669.

[4] C.W. Stirk, R.A. Athale, 1988, "Sorting with optical compare-and-exchange modules", *Appl. Opt.*, Vol. 27, **9** , pp. 1721-1726.

[5] K.E. Batcher, 1968, "Sorting networks and their applications", *AFIPS Proc.*, Spring Joint Computer Conf., Vol. 32, pp. 307-314.

[6] S.G. Akl, 1985, "Parallel sorting algorithms", *Notes and Reports in Computer Science and Applied Mathematics*, Academic Press Inc., London, pp. 61-79.

[7] D.J. Goodwill, A.C. Walker, C.R. Stanley, M.C. Holland, M. McElhinney, 1994, "Improvements in strain-balanced InGaAs/GaAs optical modulators for 1047 nm operation", *Appl. Phys. Lett.*, Vol. 64, **10** , pp. 1192-1194.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Architecture of a Field Programmable Smart Pixel Array

Ted H. Szymanski and H. Scott Hinton

Department of Electrical Engineering, McGill University
Montreal, PQ, Canada, H3A 2A7
email: teds@macs.ee.mcgill.ca, hinton@photonics.ee.mcgill.ca

Abstract. The architecture and complexity analysis of a 2 dimensional array of smart pixels which can be user programmed to implement arbitrary functions, including binary sorting or switching nodes or *HyperPlane* switching nodes, is described. A large array where each pixel can be programmed independently may require a hundred thousand bits of control memory. A programmable architecture which exploits the symmetry inherent in smart pixel arrays and reduces the control memory requirements by a few orders of magnitude is proposed. A complexity analysis indicates that mask-programmable arrays with the equivalent of 112×112 pixels and field programmable arrays with the equivalent of 64×64 pixels are achievable on a 1 cm^2 MOS die. The use of programmable arrays can automate the VLSI mask design process, can allow for cost reduction through larger production volumes, and can allow for standardization in pixel pitch, optical and opto-mechanical packaging assemblies, electronic packaging assemblies and other supporting technologies.

1. Programmable Smart Pixel Array Architectures.

The architecture of a 2 dimensional array of smart pixels which can be user programmed to implement arbitrary functions, including binary sorting or switching nodes or *HyperPlane* switching nodes, is described. Each programmable smart pixel consists of optical IO, electrical IO with neighboring smart pixels and/or bonding pads, flip-flops for internal state information and some type of programmable logic device to implement a finite state machine. The programmable logic can be implemented a number of ways, leading to classes of programmable smart pixel arrays. A "*Mask Programmable Smart Pixel Array*" is one where each pixel includes some type of VLSI mask-programmable device, such as a "*Programmable Logic Array*" (PLA). Mask-programmable PLAs are programmed at fabrication time through the addition of transistors in the VLSI masks. A "*Field Programmable Smart Pixel Array*" is one where each pixel can be user programmed in the field to implement arbitrary logic functions. The function of each pixel can be programmed by down-loading a bit-stream into an on-chip control memory, the technique used in the smart pixel arrays for the *HyperPlane* photonic backplane [5]. The dynamic programmability can also be achieved using RAM or ROM based lookup tables, or by adapting a mask-programmable PLA to become field programmable (the approach taken in this paper).

2. Reducing Memory Requirements in Field Programmable Smart Pixel Arrays.

A basic cell for a "*Programmable Smart Pixel Array*" is shown in fig. 1a. Each cell has four sides, *N, S, E, W*, and each side contains one or more electronic 1-bit wide *IO* ports, i.e., *In.N.1* denotes the input port on the North side with index 1. One internal structure of the cell is shown in fig. 1b. The internal finite state machine may consist of optical and electrical IO, a combinational PLA, followed by a set of D-Flip Flops with feedback. Each cell also has one or more global clock and reset inputs which are not shown in the figure. These cells can be stacked into 2 dimensional arrays. The ratio of electrical to optical IO bandwidth on the integrated circuit can be adjusted be varying the placement and number of electrical connections to IO bonding pads or "pin-outs". ICs with a small number of electrical pin-outs

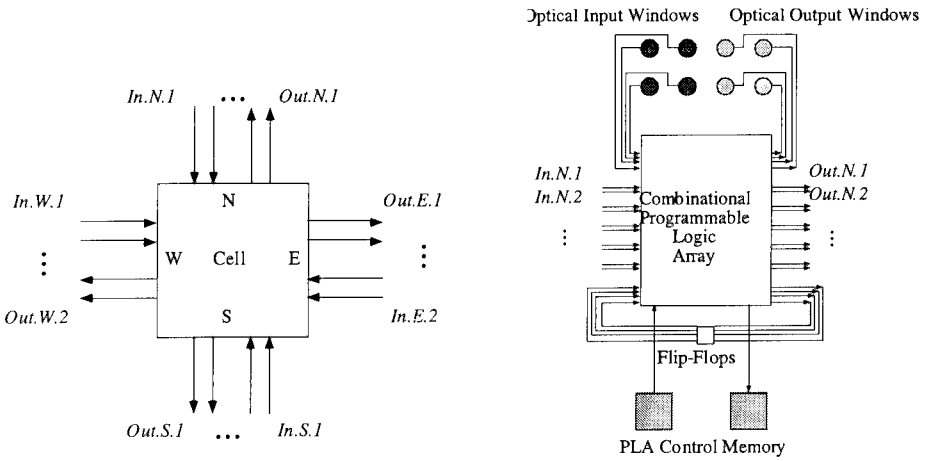


Figure I. (a) VLSI IO of a basic cell. (b) Internal structure of a basic cell.

may provide connections to pins only to cells at the perimeter of the smart pixel array through the *N,S,E,W* IO ports; ICs with many electrical pin-outs may also provide every cell with connections to pins, including cells internal to the 2D array. In a mask-programmable PLA the logic functions to be implemented by the cells are known at fabrication time and are implemented by customizing a predefined VLSI mask by inserting transistors between appropriate rows and columns in a PLA. Every cell can contain a distinct logic function since the PLA in every cell can be customized without incurring any VLSI area overhead. In mask-programmable devices the control memory shown in fig. 1b for programming the PLA is not necessary. In a field-programmable array in which cells are individually programmable, every cell will require a control memory to store the logic function for its PLA as shown in fig. 1b.

A generic model of a mask-programmable PLA is shown in figure 2. In the And-array every horizontal line represents one logic "*product*" term, i.e., the logical AND of all the inputs to which it is connected through a closed connection. In the Or-array every vertical line represents a "*sum*" term, i.e., the logical OR of all product terms to which it is connected through a closed connection. A search of the literature did not reveal any references to architectures for field programmable PLAs and hence a generic model is proposed. In a field programmable PLA a "*programmable connection*" is inserted at every intersection. Each programmable connection has 2 states, "*openned*" and "*closed*", and is controlled by a bit of control memory stored on or off the die. Closed connections are indicated by the bold boxes in fig. 2. Field programmable PLAs can be implemented in various technologies. In the GaAs FET-SEED technology one implementation of a programmable connection is a dual-gate FET transisor, implementing the logical AND function. Other implementations may chose to control the terminal voltages of a MOSFET to switch it between openned and closed states or may operate two transistors in series.

A simple individually programmable cell may contain one optical input, one optical output, one electrical IO per side and one bit of memory for the finite state machine. Such a cell would require a small PLA with 6 inputs and 6 outputs. If the PLA had a maximum capacity of 4 product terms, then 72 programmable connections per cell would be required. Each programmable connection requires only one or two transistors which represents a relatively small VLSI overhead. However, each programmable connection also requires one bit of control memory so that even a simple cell would require 72 bits of control memory. A 32-by-32 array of individually programmable cells would thus require 73,728 bits of control

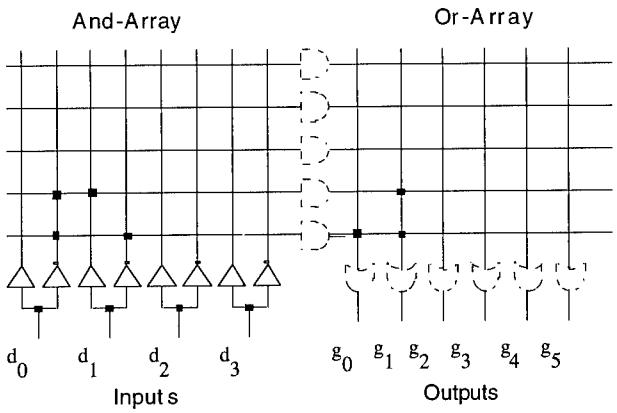


Figure II. General model of a PLA.

memory. There are three basic approaches to handle the control memory issue. If the control memory is to be stored on-chip then a large number of transistors will be required to implement the memory. If the control memory is stored off-chip and the control data is supplied through electronic pins, then a large number of electrical pins will be required. If the control memory is stored off-chip and the data is supplied on optical input windows, then a large number of optical input windows will be required. In all three cases the large size of the control memory incurs a cost. This control memory complexity must be reduced in order to yield architectures which can be manufactured with existing smart pixel technologies.

The *simplicity* and *regularity* of smart pixel arrays can be exploited to yield a field programmable architecture with a moderate control memory complexity. The programmable cells can be grouped into "*sectors*" where the logic functions of the cells in each sector are identical. The control memory used to program the cells in each sector can be shared, resulting in a significant reduction in the size of the control memory. By exploiting the symmetry in a 32×32 array of cells the control memory can be reduced in size by three orders of magnitude. A powerful programmable architecture may support a few sectors on a die, where each sector need not occupy a contigous block of cells. In the *HyperPlane* smart pixel arrays [5] a row of pixels may operate on parallel bits of data, where the most significant pixel in the row makes routing decisions that all other pixels in the row follow. In this case the 2D array should contain 2 sectors, one for most significant pixels and one for all other pixels.

3. VLSI Complexity Analysis.

Let *I, O, P* denote the number of inputs, outputs, product terms in a PLA. Let *E, L* and *D* denote the number of electrical IO, the number of optical IO and the number of flip-flops associated with the finite state machine of a programmable cell. The PLA therefore has *I=O=(E+L+D).* Each logical input signal of a mask programmable PLA in CMOS typically requires 14 λ width. Each pair of output signals adds another 14 λ width. Each pair of product terms adds 10 λ height. The pullup transistors in the And-array add 20 λ width and the pullups in the Or-array add 20 λ height. A mask-programmable PLA will require $\approx$ *(I+O)*14+20* λ width and *(P/2)*10+20* λ height. Each D-flip flop requires $\approx$ 5 NOR gates with $\approx 20\times20$ λ^2 area per gate. Each unidirectional optical IO port requires $\approx 25\times25$ λ^2. The total area of a mask programmable cell and the number of cells per side X in a 1 cm^2 die are approximately given by

$$cell-area \approx ((I+O)*14+20)\cdot((P/2)*10+20)+D\cdot2{,}000+L\cdot1{,}250.$$

$$X \approx 10{,}000/\sqrt{cell-area}\,.$$

E per cell	*L per cell*	*D per cell*	*P per cell*	*Cell-Area*	*Cell-Width (Lambda)*	*Cell-width (microns)*	*2D Array (pixels)*
4	1	1	6	12,650	113	90	111×111
8	4	4	12	50,440	225	180	112×112
4	1	1	6	20,966	145	116	86×86
8	4	4	12	119,560	346	277	64×64

Table 1. Results for various programmable smart pixel arrays.

The simplest approach to add dynamic programmability is to insert programmable connections in the PLA and share the control memory using sectors of functionally identical pixels. In this case the width of the PLA must be increased to allow the control wires to enter and exit the PLA. In general, each control wire adds $\approx 3\ \lambda$ width and $3\ \lambda$ spacing to the width of the cell. For the analysis assume that all pixels belong to the same sector and that all control wires run vertically. There will be $(2I+O)P$ control wires running vertically. Assume M layers of metalization are used to route control wires so that the increase in width is $\approx 6(2I+O)P/M$. Assuming the control memory is stored on-chip then $(2I+O)P$ D-flip flops will be needed to implement the control memory. The total area of a field programmable cell and the number of cells per side X in a 1 cm^2 die are approximately given by

$$cell-area \approx ((I+O)*14+20+6(2I+O)P/M)\cdot((P/2)*10+20)+D\cdot 2{,}000+L\cdot 1{,}250$$

$$X \approx (10{,}000-\sqrt{(2I+O)*P*2000})/\sqrt{cell-area}\,.$$

Table 1 lists the results of the analysis. The first two rows represent mask-programmable smart pixel arrays and the last two rows represent field programmable smart pixel arrays. Mask-programmable arrays with $\approx 111\times111$ pixels are achieveable through a variety of organizations, with simple cells with one pixel per cell (row 1) or with more complex cells with 4 pixels per cell (row 2). Field programmable arrays with $\approx 64\times64$ pixels are also achievable. Finally, we note a few points. Layout efficiencies may increase or decrease the pixel capacity by ≈ 25 % of the computed values. The analysis indicates that simpler cells with smaller PLAs tend to be more cost effective than complex cells with larger PLAs. For sufficiently large PLAs the area taken by the control wires becomes a dominant term and it will be more effective to place a control bit of memory at every programmable connection, yielding an architecture with individually programmable cells.

4. Conclusions.

The masks for a *field programmable smart pixel array* based upon a CMOS array of field programmable logic devices are currently being designed for proof-of-concept validation. The masks for a simpler device in the GaAS FET-SEED technology are also being designed. The CMOS die will be fabricated using the Canadian Microelectronics Corporation (CMC) fabrication runs and should be available by 1995. The developments of a *field programmable smart pixel array* would accelerate a movement towards standardization. Standardization of pixel dimensions should promote the standardization of support systems, baseplates, packaging assemblies and pin-outs.

5. References.

[1] K.W. Goosen, J.E. Cunningham, W.Y. Jan, IEEE Phot. Tech. Lett., 5, 776 (1993).
[2] T.K. Woodward, L.M.F. Chirovsky, R.A. Novotny, A.L. Lentine, "Single-Ended Operation of FET-SEED Smart Pixels", IEEE/LEOS Summer Topical Meeting on Smart Pixels - 94, pp. 24-25.
[3] M.P.Y. Desmullies, F.A.P. Tooley, J.G. Crowder, N.L. Grant, B.S. Wherrett, "Optically Interconnected exchange/bypass self-routing node arrays: logic and design layout", IEEE/LEOS Summer Topical Meeting on Smart Pixels - 94, pp. 32-33.
[4] I. Underwood, D.G. Vass, M.W.G. Snook, W.J. Hossack, L.B. Chua, "Smart Pixels using Liquid-crystal-over-silicon", Summer Topical Meeting on Smart Pixels - 94, pp. 66-67.
[5] T.H. Szymanski and H.S. Hinton, "A Smart Pixel Design for a Free-Space Optical Backplane", Summer Topical Meeting on Smart Pixels - 94, pp. 85-86.
[6] C. Mead and L. Conway, "Introduction to VLSI Systems", Addison-Wesley.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

The Scottish Collaborative Initiative on Optolectronic Sciences (SCIOS) – Devices and Demonstrators for Free-Space Digital Optical Processing

A.C. Walker, D.J. Goodwill, B.S. Ryvkin, *M. McElhinney, *F. Pottier, *B. Vogel, *M.C. Holland and *C.R. Stanley

Department of Physics, Heriot-Watt University, Edinburgh EH14 4AS, U.K.
*Department of Electrical and Electronic Engineering, University of Glasgow, U.K.

Abstract. A novel SEED-type device based on MQW GaAs/GaAlAs has been demonstrated. In addition, 16 x 16 arrays of strained-InGaAs/GaAs S-SEEDs have been successfully operated with a diode-pumped 1064 nm laser. These latter devices can be combined with CMOS electronics to create powerful smart pixel arrays. The SCIOS consortium is developing digital parallel optical processing demonstrators based on this technology.

1. Introduction

The Scottish Collaborative Initiative on Optoelectronic Sciences (SCIOS) is a partnership between Heriot-Watt, Glasgow, St. Andrews and Edinburgh Universities that has been running since 1990. It has its objectives: (i) to pursue research into the physics, materials science, fabrication and packaging technologies underlying the development of optoelectronic devices, and (ii) to combine the results of (i) with new system architecture concepts so as to realise free-space parallel optical computing and information processing demonstrators experimentally. The SCIOS technology programme includes work on III-V semiconductor devices, diffractive optics, spatial light modulators and solid-state lasers. Recent emphasis has been on smart pixel arrays including hybrid silicon/InGaAs devices. In addition, new digital free-space system architectures are being explored and a range of novel demonstrators constructed.

This paper describes recent progress in the development of digital optical switch arrays based on III-V semiconductor multiple quantum well (MQW) modulator structures and the way they can be exploited to construct parallel optical processing demonstrator systems.

2. Vertically-Integrated Detector-Modulators (VIDeMs)

SEED arrays [1] operating at 850 nm have proven to be effective as digital photonic logic planes in experiments on parallel optical processing and switching systems [2,3]. However, system speed can be limited by the degradation of SEED performance at the required high optical power [4,5] due to local heating, saturation and field screening. Woodward et al [6,7] showed that both the Joule heating associated with the photoresponse and the degradation in performance at high irradiance could be avoided by using material with a radically decreased non-radiative carrier lifetime τ in the electroabsorptive region. However, the ion-implantation method, used in refs. 6 and 7 to reduce the carrier lifetimes, broadens the exciton absorption peak and reduces the modulation performance. We have made MQW modulator structures [8] in which the carrier lifetime is already short in the as-grown structure, so that the finished device had both low photocurrent and good modulation performance. We have gone on to demonstrate how such MQW material may be used in a novel electroabsorptive optical switching device.

This device can be described as a vertically integrated detector-modulator (VIDeM), in which the two elements are connected electrically in parallel. The non-QW GaAlAs detector (lower part) has unity quantum efficiency under reverse bias. The modulator diode (top part) consists of deep GaAs/AlAs quantum wells and has extremely low quantum efficiency (η=0.4% at 15V, η=1.4% at 25V) due to the long sweep-out time across the high barriers, coupled with a short non-radiative lifetime τ. In this case, the low value of τ was, we believe, produced by low-level impurities in the MBE system at the time of growth, which fortunately did not broaden the exciton absorption feature. We hope to reproduce this effect controllably by growing at much lower temperatures.

851 nm radiation is incident on the modulator and, depending on its transmission state, a portion passes through to the detector. The quantum confined Stark effect in the MQW modulator and the Franz-Keldysh effect in the detector combine to give N-type negative resistance, leading to a factor of 3 decrease in the responsivity of the whole structure between 0V and 15V for 10μW incident optical power. This behaviour is maintained at intensities 8 times higher than for a conventional MQW SEED [4,5]. To show how this structure can be used to make a high performance reflection-mode SEED-type switch, we have recently made devices with optical output provided by a partial mirror (70% reflectivity) between the modulator and detector. A contrast ratio of 4 and a high state reflectivity of 40% are expected. Control of photo-response is likely to be an important factor in the design of future smart-pixel devices based on semiconductor modulators.

3. Strained-InGaAs SEEDs for 1064 nm

The clock-rates for demonstration parallel digital optical processing systems based on current-generation SEED arrays, operating at 850 nm wavelength, are limited by the diode laser power reaching the device array (milliwatts) The consequent demand for a better match to higher-power (> 1 watt) lasers, such as diode-pumped Nd:YLF and Nd:YAG, has prompted the development of strained InGaAs/GaAs MQW structures grown by molecular beam epitaxy (MBE) in the form of SEEDs operating at 1047-1064 nm. Previously reported devices in this materials system [9,10,11] had insufficient contrast ratio and too high operating voltage for use in switching systems. We have obtained enhanced performance, compared to our

previously reported device [11], by improved growth techniques, and have demonstrated for the first time operation of arrays of SEEDs at 1064 nm.

In our S-SEED, the *pin* MQW modulator is grown strain-balanced to a deliberately relaxed buffer layer of intermediate lattice constant, such that the InGaAs wells are in compression and balance the GaAs barriers in tension. Growth was by MBE on a [100] orientated semi-insulating GaAs substrate. To minimise the gliding of dislocations, and hence avoid broadening of the excitonic absorption peak, a low temperature was used with a low arsenic:group-III flux ratio to keep the mobility of the group-III atoms high. The buffer layer consisted of an *i*-InGaAs layer with 10-step grading of In composition from 0 to 13.5%, followed by an *i*-InAlAs/GaAs superlattice. The structure was designed to be completely relaxed at the top of the graded InGaAs layer, but X-ray diffraction measurements showed it was only 71% relaxed. The room temperature absorption peak occurred at 1061 nm rather than the design wavelength of 1047 nm. However, the variation in this wavelength was only ±0.2 nm across the central 25 mm of the 50 mm diameter substrate. The half-width at half maximum on the low energy side of this peak was 6.0 meV (5.4 nm), which is the best reported for this system. It is comparable to the figures of 6.8 meV for MOCVD-grown InGaAs/GaAsP [12] and 5.25 meV for gas-source MBE-grown in InGaAs/InGaP [13], for both of which the MQWs are strain-balanced to the substrate. It also compares well to 4.5 meV for similar GaAs/GaAlAs MQWs [1].

For initial optical testing with a tunable laser, 200 μm diameter structures were fabricated by wet etching and metallisation. The front of the sample had a partial anti-reflection coating, the back was polished but uncoated. The photocurrent characteristics showed 98% quantum efficiency at 0 V, 100% at 1 V reverse bias. The transmission-voltage performance shows a contrast ratio of 2.0 between 0 and 10 V at 1060 nm, which increases to 4.2 by the incorporation of a mirror.

S-SEED arrays, ranging in size from 16×16 ($20\ \mu m \times 20\ \mu m$ windows) to 48×96 (9216 diodes with 7 μm windows), have been fabricated. The devices incorporate a metal mirror on the top for reflection-mode operation, with the light passing through both the GaAs substrate and the temperature stabilised sapphire mount. Anti-reflection coatings have been applied at all interfaces. The whole package is compatible with our optomechanical system used for optical processing systems experiments [2]. Bistable switching with 3.0 - 4.2 contrast ratio at 10 V bias has been demonstrated for wavelengths 1054 - 1064 nm.

4. InGaAs/CMOS Smart Pixels

InGaAs SEED arrays are well suited to flip-chip assembly with a silicon electronic chip because of the substrate (GaAs) transparency at the operating wavelength. Thus smart-pixel arrays can be created with each pixel comprising: a differential detector InGaAs input stage, CMOS logic, and a differential InGaAs modulator output stage. A 4×2 smart-pixel array, capable of making numerical comparisons of two optical (bit-serial) inputs and directing them to specified outputs, has been designed as the first step towards the larger (32×32) array required for a sorting module demonstrator (see next section). The InGaAs opto-chip, which includes two diode-clamped differential inputs and two differential (modulator) outputs per node has been successfully fabricated.

The same approach can be taken to construct a high-speed spatial light modulator capable of providing parallel optical inputs to a processor at ~ 100 MHz data rate per channel. In this case the device is electrically addressed, via a serial to parallel interface, and only output modulators are required on the opto-chip. We have used the strained-InGaAs technology to fabricate a 16×16 array of differential modules. This has been flip-chip assembled (solder-bump method) onto a CMOS chip that was originally designed as the driver for a liquid-crystal SLM [14].

5. An Optical Sorting Module

The SCIOS consortium is currently constructing a digital optical sorting module based on InGaAs/CMOS smart pixel arrays, powered by diode-pumped Nd:YLF lasers and interconnected by a free-space optical perfect shuffle. The array size is 32×32 and thus it will be capable of doing a numerical sort on 1024 input words. The target specification is for a full sort, using 8-bit words, to be completed every 10 μs. This requires: an input smart pixel array operating as an electrically-addressed SLM and providing 8 binary 32×32 images in 80 ns (100 MHz) once every 10 μs; a matched smart-pixel shift register to act as the intermediate store; and a smart-pixel sorting node array capable of carrying out the compare and exchange operations demanded by the algorithm on the 10 ns per bit-plane timescale (100 MHz). Calculations of optical input requirements indicate that 1 - 1.5 watts of laser power are required per array.

Production of the optoelectronic, optical and laser components required for this system is in an advanced state. Comparisons with electronic parallel sorting hardware, which is restricted to local interconnections between nodes, indicate that this optical approach should be significantly faster.

We thank Y.P. Song, G. MacKinnon, G.S. Buller, N. Watson, N. Ross and M.R. Taghizadeh for help with fabrication of the SEEDs and holographic array generators. This work was supported by the UK Science and Engineering Research Council under the Scottish Collaborative Initiative on Optoelectronic Sciences (SCIOS). BSR acknowledges the support of an SERC Visiting Fellowship.

References

[1] A.L. Lentine, D.A.B. Miller, IEEE J. Quant. Electron., **QE-29**, 655 (1993) and references.

[2] F.A.P. Tooley, S. Wakelin, Appl. Opt., **32**, 1850 (1993).

[3] F.B. McCormick, T.J. Cloonan, A.L. Lentine, J.M. Sasian, R.L. Morrison, N.G. Beckman, S.L. Walker, M.J. Wojcik, S.J. Hinterlong, R.J. Crisci, R.A. Novotny, H.S. Hinton, Appl. Opt., **33**, 1601 (1994).

[4] A.M. Fox, D.A.B. Miller, G. Livescu, J.E. Cunningham, W.Y. Yan, IEEE J. Quant. Electron., **QE-27**, 2281 (1991).

[5] T. Sizer, R.E. LaMarche, T.K. Woodward, Appl. Phys. Lett., **61**, 420 (1992).

[6] T.K. Woodward, B. Tell, W.H. Knox, J.B. Stark, Appl. Phys. Lett., **60**, 742 (1992).

[7] T.K. Woodward, B. Tell, W.H. Knox, J.B. Stark, M.T. Asom, in *Quantum Optoelectronics Technical Digest*, 1993 (Optical Society of America, Washington, DC, 1993) Vol. 8, p. 20.

[8] B.S. Ryvkin, D.J. Goodwill, A.C. Walker, C.R. Stanley, F. Pottier, M.C. Holland, Appl. Phys. Lett., **64**, 117 (1994).

[9] T.K. Woodward, T. Sizer II, D.L. Sivco and A.Y, Cho, Appl. Phys. Lett., **57**, 548 (1990).

[10] I.J. Fritz, D.R. Myers, G.A. Vawter, T.M. Brennan and B.E. Hammons, Appl. Phys. Lett., **58**, 1608 (1991).

[11] D.J. Goodwill, A.C. Walker, C.R. Stanley, F. Pottier, M. McElhinney, Appl. Phys. Lett., **64**, 1192 (1994).

[12] J.E. Cunningham, K.W. Goosen, M. Williams, W.Y. Jan, Appl. Phys. Lett., **60**, 727 (1992).

[13] J.W. Kim, C.W. Chen, T.J. Vogt, L.M. Woods, G.Y. Robinson, D.L. Lile, IEEE Phot. Tech. Lett., **5**, 987 (1993).

[14] D.C. Burns, I. Underwood, A. O'Hara, D.G. Vass, this conference, paper MC6 (1994).

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Smart Pixels with VCSELs: Potential and Demonstration System

K. Zürl, E. Gluch, B. Schäfer, J. Schwider,
Universität Erlangen, Physikalisches Institut V, Staudtstr. 7, D-91058 Erlangen, FRG;
D. Fey, W. Erhard,
Universität Jena, Lehrstuhl Rechnerarchitektur und -kommunikation,Ricarda-Huch-Weg 16, D-07743 Jena, FRG

Abstract. A selfrouting switching network using VCSELs, microoptical array components and an opto-ASIC has been demonstrated. The setup of one entire stage with 40 channels is only 50mm long. The frame rate is 2 MHz, limited by the matrix-addressing of the VCSELs. A 7 stage 8 channel selfrouting network has been mapped onto the setup and bit error measurements at closed loop operation including re-addressing showed a BER better than 10^{-10}. Model calculations show that on the order of 10^{12} instructions per second can be achieved with such systems.

1. Introduction

In the last few years, progress had been made in building vertical cavity surface emitting laserdiode (VCSEL) arrays [1], microoptical elements like microlens arrays [2] and small hologram facets [3], and the integration of photodiodes, analog and digital electronics on silicon chips [4]. Using these devices allows for setting up small and powerful smart pixel systems, besides the approach using fast modulators like SEED arrays [5].

In this paper, we evaluate the potential of the hardware with a simplified model. A setup for a selfrouting switching network has been successfully demonstrated. The realized 40-channel-stage is only 50mm long.

2. The smart pixel model

The power of smart pixel systems does not primarily arise from *fast* data processing but from *parallel* processing (in the data planes) and *pipelining* (in the data flow direction). Therefore, a system concept has to enable "large" systems with many pixels in each data plane and with many data planes. It is mandatory to choose a modular, hierarchical system, where (a) each module can be preadjusted and (b) many modules can be arranged together with simple (final) adjustment. This implies a specified interface between the modules (fig. 1), which might be collimated beams with a pitch of 250µm. For example, for a shuffle/exchange switching network, only two types of modules are necessary:
(a) the switching module. It provides electronic switching and selfrouting logic

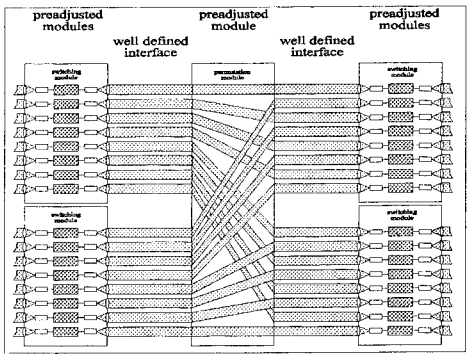


Fig.1: Modular system.

and the optoelectronic interfaces. Since the packaging in the suggested model is based on a hybrid back-to-back mounted concept (GaAs for the VCSELs, Si for the receiver and logic), some mechanical "elasticity" can be implemented e.g. where electrical bonding wires are located. Flip chip mounting is therfore not recommended (besides, thermal problems are more difficult to solve). Each module might have 8*8 channels and they can be stacked in each data plane,
(b) the permutation module, which is a *global* optical interconnection. It is identical for all stages.

If the system shall be enhanced, e.g. towards more channels of the shuffle/exchange network, both more pixels in each plane and more planes are required. This can be achieved by adding more of the (same) switching modules, however, new permutation modules have to be introduced.

Other applications, like 3D computer architectures, bit algorithms etc. are discussed in another paper [6].

3. Limitations

Three basic effects limit the performance of smart pixel systems [7,8]:
(a) the optical system, i.e. diffraction and crosstalk (due to stray light, reflections, unwanted diffraction orders etc.) limits the channel density (since only a certain amount of crosstalk is tolerable);
(b) the thermal power which can be removed per unit area from the system limits the data throughput per unit area and the connectivity;
(c) the available area per channel limits the functionality of each pixel.
Besides these three effects, which are linked, tolerances e.g. of the mechanical alignment, have to be considered.

We made calculations for the following (ideal) system:

Space variant optical interconnections, any possible connection between two adjacent pixelplanes should be posssible, the max. beam deflection angel is 45° and the max. F number of the microlenses is F# = 1.

Thus, the distance between two adjacent planes is $\sqrt{2}$ a where a is the lateral dimension of a pixelplane. Due to diffraction limit, the lateral dimension D of one optical channel (i.e. microlens, hologram facet etc.) is [7]

$$D = \sqrt{2\sqrt{2}a\lambda}$$

The pitch (and the available area) of a single pixel with one optical channel attached (which we called elementary pixel) is thus given. For a pixelplane with a=50mm, the lateral dimension of an elementary pixel is D ≈ 350μm (with λ = 850 nm) or, approx. n ≅ 20.000 elementary pixels per plane are possible.

To estimate the functionality of each pixel, an advanced technology, like it is involved at today's microprocessors like Intel's Pentium, was taken. It provides approx. 10^4 transistors/mm^2. A 8 bit macropixel, consisting of 8 elementary pixels, can have on the order of 10^4 transistors. (A two input gate needs 4 transistors, one bit static RAM memory can be realized with ≅2 transistors.) The complexity of one macropixel corresponds to 0.1 IPS (instruction per second) per clock cycle, e.g. a 32 bit adding can be carried out in 10 clock cycles. Here, the same assumptions as in chapter 2 had been made and that the area for the photodiode, preamplifier etc. is negligible.

Taking a system with a clock rate c of 100 MHz and m = 100 data planes, the overall performance is:

$$m \cdot \frac{n}{8} \cdot c \cdot 0.1 \frac{IPS}{Hz} \approx 10^{12}\, IPS$$

The heat which would have to be removed is approx. 8W/cm^2 (liquid cooling) (derived from Pentium data).

All above calculations are very rough estimations, since any constraints for mechanical mounting, power supply rails, cooling, misalignment etc. are not considered. Besides, it is critical to compare a SIMD-like architecture with commonly used MIMD architectures. So, the results here can only be considered as an estimation of the order of magnitude.

4. Experiment

To demonstrate that small systems can already today been set up, in our lab one stage of a selfrouting switching network has been set up (fig.2,3). It consists of a 10 * 10 VCSEL array (provided by R. A. Morgan), a microlens array to collimate the beams, a Kepler telescope (because the pitches of the VCSELs and the receivers are different), the holographic permutation element using close cascade technique, another microlens array and an opto-ASIC with the photodiodes and the logic. Microlenses, holograms and ASIC had been designed / manufactured at our institute. Due to problems with the drivers for the matrix-addressed VCSELs, the channel data rate was limited to 2 MBit/(s*channel).

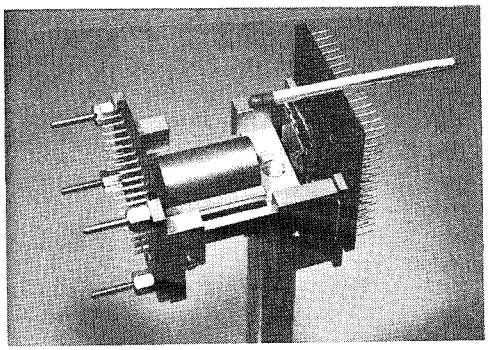


Fig. 2. Photo of the experimental setup.

A 7 stage, 8 channel selfrouting sorting network, which implemented Batcher's algorithm, has been mapped onto the setup (fig. 4). Since only one VCSEL array was available, the input data was fed onto a discrete laser diode array. A 2-D fiber bundle with different pitches on each side [9] (shape converter) conducts the light onto the first opto-ASIC (first stage with electronic exchange/bypass switches). The shuffle of the second stage had to be wired electrically, because only five stages (stage #3 to stage #7) are available in the miniaturized setup (fig. 2).

The average light power fed into the setup of each differentially coded channel was 11.2μW, the optical losses (most due to fresnel reflections) were approx. 13dB. With two interfaces to a PC (fifo-memory-like), BER measurements have been carried out at the 7 stage system, including readdressing. The BER was better than 10^{-10}.

The size of the setup (fig. 2,3) is dominated by the telescope. If the pitches of the emitters and receivers would be equal, such a stage would be only 10mm long. With a parallel addressable VCSEL array and a faster receiver (under development), a data rate of 200 Mbit/(s*channel) seems feasible.

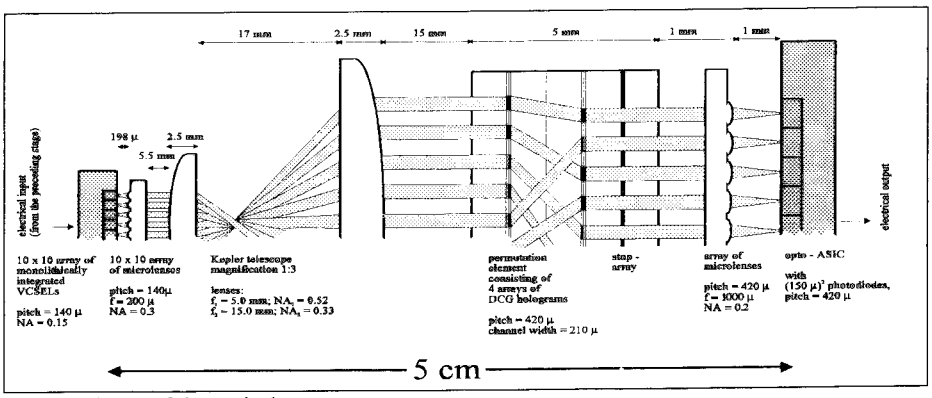


Fig. 3: Scheme of the optical setup.

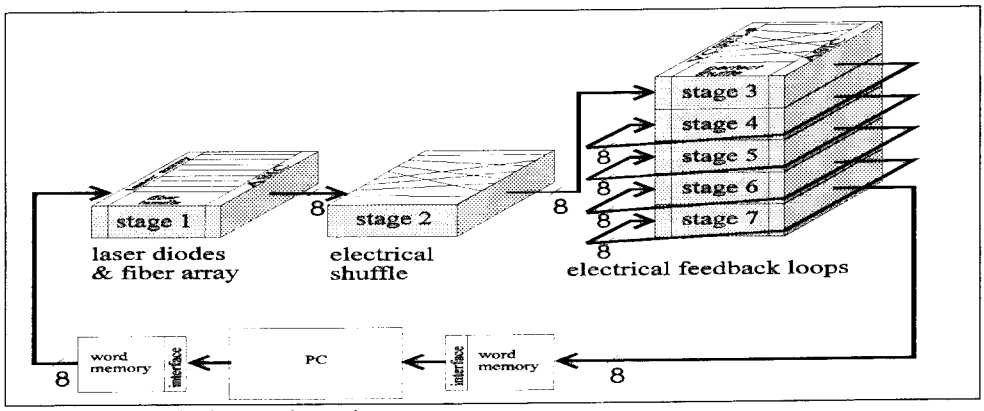


Fig. 4: Data flow in the experimental setup.

References

[1] R. A. Morgan et al., "Transverse mode control of vertical cavity to surface emitting lasers", IEEE Phot. Tech. Lett. **4**, 374, 1993

[2] S. Haselbeck, H. Schreiber, J. Schwider, N. Streibl, "Microlenses fabricated by melting a photoresist on a base layer", Opt. Engin., **32**, June 1993, p.1322ff

[3] H. Kobolla, J. Sheridan, E. Gluch, J. Schmidt, R. Völkel, J. Schwider, N. Streibl, "Holographic 2D mixed polarization deflection elements", J. Mod. Optics, 1993,vol. 40, No. 4, 613-624

[4] K. Zürl, N. Streibl, "Optoelectronic array interconnections", J. Opt. Quant. Electron. **24**, p. 405-414 (1992)

[5] F. B. McCormick, "Optical hardware design for free-space optical computing and switching", Proc. of ICO 16, Budapest, 1993, p.352ff

[6] D. Fey, W. Erhard, K. Zürl, "Optoelectronic 3-D Architectures and appropriate Algorithms for TFLOP Computing", Proc. IOC '94

[7] J. Schwider, W. Stork, N. Streibl, R. Völkel, "Possibilities and limitations of space-variant holographic optical elements for switching networks and general interconnects", Appl. Opt. 31, 7403-7410 (1992)

[8] K. Zürl, "Limitations and scaling laws in parallel optoelectronic interconnections", book chapter in: P. Lalanc, P. Chavel, "Perspectives for Parallel Optical Interconnects", Berlin, 1993

[9] U. Danzer, J. Schwider, K. Zürl, „Two dimensional optoelectronical bus type interconnection network", CLEO 94, Amsterdam, The Netherlands.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A Smart-Pixel Optical Neural Network Design using Customised Error Propagation

A. J. Waddie and J. F. Snowdon

Department of Physics, Heriot-Watt University
Riccarton, Edinburgh EH14 4AS, Scotland, UK
Tel: (0)31 451 3068; e-mail: andrew@phy.hw.ac.uk

Abstract. Smart-pixel technology provides an efficient method of implementing optoelectronic error-backpropagation neural networks. Quasi-linear detector/modulator circuits are required to provide the interface between optics and electronics. A design to perform the learning algorithm used by the S-SEED network is simulated using $1\mu m$ CMOS electronic and FET-SEED circuits. Simulations of the S-SEED neural network demonstrate the existence of 'windows' of optimal learning behaviour with respect to the neuron contrast ratio.

1. Thresholding Neuron Devices and Error Propagation Algorithm.

The development of electronic artificial neural networks has opened up considerable interest in the suitability of optics to provide the high interconnection bandwidth required to full exploit the advantages inherent in the neural paradigm [1] - [3]. A network has been devised which uses optoelectronic devices in every section of the network.

The devices employed as the thresholding neurons are S-SEEDs operating as simple buffers [6]. The S-SEED array is subdivided into three neural layers, the input, hidden and output layers described in the classic multilayer perceptron [4]. Although these layers are architecturally distinct they are contained on the same physical plane. The algorithm used to train the network is a modification of the error back propagation algorithm [4] [5]. Given N neurons the amount that the weight matrix is altered by is (z denotes an output unit)

$$\begin{aligned} \Delta W_{ij} &= \eta\Big(\sum_{i'}^{outputs}(\mathcal{R}(T^S_{i'}) - \mathcal{R}(O^S_{i'}))\mathcal{R}(O^S_j) \qquad i,j = 1..N, i \neq z \\ \Delta W_{ij} &= -\eta(\mathcal{R}(T^S_{i'}) - \mathcal{R}(O^S_{i'}))\mathcal{R}(O^S_j) \qquad i,j = 1..N, i = z \end{aligned} \tag{1}$$

$\mathcal{R}$ is the SEED reflectivity, η is an externally set learning constant, T_i is the target for output unit i and O_i is the output from unit i.

The output from the smart-pixel array which calculates the product in equation 1 is processed optically to produce the sum over the output units required to caculate the change for non-output units.

2. Smart Pixel Electronic Designs

2.1. Linear Optical Receiver and Transmitter.

A requirement for the optoelectronic implementation of the neural algorithm is an analogue receiver/modulator interface between the optics and the electronics. The receiver subcircuit is provided by a clamped FET-SEED, with the SEED held at the higher potential above the FET, operating at the λ_1 point [6]. The clamp diodes are set to restrict the operating voltage of the SEED in the linear region of the responsivity and reflectivity of the SEED.

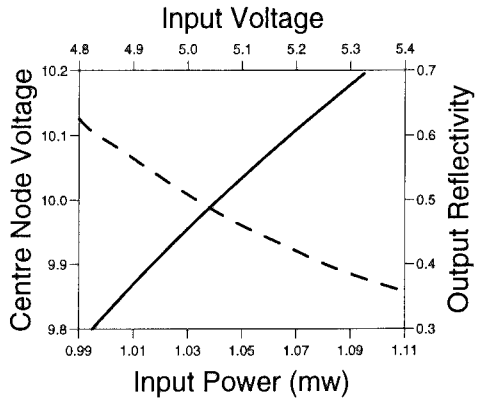


Figure 1. F-SEED Detector (Solid) and Modulator (Dashed) Responses at λ = 860 nm: Clamp diodes restrict operation of the detector within quasi-linear region of responsivity curve. Modulator read beam power = 1 mW.

The modulator subcircuit is provided by a FET-SEED with the SEED at a lower potential than the FET, again operating at the λ_1 point of the SEED. The voltage from the smart-pixel is applied to the gate node of the FET and the SEED is read.

2.2. Arithmetic Unit.

The accurate calculation of the change in weights for a particular input pattern is of fundamental importance to the efficient operation of the network. The optoelectronic interfaces described previously in conjunction with CMOS electronics provide an arithmetic unit for the calculation of the product,

$$\alpha_{ij} = (\mathcal{R}(T_{i'}^S) - \mathcal{R}(O_{i'}^S))\mathcal{R}(O_j^S) \quad (2)$$

The neural plane output vector is fanned-out by rows and then columns and the column fan-out applied as read beams to the output of the arithmetic unit to simplify the electronic design of the unit. The row fan-out is applied as the input to the smart-pixel array along with the row fan-out of the target vector. The arithmetic unit consists of a differential amplifier, a common-source amplifier and two level shifters. The circuit

was simulated using the Spice3 package for the Eurochip2 CMOS process and the results show the expected output from the modulator stage of the smart-pixel (Figure 2).

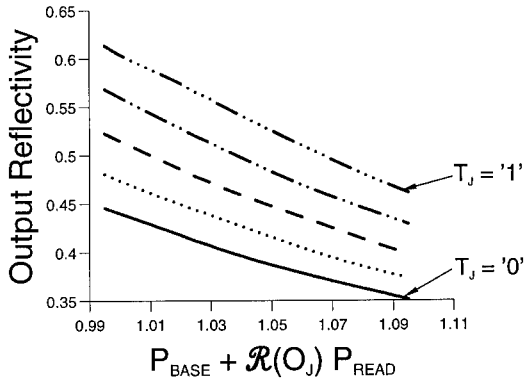


Figure 2. Smart-pixel arithmetic unit output response. $P_{BASE} = 0.9mW$; $P_{READ} = 0.1mW$. Bias power is added to the modulator read beams to ensure linear operation over the range of output voltages produced by the electronic circuit.

3. Network Simulations.

A simulation of the network was developed and a variety of simple tasks were presented to the network. To date these consist f simple Boolean functions (XOR, Half Adder), more complex Boolean functions (Full Adder, 2x2 Switching Node) and 1D pattern recognition and discrimination.

The network was configured to learn the XOR function with two input, one output and seventeen hidden units. The four input patterns of the XOR function were presented a random order to a network with an absolute zero minimum weight(Figure 3:curve 1). The effect of a non-zero minimum weight was considered with a minimum weight of 0.05. (Figure 3:curve2)

Each network configuration shows the same general behaviour, at low contrast the learning time is increasing asymptotically until at C = 1 the network requires infinite presentations to learn. For $2.5 \leq C \leq 5.5$ the learning time is minimised at around 10 presetation sets (of four patterns each) for both networks . As the contrast ratio rises the learning time begins to rise. This is due to the distance the weights have to move in the multi-dimensional weight space. At low contrast the distance moved on any one wrong output is small and so the network requires a large number of presentations to approach the correct neighbourhood in the weight space. As the contrast increases the distance moved per cycle increases towards a point which most effectively moves the weights to a correct solution. The increase in learning time observed at the upper end of the contrast ratio is produced by the weights oscillating around the correct solution in the weight space until they come within the region which will produce correct behaviour.

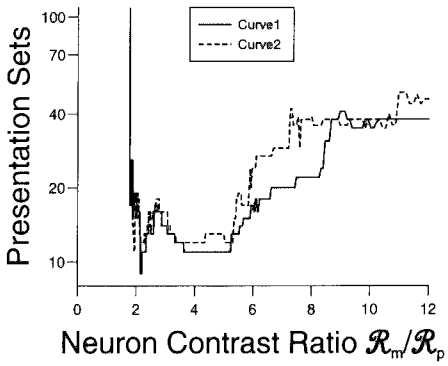


Figure 3. Optoelectronic Neural Network Simulation: Learning times with respect to output contrast ratio for different network configurations.

4. Conclusions.

We have presented the design for a smart-pixel which performs the weight update calculation for a modified error back propagation network. The pixel employs novel clamped FET-SEED detectors and modulators to produce analogue input/output stages for the analogue electronic circuit. These modulators are designed to operate considerably faster than the neural plane to minimise any bottleneck the learning process could produce.

The variation of the learning time of the network with neuron contrast ratio shows the existence of 'windows' of reduced learning time where the network weights move into the correct region in the weight space efficiently. The use of devices with non-zero minimum reflectivity in the weighting unit produces a marginal increase in the learning time for high contrast ratio weighting units.

References

[1] S. Lin, A. Grot, J. Luo and D. Psaltis, 1993, *Applied Optics*, **32**, 8, 1275-1289.

[2] C.C. Mao and K.M. Johnson, 1993, *Applied Optics*, **32**, 8, 1290-1296.

[3] P. Horan, A. Jennings, B. Kelly and J. Hegarty, 1992,*Applied Optics*,**32**,8

[4] D.E. Rumelhart and J.L. McClelland, *Parallel Distributed Processing*, 1986, MIT Press, pp. 318-362.

[5] J.F. Snowdon, 1992, *Proceedings of Int. Symp. on Opt. Appl. Sci.& Eng.*, SPIE 1773.

[6] A.L. Lentine, H.S. Hinton, D.A.B.Miller et al., 1989,*IEEE Journal of Quantum Electronics*, **25**, 8, 1928.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Dynamic studies of GaAs/AlGaAs multiple quantum well modulators using photon counting techniques

G.S. Buller, J.S. Massa, S.J. Fancey and A.C. Walker

Department of Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, U.K.

Abstract
Time-resolved photoluminescence measurements have been used to study temperature dependent carrier sweepout in biased multiple quantum well p-i-n diodes.

The multiple quantum well (MQW) p-i-n structure is commonly used in the design of optoelectronic devices, such as detectors, modulators and SEEDs. For these applications, both high speed operation, and low optical input power are often required and as such it is desirable to know which factors limit device performance in these respects. For example, the use of a MQW structure places a fundamental limitation on the response of the device, since carriers must be swept out of the quantum wells by the applied field before they can contribute to the photocurrent. The size of the device also has important implications in terms of both speed and input power. Large area devices suffer from an increased response time and input energy, due to their larger capacitance, whilst smaller devices suffer from increased surface recombination at the mesa sidewalls. In this paper we present a study of the excess carrier dynamics in biased MQW p-i-n diodes using the technique of time-resolved photoluminescence (TRPL).

The measurements were performed on a microscope based instrument (a derivative of an Edinburgh Instruments Lifemap) which has been described in detail elsewhere [1]. Sample excitation is provided at a wavelength of 765nm by a passively Q-switched picosecond AlGaAs laser diode (pulse duration <20ps)[2]. The actively quenched single photon avalanche diode (SPAD) detector [3] when coupled with the microscope optical system gives a spatial resolution of <5μm [4] and allows TRPL measurements in the spectral range 780-1100nm. The instrument uses the time-correlated single photon counting (TCSPC) technique [5] and has an instrumental full width at half-maximum (FWHM) of 50-60ps. For low temperature measurements, the samples were mounted in a continuous flow helium cryostat (Oxford Instruments model CF1104) which had been modified to allow close optical access to the sample, thus maintaining the high spatial resolution of the microscope system.

The mechanisms which govern the photoluminescence (PL) decay time from a biased MQW structure can be divided in to three categories, a) recombination, b) 2-dimensional lateral diffusion (along the plane of the wells), and c) vertical transport or sweepout through the barrier material. Recombination and diffusion will only be of significance at or near the zero internal field condition, and thus carrier sweepout will be the dominant PL decay mechanism.

For deep GaAs/$Al_xGa_{1-x}As$ quantum wells (i.e. $x > 0.2$), there are two mechanisms by which carriers escape in the presence of an electric field: thermionic emission and quantum

mechanical tunnelling. Schneider et al. [6] give the thermionic emission time $\tau_{thm}^{e,h}$ from a quantum well of width L_W for an applied field ε as

$$\frac{1}{\tau_{thm}^{e,h}} = \left(\frac{k_B T}{2\pi m_{(e,h)} L_w^2}\right)^{\frac{1}{2}} \exp-\left[\frac{H_{(e,h)}(\varepsilon)}{k_B T}\right] \quad (1)$$

where e or h refers to electrons or holes respectively, m is their effective mass in the well and T is the absolute temperature. The emission rate is determined by the barrier height, $H(\varepsilon)$, over which the particles must be emitted from a bound state in the quantum well, into the continuum of states. $H(\varepsilon)$ is given as

$$H_{(e,h)}(\varepsilon) = \Delta E_{(c,v)} - E_{(e,h)}^{(n)} - \frac{1}{2}|e|\varepsilon L_w \quad (2)$$

where $\Delta E_{c,v}$ is the conduction or valence band offset depending on whether the particles are electrons or holes and $E_{(e,h)}^{(n)}$ is the energy of the n^{th} electron (hole) subband relative to the bottom of the well.

The quantum mechanical tunnelling rate from the n^{th} sublevel has been given by Larsson et al. [7] as the product of the barrier collision frequency and the quantum mechanical tunnelling probability:

$$\frac{1}{\tau_{tun}^{e,h}} = \left(\frac{n\hbar\pi}{2L_w^2 m_{e,h}}\right) \exp-\left[\frac{2L_b\sqrt{2m_{b_{e,h}} H'(\varepsilon)}}{\hbar}\right] \quad (3)$$

L_b is the barrier thickness, $m_{b_{e,h}}$ is the particle effective mass in the barrier material, and $H'(\varepsilon)$ the effective height of the barriers in the presence of an electric field and is given as

$$H_{(e,h)}(\varepsilon) = \Delta E_{(c,v)} - E_{(e,h)}^{(n)} - \frac{1}{2}|e|\varepsilon(L_w + L_b) \quad (4)$$

where the height of the tilted barriers has been approximated by the average value.

TRPL measurements were performed on two different $GaAs/Al_{0.3}Ga_{0.7}As$ device structures, the first consisting of 71.5 periods of 100Å wells with 35Å barriers and the second consisting of 60.5 periods of 100Å wells with 65Å barriers. Figure 1 shows a 3-dimensional plot of the PL decays at a wavelength of 850nm, from the 35Å barrier device, for various applied fields between $+11kVcm^{-1}$ and $-175kVcm^{-1}$. Note the two different linear scales on the applied field axis, and that the decays have been normalised to the time-averaged PL intensity. The device mesa size was ~ 40μm×40μm and the excitation spot was determined to be sufficiently large to eliminate transverse diffusion effects. The peak photogenerated carrier density was estimated to be ~ $5\times10^{16}cm^{-3}$.

Clearly, as the reverse bias is increased from the zero applied field condition, the rate that photogenerated carriers are ejected from the wells and swept towards the corresponding doped regions increases, resulting in a reduced PL intensity and decay time. When a forward bias is applied to the device this begins to cancel the built in field (of ~$-8kVcm^{-1}$) and thus the PL decay becomes longer. To investigate the mechanisms responsible for field induced carrier emission from the wells, the mean PL decay time for each measurement was determined and the results are plotted as a function of applied field in figure 2. In the high field regime (i.e. with an applied field <$-10kVcm^{-1}$) the PL decays are predominantly single exponential, whilst for lower fields there is evidence of trap saturation and thus the decays are no longer single exponential.

In the high field regime defined above, the field dependence of the PL decay time, for both device structures, is exponential over more than a decade in applied field and two

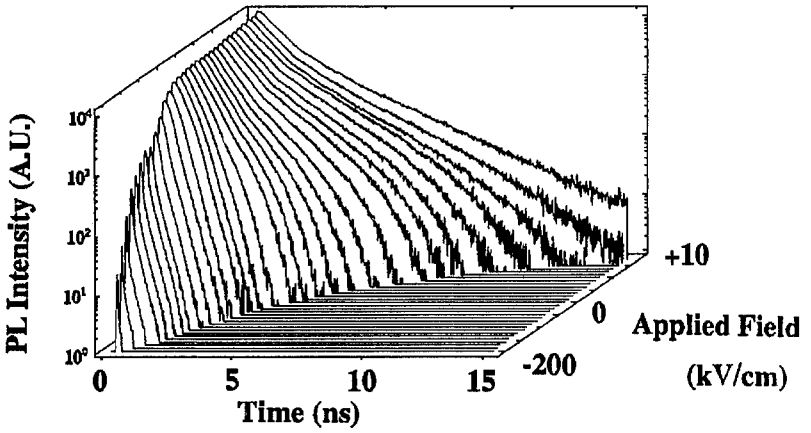


Fig. 1 *PL decays from a GaAs/AlGaAs MQW p-i-n structure as a function of applied field. The decays have been scaled to the integrated PL intensity.*

decades in decay time. This exponential dependence over such a large range in applied field would seem to eliminate tunnelling as the mechanism, since equation (3) indicates that this should be non-exponential with field. Assuming the thermionic emission model of equation (1) then the gradient of the curves for the 35Å and 65Å structures should be given as eLW/2kT. The experimentally observed values are 25.1×10^{-3}cmkV^{-1} and 31.5×10^{-3}cmkV^{-1} respectively and are in good agreement with that deduced by Cavaillès et al. [8] using a pump-probe technique on asymmetric SQW waveguides. The PL decay time for each structure appears to be directly related to the number of wells and clearly indicates the importance of retrapping of carriers, by adjacent wells, during sweepout.

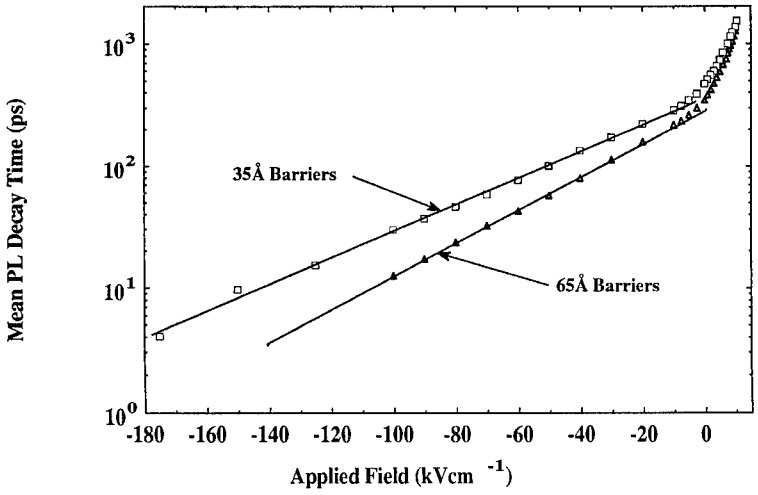


Fig. 2 *PL decay time versus applied field for the 35Å and 65Å barrier p-i-n diodes*

Figure 3 shows the PL decay time versus applied field for the 35Å barrier structure, at various temperatures. The figure indicates a general increase in the sweepout time with decreasing temperature consistent with the freezout of thermal effects. At temperatures <150K a clear minimum develops in the PL decay time at an applied field of ~25kVcm^{-1}. This would be consistent with resonant tunnelling between the heavy hole sub-bands in adjacent wells. Resonant tunnelling of electrons would be expected to occur at an applied field of ~50kVcm^{-1}. (Sweepout of a single carrier type is sufficient to terminate the PL signal).

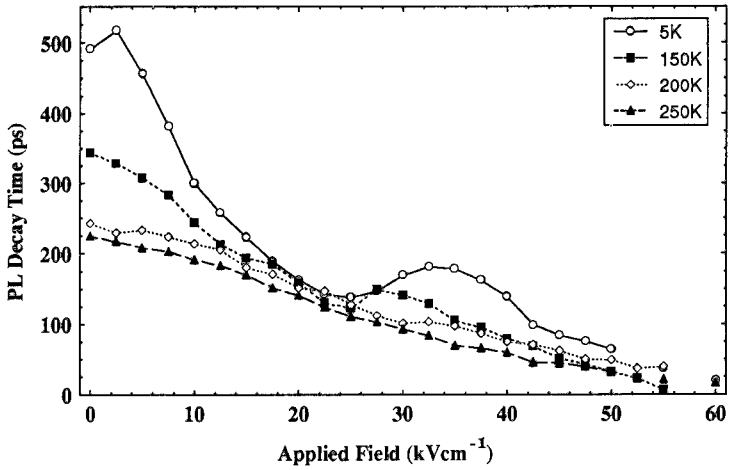


Fig. 3 *PL decay time versus applied field for the 35Å barrier p-i-n diode at various temperatures*

In conclusion, TRPL has been used to study the carrier dynamics in biased MQW p-i-n structures. In the high field limit (ie >10kVcm^{-1}), at room temperature, the sweepout mechanism appears to be of thermal origin, however, whether this is dominated by electrons or holes has not been determined. At low temperatures, there is evidence that resonant tunnelling between heavy hole sub-bands in adjacent wells is of importance.

The devices used in these experiments were on SEED array chips, commercially available from AT&T Microelectronics. The authors acknowledge the support of the Royal Society Paul Instrument Fund and the UK Science and Engineering Research Council (SERC). SJF is supported by a SERC CASE studentship with Edinburgh Instruments Ltd. The actively quenched SPADs are used by agreement of Prof. Sergio Cova and co-workers, Polytechnico di Milano, Italy.

[1] G.S. Buller, J.S. Massa, and A.C. Walker, Rev. Sci. Instrum. **63**, 2994, (1992)
[2] Z.I. Alferov, A.B. Zuravlev, E.L. Portnoi and N.M. Stel'makh, Sov. Tech. Phys. Lett. **12**, 452, (1986)
[3] A. Lacaita, M. Ghioni, and S. Cova, Electron. Lett. **25**, 841, (1989)
[4] J.S. Massa, G.S. Buller, A.C. Walker, J.L. Oudar, E.V.K. Rao, B.G. Sfez, and R. Kuselewicz, Appl. Phys. Lett. **61**, 2205, 1992
[5] D.V. O'Connor and D. Phillips, *Time-Correlated Single Photon Counting* (Academic, New York, 1983)
[6] H. Schneider and K. von Klitzing, Phys. Rev. B, **38**, 6160, (1988)
[7] A. Larsson, P.A. Andrekson, S.T. Eng, and A. Yariv, IEEE J. Quantum Electron. **24**, 787, (1988)
[8] J.A. Cavaillès, D.A.B. Miller, J.E. Cunningham, P. Li Kam Wa, and A. Miller, IEEE J. Quantum Electron. **28**, 3486, (1992)

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

The Monolithic Integration of Quantum-Well High Electron Mobility Field Effect Transistors, and Asymmetric Fabry-Pérot Optical Modulators

Garrett O'Sullivan[1], Tomás Aherne[1], Eithne Mc Cabe[1], John Hegarty[1], Paul Horan[2] , Brian Corbett[3]

[1] Dept. of Physics, Trinity College, Dublin 2, Ireland, +353-1-7022221
[2] Hitachi Dublin Lab., O'Reilly Institute, Trinity College, Dublin 2, Ireland, +353-1-6798911
[3] NMRC, LEE Maltings, Prospect Row, Cork, Ireland, +353-21-276871 extn. 4820

Abstract. We present a novel smart pixel structure which integrates a high electron mobility transistor structure within an asymmetric Fabry-Pérot optical modulator. There is no doping below the transistor layers. The structure only requires standard processing techniques. It facilitates the design of high speed, low noise smart pixel circuitry.

1. Introduction

Monolithic integration of electronic components with optical elements provides the means

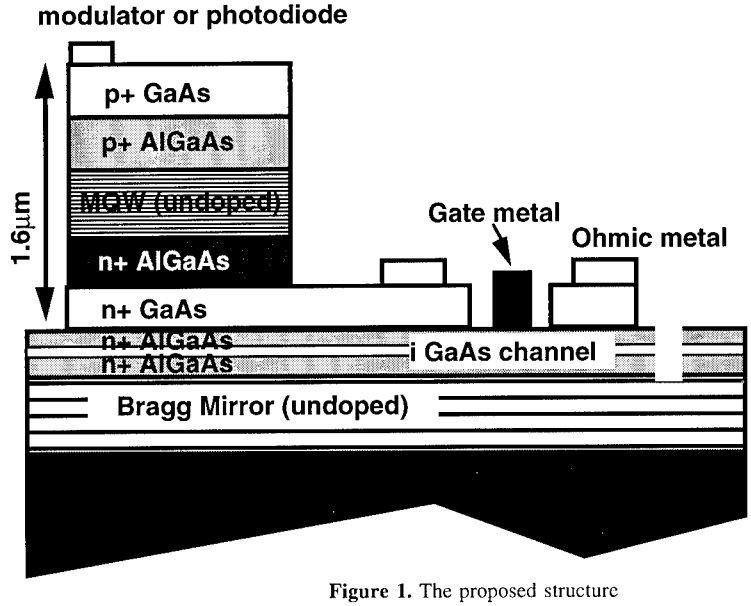


Figure 1. The proposed structure

whereby practical optoelectronic data processing becomes a reality.

Following recent interest in smart pixels [1,2,3,4] we have developed a novel wafer structure (figure 1) which combines an AlGaAs p-i-n detector/modulator structure and a quantum-well high electron mobility transistor (QW-HEMT) structure.

2. The problem of backgating

This design avoids the presence of a p-n junction under the QW-HEMT which would otherwise result in backgating [1]. Backgating reduces the output optical contrast ratio of a modulator drive circuit and gives rise to a gradual switching action. It arises when a continuous p-layer under more than one n-channel device is held at a fixed voltage. Figure 2 illustrates this problem.

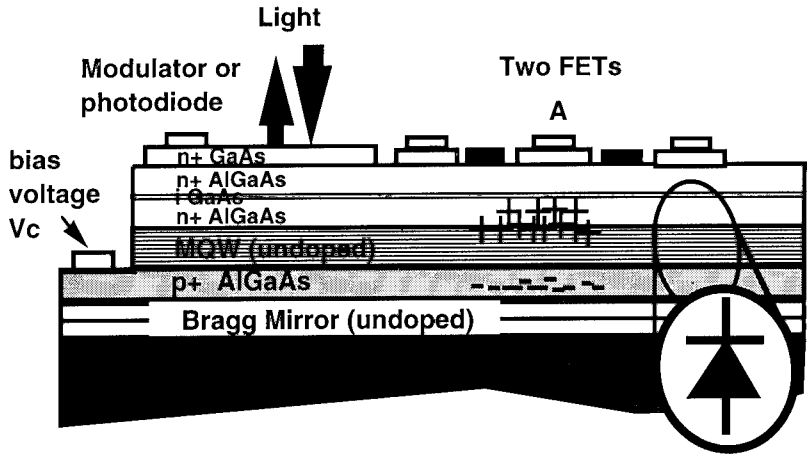


Figure 2. Illustration of backgating

The FET layers form the n-layers of the p-i(MQW)-n modulator structure. We assume the p layer is held at a constant bias and that the n-layer potential beneath point 'A' varies through the circuit action. When the voltage at 'A' rises, the diode structure beneath the FETs is reverse biassed. The depletion region associated with this diode extends upwards. The electron density in the conducting channel falls and so the circuit operation is interfered with.

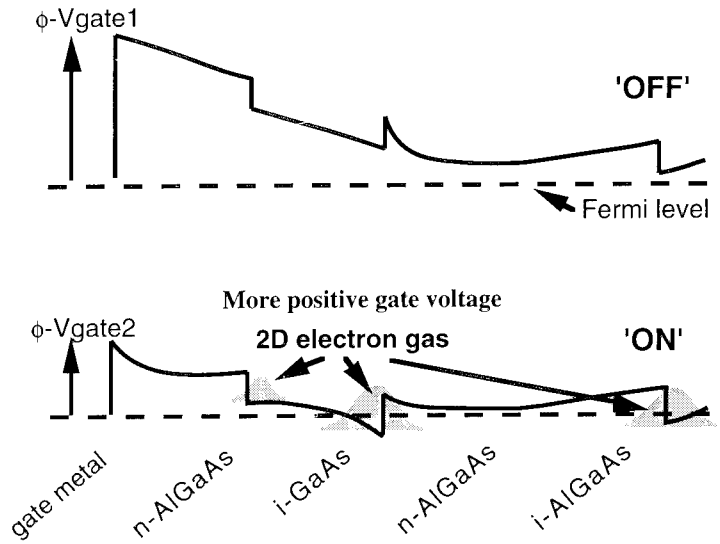


Figure 3. QW-HEMT schematic band diagrams

3. Structure description

Our solution to backgating is a modified asymmetric Fabry-Pérot modulator (AFPM) allowing high contrast reflection modulation and efficient photodetection (figure 1). The QW-HEMT layers are situated inside the AFPM cavity on top of an undoped back Bragg mirror (DBR). The QW-HEMT contacts the underside of the quantum well region directly. A deep wet etch uncovers the QW-HEMT surface for processing. This leaves the p-i-n detectors and modulators as mesas.

Two schematic conduction band diagrams of a QW-HEMT are shown in figure 3. The gate metal forms a Schottky barrier to the AlGaAs. Providing that the upper AlGaAs layer is thin enough, the density of electrons on the GaAs side of the AlGaAs-GaAs heterojuction is controlled by the voltage applied to the gate.

4. Results

Figure 4 shows the Fabry-Pérot resonance wavelength map of a 2 inch wafer of our epitaxial structure. Our required resonance wavelength (operating wavelength) is 860nm.The wafer was grown by MOCVD.

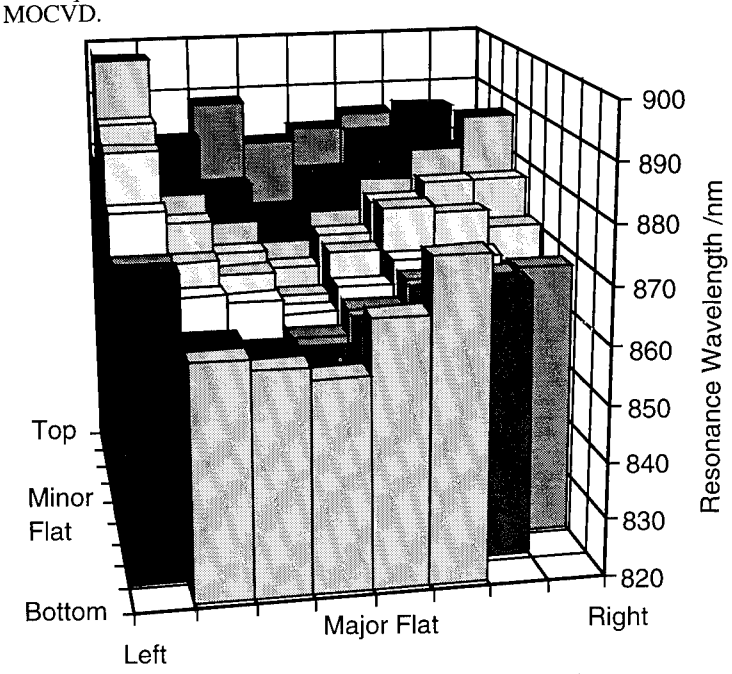


Figure 4. 2 inch wafer resonance wavelength map of complete structure

In our MOCVD grown samples the DBRs suffered from p-type impurities; this prevented transistor operation. We are now investigating MBE grown samples.

5. Discussion

We can see from the wafer map that the resonance position varies over the wafer. This certainly poses an obstacle to the use of AFPMs in commercially viable systems. Some solutions have been proposed by other groups [5,6]. The structure is flexible in that it permits etching of the front surface of the modulator in order to adjust its optical resonance wavelength. Alternatively it would be possible to deposit an additional front mirror to increase the contrast ratio of the

output modulation. For the moment we may circumvent the problem by anti-reflection coating the modulator face.

The solution to backgating used in FET-SEED systems involves the insulation of selected areas of the offending p-layer with proton bombardment. The conducting islands that remain beneath the FETs are contacted with Be implantation [3]. We expect circuit operation to be faster using our proposed design as it avoids the parasitic drain-source capacitance associated with the presence of biassed conducting material beneath FETs [7]. Our solution requires the use of the less complex processing technique of etching. Although this leaves us with a non-planar surface, optical lithography is still possible.

Electrons in HEMTs are spatially separated from their donors and move in an undoped channel so they experience less scattering. The HEMT is therefore a low noise device rendering it suitable for low noise, high speed circuitry.

Our proposal also facilitates the optimisation of the high speed performance of the AFPM [8]. This, married with the high bandwidth offered by the QW-HEMT, holds great potential for high speed smart pixels.

6. Conclusion

We have presented an alternative method of integrating FETs with multiple quantum well modulators. We suggest the use of a QW-HEMT because of its low noise and high bandwidth. We minimise FET parasitic drain-source capacitance by avoiding the presence of contacted conducting material beneath the transistor layers. We expect this will allow faster circuit operation [7].

The wafer processing required is standard; the mesa modulators and detectors are compatible with optical lithography. The mesa modulator approach has already been used to optimise AFPM bandwidth [8]. Therefore we can envisage high speed analogue smart pixel systems based on this approach.

References

[1] T.K. Woodward et al. June 1992 *IEEE Photon. Tech. Letts.* **4** (6) 614-617 " Operation of a fully Integrated GaAs-AlGaAs FET-SEED: A basic optically addressed integrated circuit"

[2] T.K. Woodward et al. August 1992 *IEEE LEOS Summer Topical Meeting on Smart Pixels, Santa Barbara, CA (IEEE, New York, NY, 1992)* **TUC3** 61-62 "GaAs/AlGaAs FET-SEED smart pixels with several transistors"

[3] L.A. D'Asaro et al. August 1992 *IEEE LEOS Summer Topical Meeting on Smart Pixels, Santa Barbara, CA (IEEE, New York, NY, 1992)* **TUC4** 63-64 "Batch Fabrication and testing of FET-SEED smart pixels arrays"

[4] B.F. Aull et al. September 1993 *Appl. Phys. Lett.* **63** (11) "Monolithic optoelectronic transistor: A new smart-pixel device"

[5] Ghisoni M et al. 1991 *Optical & Quantum Electronics* **23** (7) S915-S924 "Post-growth tailoring of the optical properties of GaAs-AlGaAs quantum well structures"

[6] Karin Z et al. 30 May 1994 *Appl. Phys. Lett.* **64** (22) 2913-2915 "Externally deposited phase-compensating dielectric mirrors for asymmetric Fabry-Perot cavity tuning"

[7] R.A.Novotny *OC'94 (Edinburgh)* private communication

[8] C.C. Barron et al. *IEEE LEOS Summer Topical Meeting on Smart Pixels, Santa Barbara, CA (IEEE, New York, NY, 1992)* **WB2** 89-90 "High-Speed Design of Asymmetric Fabry-Perot Modulators"

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Contactless Quantum Confined Stark Modulator Based on GaAs/AlAs Quantum Filters

N. T. Pelekanos[(1)(3)], B. Deveaud[(1)(a)], P. Gravey[(1)], F. Clérot[(1)], J. M. Gérard[(2)], J. Hebling[(3)], J. Kuhl[(3)]

(1) France Télécom, CNET-Lannion, BP 40, 22301 Lannion, France
(2) France Télécom, CNET-Bagneux, BP 107, 92225 Bagneux, France
(3) Max-Planck-Institut für Festkörperforschung, Heisenbergstr.1, 70569 Stuttgart, Germany

We demonstrate a novel all-optical Quantum Confined Stark modulator. The necessary electric field is switched-on optically across every period of the heterostructure, by exploiting strongly asymmetric photocarrier transfer in GaAs/AlAs layers. In an unoptimized sample, we measured at room temperature 9 meV exciton redshift at 860nm induced by optical excitation of about 50 W/cm^2. The origin of an unwanted saturation mechanism limiting the modulator performance is also discussed.

We present preliminary results on a contact-free quantum well (QW) light modulator operated all-optically, in the sense that one "write" laser beam strongly affects the optical constants in the vicinity of a QW exciton line and thus the output of a resonant "read" beam. Previous work on all-optical QW modulators has focused on n-i-p-i heterostructures [1]. The main advantage of the present design over the n-i-p-i's is that the estimated device response times are shorter by more than three orders of magnitude [2], allowing MHz modulation rates. The principle of operation of the device is as follows: every period of the heterostructure contains three QWs designed in such a way that following above bandgap photoexcitation a large fraction of the photogenerated electrons and holes tend to separate and accumulate in the exterior QWs, creating a local space-charge field having its maximum in the region in between and acting via the Quantum Confined Stark Effect (QCSE) [3] on the exciton resonance of the central QW.

In Figure 1, we show a schematic band diagram of one period of the heterostructure. The three QWs are enumerated from left to right. QW_1 is 25A of GaAs, surrounded by a graded AlGaAs (x=36-42%, 800A) and an AlAs barrier (100A). The combination of a narrow GaAs QW adjacent to an AlAs layer is the key aspect of the device. Following optical excitation with a photon energy larger than the energy gap of the AlGaAs graded barrier or QW_1 (λ_{write}<620nm or 700nm, respectively, at T=300K), this layer combination functions as a one-way "quantum filter" for the photocarriers collected in QW_1, by blocking the holes out but allowing the electrons transfer rapidly into QW_2. This is due to the fact that, for sufficiently thin GaAs layer thicknesses, the electron energy level at the Γ-point in the GaAs QW is higher than the one at the X-point in the AlAs layer. Γ–X electron transfer times on the picosecond time scale have been measured in thin GaAs/AlAs superlattices [4]. QW_2 is 150A of GaAs and QW_3 for the purposes of this report can be viewed as 80A of InGaAs (x=10%). In actuality, QW_3 consists of 70A of GaAs in which 2 monolayers of InAs islands are epitaxially inserted. The reason for this is that one projected application for this modulator is all-optical photorefractivity. The use of InAs islands is meant to eliminate lateral diffusion

of critical importance for such devices[5]. The electrons that pass through the quantum filter leaving behind holes, are able to subsequently tunnel through a thin (<50A) AlGaAs barrier into QW_3. The resulting space-charge electric field acts upon the excitonic resonance of QW_2 (λ_{read}= 860nm).

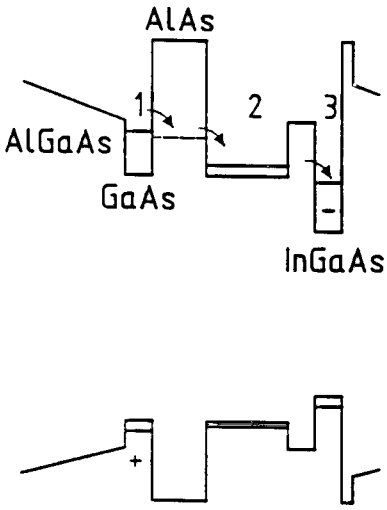


Figure 1: Schematic band diagram of one period of the heterostructure and photogeneration of electric field by rapid electron transfer from QW_1 to QW_3. The QW energy levels are denoted by full lines whereas the X-point energy level in the AlAs layer by a dotted line.

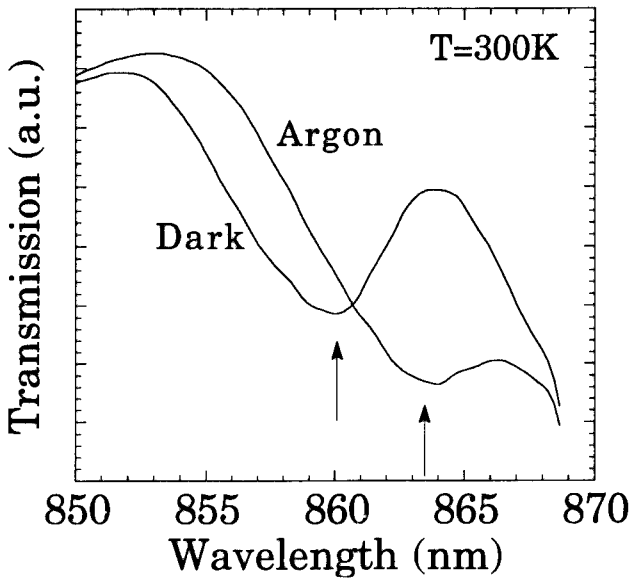


Figure 2: Redshift of the QW_2 exciton feature observed in room temperature transmission spectra, with (Argon) or without (Dark) photoexcitation by 20W/cm^2 of Argon laser blue lines.

We can directly monitor the photogenerated electric field across QW_2 by measuring the redshift of the exciton line induced by a coincident laser source. In Figure 2, we show the room temperature transmission spectra of one of our samples in the region of the QW_2 exciton, with and without photoexcitation by 20W/cm^2 of the Argon laser blue lines. The transmission exciton feature contains contributions from both heavy and light hole excitons, accounting for the asymmetric profile on the high energy side. The whole exciton feature clearly redshifts under the effect of illumination. Redshifts up to 9 meV at room temperature were registered with about 50 W/cm^2 of optical excitation. Based on available data on the QCSE-induced redshifts as a function of electric field for similar QW widths [6], we estimate that the 9 meV redshift corresponds to an effective electric field of nearly 30 kV/cm. Further increasing the incident laser intensity does not significantly increase the photogenerated electric field, indicating that the optical charging process is subject to a saturation mechanism which we will discuss next.

In Figure 3, we plot the QW_2 exciton redshift obtained in photoluminescence (PL) experiments excited by the 647 nm Krypton laser line [7], as a function of incident laser power density for various temperatures. The main observation in this figure is that the exciton redshift saturates with increasing power densities. For T=150K, for instance, the redshift saturates at a value corresponding to a photogenerated electric field of about 30kV/cm. For temperatures below 150K the curves (not shown here) overlap with the one for T=150K. On the other hand, as the temperature increases above 150K the onset of the saturation occurs at gradually smaller exciton redshifts, i.e. electric fields. The curve denoted by "1/20" is under the same conditions as the curve for T=150K but with a reduced duty cycle by a factor of 20. The fact that the two curves practically overlap excludes any scenario of heating. Finally, the curve "780nm" verifies that excitation by a laser diode with λ=780 nm does not generate electric field since it is only absorbed in QW_2 and QW_3.

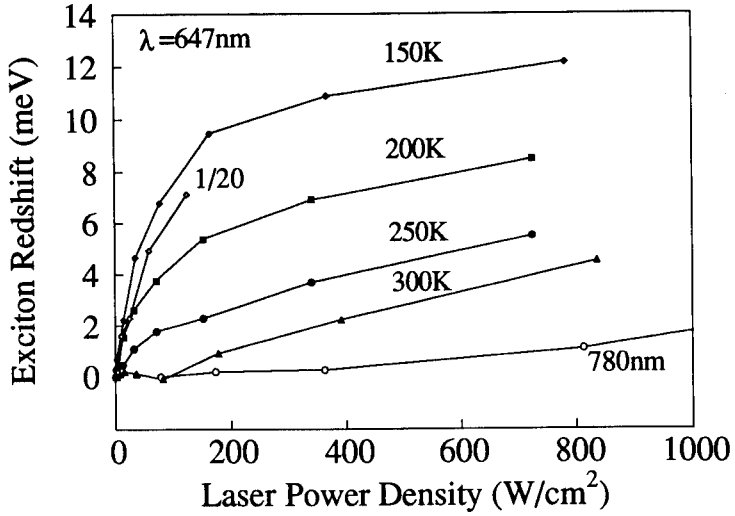


Figure 3: Exciton redshift of QW_2 as a function of incident laser power density (λ=647 nm) for various temperatures. The curve "1/20" is obtained under the same conditions as the one for T=150K but with a reduced duty cycle by a factor of 20. The curve "780nm" is obtained at T=150K with an excitation wavelength of 780 nm.

In order to optimize the modulator performance, it is important to understand the saturation mechanism. In fact, based on simple modelling calculations neglecting any saturation mechanism, we estimate that an incident laser power density of only a few W/cm^2 is sufficient to create an electric field of about 50kV/cm. In other words, comparing our experiment with the calculation, we conclude that the optical charging mechanism would be at least an order of magnitude more efficient if the saturation mechanism was absent.

We consider several possibilities for the saturation mechanism, such as: (i) quenching of the Γ–X transfer process from QW_1 to QW_2 due to electric field-induced band bending, (ii) quenching of the electron tunneling step from QW_2 to QW_3 for the same reason, and (iii) the back-transfer of electrons from QW_3 to QW_2 becomes possible at an elevated electric field. First, we note that scenarios (i) and (ii) are temperature independent, therefore they can not account for the temperature behavior of the saturation mechanism. In addition, we can exclude scenario (i) based on time-resolved PL-decay experiments [8] which indicate that the Γ–X transfer time remains smaller than 20psec (our time resolution in this case) even for photogenerated electric fields higher than 100kV/cm. Hence, the back-transfer of electrons from QW_3 to QW_2 after some value of the electric field, is most likely to be responsible for the saturation mechanism. In fact, we estimate that the electron density we can store in the lowest level of the InAs islands is only $8x10^{10}$ cm^{-2}. In order to create an electric field of 30kV/cm we need electron densities of about $2x10^{11}$ cm^{-2}. Therefore, occupation of the excited InAs electron levels must take place. The saturation mechanism may easily arise from an electric field-induced alignment of the lowest n=1 electron level of QW_2 with the first excited electron state of the InAs islands, allowing back transfer of electrons from QW3 to QW2. The zero field energy difference between the two levels is estimated to be 40-50 meV. For a spatial separation of the electronic wavefunctions of 120-150A in the samples studied, such an alignment is possible for electric fields in the range of 25-40kV/cm. This is quite commensurate with our experimental finding that the electric field saturates at 30kV/cm. Also, in this picture, the temperature behavior can be understood in a natural way since it involves two thermally distributed electron populations brought in resonance. Further investigations are under way in order to clarify this important issue.

In summary, we present a novel heterostructure device to be used as an all-optical quantum well modulator. The modulation principle is based on the Quantum Confined Stark Effect but needs no electrical contacts since it lies on the property that an incident laser is able to switch-on an electric field in every period of the heterostructure. In early experiments, we demonstrate photogenerated electric fields up to 30kV/cm with moderate laser intensities.

N.T.P. acknowledges support from the Alexander von Humboldt Foundation.

(a) Present address: Ecole Polytechnique Fédérale Lausanne, CH 1015 Lausanne, Switzerland.

[1] G. H. Döhler, Optical and Quantum Electronics **22**, S121 (1990); A. Kost, *ibid* S187.
[2] On the n-i-p-i response times see H. Ando, H. Iwamura, H. Oohashi, H. Kanbe, IEEE J. Quantum Electron. **25**, 2135 (1989).
[3] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, C. A. Burrus, Phys. Rev. B **32**, 1043 (1985).
[4] J. Feldmann, J. Nunnenkamp, G. Peter, E. Göbel, J. Kuhl, K. Ploog, P. Dawson, C. T. Foxon, Phys. Rev. B **42**, 5809 (1990).
[5] For more on the InAs islands and this aspect of the device see N. T. Pelekanos, B. Deveaud, P. Gravey, J. M. Gérard, J. Hebling, J. Kuhl, Optical Materials (1994), in print.
[6] L. Vina, E. E. Mendez, W. Wang, L. L. Chang, L. Esaki, J. Phys. C. Solid State Phys. **20**, 2803 (1987).
[7] Note that the electric field generation is significantly more efficient when exciting with the Argon blue lines compared to Krypton (λ=647 nm). The reason for this is simply that in the former case both the graded AlGaAs and QW_1 are excited, whereas in the latter case only QW_1 contributes to the optical charging process.
[8] N. T. Pelekanos, U. Strauss, W. W. Rühle, unpublished.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Smart pixel arrays: challenges of manufacturing and testing

G. Livescu, L.M.F. Chirovsky, L.A. D'Asaro, S. Hui and B. Tseng

AT&T Bell Laboratories, Murray Hill, NJ 07974

R.A. Novotny and A.L. Lentine

AT&T Bell Laboratories, Naperville, IL 60566

R.E. Leibenguth, M.W. Focht, J.M. Freund, K.G. Glogovsky, G.D. Guth and G. Przybylek

AT&T Bell Laboratories, Breinigsville, PA 18031

Abstract. Wafer level optical and electronic testing of smart pixel arrays is a must if a high yield technology is to be developed. Our results on GaAs FET-SEED based switching nodes show the uniformity and performance control levels achievable with today's technology, as well as its limits.

The fabrication of optoelectronic components must be a high yield process in order to accomplish large scale manufacturing of smart pixel arrays for applications in optical interconnects, photonic switching [1], spatial light modulators [2] and optical computing. Smart pixels consist of optical input and output devices, integrated with increasingly sophisticated circuits for gain and logic functions [3]. For example, in a recently demonstrated five stage photonic switching system [4,5], each stage consisted of an array of 4x4 routing nodes (pixels), made using the field-effect-transistor self electro-optic effect device (FET-SEED) technology [6,7]. Each node contained 25 transistors and 17 quantum well diodes; each stage (or chip) with 16 such nodes had 400 transistors and 272 diodes. Although simple FET-SEED circuits have been shown to switch in times as short as 200ps [8], and fully functional *individual nodes* operating at 400Mb/s have been demonstrated [4,5], the speed of an *entire array (or stage)* was only 200Mb/s, and that of the *entire system* was 155Mb/s, limited by the non-uniformity of the individual nodes and/or arrays. To insure uniformity of the devices and chips, wafer level electrical and optical testing is necessary. This paper presents some of our methods of mapping the performance of smart pixel arrays at the wafer level, as well as our results.

The FET-SEED technology [6,7] is based on GaAs/AlGaAs quantum well diodes and field-effect-transistors. The material grown by molecular beam epitaxy (MBE) undergoes about ten major processing steps, including: photolithography, etching and ion implantation for device definition and insulation, metal deposition for ohmic contacts, Schottky gates and electrical interconnections, and deposition of insulator layers for electrical insulation and anti-reflection coating. Planar process technology [7] allows for batch fabrication of arrays of complicated circuits, such as the 2x1 switching node schematically described in Fig.1. The node consists of an optical receiver, an inverter, a control memory and a multiplexer/driver/transmitter [4]. The electrical output A of the receiver becomes one of the inputs to the 2x1 multiplexer/driver located within the same node. The other input, B, comes from a receiver located in another node.

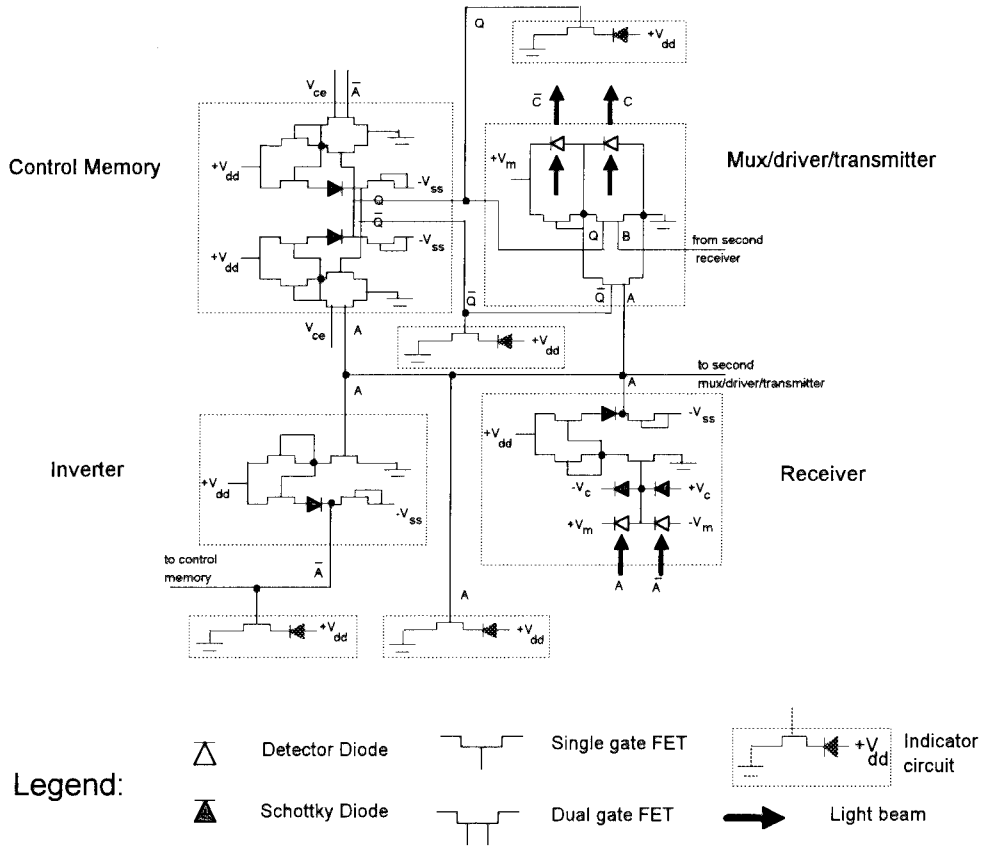


Fig.1. Schematic diagram of a single node. Notice the four "indicator lights" placed at the bottom left (the $\overline{A}$ and A), at the center ($\overline{Q}$), and the top right (Q).

Each mux/driver also has a pair of complementary electrical inputs (Q and $\overline{Q}$ in Fig.1), whose role is to control which one of the inputs A or B is regenerated as the optical output C. The control memory (set-reset latch) stores this control bit. An electrical control signal, V_{ce}, common to all the nodes in the array, is held either high, to enable writing of the memories with the control bits preceding the data bits, or low, to preserve the memory during the time the data are processed [4].

The circuits contain quantum well p-i-n diodes for detectors and modulators, Schottky clamping diodes, single-gate and double-gate field-effect transistors. Single diodes, transistors, resistors and conductors are fabricated adjacent to each array of smart pixels, for individual component characterization. By measuring their characteristics and mapping them over the area of the wafer we obtain the yields of individual devices, as well as the degree of uniformity. These yields are in general very high, of the order of 90% and better.

An example of data mapping on a 3" wafer is shown in Fig.2(a), for threshold voltages V_{th} of FET's with 10μm long gates. The standard deviation of this distribution is ΔV=100mV as shown in Fig.2(b). For 2" diameter wafers the distribution of V_{th} is narrower: ΔV is 40-70mV,

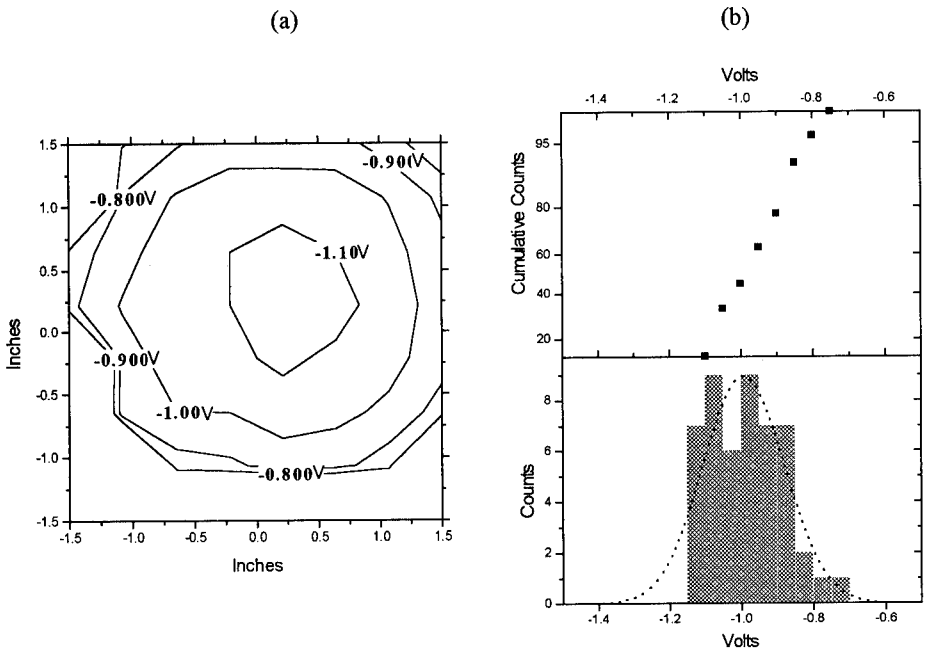


Fig.2. (a) Mapping of threshold voltages on a 3" wafer. (b) Corresponding statistical charts.

depending on the particular choice of layer and doping structure. The pattern in Fig.2(a) is circular, which indicates that the variation is likely due to thickness and/or doping variation in the MBE growth, rather than to the numerous subsequent processing steps.

When packed together into large circuits and arrays of circuits, the performance and uniformity of the individual components may suffer. Therefore it is necessary to also map the performance of the nodes. Testing the functionality of these nodes is a challenging task, as it must address both optical and electronic functions.

A simple optical test can be done using the light emitting properties of GaAs diodes. Connected in series with a FET, a light emitting diode (LED) will "light up" if the FET becomes conducting. Such a FET-LED pair thus acts as an "indicator circuit", sensitive to the voltage on the gate of the FET. By inserting indicators in critical points of the nodes, as illustrated in Fig.1, they will light-up if the voltages tested have the correct values.

The implementation of this technique for monitoring the receiver (A), inverter ($\overline{A}$) and control memory (Q, $\overline{Q}$) for an entire array of nodes is schematically described in Fig.3(a). For testing, a periodic sequence of high and low A signals is simulated electrically and coincides in time with a periodic sequence of high and low V_{ce} signals. Since only the "high V_{ce}" allows changing the state of the memory, the rate at which the Q and $\overline{Q}$ indicators light-up is half the rate at which the A and $\overline{A}$ indicators light-up.

An infrared camera and a VCR are used to record the indicator lights turning on and off. The "motion picture", a frame of which is illustrated in Fig.3(b), contains the images of all the indicators for every array on the wafer. Images like these, as well as those of especially designed "yield tester arrays" (consisting of indicator circuits only) allowed us to assess the uniformity of

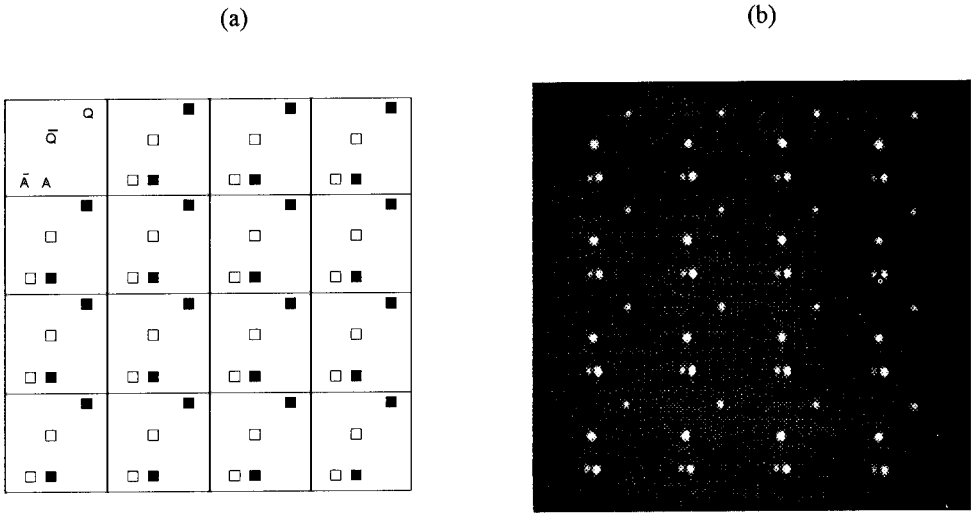


Fig.3. (a) Schematic description of the optical functional test of a 4x4 array of switching nodes. The first cell illustrates the four "indicators" which are monitored: the signal A from the receiver; the signal $\overline{A}$ from the inverter; the two control signals from the memory, Q and $\overline{Q}$. (b) The actual image of an array. In each of the A and Q pairs, one of the indicators is bright and the other is dim.

the arrays and their yield. Depending on the type of array, the functional yields vary from 5% to 50%. Note that the yields of the individual devices were larger. Our preliminary analysis of failures indicates that the vast majority of them are metal deposition errors, affecting especially the FET gates. These errors are more severe when many FET's are fabricated close together. Eliminating these errors requires improving the lift-off process used in gate fabrication, or replacing it with subtractive patterning. Using projection photolithography instead of the current contact photolithography would also help to reduce the number of metal deposition defects.

In conclusion, wafer level electrical and optical testing of FET-SEED smart pixel arrays shows the uniformity and performance levels achievable with today's technology. The high yield fabrication of densely packed FET based GaAs circuits appears to be its most challenging task.

References

[1] Hinton H S 1993 *An introduction to photonic switching fabrics* (New York: Plenum)
[2] Efron U and Livescu G. 1995 *Spatial Light modulator technology: Materials, devices and applications* (New York: Marcel Dekker Inc.) 217-286
[3] Lentine A L and Miller D A B 1993 *IEEE J. Quantum. Electron.* **29** 655-669
[4] Lentine A L, Novotny R A, Cloonan T J, Chirovsky L M F, D'Asaro L A, Focht M W, Freund J M, Guth G D, Leibenguth R E, Smith L E, Boyd G D and Woodward T K 1994, *IEEE Photon. Techn. Lett.* in print
[5] Mc. Cormick F B, Cloonan T J, Lentine A L, Sasian J M, Morrison R L, Beckman M G, Walker S L, Wojcik M J, Hinterlong S J, Crisci R J, Novotny R A and Hinton H S 1994 *Appl. Opt.* **33** 1601-1618
[6] Woodward T K, Chirovsky L M F, Lentine A L, D'Asaro L A, Laskowski E J, Focht M W, Guth G D, Pei S S, Ren F, Przybylek G J, Smith L E, Leibenguth R E, Asom M T, Kopf R F, Kuo J M and Feuer M D, 1992 *IEEE Photon. Techn. Lett.* **4** 614-617
[7] D'Asaro L A, Chirovsky L M F, Laskowski E J, Pei S S, Woodward T K, Lentine A L, Leibenguth R E, Focht M W, Freund J M, Guth G D and Smith L E 1993 *IEEE J. Quant. Electron.* **29** 670-7
[8] Livescu G, Boyd G D, Chirovsky L M F, Morgan R A and Mullally T 1994 *Opt. Lett.*, in print

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Monolithic Integrated Optoelectronic Smart Pixels

U Kehrli, D Leipold, K Thelen, O J Homan, H P Schweizer, P Seitz and B D Patterson

Paul Scherrer Institute, Badenerstr 569, CH-8048 Zurich
Tel ++41-1-492 63 50, Fax ++41-1-491 00 07

Abstract. A layer structure, grown in a single step, and a fabrication process were developed for the monolithic integration of AlGaAs/GaAs optoelectronic smart pixels. Metal semiconductor field-effect transistors (MESFETs), light-emitting diodes (LEDs) and photodiodes (PDs) were designed in such a way that the light emitted from the LEDs can efficiently be detected by the PDs. An example is presented of a fabricated optoelectronic smart pixel: a threshold circuit consisting of a dual-photodiode differential input, an inverter and a current-balanced output containing an LED. The circuit shows a switching energy of 2 pJ. The minimum switching power is < 1 nW with a contrast ratio >1000. The maximum light output of the LED is 18 μW with an overall power dissipation of 20 mW.

Monolithically integrated optoelectronic "smart pixels" are of current interest in the field of parallel optical interconnects, early vision processing and optoelectronic neural networks [1,2]. We have developed a fabrication process where, in a flexible way, different kinds of optoelectronic smart pixels can be realized in the AlGaAs/GaAs material system [3]. The material is grown in a single step by metal organic chemical vapour deposition (MOCVD). On an n^+-doped GaAs substrate, we first grow a stairstep single quantum well LED which ends with a 400 nm thick p^+-contact layer (Fig. 1). The PD/MESFET layers are grown on top of this structure. They consist of a 1 μm thick GaAs p^--buffer/absorber layer, a 200 nm thick $1 \cdot 10^{17}$ cm^{-3} n-doped GaAs channel, a 10 nm thick etch-stop layer of $Al_{0.3}Ga_{0.7}As$ and a GaAs n^+-contact layer.

The p^-n-junction between the buffer/absorber layer and the channel of the MESFET is used as a PD.

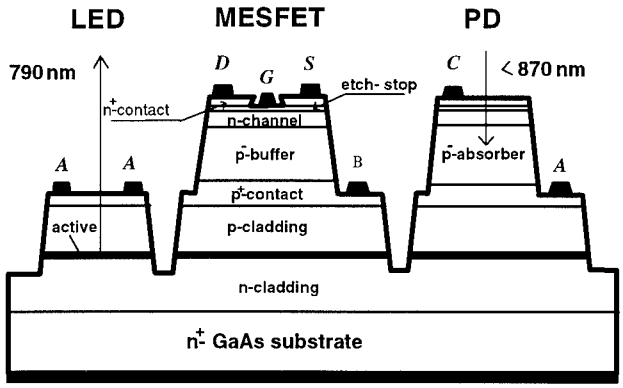


Fig. 1. Schematic layer sequence of an optoelectronic smart pixel.

The wavelength of the light emitted by the LED is set to 790 nm by the aluminium content and the thickness of the quantum well. Therefore the emitted light can efficiently be detected by the PD. The etch-stop layer ensures a homogenous MESFET threshold voltage over the wafer. The entire n-p-n layer structure is described in [4].

The fabrication process requires eight photolithographic steps and is based on mesa isolation. The backside ohmic contact is formed first by a Ge-AuNiAu metallisation. This is the common n-contact to all LEDs on the wafer. For the n-side ohmic contacts of the MESFETs and the PDs, a NiGe-AuNiAuPt [5] metallisation is evaporated and annealed at 440 C°. Then the individual MESFETs and PDs are separated by magnetron enhanced reactive ion etching (MIE) of mesas. After the deposition of a Si_3N_4 dielectric film, windows to the ohmic contacts of the MESFETs and PDs (n-terminals) as well of the LEDs and the PDs (p-terminals) are opened by reactive ion etching (RIE). The gate areas are then opened by RIE. The dielectric film is used as a mask for a selective, recessed wet-etch of the gate area [6]. This etch removes the n^+-contact layer and undercuts the dielectric film by approximately 200 nm. The next metallisation (TiPtAu) forms the Schottky gate contacts for the MESFETs. This metallisation is also used as the p-contact to the LEDs and the PDs and as a first wiring level. The LEDs are isolated by a second level of MIE mesa etching. The main wiring level (TiAl) is deposited onto the second dielectric film and patterned by wet-etching.

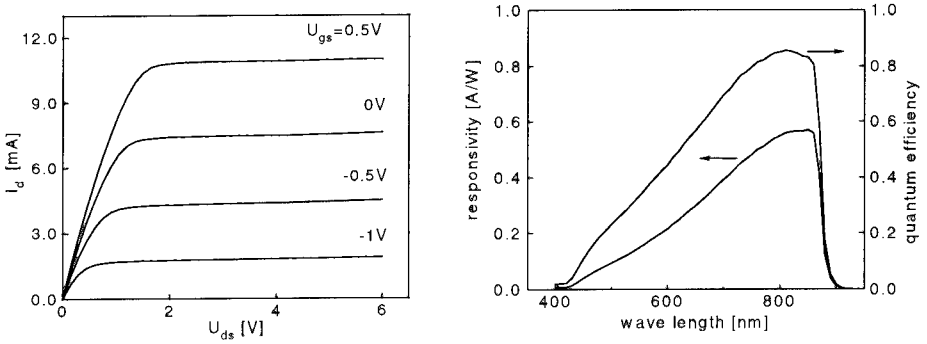


Fig. 2. Left: the drain I-V characteristic of a MESFET with a gate length of 1.5 μm and an effective width of 80 μm. Right: the quantum efficiency and responsivity vs. wavelength for a PD with an active area of 50·50 μm^2.

The fabricated individual devices are then characterised (Fig. 2). The MESFET shows a maximum transconductance of 80 mS/mm and a current density I_{ds} of 226 mA/mm for a gate length of 1.5 μm. The quantum efficiency of the PD is 0.85 at the LED peak emission wavelength of 790 nm. An efficiency of 0.008 W/A is measured for a 25·25 μm^2 LED.

In this layer design, the devices are connected to parasitic elements in three different ways. 1) The MESFET is connected to the photodiode below it which is in turn connected to an LED. A potential between the MESFET source and the p-contact of the photodiode below causes a backside channel depletion which results in a threshold voltage shift. 2) The p-contact of the photodiode is connected to an LED below, introducing an additional capacitance. 3) All LEDs share a common n-side contact. These facts must be considered in the circuit design.

The fabricated optoelectronic smart pixels are threshold circuits. The example in Fig. 3 consists of a dual-photodiode differential input connected to an inverter (M1, M2), which switches the LED driver (M3, M4, LED). The LED is turned on when the input power of the switching beam (P_{Switch}) exceeds that of the reference beam (P_{Ref}). The switching threshold can thus be controlled by the reference beam. The LED driver is current-balanced: if the LED is off, the current flows through MESFET M3, and the LED is turned on by pinching off M3.

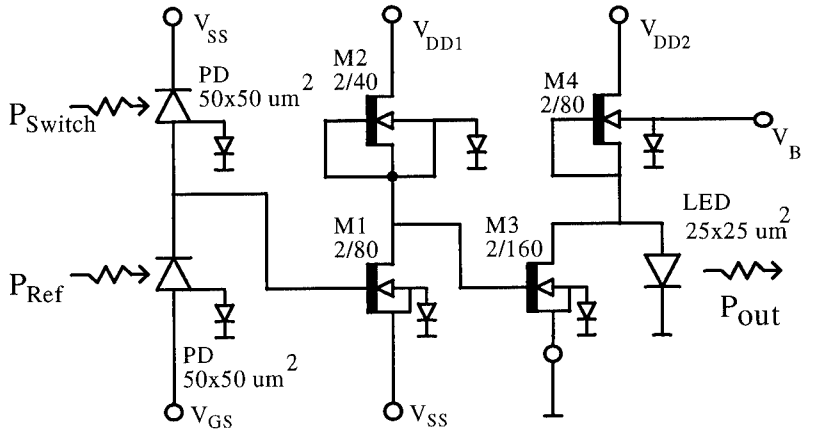


Fig. 3. Diagram of the measured threshold circuit.

The dc-behaviour of this threshold circuit is shown in Fig. 4. The output power is 18 μW in the ON state, and a contrast ratio greater than 1000 has been estimated. The minimum switching power is measured to be less than 1 nW. At an input power level of 200 nW, the circuit has a switching delay of 8.4 μsec, which corresponds to a switching energy of 1.7 pJ. The circuit occupies 200·300 μm^2 with a photodetector area of twice 50·50 μm^2. The power dissipation remains constant at 20 mW.

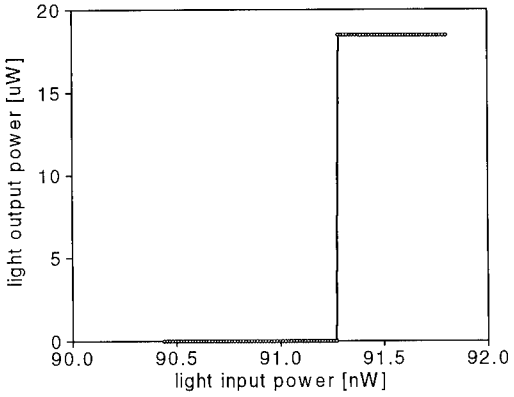


Fig. 4. The measured optical output power vs. optical input power of the threshold circuit.

At the present switching energy level of approximately 2 pJ, a maximum switching speed of 4 MHz is achievable if the entire 18 μW output power of our existing smart pixel is available at the detector. The 12·12 μm^2 LED shows no thermal rollover in the output power up to a drive current of 5 mA. A PD which is optically coupled to such an LED can have the same size. The switching energy, which scales with the PD area, would then be reduced below 100 fJ, implying an expected switching speed of 80 MHz for perfectly-coupling. In a realistic application however, the coupling is degraded by the broad emission cone of the LED and the limited aperture of the coupling optics. A Lambertian emission profile and a NA of 0.2 implies a coupling efficiency of 1% and hence a maximum switching speed of 0.8 MHz. Higher speeds require optics with a higher NA or a reduction in the divergence of the light source, i.e. by the integration of a VCSEL.

In conclusion, we have developed a layer structure grown in a single step and a fabrication process for the monolithic integration of optoelectronic smart pixels. A threshold circuits with a low switching power and switching energy have been demonstrated. The performance of such circuits is suitable for application in the fields of optoelectronic neural networks and parallel image processing. The switching speed expected in a system application is limited mainly by the efficiency and divergence of the integrated light source.

We respectfully acknowledge the design assistance of K. Engelhardt and the construction work of M. Kuhn for the optical measurement set and enlightening discussions with J.E. Epler, K.H. Gulden and M. Moser. We are also grateful for the support of Professors W. Bächtold, W. Kündig and J.Mlynek.

References

[1] Julia J Brown, J T Gardner and Stephen R Forrest, Fellow IEEE 1993 IEEE J. Quantum Electron. vol. 29 nr. 2 pp. 715-726

[2] Anthony L Lentine, Member IEEE, L M F Chirovsky and T K Woodward 1994 IEEE J. Quantum Electron. vol. 30 nr. 5 pp. 1167-1171

[3] D Leipold, U Kehrli, K Thelen, M Rossi, J E Epler, H P Schweizer, P Seitz and B D Patterson 1994 LEOS 1994 IEEE Catalog Number 94TH0606-4

[4] D Leipold, U Kehrli, B Delley, K Thelen and B D Patterson 1993 Materials Science and Engineering B21 pp. 300-303

[5] Yih-Cheng Shih, Masanori Murakami, E. L. Wilkie and A C Callegari 1987 J. Appl. Phys. 62 (2) pp. 582-590

[6] Gregory C DeSalvo, Wen F Tseng and James Comas 1992 J. Electrochem. Soc. vol. 139 no. 3 pp. 831-835

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Liquid Crystal Over Silicon Spatial Light Modulators

Douglas J. McKnight†, Mark A. Follett‡ and Kristina M. Johnson†

† Optoelectronic Computing Systems Center, University of Colorado at Boulder, Campus box 525, Boulder, Colorado 80309-0525, (303) 492 0958

‡ Boulder Nonlinear Systems, 1898 South Flatiron Court, Boulder, Co.

Abstract.
This paper describes our recent liquid crystal on silicon spatial light modulators. We present the most recent results from our 256 by 256 binary reflection mode SLM operating in an optical correlator and a 128 by 128 analog SLM that uses the Distorted Helix Ferroelectric effect. The provisional performance specifications of the analog SLM are a frame rate of $\approx$ 2.5kHz, a zero order contrast of 30:1, and $>$ 16 grey levels.

1. Introduction

Liquid crystal on silicon spatial light modulators are made by placing a thin layer of liquid crystal directly on top of a silicon chip. For a recent review of this technology see, for example, reference [1]. We have constructed a 256 by 256 binary SLM for an optical correlator application which uses a chiral smectic C (SmC*) ferroelectric liquid crystal as the light modulating layer [2]. We have also constructed a 128 $\times$ 128 analog SLM which uses the Distorted Helix Ferroelectric (DHF) [3] effect and in this paper we present some early results from this device.

2. The 256 by 256 binary SLM

The SLM is based on a 7mm by 7mm CMOS die fabricated by US2 under a 1.2 micron design rule. The active array area is a square approximately 5.53mm on a side. Each pixel in the array is addressed by a row select wire and a column data wire. The row wire activates the pixel transistor gate and the column wire presents a binary data signal to the pixel. On activation of the gate wire the data on the column wire is written to the pixel, where it is capacitively stored. The resulting electric field between the pixel mirror and a transparent electrode on a piece of cover glass drives the liquid crystal into the desired state.

Array size	256 by 256
Pixel Pitch	21.6μm
Fill-factor	79%
Flat fill-factor	60%
Diffraction efficiency	15%
Throughput into zero order	3.4%
Demonstrated frame load time	43μs
Simulated frame load time	27μs
Contrast ratio	70:1 zero order (10:1 imaged)
LC switching time	75μs

Table 1. Specifications of and results from the 256 by 256 binary SLM.

Data are transferred to the SLM over 32 parallel lines under the control of a master clock and a frame sync signal. We have demonstrated the addressing of the SLM with a master clock frequency of 48MHz which gives an image refresh rate of 23.5kHz. This is data rate of 1.6Gb/s from the driver board to the SLM. To achieve these data rates the SLM backplane utilised on-chip clock and control signal generation and data pipelining.

2.1. Results.

A cover-glass was fixed over the pixel array, spaced with polyimide balls, and the gap filled with BDH SCE13 liquid crystal by capilliary action under vacuum. The liquid crystal was aligned by means of a rubbed PVA layer on the cover glass. The device specifications are summarized in table 1.

Two of these SLMs have been used in an optical correlator as input and Fourier plane devices operating in a binary phase mode. The system was operated at 1000 correlations per second, i.e. a positive input and filter pattern were written for 500μs, followed by their inverses for 500μs. The camera shutter was set at 1ms to capture the output from both true and inverse frames. On and off-axis inputs consisting of small targets that used 0.6were used. With the laser producing approximately 5mW the correlation peaks in the output were able to saturate the camera and yield a signal to noise (pk/rms) of greater than 10dB for most of the images.

3. The 128 by 128 analog SLM

The silicon backplane for this device was fabricated on a 2 micron CMOS process through the MOSIS brokerage. Sixteen analog input lines are sampled and their voltage values routed to the pixels under the control of an on-chip controller. The analog data are generated by an external driver board which stores the data as 8-bit numbers. Fast current-output DACs convert these data into signals which are sent to a header board on which the SLM is mounted. The analog currents are converted to voltage signals by op-amps on the header board. The factor which most limits the frame-load time of the SLM is the settling time of the op-amps. By allowing 100ns for their outputs to settle we have a frame-load time of $\approx 100\mu$s.

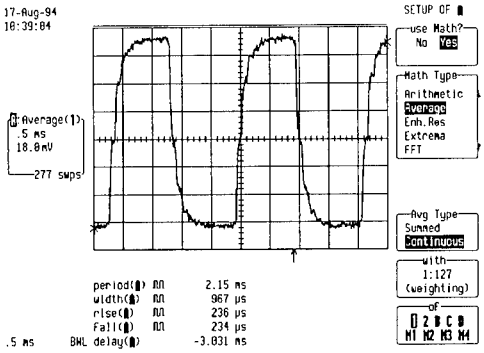


Figure 1. The output from a group of pixels on the analog SLM switching from fully–off to fully–on. The rise and fall times are $\approx 235\mu s$.

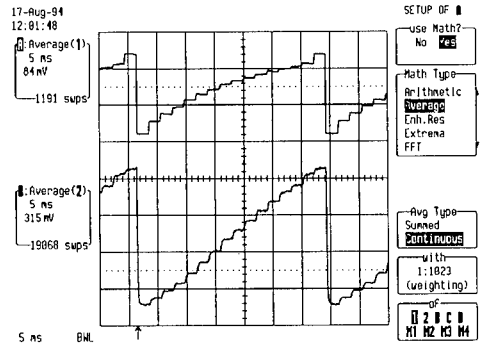


Figure 2. The output from a group of pixels on the analog SLM switching through a succession of 16 gray scales. The lower trace is the output intensity and the upper trace is proportional to the analog input voltage that was applied. The response is not quite linear so a correction was estimated and applied to the input signal to generate the approximately linear output scale shown here.

The SLM coverglass was treated with a rubbed PVA alignment layer and the cell was filled with the DHF material Hoffmann la Roche 9807. This alignment scheme is not ideal for this DHF material and so we find that there is a lot of scattering from the liquid crystal in the "off" state. We have performed experiments on test cells using different alignment materials and cooling schemes and we are confident that these techniques will yield improvements in the alignment quality in future SLMs.

3.1. Results from the analog 128 by 128 SLM.

The scattering from the alignment defects is apparent in the measurement of the contrast ratio of this SLM. The device was illuminated by a He-Ne laser (633nm) and the contrast

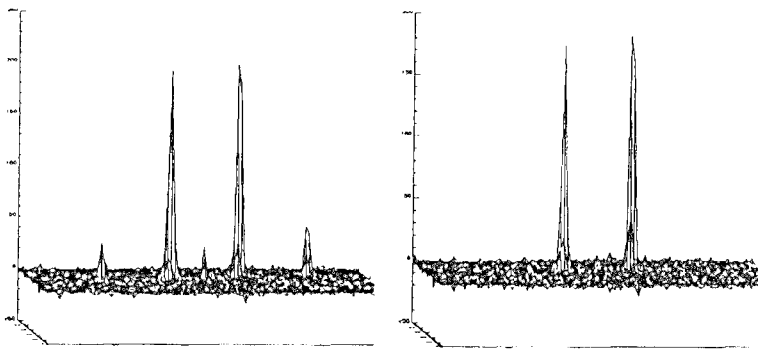


Figure 3. Contour plots of the optical Fourier transform of the analog SLM when it was displaying a binary stripe pattern (left), and a sinusoidal stripe pattern (right). The SLM was operated in phase mode.

from a group of pixels comprising about half the SLM array was measured. If only the zeroth diffracted order is collected then the contrast ratio is 30:1, however the scattering from the imperfect alignment leads to a rapid fall-off in contrast as the acceptance angle of the output optics is increased. If the optics collects the zeroth order and the $\pm$ first orders the contrast drops to 14:1 and continues to drop to $\approx$6:1 at the $\pm$5th and 6th orders. Switching speed results from the analog SLM are shown in figure 3. This measurement was made by imaging a region of the SLM array onto a photomultiplier and writing alternately 10 fully-on frames and 10 fully-off frames to the SLM. The results of a similar experiment are shown in figure 3. In this experiment a set of 16 frames are used to address the SLM and generate 16 gray levels. Figure 3.1 show the result of operating the SLM in phase mode. A binary phase stripe pattern yields the image on the left and the analog sinusoidal pattern of the same spatial frequency yields the image on the right. Using the SLM in this way is equivalent to having one bit of pure phase information and the rest of modulation in amplitude.

References

[1] K. M. Johnson et al. Smart spatial light modulators using liquid crystals on silicon. (invited paper). *Journal of Quantum Electronics*, 29(2):699, 1993.

[2] Douglas J. McKnight, Kristina M. Johnson, and Roylnnn A. Serati. 256 x 256 liquid-crystal-on-silicon spatial light modulator. *Applied Optics*, 33(14):2775–2784, 1994.

[3] J. Funfschilling and M. Schadt. Fast responding and highly multiplexible distorted helix ferroelectric liquid crystal displays. *Journal of Applied Physics*, 66(8):3877, 1989.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

High-speed liquid crystal on silicon spatial light modulators

C. C. Mao, D. McKnight, and K. M. Johnson
The Center for Optoelectronic Computing Systems
University of Colorado at Boulder
Boulder, Colorado 80309-0525

The results of a 64 x 64 high-voltage CMOS array which produced 30 V at each pixel resulting in a liquid crystal switching time of 30 μs are described.

I. Introduction.

Liquid-crystal-on-silicon (LCOS) spatial light modulators (SLMs) have applications to displays, optical beamsteering, and optical information processing. In these devices, liquid crystals (LCs) are sandwiched between a silicon VLSI backplane and a piece of glass coated with a transparent conductive electrode. Voltages are applied to metal mirrors on the silicon backplane to electrically switch a ferroelectric liquid crystal (FLC). This results in orienting the FLC optic-axis to two states. A polarized optical beam incident upon the device from the LC side is reflected from the metal mirrors, resulting in a spatial modulation of the input light's polarization. A polarizer placed at the output of the SLM can convert the polarization modulation to intensity or phase modulation.

The FLC switching speed is inversely proportional to the applied electric field E. With a standard 2 μm CMOS process, the VLSI circuitry provides voltages between 0 V and 5 V at each metal mirror for switching the liquid crystal. In general, such low voltages result in FLC switching times ranging between 0.1 ms and 0.5 ms.[1–2] Recently, McKnight et al fabricated a LCOS spatial light modulator with a 1 μm FLC layer, resulting in a 2.5 $V/\mu m$ applied E field and a 10% - 90% rise time of 75 μs.[3] If higher voltages are applied to the metal mirrors, faster FLC response times should be possible. In this paper, we describe a high-voltage (HV), 64 x 64 LCOS SLM that provides 30 V to each metal mirror resulting in a LC switching time of 30 μs.

II. High-voltage CMOS array design

The maximum voltage a transistor can switch is limited by the junction breakdown voltage. Transistor junction breakdown voltage is inversely proportional to the impurity concentration in the diffusion regions. This implies that using a lower doping concentration for the transistor drains should extend their voltage handling capability. This approach is applied to the structure shown in Fig. 1. The n-fet uses the n-well ($10^{16}/cm^3$) as its drain, which has a lower doping than n-diffusion ($10^{20}/cm^3$), The p-fet uses p-base ($10^{17}/cm^3$) as its drain, which has a lower doping than p-diffusion ($10^{19}/cm^3$). Neither of these layers is masked by polysilicon during implantation, and hence are not self-aligned. Care must be taken to ensure overlap with a thin oxide gate region to provide channel continuity. In our design, this overlap is 3 μm. The contact to these drains is surrounded by the lightly doped drain material to isolate the surrounding substrate or well.

Our previous test results showed that the HV n-fet can accommodate voltages up to 30 V, while the HV p-fet can switch voltages up to 15 V.[4] In general, a voltage of ± 10 $V/\mu m$ is required for switching FLCs with the highest speed and the largest optic-axis switching

angle. Therefore, to operate the LCOS device with a LC layer thickness of 1 μm or larger under these optimum conditions, a voltage higher than 10 V is required. To obtain bipolar voltages with peak amplitudes higher than 10 V, the circuit must operate with voltages higher than 20 V. This voltage exceeds the p-fet breakdown voltage. There are two approaches to overcome this limitation: 1) use a lighter doping p-fet drain region, 2) drop the voltage across several p-fets in series. We chose the latter approach, because we cannot specify the doping concentration of integrated circuits processed through the University of Southern California MOS Implementation Service (MOSIS).[5]

The circuit which can accommodate voltages up to 30 V is shown in Fig. 2. The HV addressing circuit includes four HV p-fets and two HV n-fets. The bias voltage, V_b, is approximately half of the HV V_{hi}. The select signals, V_1 and V_2, are 0 V and 5 V signals generated by a standard CMOS control circuit. When V_1 = 5 V and V_2 = 0 V, The output voltage, V_{out}, is charged to V_{hi}. The voltage at the node between p-fets 4 and 5 is equal to V_{hi} and the voltage on the gate of p-fet 1 is equal to V_{hi}; thus closing p-fet 1. The node between p-fet 2 and n-fet 3 is discharged to 0 V. In this case, the voltage at the node between p-fets 1 and 2 is approximately $V_{hi}/2$. The voltage across p-fets 3 and 4 is 0 V. If V_1 = 0 V and V_2 = 5 V, the resulting voltages are reversed. Therefore, this insures that the drain-to-source voltages across any p-fet do not exceed $V_{hi}/2$, and this circuit can be operated with voltages up to 30 V. Another advantage of this approach is that the control circuits can be fabricated with conventional CMOS transistors.

We have designed and fabricated a high-voltage VLSI array with 64 x 64 pixels using this circuit to convert low-voltages to high voltages (see Fig. 3). The array was fabricated in the 2-μm n-well process through the MOSIS foundry. Pixels are located on 48 μm centers and each pixel contains one HV n-fet and a metal mirror. The n-fet consists of n-well (drain), n-diffusion (source), and polysilicon (gate). The modulating mirror is a metal2 layer, which does not overlap the transistor. The pixel flat fill factor is 60%. The array is addressed line at a time using row shift registers located to the left and right sides of the array, comprised of the conventional CMOS transistors and the HV circuits for converting low voltages to high voltages. Data are input onto the columns using 16 parallel data lines and multiplexers positioned at the top and bottom of the chip.

III. Experimental results

A HV LCOS SLM was assembled and tested. The liquid crystal used is the British Drug House smectic C^* FLC SCE-13. The functionality of the device was tested by writing binary images onto the array and observing the output optical images under a polarizing microscope. A voltage equal to $V_{hi}/2$ is applied to the common electrode on the cover glass. A voltage of 0 V on the metal mirror results in the local optic-axis of the FLC to be switched to what we define as the OFF state. When a metal mirror is charged to voltage V_{hi}, the local optic-axis of the FLC is switched to what we define as the ON state. When a polarized uniform light beam illuminates the device, the reflected light from metal mirrors with V_{hi} ideally undergoes 90^0 polarization rotation, and is transmitted by an output analyzer. The reflected light from metal mirrors with 0 V is not polarization rotated, and hence extinguished by an output analyzer. Fig. 4 shows an example of output optical images from the SLM.

We also tested the response times and ON/OFF contrast ratio of the SLM. Fig. 5 gives an oscilloscope trace showing the switching times of the FLC with an applied voltage, V_{hi} =

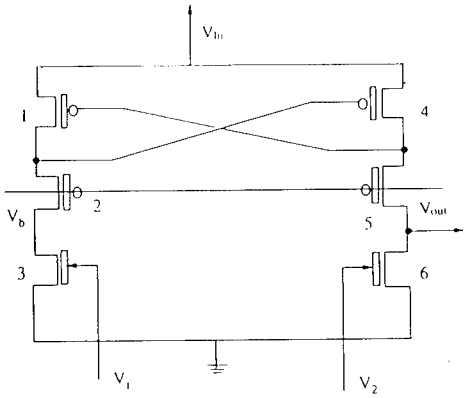


Figure 2. High-voltage addressing circuit.

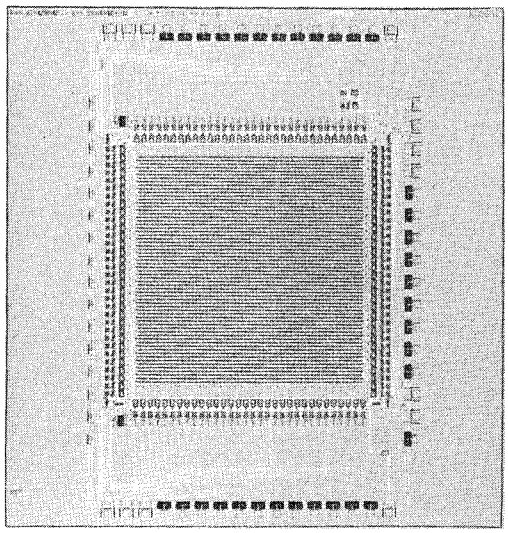


Figure 3. Photograph of the fabricated high-voltage chip.

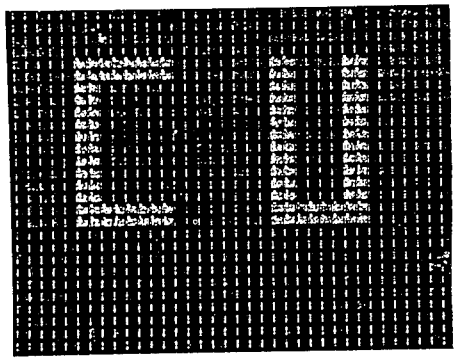


Figure 4. Photograph showing an output optical patern from high-voltage LCOS spatial light modulator.

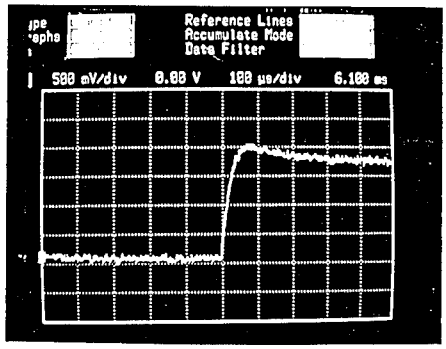


Figure 5. Oscilloscope trace showing a FLC switching time of 30 us.

25 V. The 10% - 90% FLC switching time is approximately 30 us. The measured ON/OFF contrast ratio (I_{ON}/I_{OFF}) is 13:1.

IV. Conclusions

We have designed, fabricated, and tested a high-voltage, 64 x 64 array liquid crystal on silicon spatial light modulator. The SLM can switch 30 V at the pixel mirrors, and has a 10% - 90% liquid crystal switching time of 30 μs. The contrast ratio of it is 13:1.

The authors would like to acknowledge J. Lazzaro and R. Turner for technical discussions. We are grateful for funding of this work by C. S. Draper Laboratories and the NSF/ERC through OCS grant 9015128.

References

1. L. K. Cotter, T. J. Dillon, and M. A. Handschy, Opt. Lett. **15**, 291 (1990).

2. K. M. Johnson, I. underwood, D. J. McKnight, IEEE J. Quant. Electron. **29**, 699 (1993).

3. D. McKnight, K. M. Johnson, and R. A. Serati, Appl. Opt. **33**, 2775 (1994).

4. C. C. Mao, K. M. Johnson, and Z. Y. Huang, OSA Annual Meeting Technical Digest, pp. 165, Oct. 3-8, 1993, Toronto, Canada.

5. MOSIS Project, USC/Information Sciences Institute, 4676 Admiralty Way, Marina del Rey, CA 90292-6695.

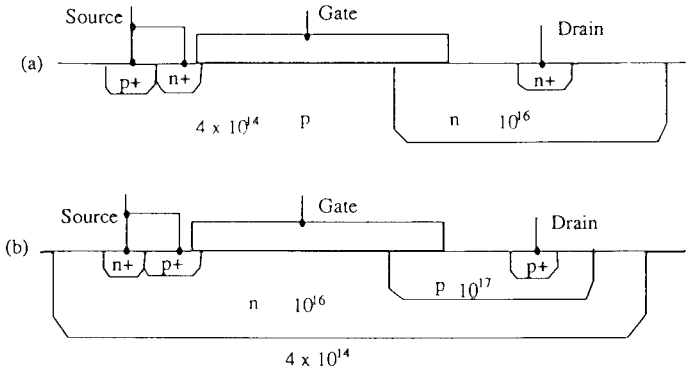


Figure 1. Structures of high-voltage fets: (a) n-fet and (b) p-fet.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Electronically Addressed Ferroelecctric Liquid Crystal over Silicon Spatial Light Modulators

D.C. Burns, I. Underwood, A. O'Hara† and D.G. Vass†
Department of Electrical Engineering (†Dept of Physics)
The University of Edinburgh, Edinburgh EH9 3JL, UK.
Tel +44 31 650 5652, fax +44 31 650 6554, email slm@uk.ac.ed.ee.

Abstract. Recent advances in electronically addressed ferroelectric liquid crystal over silicon spatial light modulators include the design of devices of increased speed, larger pixel count and greater optical efficiency. We present recent results illustrating each of these features.

1. Background

The Spatial Light Modulator (SLM) is a key component for many optical computing systems. Historically, weak SLM performance has been a limiting factor for some systems. The SLM technology of Ferroelectric Liquid Crystal over Very Large Scale Integrated (FLC/VLSI) silicon has matured considerably in recent years. Electronically Addressed SLM's (EASLM's) of medium resolution (128^2 pixels or more) have been reported by several labs [1]; a working system containing multiple devices have been reported [2].

2. Introduction

FLC/VLSI SLMs are produced by sandwiching a thin layer ($\sim 1\mu m$) of FLC between a custom designed silicon backplane and a cover glass coated on the inside with a transparent conductive electrode [1]. Most EASLM backplane designs are based on a pixel circuit consisting of one active element - a Metal- Oxide-Semiconductor Field Effect Transistor (MOSFET) which acts as a switch to control the amount of charge stored on a capacitive storage element - a small metal mirror on the surface of the silicon. The voltage thus generated produces an electric field which alters the state of a the trapped layer of FLC to produce a binary phase or amplitude modulation in an incident wavefront.This scheme provides the smallest possible pixel and thus the highest density of pixels; it is analogous to the purely electronic Dynamic Random Access Memory (DRAM) cell.
This type of pixel suffers from light induced charge leakage which is manifest as a reduction in contrast ratio with increasing intensity of incident light, thus limiting the maximium light level at which it can be operated. There is also a limit to the spontaneous polarization, P_s, of the FLC material which can be used - the pixel capacitor must store enough charge to switch it. The drive to reduce the pixel size is compromised by the need to maintain an optically-flat metal area to act as the reflective aperture (or mirror). Underlying circuit elements such as transitors or interconnect cause undulations in the overlying part of the mirror; these can cause nonuniform optical contrast across a mirror, losses due to scattering and, in coherent

systems, phase variations. A flat fill factor [5], F_3 (flat mirror area / pixel area), of around 25% has become a de facto minimum. More quantitatively if we consider the pixel area, A_{pix}, to be made up of two components - the flat mirror area, A_{mirror}, and the rest of the pixel (bus lines, transistors etc.), A_{circ}, we can write

$$A_{pix} = \frac{A_{circ}}{1 - F_3} \tag{1}$$

Within a given CMOS process (which fixes the minimum size of the circuit elements and the maximum die size) a given pixel functionality (which fixes the number of circuit elements required) there is a direct trade off between flat fill factor and pixel area (or maximum number of pixels).

3. 512×512 DRAM SLM

We have previously reported a 176×176 DRAM-style pixel array [3].The frame rate was limited to around $1kHz$ (dc balanced) by the RC time constant of the relatively high resistance polysilicon row access lines. A 512×512 pixel array (based upon the original 176×176 device) has been designed. The primary modification has been the use of aluminium to replace the polysilicon row access lines within the pixel array. The finished design has been fabricated by Austria Mikro System and is undergoing electrical testing. A design flaw has been detected in the on-chip address circuits; the drive sequence has been altered in software in order to provide a temporary workaround to this problem.

4. 256×256 SRAM SLM

An alternative to the single transistor pixel design above is based around an enhancement of the six transistor Static RAM (SRAM) cell. The enhancement involves inserting a simple logic gate between the memory node and the pixel mirror/electrode node so as toallow the addressing requirements of the FLC to be met more easily. The SRAM design overcomes all of the above disadvantages of the DRAM pixel at the expense of increased transistor count leading to increased pixel area. In particular it maintains its state indefinitely, allows the use of the fast-switching high P_s FLC materials and shows no variation of contrast ratio with incident light intensity over a wide range of intensity. We have demonstrated a fully working 256×256 pixel array built in $1.2\mu m$ CMOS technology. The contrast ratio, measured directly by imaging pixels is around $9 - 10 : 1$. The FLC switching speed has been measured at $150\mu s$ (rise) and $55\mu s$ (fall) with a supply voltage to the backplane of $V_{dd} = 6V$, see Figure 1. This asymmetry indicates a likelihood that the FLC in these particular devices is not bistable but tending towards a monostable configuration. Further investigations are underway.

5. Increasing Pixel Flat Fill Factor

We have succesfully applied a post-processing wafer planarisation technique to the 176×176 DRAM and 256×256 SRAM devices. This has allowed an optically flat mirror to be placed on top of the existing circuitry. The technique involves the deposition of a thick dielectric layer which is subsequently polished flat; this faciliti-

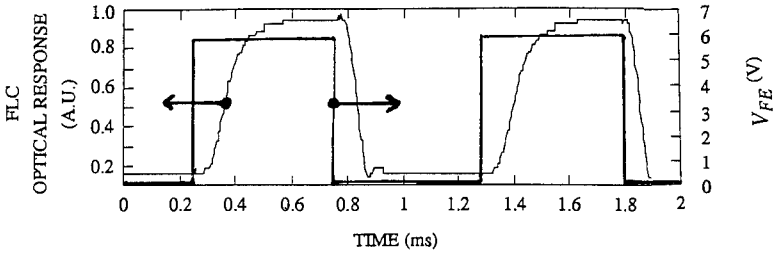


Figure 1: FLC optical response

ates the deposition of a further metal layer which is patterned to form a flat mirror covering almost the entire pixel. The planarisation technique is described in detail elsewhere [4]. FLC cell construction has been carried out succesfully on planarised backplanes. The increased flat fill factor of the backplane increases the light throughput of the finished SLM and reduces the stray light reaching the substrate. For a SLM designed with planarisation in mind, it is possible to achieve

$$A_{pix} \approx A_{circ} \tag{2}$$

coupled with a high flat fill factor

$$F_3 = \left[\frac{p-m}{p}\right]^2 \tag{3}$$

where p is the pixel pitch and m is the minimum allowed gap between adjacent mirrors. Figure 2 shows the effect of planarisation on the same image on a small part of a 256×256 SRAM SLM. The via contact leaves a small dark patch even on the planarised pixels.

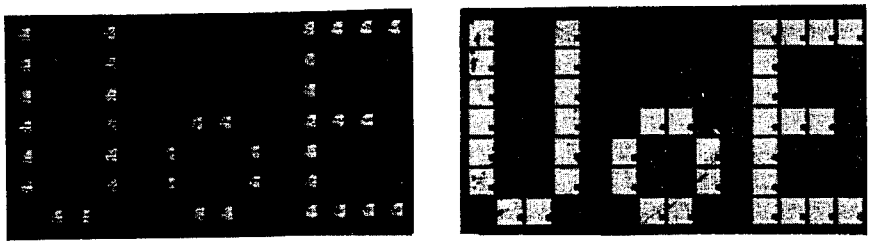


Figure 2: Part of SLM displaying pattern, unplanarised (left) and planarised

6. Summary

Table 1 summarises the present situation. High resolution devices, based on both DRAM and SRAM pixels, have been built. The planarisation process, which significantly enhances device performance, has been demonstrated on two SLM backplane designs. It is, in principle, equally applicable to all of the devices of Table 1 and to any future devices, including smart pixel SLM's. Research has commenced on a method of via filling to remove the effect present in Figure 2.

Device	176 DRAM	512 DRAM	16 SRAM	50 SRAM	256 SRAM
Date	1989	1994	1986	1988	1994
Fab process	$3\mu m$ CMOS	$3\mu m$ CMOS	$6\mu m$ nMOS	$1.5\mu m$ nMOS	$1.2\mu m$ CMOS
Pixel pitch (μm)	30	30	200	72	40
Mirror size					
Nominal (μm^2)	15×11	15×11	110×110	40×40	19×19
after planar'n	27×27	$27 \times 27^{\dagger}$	196×196	NA	37×37
F_3 (%) before / after planar'n	18 / 81	18 / 81	30 / 96	31 / NA	23 / 85
Elec address time (μs)	250	$1000^{\dagger}$	1	N/A	85
Frame rate (Hz)	1000	$250^{\dagger}$	5	20	4000

Table 1: Silicon backplanes for FLC/VLSI EASLM's († design estimate)

Given the current (minimum feature size and maximum die size) limitations of using commercial silicon vendors it is possible to pursue the DRAM approach to 1024 x 1024 and the SRAM to 512 x 512 with frame rates comparable to, or better than, those of Table 1. Access to more specialised current fabrication processes, or the ongoing process development of the commercial silicon vendors, will allow the design and fabrication of even higher performance SLM's. Even then, the highest performance is only likely to be achieved through custom post-processing to produce an optically flat and FLC-friendly surface layer.

References

[1] K.M. Johnson, D.J. McKnight, I. Underwood, "Smart spatial light modulators using liquid crystals on silicon", IEEE Journal of Quantum Electronics, Vol 29, 699-714, 1993.

[2] R. Turner, D.A. Jared, G.M. Sharp and K.M. Johnson, "Optical correlator using very-large-scale integrated circuit / ferroelectric liquid crystal electrically addressed spatial light modulators", Appl. Opt. Vol 32, 3094-3101, 1993.

[3] I. Underwood, D.G. Vass, R.M. Sillitto, G. Bradford, N.E. Fancey, A.O. Al Chalabi, M.J. Birch, W.A. Crossland, A.P. Sparks, S.G. Latham, "A high performance spatial light modulator", Proc. S.P.I.E., Vol 1562, 107-115, 1991.

[4] A. O'Hara, J.R. Hannah, I. Underwood, D.G. Vass and R.J. Holwill, "Mirror quality and efficiency improvements of reflective mode spatial light modulators by the use of dielectric coatings and chemical-mechanical polishing", Appl. Opt., Vol 32, 5549-5556, 1993.

[5] D.J. MacKnight, K.M. Johnson and R.A. Serati, "256/*times*256 liquid-crystal-on-silicon spatial light modulator", Appl. Opt., Vol 33, 2775-84, 1994.

Inst. Phys. Conf. Ser. No 139: Part V
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Sensitivity transition of LAPS-SLM and its application for optical neural network

Naoki Kato, Nobuyuki Kasama, Rieko Sekura, Katsuki Matsusita, Yasuyuki Mitsuoka and Tadao Iwaki

Corporate R & D department, Seiko Instruments Inc.,
563 takatsukashinden, Matsudo, Chiba 271, Japan

Phone: +81-473-92-7825, Fax: +81-473-92-2026

ABSTRACT

We present the sensitivity transition characteristics of LAPS-SLM. Furthermore, optical neural network that is using the LAPS-SLM as a thresholding device will be presented. The sensitivity transition coefficient is controllable by the driving waveform or the light intensities.
In the concretely system, the LAPS-SLM will be used as the weighting and thresholding device. A new learning scheme by the sensitivity transition of LAPS-SLM will be presented and discussed.

1. Introduction

Interest in optical processors, optical correlators, and optical neural network systems has grown due to their inherent high speed and parallelism. Spatial light modulators (SLMs) are basic components in these systems.(1) We have been developing a spatial light modulator using a hydrogenated amorphous silicon photoconductor and a surface stabilized ferroelectric liquid crystal light modulator (LAPS-SLM).(2),(3)

Owing to the driving pulse voltage threshold, FLC-SLMs have a definite write light intensity threshold. As is discussed concerning about the matrix displays, the accumulation effect is immanent in the devices using SSFLC mode. The accumulation effect is actualized in the SLM as the sensitivity transition. This time, we considered about the sensitivity transition characteristics of LAPS-SLM, and would propose the application of these nonlinear characteristics for the optical neural network system.

2. Device fabrication and experimental conditions

The structure of the LAPS-SLM is shown in Fig. 1. The alignment layers are provided by an oblique evaporation of SiO. The input and output windows are clamped together and the FLC material is introduced by the capillary method. The writing experiments are made on a microstage of the polarization microscope. Write light is 550nm with half-width of 40nm, and readout light is 630nm with half-width of 80nm. The polarizer and the analyzer are set in the crossed Nicols state. In these experiments, driving waveform is bipolar pulse that is shown in Fig. 2.

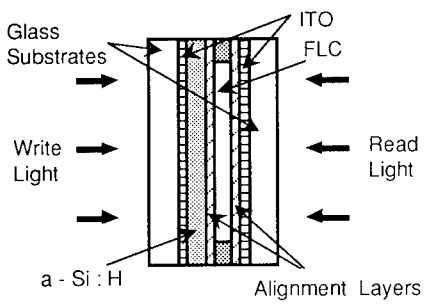


Fig. 1 Schematic illustration of LAPS-SLM

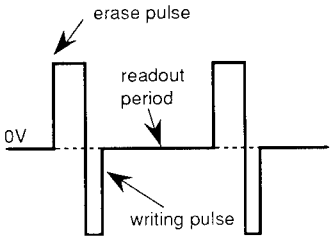


Fig. 2 Standard driving waveform
Erase pulse : 10V, 2ms
Writing pulse : -10V, 1ms
Readout period : 7ms

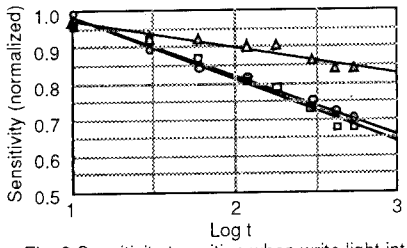


Fig. 3 Sensitivity transition when write light intensity is modified
○:200μ W/cm2 , □:400μ W/cm2 , △:600μ W/cm2

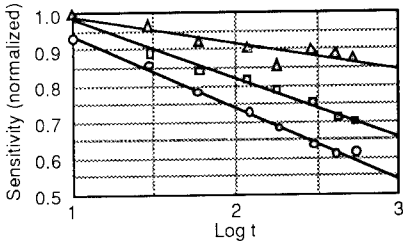


Fig. 4 Sensitivity transition when erase pulsevoltage is modified
○: 12.5V, □: 10V, △: 7.5V

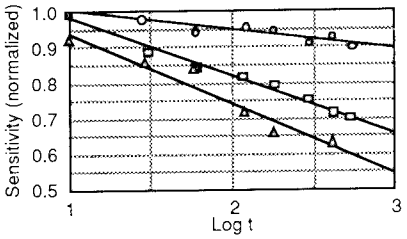


Fig. 5 Sensitivity transition when erase pulse width is modified
○: 0.75ms, □: 1ms, △: 1.5ms

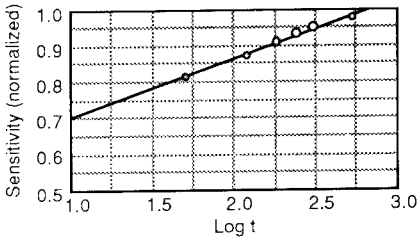


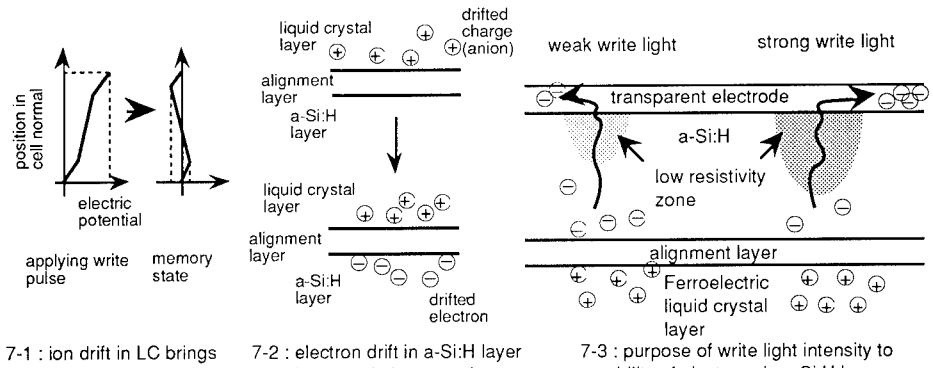
Fig. 6 Mitigation at Iw= 0 μW/cm2

3. Results and discussions

Write light intensity (I), erase pulse voltage (p) and erase pulse width (q) are modified, and the transition of the threshold write light intensities (ThI) are measured. Sensitivity is defined as ThI / THI_0, where THI_0 is the threshold write light intensity at the initial. Figs. 3, 4 and 5 show the sensitivity transition of LAPS-SLM. Under any condition, the threshold write light intensity is decreases monotonously. Fig. 3 shows the dependence of the sensitivity transition on the write light intensity (I). Driving waveform used is standard that is shown in Fig. 2 and the write light intensity is modified as 200, 400 and 600μW/cm2. Readout light is irradiated continuously, and the measurement is made individually. Fig.4 shows the dependence of the sensitivity transition on the erase pulse voltage (p). Fig. 5 shows the dependence of the sensitivity transition on the erase pulse width (q). These data show that the sensitivity transition characteristics are proportional to log$t$, and the coefficients could be controllable by I, p and q. Fig. 6 shows the mitigation characteristic after 10 minutes continuous writing. These mitigation characteristics are also proportional to log$t$. Here, I would be able to write the accumulation coefficient a=-αlog$t$ (α=ϕ(I, p, q)), and the mitigation coefficient b =βlog$t$ (β=constant).

These accumulation and mitigation characteristics are phenomenally supposed to similar to the memory and the forget in the brain[4][5], so I would consider about the application of these characteristics for the optical neural network. To control the accumulation characteristics, it seems that p or q would be adequate. However, the effect of I is slight different because as shown in Fig. 3, dependence of the accumulation on I has the minimal value. These characteristics would have a possibility to realize the negative shift when the renewing of the weight matrix.

Here, I would like to consider about the cause of the sensitivity transition by using Figs. 7. Fig. 7-1 shows the remaining of counter potential in a memory state of SSFLC. When the optical modulation is ensured, FLC molecules are turn around the cone and the reversal of the spontaneous polarization is befallen.



7-1 : ion drift in LC brings counter potential in SSFLC

7-2 : electron drift in a-Si:H layer and accumulation to surface

7-3 : purpose of write light intensity to mobility of electrons in a-Si:H layer

Figs. 7 Examination of accumuration effect

As the reversal of the spontaneous polarization, flowing of polarization reversal current and the drifting of the comprised ions are turned out simultaneously. When in the memory state, the counter potential due to the drifted ions is survived. Corresponding to the drifted ions in the FLC layer, photonically generated electrons are drifted and accumulated to the surface, as shown in Fig. 7-2. These accumulated electrons would act to decrease the write light intensity threshold level. The sensitivity transition characteristics are controllable by the write light intensity and erase pulse. It would be very important when these characteristics were applied to the actual systems. However, the dependence of the sensitivity transition on the write light intensity exhibits the extreme value shown in Fig. 3. It would be explained by the difference of the low resistivity area generated by the write light, as shown in Fig. 7-3. If the sufficiently strong write light is irradiated, the drifted electrons easily frown out and the accumulation effect would be slackened.

4. Application for optical neural networks

Recently, the optical neural networks are actively investigated. Essentially the optical systems have very well matching for neural network because of the inherent parallelisms. For instance, an optical neural network system was presented by Psaltis and Farhat.[6] However, performances of the optical neural network systems that have been presented are critically eliminated to the step of thresholding and feedback. SLMs would be in charge of the mainly part of the optical neural network systems, and in addition, LAPS-SLM possesses very well performances as the function of memory and thresholding, higher contrast, higher resolution and faster response time, to this application.

From the data shown in the Figs. 3 and 4, it seems that the LAPS-SLM would work as a memory like a brain when it was used in the optical neural network system. When the LAPS-SLM are used, vector-matrix multiplexing is managed optically, and of course the LAPS-SLM is useful on the thresholding plane. Moreover, positive application of the sensitivity transition effect of the LAPS-SLM brings the learning operation on the threshold level. If the LAPS-SLM is used as the vector-matrix multiplexer, the weight matrix on the multiplexing plane and the output vectors are binary, but the threshold levels are continuously variable and are controllable. Namely, LAPS-SLMs work as the neural memories that learn on the threshold levels. The nonlinearity is assigned from the accumulation and the mitigation times.

By fig. 8, I would like to propose the neural network system using the LAPS-SLMs as the vector-matrix multiplexer and thresholding device. The sequence of the work of this system is as follows. First, the primer weight matrix W(N) is displayed on the LCTV and is written and latched on the LAPS-SLM1. Next, the weight matrix W(N) on the LAPS-SLM1 is red out by the light of input vector Vi that is generated from the LED array and expanded by the cylinder lens CL1. Then, the optical vector-matrix multiplexing is performed.

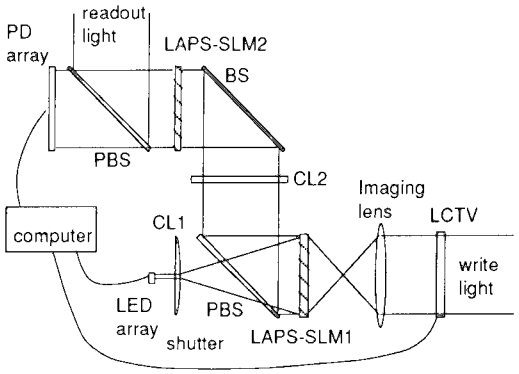


Fig. 8 Schematic illustration of optical neural network system applying the LAPS-SLM

The output light is contracted by the cylinder lens CL2, and projected on the LAPS-SLM2. At this time, thresholding are executed and the result vector $f(\Sigma W(N)Vi)$ is memorized on the LAPS-SLM2. The thresholding function f(x) on the LAPS-SLM2 is written as $f(x) = F[xi-h]$. Where F is a step function by the bistability of the FLC modulator, h is the threshold level of the LAPS-SLM2 and is controllable by the accumulating time t, the input light intensity I, driving pulse voltage p and pulse width q as shown in Figs. 3,4,5 and 6. Output is obtained by the PD array and the weight matrix on the LAPS-SLM1 is renewed by these results. The learning is carried out by this renewing of the weight matrix submitting to the delta rule. The renewing of the weight matrix is written as,

$W(N+1)=(1-b)\cdot W(N)+a\cdot\delta W(N)$.

Where W(N) is the weight matrix before renewing, a is the learning coefficient, b is forgetting coefficient and $\delta W(N)$ is the learning by the delta rule. Providing, the iteration number N is already including the effect of the accumulation and mitigation time. The learning and forgetting coefficients a and b are also carried out by the sensitivity transition characteristic of the LAPS-SLM, and written as,

$b=\beta\log t$ (β=constant), $a=-\alpha\log t$ ($\alpha=\phi(I, p, q)$).

These expressions show a new learning method introducing the effect of learning and forgetting time. In this system, convergence characteristics are affected by learning and forgetting time, and would be controllable.

Same as the vector-matrix plane, learning of the threshold level could be realized on the thresholding plane where the LAPS-SLM2 is applied.

In future, I will examine about the accurate description of the learning and experimental verification. Moreover, I would like to consider about the negative shift of the weight matrix by the effect of I on the accumulation.

5. Conclusion

The sensitivity transition characteristics of the LAPS-SLM are applied to an optical neural network system. In this system, LAPS-SLMs work as the neural memories that learn on the threshold levels.

6. References

[1] K. M. Johnson et al. 1989 *Appl. Opt.* **28** 4888
[2] N. Kato et al. 1991 *Proceedings SPIE* **1455** 190-205
[3] N. Kato et al. 1992 *Proceedings SPIE.* **1773** 275-283
[4] L. R. Peterson et al. 1959 *J. Exp. Psycol.* **58** 193-198
[5] G. Raisman et al. 1969 *Brain Res.* **14** 24-48
[6] D. Psaltis et al. 1985 *Opt. Lett.* **10-2** 98-100

VI. OPTICAL STORAGE AND SWITCHES

GUIDED WAVE TECHNIQUES

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

ALL-OPTICAL DYNAMIC MEMORIES

M.P.Petrov

A.F.Ioffe Physical Technical Institute, Russian Academy of Sciences, St.Petersburg, 194021, Russia

Abstract. The principles of operation of an all-optical dynamic memory (AODM) are considered. A considerable progress in the development of optical regenerators which are the key elements of AODM is emphasized. The analysis of tendencies in the development of AODM and numerical estimates have shown that AODM with the information capacity of more than 10^6 bits, practically unlimited storage time, bit error rate less than 10^{-9}, and acceptable optical power consumption can be built.

1. Introduction

An all-optical dynamic memory (AODM) is designed to store information in the form of optical pulses. Such a memory inevitably has to be dynamic but in the general case it is not necessary for it to be the all-optical one. However when the pulse duration becomes on the order of picoseconds and the repetition rate is tens of gigahertz the all-optical implementation of such a memory seems the only possible. An advantage of a very wide bandwidth of optical systems can be efficiently exploited in this case.

More obvious applications of AODM are in the area of digital optical computers, optical telecommunication systems, coding and decoding sytems, very broad-bandwidth signal processing.

The information capacity (N) of AODM is determined by the number of pulses which are placed along the optical path, so for the optical path of more than 1 meter it is reasonable to use an optical fiber as a stable medium where moving optical pulses are stored. For moderate storage time (10^{-6}– 10^{-2}s) the soliton system with partial compensation of losses and reduced jitter in pulse arrival times can be efficiently used [1,2]. But in the case of long (perhaps practically unlimited) storage time τ_m the optical pulses have to propagate at long distances or circulate many times around the fiber loop, running many millions of kilometers. Then to ensure a low bit error rate BER (10^{-9} or less) AODM has to be capable to support a high signal-to-noise ratio and to restore not only the intensity and shape, but also the exact time position of signal pulses. All these characteristics become poorer during a long travel of pulses because of light losses, group velocity dispersion, nonlinear optical phenomena in fibers, interaction between pulses and instabilities of various components of AODM. The problem of the AODM operation stability with a virtually unlimited storage time is very similar to that for ultralong optical data transmission systems and creates the same questions for the physics and

communication theory concerning information entropy increase.

2. Optical regenerators

To restore the signal pulse characteristics and to ensure stability of the fiber loop which is a system with feedback the optical regenerator that is really a key element of AODM has to be developed. The principles of operation of optical regenerators are very close to those of all-optical switching and logic elements [3]. Such nonlinear phenomena [4,5] as self-phase modulation (SPM), cross phase modulation (XPM), optical Kerr effect, selfrotation of the polarization ellipse, soliton interaction and stimulated Raman scattering (SRS) are widely used in manufacturing switching and logic elements. The nonlinear interferometer [6], nonlinear loop mirror (nonlinear Sagnac interferometer) [7-10], optical SRS invertor [11-13], optical switching system using soliton dragging effect [14], Kerr switching element [15], and active mode-locked laser using XPM [16] are the most well-known examples of such devices. Below two examples of all-optical regenerators which use the same principles as logic and switching elements are presented.

2.1. Nonlinear Sagnac interferometer (NSI)

NSI can perform selfswitching [7,8] or switching by a control pulse [9,10]. In the latter case timing of the output signal can be provided. The typical scheme of NSI (Fig.1) of the second type consists of an optical coupler (1) with ports 1,2,3,4 and coupling ratio 50:50, fiber loop and additional coupler (2) with port 5. The probe optical pulse can be launched into port 1 and the control one into port 5.The wavelengths of control and probe signals are assumed to be different. If there is no control pulse, the probe pulse at input 1 is completely reflected. This occurs because after the input signal propagates through coupler 1 the two output signals in ports 2 and 3 have equal amplitudes but the phase difference is $\pi/2$. Then

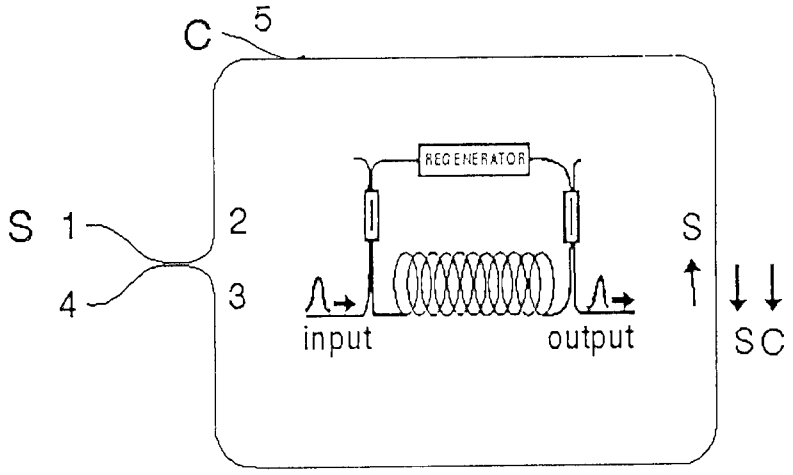


Fig.1 Diagram of a nonlinear Sagnac interferommeter.
At the insert: a schematic of AODM

these pulses travel through the loop in opposite directions and pass through the coupler. During propagation through the coupler these pulses acquire an additional phase difference $\pi/2$ resulting in a full signal compensation in port 4 and reflection of the signal in port 1. The situatuion can be completely changed when a control pulse is launched simultaneously with the probe pulse. Control pulse moves in one direction and creates a nonlinear phase shift φ_{NL} for the probe pulse moving in the same direction. Neglecting the phase shift for the counterpropagating pulse, we can consider φ_{NL} as an additional phase difference between the co- and counterpropagating pulses. Then

$$\phi_{NL} = \frac{2\pi}{\lambda} \Delta n_{NL} \cdot L_W \qquad (1)$$

where $\Delta n_{NL} = 2n_2 I_c$, $n_2 = 3.2 \cdot 10^{-20}$ m^2/W [4], I_c is the intensity of the control pulse and L_W is the walk-off length or distance along which the probe and control pulses move simultaneously providing overlap. L_W can be less than the loop length because the probe and control pulses move with different group velocities because of difference between the wavelengths of pulses and group velocity dispersion of the fiber. Under the condition that φ_{NL} approaches π the phase relationship between pulses results in switching optical energy into port 4. This type of nonlinear interferometers exhibits very good characteristics: a low control pulse intensity because of a long interaction length L_W, high temperature stability, low losses and provides an efficient way of timing of signals by control pulses. An example of such type of NSI used for regeneration of signals is described in [9].

In the experiments the output pulse duration τ_p was about 9 ps, repetition rate f_r - 5 Gbit/s, the signal pulse peak power - appr. 740 mW, switching contrast - 15 dB, BER = 10^{-9}. The loop length was 5 km. An excellent restoration of the pulse shape, intensity and time position was obtained.

The authors of [9] believe that the regenerator will operate well at bit rates above several tens of Gbit/s. Disadvantages of this system are degradation of characteristics when product $\tau_p f_r$ tends to unity and operation at two wavelengths, but the latter can be overcome under definite conditions by using the same wavelength but two different polarizations for clock and signal pulses if a birefringent fibre is used [17].

2.2. SRS regenerator

In [11-13] an all-optical SRS invertor which can be used as a regenerator has been proposed and studied. The optical invertor converts the high-level input signal into a weak output signal and a weak input signal into a high-level output one. The scheme of an SRS invertor (Fig.2) consists of the pump source, SRS amplifier (short fibre), rejection filter at the Stokes frequency and a SRS generator (long fibre). The SRS phenomenon can briefly be described as follows. If quite a strong pump pulse is launched into the fibre the amplification or generation of the Stokes pulse occurs.The Stokes signal is frequency down shifted as compared with the pump pulse. The intensity of the Stokes signals is

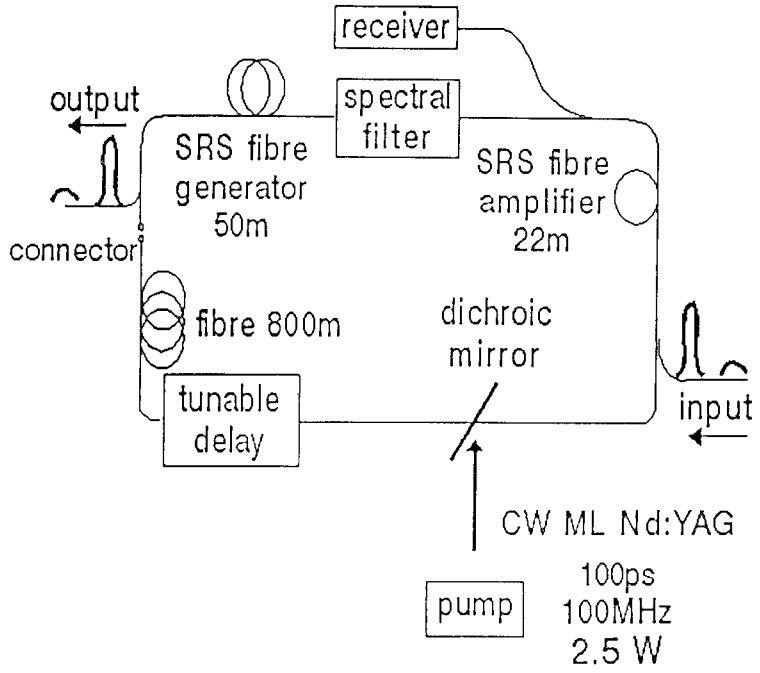


Fig.2 Schematic diagram of AODM with a SRS invertor. a)Connector is open. The system is operating as an isolated SRS invertor. b)Connector is closed. The system is operating as AODM.

$$I_S(L) = I_S(0) \exp (gI_pL) \qquad (2)$$

where L is the interaction length between the pump and Stokes pulses, I_p is the pump intensity, $g = 0.9\ 10^{-11}$ cm/W [4] and $I_S(0)$ is the initial Stokes intensity. Here we neglect for simplicity absorption and depletion of the pump signal.

When there is no input Stokes signal $I_S(0)$ is determined by noise scattering, and the fiber can serve as a generator of the Stokes signal if L is long enough. If there is an input signal $I_s(0)$ the fibre operates as a Stokes amplifier (a nonlinear amplifier in the general case).

Then the principle of operation of the SRS invertor can be easily understood. Let a signal pulse at the Stokes frequency be applied at the input. This pulse is amplified in a short fibre and causes pump depletion. But the amplified Stokes pulse is rejected by the filter and does not reach the output. So there is no Stokes signal at the output. In the opposite case if there is no Stokes signal at the input a powerful pump pulse propagates through the first (short) fibre undepleted and then it enters the second (long) fibre where the Stokes signal is generated. So the Stokes signal appears at the output. If this process is repeated twice the input pulse is regenerated. The SRS regenerator ensures timing by synchronous pumping. Restoration of the intensity and stabilization of the signal shape are due to periodical generation of the output Stokes pulses.

3. Selected schemes of AODM

a) Dynamic optical memory with electrooptic regenerator [18,19].
The scheme is not quite all-optical because the regenerator includes the electrically controlled optical modulator using $LiNbO_3$ crystal. But the characteristics of this system approach those of the all-optical one. The system operates like an actively mode locked ring fiber soliton laser. A one-million-kilometer trip of optical solitons around a 510-km long fiber loop containing ten erbium-doped fiber amplifiers has been demonstrated (storage time more than 5 s). The repetition rate was 5-10 Gbit/s, pulse duration - 30-40 ps and peak power P_p = 0.6 - 2 mW. BER was less than 10^{-9}. The authors [18,19] pointed out that the electrooptical regenerator ensures excellent suppression of jitter and restoration of the shape of solitons in spite of a rather short distance between them. So the authors claimed that solitons can be transmitted at unlimited distances. Unfortunately, the papers [18,19] do not contain data about the absolute stability of synchronization between reshaped signals and clock regenerator which is important for memory applications. Potentially the system with the fibre loop of 510 km can contain more than 10^7 bits of information.
b) AODM using nonlinear Sagnac interferometer [20].

This scheme exploits the signal and control pulses at the same frequency but with orthogonal polarizations. If the output signal pulse is used as a control one, the system operates as an invertor or a shift register. The repetition rate is 100 MHz, pulse duration is appr. 15 ps. The timing is provided by the external Nd:YAG pump pulse laser. The optical losses are compensated for by an erbium doped fiber amplifier. The length of the loop in NSI was 450 m, so the memory volume was 254 pulses. It demonstrates a stable operation for hours in the regime of optical invertor.
c) AODM with SRS invertor [21,22] (Fig.2).
This example of AODM is one of the most stable and simple because a SRS regenerator used in the scheme is only slightly affected by temperature and other outside factors, but provides a high signal-to-noise ratio. It has experimentally been demonstrated that τ_m > 10 min which means that signals travel for more than 100 million kilometers around the fibre loop. The repetition rate was about 100 MHz, pulse duration was 70-100 ps, and the number of circulating pulses was about 600. The signal-to-noise ratio was more than 50. A Nd:YAG laser at λ = 1.06 μm is used as a pump source and the signal wavelength corresponded to the Stokes wavelength (1.12 μm). The system exhibits a stable operating regime on variation of the optical path length by a tunable delay line in the interval appr. equal to the pulse duration as well as in case of the pump power fluctuations.

4. Estimates of general characteristics of AODM

All the most important characteristics of AODM such as N, τ_m, BER, and average signal power consumption P_{av} are related to each other. Let us estimate roughly some characteristics of the system with regenerator based on nonlinear cross phase modulation and operating with soliton-like pulses.
P_{av} and N can be estimated from

$$P_{av} \cong P_p \tau_p f_r$$

$$N \cong L_t f_r / V_{gr} \quad (3)$$

where L_t is the total length of a fiber memory system and V_{gr} is the group velocity. BER is a function of the peak pulse power P_p. P_{av} has to be well below 1 W because of the problem of a limited dynamic range of the signal amplifier. It is desirable that the product $\tau_p f_r$ be on the order of 10^{-1} or less to prevent soliton-soliton interaction or influence of counterpropagating stream of pulses in NSI. So the reasonable magnitude of P_p is much less than 10 W. It can be made low enough by using sufficiently long walk-off length L_w. Assuming that a necessary critical value of φ_{NL} equals π the relationship for P_p can be found from (1)

$$P_p = \lambda S_{eff} (4 L_W n_2)^{-1} \quad (4)$$

where S_{eff} is the effective crosssection of the fiber core. But in reality P_p cannot be too low for very short pulses because it is difficult to get very long L_w and because a low BER and hence a high signal/noise ratio is required. For the reasonable value of $S_{eff} \sim 50\ \mu m^2$, $\lambda = 1.5\ \mu m$, and $L_w \sim$ 500 m, P_p is 1.1 W. Then for $f_r = 50$ Gbit/s and $N = 10^6$, the total fiber length L_t is appr. 4 km. If a fiber with a proper group velocity dispersion is used, the pulse duration can be as short as 1-3 ps. BER can be less than 10^{-9} in this case. For much higher N and L_t a number of regenerators can be used. These estimates are very rough, but more precise calculation for dispersion-shifted fibers give even more optimistic parameters.

5. Conclusion

Feasibility of AODM using nonlinear optical phenomena in fibers has been experimentally demonstrated with information capacity $N \sim 10^{2-3}$ and storage time $\tau_m > 10$ min. Efficient optical regenerators have been developed, so improvement of AODM characteristics up to $N > 10^6$, practically unlimited τ_m, repetition rate 10-50 $Gbits^{-1}$ and optical energy consumption less than 1 W at a low bit error rate seems to be absolutely realistic. One of the major problems here is to find application areas where AODMs will be most efficiently and optimally used.

The authors are thankful to Dr. V.I.Belotitskii and Dr. V.V.Spirin for helpful discussions.

References

[1] Mollenauer L F, Lichtman E, Neubelt M J, and Harvey G T 1993 *Electr. Lett.* **10** 910-911
[2] Widdowson T, Malyon D J, Shan X and Wakinson P J 1994 *Electr. Lett.* **3** 661-662
[3] Petrov M P 1993 in *Optical Information Proccessing,* Gulyaev Y V and Pape D R, Eds. Proc. SPIE **2051**, 198-211

[4] Agrawal G P 1989 *Nonlinear Fiber Optics* (Academic Press)
[5] Shen Y R 1984 *The principles of nonlinear optics* (John Wiley and Sons)
[6] Nayar B K, Finlayson N, Doran N J, Davey S T, Williams D L and Arkwright J W 1991 *Opt. Lett.* **16** 408
[7] Nayar B K, Blow K J and Doran N J 1991 *Opt.Comput. and Processing* **1** 81-89
[8] Blow K J, Doran N J, Nayar B K and Nelson B P 1989 *Opt.Lett.* **14** 754
[9] Jinno M and Abe M 1992 *Electr. Lett.* **28** 1350-11352
[10] Blow K J, Doran N J, Nayar B K and Nelson B P 1990 *Opt.Lett.* **5** 248-250
[11] Petrov M P and Kuzin E A 1985 *Preprint 975* (A F Ioffe Phys.Tech.Institute Leningrad USSR)
[12] Belotitskii V I, Kuzin E A, Petrov M P and Spirin V V 1988 *Zh.Tekh.Fiz.* **58** 1325-1329(in Russ)
[13] Petrov M P and Kuzin E A 1989 in *Optical Proccessing and Computing*, Eds. Arsenault H, Szoplick T and Macukow (Academic Press, Inc) 33-57
[14] Islam M N, Soccolich C E and GordonJ P 1992 *Optical and Quantum Electr.* **24** 1215-1235
[15] Morioka T and Sanuwatari M 1988 *IEEE J. on Selected Areas in Commun.* **6** 1186-1198
[16] Lucek J K and Smith K 1993 *Opt.Lett.* **18** 1228
[17] Whitaker N A Jr., Avramopoulos H, French P M W, Gabriel M C, La Marche R E, Di Giovanni D J and Presby H M 1991 *Opt.Lett.* **16** 1838
[18] Nakazawa M, Yamada E, Kubota H and Suzuki K 1991 *Electr. Lett.* **27** 1270-1272
[19] Nakazawa M, Suzuki K, Yamada E, Kubota H, Kimura Y, and Takaya M 1993 *Electr.Lett.* **29** 729-730
[20] Whitaker N A Jr., Gabriel M C, Avramopoulos H and Huang A 1991 *Opt.Lett.* **16** 1999-2001
[21] Petrov M P , Kuzin E A, Belotitskii V I and Spirin V V 1992 *Pis'ma v Zh.Tekh.Fiz.* **18** 29
[22] Belotitskii V I, Kuzin E A, Petrov M P and Spirin V V 1993 *Electr.Lett.* **29** 49-50

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

A Novel Information Storage Device Utilizing Photostimulated Luminescence in Eu-Doped KCl Phosphors

Hidehito NANTO, Fumitaka ENDO, Yoshiaki HIRAI, Mitsuo IKEDA, Shouichi NASU* and Nozomu TAKEUCHI**

Electron Device System Research Laboratory, Kanazawa Institute of Technology, 7–1 Oogigaoka, Nonoichi–machi, P.O.Kanazawa–South, 921 Japan
**Department of Electrical Engineering, Kanazawa Institute of Technology, 7–1 Oogigaoka, Nonoichi–machi, P.O.Kanazawa–South, 921 Japan*
***Faculty of Technology, Kanazawa University, 20–Kodatsuno, Kanazawa, Ishikawa* 920 Japan

Abstract. Intense photostimulated luminescence (PSL) with a peak at 420 nm is observed in ultra–violet(UV)–light–irradiated europium–doped potassium chloride (KCl:Eu) crystalline phosphors. The PSL characteristics of UV–irradiated KCl:Eu phosphor for optical memory application are studied. The excitation and emission mechanisms of the 420 nm PSL, which are consistent with the results obtained, are discussed.

1.Introduction

A new type of optical memory based on the photostimulated luminescence (PSL) phenomenon in electron trapping phosphor materials for optical storage, has been studied in the fields of optical parallel Boolean logic operations [1], optical associative memory [2] and optical neural networks [3,4]. The electron trapping phosphor materials can emit different output photons which correlate spatially in intensity with the input photons. Consequently, the electron trapping phosphor materials can be used to store optical information as trapped electrons and the information stored can be read out by a laser beam scanning of the phosphor materials [4,5]. The unique features of the electron trapping phosphor materials which exhibit the PSL phenomenon provide the potential for high bit storage densities, high data transfer and fast recovery speeds[6]. Important characteristics of good electron trapping phosphor materials for optical memory are high PSL brightness for low noise, short luminescence lifetime for minimum read out time and low light scattering for high bit storage densities. Especially, the electron trapping phosphor materials using transparent single crystals or thin films provide an efficient PSL and low light scattering.

As a result of surveying many possible transparent alkali halide phosphors in order to obtain a new electron trapping phosphor material with high PSL brightness and low light scattering, we found that transparent Eu–doped potassium chloride (KCl:Eu) crystalline phosphors exhibit an efficient PSL for optical stimulation with visible light after ultra–violet (UV) light irradiation at room temperature (RT). In this paper, we report the PSL characteristics of KCl:Eu crystalline phosphors for optical memory utilizing the PSL phenomenon. The PSL excitation and emission mechanisms are also discussed.

2.Experimental

Crystalline phosphors of KCl:Eu used in the present study were grown by the Bridgman method from a molten mixture of reagent grade KCl and $EuCl_3 \cdot 6H_2O$ (0.1 mol%). The Eu^{2+} concentration in the grown crystals was determined to be about 1.4×10^{18} ions/cm^3 from the intensity of the Eu^{2+} characteristic optical absorption bands at about 243 nm and 343 nm using Smakula's formula. The concentration of Eu^{2+} ions in the crystal is about one order of magnitude smaller than that of $EuCl_3$ in the original mixture. The KCl:Eu samples of

typically 5×5×1 mm^3 in size were kept at 550 ℃ for 30 min and then quenched to RT to disperse the aggregation of Eu impurity.

The UV−light irradiation for optical excitation was carried out using an Hg−lamp. The PSL measurements were carried out using a Hitachi F−3010 spectrofluorometer at room temperature (RT). The optical absorption spectra were measured with a Hitachi U−2000 spectrophotometer. The PSL spectra were corrected for the diffraction of the grating and optical response of the photomultiplier. Details of the system to measure luminescence have been described elsewhere [7].

3.Results and Discussion

An intense PSL peak at about 420 nm was observed when the UV−light irradiated specimen was stimulated with 580 nm light at RT. A typical 420 nm PSL emission spectrum (solid line) as well as stimulation (read−out) spectrum (dashed line) for the 420 nm PSL peak from 240 nm UV−light irradiated KCl:Eu sample is shown in Fig.1. Excitation (write−in) spectrum (dotted line) for the 420 nm PSL peak is also shown in Fig.1. The 420 nm PSL emission is assigned to an inner ionic transition ($4f^6 5d \rightarrow 4f^7$) of isolated Eu^{2+} ions which occupy cation sites in KCl crystal, since photoluminescence with a peak at about 420 nm was observed when the specimen was excited with 243 nm or 343 nm light .This wavelength corresponds to that of the characteristic optical absorption band maximum of isolated Eu^{2+} ions [8]. The optical absorption spectra before (dashed line) and after (solid line) UV−light irradiation are also shown in Fig.2. The broad optical absorption band with a peak at about 560−580 nm, which is created by UV−light irradiation is due to the F centers [9] and/or the Z_1 centers, since it is believed that alkali halide crystals doped with Eu, Sm, Ca, Ba and Sr etc. (impurities which have low second ionization potential) give rise to Z_1 centers when F centers are optically bleached at RT [10,11]. It is likely that the F centers are bleached during UV−light irradiation, because the Hg−lamp was used as the UV−light source. Since this absorption band is in good agreement with the stimulation spectrum as shown in Fig.1, one can point out that there is a close relationship between the 420 nm PSL peak and the F centers and/or Z_1 centers. The excitation spectrum (dotted line) as shown in Fig.1 agrees with high−energy absorption band of Eu^{2+} ions as shown in Fig.2. This result suggests that the 420 nm PSL is only excited by Eu^{2+} high−energy absorption band. It has been reported that Eu−doped KCl single crystals exhibited a photoconductivity during UV−light irradiation as shown in Fig.3 [12]. The photoconductivity spectrum as shown in Fig.3 is coincident with the excitation spectrum of 420 nm PSL peak as shown in Fig.1. The result as described above is, thus, consistent with the following excitation and emission mechanism of the 420 nm PSL peak.

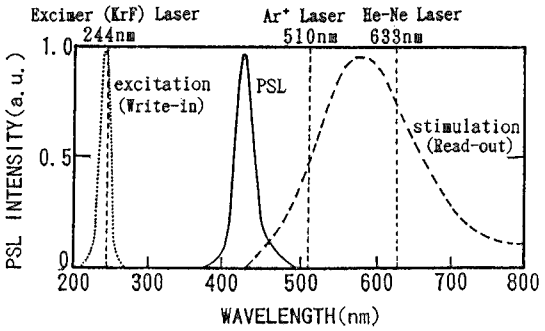


Fig.1 Typical PSL emission spectrum (solid line) when UV−irradiated KCl phosphor was stimulated with 560 nm light. The excitation spectrum (dotted line) and stimulation spectrum (dashed line) for the 420 nm PSL peak are also shown in the figure.

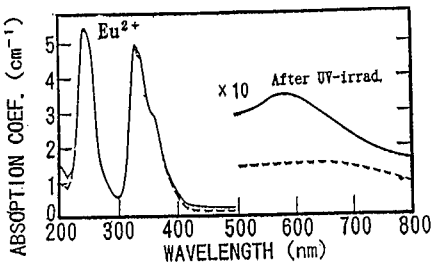


Fig.2 Optical absorption spectra before (dashed line) and after (solid line) UV–light irradiation.

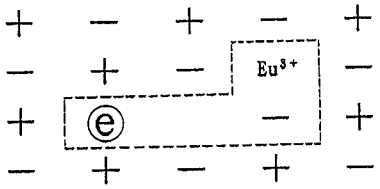


Fig.4 Possible configrational model for the electron traps in UV–ray irradiated specimen.

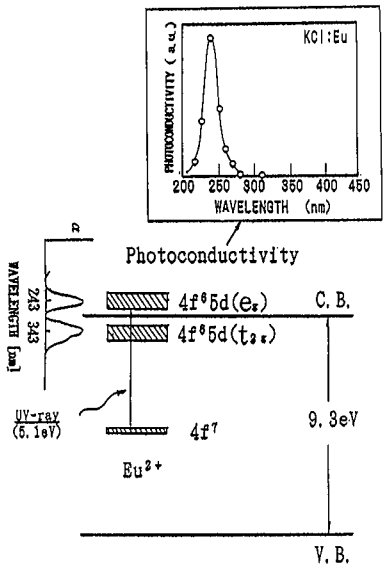


Fig.3 Photoconductivity spectrum [12] of KCl:Eu phosphor measured by UV–ray irradiation and
energy band diagram for excitation mechanism in UV–ray irradiated specimen.

Due to the overlap between the Eu^{2+} high–energy absorption band and the KCl conduction band, Eu^{2+} photoionization and subsequent free electrons occur during UV–light irradiation as schematically shown in Fig.3. It is, therefore, likely that some of the electrons excited from Eu^{2+} ions to the conduction band by UV irradiation are trapped at anion vacancies neighboring Eu^{3+} ions to produce the F centers strongly perturbed by neighboring positive ion vacancies and Eu^{3+} ions as shown in Fig.4, like Z_1 centers. By subsequent stimulation with 560–580 nm light, electrons which are optically released from the complex centers recombine with the Eu^{3+} ions leading to the excited Eu^{2+} ions from which the 420 nm PSL is emitted. Obviously, transparent KCl:Eu crystalline phosphors can be used to store optical information using the UV–light such as a KrF laser (248 nm) and the information can be read out using visible light ,a He–Ne laser or Ar^+ ion laser, because the wavelengths of peak maximum of excitation and stimulation spectrum are 240 nm and 580 nm, respectively, as shown in Fig.1. In addition to storage, the KCl:Eu phosphor is a subject of interest as hardware devices capable of performing multiplication, addition and subtraction within a dynamic range covering a few orders of magnitude, since the 420 nm PSL emission intensity is proportional to the UV–light irradiation dose as shown in Fig.5. This result also suggests that the PSL phenomenon in KCl:Eu phosphors is applicable to an analog memory device for a learning optical neural network.

The lifetime of the 420 nm PSL was about 1.6 μsec. This lifetime is fairly good for high speed operation of the optical memory. We confirmed that the information stored in UV–light irradiated specimen exhibited excellent fading characteristics at RT in the dark. Figure 6 shows the fading characteristics of the PSL intensity in UV–light irradiated sample at RT in the dark. It can be seen that the PSL intensity decreases with increasing the passage time in the early stage and then saturates to a constant PSL intensity. One can, thus, expect that stored UV write–in information is fairly stable at RT in the dark.

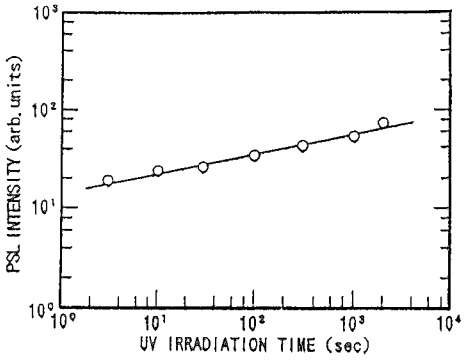


Fig.5 PSL intensity as a function of UV write–in intensity in KCl:Eu crystal.

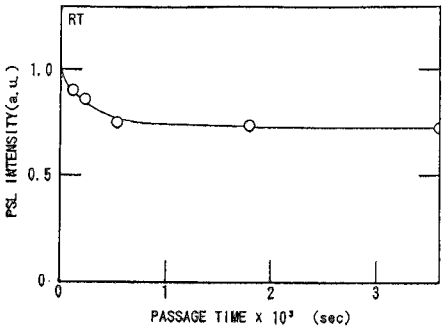


Fig.6 Fading characteristics of the PSL intensity in UV– irradiated KCl:Eu crystal at RT in the dark.

4.Summary

We have observed an intense PSL with a peak at about 420 nm when UV–light irradiated KCl:Eu crystalline phosphors are stimulated with 560–580 nm light. The PSL intensity is proportional to the UV irradiation dose. The KCl:Eu phosphor as the electron trapping phosphor material is one of the most attractive candidates for erasable and rewritable optical memory utilizing the PSL phenomenon, since excimer lasers such as KrF (248 nm) and KrCl (222 nm) lasers and visible–emitting lasers such as He–Ne and Ar^+ ion lasers can be used as write–in and read–out light sources, respectively.

Acknowledgements

We wish to thank Dr. S.Taniguchi, Prof. K. Inabe and Mr. S. Nakamura for their fruitful discussion and H. Enomoto, M. Totsuka, M. Ikeda and M. Ito for their technical assistance. This work is partly supported by Nakatani Electronic Measuring Technology Association of Japan and a Grant–in–Aid for General Scientific Research (No.05245214 and No.06236101) from the Ministry of Education, Science and Culture of Japan.

References

[1] Jutamulia S, Storti G M, Lindmayer J and Seiderman W 1990 *Appl. Opt.* 29, 4806.

[2] Jutamulia S, Storti G M, Lindmayer J and Seiderman W 1991 *Appl. Opt.* 30, 2879.

[3] Itoh F, Kitayama K and Tamura Y 1990 *Opt. Lett.* 15, 860.

[4] Jutamulia S, Storti G M, Lindmayer J and Seiderman W 1991 *Appl. Opt.* 30, 1786.

[5] Nanto H 1993 *Kotai Butsuri* 28, 49.

[6] Lindmayer J 1988 *Solid State Technol.* 8, 135.

[7] Nanto H, Usuda T, Murayama K, Sokooshi H, Nakamura S, Inabe K and Takeuchi N 1991 *Appl. Phys. Lett.* 59, 1838.

[8] Rubio J O, Murrieta H S, Hernandez J A and Lopez F J 1981 *Phys. Rev.* B24, 4847.

[9] Rubio J O, Flores M C, Murrieta H S, Hernandez J A, Jaque F and Lopez F J 1982 *Phys. Rev.* B26, 2199.

[10] Gartia R K and Acharya B S 1978 *Phys. Stat. Sol. (a)* 47, K165.

[11] Parfianovich I A, Penzina E E, Khulugurov V M, Metzik V M and Sobolev L M 1978 *Phys. Stat. Sol. (b)* 89, K91.

[12] Aguirre de Carcer I, Cusso F and Jaque F 1988 *Phys. Rev.* B38, 10812.

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

AlAs/GaAs Multilayered Structures as Low Intensity Optical Media

Steffen Knigge, Markus Wicke, Dirk Niederland, and Dieter Jäger

Fachgebiet Optoelektronik, FB 9, Gerhard-Mercator-Universität Duisburg, Kommandantenstr. 60, 47057 Duisburg, Germany

Abstract. The nonlinear optical properties of a hybrid AlAs/GaAs multilayered heterostructure are presented. Experimentally, optical bistability at optical intensities as low as 30 $mWcm^{-2}$ is obtained when a bias voltage is applied perpendicular to the layers.

1. Introduction

In 1991 He and Cada [1] predicted optical bistability in a multilayered periodical AlAs/GaAs structure at optical intensity levels of 10 $kWcm^{-2}$. In the following year Cada *et al* [2] achieved the experimental proof of this prediction in a 30 period GaAs/AlAs structure and in 1993 He *et al* [3] reported switching times < 50 ps for optical bistability in a heterostructure. Establishing a dc electric field perpendicular to the layers the authors achieved electrically induced optical bistability in a similar multilayered structure and the threshold intensity for the bistability has been lowered by three orders of magnitude down to 1.4 Wcm^{-2} [4]. A phenomenological model for this hybrid case was presented proposing two states of the device, charge transport and charge storage state [5]. Formulating equations for this electrically induced optical bistability, Ivanov and Haug [6] determined optical switching times for this bistability < 100 ns leading to switching energy densities < 1 $fJ\mu m^{-2}$.

2. Experimental results

In this paper, a 20 period AlAs/GaAs multilayered structure is investigated (Fig. 1). The layers are grown by molecular beam epitaxy (MBE) on an i-GaAs substrate with nominal thicknesses of 58 nm for GaAs and 69 nm for AlAs, respectively. Gold contacts are evaporated on the top and the bottom to establish a dc electric field in the structure. The load resistance can be magnified with the external resistor R_{ext}.

First, as an electro-optical characterization, the reflection spectra of the structure with and without an applied voltage V are shown at a temperature of $\vartheta = 13.5$ ^0C in

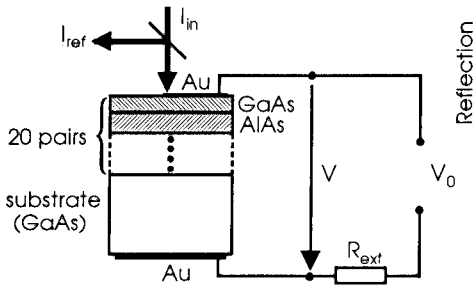


Fig. 1: Sketch of the hybrid multilayered structure.

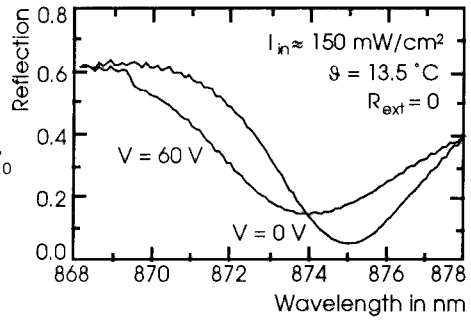


Fig. 2: Reflection spectra measured with and without an impressed voltage.

Fig. 2. It is remarkable, that for wavelengths λ = 868 nm to 874 nm the reflection $\mathcal{R}$ is lowered and for λ = 874 nm to 878 nm $\mathcal{R}$ is magnified by an applied electrical field perpendicular to the layers.

In a further experiment, $\mathcal{R}$ is measured versus the incident optical intensity I_{in} at λ = 875 nm and an impressed voltage V = 60 V (Fig. 3). Starting at low intensities, first $\mathcal{R}$ decreases with increasing I_{in}. Moreover, at a threshold intensity I_{th} = 96 mWcm^{-2} the reflection switches to a higher state. As can be seen from Fig. 3, a switching contrast as high as 5.0 dB is achieved. Beyond that point, $\mathcal{R}$ is constant with increasing I_{in}. Lowering I_{in} again, $\mathcal{R}$ remains almost constant at the higher level and is switching down to the lower state at I_{in} = 12 mWcm^{-2}. Hence, a hysteresis loop is formed.

In another experiment the electro-optical modulator properties of the device are compared with the properties under conditions for optical bistability. The solid line in Fig. 4 shows the contrast ratio of the electro-optical modulation, K_{mod} = 10 * lg[$\mathcal{R}$(60V)/$\mathcal{R}$(0V)], as a function of λ at I_{in} = 150 mWcm^{-2}. In agreement with Fig. 2 one can find a change of the sign of K_{mod}. K_{mod} has a negative value for λ = 868 nm to 874 nm and a positive one for λ = 874 nm to 878 nm, respectively. As shown in Fig. 4, K_{mod} reaches the peak value of 5.9 dB at λ = 875.3 nm. The crosses in Fig. 4 represent the switching contrast at I_{th} with V = 60 V for different wavelengths. In the wavelength range of λ = 874 nm to 878 nm upswitching of $\mathcal{R}$ occurs as sketched in Fig. 3. In this region the counter clockwise hysteresis loops give a positive value of the switching contrast. In contrast to this counter clockwise hysteresis, clockwise hysteresis loops occur in the range of λ = 868 nm to 874 nm. In this case, starting at low intensities, $\mathcal{R}$ increases with increasing I_{in}. At I_{th} $\mathcal{R}$ switches down to a lower state. Beyond this point of I_{in}, $\mathcal{R}$ stays almost constant. Lowering I_{in} $\mathcal{R}$ switches back to the first state at $I_{in} < I_{th}$.

Using a second device, I_{th} is determined as a function of the voltage V at room temperature. In Fig. 5 the contrast ratio for upswitching of $\mathcal{R}$ at I_{th} is shown in the voltage range of V = 40 V to 115 V. As can be seen, I_{th} decreases with increasing voltage. As a key result, I_{th} is lowered down to 30 mWcm^{-2}.

In a last experiment the influence of the load resistance on the optical bistability is observed using the second device. Increasing the load resistance by the additional, external resistor R_{ext}, the contrast ratio at I_{th} is decreasing (Fig. 6). It is remarkable

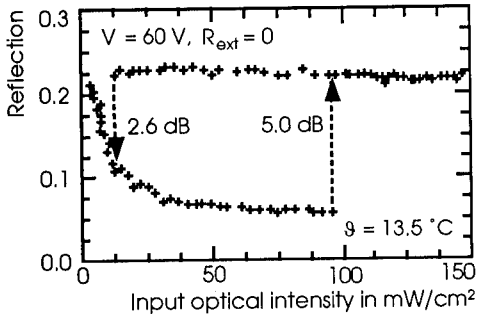


Fig. 3: Optical bistability measured at λ = 875 nm.

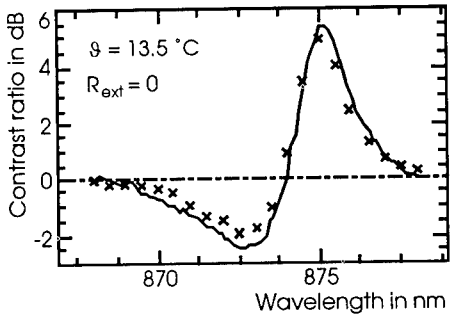


Fig. 4: Contrast ratio of electro-optical modulation within a voltage swing of 60 V (solid line) and contrast ratio of optical bistability at I_{th} with V = 60 V (crosses).

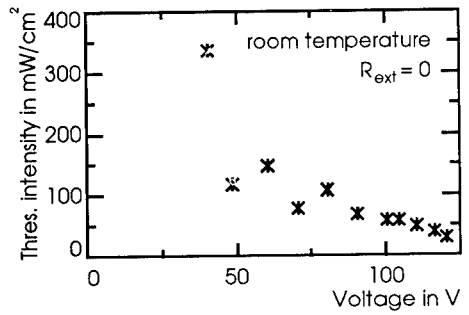


Fig. 5: Threshold intensity I_{th} for the optical bistability versus the applied voltage.

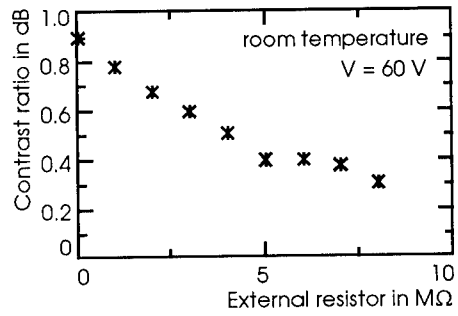


Fig. 6: Contrast ratio of optical bistability as a function of the external resistor

that the peak value of the optical switching contrast is achieved at $R_{ext} = 0$.

3. Discussion

Our measurements show electro-optical modulation and electrically induced optical bistability using an AlAs/GaAs multilayered structure. As depicted in Fig. 3 optical bistability with a switching contrast of 5 dB is realized at λ = 875 nm and compared with the results in [4], the nonlinearity of the structure is extremely enlarged and the switching contrast is enhanced by more than a factor of 5.

In Fig. 4, the contrast ratio of optical bistability is identical to the behaviour of $K_{mod}(\lambda)$. Hence, the switching contrast of the bistability and the electro-optical modulation are in coincidence. So for a given optically induced change of the optical parameters the switching contrast can easily be enlarged by a modified structure, e. g. a distributed-feedback/Fabry-Perot etalon configuration.

As can be seen from Fig. 5, the optical threshold intensity is lowered by increasing the applied voltage perpendicular to the layers. Hence, the responsivity of the structure can be improved by an increasing electrical field in the structure. As a consequence, for the first time optical bistability is achieved at an optical threshold intensity as low as 30 $mWcm^{-2}$ in such a multilayered structure. Lowering the spot diameter from 60 μm

down to 20 μm an optical threshold power of about 100 nW is expected for bistability. Therefore an array of 100 * 100 elements can easily be switched using a laser diode with an optical power < 10 mW.

In Fig. 6 one can see, that the optical switching contrast ratio is lowered by an additional external resistor R_{ext}. As a key result, the contrast ratio of the optical bistability could be enhanced by a decreasing load resistance. Furthermore, a lowering of the load resistance decreases the RC time constant of the device. As mentioned by Ivanov and Haug [6] this RC time constant is the limiting factor for the dynamical properties of the device. As a consequence, a decrease of the internal load resistance will most probably improve the optical switching contrast and the dynamical behaviour of our device.

4. Conclusion

In summary, extremely enlarged nonlinear optical properties and highest optical responsivity are measured at a hybrid multilayered heterostructure. These fantastic properties of this nonlinear medium are leading to various potential applications for low intensity digital optical computing devices.

Acknowledgments

We appreciate stimulating discussions with H. Haug and A. L. Ivanov of the university of Frankfurt and G.-H. Döhler and P. Kiesel of the university of Erlangen-Nürnberg. The authors thank C. Chaix of Riber for the MBE growth of a multilayered structure. This study is part of a project which is financially supported by the Volkswagen-Stiftung, Hannover.

References

[1] He J, Cada M 1991 Optical bistability in semiconductor periodic structures *IEEE J. Quant. Electron.* **27**, 1182–1188

[2] Cada M, He J, Acklin B, Proctor M, Martin D, Morier-Genoud F, Dupertuis M A and Glinski J M 1992 All-optical reflectivity tuning and logic gating in a GaAs/AlAs periodic layered structure *Appl. Phys. Lett.* **60**, 404–406

[3] He J, Cada M, Dupertuis M A, Martin D, Morier-Genoud F, Rolland C and SpringThorpe A J 1993 All-optical bistable switching and signal regeneration in a semiconductor layered distributed-feedback/Fabry-Perot structure *Appl. Phys. Lett.* **63**, 866–868

[4] Knigge S, Zumkley S, Wingen G, Humbach O, Chaix C and Jäger D 1992 Experiments on optoelectronic bistability in distributed AlAs/GaAs-Bragg reflectors *Microelectronic Engineering* **19**, 52–56

[5] Knigge S, Pint H, Wicke M, Wingen G, Zumkley S, Bugge F, Chaix C and Jäger D 1993 Nonlinear hybrid AlAs/GaAs Bragg reflectors for photonic switching applications In *Optical Information Processing*, Y. V. Gulyaev and D. R. Pape (Eds.), SPIE Vol. 2051, 148–151

[6] Ivanov A L and Haug H 1994 Modelling of a low-intensity electro-optical semiconductor switching device due to intrinsic photoconductivity *J. Appl. Phys.* **76**(4)

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Large Reduction of Saturation Carrier Densities by Strain in InGaAs/AlGaAs Quantum Wells

M Moloney and J Hegarty

Department of Physics, Trinity College, Dublin 2, Ireland.

L Buydens and P Demeester

I.M.E.C., Rijkuniversiteit Gent, Sint-Pietersnieuwstraat, 41, B-9000 Gent, Belgium.

Abstract. Compressive strain in InGaAs/Ga(Al)As quantum wells leads to a reduction in the saturation carrier density for the bandedge exciton due to a decrease in the hole density of states. An Indium concentration of 11% gives a reduction by a factor of 9 from GaAs thereby opening a new path towards viable nonlinear optical devices. The relevant issues of carrier lifetimes and strain relief are discussed.

The concept of all-optical computing based on nonlinear optical devices has been in vogue for several years. Considerable effort has been spent on developing suitable architectures and materials with sufficient nonlinearity to be useful in device arrays. Unfortunately to date even the strongest known nonlinearity is still insufficient to be useful. One of the most promising materials is a direct gap quantum well such as the prototypical GaAs/GaAlAs structure with its strong saturable exciton resonance at the bandedge. Saturation of the exciton absorption through band filling effects at room temperature [1] is a function of the densities of states of the electrons and holes making up the exciton. Any reduction in the density of states would lead to more effective bandfilling for fixed carrier concentration and hence to a lower saturation density and possibly lower saturation intensity, the latter being most important for devices.

We have shown that in an InGaAs/GaAs quantum well Fabry-Perot device structure [2] compressive strain in the InGaAs wells gives a factor of 2 reduction in saturation carrier density. Jin et al [3] have also demonstrated a reduction in a well structure containing 15% Indium. More recently [4], we have shown that with proper control of strain, the saturation density can be reduced by up to an order of magnitude compared to GaAs wells. This occurs at an In concentration of about 10% and indicates that further increase in In concentration without strain relief could lead to further reduction in saturation values. This represents a significant step towards a set of materials with sufficient nonlinearity for useful devices. In this paper, we summarise these key results and discuss the problems that need to be resolved before the effects can be used. In particular, carrier lifetimes in Indium doped GaAs are short [5] for reasons not quite understood. Since carrier lifetimes are directly involved in saturation *intensities*, application to nonlinear devices will require this material parameter to be optimised.

Samples of $In_xGa_{1-x}As$/GaAlAs quantum wells were grown by metal organic vapour phase epitaxy (MOVPE) [6], with x ranging from 0 to 0.15. Low temperature luminescence and absorption measurements together with the known well widths gave accurate values for the Indium concentrations. A series of samples with x = 0, 0.03, 0.11 and 0.15 labelled S1 to S4, respectively in the figures were used for detailed measurements. Details of the samples are given in Table 1. Samples with x = 0.03 and 0.11 were fully strained. The x = 0.15 sample was partially strain relieved, as shown by cross-hatching under a Nomarskii microscope. Even though the GaAs substrate has a larger bandgap than the InGaAs wells, bandtailing in the

substrate prevented nonlinear measurements in transmission for low Indium concentrations so the substrate was etched away in all samples. All measurements were made at room temperature.

Table 1

Sample Code	Indium Concentration	No. of Periods	Well/Barrier Width (nm)	Barrier Composition
S1	0.00	10	8/8	$Al_{0.2}Ga_{0.8}As$
S2	0.03	10	8/8	$Al_{0.2}Ga_{0.8}As$
S3	0.11	10	8/8	$Al_{0.2}Ga_{0.8}As$
S4	0.15	10	8/8	$Al_{0.2}Ga_{0.8}As$

Details of InGaAs/AlGaAs samples.

The saturation carrier density N_{sat} for the lowest energy exciton resonance was calculated from the dependence of the peak excitonic absorption coefficient $\alpha(N)$ on incident resonant light intensity I according to the empirical saturation equation: $\alpha(N) = \frac{\alpha_0}{1 + N/N_{sat}}$ where N is the carrier density and α_0 is the saturable low intensity absorption coefficient. $\alpha(N)$ is measured from the intensity dependent transmission spectra near the exciton resonance using the technique described in [4]. The carrier density N is determined from $N = \alpha(N) I\tau/h\nu$ where τ is the carrier lifetime, $h\nu$ is the photon energy and I is the incident intensity. The lifetime τ is measured by the standard pump/probe technique over the intensity range used to measure $\alpha(I)$.

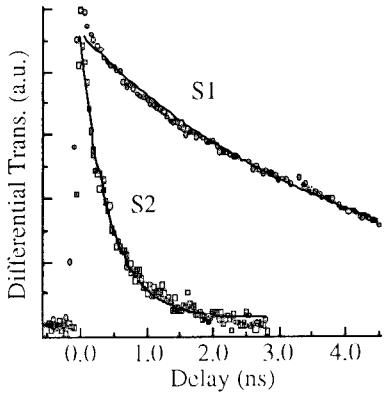


Figure 1 Pump/probe decays at peak of heavy hole exciton in S1 (GaAs) and S2 (3% In).

Figure 1 shows decays of the probe signal for the case of x = 0 (S1) and x = 0.03 (S2) as a function of delay after the pump where both pump and probe are resonant with the lowest energy InGaAs exciton. The carrier lifetime is 4 nsec for the GaAs well and only 420 psec for the 3% In case. In fact, all of the InGaAs wells gave approximately the same lifetime [5] indicating low quantum efficiency. This is attributed to either thermal activation over the barriers or quenching due to defects at the GaAlAs barrier interface. With the known lifetimes, we can now calculate N_{sat} from the following measurements.

Figure 2(a) shows the absorption spectra as a function of I for the 3% Indium sample (S2). The heavy hole resonance at 854 nm is clearly seen and its absorption decreases strongly for input intensities up to 14 kW cm^{-2}. The variation of the peak absorption coefficient with N is shown in Figure 2(b) for two Indium concentrations, 0.03 and 0.11, respectively. It is clear that saturation of the exciton occurs at much lower carrier densities for the higher Indium concentration. We calculate N_{sat} to be 7 x 10^{17} and 1 x 10^{18} cm^{-3} for x = 0.03 and 0.11, respectively. Figure 3 shows how N_{sat} varies with Indium concentration for all four samples.

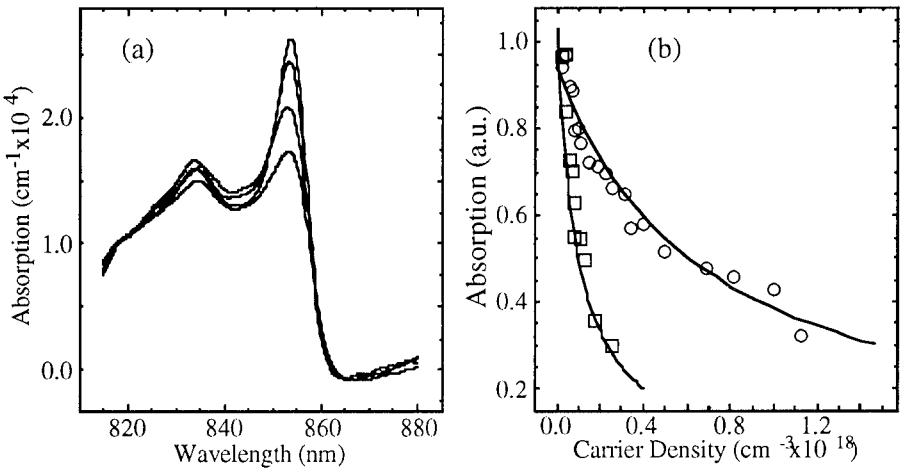


Figure 2 (a) Absorption spectra for S2 (3% In) for I = 0.009, 2.23, 9.5 and 14 kW cm^{-2}. (b) Peak exciton absorption coefficient as a function of carrier density for S2 (○, 3% In) and S3 (□, 11% In), normalised to each other.

It is clear that increasing the Indium content leads to a very significant decrease in N_{sat}, almost an order of magnitude for x = 0.11 compared to GaAs (x=0). The fact that improvements in N_{sat} of this size can be achieved by merely engineering the composition is a hopeful sign that all-optical nonlinearities may be eventually useful. The decrease in N_{sat} has been correlated with the reduction in the density of hole states [4] brought about by a decrease of the effective mass near k=0 due to compressive strain [7].

In Figure 3, the greatest change in N_{sat} is between x = 0.03 and 0.11. Sample S4 with x = 0.15 is similar to S3 and this can be explained by the presence of observed strain relaxation in this sample. Also included in Figure 3 are two other data points, one for an InGaAs asymmetric Fabry-Perot modulator structure (AFPM) [2] with a small reduction in N_{sat}, even though x = 0.1. This can also be explained by the fact that this structure has a much greater number of periods than S1-S4 and both well and barrier layer widths are much greater than the layer widths of S1 - S4. Consequently, this structure is much more sensitive to strain relaxation which does indeed occur. Finally, the work of Jin et al [3] is included in which x = 0.15, similar to S4. In their case, N_{sat} had a much higher value than in S4, again possibly due to strain relief. These results point to the importance of careful control of strain. Control becomes more critical as the indium content increases since the critical thickness before the onset of strain relief also decreases.

A reduction in the hole density of states should also reduce the absorption coefficient at the bandedge. In devices requiring a certain amount of absorption this could be compensated for by an increase in the number of layers. Because there is a critical thickness for strained structures, this is an undesirable feature. However, no evident reduction in the *exciton*

absorption strength was observed up to x = 0.15 showing that it is possible to get significant reduction in N_{sat} with little reduction in the excitonic value of α_o.

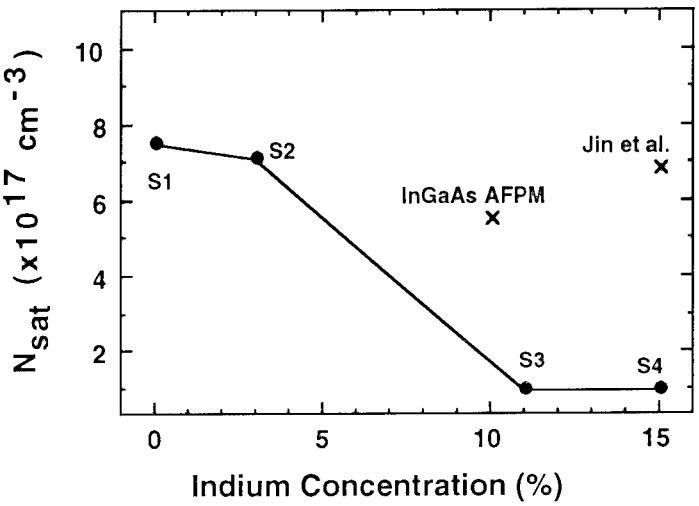


Figure 3 Plot of saturation carrier density against indium concentration for all measured samples.

To translate the effect into device usefulness the saturation *intensity* I_{sat} must also be similarly reduced. Since I_{sat} scales linearly with τ then it is necessary to be able to independently control τ and eliminate any nonradiative decay. Lifetimes in these InGaAs wells are much shorter than in GaAs and to date this seems to be a general problem [8] due to defects, low barrier heights or some other as yet unknown factor. It should be pointed out that these lifetimes are measured at densities much lower than typical in lasers where defects may be saturated, recovering high quantum efficiency.

In summary, this paper has outlined a path towards enhancing resonant optical nonlinearities by compressive strain in InGaAs/GaAlAs quantum wells. An order of magnitude reduction in saturation carrier density has already been observed with 11% Indium. This indicates several avenues for further research including control of carrier lifetimes to also reduce saturation intensities, investigation of structures with higher Indium content, control of strain relief, and other strained systems with similar bandgap curvature engineering possibilities.

References:

[1] Schmitt-Rink S, Chemla D S and Miller D A B 1989 *Adv in Phys* **38** 89.
[2] Moloney M H, Heffernan J F, Hegarty J, Grey R and Woodhead J 1993 *Appl Phys Lett* **63** 435.
[3] Jin R, Okada K, Khitrova G, Gibbs H M, Pereira M, Koch S W and Peyghambarian N 1992 *Appl Phys Lett* **61** 1745.
[4] Moloney M H, Hegarty J, Buydens L and Demeester P 1994 *Appl Phys Lett* **64** 997.
[5] Moloney M H, Hegarty J, Buydens L, Demeester P, Grey R and Woodhead J 1993 *Appl Phys Lett* **62** 3327.
[6] Buydens L, Demeester P, Yu Z Q and VanDaele P 1992 *Jnl Appl Phys* **71** 3249.
[7] Suemune I 1991 *IEEE Jnl Quantum Electron* **27** 1149.
[8] Unpublished.

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Dynamical behaviour of opto-optical logic switching devices employing n-i-p-i based smart pixels

P. Kiesel, M. Kneissl, P. Riel, K. Reingruber, K.H. Gulden, E. Greger, B. Knüpfer, X.X. Wu[×], J.S. Smith[×], and G.H. Döhler

Institut für Technische Physik, Universität Erlangen, Erwin-Rommel-Str. 1, 91058 Erlangen, Germany
[×]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA

Abstract. We report on experimental results on the dynamical behaviour of n-i-p-i based smart pixels, composed of photoconductive switches and electroabsorptive n-i-p-i modulators. For the photoconductive switch we present switching times of 1.9 ns at an optical power of 880 μW, corresponding to a switching energy of 1.7 pJ. The contrast of the electronic output signal is larger than 10^7 and a maximum dc gain exceeding 10^6 is achieved. For the opto-optical switching contrast ratios of 4:1 at 1.6 mW output power are shown with switching energies of 2.4 fJ/μm² (1.7 pJ). The opto-optical gain is tunable from 10-10^6.

1. Introduction

Recently we have demonstrated a new smart pixel concept [1] composing of a photoconductive switch with high electrical gain [2] and a high contrast electro-optical n-i-p-i modulator [3]. A schematic picture of our smart pixel concept is shown in figure 1. The switch is connected to an electro-optical modulator, whose output P_{out} changes between its 'on'- and 'off'-value depending on whether the switch is 'open' or 'closed'. One advantage of such a hybrid concept is that both elements (switch and modulator) can be optimized separately. Therefore the switches can have high responsivity and high photoconductive gain and the modulators can have high contrast and low insertion loss. Due to the high photoconductive gain of the switch only a small optical input P_{sw} is necessary to control the much larger output power of the n-i-p-i modulator. So pure opto-optical logic devices with high optical gain are realized requiring only a dc voltage for operation.

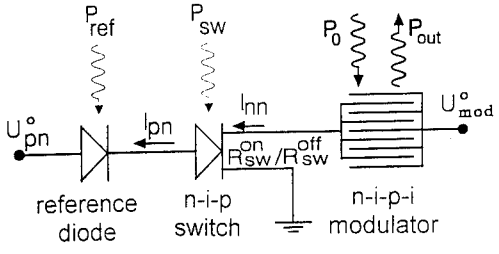


Fig. 1: Schematic diagram depicting our smart pixel concept. The smart pixel consists of an opto-electrical switch, a reference diode and a high contrast electro-optical n-i-p-i modulator. In this circuit the large modulator output power P_{out} is controlled by a small input power P_{sw} incident onto the switch.

2. Opto-electrical switching

The photoconductive switch is basically a pin-FET, integrated into a single device. The goal of designing this switch is to obtain the largest possible photoconductive response in an n-channel with a minimum number of photons incident onto the switch. Apart from optimizing the quantum efficiency of the pin diode this goal can be achieved by minimizing the capacitance of the switch. Therefore we use a sophisticated sample design with a small detection area which surrounds a spatially separated large absorption area. Figure 2 shows a SEM picture of the switch which has been monolithically integrated with the reference diode. The large absorption area is always depleted and thus doesn't contribute to the device capacitance [4]. Due to the 'giant ambipolar diffusion' constant of such structures [5], the diffusion of the photo-generated carriers from the inner absorption - to the surrounding detection area is fast enough (e.g. $\tau_{diff} < 50\,ps$, if the diameter of the absorption area is less than $20\,\mu m$). In figure 3 experimental results on the dynamical switching behaviour of the photoconductive switch are shown. With an optical power P_{sw} of $880\,\mu W$ the switching time of the n-layer current is 1.9 ns. This corresponds to a switching energy of 1.7 pJ (= 2.4 $fJ/\mu m^2$ referring to the area of this device ($\varnothing = 30\,\mu m$)). For this low switching time the opto-electrical gain is still 40. Figure 4 shows the dc-switching behaviour of the n-layer current I_{nn} vs. the optical power P_{sw} at the photoconductive switch. The curves correspond to different optical power levels P_{ref} incident onto the reference diode, shifting the switching point along the power axis. For the smallest reference power the contrast between the 'on' and 'off'-values of the n-layer current is larger than 10^7, corresponding to a high photoconductive gain of 10^6.

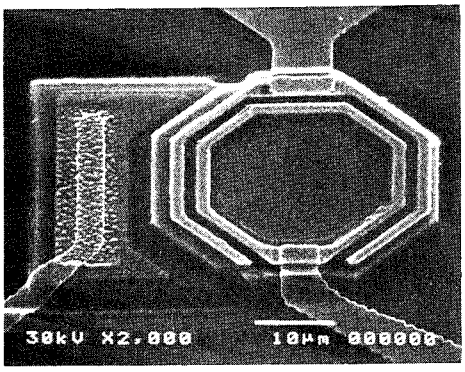


Fig. 2: SEM picture of a monolithically integrated electro-optical switch. The total diameter of the ring structure is 30 µm. The two metal rings are 2 µm thick and separated by 1 µm.

3. Opto-optical switching

By combining the photoconductive switch with a high contrast n-i-p-i modulator, as described in Ref. [3], we obtained opto-optical switching. Depending on whether the switch is in the high or low resistance state the voltage U^0_{mod} (see figure 1) drops either across the switch or the modulator. If the voltage drops across the switch, the modulator is in its transparent state and then the optical output signal P_{out} is high. Our hybrid smart pixel concept allows for opto-optical switching with high optical gain and low optical switching energies. It is possible to achieve various

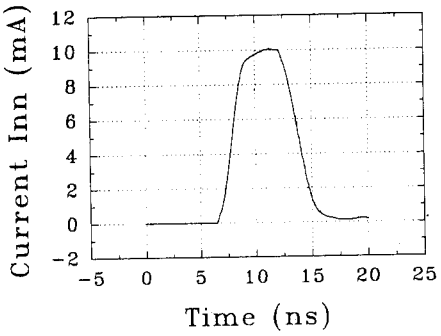


Fig. 3: Dynamical switching behaviour in the n-layer current I_{nn} of the photoconductive switch. In this case I_{nn} is switched on and off by pulses of the same optical power $P_{sw} = P_{ref} = 880\,\mu W$. The switching time τ_{sw} is 1.9 ns, corresponding to a switching energy of 1.7 pJ for this Ø = 30 µm device.

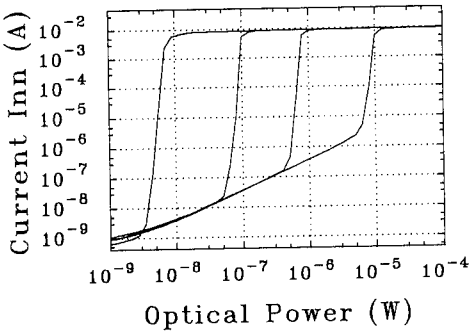


Fig. 4 : Double logarithmic plot of the n-layer current I_{nn} vs. the optical power at P_{sw} for different values of P_{ref}. If the optical power at the switch P_{sw} exceeds the reference power P_{ref} at the reference diode, the voltage distribution between switch and reference diode changes and simultanously the n-layer current of the switch rises from its depletion value up into the mA range.

logical functions. Here we show results of a NOR gate. Figure 5 shows switching behaviour of the optical output P_{out} of the n-i-p-i modulator vs. the optical intput power P_{sw} incident onto the switch. In this case the optical output signal changes from 400 µW to 1.6 mW corresponding to an on/off ratio of 4:1. As shown in figure 5 the switching point can be adjusted by changing the reference power P_{ref}. Depending on the switching power an opto-optical gain ranging from 10 to about 10^6 has been achieved. Figure 6 shows the dynamical behaviour of the smart pixel. With pulses of $P_{sw} = P_{ref} = 2\,\mu W$ incident either onto the switch or reference diode the optical output can be switched 'on' or 'off'. The corresponding switching time is $\approx 1\,\mu s$ leading to a switching energie of 2 pJ. The optical gain in this case is 750.

4. Conclusions

We have demonstrated low optical switching energies using a new n-i-p-i based smart pixel concept. Large values for the electrical gain ($> 10^7$) could be obtained with switching energies as low as 1.7 pJ corresponding to 2.6 $fJ/\mu m^2$. The switching power P_{sw}, the electrical and optical gain and the switching times are externally adjustable. We have demonstrated opto-electrical switching at switching times ≤ 1.9 ns for a switching power of 880 µW. By forming a smart pixel composed of an opto-electrical switch with photoconductive gain and an electroabsorptive n-i-p-i modulator we were able to obtain opto-optical switching with low energies (≤ 2 pJ). The optical gain is adjustable from 30 to 10^6, the contrast of the output signal was 4.5 : 1.

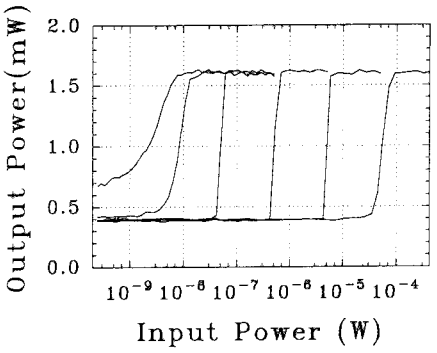


Fig. 5: Opto-optical switching of the smart pixel. For various optical power levels P_{ref} on the reference diode the diagram shows the output power of the n-i-p-i modulator controlled by the much lower input power on the photo-conductive switch.

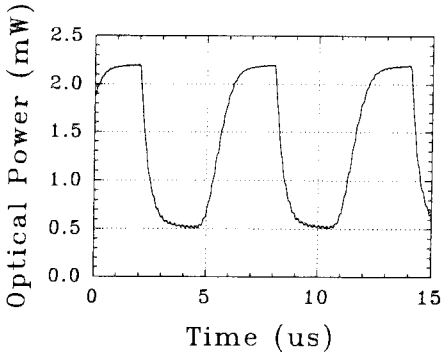


Fig. 6: Time dependence of the optical output signal P_{out} from the n-i-p-i modulator modulated by the input signal P_{sw} on the switch. For a switching power $P_{sw} = 2\mu W$ switching times of $\tau_{sw} \approx 1\,\mu s$ are achieved ($E_{sw} = 2\,pJ$). The optical gain is 750 and the on/off ratio of the modulated signal is 4.5:1.

Finally we note that the high saturation current of our switch would also be sufficient to operate another kind of "smart pixels", composed of our switch and a vertical cavity surface emitting laser structure (VCSEL).

This work has partly been supported by the German Ministry of Research and Technology (BMFT) under (TK 0567 0). The collaboration with Micronic laser Systems AB regarding the mask written by the Laser Pattern Generator LBG 15P is gratefully acknowledged.

References

[1] P. Kiesel, K.H. Gulden, A. Höfler, M. Kneissl, B. Knüpfer, P. Riel, X. Wu, J.S. Smith, and G.H. Döhler, "Bistable opto-optical switches with high gain based on n-i-p-i doping superlattices", Appl. Phys. Lett. **62**, 3288-3290 (1993)

[2] A. Höfler, K.H. Gulden, P. Kiesel, M. Kneissl, B. Knüpfer, P. Riel, G. Tränkle, G. Weimann, G. Tränkle, and G.H. Döhler, "Low Power (Bistable) Opto-electrical Threshold Switches with High gain Based on n-i-p-i Doping Superlattices", Appl. Phys. Lett. **62**, 3399-3401 (1993)

[3] P. Kiesel, K.H. Gulden, A. Höfler, M. Kneissl, B. Knüpfer, N. Linder, P. Riel, X. Wu, J.S. Smith, and G.H. Döhler, "High speed and high contrast electro-optical modulator based on n-i-p-i doping superlattices", Superlattices and Microstructures **13**, 21-24 (1993)

[4] P. Riel, P. Kiesel, M. Ennes, Th. Gabler, M. Kneissl, G. Böhm, G. Tränkle, G. Weimann, K.H. Gulden, X.X. Wu, J.S. Smith, and G.H. Döhler, "N-i-p-i-based high-speed detectors and bistable switches with gain", in: "Quantum Well and Superlattice Physics IV", SPIE Proc. **1675**, 242-254 (1992)

[5] K.H. Gulden, H. Lin, P. Kiesel, P. Riel, K.J. Ebeling, G.H. Döhler, "Giant Ambipolar Diffusion in n-i-p-i Structures", Phys. Rev. Lett. **66**, 373-376 (1991)

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical bistability above the bandgap energy in a GaAs Fabry-Perot

C. Bagnoud†, N. Collings

Institute of Microtechnology, University of Neuchâtel
CH-2000 Neuchâtel, Switzerland

Abstract
Above the bandgap energy, the combination of the absorptive and dispersive nonlinearities lead to a new type of bistable device. We present this device and determine the changes of the nonlinear refractive index and of the absorption with respect to the incident light power.

1. Introduction

We have presented earlier the bistable switching of an integrated nonlinear Fabry-Perot device (NLFP) with a low threshold (~ 1 mW) and high contrast [1]. The structure, an optical nonlinear spacer of approximately 2 μm length of intrinsic GaAs sandwiched between two AlGaAs dielectric mirrors, was designed for operations at 885 nm. Since the stopband of the dielectric mirrors is broad (115 nm for the back mirror), the NLFP can also be operated on a broad spectrum of wavelengths, even well below the wavelength of the bandgap ($\lambda_{gap} \approx 871$ nm).

2. Optical bistability above the gap energy

Below λ_{gap}, the Fabry-Perot resonance has a very low finesse at low incident power, since the back mirror is masked by the high absorption cavity. At higher incident power, a decrease in the absorption is induced, increasing finesse, and first a positive then a negative change of the refractive index will shift spectrally the resonance [2]. Figure 1a reproduces the changes in the reflectivity spectra due to the increase of the incident power. Optical bistability is observed at wavelength $\lambda = 853$ nm for 55 and 61 mW incident power (jump from high to low reflectivity and a quasi-infinite differential reflectivity). The finesse of the resonance increases from less than 2 at low power to more than 9 at 61 mW. The same behavior for the finesse was found in simulation (plane wave model, absorption and refractive index changes interpolated from experimental data [2]).

Figure 1b shows the bistable switching behavior above the gap ($\lambda \approx 853$ nm) at high incident power. A 55 mW power beam, to which a short pulse (100 ns) of 12 mW is added,

induces latched switching which lasts for more than 300 ns. The switch-off is caused by the positive refractive index change induced by heating. The measured contrast at switch-on is 3.5:1. Switch-on time is detector limited (10-20 ns).

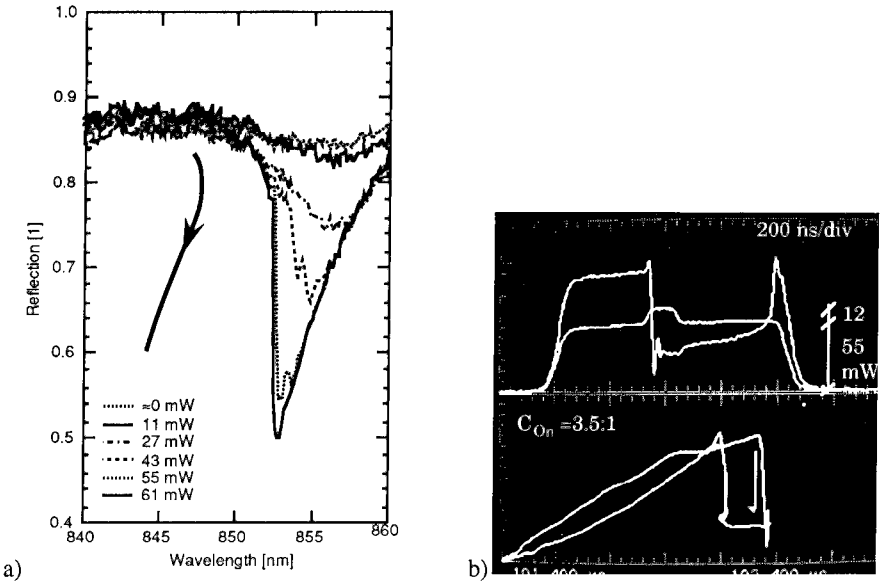


Fig. 1 a) Spectral reflectivity measurements of a Fabry-Perot resonance, for incident powers from 60 μW to 61 mW on a 6 μm spotsize. The change in the resonance wavelength is sketched at the left of the figure.
b) Bistable (latched) switching above the gap energy. Top: power of input and output beams vs. time. Bottom: output power vs. input power.

3. Absorption and refractive index change investigated above the energy of the gap

We deduced the absorption and the refractive index change variations from nonlinear spectral reflectivity measurements (SRMs) (Fig. 2). The optical constants vary with respect to different incident powers of the laser beam (0 to 15 mW) and with the different operating wavelengths (between 847 and 867 nm in our measurements). From fitting the formula for the reflectance of the Fabry-Perot with absorption to the SRMs, we determined the minimum of the low finesse resonances and their corresponding absorption. The nonlinear refractive index change is deduced from the wavelength shift of the resonance [3]. The internal light intensity in the low finesse cavity is determined from the incident intensity I_i, the reflectivity of the device R, and the absorption following the relation

$$I = I_i \frac{1-R}{\alpha L},$$

where L is the length of spacer. This relation is obtained from the equation for the reflectivity of a Fabry-Perot including the spatial averaging of the intensity in the cavity [4].

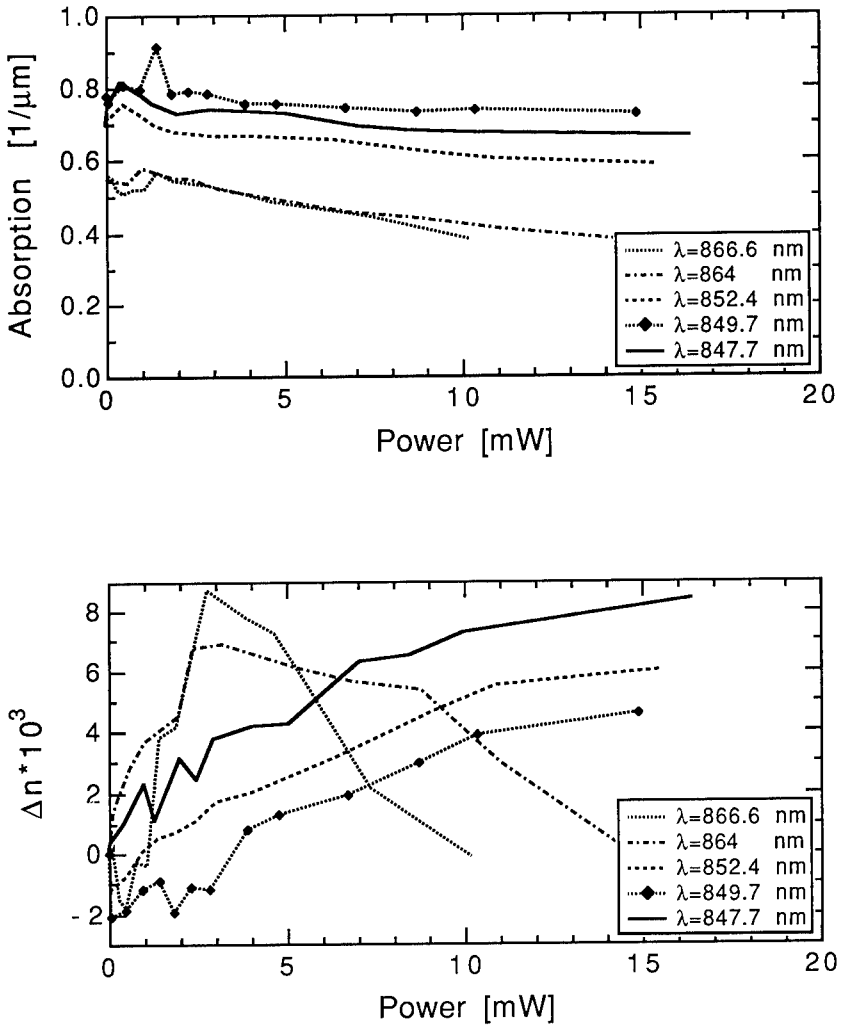


Fig. 2 Changes in absorption (top) and in the refractive index (bottom) with respect to the incident power on a 12 μm spotsize. The points are determined from nonlinear SRMs at photon energies above the gap on the sample #337. The number of points indicated for λ = 849.7 nm is the number of SRMs used to get the curve.

Phenomenological coefficients $\tilde{\alpha}_2$ and $\tilde{n}_2$ are deduced, characterizing the linear absorption and refractive index change at different wavelengths with respect to the internal intensity. The values of the phenomenological coefficients are $\tilde{\alpha}_2$ = 322 cm/kW (λ = 853 nm) and $\tilde{n}_2 \approx 2 \cdot 10^{-3}$ cm^2/kW (λ between 864 and 867 nm, incident power > 2.5 mW). $\tilde{n}_2$ is 30 times higher than the phenomenological Kerr coefficient $n_2 \approx -7 \cdot 10^{-5}$ cm^2/kW determined at a

photon energy below the gap. This is related to the 30 times higher absorption of $7 \cdot 10^3$ cm^{-1} measured above the gap in comparison to $\alpha = 250$ cm^{-1} below the gap [4].

4. Discussion

The above gap bistability requires higher powers (61 mW) than the below gap bistability (1 mW) due to the requirements that the absorption be bleached and that the dispersive refractive index change be sufficiently negative. However, there may be an application for such a type of bistability in cascaded systems where an "upconverter" etalon is required [5]. Two-wavelength logic-gate etalons are operated with a signal wavelength which is longer than the pump (bias) wavelength. For example, the signal beam can be on-resonance and be more efficiently absorbed than the pump wavelength which is detuned from resonance towards shorter wavelengths. The upconverter would work in the opposite sense with signal wavelengths shorter than the pump wavelength and thus allow cascadability. We have here shown (as is predicted from the curves in [2]) that it is possible to bleach the absorption above gap over a range of wavelengths. Therefore, it is possible to chose the pump wavelength detuned from resonance as before and then position the signal wavelength shorter than this. In order that the pump wavelength of the upconverter is the same as the signal wavelength of the logic-gate etalon, the bandgaps of the material chosen in the two cases would have to be different (as suggested in [5]).

† Present affiliation of C. Bagnoud is the Dept. of Electrical and Computer Engineering, UCSD, La Jolla, California 92093-0407, USA.

References

[1] Acklin B et al 1991 *OSA Proceedings on Photonic Switching, vol. 8*, H. Scott Hinton and Joseph W. Goodman, eds., (Optical Society of America, Washington, DC, 1991) 231-234
[2] Lee Y H et al 1986 *Phys. Rev. Lett.* **57**, 2446-2449
[3] Sfez B G et al 1990 *Appl. Phys. Lett.* **57**, 324-326
[4] Bagnoud C 1993 *The study of bistable GaAs Fabry-Perot devices for operation in array format* (Ph.D. thesis, University of Neuchâtel, Switzerland)
[5] Richardson D, Gibbs H M, and Koch S W 1991 *IEEE J.Quant. Elect.* **27**, 804-808

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

An optical set-reset flipflop semiconductor laser with two mutually complementary outputs

Masanobu Watanabe, Seiji Mukai, and Hiroyoshi Yajima

Optical Information Section, Electrotechnical Laboratory,
Umezono, Tsukuba, Ibaraki, 305 Japan

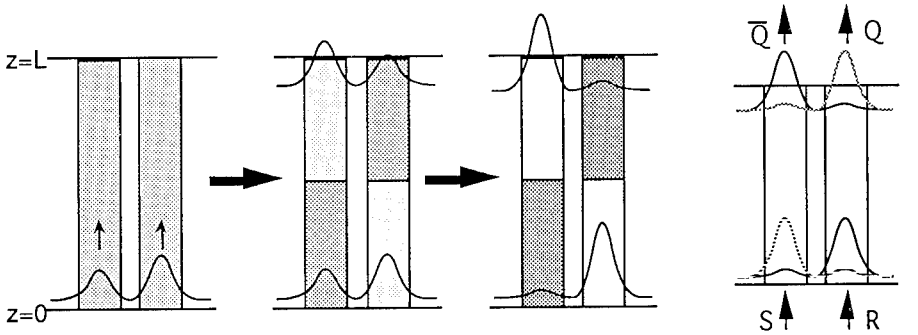
Abstract

Theory and experiment on crosscoupled-mode bistability in a twin-stripe laser is reported. The laser has two output ports complementary to each other, which is analogous to a set-reset flipflop in electronics. Limitation on the cavity length is also discussed.

1. Introduction

A crosscoupled field which diagonally crosses from one stripe to the other in a twin-stripe laser was first experimentally observed by White and Carroll in 1983 [1] and the mechanism for its generation was theoretically explained by Watanabe et al in 1990 [2]. The theory showed that such an asymmetric field with its pattern at a facet being the mirror image of that at the other facet can be obtained with completely symmetric structure and symmetric current injection. Consequently, the laser should show bistability between two crosscoupled modes. The optical power distribution of the bistable states is similar to the voltage distribution of a set-reset flipflop in electronics [3], and hence, useful for optical switching and logic operation.

Here, theory and experiment on crosscoupled-mode bistability in a twin-stripe laser is reported. Switching indicating bistability between two crosscoupled modes were observed.



(a) Fluctuation leads to slightly asymmetric pattern at a facet

(b) After one-way propagation

(c) Final stable state

(d) Set-reset flip-flop

Fig. 1 Process to the crosscoupled mode operation

2. Theory

Fig.1 illustrates top views of a twin-stripe laser to show how a cross-coupled mode [1] builds up [2-4]. Assume that the light power and the carriers are mainly in each of the two waveguides under the stripes, the current is *uniformly* injected into the stripes, and the cavity length is near to the coupling length of the twin waveguide. The explanation starts with symmetric light pattern at both facets. Suppose a fluctuation which leads to a slightly asymmetric pattern near a facet (z=0) as shown in Fig.1(a). This field pattern changes during one-way propagation to the other facet (z=L), where it becomes nearly the mirror image (left and right is reversed) of the light pattern at z=0, as shown in Fig.1(b). During this propagation, this light consumes more carriers at the right-lower and left-upper regions of the twin waveguide than at the other regions. We should now examine how the resultant diagonal carrier distribution reacts on the light field in the next step.

Intuitively, one may expect that the carrier distribution will enhance the light power at the regions with high carrier density, which makes the light pattern return to the symmetric shape because of negative feedback. On the contrary, however, it was theoretically shown that the diagonal carrier distribution enhances the light power at the regions with *low* carrier density and makes the original light pattern more asymmetric, if the cavity length is shorter (longer) than the coupling length for twin-stripe lasers with low (high) inter-stripe gain [2-4]. This happens because of both lateral and longitudinal resonance, and the light confinement difference between the resonant lateral modes.

An example of light power distribution of resonant lateral modes are shown in Fig.2 which were calculated for the diagonal carrier distribution shown in Fig.1(b). Here, the interstripe gain was set to be zero, the cavity length L was set to be 95 % of coupling length, and the carrier density difference between two regions was set to be 2.4 % of the average carrier density. In this example, the mode in Fig.2(a) has the higher gain than the other because of the higher light confinement, in spite that the highest peak is in the low carrier density side. Therefore, the diagonal carrier distribution supports the mode with its highest peak in the low carrier density side, which leads to positive feedback.

The light and the carrier distributions enhance the asymmetries of each other due to the positive feedback and finally, both have substantially asymmetric patterns as illustrated in Fig.1(c). If the original light fluctuation is such that the left peak (instead of the right peak as shown in Fig.1(a)) becomes larger at z=0, then another stable state with mirror images of the light and the carrier distributions shown in Fig.1(c) is obtained.

Thus, there are two stable states drawn in Fig.1(d) in solid and broken curves, respectively. The laser can be switched from a crosscoupled state to the other by light injections. The operation is similar to a set-reset flip-flop in electronics particularly in that the laser has two mutually complementary outputs, while bistable semiconductor lasers reported so far have usually only one output and hence require an additional inverter to get the complementary output. Thus, twin-stripe lasers in crosscoupled-mode operation should be useful for optical switching and logic operation.

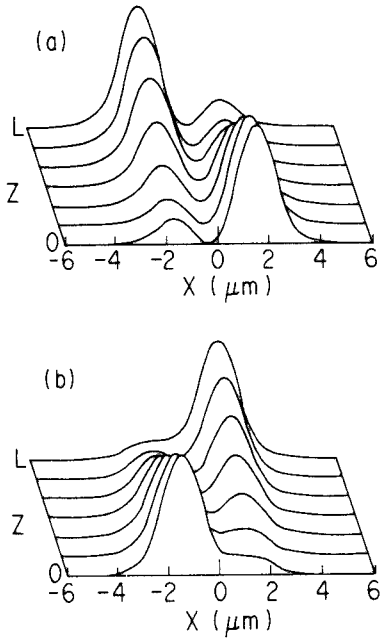


Fig.2 Power distribution of resonant lateral modes

3. Experiment

Fig.3 shows the cross section of the lasers made for measurement which has a similar structure with that in [5]. It has an ordinary double heterostructure with a 0.1 μm thick GaAs active layer sandwiched by two $Al_{0.35}Ga_{0.65}As$ cladding layers. Two 2-μm-wide stripe anodes with 4 μm spacing were formed on 4-μm-wide mesas made by wet etching. The wafer was cleaved to make lasers with typical cavity length of 500 μm which is estimated to be near to the coupling length.

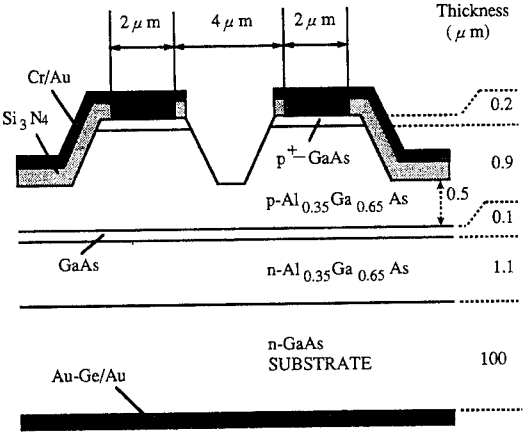


Fig.3 Cross section of the measured laser

Fig.4(a) and (b) show near-field patterns measured with the left and right currents fixed at 90mA and 95mA, respectively [6]. They show the two crosscoupled modes whose patterns are the mirror images of each other. When we measured the field patterns with pulse drive, Fig.1(a) was obtained at some pulses and (b) was obtained at the other pulses. This clear switching with nominally identical pulses indicates bistability between two crosscoupled modes.

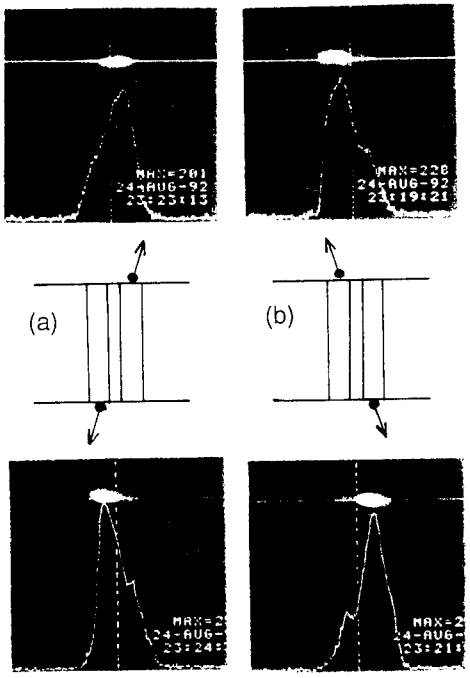


Fig.4 Measured near-field patterns

4. Cavity length limitation

The cavity length should be near to the coupling length of the twin waveguide to give a highly asymmetric light patterns. To reduce the laser length for crosscoupled mode generation, the coupling between the two waveguides should become stronger. The strongest limit is that the two waveguides are combined into one. Therefore, the possibly shortest laser length can be estimated by the propagation constant difference between the fundamental and the first-order lateral modes of a single waveguide.

For a simple slab waveguide, the larger the permittivity difference between the core and the clad, the shorter the minimum coupling length and the thinner the required core thickness. For GaAs core and $Al_{0.35}Ga_{0.65}As$ clad whose permittivities are 13.1 and 11.4, respectively, the minimum coupling length is 2.7 μm while the required core thickness is 0.2 μm. Here, the permittivity reduction of around 0.1 in the active layer was considered assuming the carrier density of $2\times10^{18}/cm^3$. In this very limit condition, however, the crosscoupled mode does not have a clear asymmetric pattern, as explained below.

To see how far we should stay away from the limit condition, a simple calculation was made for looking at the supermode patterns and the coupling length for various coupling strength. Fig.5 shows some of the results for simple GaAs - $Al_{0.35}Ga_{0.65}As$ twin slab waveguides where the core thickness was fixed to 0.1 μm (0.2 μm when they are combined). The coupling was changed by the waveguide separation. For large separation (>0.4μm), the crosscoupled mode can have a clear asymmetric pattern with its power predominantly at one side because the two supermodes which compose the crosscoupled mode have similar power distribution to each other and can well cancel each other at the other side. At the limit condition with zero waveguide separation, the crosscoupled mode does not have a clear asymmetric pattern, because the first-order mode is at the transition point to cut-off and its power spreads to outside much more than the fundamental mode. Consequently, around 10 μm cavity length is required for clear spatial switching due to crosscoupled mode operation in a GaAs - $Al_{0.35}Ga_{0.65}As$ twin slab waveguide.

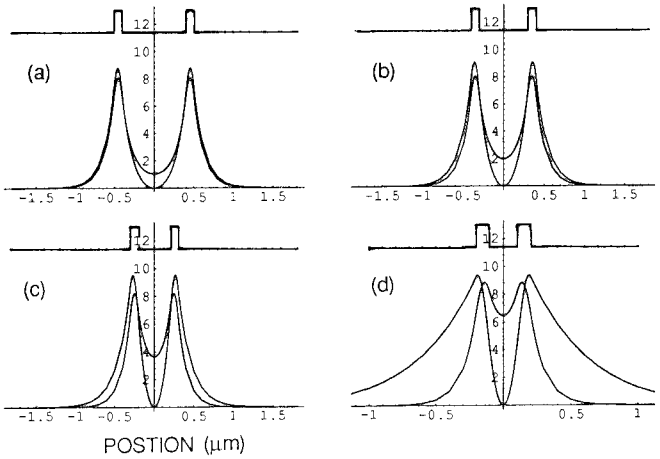


Fig.5 Power distribution of the fundamental and the first-order super modes for various waveguide separation. The permittivity distribution is also shown at the top of each figure. The separation, d, and the resultant coupling length, Lc, are: (a) d=0.8μm, Lc=86μm, (b) d=0.6μm, Lc=39μm, (c) d=0.4μm, Lc=17μm, (d) d=0.2μm, Lc=7μm.

5. Summary

Mechanism for the crosscoupled mode generation was theoretically explained and two crosscoupled modes were observed experimentally with nominally identical current pulses. This indicates bistability which is useful for set-reset flipflop operation with two mutually complementary outputs. The required cavity length to obtain clear spatial switching was estimated to be around 10 μm for a GaAs-$Al_{0.35}Ga_{0.65}As$ twin slab waveguide.

References

[1] I.H. White and J.E. Carroll, Electron. Lett., **19**, 337 (1983).
[2] M. Watanabe, I.H. White, and J.E. Carroll, IEEE J. Quantum Electron., **QE-26**, 1942 (1990).
[3] M. Watanabe, I.H. White, and J.E. Carroll, ibid., **QE-28**, 395 (1992).
[4] M. Watanabe, S. Mukai, K. Matsubara, and H. Yajima, ibid., **QE-29**, 2855 (1993).
[5] M. Watanabe, S. Mukai, H.Itoh, and H. Yajima, J. Appl. Phys., **68**, 2599 (1990).
[6] M. Watanabe, H. Fujiki, S. Mukai, M. Ogura, H. Yajima, K Shimoyama, and H Gotoh, to be published in IEEE J. Quantum Electron., **QE-30**, no.10 (1994).

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Resonant-tunneling Triangular-barrier Optoelectronic Switch (R-TOPS) for Optical Logic

H. Sakata, K. Utaka and Y. Matsushima

KDD R&D Laboratories
2-1-15 Ohara, Kamifukuoka-shi, Saitama 356, Japan
Phone:+81-492-66-7831 FAX:+81-492-66-7516

Abstract. We report a novel optical bistable device, which consists of a double-barrier resonant-tunneling diode and a triangular-barrier phototransistor. Clear negative differential resistance and optical bistability with high contrast and high sensitivity are demonstrated.

1.Introduction

An optical bistable device is a key for optical computing, and a lot of devices have been studied so far for this purpose. As a photodetecting element in the devices several kinds of structures such as a conventional heterostructure phototransistor (HPT) have been studied [1]. As another candidate, a triangular-barrier phototransistor (TBP) with a very thin gate layer, which has been developed in GaAs / AlGaAs [2], is very promising from a view point of high speed and high sensitivity. On the other hand, a resonant-tunneling diode (RTD) has a potential to give rise to unique functions attributed to its N-shaped negative differential resistance (NDR) and high speed operation [3]. For these reasons, the integration of a double-barrier RTD (DB-RTD) and a TBP seems very interesting to make a novel optical bistable device.

On the other hand, 1 μm-wavelength range operation of these devices is desirable for the matching to optical communication. In order to fabricate the devices with InP-related materials for this purpose, a gas source molecular beam epitaxy (GSMBE) should be a promising technique [4], because sharp interfaces and a good δ-doped layer are attained, which are inevitable for the device.

In this paper, we propose Resonant-tunneling Triangular-barrier Optoelectronic Switch (R-TOPS) grown by GSMBE, and the fabrication and the fundamental experimental results are reported at 1μm-wavelength range operation.

2.Device structure

Figure 1 shows a band diagram of the device fabricated. It was composed of InGaAs / (In)AlAs for 1.5 μm wavelength range operation. The TBP had n^+-i-δp^+-i-n^+ structure, and an i-InAlAs region was designated as a source, a δp^+-InGaAs region as a gate, and an i-InGaAs

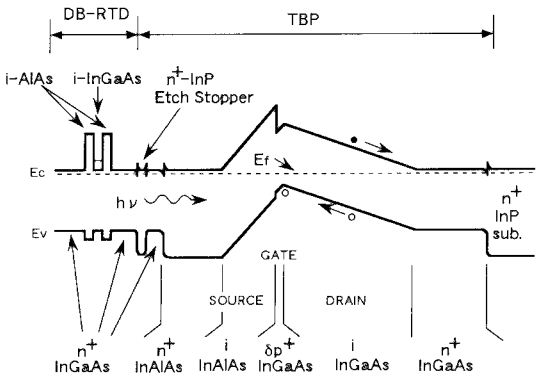


Fig.1. Band Diagram

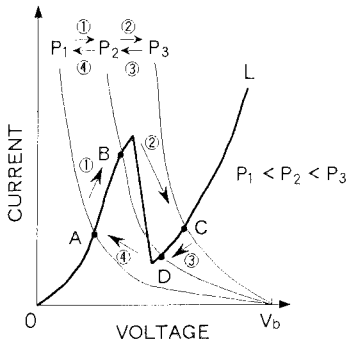


Fig.2.Schematic load-line graph for R-TOPS

region as a drain. When an input-light is incident to the device with the drain positively biased and is absorbed in the drain, the generated holes move to the gate layer and lower the potential barrier of the gate. This makes majority carriers, electrons in this case, flow over the potential barrier from the source to the drain. The majority carrier flow is due to thermionic emission because of the ultrathin gate layer, so the current versus voltage characteristic shows an exponential-like curve, and high optical sensitivity and high speed operation are expected [2]. As for the DB-RTD a well layer and barrier layers were i-InGaAs and i-AlAs, respectively, and had i-InGaAs spacer layers outside the barrier layers. It was reported that this structure shows a large peak-to-valley current ratio at room temperature because of the very large barrier height [5]. The TBP and the DB-RTD were integrated with an n+-InP etch stop layer in-between. Since the DB-RTD has N-shaped NDR characteristics, and the TBP has exponential-like dependence of current upon voltage varied by input-light power, it is predicted that the R-TOPS shows a new class of bistability. And the output-current change is expected to be large because the TBP has a small differential resistance. Figure 2 shows a schematic load-line graph for the R-TOPS. A thick curve is the characteristic of the DB-RTD, and three thin curves are the ones of the TBP at different input-light powers P_1, P_2 and P_3, which show exponential-like dependence of current upon voltage due to thermionic emission. By changing the input-light power, the intersections vary as shown in Fig.2.

The device was grown on a (100) n+-InP substrate by GS-MBE, in which 100% AsH_3 and PH_3 were used for group V sources, and silicon and beryllium were used as n-type and p-type dopants, respectively. The growth temperature was 530°C for InGaAs and (In)AlAs. The device structure consisted of n+-$In_{0.53}Ga_{0.47}As$ (0.098μm, 2.3×10^{18} cm^{-3}), i-$In_{0.53}Ga_{0.47}As$ (0.79μm), δp+-$In_{0.53}Ga_{0.47}As$ (59Å, 3.7×10^{18} cm^{-3}), i-$In_{0.52}Al_{0.48}As$ (0.041μm), n+-$In_{0.52}Al_{0.48}As$ (0.10μm, 2.2×10^{18} cm^{-3}), n+-$In_{0.53}Ga_{0.47}As$ (0.034μm, 2.3×10^{18} cm^{-3}), n+-InP (56Å, 4.2×10^{18} cm^{-3}), n+-$In_{0.53}Ga_{0.47}As$ (0.099μm, 2.3×10^{18} cm^{-3}), i-$In_{0.53}Ga_{0.47}As$ (12Å), i-AlAs (30Å), i-$In_{0.53}Ga_{0.47}As$ (32Å), i-AlAs (30Å), i-$In_{0.53}Ga_{0.47}As$ (12Å) and n+-$In_{0.53}Ga_{0.47}As$ (0.099μm, 2.3×10^{18} cm^{-3}). The device was etched into a mesa with 60 μmϕ. Contact to the top layer was formed by evaporating Au / Sn in the shape of an open ring structure using lift-off technique with SiN_x passivation. Input-light was illuminated to the top of the device with a lensed fiber.

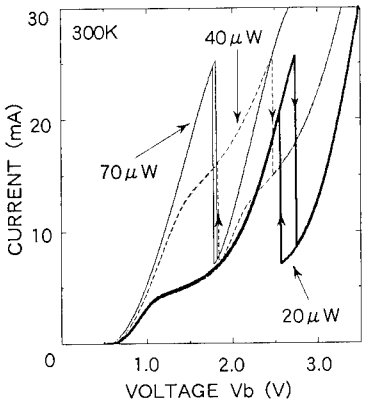


Fig.3.I-V characteristics at different input-light powers

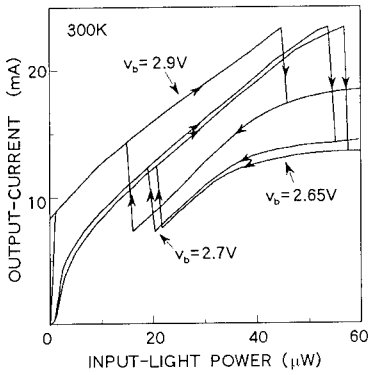


Fig.4.Output-current vs. input-light power at different bias voltages

3.Characteristics

Experimental I-V characteristics at different input-light powers at room temperature are shown in Fig.3. The substrate side was normally positively biased at V_b, and the wavelength of input-light was 1.55 μm. Clear N-shaped NDR with bistability ('electronic bistability') was observed, and a peak-to-valley current ratio and a peak current density were 3.6 and 910 A/cm^2 , respectively. The peak and the valley currents were determined by those of the RTD, so these were almost independent of the input light power. On the other hand, as the input-light power increased, the peaks moved to lower voltage side, which was due to the decrease of series resistance of the TBP. At the same time , hysteresis widths were also varied. This was attributed to the slope characteristics of the TBP depending on the input-light power. These properties show that both TBP and DB-RTD worked well as expected, even when they were integrated. The bistability in the relation between input-light power and output-current ('optoelectronic bistability') at room temperature was obtained as shown in Fig.4. It operated at the input-light power as small as 50μW or even smaller, and the output-current change was about 10mA, which was large enough to drive other optical devices such as a laser diode. Since the hysteresis width was about 30μW, switching could be attained by the input-light swing of no more than 20μW with an appropriate bias light. We also realized 'optical bistable' operation by connecting the R-TOPS and a laser diode as shown in Fig.5. A high contrast ratio about 13dB and an intrinsic optical gain of about 6dB from an input fiber to an output fiber were obtained. These characteristics show the possibility of optical XOR logic and memory operation. In practice, dynamic property was measured as shown in Fig.6. An upper trace is the input-light, and a lower trace is the output-current. The input-light was modulated by both set and reset pulses with a bias input-light. The repetition rate was 500 kHz, and the pulse width was 150ns. Due to the utilization of N-shaped NDR, the set operation was done by the decrease of input-light power, and the reset one by the reverse. Good dynamic memory operation was obtained with a switching energy of as small as 4.5pJ, corresponding to 1.6 fJ/μm^2. It is possible to improve the switching speed of this device by optimizing device parameters and measurement circuit.

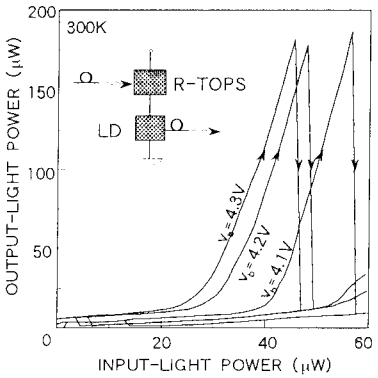


Fig.5.Optical bistable characteristics at different bias voltages

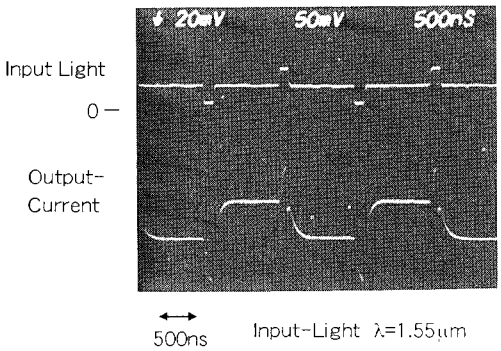


Figure 6. Dynamic memory operation

4.Conclusion

In conclusion, we fabricated a novel R-TOPS by GSMBE, and clear N-shaped NDR and bistability were obtained. And dynamic memory operation was also demonstrated. By introducing light emitting or modulating elements, we can realize all optical bistable device. The R-TOPS is expected to play an important role for optical computing.

Acknowledgment

The authors would like to express their thanks to Drs. Y. Urano, K. Sakai and Y. Mimura for encouragement.

References

[1] Y.Kawamura, H. Asai, S.Matsuo, and C. Amano: ' InGaAs-InAlAs Multiple Quantum Well Optical Bistable Devices Using the Resonant Tunneling Effect', IEEE J. Quantum Electron., 1992, **QE-28** , pp.308-314

[2] C.Y.Chen, A.Y. Cho, P. A. Garbinski, C. G. Bethea, and B. F. Levine : 'Modulated barrier photodiode: A new majority-carrier photodetector', Appl. Phys. Lett. 1981,**39** ,pp. 340-342

[3] F.Capasso, S.Sen, F.Beltram, L. M. Lunardi, A. S. Vengurlekar, P. R. Smith, N. J. Shah, R. J. Malik, and A. Y. Cho:'Quantum Functional Devices: Resonant-Tunneling Transistors, Circuits with Reduced Complexity, and Multiple-Valued Logic", IEEE Trans. Electron. Devices, 1989, **ED-36**, pp.2065-2082

[4] L. Goldstein: 'Optoelectronic Devices by GSMBE', J. Crystal Growth, 1990, **105**, pp.93-96

[5] T.Inata, S. Muto, Y. Nakata, S. Sasa, T. Fujii, and S. Hiyamizu: 'A Pseudomorphic $In_{0.53}Ga_{0.47}AlAs$ Resonant Tunneling Barrier with a Peak-to-Valley Current Ratio of 14 at Room Temperature' Jpn. J. Appl. Phys., 1987, **26**, pp.L1332 -L1334

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Triangular-barrier Optoelectronic Switch (TOPS) for Optical Logic

H. Sakata, K. Utaka and Y. Matsushima

KDD R&D Laboratories
2-1-15 Ohara, Kamifukuoka-shi, Saitama 356, Japan
Phone:+81-492-66-7831 FAX:+81-492-66-7516

Abstract. We report a novel optical functional device, Triangular-barrier Optoelectronic Switch (TOPS), which consists of a triangular-barrier phototransistor with avalanche multiplication. Clear differential gain, bistability and latch characteristics with high sensitivity and high gain were obtained by only changing the bias voltages.

1.Introduction

In order to realize optical computing, various kinds of functions are needed for optical devices such as logic, memory and so on. For these purposes, an optoelectronic switch showing nonlinear optical response is quite promising, and several kinds of devices have been studied so far. Among those devices, the p-n-p-n double-heterostructure optoelectronic switch (DOES) exhibited optically controlled switching [1,2]. The heterojunction bipolar phototransistor (HPT), which consisted of n^+-p^+-i-n^+ structure, also showed similar characteristics [3]. These devices utilized avalanche breakdown in the reverse-biased region for switching. And once the breakdown occurred, it was hardly possible to reset it to the off-state optically, unless a relatively high impedance resistor of about 1kΩ was connected externally, which was a disadvantage for high speed operation [2]. This was because the on-state I-V characteristics of these devices were almost independent of input-light due to inherent properties of p-n junctions. If some structure whose I-V characteristics depend on input-light power can be utilized, switching may be fully controlled by light, in other words, optical reset is also expected. One of the possible candidates for such a structure is n^+-i-δp^+-i-n^+ structure.

This structure as a triangular-barrier phototransistor (TBP) with a very thin gate layer has been developed in GaAs / AlGaAs, and is very promising from a view point of high speed and high sensitivity [4]. But only conventional functions as a phototransistor were reported. In order to apply a TBP to an optoelectronic switch, it is important to introduce avalanche multiplication into a TBP . In this paper we propose Triangular-barrier Optoelectronic Switch (TOPS) for 1μm wavelength range operation. It consisted of InGaAs / InAlAs grown by gas source molecular beam epitaxy (GSMBE), which is a promising technique for sharp interfaces and a good δ-doped layer [5]. We successfully demonstrated flexible optical

functions of the TOPS associated with avalanche multiplication. To our knowledge, this is the first report on optically controllable optoelectronic switch by a TBP.

2.Device structure

Figure 1 shows a band diagram of the TOPS biased. It had n^+-i-δp^+-i-n^+ structure, and was composed of InGaAs / InAlAs for 1.5 μm wavelength range operation. The device structure consisted of n^+-$In_{0.53}Ga_{0.47}As$ (0.10μm, 2.0×10^{18} cm^{-3}), i-$In_{0.53}Ga_{0.47}As$ (0.84μm), δp^+-$In_{0.53}Ga_{0.47}As$ (82Å, 1.2×10^{18} cm^{-3}), i-$In_{0.52}Al_{0.48}As$ (0.041μm), n^+-$In_{0.52}Al_{0.48}As$ (0.10μm, 2.3×10^{18} cm^{-3}) and n^+-$In_{0.53}Ga_{0.47}As$ cap layer (0.034μm, 2.0×10^{18} cm^{-3}). An i-InAlAs region, a δp^+-InGaAs region and an i-InGaAs region were designated as a source, a gate and a drain, respectively. The device was grown on a (100) n^+-InP substrate by GSMBE, in which 100% AsH_3 was used for a group V source, and silicon and beryllium were used as n-type and p-type dopants, respectively. The growth temperature was 500°C, and the typical growth rate was 0.4μm/h. The device was formed into a mesa structure with 60 μmϕ . Contact to the top layer was achieved by evaporating Au / Sn to form an open ring structure using lift-off technique with SiN_x passivation. An input -light was introduced to the top of the device with a lensed fiber, and the wavelength of the input-light was 1.55 μm.

If the input-light is incident to the device when the drain is biased positively, electrons and holes are generated in the drain, and the holes move to the gate layer and accumulate in it. Resultant lowering of the potential barrier of the gate increases the majority carrier flow of electrons over the potential barrier from the source to the drain regions due to thermionic emission. So the current vs. voltage characteristics show exponential-like curves. The avalanche multiplication can occur in the drain when a device structure is optimized and an electric field in the drain is large enough. Multiplied holes further lower the potential barrier, so the majority carriers flow moreover. This positive feedback phenomenon between the hole generation and the electron flow through the avalanche multiplication gives rise to optoelectronic switching. Since the output current of the TBP is quite dependent on the input-light power, we can control the switching by the input-light change. It should be noted that we can stop the avalanche multiplication by decreasing the input-light power. And various functions can be chosen by only changing bias voltages.

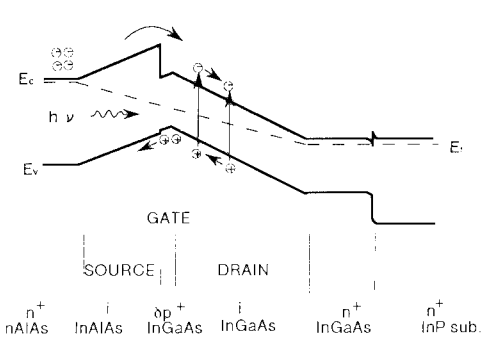


Fig.1. Band Diagram with a bias voltage.

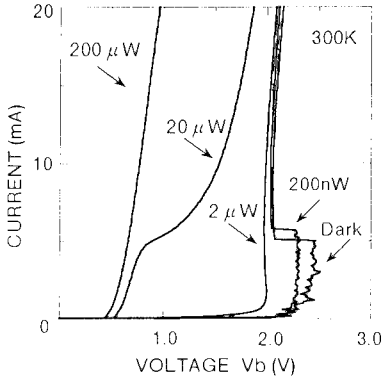
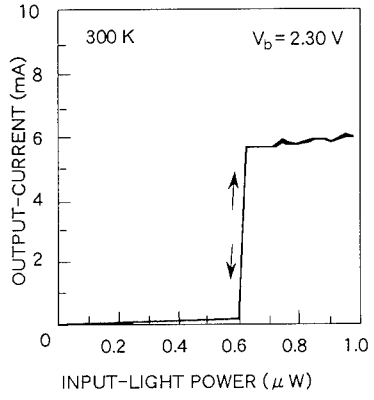


Fig.2. I-V characteristics at different input-light powers (λ=1.55μm)

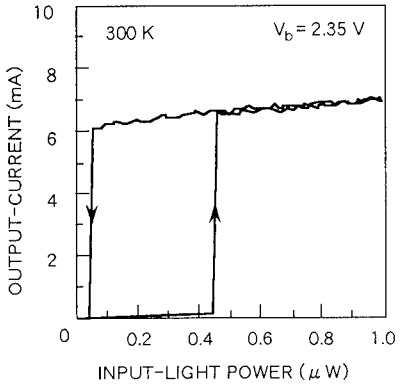
3.Characteristics

Experimental I-V characteristics at different input-light powers observed at room temperature are shown in Fig.2. Ordinary I-V characteristics of a TBP were obtained as the input-light power was more than 2μW. However, a significant S-shaped negative-differential-resistance (NDR) was observed as the input-light power was less than 2μW. These NDR characteristics had almost the same traces for the cases when the input-light power increased and decreased.

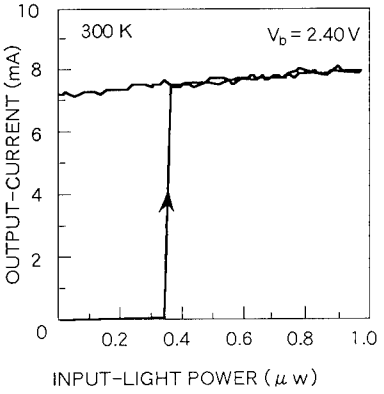
The input-light power versus output-current characteristics by changing the bias voltages at room temperature were shown in Fig.3 (a)~(c). We obtained three types of functional operations by just changing the bias voltages V_b without any additional external resistance. At a bias voltage of 2.30 V, differential gain characteristic was observed as shown in Fig.3 (a). It is noted that the step was very sharp at the input-light power of as small as 600nW, and the output current was about 6mA. A corresponding optical gain was estimated to the 7400. We can apply this mode for AND or OR logic. When the bias voltage increased to be 2.35 V, it was changed to bistable characteristic as shown in Fig.3 (b). The turn-on



(a) Differential gain characteristic.



(b) Bistable characteristic.



(c) Latch characteristic.

Fig.3. Input-light power vs. output-current characteristics by changing bias voltages V_b

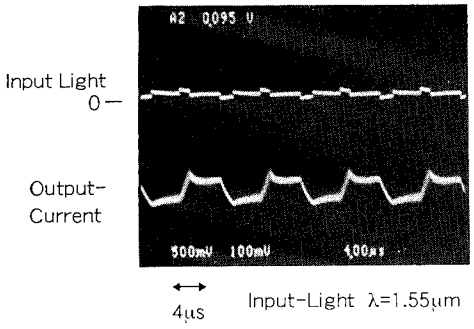


Fig.4. Dynamic memory operation by bistable characteristic

input-light power was decreased to be 450nW, so the estimated optical gain was 12000. A hysteresis width was varied depending on the bias voltage. So this mode is good for memory operation. At a larger bias voltage of 2.40 V, latch characteristic was attained as shown in Fig.3 (c). An optical gain in this case was as large as 17000 at 350 nW turn-on input-light power, so it can work as switch operation. We can reset the holding state by decreasing the bias voltage in this latch mode.

Since an output current density of the TOPS was as large as 200 A/cm^2, we can realize all optical functional devices such as an optical logic device, an optical memory and so on by integrating a light-emitting or a light-modulating devices with it.

Preliminary dynamic memory operation based on the bistable characteristic was obtained at the repetition rate of 100kHz as shown in Fig.4. An upper trace is the input-light, and a lower trace is the output-current. The input-light was modulated by both set and reset pulses with a bias input-light. And it is possible to improve the switching speed of this device by optimizing device parameters and measurement circuit.

4.Conclusion

We demonstrated a novel optical functional device, Triangular-barrier Optoelectronic Switch (TOPS), which consisted of a triangular-barrier phototransistor with avalanche multiplication. Clear differential gain, bistability and latch characteristics were obtained by only changing the bias voltages, while the input-light power was less than 600 nW and the optical gain was more than 7000 . We can apply the TOPS to optical functional devices for optical logic and optical memory. The TOPS is expected to play an important role for optical computing.

Acknowledgement

The authors would like to express their thanks to Drs. Y. Urano, K. Sakai and Y. Mimura for encouragement.

References

[1] G. W. Taylor, J. G. Simmons, A.Y, Cho, and R. S. Mand: 'A new double heterostructure optoelectronic switching device using molecular beam epitaxy ', J. Appl. Phys. 1986,**59** ,pp. 596-600

[2] G. W. Taylor, R. S. Mand, J. G. Simmons, and A.Y, Cho: 'Optically induced switching in a p-channel double heterostructure optoelectronic switch', Appl. Phys. Lett. , 1986,**49** ,pp. 1406-1408

[3] F.Y. Huang and H.Morkoç: 'GaAs/InGaAs/AlGaAs optoelectronic switch in avalanche heterojunction phototransistor vertically integrated with a resonant cavity', Appl. Phys. Lett. 1994,**64** ,pp. 405-407

[4] C.Y.Chen, A.Y. Cho, P. A. Garbinski, C. G. Bethea, and B. F. Levine : 'Modulated barrier photodiode: A new majority-carrier photodetector', Appl. Phys. Lett. 1984,**39** ,pp. 340-342

[5] L. Goldstein: 'Optoelectronic Devices by GSMBE', J. Crystal Growth, 1990, **105**, 93-96

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

All-optical bistable devices (ORION) using laser diodes coupled to interference filters of narrow spectral bandwidth

Yoshinobu Maeda[1,2], Masatoshi Migitaka[2], Hong K. Kim[3], Yabo Li[3], Cheng C. Li[3], Dietrich W. Langer[3]

[1]Light and Material Group, PRESTO, JRDC
[2]Department of Information and Control Engineering, ToyotaTechnological Institute, Hisakata, Tempaku, Nagoya 468, Japan
[3]Department of Electrical Engineering, 348 Benedum Hall, University of Pittsburgh, Pittsburgh, PA 15261, USA

Abstract. All-optical bistable devices were demonstrated by using the longitudinal mode hopping of the laser diode and the narrow transmission spectrum of an interference filter. Since a hysteresis characteristics exists in the relationship between the wavelength and the injection current of the laser diode, an optical bistability was observed in this system. Optical switch-on and -off were confirmed by directly injecting a dye laser. In addition, Optical characteristics of Er-doped GaAs and Er-doped silicate glass filters were investigated.

1. Introduction

Optical bistability in laser diodes (LDs) is a most interesting subject because of its many advantages; for example, such diodes have optical gain, can provide large fan-out and can have operating times less than a sub-nsec [1, 2]. Bistable optical devices were confirmed by using the longitudinal mode hopping of the laser diode and the narrow absorption band of erbium in an yttrium aluminum garnet crystal (Er:YAG), and referred to as ORION (Optical logic devices using the Red shift of a laser diode and Inversion in Optical absorbers (or filters) of a Narrow spectral bandwidth) [3]. However, when using an absorber such as an Er:YAG crystal as reported earlier, no absorption wavelength can be utilized other than that specific to the absorber. Furthermore, there was a disadvantage that crystals like YAG hardly conform to semiconductor. This paper shows that the ORION device was successfully demonstrated by using an artificial dielectric interference filter in place of the natural absorber such as the Er:YAG. By using the interference filters in this study, it becomes possible to use an arbitrary wavelength and form the filter on semiconductors using vacuum evaporation with ease. The optical switch-on and -off phenomena were observed by direct injection of an external dye laser beam into the LD. In addition, the natural absorption filters such as the Er-doped GaAs and the Er-doped silicate glass were investigated.

2. Interference filters

The light source was a high power AlGaAs laser diode operating in the single transverse mode. The threshold current is 50 mA, the operating current at optical output of 30 mW is 100 mA and the slope efficiency is 0.6 mW/mA. Its oscillation wavelength was kept stable by adjusting the temperature of the laser-head. The dielectric interference filter was formed by vacuum evaporating 11 layers of 866 Å of TiO_2 and 1335 Å of SiO_2 alternately on a glass substrate heated to 300 °C, then placing 2670 Å of SiO_2 as a spacer, and adding another 11 layers of TiO_2 and SiO_2 alternately with the same thicknesses. The refractive indices of TiO_2 and SiO_2 films were 2.25 and 1.46, respectively. Optical wave-forms were determined with a digital oscilloscope with a bandwidth of 1 GHz and a photo-detector with a rise time of 90 psec. An external dye laser excited with a nitrogen (N_2) laser was injected through a beam splitter into the LD. The dye laser produced a beam with a central wavelength of 546 nm, a spectral width of 2 nm and a pulse width of 500 psec.

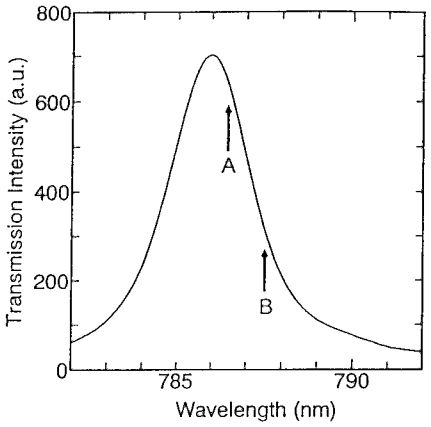


Fig. 1 Transmission spectrum of the interference filter. Points A and B indicate central wavelength of A and B modes of the laser diode.

Figure 1 shows a transmission spectrum of the interference filter, with its center at 786.8 nm and a characteristics of the half-width being about 3 nm. Point A and B indicate central wavelength of A and B modes of the LD. Figure 2 shows relationship after transmission through the interference filter between the laser power and the injection current of the LD in continuous operation measured with an optical power-meter. It was observed that the bistable curve fell rapidly at 82 mA as the injection current increased, and in contrast, rose rapidly 78 mA when it decreased. Figure 3 (a) exhibits the optical intensity variation after transmission through the filter when the pulsed dye laser in Fig. 3 (b) was injected into the LD during continuous oscillation at an injection current of 79 mA. The optical switch-on and-off are observed when the trigger light pulse of Fig. 3 (b) is applied. The rise and fall time are around 1 nsec and these are limited by the bandwidth of

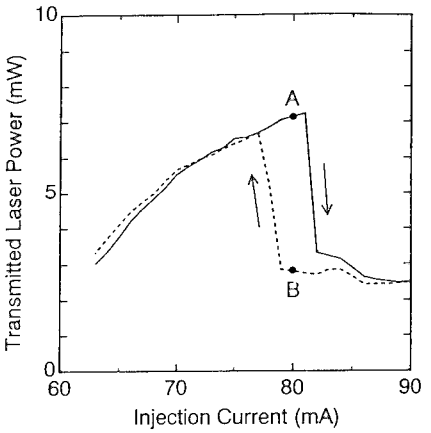


Fig. 2 Laser intensity after transmission through the filter versus injection current of the laser diode characteristics showing hysteresis.

the oscilloscope.

3. Er-doped GaAs and Silicate glass filters

In this section, optical characteristics of Er-doped GaAs films and Er-doped silicate glass films were investigated. Details of the sample preparation were reported in Refs [4, 5]. Er-doped GaAs was grown by metalorganic chemical vapor deposition. Typical photo-luminescence (PL) spectra of Er-doped GaAs epitaxial layers were obtained at 5.5, 77 and 300K. At the temperature of 5.5K, the PL spectrum exhibits two distinct peaks at 1.54 and 1.55 μm. The spectral bandwidth of main peak at 1.54 μm is estimated about 10 nm. The emission lines are related to the internal 4f-4f transition $^4I_{13/2}$ - $^4I_{15/2}$ of Er^{3+}. Figure 4 shows a room-temperature PL spectrum of an Er-doped silica glass film deposited on a Si substrate using RF magnetron sputtering. The spectrum is peaked at 1.54 μm. Another peak (~1.55 μm) is noted the same as the Er-doped GaAs films. The spectral band-width of main peak at 1.54 μm is estimated about 10 nm. From these experimental results, Er-doped GaAs and Er-doped silicate glass have potential applications in optical filters of narrow spectral bandwidth suitable for ORION devices.

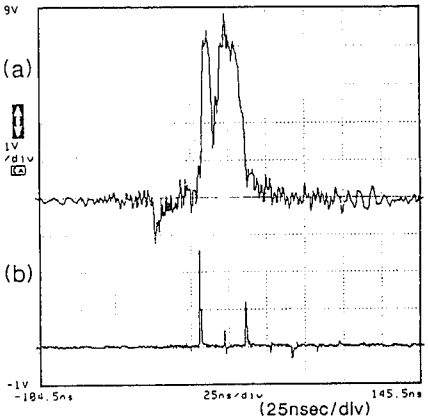


Fig. 3 (a) : Optical intensity variation after transmission through the interference filter by applying the trigger light pulse into the laser diode. (b): Trigger light pulse of dye laser.

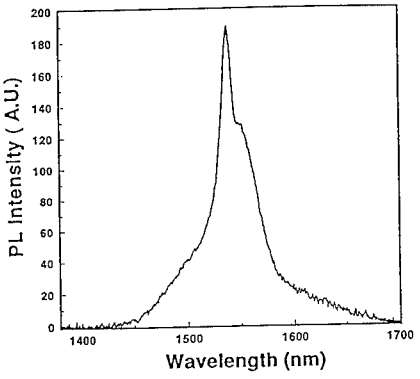


Fig. 4 Room-temperature PL spectra of an Er-doped silica film deposited on a silicon substrate. The film was deposited in an Ar/O_2 (95/5) ambient and annealed at 850°C for 30 min.

4. Switching time of ORION devices

We take consideration on the switching time of the ORION device. As mentioned above, the switching time of this device is duration of giving rise to the longitudinal mode hopping by injecting carrier into the LD. It was confirmed however from experimental to perform the switching response against laser with pulse width of 500 psec. Detailed switching mechanism has to be still explored whether the switching action simply takes place during this 500 psec or after carrier accumulation after the 500 psec. Also for this purpose, accurate switching time must be measured hereafter, by using a laser source of short pulse and a high speed photodetector.

In general, the longitudinal modes of laser diodes are altered by the temperature and injection carrier density in the oscillation area. In DC operation, the injection carrier density is constant for larger than the threshold current and thus the increase in the refractive index caused by temperature rising becomes dominant in slight shifting of a series of longitudinal modes to longer wavelength. If the current is increased further, the oscillation mode is skipped from the center longitudinal mode to the mode with the highest intensity in the series of longitudinal modes on the longer wavelength side [6]. In high-speed communication using shorter pulse width than the attenuation time of the relaxation oscillation, there are as many as several number of longitudinal modes. The devices in this work is greatly dependent on the oscillation wavelength of the laser diode, or the longitudinal modes. Therefore, the changes of the longitudinal modes due to relaxation oscillation is an important factor to determine the characteristics of the device. In addition, the longitudinal hopping has so far been treated as an undesirable event because it causes noise, and there has necessarily been to design suitable from the standpoint of users. It is expected therefore that devices with excellent switching time and characteristics could be fabricated if LDs optimized for making use of the longitudinal mode hopping would be designed.

5. Conclusions

A bistable optical device was demonstrated by using the longitudinal mode hopping of LD and a narrow transmission spectral line of an interference filter. Since a hysteresis characteristic exists in the relationship between the wavelength and the injection current of the LD, an optical bistability was observed in this system. In addition, the optical switch-on and -off phenomena were observed by directly injecting a pulse 500 psec wide into an external dye laser. The Er-doped GaAs and Er-doped silicate glass have potential applications in optical filters of narrow spectral bandwidth. It was confirmed that the device could convert changes in the wavelength of the LD into changes in intensity transmitting through the filter and would be applied to an all-optical logic device.

References

[1] Lasher G J 1964 *Solid-State Electron.* **7** 707-712
[2] Adams M J 1986 *Int. J. Electron.* **60** 123-142
[3] Maeda Y 1993a *Rev. Laser Eng.* **21** 577-582; 1993b *Appl. Phys. Lett.* **62** 3393-3395; 1994 *Appl. Opt.* **33** 267-269
[4] Fang X M, Li Y and Langer D W 1993 *J. Appl. Phys.* **74** 6990-6992
[5] Kim H K, Li C C, Fang X M, Li Y and Langer D W 1994 *J. Luminescence* **60** & **61** 220-222
[6] Ikegami T and Suematsu Y 1968 *IEEE J. Quantum Electron.* **QE-4** 148-151

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Carrier heating induced optical bistability in degenerate semiconductors and quantum-well structures

V I Tolstikhin[1], and M Willander

Department of Physics and Measurement Technology, Linköping University S-581 83, Linköping Sweden

Abstract. Carrier heating induced suppression of band filling is shown as a fast and effective mechanism of optical nonlinearity leading to all-optical bistability in degenerate semiconductors. The regenerative pulsations in bistable etalon are also discussed those are possible due to competition between carrier generation and heating influences on band filling.

1. Introduction.

Photonic switching in III-V semiconductors in the spectral range near the fundamental absorption edge attracts much attention because of its great capability for optical computing and data processing [1]. The major characteristic of the related digital systems is their speed, which is mostly determined by the physical mechanism of the optical nonlinearity of the active medium. Recently, it has been shown that suppression of band filling by the optically induced carrier heating in degenerate semiconductor leads to all-optical switching in a picosecond time-scale [2,3]. Herein we study this mechanism of nonlinearity and related optical bistability as applied to both bulk and multiple quantum well (MQW) GaInAs/InP structures and also discuss the possibility of bistable Fabry-Perot etalon to selfpulsate.

2. Mechanism of nonlinearity

Consider the degenerate n-type semiconductor (or MQW structure) with thermalized electrons having their Fermi quasilevel Φ_e well above the conduction band (or ground subband) bottom. For the energy states below Φ_e the probability of direct interband transitions is reduced by band filling and so the fundamental absorption edge, the sharper the lower carrier temperature, is effectively shifted towards the higher photon energies [4]. However, any increase in the electron effective temperature T_e leads to more occupation of the high-energy states and so suppresses the band filling and smears out the fundamental absorption edge. As a result, for the spectral range corresponding to the direct transitions below the Fermi quasilevel Φ_e the interband absorption coefficient α_{cv} increases with effective temperature as well as the related carrier contribution to the band edge refractive index δn_{cv} is influenced by T_e.

[1] Also at the Institute of Radio Engineering and Electronics of the Russian Academy of Sciences, 11 Mokhovaya, Moscow 103907 Russia

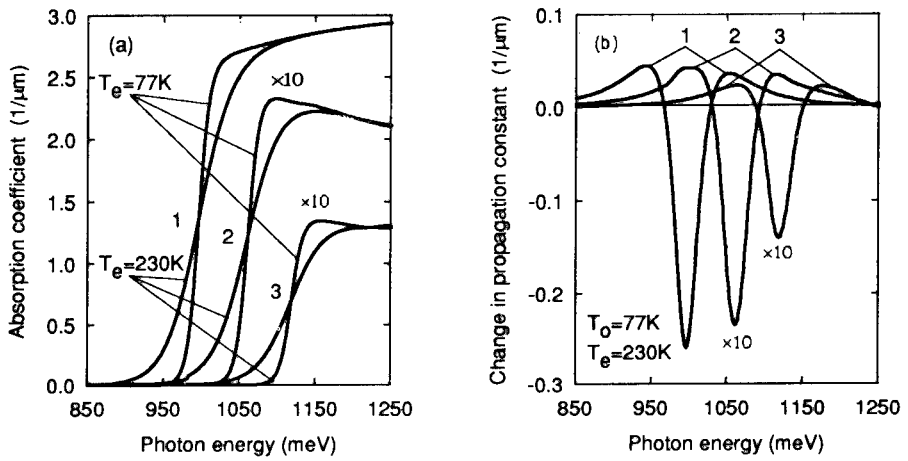


Fig. 1 The influence of an effective electron temperature on the interband absorption coefficient (**a**) and change in propagation constant (**b**) spectra of a bulk $Ga_{0.47}In_{0.53}As$ (1) and $Ga_{0.47}In_{0.53}As/InP$ MQW structures with well width: a = *80Å* (2), *65Å* (3) and period d = *300Å*. Carrier concentration (reduced to a bulk value) is equal to 4.10^{18} cm^{-3} throughout.

These effects are clearly seen from the Fig. 1 showing the spectra of the interband absorption coefficient α_{cv} and related carrier heating induced change in propagation constant, $\delta\beta_{cv} = (\omega/\pi)\mathcal{P}\int_0^\infty d\omega' \left[\alpha_{cv}(T_e, \hbar\omega') - \alpha_{cv}(T_0, \hbar\omega')\right] / \left(\omega'^2 - \omega^2\right)$, computed for both bulk and MQW semiconductors. The applied two-band model does not take into account the Coulomb electron-hole interaction because of the actual photon energies are well above the excitonic absorption region. In the case of MQW structure no coupling between separate quantum wells is assumed, and the absorption coefficient is defined as $\alpha_{cv} = \gamma_{cv}/d$, where d is period of the structure, and γ_{cv} is probability of the interband absorption related to a single quantum well. The normalization constant proportional to the square of the momentum matrix element of a bulk material has been extracted from Ref. [5]. Note, that for MQW structures the values of α_{cv} above the interband absorption edge as well as the peak absolute values of $\delta\beta_{cv}$ are significantly less than for a bulk semiconductor. This is due to lower density of states in a MQW structure as compared to a bulk material. The reduction in α_{cv} directly above the fundamental absorption edge is estimated as $\cong (k_{eF} d)^{-1} \sqrt{1 + (m_e/\mu_{eh})(\Phi_e/\varepsilon_g)}$, where m_e and μ_{eh} are the electron and reduced effective masses, respectively, and $k_{eF} = \sqrt{2m_e\Phi_e}/\hbar$. Under the strong degeneracy conditions this factor is quite small and therefore the same variations of $\delta\beta_{cv}$ require for the larger change in T_e in the case of a MQW structure compared to a case of a bulk semiconductor. That is why we restrict herein our further consideration by the bulk semiconductors only.

Temperature dependencies of α_{cv} and $\delta\beta_{cv}$ lead to absorptive and dispersive optical nonlinearities, respectively, provided the carrier heating is caused by an optical excitation. In such a case for any time-scale beyond the very short stage of carrier thermalization the optical response can be treated as linear, but all the material parameters are assumed to depend parametrically on T_e. Regarding the possible mechanisms of a photoinduced carrier heating in the narrow-gap semiconductors under consideration, the most important ones are: intraband absorption, generation of energetic carriers, and Auger recombination [2,3]

3. Model of the active medium

To describe the carrier heating in optically excited semiconductor, the processes of light-carrier interactions and nonequilibrium carrier relaxation should be considered in self-consistent manner. Our approach as concerning to a bulk semiconductor is based on the following assumptions: (1) Electrons are thermalized i.e. completely described by their concentration N_e and effective temperature T_e. (2) Heavy holes, contrary to light electrons, remain in the equilibrium, and therefore are described by only concentration N_h. The last is rather small compared to N_e but should be taken into account because of the high sensitivity of the interband absorption to changes in N_h. (3) Any nonequilibrium state is quasineutral, i.e. $N_e = N_0 + N_h$, where N_0 is the concentration of ionized impurities. (4) Light absorption within the actual spectral range results from the both direct interband and inderect (impurity, hole, electron and LO-phonon assisted) intraband transitions. (5) The nonequilibrium carrier recombination is due to CHCC Auger process. (6) Electron to lattice energy transfer is governed by the inelastic LO-phonon scattering which is restricted by the phonon bottleneck effect i.e. hot phonon effects also are taken into account. Then, the description of the photoexcited semiconductor is reduced to a couple of the diffusion equations written for the hole concentration N_h and electron temperature T_e. Fig. 2 shows the parametrical dependence of the total absorption coefficient α on the local light intensity $\Im$ related to their steady state homogeneous solution. It can be seen, that the effective temperature is a S-shaped function of the local light intensity in a spectral range corresponding to direct transitions below the Fermi quasilevel, that is due to rapid growth of α with T_e. The multivalued shape of the steady state homogeneous $T_e(\Im)$ curve is the necessary condition for the optical bistability due to increasing absorption, has been discussed previously [2,3]. The true steady state distributions of carrier parameters and thus the real possibility of the bistable response and related switching phenomena depend on the relation between the hole diffusion and electron thermal conductivity lengths, ℓ_N and ℓ_T, the values of absorption and propagation constants, $\alpha^{-1}$and β^{-1} and also the size of active region ℓ.

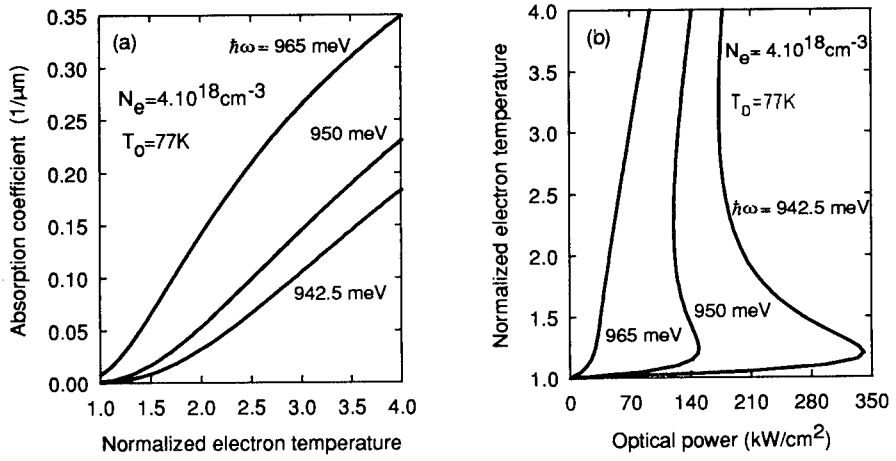


Fig. 2 Nonlinear absorption characteristics of a bulk $Ga_{0.47}In_{0.53}As$: absorption coefficient vs electron temperature (**a**) and electron temperatute vs local optical power (**b**).

4. Nonlinear response of the stationary excited Fabry-Perot etalon.

If the absorption is small due to the band filling, the conditions $\alpha\ell << 1 << \beta\ell$ seem to be quite reasonable. Assuming also the inequalities $1 < \beta\ell_T << \beta\ell_N$ we can treat the carrier parameters distribution as uniform and use the mean-field approximation for optical power within the nonlinear Fabry-Perot etalon formed by a semiconductor layer. Hence, the description of nonlinear response reduces to a couple of rate equations written for a hole concentration N_h and electron temperature T_e. Then, usually cooling time τ_T is as small as $\cong 1$ ps (even if being lagged by a phonon bottleneck effect), while the recombination time τ_N is anyway more than 1 ns. Such a great difference makes it possible to distinguish the fast and slow stages in the dynamics of the photoexcited semiconductor, governed by the relaxation of T_e and N_h, respectively. The fast, within a time scale $\tau_T < t < \tau_N$, response of a nonlinear etalon is shown in Fig. 3, displaying the bistability of the both $T_e(\Im_{in})$ and $\Im_r(\Im_{in})$ characteristics. But, in a time scale more than τ_N such a response under the stationary excitation may be unstable due to the competition between carrier generation and heating influences on band filling. This instability and related regenerative pulsation are of the same nature as ones pointed by McCall [6] and later observed by Mackenzie e.a. [7] in semiconductor etalon with competing electronic (as fast) and lattice (as slow) dispersive nonlinearities. The only one difference is that in our case both competing nonlinearities are electronic (caused by carrier concentration and temperature influences on band filling), but that is why period of selfpulsations (determinined by τ_N) may be reduced to a few ns.

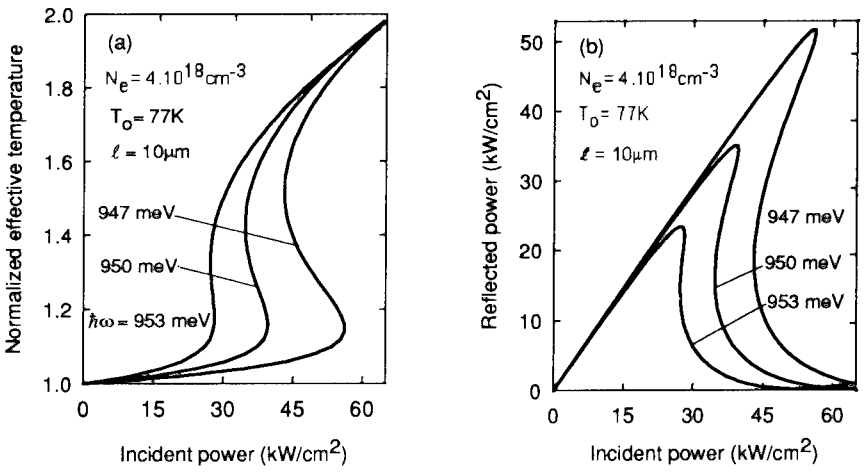


Fig. 3 The bistable steady state characteristics of the asymmetric Fabry-Perot etalon formed by the uniform layer of a bulk $Ga_{0.47}In_{0.53}As$: electron temperature (**a**) and reflected power (**b**) vs incident power

References

[1] Koch S W, Peghambarian N, and Gibbs H M 1988 *J. Appl. Phys.*, **63**, R1-R11
[2] Pishchalko V D, and Tolstikhin V I 1988 *Sov. Phys. Semicond.*, **22**, 1275-1277
[3] Tolstikhin V I 1993 *Tech. Phys. Lett.*, **19**, 62-65
[4] Moss T S, Burrell G J, Ellis B 1973 *Semiconductor Opto-Electronics* (London: Butterworths)
[5] Bacher F R, Blakemore J S, Ebner J T, and Arthur J R 1988 *Phys. Rev.*, **B37**, 2551-2557
[6] McCall S L 1978 *Appl. Phys. Lett.*, 32, 284-286
[7] Mackenzie H A, Reid J J E, Al-Attar H A, and Abraham E 1986 *Optics Commun.*, 60, 181-185

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Photonic Devices Based on Excited-State Nonlinear Absorption

Chun-Fei Li and Miao Yang

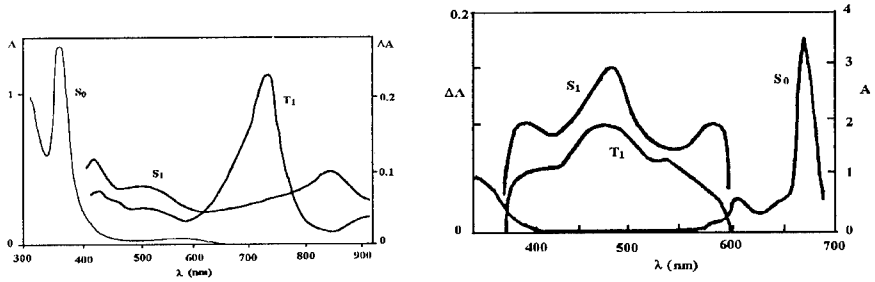
Department of Physics, Harbin Institute of Technology, Harbin 150001, China

Abstract A series of all-optical devices based on nonlinear excited-state absorption working at non-resonant frequency are proposed. Experimental and theoretical results obtained with C_{60} and metal-organic materials using ns and ps laser at 532 nm are presented in this paper.

All-optical devices are necessary for the future optical communication and optical computing. Although semiconductor devices using resonant nonlinearity usually exhibit large nonlinear refractive index, their large linear absorption may cause low device throughput and undesirable thermal effects. Here we demonstrate a new kind of devices based on the excited-state nonlinear absorption working at non-resonant wavelength using organic materials, which are advantageous for their low linear absorption, fast response and mirrorless structure.

1. Excited-State Nonlinear Absorption

The reverse saturable absorption (RSA) of C_{60} and copper phthalocyanine (CuPc) have been investigated. The absorption spectra of the ground state and the differential absorption spectra of excited states are shown in Fig.1(a) for C_{60}[1] and Fig.1(b) for CuPc[2]. One can see that their absorption cross section of the ground-state are all smaller than that of triplet and singlet excited states at 532 nm wavelength, so the RSA can be observed.



(a) (b)

Fig. 1. Absorption spectra of the ground state and the excited states for C_{60} (a) and CuPc (b).

The energy-level diagram for the metallo-organic compounds or C_{60} is shown in Fig.2, where S_0 is the ground-state electronic level; S_1 and S_2 are the first and second singlet excited-state electronic level, respectively; T_1 and T_2 are the first and second triplet excited-state electronic level, respectively. There are many vibronic sub-levels above each electronic level.

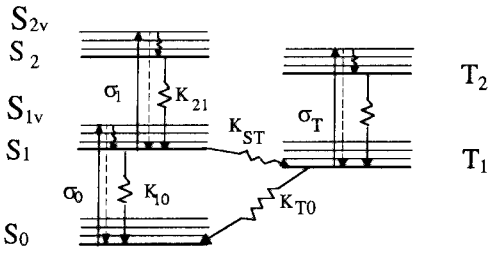


Fig.2 Energy-level diagram for C_{60} and metallo-organic materials.

The molecules in S_0, S_1 and T_1 simultaneously absorb the photons with same frequency ν with absorption crass section σ_0, $\sigma_1$and σ_T, and transit to S_{1V}, S_{2V} and T_{2V}, respectively. They rapidly relax from there to S_1, S_2 and T_2, then nonradiationally relax down to next lower levels with transition probability k_{ST}, k_{10}, k_{21} and k_{T0}, respectively. Because the lifetimes of S_{1V}, S_{2V}, T_{2V} and S_2, T_2 are very short (<ps), the populations in these levels can be neglected. The rate equations for describing time-variation of populations n_0, n_1, n_T in S_0, S_1, T_1, and a light-propagation equation for describing the variation of photonic flux ϕ along z direction in the sample can therefore be written as:

$$\partial n_1 / \partial t = \sigma_0 \phi n_0 - \sigma_1 \phi n_1 - k_{10} n_1 - k_{ST} n_1 \quad (1)$$

$$\partial n_T / \partial t = -\sigma_T \phi n_T - k_{T0} n_T + k_{ST} n_1 \quad (2)$$

$$N = n_0 + n_1 + n_T \quad (3)$$

$$\partial \phi / \partial z = -(\sigma_0 n_0 + \sigma_1 n_1 + \sigma_T n_T)\phi \quad (4)$$

where N is the total population, $\phi = I / h\nu$. Assume the intensity of incident pulsed light is a Gaussian temporal function,

$$I = I(t,z) = I_m(z) e^{-c(\frac{t}{\Delta t})^2} \quad (5)$$

here $I_m(z)$ is the peak intensity at z, Δt is the pulse width of laser, and c is a normalized constant. Using Eq.(1)-Eq.(5), the curve of energy transmittance T=F(L)/F(0) versus the incident fluence F(0), namely RSA characteristics, can be calculated. The curves of T-F(0) have been simulated by calculating numerically, which are well agree with the experimental results as shown in Fig. 3.

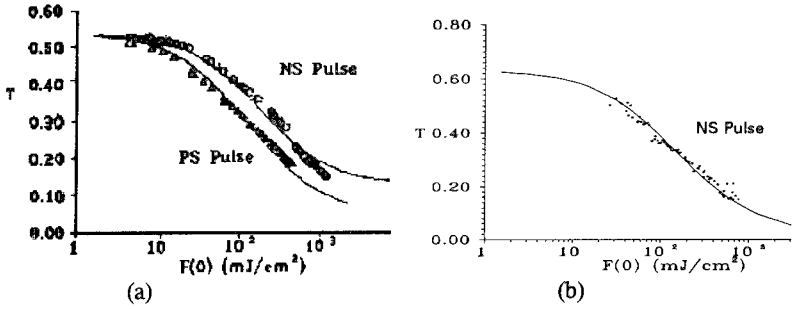


Fig.3. Experiments data and simulations of RSA for solutions of C_{60} (a) and CuPc (b) with ns/ps laser pulses.

2. Photonic Devices Based on Excited-State Nonlinear Absorption

The transient absorptive **optical bistability** was experimentally obtained using a CuPc solution and laser pulses with width of 15 ns at 532 nm. The comparison between calculated result (1) and experimental data (2) are illustrated in Fig.4.

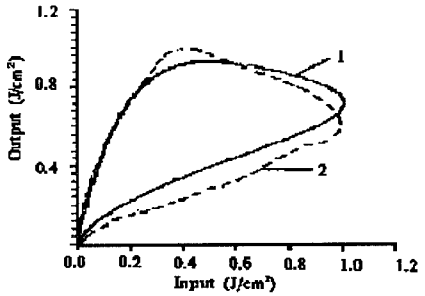


Fig.4. Experiments data and simulations of transient optical bistability in CuPc solution.

The experimental data of **optical limiting**, i.e. output fluence F(L) versus input fluence F(0) were obtained using C_{60} and CuPc solution and a 8 ns pulsed laser at 532 nm. The simulations are also consistent with experimental data as shown in Fig.5.

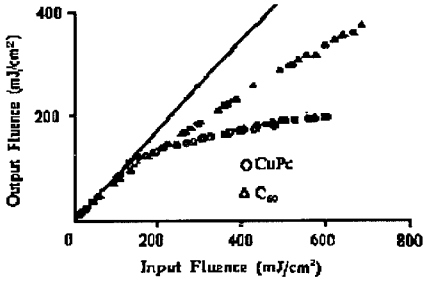
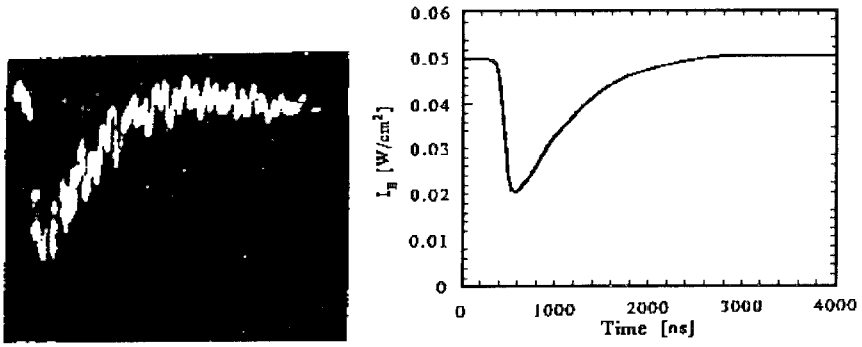


Fig.5 Experiments data and simulations of optical limiting in CuPc and C_{60} solutions.

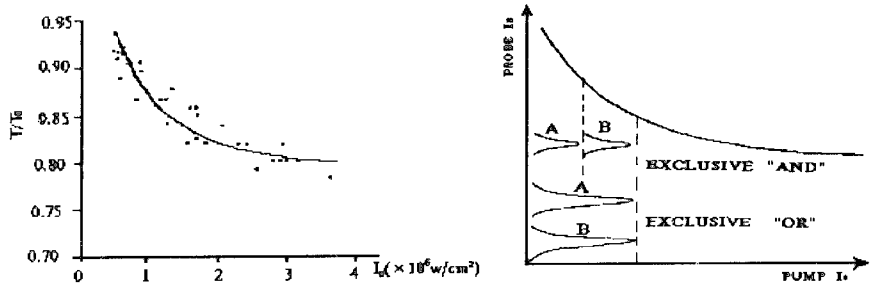
When a Nd:YAG 15 ns pulsed laser with $I_0 = 3x10^5$ W/cm^2 at 532 nm pumps a C_{60} solution solution, a large population can be accumulated in T_1. If in the same time a cw laser diode beam at 747nm absorption peak wavelength of excited state T_1 passes through the C_{60} solution. Photons will be intensively absorbed by molecules in T_1. So the output intensity switches off. This **all-optical switching** process (photograph) is shown in Fig.6(a). The simulation can be made by using Eq.(1)-(3) for pump beam and Eq.(4) for probe beam as shown in Fig.6(b). The switch-on time depends on the incident laser pulse-width about 15ns; the switch-off time mainly depends on the relaxation time of T_1 about 300 ns.



(a) (b)

Fig.6 All-optical switching in C_{60}. (a) Experimental photograph, (b) Simulation curve.

Fig.7(a) gives the experimental and simulated results of **optical modulation**, i.e. relative transmission of probe beam T/T_0 versus peak pumping intensity, T_0 is the transmission of probe beam without the pump beam I_0. A design for "Exclusive AND" and "Exclusive OR" logic gates based on the modulation characteristics at different working points is shown in Fig.7(b).



(a) (b)

Fig.7 The modulation characteristics (a) and A design of exclusive "and" and "or" logic gates (b).

The excited-state photonic devices (ESPD) have many advantages: with fast response time (ns to sub-ps); small total linear absorption coefficient; small working power for probe beam, and mirrorless structure. So this kind devices promise to be use in the all-optical computing, the all-optical communication and the laser protection.

References

[1] Ebbesen T W and Tanigaki K et al 1991 *Chem. Phys. Lett.* **181** 501-503
[2] Hercher M and Chu M et al 1968 *IEEE J. Quantum Electron.* **QE-4** 954-968
[3] Ohtanl Hiroyuki and Kobayashi Takayoshi et al 1984 *J. Phys. Chem.* **88** 4431-4435
[4] Li Chunfei and Zhang Lei et al 1984 *J. Opt. Soc. Am. B* **11**
[5] Li, Chunfei and Zhang Lei et al 1994 *Phys. Rev. A* **49** 1149-1157
[6] Li Chunfei and Yang Miao 1993 *International J of Non. Opt. Phys.* **2** 551-558

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Polychromatic all-optical gate based on second-order nonlinearity

V P Torchigin and A E Kobyakov

Scientific Computer Centre, Russian Academy of Sciences, Leninsky pr. 32 A, Moscow, 117334, Russia, e-mail: user@comcp.msk.su

Abstract. A possibility of using known quasi-phase-matched second-harmonic generators as main components of an all-optical gate capable of performing logical operations on WDM signals is considered. Main parameters of the gate are estimated.

I. Introduction

To date, the wave-division-multiplexing (WDM) techniques are widely used allowing one to transmit simultaneously many optical signals having different wavelengths through the single optical channel. The opportunities of such an approach would be much greater if polychromatic all-optical gates capable of direct processing of WDM signals were available. Now, no gates of such kind are known even in theory.

One should note that, at present, there are also no all-optical gates for processing usual optical signals which were able to compete with electronic gates in such basic parameters as dissipated power and switching frequency. Nevertheless, carrying out careful analysis, one can find that, to date, there are relatively small-power and superfast all-optical devices that could be used as the main nonlinear component of the WDM gate with the overall rate of signal processing near the maximal bit-rate in a short optical communication line. These are integrated-optic quasi-phase-matched second-harmonic generators (SHG).

A simple and practical method is described for constructing WDM gates consisting of SHGs [1-3] and linear wavelength selective couplers. All these components are tested experimentally in many optical devices.

2. Operating principles

The main idea of the gate is that the known quasi-phase-matched SHG based on an integrated-optic domain-inverted waveguide [1,2] can be used as an optical parametric travelling wave amplifier (OPTWA) without being changed. For an OPTWA, the condition for quasi-phase-matching (QPM) can be written as follows:

$$\Delta k = 2\pi/\Lambda \tag{1}$$

where

$$\Delta k = (n(\omega_p)\omega_p - n(\omega_s)\omega_s - n(\omega_i)\omega_i)/c \tag{2}$$

$n(\omega)$ is the effective refractive index of a waveguide, Λ is the period of the domain structure in an SHG, c is the velocity of light in vacuum, ω_p, ω_s, and ω_i are carriers of pump, signal, and idle waves, respectively, and

$$\omega_p=\omega_s+\omega_i \tag{3}$$

In the particular case when $\omega_s=\omega_i=\omega_p/2$ equation (1) is valid for a quasi-phase-matched SHG. For a normal dispersion medium , in particular, for the domain-inverted waveguide $n(\omega)$ is an increasing function of ω.

One can verify the following two statements concerning the value of Δk in (1) provided that $n(\omega)$ is an arbitrary increasing function:

1. If ω_p, ω_s, and ω_i increase by positive value $\Delta\omega$, $\Delta\omega/2$, and $\Delta\omega/2$, respectively, then Δk becomes greater.

2. If ω_p is constant, $\omega_s>\omega_i$, and the difference $\omega_s-\omega_i$ increases by positive value $\delta\omega$ then Δk becomes less.

Thus, for each $\Delta\omega>0$ we can find such a value of $\delta\omega>0$ that Δk is constant. Actually, the increment in Δk caused by increase in ω_p of $\Delta\omega>0$ can be compensated by a decrease in Δk due to an increase in the value of $\omega_s-\omega_i$ by $\delta\omega>0$. In other words, any SHG with $\omega_s=\omega_p/2$, for which equation (1) is satisfied, can be used as an OPTWA with other values of ω_p and ω_s for which (1) is also valid.

In the particular case when $n(\omega)$ is a linear increasing function and when for $\omega_p=\omega_{p0}$ $\omega_s=\omega_i=\omega_{p0}/2$, then to satisfy quasi-phase-matching condition (1) for pump wave carrier greater than ω_{p0} signal and idle wave carriers should correspond the following functions:

$$\omega_s=(\omega_p+[\omega_p^2-\omega_{p0}^2]^{1/2})/2 \tag{4},$$

$$\omega_i=(\omega_p-[\omega_p^2-\omega_{p0}^2]^{1/2})/2 \tag{5}.$$

Here ω_{p0} is the carrier of the pump for which the OPTWA operates in the degenerate regime. The same OPTWA can operate in the SHG-regime to obtain the second harmonic with the carrier ω_{p0}. According to (4) and (5) the dependence of ω_s and ω_i on ω_p is depicted in figure 1. Because values ω_s and ω_i are determined uniquely from the system of equations (1)-(3), for $\omega_p>\omega_{p0}$ each value of ω_p corresponds to the only pair of values ω_s and ω_i. There are no any waves with carriers ω_s and ω_i different from pointed above that could interact with the pump ω_p.

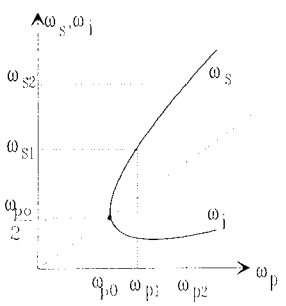


Fig. 1 Dependence of the carriers ω_s and ω_i of signal and idle waves on the carrier ω_p of the pump in an OPTWA if the QPM condition is satisfied

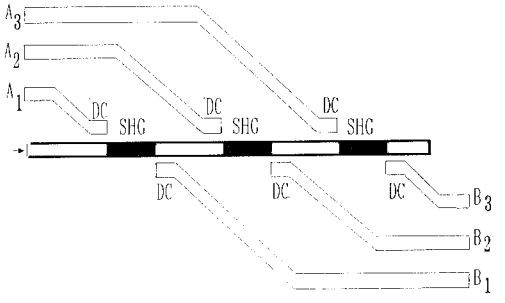


Fig. 2 Schematic diagram of the gate

If several pumps with different carriers enter an SHG simultaneously then the result of the change in the refractive index n is equal to the sum of the effects when each pump enters an SHG separately. The character of modulation performed by one pump does not depend on presence of other pumps. Input control signals with carriers ω_{sm} (m=1, 2,...,*M*) interact only with corresponding pumps with carriers ω_{pm} for which condition (1) is valid.

As the waves propagate along a lightguide, the interaction mentioned manifests itself in an increase in the intensity of the waves with carriers ω_s and ω_i and in a decrease in intensity of the wave with carrier ω_p [4]. The intensity of the wave with carrier ω_p falls to zero after some distance. If the length of an SHG is equal to this distance, there is no wave with carrier ω_p at the output of an SHG. This situation can take place if a small control signal *A* with carrier ω_s enters an SHG.

3. Gate design

A schematic diagram of the proposed gate is shown in figure 2. Several identical SHGs with the periodical domain-inverted structures [1,2] are connected in series. Optical power is delivered to the input of the left SHG in the form of a periodical sequence of pulses with carrier frequency ω_{pm}. Frequency selective directional couplers (DC) are used both for entering logical signals A_{jm} (in fig. 2 j=1,2,3) with carriers $\omega_{sm}<\omega_{pm}$ and for extraction of result signals B_{jm} with the same carriers from the SHG. The coefficient *H* of cross-transmission of DCs is close to *1* for signals with carrier frequency ω_{sm} and to *0* for signals with carrier ω_{pm}. As the localization of the field of a light wave in the cross-section of a lightguide decreases for lower frequencies, coupling indexes for waves with carriers ω_s and ω_i essentially exceed that for a wave with carrier ω_p. In this case the usual linear DC having $H\cong0$ for the wave with carrier ω_p has $H\cong1$ for waves with carriers ω_s and ω_i.

We can see that the signal with carrier ω_s is present at the output B_1 only when control signal A_1 with carrier ω_s is present at the input, that is the output signal corresponds to logical function $B_1=A_1$. Analogously, a signal with carrier ω_s is present at the output B_2 only in the absence of signal A_1 and the presence of signal A_2, i.e. the resulting signal at output B_2 corresponds to logical function $B_2=\overline{A}_1A_2$.

Generally, the gate have *N* logical inputs A_1, A_2,...A_N for control signals and the same number of logical outputs for result signals B_1, B_2...B_N. The resulting signals correspond to the following functions: $B_1=A_1$, $B_2=\overline{A}_1A_2$, $B_3=\overline{A}_1\overline{A}_2A_3$...

4 Main parameters

The power required for the gate can be evaluated from an analysis of the power characteristics of SHGs which are the main components of the gate and the only nonlinear devices whose operation depends on the intensity of light signals. Present technology makes it possible to produce SHG with a power of tens milliwatts. Standard GaAs/AlGaAs lasers can be used as a power supply. For a gate the power of the input control signal can be 10^1-10^2 times less than that of the pump, that is, less than power of the other known waveguide all-optical gates by a factor of 10^3.

With regard the rate of processing optical pulses, the greatest restriction is imposed by the difference in group velocities of pulses with carriers ω_p and ω_s in an SHG. Assuming that the difference in velocities is equal to 8% [5] and the length of an SHG is ~1cm, we obtain that the pulse walk-off is 10 ps and pulsewidth must be greater, e.g. 20 ps. Besides, minimal

pulsewidth τ is limited by the bandwidth of an amplifier. Bandwidth ΔF of an SHG having a length of about 1 cm is 0.2 nm [1], or $6*10^{10}$ Hz. As $\tau=1/\Delta F\cong 20$ ps, the maximal rate of the pulses is about 25 GHz. This is essentially less than the bandwidth of the optical transmission channels, but the gate can process M sets of signals simultaneously and the maximal rate of the pulses is shown to be M times higher.

The maximal value of M is determined by the expression $M=W_t/W_1$, where W_t is the total bandwidth for the various pumps determined by the frequency capabilities of the nonlinear material, waveguides, and couplers, W_1 is the bandwidth for a single pump. For example, if W_t is equal to 10% of carrier ω_1, $\omega_1=6\pi*10^{14}$ s^{-1}, $W_1=50$ GHz then $M=600$. Naturally, the total power of a pump also increases by M times. Thus, taking into consideration the whole bandwidth of WDM signals and possibility of simultaneous processing M sets of signals, we can see that the number of logical operations performed by the gate is comparable with the maximal number of bits transmitted through the optical communication lines.

Contrast ratio κ of the gate is equal to the ratio of the output result signals corresponding to logical *1* and *0*. When the pump is present and the input signals are amplified by γ dB, the output signal corresponds to logical *1*. When the pump is absent and the input signals pass to the output with an attenuation of σ dB, the signal corresponds to logical *0*. The value of σ is 0.5 - 1.0 dB/cm, $\kappa=\gamma+\sigma$, and the value of γ is restricted only by the level of noise and the parasitic signals reflected from the outputs and can exceed 20 dB. Thus, the gate has satisfactory parameters from the viewpoint of a circuit designer.

5. Applications

It appears likely that the principal application of the gate described here will be in WDM systems. Another application of the gate is its use in all-optical supercomputers. Such a gate enables an implementation of a polychromatic computer comprising a set of N monochromatic supercomputers working concurrently. Each computer uses common optical hardware (gates and transmission lines) and operates in its own wavelength independently on the others. In this case total performance is increased N times without the hardware being enlarged.

6. Conclusion

The analysis performed shows that the present state of the art for creatiing optical quasi-phase-matched second- harmonic generators makes it possible to construct an all-optical gate without essential additional investigations. The gate can process several sets of optical pulses simultaneously, has satisfactory parameters from the viewpoint of circuit designer, and is significantly superior to any other available in operating rate and consumption power. Moreover, no new materials or technologies are required for its fabrication.

References

[1] Arvidsson G, Laurell F and Webjourn J 1993 *Proc. of the 6-th European Conf. on Integr. Optics ESIO-93 Apr. 18-22 1993 Neuchatel Switzerland* 3-23 - 3-25
[2] Yamamoto K, Mizuuchi K and Taniuchi T 1992 *IEEE J. of Quantum Electronics* **28** 1909-1914
[3] Fujunura M, Kintaka K, Juhara T and Nishihara H 1992 *Electron. Letters* **28** 1868-1871
[4] Bloembergen N *Nonlinear optics* (W. A. Benjamin, Inc. New York-Amsterdam, 1965)
[5] Colucci G, Romano D, Bava G P and Montrosset I 1992 *IEEE J. of Quantum Electronics* **28** 729-738

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Diffraction kinetics of electronic and thermal transient gratings in GaAs epilayers and GaAs/GaAlAs multi-quantum wells

N. Gouaichault, R. Grac, L. Legratiet, E. Bedel, C. Fontaine, J.L. Iehl, M. Pugnet and J.H. Collet,
LAAS-CNRS, 7, Avenue du Colonel Roche- 31077 - Toulouse Cedex- France

Abstract:

Diffraction kinetics of transient gratings in GaAs epilayers and GaAs/GaAlAs multi-quantum wells at room temperature are reported. Samples are prepared by lift-off epitaxy. The results are interpreted in terms of time separation of the electronic and thermal contributions.

1. Introduction

The aim of this work was to study the subnanosecond diffraction capabilities of GaAs epilayers and GaAs/GaAlAs multi-quantum wells at room temperature, using a forward geometry. The epitaxial structures were grafted onto glass slides by means of the lift-off epitaxy technique. Diffraction kinetics of transient gratings were measured using the first order diffraction in the Raman-Nath configuration.

2. Experimental set-up

Pulses at the wavelength 0.532 μm were generated by frequency doubling in KDP crystals the fundamental harmonic of an actively mode-locked Nd:YAG laser. The duration of the pulses was approximately 30 ps. These pulses synchronously pump a dye laser operating in the range (825-850 nm). The duration of the pulses was approximately 15-20 ps. Two pump pulses (P1 and P2) interacted inside the sample to produce transient gratings.The gratings were then read by a probe pulse (S) generated by the dye laser and directed in normal incidence to the sample surface (Raman-Nath configuration). The probe pulse was delayed with respect to the pump pulses P1 and P2. The pump and probe beam sections were limited by a circular diaphragm stuck on the sample. The diameter of the hole was 1mm. We measured the energy of the *first order* diffracted pulse versus the probe delay. In a first set of experiments, gratings were generated at the wavelength 532 nm and probed in the infrared. The external angle between the two pump beams was $2\theta = 0.1$ radian, corresponding to a grating period Λ of 5.3 μm. A typical diffraction efficiency kinetics is shown in Fig. 1. Then, in a second set of experiments, we use degenerate pumps and probe in the infrared. The external angle was $2\theta = 0.18$ radian and $\Lambda = 4.6$ μm. The diffraction efficiency kinetics are reported in Fig 2.

3. Structure preparation

Active multi-quantum well epitaxial structures on transparent substrates were achieved in two steps. First, the structures were grown on GaAs substrates by molecular beam epitaxy. They consist of 80 periods of alternating GaAs wells and (Ga,Al)As barriers, whose thicknesses are 9.6 nm and 13 nm respectively. The quantum wells are sandwiched between two 0.2 μm thick $Ga_{0.7}Al_{0.3}As$ layers. Second, the epitaxial structures were removed from their substrates and grafted onto glass slides by means of

the lift-off epitaxy technique. The structures areas were around 4 mm^2. The optical transmission spectra of the structures was recorded. The absorption threshold of the structure were around 850 nm at room temperature (excitonic resonance).

4. Experimental results with gratings generated at 532 nm

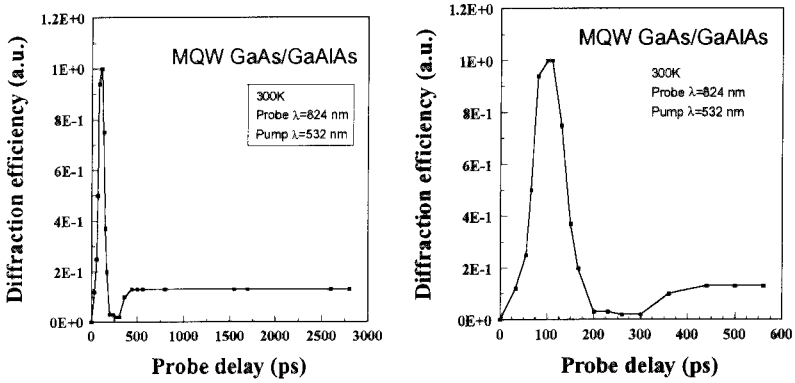


Fig. 1: Diffraction kinetics in the first order Raman-Nath configuration measured on multi-quantum wells GaAs/GaAlAs. The pump pulse energies for P1 and P2 are respectively 4 mJ/cm^2 and 2 mJ/cm^2 . In the left curve, the time axis has been expanded.

The diffraction kinetics exhibit a peak during the first 200 ps and then a minimum within the interval (200-300 ps). The diffracted pulse is rather unstable in this late interval, then each reported data point has been averaged over 20 pulses. After 300 ps the energy of the diffracted pulse rises up to a stable value. We interpret these results in the following manner:

The peak in the diffraction kinetics of Fig.1 is due to an electronic process. The presence of the electron-hole plasma induces a *decrease* of the optical index in the illuminated fringes. Moreover, the crystal is heated by carrier relaxation and non radiative recombination, which results an *increase* of the optical index in the illuminated fringes. The decay times of these two opposite contributions are very different. A change of sign of the optical index variation takes place during the interval (200-300 ps). This transient extinction of the index grating results in a pronounced minimum observed in the diffraction kinetics. In the following, we will briefly discuss the order of magnitude of the electronic and thermal contributions to the diffraction efficiency in order to test the coherence of this interpretation. The absorption coefficient in $Ga_{0.7}Al_{0.3}As$ at 532 nm is around $6x10^4$ cm^{-1} [1]. Then, we will consider here that only the 0.2 μm thick $Ga_{0.7}Al_{0.3}As$ layer is excited. The plasma density generated in this layer during the picosecond excitation is limited by Auger recombination. We estimate the plasma density at $\rho=5x10^{19}$ cm^{-3} in the illuminated fringes (taking the Auger coefficient equal to $3x10^{-30}$ cm^6 s^{-1}). The plasma induced optical index variation can be written [2]:

$$\Delta n_p = -\frac{1}{2n_b}\frac{e^2}{\mu\varepsilon_0\omega^2}\left[\frac{\omega_g^2}{\omega_g^2-\omega^2}\right]\rho$$

where n_b is the optical index, μ is the reduced mass, ω and ω_g are the angular frequencies of light and bandgap respectively. At the peak of the curve in Fig.1, the plasma contribution to the diffracted signal is dominant, then the above-mentioned

formula can be used to estimate the optical index variation. We obtain a maximum index variation of 0.1 to 0.2. The other contribution to the optical index variation arises from temperature rise ΔT dynamics which is governed by:

$$\frac{d(\Delta T)}{dt} = -\left(\frac{2\pi}{\Lambda}\right)^2 D_{th}\Delta T + \frac{1}{Mc_p}\frac{\rho(t)E_\rho}{\tau}$$

where D_{th} is the heat diffusion coefficient, c_p is the specific heat capacity, M is the crystal density and $\rho(t)$ is the time dependent electron-hole plasma density. The recombination of the carriers occurs with a time constant τ and releases a quantity of energy E_ρ per electron-hole pair to the lattice. For probe delay above 400 ps, the grating decay time τ_T is only governed by heat diffusion, and is given by:

$$\frac{1}{\tau_T} = \left(\frac{2\pi}{\Lambda}\right)^2 D_{th} = \left(\frac{2\pi}{\Lambda}\right)^2 \frac{\kappa_T}{Mc_p}$$

where κ_T is the thermal conductivity. Assuming a total conversion of optical into thermal energy, the initial temperature rise is estimated to $\Delta T = \alpha E_{opt} / Mc_p$ and the optical index rise reads: $\Delta n_T = (dn/dT)\,\Delta T$. For numerical calculations we use: $n_b = 3.4$ at 824 nm and $n_b = 3.87$ at 532 nm [1], the electron effective mass $m_n = 0.092$ [3], the hole effective mass $m_h = 0.5$ [4], the bandgap $E_g = 1.8$ eV, $\kappa_T = 0.122$ J s^{-1} $cm^{-1}K^{-1}$, the specific heat capacity $c_p = 0.364$ J g^{-1} K^{-1}, M = 4.88 g cm^{-3}, and $dn/dT = 1.9\ 10^{-4}$ [5].

The first order diffraction efficiency in the Raman-Nath regime takes the following forms[6], for the peak value and the plateau value respectively:

$$\eta_{peak} = AJ_1^2\left[\frac{2\pi(\Delta n_p + \Delta n_T)d}{\lambda}\right] \approx A\left[\frac{2\pi(\Delta n_p + \Delta n_T)d}{\lambda}\right]^2$$

$$\eta_{plateau} = AJ_1^2\left[\frac{2\pi\Delta n_T d}{\lambda}\right] \approx A\left[\frac{2\pi\Delta n_T d}{\lambda}\right]^2$$

Where λ is the probe wavelength, J_1 is the Bessel function of the first order, d is the grating thickness, and A is a constant. Then, we obtain the following ratio:

$$\frac{\eta_{peak}}{\eta_{plateau}} \approx \left[\frac{\Delta n_T + \Delta n_p}{\Delta n_T}\right]^2$$

We have neglected here the contribution of the absorption gratings. This is supported by the fact that the probe transmission modulation is lower than 2%, therefore the calculated diffraction efficiency is lower than 10^{-4}, which is negligible with respect to the measured value, around 10^{-3}. Using the above-mentioned parameters, except for the Auger coefficient which is taken between 10^{-30} cm^6 s^{-1} and 10^{-29} cm^6 s^{-1}, the numerical calculation gives:

$3 < \eta_{peak}/\eta_{plateau} < 20$.

This result is in reasonable agreement with the experimental data (see Fig.1). In Fig. 1, the transition between the electronic and thermal gratings has been clearly observed because of the high density and temperature of carriers generated in the $Ga_{0.7}Al_{0.3}As$ crystal. In the following experiments, we will use excitation wavelengths close to the bandgap, then E_ρ and ΔT will decrease.

5. Diffraction results in degenerate pump-probe experiments

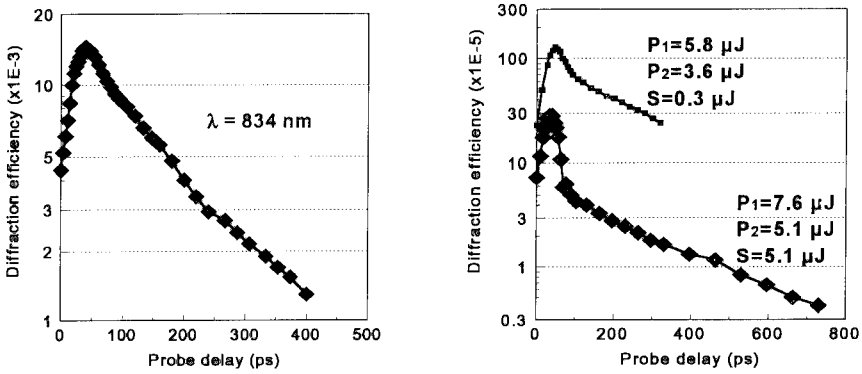


Fig. 2: Diffraction kinetics in the first order Raman-Nath configuration measured on multi-quantum wells GaAs/GaAlAs. Left: the pulse energy for P1 , P2 and the probe S were respectively 6.7 μJ, 4.6 μJ and 0.23 μJ. Right: the wavelength was λ = 815 nm; the diffraction efficiency decreases with the probe pulse energy.

In these experiments, pump and probe beams are produced by the same picosecond dye laser. The total pump energy of the order of 1 mJ/ cm^2. The diffraction efficiencies are defined with respect to the transmited pulse. The Fig.2 (Left) shows the diffraction efficiency kinetics for λ = 834 nm. This wavelength corresponds to a rather homogeneous excitation of the wells. The diffracted signal decay time is 170 ps, corresponding to a grating lifetime in the structure of 340 ps. Taking into account the ambipolar diffusion, we estimate a carrier lifetime greater than 500 ps. We show in Fig. 2 (Right) two kinetics at 815 nm. In this case, the excitation of the wells is more inhomogeneous than at 834 nm. The main difference between the two kinetics lies in the values of the probe pulse energy. In the experiment reported in the upper curve, the probe energy is lower than the pump energy while in the case of the lower curve, pump and probe energies are of the same order of magnitude. In the last case, a partial grating erasure by the probe pulse is probable. The maximum diffraction efficiency is roughly an order of magnitude greater at 834 nm than at 815 nm (due to the influence of the excitonic resonance at 850 nm). No thermal gratings have been detected in these degenerate experiments, in contrast with the previous results (Fig. 1). This is due to the decrease of the pump pulse energy used, and to the improvement of the excitation homogeneity. To conclude, the good reliability of the lift-off epitaxy technique allows its use for all-optical epitaxial device engineering.

References

[1] Aspnes D E , Kelso S M , Logan R A and Bhat R 1986 *J. Appl. Phys.* **60** 754-767
[2] Auston D H, Mc Afee S, Shank C V, Ippen E P, Teschke O 1978 *Solid St.Electron.* **21** 147-150
[3] Adachi S 1985 *J. Appl. Phys.* **58** R1-R29
[4] Blakemore J S 1982 *J. Appl. Phys.* **53** R123-R181
[5] Landolt-Börnstein 1986 *Zahlenwerte und Funktionen aus Naturwissenschaften und Technik* (Springer Verlag)
[6] Born M and Wolf E 1964 *Principles of Optics* p 610 (Pergamon Press)

Aknowledgements: The authors wish to thank A. MUNOZ-YAGUE (LAAS-CNRS), P. GRAVEY and J.-Y. MOISAN (CNET Lannion) for fruitful discussions and FRANCE TELECOM for financial support.

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Design Examples of Wavelength Conversion and Filtering Components

Jun-ichi Mizusawa

NTT Telecommunication Networks Laboratories
Address: 3-9-11, Midori-cho, Musashino-shi, Tokyo 180, Japan
Tel: +81 422 59 2551 Fax: +81 422 59 2552
E-mail: mizu@tnlab.ntt.jp

Abstract Since we expect future photonic networks to use wavelength multiplexed technologies, we designed two key components (wavelength conversion and filtering components) using commercially available popular technologies. Design parameters and design results are introduced.

1. Introduction

From the perspective of a photonic switching network system designer , the first step in the implementation of photonic networks should be the introduction of frequency multiplexing technologies into subscriber access fiber lines. Figure 1 illustrates this idea. Frequency multiplexing over a subscriber-access fiber line facilitates transporting a variety of media signals over a single fiber. By making line usage more efficient, this method encourages fiber installation, which is indispensable for future fiber networks.

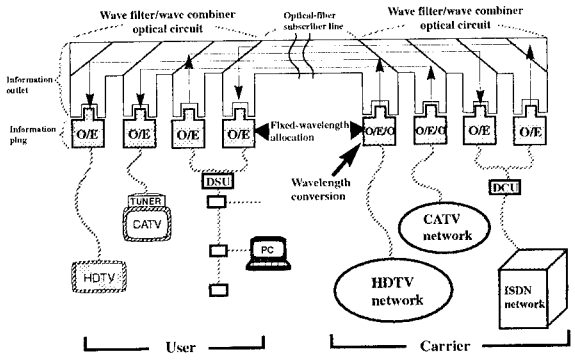


Figure 1. Shared Usage of a Subscriber Access Fiber.

If this proposal is accepted, the next step would be the introduction of a switching system based on frequency multiplexing, such as the one shown in Fig. 2. Since the switching matrix has a very large traffic capacity, it should use time-division multiplexing(TDM) technology.

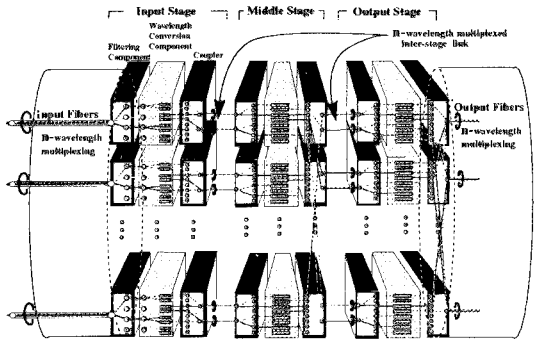


Figure 2. Experimental Switching Fabric Design.

The key functions in this system are wavelength conversion, filtering, and coupling.

As a system designer, I tried to find devices that perform these functions, but it is quite difficult to get a number of components that can be assembled as part of a switching system. Most research-level devices are not yet practical in terms of cost, reliability, and availability. I therefore designed two of these components: a wavelength converter and filters. The purpose of this design exercise was to determine the availability of existing commercial technologies. Of course, they will be improvd as new device technologies become available. The purpose of this work was to give device researchers an example of the design parameters that photonic switching system designers may want the future.

The system design parameters are as follows. The values were selected mostly based on my knowledge and experience. The main objective was to find a set of values that can be implemented as an experimental fiber access line and a switching system based on frequency multiplexing. Therefore, an economical component design was the first priority.

(a) Wavelength band: 1100 - 1600 nm
(b) Number of multiplexed waves: 8
(c) Minimum adjacent wavelength separation: 20 nm
(d) Fiber mode: multi-mode (50-micron core diameter)
(e) Maximum signal speed: 622 Mbps (>1 GHz)
(f) Transmission distance: 5-10 km

The selected laser diode (LD) wavelengths were 1175-1190, 1210-1225, 1275-1285, 1305-1310, 1330-1340, 1505-1525, 1545-1550, and 1570-1580 nm.

2. Wavelength Conversion Component

This component converts a signal of one wavelength into another wavelength with the control of wavelength selection signal. The incoming fiber signal is received by a photo detector (PD),

the electronic signal is amplified, an LD driver activates the appropriate LD (out of eight), and then the output light is guided to an output fiber through a coupling device.
(a) Input signal level: >-35 dBm (applied InGaAs-APD)
(b) Electronic amp. gain: >40 dB
(c) Modulation transparency: digital/analog (by bias selection)
(d) LD output signal level: 0 to 3 dBm (using a commercial Fabry-Perot resonator-type LD bare chip)
(f) Fiber output signal level: -10 to -15 dBm
(g) Signal selection port: TTL level (one out of eight selection)
(h) Size: 135x70x40 mm (excluding connectors and pigtails)
(i) Power: +5,-5,+80 volts
(j) Other: power-feed connector, channel-selection monitor LED, input-signal monitor LED.

3. Filtering Component

This component receives the signal from the input fiber, the signal contains eight multiplexed wavelengths. The selected wavelength signal is guided to the output port (out of eight) that corresponds to the wavelength. This component has no active devices. The input light signal is converted into a concurrent beam by multiple lenses. Each time the beam reaches a filter, it is divided into two wavelength signals: one passes through to the output port and the other is reflected to the next output port.Each output port has a lens circuit that guides the light to an output fiber.
(a) Number of newly designed lens types: eight
(b) Filter production technology: dielectric multi-coated filter
(c) Number of filter types: 16
(d) Insertion loss: less than 10 dB (planned)
(e) Size: 26x26x18 mm (excluding connectors and pigtails)

4. Design Evaluation

As stated, the purpose of this design exercise was to determine the availability of existing commercial technologies. I found that it is possible to design both wavelength conversion and filtering components if the system design parameters listed in Sec.1 are accepted. From these experimental components, it is also possible to construct a subscriber-access fiber line and a switching system based on frequency multiplexing, even though technical elaboration is necessary.

Figure 3 shows the LD drive current vs. the optical output power characteristics of the wavelength conversion component. The commercial technologies used were Fabry-Perot resonator-type LDs and an electronic circuit. The horizontal axis includes the coupling device

loss, which is approximately 15 dB. Although Figure 3 only shows the measurements for one sample component, it clearly shows characteristic changes in LD drive current vs. optical output power for each wavelength. The electronic circuit should therefore compensate for this characteristic difference. To prevent over-driving the current on a LD, a current limiting circuit is included.

Figure 4 shows the average insertion loss distribution for each frequency as measured for forty experimental filtering components. The commercially available technology used here was dielectric multi-coated filter production. The production technology was improved a bit to meet the design targets. The results show that it is possible to design a very small passive filter. The component size is 26x26x18mm. Assembly work has proved that the alignment of optical beam direction is critical. The figure shows the results of manual beam adjustment. The insertion loss could be reduced to 10 dB by improving the optical beam alignment.

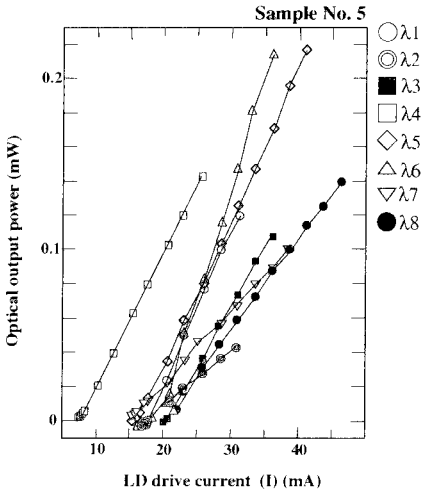


Figure 3. Wavelength Conversion Component.

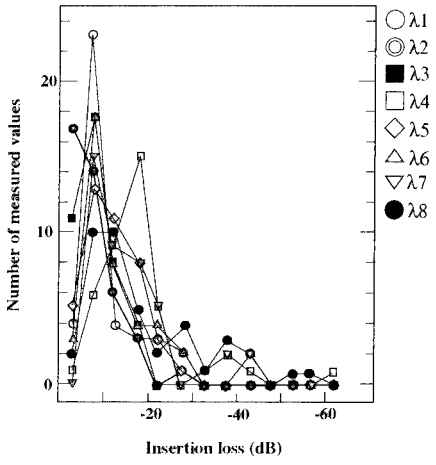


Figure 4. Filtering Component.

This design exercise showed that :

(a) Both components were produced more than forty, their cost was reasonable.

(b) Eight wavelengths provide an adequate degree of multiplexing at present. With more, it would be difficult to use commercial LD devices and coupling device loss would increase.

(c) There is a technical bottleneck in the connection between the LD and the coupling device. Laser diode alignment deviation causes a loss of more than 10 dB. A simple method to solve this problem is required.

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd.

Waveguide type rotating phase plate as optical frequency shifter

Hiroaki Inoue, Shinji Nishimura, Tatsuo Kanetake and Hideaki Takano

RWCP Optoelectronics Hitachi Laboratory,
c/o Central Research Laboratory, Hitachi, Ltd.,

Abstract

A large number of activities on 2 dimensional optical devices for free space optical interconnection have been done in order to attain a large interconnection through-put for optical computing. In addition to the space domain processing, optical frequency domain data processing can realize the larger through-put by means of a multi-dimensional interconnection technology.

Introduction

Here, an optical frequency shifter would be one of the most important devices, because it may realize the multi-dimensional interconnection as well as the FDM crossconnect technology in communication network. (Shown in Fig.1) Some optical frequency (wavelength) conversion devices such as semiconductor amplifier utilizing FDM effect and bistable laser diodes are reported so far. These have advantageous features of high conversion efficiencies (>0dB) and large frequency (wavelength) shift (>10nm). However, there are also drawbacks on signal bandwidth (<GHz) and signal modulation code (only applicable to intensity modulated signal). Rotating half wave-plate [1], in principle, can realize bit-rate free and modulation code free optical frequency conversion since it stands on Doppler effect.

Principle

In this paper, we report the fundamental modulation characteristics of waveguide type optical frequency shifter based on the rotating phase plate, as our results of this project in the last fiscal

year. We successively demonstrate parallel and perpendicular phase modulation characteristics for the effective azimuth control of the phase plate.

Figure 2 shows the principle of the optical frequency shifting by the rotating half-wave plate. The angular frequency of right (left) circularly polarized incident light (ω) is converted to $\omega-\omega p$ on the rotating phase plate in clockwise (counter clockwise) due to Doppler effect. Here, ωp denotes the angular frequency of the rotating phase plate. If the phase plate is a half-wave plate, the output light becomes the left (right) circularly polarized light with the angular frequency of $\omega-2\omega p$. This process does not depend on the signal bit-rate and modulation code of incident light. Therefore, the optical devices based on this principle can be applicable to the bit-rate free and the modulation code free (not only to intensity modulated, but also to FSK or PSK coded signals) optical frequency shifter. The rotating phase plate could be realized by applying the rotating electric field to EO material effectively.

Design

A schematic view of the fabricated device are shown in Fig.3,4. The waveguide was formed by RIBE, after successive growth of $Al_{0.3}Ga_{0.7}As$ lower clad, GaAs guide and $Al_{0.3}Ga_{0.7}As$ upper clad layers on a (110) GaAs substrate being 6° off towards <111>B. All layers were undoped and grown by MBE. Because the light propagation direction should be in the 3-fold axis of GaAs(<111>), the waveguide was designed to be S-shaped, which was inclined to $(\bar{1}10)$ cleaving facets by 35° 3 Schottky electrodes (Cr/Au), one of them being on the waveguide and two of them being besides the waveguide, were evaporated to apply the rotating electric field to the waveguide.

Experiment

A laser diode (λ=1.55μm) was used as the input light source. Both of TE and TM modes were launched to evaluate the phase modulation characteristics induced by the parallel and perpendicular field, respectively.(shown in Fig.5) As shown in Fig.6 Obtained Vπ values for the parallel and perpendicular field were 10V and 13V, respectively Therefore, by applying sinusoidal electric fields to the electrodes with adequate phases, the effective rotating azimuth of the phase plate can be

induced as the waveguide frequency shifter because of necessity of the high voltage and high frequency electric driver. However, by introducing large EO effect by quantum size effect such as QCSE [2], we believe the issue can be overcome. Therefore extensive research on new nonlinear optical materials is strongly requested.

Conclusion

In conclusion, the fundamental modulation characteristics of waveguide type optical frequency shifter based on the rotating phase plate have been reported as an optical device for multi-dimensional interconnection for the first time. Demonstrated parallel and perpendicular phase modulation characteristics for the effective azimuth rotation of the phase plate indicated the possibility of the bit-rate and modulation code free frequency shifter.

References

1) C.F.Buhrer, D.H.Baird and E.M.Conwell, "Optical frequency shifting by electro optic effect", Appl. Phys. Lett., vol.1, p.46, (1962)

2) S.Nishimura, H.Inoue, H.Sano and K.Ishida, "Electrooptic effect in an InGaAs/InAlAs multiquantum well structure", IEEE Photon. Technol. Lett., vol.4, p.1123, (1992)

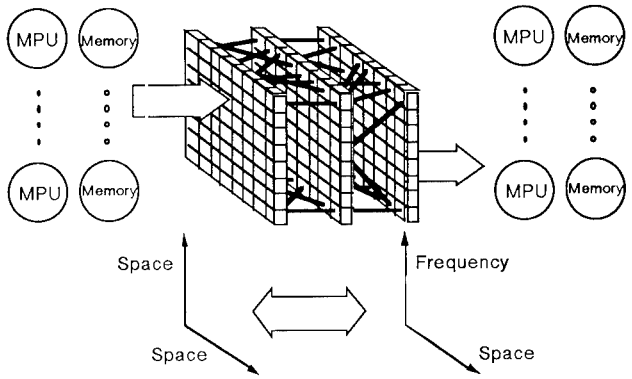


Fig.1 2-dimensional optical interconnection

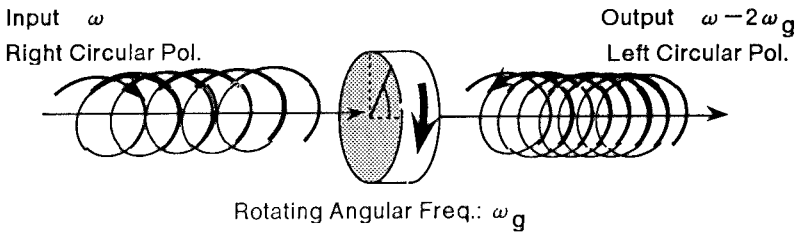


Fig.2 Principle of the frequency conversion applied rotating phase plate

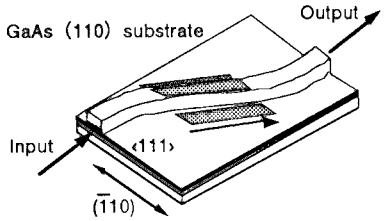


Fig.3 Device structure of the waveguide type optical frequency converter

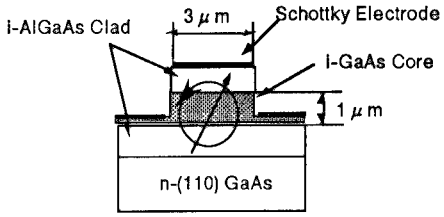


Fig.4 Sectional view of the frequency converter

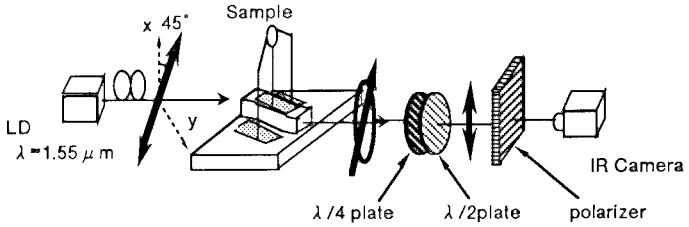


Fig.5 Measurement setup for the phase modulation

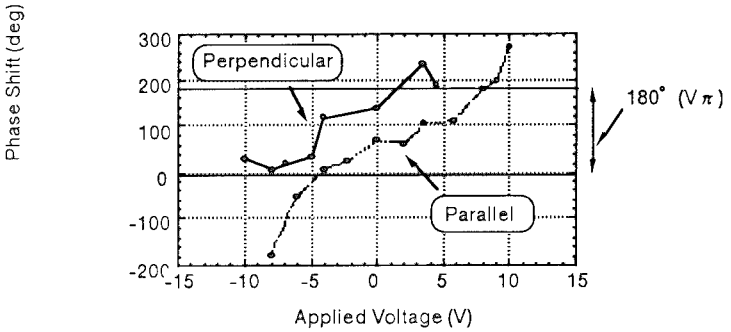


Fig.6 Measurement result of the fundamental phase modulation characteristics

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Sequences of bifurcations and transition to Chaos in an Optical-Processing Element

Ana González-Marcos & J.A. Martín-Pereda

E.T.S.I. Telecomunicación. Ciudad Universitaria, 28040 - Madrid, Spain

Abstract. Digital chaotic behaviour in an Optical-Processing Element is reported. It is obtained as the result of processing two fixed trains of bits. Period doublings in a Feigenbaum-like scenario have been obtained. A new method to characterize digital chaos is reported

1. Introduction

An optically-programmable digital circuit has been already reported by us, [1]-[2], as a Programmable Logic Gate. A brief description on its method of operation, as well as the way it has been implemented, can be found there.

As is very well known from the literature, there are several situations where a chaotic behaviour arises from electrical and electronic circuits. Most of the results concern, and are related, to analogue signals. Their characteristics have been studied by conventional methods employed in any other nonlinear phenomena.

A very different situation is present when the circuit operates with digital signals and the possible chaotic result is a signal composed of "zeroes" and "ones". Hence, the main objective of this paper is to present a new method to obtain the above mentioned type of chaotic signals as well as an alternative way to their study.

2. General structure of the Optical-Processing Element with feedback.

The general scheme of the Cell has been previously reported in several places [1]-[3]. A logic behaviour was reported showing the possibility to obtain up to fourteen pairs of logic functions from two digital inputs. Two control gates allow the change from one type of logic output to another. The cell was implemented with optical components although the non-linear devices **P** and **Q**, namely an "on-off" and a "SEED-like", were simulated with optoelectronic methods (see Fig. 1). The output of each one of them corresponds to the two final outputs, O_1 and O_2, of the cell. The possible inputs to the circuit are four. Two of them are for input data, $\mathbf{I}_1$ and $\mathbf{I}_2$, and the other two, $\boldsymbol{g}$ and $\boldsymbol{h}$, for control signals. The corresponding inputs to the non-linear devices, **P** and **Q**, are based on these signals plus some others coming from inside the cell.

The practical implementation we have carried out of the processing element has been based on an optoelectronic configuration. Lines in Fig. 1 represent optical multimode fibers. The indicated blocks, placed in order to combine the corresponding signals, are conventional

optical couplers. In this way, the inputs arriving to the above mentioned **P** and **Q** devices, are multilevel signals. These devices have been simulated electronically. Optical signals were converted to electrical by conventional photodiodes and, after processing, converted again to optical signals by LEDs. More details can be seen in reference [3].

A new situation appears when some type of feedback is added to the cell. Moreover, in order to have the possibility to work with some more parameters, a time delay has been added to the feedback. Another time delay has been introduced inside the own cell. This time corresponds to response time of the non-linear devices that, in our previous case, were optoelectronic simulations. The general configuration appears in Fig. 1, where these delays, as well as the whole cell configuration, is shown.

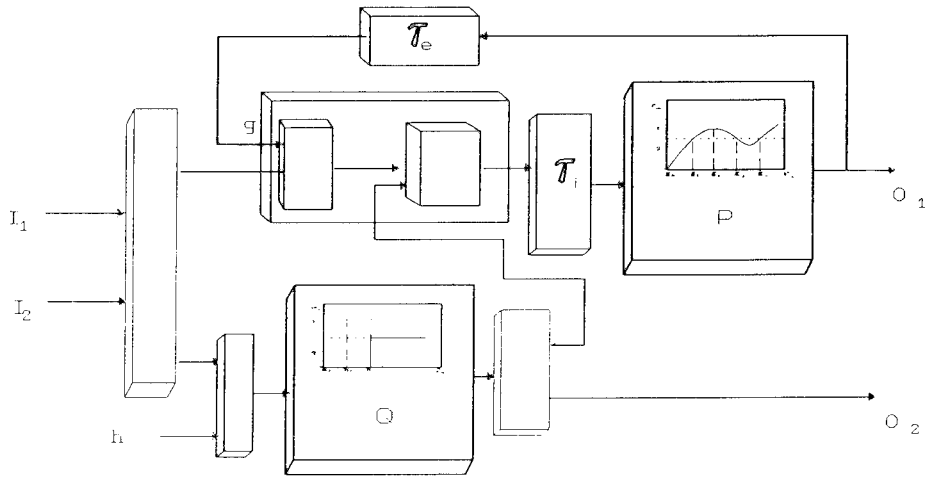


Figure 1.- **Optical-Processing Element with Feedback. White boxes are 2 x 2 or 2 x 1 couplers.**

As it can be shown, there are several possibilities to add feedback to the cell. Any connection between one of the two possible outputs, $\mathbf{O}_1$ or $\mathbf{O}_2$, and any of the four different inputs, namely, $\mathbf{I}_1$, $\mathbf{I}_2$, $\boldsymbol{g}$ and $\boldsymbol{h}$, should give feedback. But results, depending on the adopted configuration, have to be different. Because the **P**-device output has more possible different output functions, depending on its control signal, namely seven, than the **Q**, its output has been used for feedback. This signal will be the control signal $\boldsymbol{g}$ for device **P**. Figure 1 shows the employed circuit. A computer simulation has been employed for the rest of present work.

3. General behaviour of the cell.

The first analysis that we have performed considered null delay times. This situation has no analytic solution and no data were obtained. The circumstances are strongly different if we introduce finite delay times, namely, internal and external delays.

According to previous studies [4], the situation with more probability to give a periodic or even chaotic solution is when the internal delay time is shorter than the external one. In any

case, input has been a regular train of pulses. The input to the non-linear device is a multilevel signal corresponding to the addition of the two periodic inputs. The period of this signal corresponds, in the case studied, to a time of 14 milliseconds.

If the ratio between internal delay time and external delay time is smaller than 1 ms, we obtain a periodic situation. The period of this signal is strongly dependent on the ratio value. In the particular case, where external delay time is 200 ms and internal delay are 2, 4 and 12 ms, obtained results are summarized in Table I. An interesting result is the duplication in period time when the ratio between delays gets smaller. In our case, it goes from 70 to 280. Hence, frequency doubling has been obtained. This result is one of the best indications of a possible route to chaos.

TABLE I.- Characteristics of the output signals, according to the delay times.

t_p	τ_e	τ_i	τ_i/τ_e	Period
14	200	2	0.01	280
14	200	4	0.02	140
14	200	12	0.06	70

Values given at Table I do not correspond to the real transition points between different periods. These values are in a range where the period remains constant. If we calculate the equivalent to the Feigenbaum ratio for the indicated values, a value of 4 is obtained. But if higher order transition points are taken into account, a number, closer to 4.6, has been obtained.

As it is can be seen in Table I, as the internal delay time goes to smaller values, the period of the output signal gets higher and, eventually, becomes chaotic. This situation has been obtained only by computer simulation with internal delay time zero. Experimentally, we have not tried to obtain it. A sample of such a situation is shown in Fig. 2.

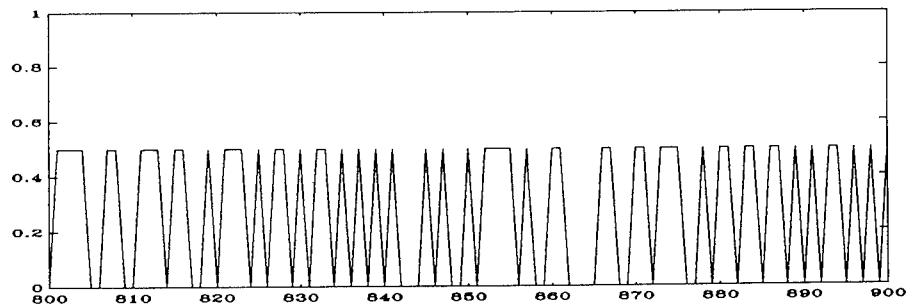


Figure 2.- Output from the logic cell when a chaotic behavior is present.

In order to characterize the obtained chaotic signal, conventional methods are very difficult to apply. A problem related with the above results is the presence of a digital signal with just to two values, "0" and "1". Methods employed with analogue signals are not

applicable here. Hence a new technique has to be implemented. The method we have adopted is to group sets of four bits and to convert them to their corresponding hexadecimal value. Hence, for example, "0010" would be a "2", "1001" a "9" and "1110" a "14".

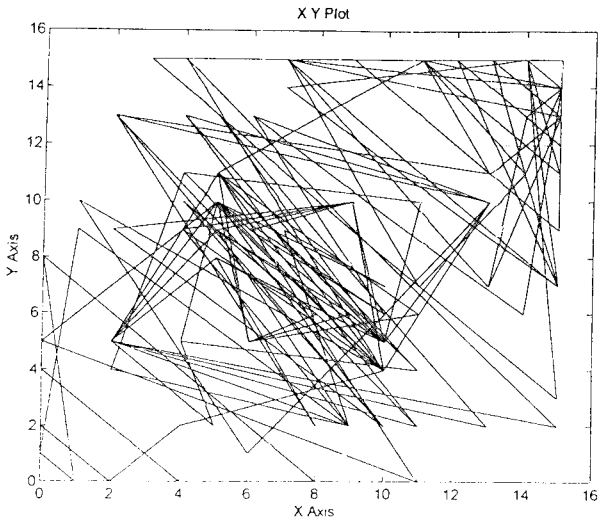


Figure 3.- Diagram t_{i+1} vs. t_i for a digital chaotic signal as in Fig. 2.

A diagram, similar to the t_{i+1} versus t_i in analogue signals, has been obtained here. In the case of periodic signals, a regular configuration is obtained. But in the case of chaotic signals, no definite pattern is obtained. This situation appears in Fig 3.

4. Conclusions

A new type of digital chaotic signal has been presented. It is the result of a feedback in an optical processing logic cell, previously reported. According to the reported results, a new possibility to study digital chaos has been employed . It is based in the conversion from binary to hexadecimal signals. Diagrams t_{i+1} vs. t_i can be obtained with our method.

References

[1] González-Marcos, A. & J.A. Martín-Pereda. "Quasi-chaotic digital behaviour in an optically processing element". **SPIE**, 2038, 67-77. (1993).

[2] J.A. Martín-Pereda & A. González-Marcos. "Optical Programmable Processing Element using Optical Fibers". IEEE Lasers and Electro-Optics Society, **LEOS'92**. Boston, 15-20 November, 1992.

[3] A. González-Marcos, *PhD Thesis*. Universidad Politécnica de Madrid. 1993.

[4] A. Neyer and E. Voges, "Dynamics of Electrooptic Bistable Devices with Delayed Feedback", **IEEE J. Quantum Electron.**, QE-18, 2009-2015. 1982.

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

New optical effects stimulated by anisotropy in information processing schemes

A.M. Goncharenko[1], N.A. Khilo[1], N.N. Korneyev[1], V.N. Belyi[2], A.G. Khatevich[2]

[1]Division for Optical Problems in Information Technologies, Academy of Sciences of Belarus, P.O. Box No. 1, 220072 Minsk, Belarus
[2]Laboratory of Crystallooptics, Institute of Physics, Academy of Sciences of Belarus, Fr. Scaryna ave., 70, Minsk 220072, Belarus

Abstract. The possibility of self-focussing and diffractionless propagation of light beams in linear biaxial gyrotropic crystals has been shown. Asymmetrical components of Kerr-like nonlinearity tensor being taken into account may lead to instability of the slow component of fibre modes and also to self-rotation of the polarisation plane of linearly-polarised incident light.

1. Focussing and diffractionless propagation of light beams in biaxial gyrotropic crystals

Further development of methods of forming spatial structures of light beams is based on the use of different spatially nonuniform elements - lenses, waveguides and cavity structures, etc. There is quite a different way to control light beams spatial structure which is based on the use of anisotropic media. Induced by anisotropy, changing of surfaces of eigenmode wave vectors is more radical in acoustics and results in phonon self-focussing and self-trapping [1]. Observation of effects of such kind in optics is more difficult owing to conservation of the positively curved wave vector surface in the presence of anisotropy. The exception here is light propagation in the vicinity of optical axes in biaxial crystals. The availability of small perturbations removing the degeneracy of phase velocities, e.g. gyrotropy, tends to substantial reconstruction of the wave vector surface structure in the vicinity of axes and to the appearance of regions with negative curvature (or close to zero) which makes possible, in its own turn, photonic self-focussing and self-trapping.

1.1 The structure of the surface index of refraction in the vicinity of binormals

Let us represent beams in anisotropic crystals in terms of the integral

$$E(r,z)=\frac{1}{(2\pi)^2}\int A(q)e^{i(k_z z+qr)}d^2q \quad , \tag{1}$$

where the wave vector $\vec{k}$ of partial wave is taken as $\vec{k}=k_z(q)\vec{n}+\vec{q}$, where $\vec{n}$ is the unit vector in the beam axis direction. For weakly diverging beams the following expansion is valid [2]:

$$k(q)\approx k_0 n_0-\frac{1}{v}u_1 q-\frac{1}{2v}\hat{u}_2 qq-\frac{1}{6v}\hat{u}_3 qqq \tag{2}$$

where tensors of m-rank $\hat{u}_m = \partial^m \omega / \partial k^m$, ω - light frequency, v, n_0 - phase velocity and index of refraction in the beam axis direction, $k_0 = \omega/c$; u_1 - the vector of group velocity, $\hat{u}_2 / v$ - curvature tensor of wave vector surface.

From Maxwell's equations for gyrotropic anisotropic media, when the beam axis and binormal exactly coincide one may obtain the following expressions for u_1q and u_2qq:

$$u_1 q = \frac{c^2}{2v}\varepsilon_0^{-1} q_1 \quad , \quad u_2 qq = \frac{vpq^2}{k_0 n_0} \quad , \tag{3}$$

where

$$\varepsilon_0^{-2} = \left(\varepsilon_3^{-1} - \varepsilon_2^{-1}\right)\left(\varepsilon_2^{-1} - \varepsilon_1^{-1}\right) \quad , \quad p = \frac{n_0^2}{2}\left(\varepsilon_1^{-1} + \varepsilon_3^{-1} \pm \frac{\varepsilon_0^{-1}}{G}\right) \quad , \tag{4}$$

$v^2 / c^2 = \varepsilon_2^{-1} \pm G$, G - projection of gyration vector on binormal direction, $q^2 = q_1^2 + q_2^2$, upper and lower signs refer to fast and slow waves accordingly. With (3)-(4) being taken into account it follows that principal values of curvature will be $u_{11} = u_{22} = p / k_0 n_0$.

From this relationship it follows that in the range of the crystal parameters, where $\left(\varepsilon_3^{-1} - \varepsilon_2^{-1}\right)\left(\varepsilon_2^{-1} - \varepsilon_1^{-1}\right) > \left(\varepsilon_3^{-1} + \varepsilon_1^{-1}\right)G$, u_{11} and u_{22} become negative, that means concave formation at wave vector surface of the slow wave.

1.2 Gaussian beam propagation in a crystal with negative and zero curved wave-normal surface

We consider two kinds of scheme for propagation inside the crystal, when the incident beam is either a divergent or a convergent gaussian one, accordingly.

In the first case the output gaussian beam radius and its phase front curvature radius are determined from the expressions:

$$\omega^2(\bar{z}) = \omega_0^2\left(1 + \left(\frac{2\bar{z}}{k\omega_0^2}\right)^2\right) , \; R(\bar{z}) = n_0 \bar{z}\left(1 + \left(\frac{k\omega_0^2}{2\bar{z}}\right)^2\right) \quad , \tag{5}$$

where $\bar{z} = z_1 + z_2 p / n_0$, z_1 is the distance from waist to the front face of the crystal, z_2 is the distance passed by light inside the crystal of length L. Beam parameters at the output of the crystal $(z_2 = L)$ are determined by (5), only there R is replaced by $n_0 R$.

From (5) it follows that the availability of the crystal parameter $p \neq 1$ implies the character of beam diffraction. We consider regime, when $p \leq 0$.

1.2.1 Diffractionless beam propagation

Two variants may occur: a) $z_1 \neq 0$. Then $\bar{z} = z_1$, ω^2 and also R do not depend on distance passed inside the crystal and therefore diffractionless propagation occurs. b) $z_1 = 0$. Then $\bar{z} = 0$, and $\omega^2 = \omega_0^2$, $R = 0$, that is extending the waist of the gaussian beam on the crystal length occurs.

1.2.2 Beam focussing

Let p < 0. The parameter $\bar{z}$ will be decreased while light propagates inside the crystal that is the beam radius will be decreased. The specific manifestation of a focussing effect herewith will essentially depend on the parameters z_1, p and L. If $\bar{z}$ tends to zero while $z_2 = L$, then focusing at the back face of the crystal occurs. If $\bar{z} = 0$ while $z_2 < L$, then the beam first is focussed inside the crystal at the distance $z_2 = -z_1/p$ from the input. Then the beam is diverged diffractionally but has negative phase front curvature that tends to give secondary focussing outside the crystal at the distance $F = -(z_1 + Lp/n_0)$ from the output of the crystal.

We investigate now the case when a convergent gaussian beam hits the crystal surface. For its formation one must supplement the optical scheme considered above with a lens with focal length F.

The solution of the diffraction problem in this case has the form of (5) with the replacement $\bar{z} \to \bar{z}_1$, where $\bar{z}_1 = z_1 - F + z_2 p/n_0$.

For the case when $z_1 < F$. Focussing by lens occurs inside the crystal while when $z_1 > F$ the problem is reduced to that considered above. We consider two cases here.

1) Let $z_1 < F$, $p < 0$. Then $\bar{z}_1 < 0$ corresponds to propagation inside the crystal and diffraction divergence of a beam with concave wave front. The beam focussing, having stopped inside the crystal goes on outside, and the waist is at a distance of $F_1 = -\bar{z}_1(L)$ from beam output.

2) Let $z_1 < F$, $p = 0$. In this case focussing inside the crystal is stopped and a beam with a negative wave front curvature radius and free from diffraction is formed. Further focussing occurs after its output from the crystal. Thus in the case presented effects of suppression of the divergence and the focussing are realised subsequently.

The study carried out shows that biaxial gyrotropic crystals possess qualitatively new optical properties for light propagating in the vicinity of the optical axes. The investigation of effects of focussing and diffractionless propagation is attractive for realisation of optical interconnections in optical information processing schemes.

2. Light propagation in birefringent optical fibres with anisotropic Kerr-like nonlinearity

Assume a nonlinear polarisation of the form

$$P_i^{nl} = \chi_{ijkl}(\omega = -\omega + \omega + \omega) E_j^* E_k E_l \qquad (6)$$

The tensor χ_{ijkl} is known to possess frequency-reversal property of symmetry and therefore in the notation presented is symmetrical in the last two indexes [3,4]. We write it as a sum $\chi_{ijkl} = \chi^s_{ijkl} + \chi^a_{ijkl}$, where χ^s_{ijkl} and χ^a_{ijkl} are symmetrical and asymmetrical on index reversal, e.g. $\chi^a_{ijkl} = -\chi^a_{ijkl}$. For such a representation each of the tensors χ^s_{ijkl} and χ^a_{ijkl} is symmetrical in the first pair of indexes. So in further calculations we use 6-dimensional

indexes $\chi_{ij}(i,j=1\cdots6)$. We examine several optical effects determined by χ_{16}^{s} and χ_{26}^{s} being taken into account.

2.1 Account of χ_{16}

From equations for coupled modes it follows that energy is conserved if the component χ_{16}^{a} is purely imaginary, that is $\chi_{16}^{a}=i\chi_{16}$, that is clear from the general condition of hermitivity for tensors for nonabsorbing media. By choosing an appropriate circularly polarised basis one can show that nonlinear interaction influences only the mode phases. Therefore one may obtain the equation for the phase difference $\psi=\varphi_{+}-\varphi_{-}$:

$$\frac{d\psi}{dz}=-\chi_{12}\left(|a_{+}|^{2}-|a_{-}|^{2}\right)+\chi_{16}\left(|a_{+}|^{2}+|a_{-}|^{2}\right) \quad . \tag{7}$$

As appears from (7), a nonlinear phase shift and, therefore, self-rotation of the polarisation plane, in comparison with the case of $\chi_{16}=0$ [5,6], occurs for linearly polarised incident light.

Account for χ_{12}^{a}

The coupled mode equations have the form

$$-i\frac{da}{dz}=\left(\chi_{12}^{c}+i\chi_{12}^{a}\right)a^{*}b^{2}+\left(2\chi_{66}|b|^{2}+\chi_{11}|a|^{2}+k\right)a \quad , \tag{8a}$$

$$-i\frac{db}{dz}=\left(\chi_{12}^{c}-i\chi_{12}^{a}\right)b^{*}a^{2}+\left(2\chi_{66}|a|^{2}+\chi_{22}|b|^{2}-k\right)b \quad , \tag{8b}$$

where k is proportional to the linear anisotropy. Questions concerning stability of the slow mode are studied below.

We introduce both the parameters of anisotropy $\Delta_{11}=(\chi_{22}-\chi_{11})/2$, $\Delta_{12}=2\chi_{66}+\chi_{12}-(\chi_{11}+\chi_{22})/2$ and a parameter x_0^2, characterising the relative level of excitation of orthogonally-polarised modes. When $x_0^2=0$, the slow mode is excited, which was stable without account of the Kerr-like anisotropy [5]. Now it loses stability when $(2\bar{k}-\Delta_{11})/\Delta_{12}<1$, where $\bar{k}=k/\alpha_0^2$, α_0^2 - the total intensity.

Under self-consistent change of x_0^2 and α_0^2, when $x_0^2=\left(1+(\Delta_{11}-2\bar{k})/\Delta_{12}\right)/2$, the slow mode keeps the stability while intensity increases. The polarisation of the mode is elliptical and is changed from linear on the Y-axis (polarisation of slow mode) when $2\bar{k}=\Delta_{11}-\Delta_{12}$ to linear on the X-axis when $2\bar{k}=\Delta_{11}+\Delta_{12}$.

References

[1] Newberry B.P., Thompson R.B., 1985, J. Acoust. Soc. A., **85**, 290-2300.
[2] Khatkevich A.G., 1972, J. Prykladnoj Spectrosc. (USS), **17**, 237-244.
[3] Blombergen N, 1965, Nonlinear Optics, N.Y.: Benjamin.
[4] Shen Y.R., 1984, The Principles of Nonlinear Optics, N.Y.: John Wiley & Sons, Inc.
[5] Winful H.G., 1985, Appl. Phys. Lett., **47**, 213.
[6] Stolen R.H., Batineau J, Ashkin A, 1982, Opt. Lett., **7**, 512.

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Architectures for space-wavelength photonic switching using collinear guided wave acoustooptics

V. V.Proklov, A.U.Bashlakov and V.A.Birjukov

Institute of Radioengineering and Electronics,Russian Academy of Sciences,11 Mochovaya str., 103907 Moscow, Russia

Abstract. There are analyzed potential parameters of the planar AO modulators for space- and wavelength photonic switching. We consider different schemes of superfast AO switching with capability up to 10^{15} switch./sec within one multichannel guided wave device. On the basis of the experimentally acheved parameters of the planar Ti:$LiNbO_3$ AO modulators we discuss a few prospective applications in high speed digital multipliers, commutators for optoelectronic super computers and associative memory systems.

1. Introduction

Recent developments in acoustooptics show a fairly good set of main parameters of the acoustooptic modulators (AOM) appropriate to perform optical beam control and the optical and radio signal processing [1]. The most important of these are the amplitude dynamic range (generally N_A~200÷250), the spatial deflection to many resolvable directions (N_Q~500), the number of resolvable sound frequencies (N_S~500) and the polarization planes ($N_P \leq 2$). The typical risetime for all kinds of the mentioned switchings is about the sound transit time (T_S~10μs), which yields the estimation for potential switching speed up to $S \sim (N_A \cdot N_\Theta \cdot N_f \cdot N_P)/T_S \approx 10^{13}$ switch/sec. This potential in switching capability of an AOM seems very high . Unfortunately it is a big problem to realize all its profits because of the absence of special algorithms and related pre- and post- electronic circuits. One of the appropriate ways for that is the design of systems with spatially distributed signals inside multichannel AOMs. The most progressive way is the implementation of the recently developed planar AOMs based on Ti-diffused $LiNbO_3$ waveguides and of surface acoustic wave (SAW) propagation [2]. In this case the additional advantages like the simplicity and the reproducibility of the planar technology, the small sizes and advanced compatibility with ordinary electronics should be acheved.

This paper takes deals with the analysis of the potential characteristics of some prospective schemes of AO switching devices based on multichannel planar AOMs, directed to implementations in digital matrix-vector multipliers, associative optoelectronic memory and optical channel commutators for fiber-optic telecommunications and optoelectronic supercomputers.

2. Switching rate analysis for planar acoustooptic modulators

In [2] very high performances have been demonstrated for for special a type of AOM based on the light guided modes interaction with SAW under its collinear propagation in planar Ti-diffused waveguides in YX-$LiNbO_3$, which results in the radiative substrate mode with 90°-rotation of the plane polarization (see Fig. 1). Under fixed light frequency ν_0 it showed a very high resolution of radio (sound) frequencies for a set of moderate other parameters in multichannel planar AOMs (λ_0=0.63μm): the central frequency $f_0 \approx$ 550MHz, the bandwidth $\Delta f \approx 50 \div 100$MHz, the radio frequency resolution $\delta f_S \geq 50$kHz the SAW transit time $T_S \leq 20$μs, the number of resolvable sites of frequency $N_S \approx 10^3$, the linear dynamic range for input signals DD≤35 dB and the diffraction efficiency $\eta_0 \approx$ 10%/W [3,4]. The simple estimation shows that it supports the switching speed up to $S_1 \approx N_S/T_S \sim 10^8$ operations per second. The expanded version with many light frequencies N_L

(so called optical frequency multiplexed regime) with $N_L \approx N_S \sim 10^3$ will obviously demonstrate more enhanced switching capability $S_{1max} \sim 10^{11}$ switch./sec. (see Fig. 2). According [3] it might be simply designed up to 20 separated AO channels on one chip of $LiNbO_3$, which means the respective growth of the switching capability limit up to $S_{20max} \sim 10^{13}$ switch./sec.

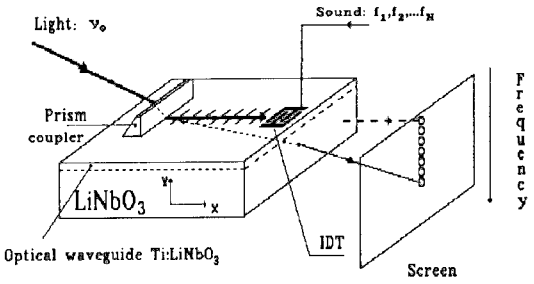


Fig.1. Single channel planar spatial frequency AO switch on base of Ti:$LiNbO_3$ with collinear wave interaction.

Under all of earlier estimations it has been proposed that the time for one command to perform many parallel operations is limited by the transit time of SAW ($T_S \sim 10\mu s$), but sometimes it might be changed to the light transit time $T_L \sim 0.1ns$ with the proportional growth of the potential switching capability up to 10^{16} switch./sec.

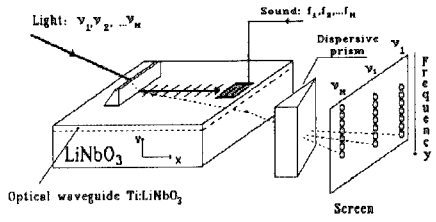


Fig.2 The same as Fig.1 with multiwavelength optical input.

Here we shall consider a several commutator schemes optimized for utilizing planar multichannel AOMs. Fig. 3 shows the scheme for an AO digital matrix-vector multiplier based on the DMAC-algorithm in the spectrum domain [5]. It works under the special relations between the multifrequency light and sound waves (when the set of light frequencies $\nu_i = \nu_0 q^{-i}$, with i=1,2...,N, is fitted to the SAW frequencies $f_i = f_0 q^i$ with i=1,2...,N, where q is a·base of the harmonic sequence).The result of the DMAC for two spectral binary codes to perform digital multiplication of the "optical matrix" $\tilde{a}_{mk} = a_{mk}(\nu_1, \nu_2 \ldots \nu_N)$ and the "sound vector" $\tilde{b}_k = b_k(f_1, f_2 \ldots f_N)$ .The output

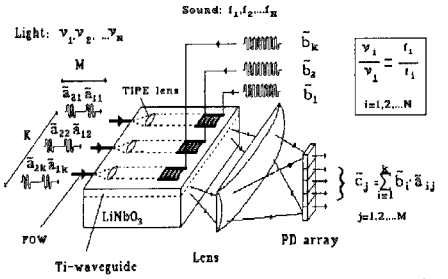


Fig.3 Digital matrix-vector multiplier scheme using multichannel planar AO modulator on Ti:$LiNbO_3$ with multifrequency SAW and multiwavelength optics.

shows consecutively in time the resulting vector components $\tilde{c}_j = \sum_{i=1}^{N} \tilde{b}_i \tilde{a}_{ij}$ in mixed code presentation. Looking for optimal matching of all related signals and taking into account the real planar AOM parameters one can expect the interbit switching capability for this scheme as high as ~ 10^{12} oper./sec. with the light pulses duration $T_L \approx 1ns$ [5]. It seems to be very attractive as an arithmetical processor for fast solution of linear algebraic multiple equations with rates up to 10^{10} multiplication and adds per second for 16-bit coded operands. But recently this kind of development is seriously limited without adequate speed of ADC function for decoding of the AOM outputs (see, for example, [6]).

If the optical input of this scheme is transformed to the single light frequency ν_0 (Fig. 4) it gives "immediately" (after the light transit time T_L) the results of the inner vector

products of the "optical vector" $\{a_k\}$ by the series of "acoustical vectors" $\{b_k\}_j=\{b_k(f_1,f_2...f_1)\}$. With T_{1min}~0.1ns it gives maximum capability as high as $S_{max} \sim (N_s \cdot K(K-1))/T_{L\,min} \approx 4 \cdot 10^{15}$switch/sec. This type of scalar vector multiplier is very compatible to perform the fast comparison of the "optical word" with the archives of many "acoustical words" inside of the associative memory device for a new generation of the optoelectronic supercomputers [7].

The simplest high fidelity planar AO switch for associative memory systems might be built on the scheme of Fig. 5. It uses the most efficient AO transition inside of planar waveguide between a pair of the lowest optical modes with orthogonal planes of polarization $TE_0 \leftrightarrow TM_0$ (η_0=1%/mW). The diffracted light might be selected from the input one by using thin metallic film polarizer at the ends of every AO channel so the j-output will consist of the result of the desirable inner product $c_k = \sum_{i=1}^{N} b_k(f_i)a(\nu_i)$ if the optical and sound frequencies satisfy the ratio (ν_i/f_i) = const. The capability limit for the case of 20 channel AOM is estimated as $S \approx 20 \cdot 10^3/10^{10} \approx 2 \cdot 10^{14}$switch./sec. This kind of the AO commutators seems be very reliable in a form compatible with fiberoptic and integrated optic peripheral devices.

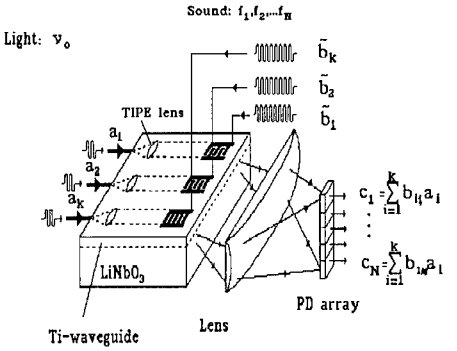


Fig.4 Scheme of AO commutator for associative memory system utilising frequency multiplexed acoustic signals.

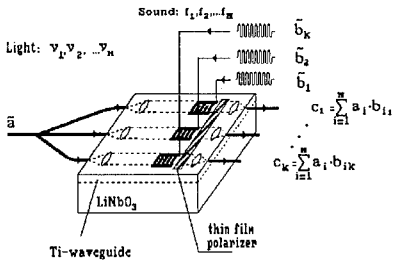


Fig.5. The same as Fig.4 utilizing frequency multiplexed light signals.

The scheme of Fig. 5. must work also in the regime of the optical channel commutation generally applicable for optoelectronic computers and fiber optic telecommunication systems. In fact, if K=N each input optical channel defined by frequency ν_i might be driven to any of spatial outputs by appropriate choice of SAW sound frequency f_i (because of n_i/f_i=const). That means the distribution of all sound frequencies between all AOM channels will fix respectively the distribution of all frequencies of the light between output channels.

Following [2] it is easy to find out the optimal relations of the planar AO commutator parameters:

$$\eta_0 = \sin^2\left[M_2 \cdot F \cdot P_S \frac{v_S T_S k}{p \cdot d}\right]^{1/2}, \qquad (1)$$

$$k = p \cdot f_0 T_S, \qquad (2)$$

where $p = \frac{\Delta f}{f_0} = \frac{\Delta \nu}{\nu_0}$ is the fractional frequency bandwidth of AOM, M_2 is the material AO merit of quality, v_S and f_0 are the velocity and the central frequency of SAW, F is the overlapping integral of sound and light waves, P_S and d are the power and the beam

width of SAW. From (1) and (2) one can estimate the appropriate set of parameters. For example in the case of a 20-channel AOM based on Ti:$LiNbO_3$ with driving power $P_S \approx$ 0.1W per channel it gives the switching time $T_S \sim 0.1 \mu s$. This scheme with planar AOMs suffers from excess optical losses (K-times) due to optical splitting into K-channels on the input side and the limited number of the planar channels per chip ($K_{max} \approx 20$) . Both of these parameters probably might be developed with the implementation of channelized AOMs based on directional mode coupling in Ti-diffused strip-line optical waveguides combined with TIPE acoustic waveguides on $LiNbO_3$ (see, for example, [8,9]). In this case the number of AO channels per chip is limited by the diameter of standard fibers (not less ~ 125μm). One can estimate as $K_{max} \approx 160$. According to [8] the appropriate length of the AO strip - type coupler L_C ~ 1 cm, which yields the limited number of the sequential couplers along the chip as $K_2 \approx 8$. So one can propose the scheme shown in Fig. 6 with the compromised number of parallel ($K_1 \leq 20$) and sequential ($K_2 \leq 8$) AO couplers to perform arbitrary optical commutations for 160 channels during time $T_S \approx 3\mu s$.

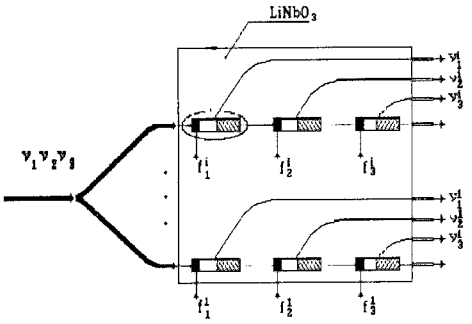


Fig.6. Scheme of space division commutator for frequency multiplexed signals based on AO directional coupler utilizing TIPE channel waveguides in $LiNbO_3$.

3. Conclusion

There were shown above very high limits of the potential switching capability in the space and wavelength domains for a few schemes of AO commutators based on Ti:$LiNbO_3$ planar waveguides with good promises for some actual applications. The results of the paper showed the necessity of the realization and the detailed investigation of full scale experimental prototypes of the proposed AO commutators to find out the supposed restrictions due to the second order effects, the technological imperfections, etc.

4. Acknowledgments

The authors wish thank Dr. E.M.Korablev and Dr. O.A.Byshevsky-Konopko for helpful discussions and Mr. V.N.Chesnokov and Mrs. I.A.Prusova for great help with the preparation of this presentation.

References

[1] Magdich L N and Molchanov V Ya 1978 *Acoustooptic devices and applications* (Moscow: Sov.radio).
[2] Proklov V V and Korablev E M *Proc. IEEE 1992 Uitrasonic Syposium, Tucson AZ* (New York: IEEE,Inc.) **1** 173-8
[3] Proklov V V 1986 *Proc. Intern. Symp. on Surface waves in solids and layered structures, Novosibirsk* (Moscow, VINITI Press) **1** 148-63
[4] Proklov V V 1992 J.Acta physica slovenia **42** 131-41
[5] Gulyaev Yu V Proklov V V Sokololvsky S V and Bashlakov A U 1991 *Int. J. on Optical Computing* **1** 141-52
[6] Goutzoulis A P 1986 *Proc. SPIE on Optical Information Processing* **639**
[7] Burtsev V S 1992 *Tech. Digest ICO Int. Top. Meeting on Optical Computing, Minsk* (Minsk:Bel.Acad.Sci.Publ.) **1** 30 B8
[8] Goto N Miyazaki Y and Akao Y 1983 *Trans. IEICE Japan*, **E 66** 442-9
[9] Hinkov V P Opitz R and Sohler W 1988 J. of Lightwave technology **6** 903-8

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 British Crown Copyright

An acousto-optic FM demodulator with inherent AM and laser noise rejection

P Brooks

EE Department, Royal Naval Engineering College, Manadon, Plymouth, PL5 3AQ, UK

C D Reeve

SECEE, University of Plymouth, Drakes Circus, Plymouth, PL4 8AA,UK

Abstract. This paper describes an acousto-optic FM demodulator that offers improvements in noise rejection performance over early systems. Theory is developed and the system is modelled and results are presented. It is shown that the system is capable of giving up to 25 dB of AM rejection.

1. Introduction

The ability to demodulate FM signals has many applications, especially in the area of communications and equipment test. Conventional methods such as the PLL have provided general purpose solutions but in many applications they do not have the required performance. The acousto-optic FM demodulator [1,2] offers great promise but its potential performance has been limited by AM signal content and, more importantly, laser noise [3]. The system suggested here overcomes these limitations for only a modest increase in system complexity.

2. System architecture

A schematic diagram of the system is shown below in figure 1.

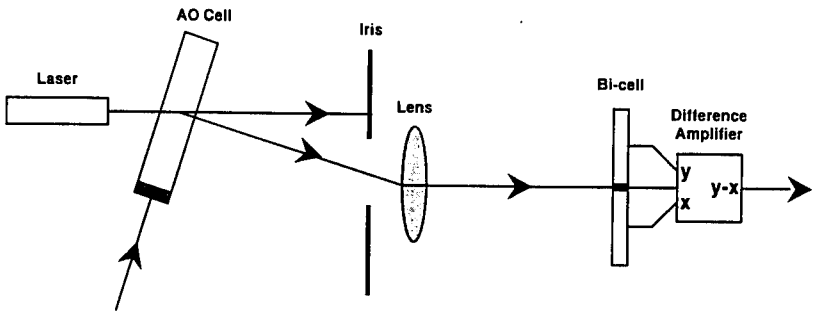


figure 1 System Schematic

It consists of a laser, which illuminates an aperture in an acousto-optic (AO) cell. The AO cell is driven by the signal of interest via an RF amplifier. The AO cell deflects the AO cell through an angle proportional to the frequency of the applied signal. The deflected beam is focused onto a photodetector assemble consisting of a bi-cell photodiode via a simple lens. The bi-cell is arranged so that when an undeviated carrier is used to drive the AO cell the deflected beam illuminates each half of the bi-cell equally. Any change in frequency will cause the beam deflection to change and hence the illumination on each half of the diode will vary.
The resulting photocurrent from each half of the bi-cell is amplified by a transimpedance amplifier and the resulting voltage outputs go to a sum-difference amplifier. The resulting outputs give the AM and FM contents of the signal respectively.

3. System Analysis

The deflected intensity profile from the AO cell can be described by a truncated Gaussian thus:

$$P = \int_{-2\sigma_t^2}^{+2\sigma_t^2} \int_{-2\sigma_t^2}^{+2\sigma_t^2} \frac{P}{2\pi\sigma_t^2} e^{-\frac{x^2+y^2}{2\sigma_t^2}} dxdy$$

Where P is the total power in the deflected beam, σ_t is the beam width standard deviation of the beam at the photodetector assembly. P is a function of both laser intensity and AM modulation on the carrier. Since our detector is a bi-cell device and is positioned such that the undeviated beam falls on each have of the bi-cell equally then

$$P_1 = \int_{-2\sigma_t}^{+2\sigma_t} \int_{-2\sigma_t}^{\delta x} \frac{P}{2\pi\sigma_t^2} e^{-\frac{x^2+y^2}{2\sigma_t^2}} dxdy$$

and

$$P_2 = \int_{-2\sigma_t}^{+2\sigma_t} \int_{\delta x}^{+2\sigma_t} \frac{P}{2\pi\sigma_t^2} e^{-\frac{x^2+y^2}{2\sigma_t^2}} dxdy$$

where P_1 and P_2 is the optical power on each half of the bi-cell and δx is the beam displacement at the detector caused by a change in applied frequency δf and is given by

$$\delta x = \frac{\lambda l}{V_a} \delta f$$

Where λ is the wavelength of the laser, V_a is the acoustic velocity in the AO cell and l is the focal length of the lens between the AO cell and detector assembly.
For small frequency deviations and hence small beam displacement ($\delta x \rightarrow 0$) we can express each integral as a simple linear function [3] thus allowing us to express the optical power in each half of the diode thus

$$P_1 = \frac{P}{2}(1 + K)$$

and

$$P_2 = \frac{P}{2}(1 - K)$$

where K is the ratio of the signal power ($P_{(1\ or\ 2)}$-P/2) to the background invariant power P. If the signal from each half of the photodiode is amplified by a transimpedance amplifier of gain G and then the difference voltage (ΔV) taken we obtain.

$$\Delta V = GR_\lambda PK$$

Where R_λ is the photodiode responsivity measured in A/W, assumed to be equal for each half of the photodiode.

The value of K can be expressed as a function of acoustic transit time τ and frequency deviation [3] thus

$$K = \sqrt{2\pi}\tau\delta f$$

Therefore

$$\Delta V = GR_\lambda P\sqrt{2\pi}\tau\delta f$$

The output signal from the difference amplifier is a direct and linear function of both the applied frequency and the incident optical power.

The above relationship can be compared to the output signal from a single knife edge system, given by

$$V = GR_\lambda \, P/2(1+K)$$

For small deviations ($\delta x < \sigma_t$) then $K << 1$ and the effect of variations in optical power P, a function of laser intensity and AM modulation on the AO drive signal will have a far greater effect than in a bi-cell scheme. We define the AM rejection as the inverse of the AM gain, the AM rejection ratio (AMRR) of the bi-cell scheme relative to the single knife edge system is

$$AMRR = \frac{GR_\lambda \, P/2(1+K)}{GR_\lambda PK} = \frac{1}{2K}$$

The AMRR is inversely proportional to applied frequency. This is intuitively correct: with no frequency deviation both diodes receive equal power, any variations due to intensity fluctuations or AM signal content are common mode and hence cancel. If a frequency offset occurs the beam will be displaced and there will be a difference in the respective photocurrents and so any intensity variations cannot be totally removed.

If we express the AMRR in terms of dB and we use 'normal' system parameters (τ = 800 ns) then we can write

$$AMRR(dB) = 54 - 10\log(\delta f)$$

This remains positive for frequency deviations of 325 kHz, this is close to the linear operating range of the system [2].

Total independence from AM signal content may be achieved by taking the ratio of the difference (ΔV) and sum (ΣV) signals thus

$$\frac{\Delta V}{\Sigma V} = \frac{GR_\lambda PK}{GR_\lambda P} = K$$

This has no AM content and is a direct function of FM deviation.

This can be achieved by an analogue multiplier configured as a divider but this restricts both system bandwidth and frequency resolution.

4. Experimental Results

The system used consisted of a stabilised He-Ne laser illuminating an Isomet OPT-1 TeO_2 cell the deflected beam was imaged onto a bi-cell diode via a simple f-f lens system. The

photocurrent from each diode was amplified by a low noise transimpedance amplifier and the sum and difference signals taken on a digital oscilloscope. The system was aligned such that a 45 MHz carrier caused a beam deflection which illuminated each half of the bi-cell equally. Any change in frequency caused a change in the respective illumination intensities which was measured. The 45 MHz carrier was AM modulated at 1 MHz while the carrier frequency was manually shifted from 45 MHz in 500 Hz steps. The resulting level of AM breakthrough was measured and the AMRR was calculated.
Figure 2 shows the AMRR of a practical system compared to theoretically modelled results. At low deviation (< 1 kHz) we see a limiting value of around 20 dB falling to 10 dB at 25 kHz.

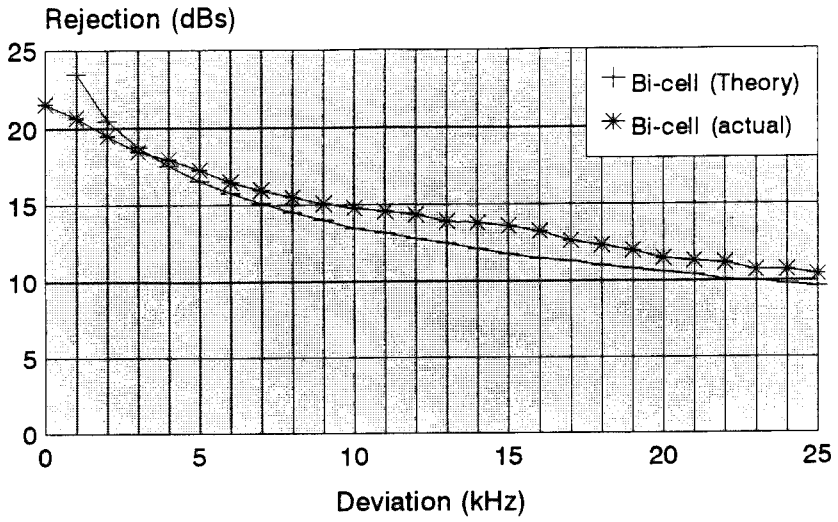


figure 2 Theoretical and Experimental Results

5. Conclusions

It has been shown, both in theory and practice that the Bi-cell acousto-optic FM demodulator is capable of offering high levels of AM and intensity noise reduction. For narrow deviation signals the high levels of AMRR (> 20 dB) is far greater than the 3 dB offered by the original system. The difference/sum ratio system offers higher levels of AMRR at the expense of increased cost, reduced bandwidth and decreased frequency resolution. At present, the difference amplifier gives good performance with only a very small increase in price over the original single diode system.

References

[1] POON T C & PIEPER R J. 1985 *IEEE Trans Ed.* **Vol. E-28**, No.1.

[2] REEVE C D 1989 *RNEC Research Report* **89026**

[3] BROOKS P & REEVE C D *RNEC Research Report* **93028**

Copyright © HMSO, London, 1994

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Integrated-optic commutator for high-speed local communication network with cross-talk minimization in optical directional couplers

Igor G.Voitenko and Toyohiko Yatagai*

Institute of Applied Optics, Byelorussian Academy of Sciences, 212793 Mogilev, Belarus
*Institute of Applied Physics, University of Tsukuba, Tsukuba, Ibaraki 305, Japan

Abstract. The performance of an 8×8 polarisation-independent Ti:$LiNbO_3$ switch matrix is described on the basis of a new type of directional coupler with its full complement of the local area network. A reverse *Db* electric field configuration is used for both the "bar" and "cross" state and no attempt is made to tune individually the two states. Some possible integrated-optical devices and the optical versions of multistage nonblocking interconnection networks are discussed.

1. Introduction

The importance of the interconnection networks in tightly coupled multiprocessor systems has been pointed out by many authors [1-3]. These networks become important when the high number of system elements to be connected makes the architectures based on a shared bus inefficient. In particular, in large structures, multistage networks, increasing according to $O(N\log_2 N)$, turn out to be important [4]. The possibility of achieving computer systems, or parts of them, with optical components, seems to be particularly attractive due to the various advantages, such as electromagnetic interference immunity, intrinsic parallelism of the optical systems, and the consequent large space-bandwidth product [5].

Optical couplers on the basis of tunnel coupled waveguides are the key elements in a number of modulation, commutation and logic schemes. In designing such schemes the problem of a cross-talk minimization holds. It was shown [6] that the cross-talk decreased together with the waveguide coupling. But this decrease leds to the sharp increase of a device length. The cross-talk in electrooptic switches can be compensated also by using a two-sectional (Db) electrode scheme with the sections of different length [7]. However, to get the complete switching from a waveguide to another waveguide (Figure 1(a)), the opposite polarity voltage is needed. Such peculiarity restricts the device usefulness in some schemes. The cross-talk can be minimized through optimization of adjusting sections in which bent single-mode waveguides are used [8,9]. A loss inequality of the even and odd modes at the adjusting sections, however, was not taken into account.

In this paper we show that this inequality is a main cause of the cross-talk.The dependence of the switch cross-talk on the adjusting section parameters is analyzed. An adjusting sectional scheme which permits to eliminate the crosstalk is proposed. A device on a base of diffused channel waveguides is analyzed.

2. Cross-talk minimization and directional coupler design

2.1 Theory

We deal with weakly inhomogeneous waveguides. Modes of the waveguide are described approximately by the scalar equation

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial z^2}+k_0\psi(x,y,z)=0 \tag{1}$$

which is solved for $y<0$ with the boundary condition $\psi_{y=0}=0$ ($y=0$ is the substrate surface equation, k_0 is wave number of free space). A waveguide structure with the plane of symmetry $x=0$, as shown in Fig.1, and the waveguides at input $(z<0)$, output $(z>l+2L)$ and at the coupling section be parallel and single-mode, as shown in Fig.1(b). We use following expressions for the fields of the even and odd modes of the waveguide structure:

$$\psi_{ei}=S_i\varphi_{ei}(x,y)\langle\varphi_{ei}^2\rangle^{-0,5},\ \psi_{oi}=A_i\varphi_{oi}(x,y)\langle\varphi_{oi}^2\rangle^{-o,5},\ z<0, \tag{2}$$

$$\psi_e=S\,\varphi_e(x,y)\langle\varphi_e^2\rangle^{-0,5},\ \psi_o=A\,\varphi_o(x,y)\langle\varphi_o^2\rangle^{-o,5},\ l+L\ z<L, \tag{3}$$

$$\psi_{et}=S_t\varphi_{et}(x,y)\langle\varphi_{et}^2\rangle^{-0,5},\quad \psi_{ot}=A_t\varphi_{ot}(x,y)\langle\varphi_{ot}^2\rangle^{-0,5},\quad z>2L+l, \tag{4}$$

where S_i and A_i are amplitudes, $\langle f^2\rangle^{-0,5}$ are normalized constants, $\varphi_i(x,y)$ are transversal distributions of mode fields.

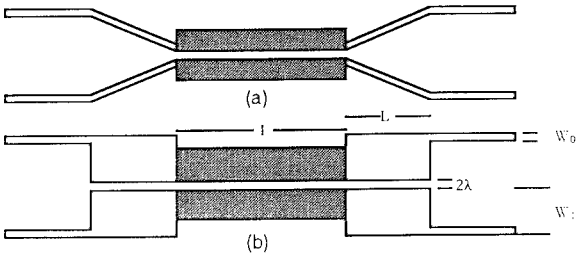


Fig.1. Switch structure. (a):traditional structure; (b):novel structure.proposed.

We consider a device in terms of diffused channel waveguides. We estimate propagation constants: $\beta_1=(h_i^2k_o^{-2}-\varepsilon_o)\Delta\varepsilon^{-1}$, which are shown in Table 1. For weakly guiding waveguides, we have

$$h_i - h_j = k_o(\beta_i-\beta_j)\Delta\varepsilon(2\sqrt{\varepsilon_s}^{-1}) \tag{5}$$

Using Table 1 and Eq. 5, we come to $r_o(\phi)$ and $r_{e1}(\phi)$ dependence presented in Fig.2.

Table 1 Propagation constants.

$\beta_e-\beta_o$	$\beta_{ee}-\beta_{oe}$	$\beta_{ee}-\beta_{eo}$	$\beta_{eo}-\beta_{oo}$	a_e	a_o
0.0170	0.1500	0.0046	0.1842	0.8540	0.8890

The condition $r_{o1}=r_{e1}$ =0,983 holds at F=2,8664. In the case $\Delta\phi_1$=-0,162 and PP_o^{-1}=0,935 (-0,29dB). Sufficient conditions of zero cross-talk are $r_{o1}=r_{e1}$, $r_{o3}=r_{e3}$, $\exp[i(\Delta\phi_1-\Delta\phi_3)]=1$. In addition x =0 for the "cross" state and for the "bar" state x is found from an equation $\tau_{11}+\exp(-i\Delta hl)\tau_{11}^{*}\Delta=0$.

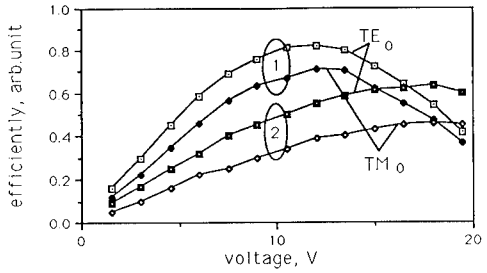


Fig. 2 Optical output power of the directional coupler vs. input voltage.
1 -novel type of directional coupler; 2 -traditional type of directional coupler.

2.2 Experiment

In our experiments, titanium is deposited in a high vacuum plant through vapour deposition to a thickness of 40nm The *Ti* indiffusion is carried out at 1040°C for 4*h*. *MgO* is deposited nonreactively through vapour deposition at a rate of 20 nm/min to avoid dissociation. The strip thicknesses are 30nm indiffusion is carried out for 2*h* at 910°C with a flow of dry synthetic air at a rate of 2,5 l/min.

The fabrication parameters for the two-mode waveguides were: titanium strip width W=4mm and coupler length L=8mm. In the initial experiments the two-moded intersection region was realized by increasing the width of the center titanium stripe to $2W$ and keeping the titanium thickness constant. In the center titanium stripe a *MgO* stripe (2mm width and thickness 30nm) was deposited and the second indiffusion was made.

The directional coupler was fabricated using quartz buffer layers and *Cr-Au* electrodes. The polished waveguide edge faces were butt-coupled to polarization-preserving fibers using index-matching epoxy. The far-end cross-talk attenuation or on/off ratio varies between 19 and 29dB. For the 8mm coupling length device and 2,5mm gap, more then 80% efficiently was obtained (wavelength is 1,3mm and 12V switching voltage). Optical output powers are plotted for TE and TM modes, with waveguide gap of 2,5mm in Fig.4.

3. Optical multistage interconnection network architecture

A popular method for realizing the switching or commutator fabric is the Batcher or Batcher-banyan network [10]. This multistage network where each stage performs a fixed

permutation on the incoming lines, and then routes them through a column of 2x2 switching elements. According to the standard bit-controlled routing algorithm for these networks. switches in the first stage are controlled by the most significant bit of the destination tag, those in the second by next bit, and so one. By convention, the control bit of the top input decides the state of the switching elements. If it is zero, the switch is the straight connection state; otherwise it is crossed. These switches route the input with the lower tag to the top output and the one with the higher tag to the bottom output.

A strictly nonblocking crossbar switch can always establish a connection between an input and output without prejudice to connections already made. This architecture requires n switches. In a guided wave device with a large aspect ratio there are $2n-1$ switch cell lengths from end to end of the switch. Such an architecture can have fewer switches and be shorter then the equivalent strictly nonblocking switch. We have designed a switch based on the 23 switch 8x8 design shown in Fig.3.

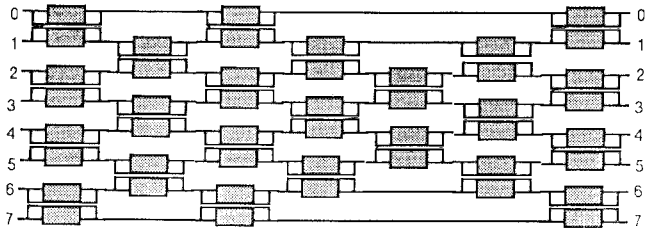


Fig.3. Nonblocking 8x8 integrated-optical switch matrix design.

References

1. K. E. Batcher, Sorting networks and their applications, in 1968 Spring Joint Computer Conf., AFIPS Proc., **32** pp. 307-314 (1968).
2. T.Y. Feng, A survey of interconnection networks, IEEE Comput. Mag., **4**, pp. 12-27, Dec. 1981.
3. D.P. Agraval, Advanced computer architecture, Tutorial Text, Computer Society Press, 386 pp., 1986.
4. H.S. Stone, Parallel processing with the perfect shuffle, IEEE Trans. Comput., **C-20**, no. 2, pp. 153-161, Feb. 1971.
5. E. Marom, Integrated optics switch array network decomposition, Opt. Comm., **58**, pp. 92-94, 1986.
6. K.L.Chen and S.Wang, Cross-talk problems in optical directional couplers, Appl. Phys. Lett., **44**, no.2, pp. 166-168, 1984.
7. J.P. Donelly, L.A. Molter and H.A. Haus, The extinction ratio in optical two-guide coupler Db switches, IEEE J Quant. Electron., **25**, no. 5, pp. 924-932, 1989.
8. H. A. Haus and N. A. Whitaker, Elimination of cross-talk in optical directional coupler, Appl. Phys. Lett., **46**, no. 1, pp. 1-3, 1985.
9. J. P. Weber, L. Thylen and S. Wang, Cross-talk and switching characteristics in directional couplers, IEEE J. Quant. Electron., **24**, no. 3, pp. 537-548, 1988.
10. D. H. Lawrie, Access and alignment of data in an array processor, IEEE Trans. Comput., **C-24**, pp. 1145-1155, Dec. 1975.

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Optical Characteristic of Photorefractive Three-dimensional Waveguides in Lithium Niobate

Kazuyoshi Itoh, Osamu Matoba, and Yoshiki Ichioka

Osaka University, Department of Applied Physics
2-1 Yamadaoka, Suita, Osaka, 565, Japan

Abstract. We present in this paper an experimental and numerical analysis of photorefractive waveguides fabricated in a lithium niobate crystal. A parametric model of 3-D refractive index distribution of the photorefractive waveguide is used. The maximum refractive index change is estimated at $\Delta n_{max} = 1.2\times10^{-3}$ by comparing the near-field patterns in the rear face of the crystal that are calculated by the beam propagation method with the experimental observations.

1. Introduction

We proposed a novel approach to three dimensional optical interconnections for optical neural networks[1]. Photorefractive waveguides are fabricated by simply focusing and scanning a laser beam in the photorefractive material as shown in Fig. 1. Figure 2 shows our concept to implement high-density optical interconnections for neural networks. Resultant waveguides with variable index profiles may be used for the optical dynamic interconnections. We believe in the possibility that high-density optical interconnections with a self-organization ability be realized by this approach. In the preliminary experiment, the photorefractive waveguides were successfully fabricated by scanning a focused laser beam in a lithium niobate (LN) crystal and the guided light on the rare face of the crystal was clearly observed. We have proposed two dimensional model of refractive index distribution of the photorefractive waveguide.[2] The accurate knowledge of the three dimensional refractive index distribution of the photorefractive waveguide is necessary to evaluate the optical characteristics of the waveguide. However, it is difficult to measure the complicated structure of the three dimensional refractive index distribution inside the crystal. In this paper, we propose a three dimensional model for the refractive index change. The maximum refractive index change is determined by comparison between experimental and numerical results of the near-field patterns of curved waveguides.

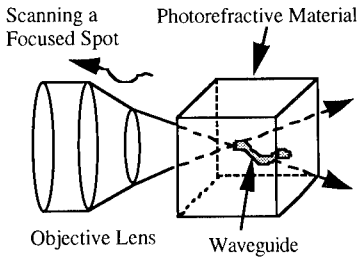


Fig. 1 Fabrication of 3-D photorefractive waveguide.

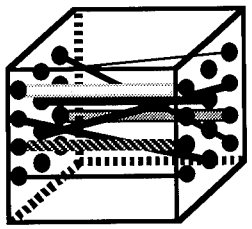


Fig.2 High-density optical interconnection using photorefractive waveguide.

2. Model of 3-D refractive index distribution of photorefractive waveguide

We propose a parametric model of a three dimensional refractive index distribution of the photorefractive waveguide that is based on the measured refractive-index distribution [1]. Figure 3(a) shows the transverse refractive index distribution in the plane perpendicular to the propagating direction of a straight waveguide. The field of view of this figure covers a rectangular area of approximately 70 μm × 70 μm. In Fig. 3(a), the transverse axis is parallel to the *c* axis and the longitudinal axis is parallel to the *b* axis of the LN crystal. Bright areas indicate those of the higher refractive index and the dark areas those of the lower index. A pair of waveguides are fabricated at both sides of the exposed region along the *c* axis.

Let the *x*-axis be parallel to the *c* axis and the *y*-axis be parallel to the *b* axis of the LN crystal, and the *z*-axis coincide with the optical axis of the experimental optical system. The proposed model for the refractive index distribution of a photorefractive waveguide in the LN crystal is,

$$\Delta n(x,y,z) = \alpha \sum_{i=1}^{3} a_i \exp\left\{\frac{-(x+c_i+f(z))^2}{b_i^2} + \frac{-(y+e_i+g(z))^2}{d_i^2}\right\}, \qquad (1)$$

where α is a variable parameter and depends on the power of fabrication beam, a_i, b_i , c_i , d_i, and e_i are determined by fitting $\Delta n(x,y,z)$ to the experimental data, $f(z)$ and $g(z)$ denote the deviation of the focused spot with regard to the *x*-axis and *y*-axis, respectively. Figure 3(b) shows the model refractive index distribution fitted to the experimental result shown in Fig. 3(a). By changing f(z) and g(z), we can design varieties of 3-D waveguides.

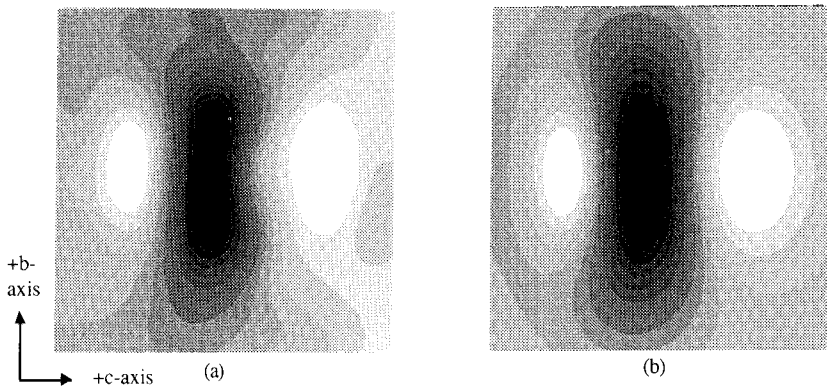


Fig. 3 Refactive index distributions in the plane perpendicular to the propagating direction of a straigth waveguide; (a) the experimental result, and (b) the model fitted to the experimental result are displayed.

3. Experimental and numerical results

The experimental setup is shown in Fig. 4. A LN crystal that nominally contains impurities less than 10 ppm (The Fe ion is less than several ppm.). The thickness of crystal is approximately 2.0 mm. A linearly polarized argon ion laser beam (λ=514.5nm) is used to fabricate the waveguide and to excite the guided light. When we fabricate the photorefractive waveguide, the electric field vector of the argon ion laser is set perpendicular to the *c* axis (ordinary ray) . When the light is to be guided, a half-wave plate is inserted to excite the extra-ordinary ray and the optical beam power is reduced to approximately 1/100 than that of the fabrication beam. An argon ion laser beam is focused by a microscope objective lens L1. The

numerical aperture of the focused beam is approximately 0.09. To fabricate the 3-D waveguide, the longitudinal position of the focus and the lateral position of the crystal are controlled by the translators. The microscope objective lens L2 is placed behind the crystal for observation of near-field pattern on the rear face of the crystal. The near-field pattern is obtained by the CCD image sensor.

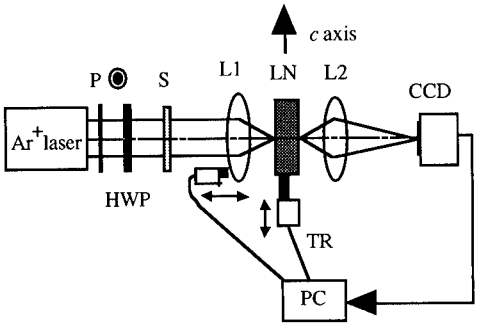


Fig. 4 Optical system for fabricating and testing the photorefractive waveguides; P denotes polarizer, S shutter; HWP half wave-plate, TR translator, PC personal computer, L's microscope objective lenses, respectively.

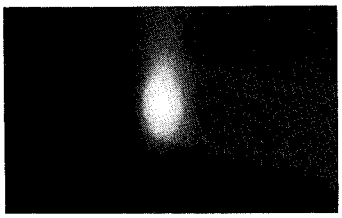
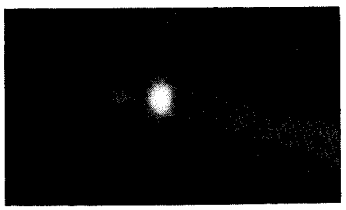
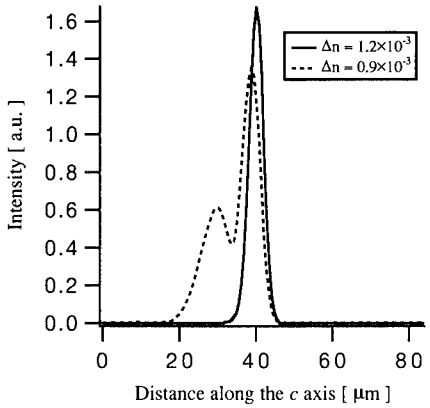


Fig. 5 Near-field pattern of the curved waveguide.

(a)



(b)



(c)

Fig. 6 Numerical results of near-field pattern in the case of (a) $\Delta n_{max} = 1.2\times10^{-3}$ and (b) $\Delta n_{max} = 0.9\times10^{-3}$, and (c) a simultaneous plot of these intensity distributions taken along the central line parallel to the *c* axis. The solid and broken curves indicate patterns taken with $\Delta n_{max} = 1.2\times10^{-3}$ and $\Delta n_{max} = 0.9\times10^{-3}$, respectively.

We compared the results of a numerical analysis of near-field patterns of the curved waveguides with the experimental results. In the curved waveguide with the insufficient

refractive index change, a part of guided light may be radiated neat the curved portion. Thus, an accurate index distribution of the waveguide can be obtained by this comparison. We fabricated an s-shaped waveguide whose transverse displacement, $|f(z_{in}) - f(z_{out})|$, between the input(z_{in}) and output(z_{out}) ends of the waveguide is 30μm [$g(z_{in})$=$g(z_{out})$]. We used the waveguide that is located at the negative side along the x-axis. The experimental result of near-field pattern of the curved waveguide is shown in Fig. 5. The field of view of this figure covers a rectangular area of approximately 50 μm × 85 μm. The radiation loss along the x-axis is almost negligible. In numerical simulation, various near-field patterns are calculated using the beam propagation method[3] on the assumption that the incident beam is a Gaussian beam. Figure 6 shows the numerical result of near-field patterns of curved waveguide in the case of $\Delta n_{max} = 1.2\times10^{-3}$ and 0.9×10^{-3} , respectively. The intensity distribution of guided beam in the curved waveguide on the x-z plane is shown in Fig. 7. In the case of the lower value of Δn_{max}, a part of the guided light is radiated along the x-axis at the curved portion of the waveguide. The result of $\Delta n_{max} = 1.2\times10^{-3}$ shows agreement with the experimental result except that the width of the intensity profile along the x-axis is narrower than experimental result. The maximum refractive index change, Δn_{max}, is estimated to be equal to or larger than 1.2×10^{-3}.

4. Conclusion

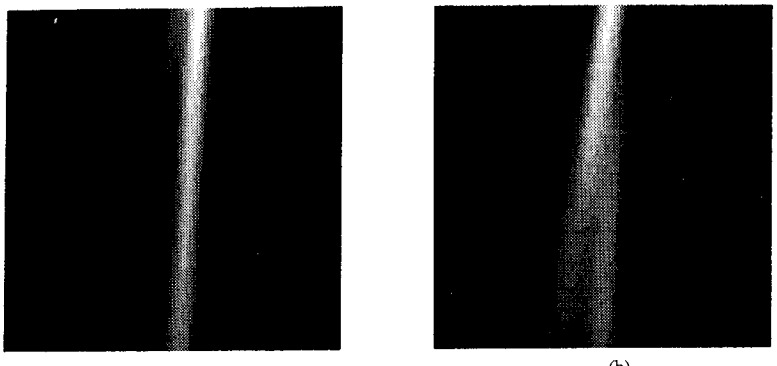
We proposed the three dimensional refractive index distribution of photorefractive waveguide and estimated the maximum refractive index change by making a comparison between the experimental results and numerical analysis. The comparison showed that Δn_{max} is equal to or larger than 1.2×10^{-3} .

Acknowledgment

This research has been supported in part by a grant-in-aid from the Ministry of Education, Science, and Culture, Japan.

References

[1] Itoh K, Matoba O, and Ichioka Y 1994 *Opt. Lett.* **19** 652-654
[2] Matoba O, Itoh K, and Ichioka Y *Opt. Rev. (to appear)*
[3] Feit M D and Fleck, Jr J A 1978 *Appl. Opt.* **17** 3990-3998



(a) (b)

Fig. 7 Intensity distributions of guided beam in the x-z plane; (a) $\Delta n_{max} = 1.2\times10^{-3}$ and (b) $\Delta n_{max} = 0.9\times10^{-3}$

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Fiber-optical loop memory structure for optoelectronic computer

V.A. Pilipovich, A.K. Esman, I.A. Goncharenko, V.S. Posedko, I.F. Solonovich

Institute of Electronics, Academy of Sciences of Belarus,22 Lagoiski Trakt, Minsk-90, 220841 Republic of Belarus

Abstract. The method of enhancing the information capacity and reliable storage of information in circulator fiber-optic memory is presented. The method is based on creation of an additional built-in channel with contrary-directed circulation of signals. This channel can be used for transmission of both information and auxiliary signals: address words, clock signals, correcting sequences, etc. The possibility of compensating the information signal losses be means of stimulated Raman scattering is considered.

1. Introduction

Increasing of the information capacity of a fiber-optic memory device (FOMD) is normally achieved either by increasing the length of fiber circulation loop accompanied by added complexity of the synchronization circuit or by using conventional methods of channel multiplication typical of fiber-optic communication links,i.e. spectral,polarization and temporal methods [1,2]. However, closed nature of circulator memory systems makes it also possible to create additional information channels that can be employed for attaining both reliable addressing, synchronization and noise-immune encoding and increased information capacity of the FOMD [3,4]. In this case,the use of traditional multiplication and synchronization techniques is not excluded.

This paper is concerned with the proposed method of increasing the information capacity and reliability of information storage in a circulator fiber-optic memory by creating an additional built-in channel with contrary-directed signal circulation.

2. Organization of the contrary-directed channel

In accordance with the method proposed, two contrary-directed circulation channels - main and additional ones -are organized in a single fiber FOMD loop (generalized structural diagram of the FOMD is shown in Fig.1). Information signal sequence circulates in the main channel, while the additional channel

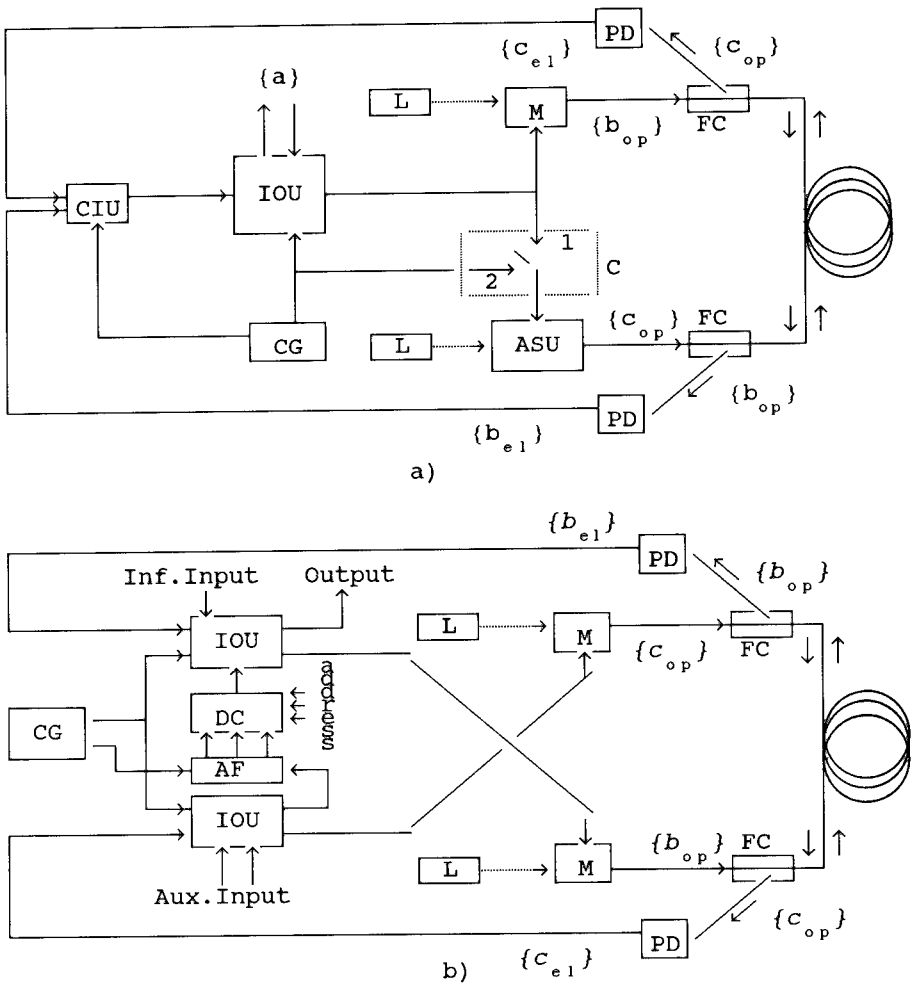


Fig 1. Structural diagram of the fiber-optic memory loop with contrary directed circulation of the signals. Here, PD is photodetector, M is modulator, ASU is unit for creation of auxiliary signals, C is commutator, CIU is unit for correction of information,L is laser,IOU is information input/output unit, FC is fiber-optical coupler, CG is generator of the clock pulses, AF is address former, DC is digital comparator.

can be used to transmit both part of the information sequence and auxiliary signals (verifying, synchronizing and addressing words) formed by the auxiliary signal unit (ASU).In the latter case, optical signals circulating in both channels can be transmitted at both the same and different carrier frequencies.

In the case shown in Fig.1a,the additional channel serves to transmit a verifying correcting sequence and ensures enhanced reliability of storage of the main information sequence. The input information sequence $\{a\}$ is subjected to

convolutional symmetrical coding and is converted to the corresponding sequences of information $\{b_{op}\}$ and test $\{c_{op}\}$ optical signals. In this case,the commutator is switched to position 1 (see Fig. 1a), and the ASU acts as a convolutional encoder consisting of two series connected phase modulators whose control voltage is set equal to zero or half-wave value depending on the value of the information signal.

Separated information and test signal flows arrive simultaneously in the fiber-optic circulation loop through couplers placed on its opposite ends. The contrary-directed flows $\{b_{op}\}$ and $\{c_{op}\}$ also emerge simultaneously at the loop outputs and, being converted to the corresponding sequences of electrical signals $\{b_{el}\}$ and $\{c_{el}\}$, are used for decoding information and correcting errors without any demultiplexing required.The corrected information sequence is again subjected to encoding.

The use of the contrary-directed channel in the fiber-optic circulator makes it possible to increase error-free storage probability owing to a two-fold reduction of the character repetition rate in the optical tract or to achieve a two-fold increase in the stored information volume at the same clock frequency.

To ensure a reliable access to any segment of circulating information sequence and to improve its storage security, the additional channel can be used as an addressing channel. In this case (Fig.1b), writing of the information sequence and the auxiliary (addressing) sequence into the loop memory is achieved by feeding them synchronously to the information (Inf. input) and auxiliary (aux.input) input terminals of the input/output unit (IOU), respectively, and further to control input terminals of the corresponding electrooptical modulators. As the addressing sequence circulates in closed loop, its segments (address words) are also synchronously written into an *n*-bit address former (AF) thereby forming the address of a coming information sequence. Synchronous circulation of signal flows along identical channels allows simultaneous feeding of the IOU input terminals independent of any external influences.

Access to the required information signal (rewriting, retrieval) isrealized through setting appropriate address on the address inputs of a digital comparator (DC). When the contents of the AF coincides with the set address during the next passage of the address sequence along the circulation loop, the DC generates a control signal used to strobe information signals.

In order to reduce error probability in forming address words,it is advantageous to use a pseudo-random sequence (PRS) as address sequence.The PRS length is $M = 2^{**}N - 1$,where N is the PRS generator bit capacity,and is determined by the memory information capacity; the address former represents an n-bit shift register, where n is not less than N.In case a redundant length of the PRS address words is used,i.e.for $n > N$, addressing is free of errors even if the received address word has several incorrect ("spoilt") bits.

The additional channel can also be used for transmission of a sequence of clock signals whose time parameters are matched with those of the information signal. In this case,the information and synchronizing sequences circulate in synchro-

nism in opposite directions along the same loop,which excludes their mistiming caused by the influence of external temperature and mechanical factors on the fiber-optic tract thus improving reliability of the information signal synchronization. In this case, the ASU represents a modulator that is identical to the main channel modulator controlled by the clock pulse generator (commutator in Fig.1a is switched to position 2). The clock signals at the output end of the additional channel are used for synchronous detection of the output information sequence. The repeated input of the restored information signals into the circulation loop is performed synchronously with pulses of a quartz clock.

The contrary-directed channel in the fiber-optic circulator can be employed for amplification and compensation of attenuated information signals propagating in the light guide, making an electrooptical regenerator superfluous.Moreover, by closing the loop with the couplers, an all-optical FOMD with multiple propagation of signals in the loop is obtained. Even if the electrooptical regenerator is employed, optical amplification allows formation of a considerably longer loop and, hence, obtaining of a higher FOMD capacity.

Optical amplification in a fiber can be accomplished by using various nonlinear optical effects, for example,stimulated Raman scattering (SRS) [5].

3. Conclusions

The bi-directional circulation regime proposed makes it possible to find an effective solution to the problems of increasing the information capacity and attaining reliable storage of information in the FOMD circulation loop.Moreover, it allows optical amplification of information signals to be achieved. Amplification of information signals can be effected by using both continuous wave pumping and synchronizing signals of appropriate power. Thus, the contrary-directed channel can be simultaneously used for both synchronization and amplification of information pulses. The required optical power of the pump wave is of the order of some hundreds of milliwatts [4].

This work has been financially supported from the Fundamental Research Foundation of the Republic of Belarus.

References

[1] Belovolov M I, Golovin N I, Golovina T N, Dianov E M, et al. 1985 *Kvantovaya Elektronika* **12** 214-216
[2] Pilipovich V A,Esman A K,Vizner A A 1990 *Doklady AN BSSR* 34 318-320
[3] Pilipovich V A, Esman A K, Goncharenko I A, Posedko V S and Solonovich I F 1993 *ISFOC'93 Conference Proceedings.-Information Gatekeepers Inc.* 97-100
[4] Pilipovich V A, Esman A K, Goncharenko I A, Posedko V S and Solonovich I F 1993 *Sov.Lightwave Commun.* **3** 169-175
[5] Aoki Y 1988 *J.Lightwave Technology* **6** 1225-1239

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

All-optical arithmetic operations by means of spatial soliton interactions

F.Garzia, C.Sibilia, M.Bertolotti

Dipartimento di Energetica, Universita' di Roma "La Sapienza",
Via A. Scarpa 14, 00161 Roma, Italy, Tel.+39-6-49916541,
Fax +39-6-44240183, E-mail bertol88@itcaspur.caspur.it

Abstract. A scheme of a very fast way to make all-optical arithmetic operations realized in a nonlinear planar waveguide is presented. The scheme is based on interaction properties of solitons.

1. Scheme of the device

Our scheme is based on a couple of two propagating spatial solitons in a nonlinear waveguide, under the hypotesis of transverse confinement realized by a third order nonlinear medium, and where the relative phase is properly varied, so that the interaction properties of solitons can be utilized to perform the required process.
Our analysis starts from an adimensional nonlinear Schroedinger equation[1]:

$$i\frac{\partial u}{\partial Z} = -\frac{1}{2}\frac{\partial^2 u}{\partial X^2} - |u|^2 u. \qquad (1)$$

Under proper conditions the propagation is described from the following system of coupled differential equations[2]:

$$q_{zz} = -4\exp(-2q)\cos(2\Phi)$$

$$\Phi_{zz} = 4\exp(-2q)\sin(2\Phi) \qquad (2)$$

where q=q(z) is the relative distance, along the propagation direction, at the generically z coordinate, and Φ is the relative phase of solitons.
In Ref.[3] the influence is studied of the change of relative phase of two input solitons, of equal input intensity, on the intensity of the soliton whose phase has not been changed. The intensity is found to vary according to the law:

$$I_0 = k_I \Phi \qquad (3)$$

where k_I is a constant depending on initial distance with respect to the other soliton.
This means that the phase information carried by a pulse can be transferred to the intensity of another pulse after propagation (intensity modulation by change in phase).
By the help of this property all arithmetic operations can be realized. For example the adder can be realized when two soliton beams enter in the same waveguide with a relative phase equal to zero and their interaction in the first part of propagation is avoided by means of a proper shield. A beam of intensity I_1 propagated normally to the soliton path, changes via a nonlinear interaction the refractive index of a lenght L of the soliton path (see Fig.1), inducing a phase-change equal to $k_\phi L I_1$. A second beam of intensity I_2 interacts with the same soliton for the same length L inducing a further

phase change equal to $k_\phi L I_2$. The total phase change is $k_\phi L(I_1 + I_2)$. If now we let the two solitons to interact, they interfere according to the intensity modulation effect, and the intensity of one soliton becomes $I_0 = k_i k_\phi L(I_1 + I_2)$, that is proportional to the sum of intensities of input beams, realizing an optical addition.
In a similar way it is possible to realize an optical subtraction. In fact if we consider the same structure of the adder where the zone interested by the second beam I_2 is composed by a negative nonlinear refractive index coefficient, the phase change induced over the soliton beam is equal to $-k_\phi L I_2$. The intensity of soliton after interaction becomes: $I_0 = k_i k_\phi L(I_1 - I_2)$.
The division is made in a quite similar way by using only a beam of intensity I_1. If we reduce the length L by a factor N, the phase change induced over the soliton is equal to $k_\phi I_1 L / N$, and the output intensity becomes $I_0 = k_i k_\phi I_1 L / N$. This process is quite critical since it depends of our capability of reducing the length L properly. See figs. 2.
The multiplication is made by increasing N times the length L, so that the output intensity becomes $I_0 = k_i k_\phi L N I_1$.

2. Dimensioning of the device

To dimension these devices it is necessary to match different parameters. The medium ought to have a large nonlinear refractive index n_2 to use low power level, with a fast response time to make a quick device.
Once chosen the Kerr medium the linear and nonlinear refractive indices n_0, n_2 are determinated and the wavelength λ_0 necessary to excite the nonlinearity.
The structure of the waveguide determines the spot size a_0 necessary to excite the minimum number of modes. The ideal situation is the excitation of only the fundamental mode.
The power necessary to excite a soliton beam is thus:

$$P_s = \frac{\lambda_0^2}{4 a_0^2 n_0 n_2}. \tag{4}$$

Since the material is characterized by a well-defined response time of the nonlinearity, the duration of the pulses emitted by the light source must be greater or almost equal to this response time.
The intensity modulation effect strictly depends of the initial soliton beam spot size[3]: to let the effect take place it is necessary to make the two soliton propagate for a distance almost equal to ten times the beam width: we use a security coefficient equal to ten.
The time of the intensity modulation effect is immediately given as:

$$T_m = 100 \frac{a_0 n_0}{c}. \tag{5}$$

The phase varying soliton that crosses through the zone enlighten by the input beams experiences a phase variation with respect to the other soliton equal to:

$$\Delta\varphi = \Delta n \cdot k \cdot L, \tag{6}$$

where Δn is the induced refractive index variation, k is is the wavevector in the medium and L the computation length.
Since the intensity modulation by phase change effect is linear only for relative phase values included in the range $\pi/2 \div \pi$, that is for a maximum phase variation of $\pi/2$ (provided one desires to use all the phase dynamic available), we can calculate the computation length L from Eq.[6] imposing $\Delta\varphi = \pi/2$, that gives:

$$L = \frac{\pi/2}{k \cdot \Delta n} = \frac{a_0^2}{\lambda_0}. \tag{7}$$

The time necessary to pass through L, said computation time, is:

$$T_{P_s} = \frac{n_0}{c}\frac{a_0^2}{\lambda_0}, \tag{8}$$

The time expressed by Eq.[8] is the minimum time reachable for a phase change of $\pi/2$, since it is calculated for a controller beam power equal to P_s. A higher power of the controller beam would generate a transversal soliton that collides with other solitons deflecting them from projected trajectories generating malfunctions of the device. For this reason it is necessary to decrease the power level. Reducing it N times it is immediate to see that the computation time becomes N times greater.
Generally it is sufficient to reduce the power by a factor of two without appreciably slow down the device.

3. A practical example.

Let's see a practical example. We consider a waveguide composed by a thin film of CS_2, that is an organic compound whose nonlinear mechanism is represented by the molecular-orientation Kerr effect whose response time is of about 1 ps. The optical parameters are:

$$n_0 = 1.6, \quad n_2 = 4 \cdot 10^{-18} m^2 / W. \tag{9}$$

Since the material is liquid, it is confined by means of two parallel quartz plates. The thickness of the waveguide is of about 10 μm and we have 10 propagating modes. It is possible to excite the fundamental mode if a narrow converging beam is used. If we use a spot size of 100 μm with a λ=532 μm, that is the wavelength at which the material has a low absorption coefficient, we can calculate from Eq.[4] the critical power necessary to obtain a soliton beam:

$$P_s \cong 1 \cdot 10^{11} W / m^2. \tag{10}$$

To have this power level it is necessary to use a pulsed source. The ideal candidate at λ=532 μm is a Nd:YAG laser. Since the typical pulse duration's of this laser are in the range 10÷100 ps, we are sure that the medium is fast enough to support and sustain a soliton beam.
The computation time is given by Eq.[8]:

$$T_{P_s} \cong 10 ps. \tag{11}$$

The time necessary to perform a single addition is:

$$T_O = T_m + 2T_{P_s} = 80 ps \tag{12}$$

The repetition rate of the operations depends obviously on the source and on the thermal transmission features of the heat generated by the absorption of medium.

4. Conclusions

We presented a flexible scheme to realize all-optical arithmetic operations that, given two input beams, is able to generate an output beam whose intensity is proportional to the sum (or to the difference if properly designed) of the input intensities. The scheme, properly modified, is able to realize the division or the multiplication of the intensity of an input beam.

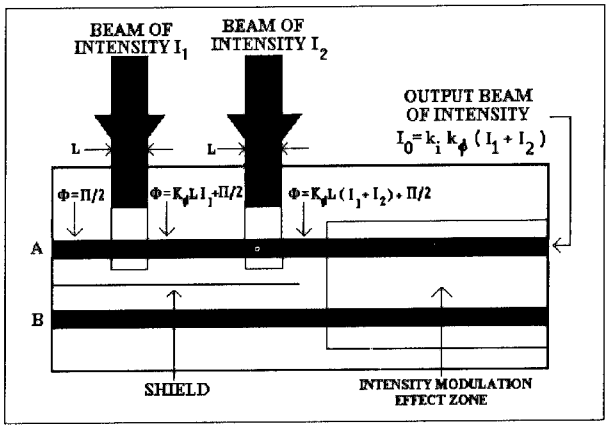


Fig.1 Scheme of the optical adder.

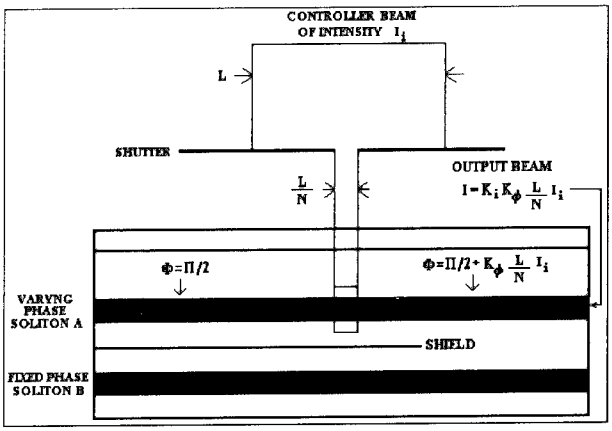


Fig.2 Scheme of the optical divider.

References

[1] Zakharov V E and Shabat A B 1972 *Sov.Phys. JETP* **34** 62-69
[2] Gordon J P 1983 *Opt. Lett.* **8** 596-600
[3] Garzia F, Sibilia C, Bertolotti M, Horak R and Bajer J 1994 *Opt.Comm.* **108** 47-54
[4] Maneuf S, Desailly R and Froehly C 1988 *Opt.Comm.* **65** 193-198
[5] Reynaud F and Barthelemy A 1990 *Europh. Lett.* **12** 401-409
[6] Shalaby M, Reynaud R and Barthelemy A 1992 *Opt. Lett.* **17** 778-785

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

NEW TYPES OF OPTICAL SOLITON-LIKE STRUCTURES AND THEIR POSSIBLE APPLICATIONS TO OPTICAL COMPUTING

A.V.Fedorov, S.V.Fedorov, G.V.Khodova, N.N.Rosanov, V.A.Smirnov and N.V.Vyssotina

Institute of Laser Physics, Research Center "S.I.Vavilov State Optical Institute", St.Petersburg, 199034, Russia

Abstract. Basic properties and possible applications to optical computing of new types of spatial and spatio-temporal solitons and their bound structures are considered in the following nonlinear optical schemes: (a) medium with nonlinearity of refractive index: (b) wide-aperture laser with saturable absorption; (c) optical fibre with sections of gain and absorbing media; (d) wide-aperture nonlinear interferometer.

1. Introduction

Optical solitons in nonlinear optical fibres have long been attracting the attention of investigators in connection with their applications to optical communications, superfast switchings and optical computing [1]. At present, the number of optical soliton-like structures has considerably increased. Thus, possibilities of applications of spatial solitons generated in integrated optical (planar) circuits are discussed [2]. Attractive are the opportunities to use three-dimensional spatial solitons - "light bullets" - for information processing and computing [3]. A number of schemes for all-optical information processing has been proposed, switching waves and diffractive autosolitons (spatio-temporal solitons) in wide-aperture nonlinear interferometer being used there [4-6].

In the present work, we compare general properties of bright soliton-like light structures in a number of nonlinear optical systems, mainly considering the stability of their characteristics. We come to the conclusion that these structures are most stable in bistable interferometers. In this connection, we analyze more thoroughly the all-optical scheme of full adder proposed in [5] and essentially using the properties of diffractive autosolitons. A number of stages of a multidigit summing cycle has been modelled in [7].

2. Solitons in nonlinear transparent media

Here we discuss 2-D spatial solitons in a medium with saturable nonlinearity of refractive index. Computer simulations for interaction of 2-D solitons were done in [8]. Colliding nonfundamental solitons decay into fundamental ones. In collisions, fundamental solitons pass through each other without noticeable changes, if the interaction region is sufficiently small. In the mode of strong interaction, repulsion of solitons with partial power transfer or their attraction and convergence are observed. Due to continuous spectrum of propagation constant or the maximal radiation intensity for fundamental solitons, noise causes the drift of soliton parameters.

3. Solitons in wide-aperture lasers and active fibers

For lasers with saturable absorption, hard excitation of lasing is typical. In these conditions, generation of switching waves and laser solitons - islands of lasing over the nonlasing background - is possible for wide-aperture lasers [9]. The maximal radiation intensity for a laser soliton has a certain value, for which radiation losses are compensated by saturable gain.

In transversely 2-D geometry, there are stable laser solitons with different topological charge: $m = 0$ (regular wave front), and $m = \pm1, \pm2, \ldots$ (vortices). In a laser with infinite aperture, there are laser solitons propagating across the aperture with arbitrary velocity. This corresponds to arbitrary angle between the axis of radiation propagation in the soliton and the cavity axis. In an actual laser with finite aperture a soliton is reflected by the mirror edge, if the angle between the axes is smaller than some critical value. The result of collision of transversely 1-D solitons depends on their relative velocity and the phase difference. With the relative velocity decrease, the following regimes take place consecutively: passing of the solitons through each other; generation of new solitons; convergence of the solitons into one soliton; their repulsion.

The equation for propagation of pulses in one-mode nonlinear fibre with intervals with saturable gain and absorption coincides with equations for transversely 1-D laser with saturable absorption. Therefore, the results presented above are valid for this case also. The hard type of lasing excitation suppresses radiation noises, and the fixed value of the maximal intensity excludes the drift of the soliton energy. This makes laser solitons promising for optical communications.

4. Diffractive autosolitons in nonlinear interferometers

In a wide-aperture interferometer excited by a wide beam of cw- external radiation, there is a discrete spectrum of spatio-temporal solitons - "diffractive autosolitons" (see review [6]). The maximal intensity and width of an autosoliton have definite values. Bound states of diffractive autosolitons have discrete spectrum of distances between them. Asymmetric structures move in the transverse direction with constant velocity determined by the intensity of the holding radiation. The following section presents consideration of possible use of properties of diffractive autosolitons.

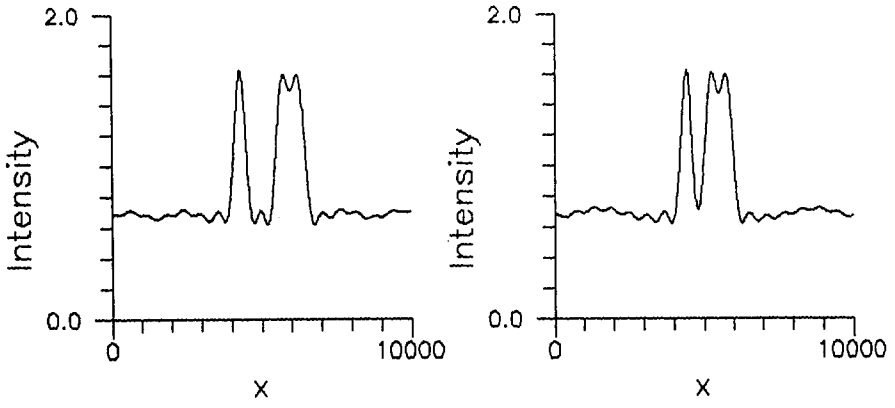


Fig. 1. Generation of asymmetric structure of two coupled autosolitons. (a) $t = 0$; (b) $t = 100$.

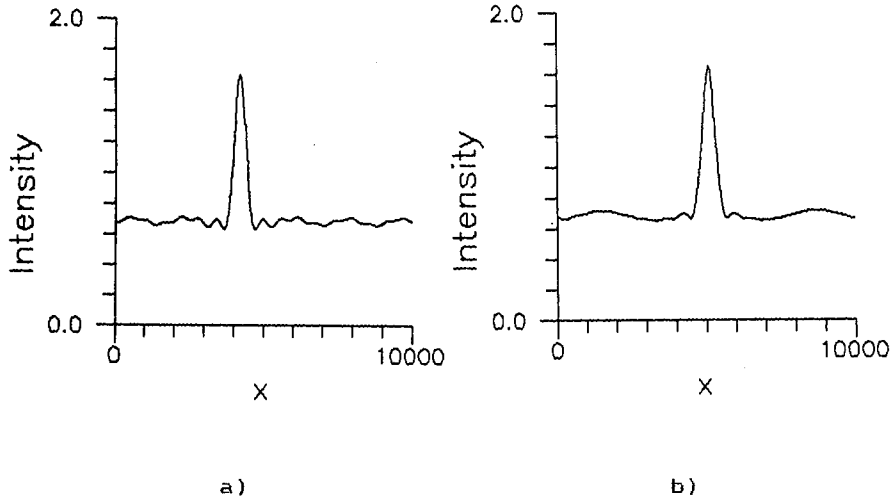


Fig. 2. Shift of autosoliton. (a) $t = 0$; (b) $t = 400$.

5. Computer simulations of optical full adder

One of the stages of the all-optical full-adder operation proposed in [5,6] is doubling of spatial modulation period of holding radiation. In the digits with unities in both summands, it provides generation of an asymmetric coupled structure of two autosolitons with different widths from the autosolitons stored in independent cells. If only one of the summands has unity in the considered digit, the change of holding radiation modulation produces necessary shift of the autosoliton representing the unity to the center of the digit for its future processing.

Fig. 1 presents the results of computer simulations for the case when both summands have unities in the corresponding digit. The unities are stored in the form of a narrow and a wide autosolitons located in the maximums of holding radiation intensity (Fig. 1a). The holding intensity is described by

$$I_i = I_{io} + I_{mod}\cos\{k(x - x_o)\}. \qquad (1)$$

where (λ is the radiation wavelength and I_{sat} is the saturation intensity)

$$I_{io} = 0.95I_{sat},\ I_{mod} = 0.02I_{sat},\ k = 3.35\cdot 10^{-9}\lambda^{-1},\ x_o = 4200\lambda. \qquad (2)$$

In the considered system, autosolitons move in the gradient of intensity in the direction of higher intensities. The change of the parameters of holding radiation modulation to

$$k = 1.675\cdot 10^{-9}\ \lambda^{-1},\quad x_o = 5100\ \lambda, \qquad (3)$$

places the autosolitons on the slopes of the intensity maximum and makes them move towards each other. Fig. 1b shows the settled intensity distribution in 100 round-trip times after modulation switching. This spatial structure of light becomes almost stationary in about 40 round-trip times.

Fig. 2 illustrates the shift of an autosoliton after the change of holding radiation modulation. Initially, the autosoliton is stored in the intensity maximum (Fig. 2a), the modulation being given by (1) with parameters (2). After changing the parameters of the modulation to (3), the autosoliton moves to the intensity maximum, reaching it in about 150 round-trip times. Fig. 2b shows the intensity distribution in 400 round-trip times after the switching.

Thus, computer simulation confirmed the possibility of realization of all stages necessary for optical adding of multidigit numbers within the method proposed earlier.

References

[1] Hasegawa A 1989 *Optical Solitons in Fibers* (Berlin:Springer)
[2] Wright E M 1992 *Optical and Quantum Electron.* **24** N 11
[3] Silberberg Y 1990 *Opt. Lett.* **15** 1282-4
[4] Rosanov N N and Fedorov A V 1990 *Opt. and Spectr.* **68** 565-6
[5] Rosanov N N 1992 *Optics and Spectr.* **72** 447-53
[6] Rosanov N N 1992 *Proc. SPIE* 1840 130-43
[7] Rosanov N N, Fedorov A V and Khodova G V 1993 *Proc. SPIE* **2051** 349-56
[8] Vyssotina N V, Rosanov N N and Smirnov V A 1994 *Optics and Spectr.* **76** 774-84
[9] Rosanov N N and Fedorov S V 1992 *Opt. and Spectr.* **72** 1394-9

Inst. Phys. Conf. Ser. No 139: Part VI
Paper presented at Opt. Comput. Int. Conf., Edinburgh, 22–25 August 1994
© 1995 IOP Publishing Ltd

Asymmetric Light Bullet Dragging Logic

Robert Mcleod, Steve Blair, and Kelvin Wagner
Optoelectronic Computing Systems Center
University of Colorado, Boulder CO, 80309-0425
1-303-492-4661 kelvin@boulder.colorado.edu

Abstract

The asymmetric dragging interaction between three-dimensional optical solitons may allow cascadable, phase-insensitive, NOR gates with gain to be implemented at ultrahigh speeds in massively parallel three-dimensional bit-level systolic-array architectures.

In media that have both a self-focusing non-linearity ($n_2 > 0$) and negative group velocity dispersion, a pulse can collapse in both space and time forming a stable "light bullet" in 3+1 dimensions.[1–2] These collapsed optical pulses have high peak powers but small total energies, making them attractive for nonlinear optics applications. These three dimensional solitons can be made to interact to produce a logic gate in the same way that one-dimensional spatial (or temporal) solitons do. We are investigating the asymmetric interaction between two orthogonally polarized solitons brought into coincidence both spatially and temporally at the boundary of a nonlinear medium and propagating at slightly different angles. This geometry permits a weak signal to spatially drag a strong pump by well over a beam width, allowing them both to be blocked by a spatial aperture.[3] This results in a phase insensitive inverter with gain that can be cascaded to implement a high contrast NOR gate that transmits an uncorrupted pump through the aperture if the signals are not present, and drags the pump out of the spatial aperture so that it is blocked if they are present. A beam propagation simulation of an asymmetric light-bullet dragging interaction is shown in Figure 1. This shows a dragging of a $5\mu m \times 5\mu m \times 4\mu m$ 1 pJ pump (I_p=60MW/cm^2) by a .25pJ signal (I_s=15MW/cm^2) in about 0.7mm of propagation distance using a rather large saturating nonlinearity of $n_2^E = 10^{-16}\mathrm{m}^2/\mathrm{V}^2$. Simultaneous multidimensional dragging by two signals one tilted in $\hat{x}$ and the other tilted in $\hat{y}$, is shown in Figure 2, demonstrating the single stage implementation of a NOR gate. The signals are in the same polarization and interfere, but numerical studies show that the assymetrical interaction can always function as a NOR gate for any relative phase of the two signal solitons.

In order to optimize these light bullet dragging gates without the huge computational overhead of the 3+1 dimensional beam propagation simulations, we are examining the properties of 1D asymmetric spatial soliton dragging interactions [4], as illustrated in Figure 3. The figure illustrates asymmetric soliton dragging showing propagation of a pump alone on the left passing through a spatial aperture, the dragging of a pump by a signal in the middle, and a signal alone propagating at an angle on the right. The top plane represents the spatial evolution of the intensity profile for the pump polarization and the bottom plane represents intensity for the orthogonal signal polarization. The inset schematically shows the waveguide device geometry of an inverter in both states.

As an example we illustrate the output contrast ratio (wrt the fundamental soliton power) in a two-level system saturating nonlinear medium (in this case I_{sat} is 4 times the nonsaturated fundamental soliton peak intensity), as a function of interaction angle, propagation distance and pump-to-signal beam ratio in Figure 4. It is clear that high gain and good contrast are achievable when linear resolvability ≈ 1 for propagation distances ≥ 5.

Layered sandwiches of nonlinear media, aperture arrays and linear media suggest the possibilty of computing in 3-D with asymmetric light bullet dragging gates. But getting the signals and pumps to the desired interaction sites without disruption by unwanted signals may be difficult. It may be necessary to systolisize at the bit level in 3-dimensions by pulsing the clock pumps and signals so that they pass through each other in intervening layers of linear media until arriving at the desired dragging logic site, and to program the functionality of the array

of logic gates by the presence and absence of clock pumps in the space-time lattice of possible light bullet locations. An example of a 1-D spatial soliton dragging logic cascaded majority logic circuit is illustrated in Figure 5, demonstrating that more complex functions than simple NOR gates can be implemented without intervening interconnections.

Ultrafast, massively parallel, low latency, all optical NOR gates with gain, cascadability, input-output isolation, and phase insensitivity have been proposed, numerically demonstrated and parametrically optimized. This optical switching interaction opens up new architectural possibilities for computing in 3+1 dimensions that may allow the realization of volume parallel digital optical computers.

The authors acknowledge support of the NSF young investigator program ECS 9258088 and a DoD fellowship DAAL03-92-G-0351.

References

1. Y. Silberberg, Collapse of optical pulses, *Opt. Lett.*, vol. 15(22), p 1282, 1990.
2. J. E. Rothenberg, Space-time focusing, *Opt. Lett.*, vol. 17(19), p 1340, 1992.
3. K. Wagner and B. Mcleod, Spatial Soliton Dragging Gates and Light Bullets, OSA Topical Mtg. on Optical Computing, Palm Springs CA, March 1993.
4. S. Blair, K. Wagner, and R. McLeod, Asymetric Spatial Soliton Dragging, Accepted for publication Optics Letters, 1994.

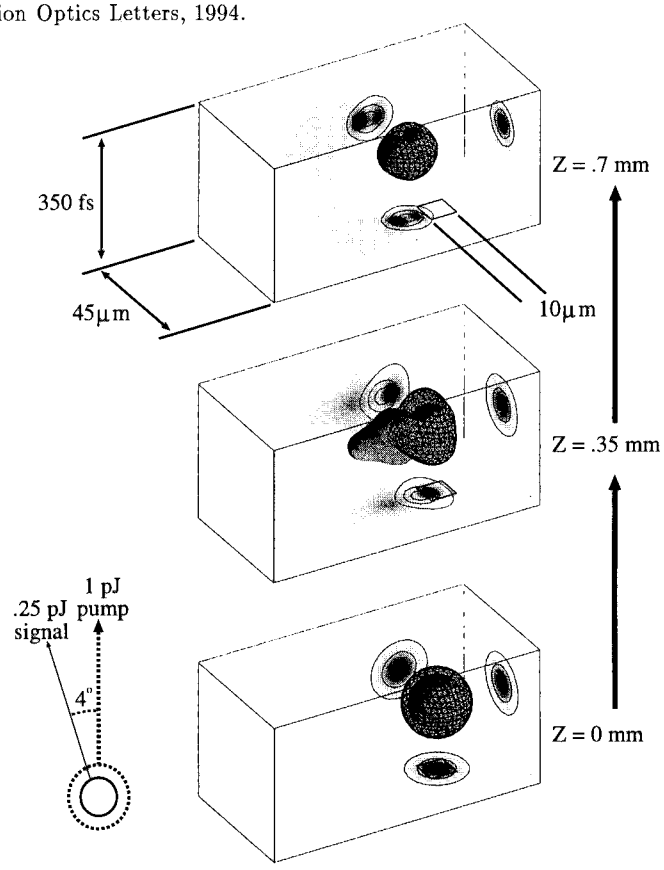


Figure 1: Light bullet dragging resulting in a contrast of 32. In the bottom frame, the 1/4 pJ signal soliton, rendered as a solid isosurface at 1/10 of its peak intensity, is initially overlapping the 1 pJ pump light bullet, represented by a grid isosurface at the same intensity. As the two solitons propagate upward, with the signal tilted at an initial four degree angle, the pump is pulled out of the 10 micron aperture to implement the inversion or switching operation.

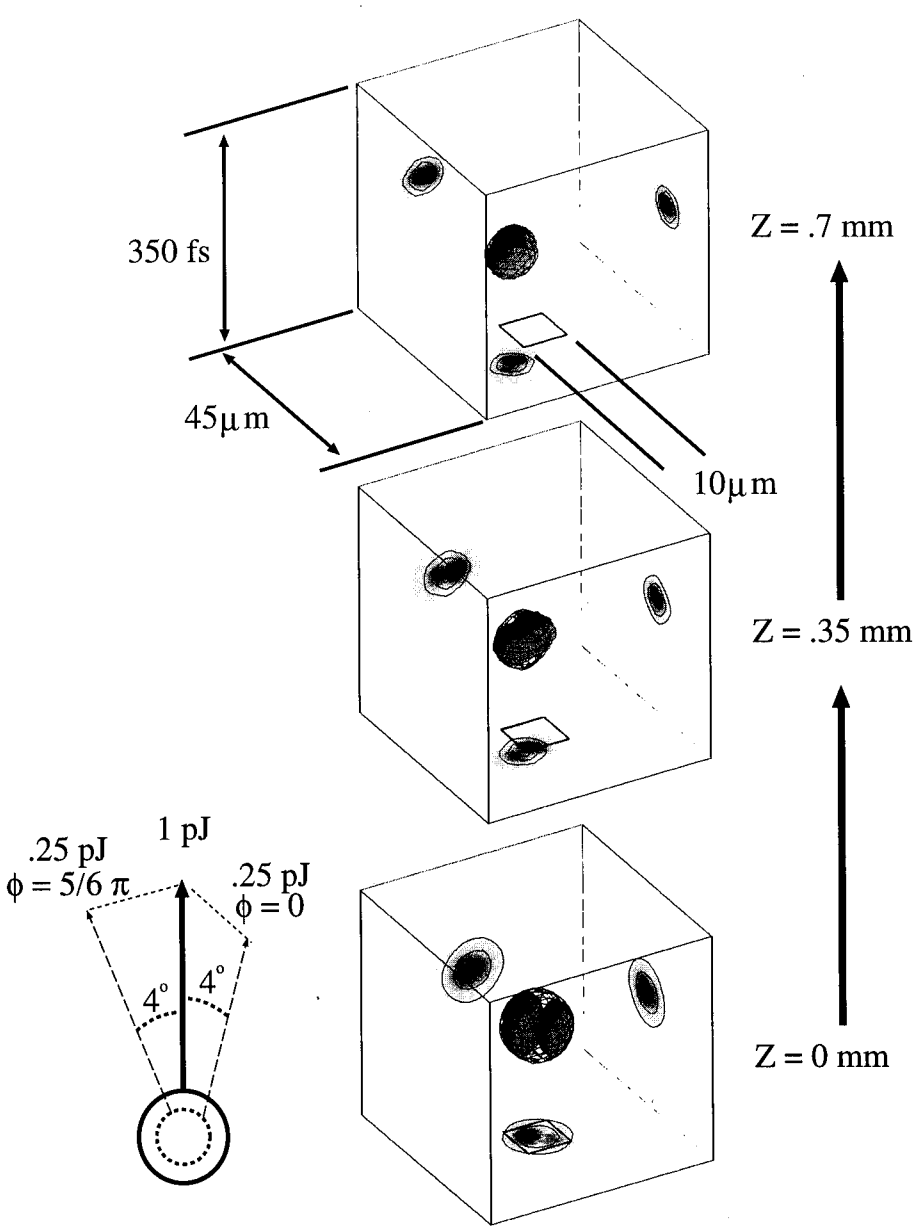


Figure 2: Single-stage two-input NOR light bullet dragging logic gate. All parameters are the same as the previous figure, but now two signals, each tilted at an initial four degree angle in two orthogonal directions, interact with the pump and drag it on the diagonal between them. The two signals are at nearly the worst case of $5/6\ \pi$ out of phase.

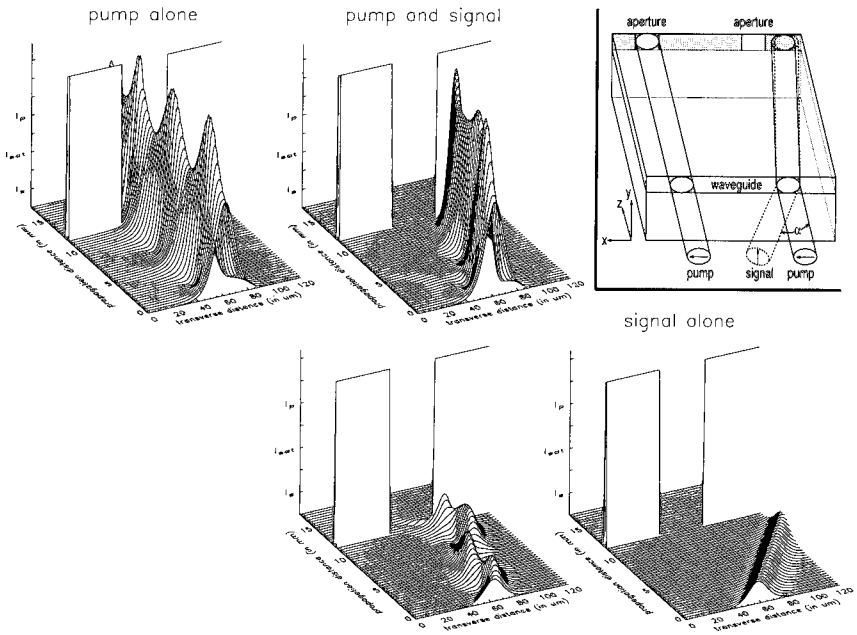


Figure 3: Illustration of asymmetric soliton dragging

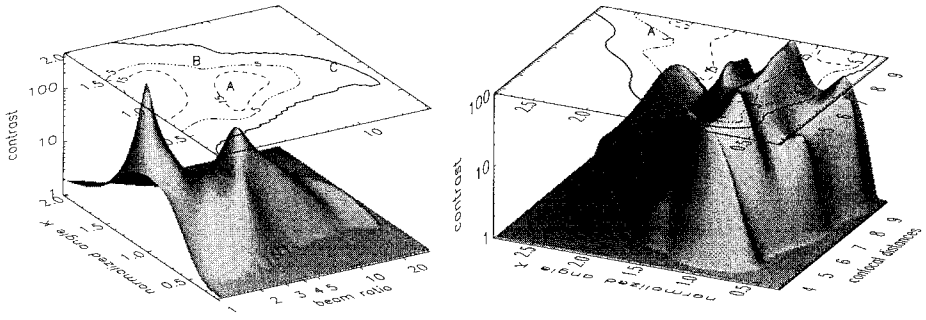


Figure 4: Output contrast versus interaction angle and pump-to-signal ratio for a propagation distance of 5 confocal distances, and contrast versus interaction angle and propagation distance for a beam ratio of 5.

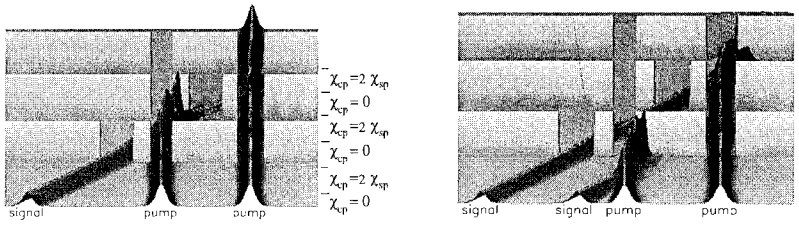


Figure 5: Cascaded 3-level simulation of majority logic circuit showing that the first pump implements a NOR that is blocked by either signal, and the second pump is only blocked when 2 or more off-axis signals are present.