DTIC FILE COPY

AD-A220 879

OPTICAL COMPUTING

DTIC
ELECTE
APR 25 1990
S D
Co D

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

1989 TECHNICAL DIGEST SERIES VOLUME 9

FEBRUARY 27–MARCH 1, 1989
SALT LAKE CITY, UT

90 04 23 084

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS n/a
2a. SECURITY CLASSIFICATION AUTHORITY n/a	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release- distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE n/a	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) N00014-89-J-1986	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Optical Society of America 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research Code 1513:RAR
6c. ADDRESS (City, State, and ZIP Code) 1816 Jefferson Place, N.W. Washington, D.C. 20036	7b. ADDRESS (City, State, and ZIP Code) Department of the Navy Office of Naval Research, Code 1513:RAR 800 North Quincy Street, Arlington, VA22217-5
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Organization of the 1989 Optical Computing Topical Meeting

12. PERSONAL AUTHOR(S)

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 02/15/89 TO 12/31/89	89/12/31	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The main objective of the meeting was to report on the current state of the art in optical computing and to look into possible future directions that are likely to emerge from an interation between new devices, novel architectures, and nontraditional applications for optical computing.

The meeting consisted of invited papers and both oral and poster contributed papers in all areas of research in materials, devices, architectures and algorithms relevant to optical computing. Topics of interest included: optical interconnections for computing, optical computing systems, architectures and algorithms, digital or analog, hybrid optical/electronic processor, optical "neural" processors, including optical associative memories, and optical artifical intelligence.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Jarus W. Quinn, Executive Director	22b TELEPHONE (Include Area Code) 202/223-8130 22c. OFFICE SYMBOL

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

Unclassified

OPTICAL COMPUTING
1989 TECHNICAL DIGEST SERIES, VOLUME 9

CONFERENCE EDITION

Summaries of papers presented at the
Optical Computing
Topical Meeting

February 27–March 1, 1989

Salt Lake City, Utah

PRICE-$70.00 per Optical Society of
America, 1816 Jefferson Place, NW
Wasington, DC 20036
TELECON 4/23/90 VG

Cosponsored by

Optical Society of America

Air Force Office of Scientific Research

Defense Advanced Research Projects Agency

National Science Foundation

Office of Naval Research

SDIO/Innovative Science and Technology Office

Optical Society of America
1816 Jefferson Place, N.W.
Washington, D.C. 20036
(202) 223-8130

Accesion For
NTIS CRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By 70.00
Distribution /
Availability Codes
Dist | Avail and/or Special
A-1 | 24

Articles in this publication may be cited in other publications. In order to facilitate access to the original publication source, the following form for the citation is suggested:

Name of Author(s), Title of Paper, Optical Computing, 1989 Technical Digest Series, Vol. 9, (Optical Society of America, Washington, D.C. 1989) pp. xx-xx.

ISBN Number
Conference Edition 1-55752-080-1 (softcover)
Postconference Edition 1-55752-081-X (hardcover)
1989 Technical Digest Series 1-55752-105-0 (hardcover)

Library of Congress Catalog Card Number
Conference Edition 88-62729
Postconference Edition 88-62730

Copyright © 1989, Optical Society of America

Individual readers of this digest and libraries acting for them are permitted to make fair use of the material in it, such as to copy an article for use in teaching or research, without payment of fee, provided that such copies are not sold. Copying for sale is subject to payment of copying fees. The code 1-55752-105-0/89/$2.00 gives the per-article copying fee for each copy of the article made beyond the free copying permitted under Sections 107 and 108 of the U.S. Copyright Law. The fee should be paid through the Copyright Clearance Center, Inc., 21 Congress Street, Salem, Mass. 01970.

Permission is granted to quote excerpts from articles in this digest in scientific works with the customary acknowledgment of the source, including the author's name and the name of the digest, page, year, and name of the Society. Reproduction of figures and tables is likewise permitted in other articles and books provided that the same information is printed with them, permission of one of the original authors is obtained, and notification is given to the Optical Society of America. Republication or systematic or multiple reproduction of any material in this digest is permitted only under license from the Optical Society of America; in addition, the Optical Society may require that permission also be obtained from one of the authors. Address inquiries and notices to Director of Publications, Optical Society of America, 1816 Jefferson Place, N.W., Washington, DC 20036. In the case of articles whose authors are employees of the United States Government or its contractors or grantees, the Optical Society of America recognizes the right of the United States Government to retain a nonexclusive, royalty-free license to use the author's copyrighted article for United States Government purposes.

This work relates to the Department of Navy Grant issued by SDIO/Innovative Science and Technology Office and the Office of Naval Research. The U.S. Government has a royalty license through the world in all copyrightable material contained herein.

The views and conclusions contained in this document are those of the author(s) and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

The material is based upon work supported by the National Science Foundation under Grant No. EET-8901310. The Government has certain rights in this material. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

SUNDAY, FEBRUARY 26, 1989

SALON D

6:00 PM-9:00 PM REGISTRATION/RECEPTION

MONDAY, FEBRUARY 27, 1989

SALON D

6:30 AM-8:15 AM BUFFET BREAKFAST

SALON FOYER

7:00 AM-6:00 PM REGISTRATION/SPEAKER CHECKIN

SALON F

8:15 AM-8:30 AM
OPENING REMARKS
Alexander A. Sawchuk, *University of Southern California*

8:30 AM-9:30 AM
MA NEURAL SYSTEMS: 1
C. Lee Giles, *Air Force Office of Scientific Research, Presider*

8:30 AM (Invited Paper)
MA1 Optical Implementations of Neural Computing, Ravindra A. Athale, *BDM Corp.* Different approaches to optical implementation of neural models for computations are reviewed. (p. 2)

9:00 AM (Invited Paper)
MA2 Electronic vs Optical Implementations of Neural Networks, Jay P. Sage, *MIT Lincoln Laboratory.* We address, for the optical community, the relative advantages and limitations of electronic neural network implementations in contrast to optical implementations. (p. 5)

SALON F

9:30 AM-10:00 AM
MB NEURAL SYSTEMS: 2
Henri H. Arsenault, *Laval University, Presider*

9:30 AM
MB1 Implementation of Dynamic Hopfield-Like Networks Using Photorefractive Crystals, Jeff Wilde, Lambertus Hesselink, *Stanford U.* We present an architecture for optically implementing a digital associative memory using a coherent optical system. Bipolar information is holographically phase encoded in a photorefractive crystal. (p. 10)

9:45 AM
MB2 Classification of Normal and Aberrant Chromosomes by an Optical Neural Network in Flow Cytometry, S. Noehte, R. Manner, M. Hausmann, H. Horner, C. Cremer, *U. Heidelberg, F. R. Germany.* An optical neural network capable of classifying normal and aberrant chromosomes in flow cytometry at a rate of 10 kHz is described. (p. 14)

SALON D

10:00 AM-10:30 AM COFFEE BREAK

SALON F

10:30 AM-11:30 AM
MC OPTICAL ARTIFICIAL INTELLIGENCE AND ADAPTIVE SYSTEMS
W. Thomas Cathey, *University of Colorado, Presider*

10:30 AM (Invited Paper)
MC1 Optical Artificial Intelligence Based on Semantic Network Architecture, Toyohiko Yatagai, *U. Tsukuba, Japan.* We propose an optical architecture for context-sensitive association by a modifying learning matrix. An MSLM is used for recording the matrix and matrix-vector multiplication. (p. 20)

11:00 AM
MC2 Optical Matrix Encoding for Constraint Satisfaction, Gary C. Marsden, Fouad Kiamilev, Sadik Esener, Sing H. Lee, *UC-San Diego.* Constraint satisfaction problems are representable in matrix form. Consistent labelling algorithms requiring a limited dynamic range have been developed to take advantage of the parallelism of optics. (p. 24)

11:15 AM
MC3 Adaptive Optical Filtering Architecture, Michael G. Price, Joseph C. Harsanyi, *Systeka, Inc.;* Alan E. Craig, John N. Lee, *U.S. Naval Research Laboratory.* Adaptive filtering is applied to narrowband interference rejection for wideband receiver systems. A time/space integrating optical architecture using a spatial light modulator is described. (p. 28)

SALON F

11:30 AM-12:30 PM
MD NEURAL SYSTEMS: 3
Demetri Psaltis, *California Institute of Technology, Presider*

11:30 AM
MD1 Optical Associative Memory Utilizing Electrically and Optically Addressed Liquid Crystal Spatial Light Modulators, Kristina M. Johnson, M. Kranzdorf, B. J. Bigner, L. Zhang, *U. Colorado.* We present new results on performing optical associative memory with the polarization-based optical connectionist machine. Both electrically and optically addressable spatial light modulators have been incorporated into the design of this system. (p. 32)

11:45 AM
MD2 Competitively Inhibited Optical Neural Networks Using Two-Step Holographic Materials, Michael Lemmon, B. V. K. Vijaya Kumar, *Carnegie Mellon U.* Competitively inhibited networks can be used as MAP predictors on a variety of problems. An optical implementation of these is proposed based on two-step holographic materials. (p. 36)

12:00 M
MD3 Adaptive 2-D Quadratic Associative Memory Using Holographic Lenslet Arrays, Ju-Seog Jang, Sang-Yung Shin, Soo-Young Lee, *Advanced Institute of Science & Technology, Korea.* Optical implementation of adaptive 2-D quadratic associative memory that requires parallel N^2 weighted interconnections is described by using holographic lenslet arrays and spatial light modulators. (p. 40)

12:15 PM
MD4 Self-Pumped Optical Neural Networks, Yuri Owechko, *Hughes Research Laboratories.* Optical neural network architectures are described which store each connection weight in a continuum of spatially distributed photorefractive gratings. This approach reduces crosstalk and fully utilizes the spatial light modulator. (p. 44)

12:30 PM-2:00 PM LUNCH BREAK

SALON F

2:00 PM-2:45 PM
ME SLMs AND OPTICAL DEVICES: 1
Arthur Fisher, *U.S. Naval Research Laboratory, Presider*

2:00 PM
ME1 Sixty-Four-Element Hybrid PLZT/Silicon Spatial Light Modulator Array, I. Bennion, M. J. Goodwin, C. J. Groves-Kirkby, A. D. Parsons, *Plessey Research Caswell, U. K.* A 64-element (8 × 8) hybrid PLZT/Si electrooptic modulator array is described, with uses in optical computing, signal processing, and interconnection. (p. 50)

2:15 PM
ME2 Dual Beam Recrystallization of Si on PLZT, Ali Ersen, Samhita Dasgupta, T. H. Lin, Sadik Esener, Sing H. Lee, *UC-San Diego.* We report the results of two beam (Ar⁺ and CO_2) recrystallization of silicon on PLZT toward the fabrication of spatial light modulator arrays. (p. 54)

2:30 PM
ME3 Parallel Readout of Optical Disks, Demetri Psaltis, Alan A. Yamamura, Mark A. Neifeld, *California Institute of Technology;* Seiji Kobayashi, *Sony Corp., Japan.* In the context of parallel access of optical disks, we examine available systems including a Sony magnetooptic system. consider optical limitations, and describe suitable uses. (p. 58)

SALON F

2:45 PM-3:30 PM
MF SLMs AND OPTICAL DEVICES: 2
Sing H. Lee, *University of California-San Diego, Presider*

2:45 PM
MF1 High-Speed Optically Addressed Spatial Light Modulator for Optical Computing, R. A. Rice, W. Li, G. Moddel, *U. Colorado.* We present the resolution and response time characteristics for an optically addressed spatial light modulator using a hydrogenated amorphous silicon photosensor and ferroelectric liquid crystal modulator. (p. 64)

3:00 PM
MF2 Optical Nonlinear Neurons and Dynamic Interconnections Using the Field Shielding Nonlinearity in CdTe, William H. Steier, Mehrdad Ziari, *U. Southern California.* An optical neuron with a soft threshold response and an optical dynamic interconnection with microsecond speed and modest switching energy has been demonstrated in the infrared using CdTe. (p. 67)

3:15 PM
MF3 Photorefractive Neuron by Two-Wave Mixing, V. Hornung-Lequeux, P. Lalanne, J. Taboury, G. Roosen, *Institute of Theoretical and Applied Optics, France.* Theoretical analysis and experimental studies show that photorefractive two-wave mixing in barium titanate is well suited for the implementation of an all-optical input-output neural response. (p. 71)

SALON D

3:30 PM-4:00 PM COFFEE BREAK

SALON F

4:00 PM-5:00 PM
MG SLMs AND OPTICAL DEVICES: 3
Bernard Soffer, *Hughes Research Laboratory, Presider*

4:00 PM
MG1 Photorefractive Spatial Light Modulation by Electrocontrolled Beam Coupling in SBN:Ce Crystals, Jian Ma, Liren Liu, Shudong Wu, Zhijiang Wang, *Shanghai Institute of Optics and Fine Mechanics, China.* Dynamic incoherent-to-coherent image conversion is proposed, which is based on the effects of electrocontrolled two-beam coupling in SBN:Ce and image spatial modulation of coupling gain. Either a negative or positive coherent replica is obtained by altering the electric field. (p. 76)

4:15 PM
MG2 InP/InGaAs-Based Charge-Coupled Devices for MQW Spatial Light Modulator Applications, K. Y. Han, R. Chang, C. W. Chen, J. H. Quigley, M. Hafich, G. Y. Robinson, D. L. Lile, *Colorado State U.* We present the results of electrical and optical performance characterization of InP and InGaAs charge-coupled device-based MQW spatial light modulators. (p. 80)

4:30 PM
MG3 Optical Space-Variant Logic Gate Using a New Hybrid BSO Spatial Light Modulator. Ji Zhang, Weiwei Liu, Licheng Zhong, Yili Gou, *Tsinghua U., China.* We propose a new hybrid BSO spatial light modulator for encoding input patterns. (p. 84)

4:45 PM
MG4 High-Speed Parallel Optical Processors of Photorefractive GaAs, Li-Jen Cheng, Duncan T. H. Liu, *California Institute of Technology.* We report the first, we believe, demonstration of several basic computing processes using an interferometric technique with a GaAs phase conjugate mirror. (p. 87)

SALON F

5:00 PM-6:00 PM
MH SYMBOLIC SUBSTITUTION
Karlheinz Brenner, *University of Erlangen-Nuremberg, F. R. Germany, Presider*

5:00 PM
MH1 Design of a Symbolic Substitution-Based Optical Random Access Memory, Miles J. Murdocca, Binay Sugla, *AT&T Bell Laboratories.* Symbolic substitution is used in the design of an optical random access memory. The design is near optimal in gate count and circuit depth. (p. 92)

5:15 PM
MH2 Massively Parallel Optical Computer, Ahmed Louri, *U. Arizona.* We present a new optical architecture for supporting massively parallel computations. The system processes 2-D arrays as basic data objects. The processing is based on the optical symbolic substitution (SS) logic. New SS rules are introduced. Implementation issues and performance analysis are also considered. (p. 96)

5:30 PM
MH3 Uses of Optical Symbolic Substitution in Image Processing: Median Filters, Abdallah K. Cherri, Mohammad A. Karim, *U. Dayton.* One- and two-dimensional optical symbolic substitution median filters are used to eliminate noise from 2-D input images. (p. 100)

5:45 PM
MH4 Parallel Addition and Subtraction in One Computing Cycle Using Optical Symbolic Substitution, G. Pedrini, R. Thalmann, K. J. Weible, *U. Neuchatel, Switzerland.* An optical symbolic substitution system is presented which performs both addition and subtraction in parallel within one computing cycle. Technique and experimental results are presented. (p. 104)

TUESDAY, FEBRUARY 28, 1989

SALON D

6:30 AM-8:00 AM BUFFET BREAKFAST

SALON FOYER

7:00 AM-6:00 PM REGISTRATION/SPEAKER CHECKIN

SALON F

8:00 AM-9:00 AM
TuA OPTICAL INTERCONNECTIONS: 1
H. John Caulfield, *University of Alabama in Huntsville, Presider*

8:00 AM (Invited Paper)
TuA1 Optical Computing Research at MCC, Steve Redfield, *Microelectronics & Computer Technology Corp.* MCC has been looking at the use of optics in computing systems to overcome barriers that are inadequately addressed by electronics. The history, motivation, and successes of these efforts are presented. (p. 110)

8:30 AM
TuA2 Modified Brewster Telescopes, Adolf W. Lohmann, Wilhelm Stork, *U. Erlangen, F. R. Germany.* For a 1-D perfect shuffle of 2-D data arrays do we need a 2:1 anamorphic imaging system? Modified Brewster telescopes are suitable. (p. 114)

8:45 AM
TuA3 Optical Implementations of Interconnection Networks for Massively Parallel Architectures, Julian Bristow, Aloke Guha, Charles Sullivan, *Honeywell, Inc.* The connectivity requirements of massively parallel architectures have been examined. Guided wave optical interconnections offer advantages over free-space implementations. Performance of the interconnection medium is reported. (p. 118)

SALON F

9:00 AM-10:00 AM
TuB OPTICAL INTERCONNECTIONS: 2
John F. Walkup, *Texas Tech University, Presider*

9:00 AM
TuB1 Implementation of Dynamic Holographic Interconnects with Variable Weights in Photorefractive Crystals, A. Marrakchi, J. S. Patel, *Bellcore.* The double-exposure and time-average holographic techniques are applied to elementary gratings in photorefractive crystals, resulting in a variable interconnection strength. In fan-in and fan-out situations where multiple phase gratings are frequency multiplexed, it is possible to alter separately each interconnection without affecting the others. (p. 124)

9:15 AM
TuB2 Energy Efficiency of Optical Interconnection Using Photorefractive Dynamic Holograms, Arthur Chiou, Pochi Yeh, *Rockwell International Science Center.* The energy efficiency (η) of 1 to $N \times N$ reconfigurable optical interconnections and $N \times N$ crossbar switches using photorefractive dynamic holograms is investigated experimentally for $N = 4$, 8, and 16. The results, for a $BaTiO_3$ crystal, on energy distribution, crosstalk, and the dependence of η on N are presented and discussed. (p. 128)

9:30 AM
TuB3 Free-Space Optical Interconnection Scheme, Alex Dickinson, Michael E. Prise, *AT&T Bell Laboratories.* We describe mechanisms for performing free-space intermodule optical interconnections within a digital electronic computer utilizing large arrays of light beams. A particular architecture and its ongoing implementation with integrated components is discussed. (p. 132)

9:45 AM
TuB4 Microoptic Systems: Essential for Optical Computing, J. L. Jewell, S. L. McCall, *AT&T Bell Laboratories.* Modern computers require compact systems or subsystems. Advantages of using microoptics in array-based optical computers are cited, and some technological progress is reviewed. (p. 136)

SALON D

10:00 AM-10:30 AM COFFEE BREAK

SALON F

10:30 AM-11:30 AM
TuC OPTICAL INTERCONNECTIONS: 3
J. Shamir, *University of Alabama in Huntsville, Presider*

10:30 AM
TuC1 2-D Optical Trimmed Inverse Augmented Data Manipulator Networks, T. J. Cloonan, M. J. Herron, *AT&T Bell Laboratories.* Two-dimensional trimmed inverse augmented data manipulator networks are defined and analyzed. An optical implementation is then described using computer-generated binary phase gratings. (p. 142)

10:45 AM
TuC2 Alignment and Performance Tradeoffs for Free-Space Optical Interconnections, Dean Z. Tsang, *MIT Lincoln Laboratory.* Efficiency, speed, and alignment sensitivity tradeoffs of free-space optical interconnections have been evaluated and an 18.8% efficient 1-Gbit/s free-space link has been demonstrated. (p. 146)

11:00 AM
TuC3 Optical Holographic Interconnection Networks for Parallel and Distributed Processing, Freddie Lin, *Physical Optics Corp.* An optical volume holographic approach is proposed to relieve the bottleneck and complexity of interconnection networks for large-scale multicomputer systems. (p. 150)

11:15 AM
TuC4 Light Effective Perfect Shuffle Using Fresnel Mirrors, Yunlong Sheng, Henri H. Arsenault, *U. Laval, Canada.* A reliable 2-D optical perfect shuffle using self-luminous inputs and light effective Fresnel mirrors is introduced. Spatial light modulators are used for the exchange box. (p. 154)

SALON F

11:30 AM-12:30 PM
TuD OPTICAL COMPUTING SYSTEMS AND COMPONENTS
Satoshi Ishihara, *Optoelectronic Industry & Technology Development Association, Japan, Presider*

11:30 AM
TuD1 Techniques for Array Illumination, Norbert Streibl, *U. Erlangen-Nuremberg, F. R. Germany;* Jurgen Jahns, *AT&T Bell Laboratories.* In an optical computing system comprising free-space interconnections, the uniform illumination of 2-D arrays of nonlinear devices is a crucial task. Various techniques using Fraunhofer diffraction, Fresnel diffraction, and spatial filtering are compared. (p. 160)

11:45 AM
TuD2 Array Illuminator Using a Grating Coupler, Mitsuo Takeda, *U. Electro-communications, Japan;* Toshihiro Kubota, *Kyoto Institute of Technology, Japan.* An integrated optical array illuminator is proposed. The principle and preliminary experiments of an array illuminator using grating couplers are presented. (p. 164)

12:00 M
TuD3 Substrate Mode Holograms for Optical Interconnects, Raymond K. Kostuk, Masayuki Kato, Yang-Tung Huang, *U. Arizona.* The advantages and design considerations for free-space holographic interconnects are discussed. Substrate mode holograms for this application are introduced and experimentally demonstrated. (p. 168)

12:15 PM
TuD4 Hybrid Acoustooptic Spectrum Analyzer for Radio Astronomy with Semiconductor Lasers, N. N. Evtihiev, V. V. Perepelitsa, *Moscow Engineering Physics Institute, USSR;* N. A. Esepkina, S. V. Pruss-Zhukovsky, O. N. Vlasov, S. K. Kruglov, *Leningrad Polytechnic Institute, USSR.* A hybrid acoustooptic spectrometer with a semiconductor laser is investigated. The whole system is controlled by computer, and it provides high SNR and a large frequency band. Characteristics of phased array diode lasers in the spectrometer are presented. (p. 172)

12:30 PM-2:00 PM LUNCH BREAK

SALON F

2:00 PM-2:45 PM
TuE OPTICAL COMPUTING SYSTEMS: 1
Steven C. Gustafson, *University of Dayton, Presider*

2:00 PM (Invited Paper)
TuE1 Perspectives in Optical Computing, John Neff, *DuPont Corp.* (p. 176)

2:30 PM
TuE2 Energetic Advantage of Analog Over Digital Computing, H. John Caulfield, *U. Alabama in Huntsville.* I show that some analog computers have no minimum energy per calculation. This arises from the quantum mechanical nature of the photon-computer interaction. (p. 180)

SALON F

2:45 PM-3:30 PM
TuF OPTICAL COMPUTING SYSTEMS: 2
George Eichmann, *CUNY-City College, Presider*

2:45 PM
TuF1 Tantalus and Optical Computing, W. Thomas Cathey, *U. Colorado.* Optics promises several advantages over electronics in computation. We explore these promises, the ones that are likely to be fulfilled, and those that remain tantalizing but unattainable. (p. 186)

3:00 PM
TuF2 The Mock Counter, Ann B. Yadlowsky, Harry F. Jordan, *U. Colorado.* An optoelectronic emulation of an optical counter is described. It is a first step toward a complete bit-serial optical computer based on fiber optics. (p. 189)

3:15 PM
TuF3 Cascade Connective Optical Logic Processor Using 2-D Electrophotonic Devices, S. Kawai, Y. Tashiro, H. Ichinose, K. Kasahara, K. Kubota, *NEC Corp., Japan.* A new optical logic algorithm for an optical processor with cascade connectability is presented. Construction and operation of a compact optical processor have been successful. (p. 193)

SALON D

3:30 PM-4:00 PM COFFEE BREAK

SALON F

4:00 PM-5:00 PM
TuG OPTICAL COMPUTING SYSTEMS: 3
William Miceli, *U.S. Office of Naval Research, Presider*

4:00 PM
TuG1 Computational Origami: the Folding of Circuits and Systems, A. Huang, *AT&T Bell Laboratories.* A technique which regularizes and folds circuits and systems to match the parallelism of optics is presented. (p. 198)

4:15 PM
TuG2 All-Optical Game of Life Computer, Lawrence H. Domash, *Foster-Miller, Inc.;* Mark Cronin-Golomb, *Tufts U.* A photorefractive all-optical cellular automaton computer is proposed, based on the computationally universal Game of Life model. The design addresses generic problems of thresholding, binarization, storage, timing, and error propagation. (p. 202)

4:30 PM
TuG3 Optical Disk-Based Correlation Architectures, Demetri Psaltis, Mark A. Neifeld, Alan Yamamura, *California Institute of Technology.* We describe and experimentally demonstrate three optical image correlator architectures that are implemented using optical memory disks. More than 10,000 image correlations per second is achievable. (p. 206)

4:45 PM
TuG4 Proposal for an Optical Content Addressable Memory, Miles J. Murdocca, *AT&T Bell Laboratories;* John Hall, Saul Levy, Donald Smith, *Rutgers U.* A content addressable memory design that demands high throughput is proposed for arrays of optically nonlinear logic gates interconnected in free space. (p. 210)

SALON F

5:00 PM-6:00 PM
TuH OPTICAL COMPUTING SYSTEMS: 4
Adolf W. Lohmann, *University of Erlangen-Nuremberg, F. R. Germany, Presider*

5:00 PM
TuH1 Optical Outer Product Look-up Table Architectures for Residue Arithmetic, Mark L. Heinrich, Ravindra A. Athale, Michael W. Haney, *BDM Corp.* An optical outer product architecture that minimizes gate count is described to implement arbitrary integer-valued functions of two variables in a single gate delay using a position-coded residue representation. (p. 216)

5:15 PM
TuH2 Optical Transversal Filter with Variable Weights, Debra M. Gookin, Mark H. Berry, *U.S Naval Ocean Systems Center.* This fiber optic tapped delay line transversal filter uses computer controlled integrated optical two-by-two couplers to vary the tap weights. (p. 220)

5:30 PM
TuH3 Integrated Optoelectronic Cellular Array for Fine-Grained Parallel Processing Systems, M. Hibbs-Brenner, S. D. Mukherjee, M. P. Bendett, *Honeywell, Inc.;* A. R. Tanguay, Jr., *U. Southern California.* The design, scalability, and potential uses of an optoelectronic cellular array are described. Results are presented on the fabrication of the integrable components. (p. 223)

5:45 PM
TuH4 Optoelectronic Parallel Processing Arrays: System Architecture and Progress Toward a Prototype, Timothy J. Drabik, Thomas K. Gaylord, *Georgia Institute of Technology.* A VLSI-based optoelectronic parallel processing array methodology is presented and contrasted with all-optical approaches. Experimental results relating to a prototype system are discussed. (p. 227)

TUESDAY, FEBRUARY 28, 1989—*Continued*

SALON D

7:30 PM–9:30 PM
TuI POSTER SESSION
Refreshments served

TuI1 Programmable Emulation with the Optical Reconfigurable Logic Array, F. F. Zeise, P. S. Guilfoyle, *OptiComp Corp.* We discuss an optically implemented reprogrammable logic array using control logic to compute ALU primitives for emulating a general purpose programmable computer. (p. 232)

TuI2 Optical Multiple-Valued Logic Using Composite Bistable Laser Diodes or Light-Emitting Diode Circuits, Shutian Liu, Chunfei Li, Jie Wu, Yudong Liu, *Harbin Institute of Technology, China.* Using composite bistable laser diode/light-emitting diode circuits, we demonstrate optical multiple-valued logic functions that have the potential for optical signal processing and optical computing. (p. 236)

TuI3 Optical-Holographic-Associative-Memory-Based Parallel Register Transfer Processor, George Eichmann, Andrew Kostrzewski, Dai Hyun Kim, Yao Li, *CUNY–City College.* Optical-holographic-associative-memory-based register transfer microoperations are proposed. Experimental results obtained with a hybrid optical parallel digital register transfer processor are presented. (p. 240)

TuI4 Optical Network Design for a Bit-Serial Parallel Processor, Adolf W. Lohmann, Gregor Stucke, *U. Erlangen-Nuremberg, F. R. Germany.* A bit-serial parallel processor under SIMD control may be extended into a hybrid system with an optical network for shuffling and cyclic shifting. (p. 244)

TuI5 Symbolic Substitution-Based Parallel Adder/Subtracter, S. Barua, *California State U.* A highly parallel optical adder/subtracter based on symbolic substitution is presented. The architecture performs addition/subtraction in two stages regardless of the length of the operands. (p. 246)

TuI6 Using a Symbolic Substitution Method on Optical Matrix Multiplication, Kuo-fan Chin, Minxian Wu, Shaomin Zhou, *Tsinghua U., China.* The method of symbolic substitution combined with an outer product of matrices is proposed for solving optical matrix multiplication, with high accuracy and calculating speed. (p. 250)

TuI7 Correlation Algorithm and Architecture for Optical Complex Discrete Fourier Transformation, Hongxin Huang, Liren Liu, Zhijiang Wang, *Shanghai Institute of Optics and Fine Mechanics, China.* A multichannel incoherent optical correlator for performing complex DFT is proposed. The matrix-code method of complex DFT is discussed, and some properties are demonstrated. (p. 254)

TuI8 GaAs Waveguide Microlenses and Lens Arrays with Uses in Data Processing and Computing, T. Q. Vu, C. S. Tsai, *UC-Irvine.* We report the first, we believe, successful fabrication of planar waveguide microlenses and lens arrays in GaAs by using ion milling. The single lenses and lens arrays of analog Fresnel, chirp grating, and hybrid types fabricated have shown high efficiencies and near diffraction-limited spot sizes. (p. 258)

TuI9 Intelligent Optical Processors, Anjan Ghosh, *U. Iowa.* Ideas of bimodal optical computing and matrix-preconditioning are combined to develop intelligent optical processors that adapt the computations depending on data to produce accurate solutions in a given time. (p. 262)

TuI10 Uses of a Polarization-Based Optical Processor, Abraham P. Ittycheriah, John F. Walkup, Thomas F. Krile, *Texas Tech U.* A cascadable polarization-based optical processor is used to perform logic functions and Walsh and Haar transforms. Critical parameters are presented. (p. 266)

TuI11 Guided Wave Vector-Matrix Multiplier, Mark H. Berry, Debra M. Gookin, *U.S. Naval Ocean Systems Center.* A versatile fiber optic and integrated optic system for forming vector-matrix products is described. Large matrices can be input and multiplied in nanoseconds. (p. 270)

TuI12 Image Processing Using Polarization-Encoded Optical Shadow-Casting. 2: Edge Detection, A. A. S. Awwal, Mohammad A. Karim, *U. Dayton.* A powerful space-variant image processing technique, namely, edge detection, is performed using a polarization-encoded optical shadow-casting system. (p. 272)

TuI13 Polarization-Based Optical Computing Using Liquid Crystals, Johannes D. Roux, F. Wilhelm Leuschner, *U. Pretoria, South Africa.* The demonstration of a novel parallel optical logic gate using polarization-encoded logic and liquid crystal spatial light modulators is reported. (p. 276)

TuI14 Single-Element 2-D Bragg Cells for Optical Computing, Jolanta I. Soos, Douglas C. Leepa, Ronald G. Rosemeier, *Brimrose Corp. of America.* The cubic crystallographic symmetry of gallium phosphide makes this crystal a good candidate for single-element 2-D acoustooptic deflectors. (p. 280)

TuI15 Multiplexed Waveguide Hologram for Optical Processing and Computing, Freddie Lin, *Physical Optics Corp.* Waveguide phase holograms which have high-density storage capacity and a large number of multiplexed channels are useful in optical computing and signal processing. (p. 281)

TuI16 Dynamic Optical Interconnection for Reconfigurable Neural Networks, Bradley D. Clymer, Qiu-Shi Ren, *Ohio State U.* An optical interconnection system which allows dynamic reconfiguration of a neural network is discussed. We present a novel use of a two-level holographic reconstruction system. (p. 285)

TuI17 Entropy-Optimized Filter for Pattern Recognition, Uri Mahlab, Michael Fleisher, Joseph Shamir, *Technion, Israel.* We introduce the entropy function as a new concept in the generation of spatial filters for pattern recognition and classification. Computer simulation experiments demonstrate efficient recognition of single patterns or classes even when these are submerged in high-level random noise. (p. 289)

TuI18 Synchronous Discrete Neural Networks for Minimization, Hyuk Lee, *Polytechnic Institute of New York.* A general neural minimization algorithm, which can be applied to arbitrary types of polynomial energy function, is presented. The algorithm can be operated in a synchronous as well as asynchronous way. The synchronous algorithm can be implemented by highly parallel optical systems. (p. 292)

TuI19 Optical Error Correction by Adaptive Thresholding, David Kagan, *California Institute of Technology;* Harry Friedmann, *Bar-Ilan U., Israel.* Using a control beam, which is proportional to a detected error, to shift the soft thresholding output curves of a nonlinear optical device, error correction can be made. (p. 296)

TuI20 Neural Network Models Based on Optical Resonator Designs, Steven C. Gustafson, Gordon R. Little, *U. Dayton.* Neural network models consistent with optical resonator designs that may include dynamic holograms and thresholded phase conjugate mirrors are considered. (p. 300)

TuI21 Implementation of the NETL Knowledge-Base System with Programmable Optoelectronic Multiprocessor Architecture, Fouad Kiamilev, Sadik Esener, *UC-San Diego.* Programmable optoelectronic multiprocessor architecture is well suited for implementing the massively parallel NETL knowledge-based system, because the symbolic data structure can be directly mapped onto POEM hardware. (p. 303)

Tul22 Trainable Optical Network for Pattern Recognition, John H. Hong, Pochi Yeh, *Rockwell International Science Center.* An optical implementation of a single layer network for pattern recognition is described, in which both subtractive and additive changes of the weights can be made. (p. 307)

Tul23 Optical Implementation of Association and Learning Based on PRIMO/Light Valve Devices, U. Efron, Yuri Owechko, *Hughes Research Laboratories.* An optical outer product associative memory system is proposed, based on the PRIMO and light valve devices, that is capable of supervised learning. (p. 311)

Tul24 Thermal Nonlinear Microcavity and Optical Computing, C. Godsalve, E. Abraham, *Heriot-Watt J., U.K.* Characteristics of thermal nonlinear microcavities with respect to spot size, radius, height, and thermal conductivities are analyzed, and their use in optical computing is considered. (p. 315)

Tul25 Two-Beam Coupling Polarization Properties in BSO Using Alternating Electric Fields, G. Pauliat, G. Roosen, *Institute of Theoretical & Applied Optics, France.* The coupled-wave theory is used to predict all the properties of the amplified beam. In some specific experimental conditions these characteristics are time independent. (p. 318)

Tul26 Band-Tunable Multichannel Scale Invariant Pattern Recognition System with Zone Plates, Minhua Liang, Shudong Wu, Liren Liu, Zhijiang Wang, *Shanghai Institute of Optics & Fine Mechanics, China.* A new system of scale invariant pattern recognition, which has a large scale-tunable and movable range and utilizes zone plates, is investigated. (p. 322)

Tul27 Hybrid Optical Processing for Measuring the Refractive-Index Profile in Single-Mode Fibers, T. Nobuyoshi, *Okayama-Rika U., Japan.* A new measuring technique for the calculation of the refractive-index profile in single-mode fibers is proposed through the use of hybrid optical processing. (p. 326)

Tul28 Hardware Requirement for 2-D Image Convolution in Electrooptic Systems, M. Mary Eshaghian, D. K. Panda, V. K. Prasanna Kumar, *U. Southern California.* We show a lower bound on the volume requirement of any electrooptical chip for 2-D image convolution. The results are based on a parallel model of computation using optical interconnects. (p. 330)

Tul29 Optical Systems Tolerances for Symmetric Self-Electrooptic Effect Devices in Optical Computers, Nick C. Craft, *Heriot-Watt U., U. K.;* Michael E. Prise, *AT&T Bell Laboratories* We investigate the optical tolerance requirements of differential optical logic devices such as the symmetric SEED and describe a two-element array generation technique. (p. 334)

Tul30 Tolerance Analysis and Design of Optical Processors, J. F. Snowdon, B. S. Wherrett, *Heriot-Watt U., U.K.* A tolerance design methodology for optical processors is presented. Examples of both fixed pipeline and programmable processor architectures are analyzed and design strategies discussed. (p. 338)

Tul31 Optimization of Binary Phase Only Filters with a Simulated Annealing Algorithm, Myung Soo Kim, Michael R. Feldman, Clark C. Guest, *UC San Diego.* A simulated annealing algorithm to encode optimum binary phase only filters is introduced. It is shown that similar patterns indistinguishable with conventional encoding methods, are clearly distinguished with the optimized filter. (p. 342)

Tul32 Computer Holographic Elements Using PostScript Laser Printers, Lawrence Domash, Philip Levin, *Foster Miller, Inc.* The PostScript software environment and industrial laser typesetters with 10-μm resolution are capable of conveniently producing masks for a variety of diffractive optical elements. (p. 346)

SALON D

6:30 AM-8:00 AM BUFFET BREAKFAST

SALON FOYER

7:00 AM-5:30 PM REGISTRATION/SPEAKER CHECKIN

SALON F

8:00 AM-9:00 AM
WA DIGITAL OPTICAL COMPUTING: 1
John A. Neff, *DuPont Corporation, Presider*

8:00 AM (Invited Paper)
WA1 Digital Optical Computing with Fibers and Directional Couplers, Harry F. Jordan, *U. Colorado.* We discuss the goal of the Digital Optical Computing program of the University of Colorado's Center for Optoelectronic Computing Systems: to design and demonstrate a prototype of a stored program optical computer using the knowledge base developed in connection with electronic digital computers. (p. 352)

8:30 AM
WA2 Dynamic Optical Processing for Parallel Digital Addition and Subtraction, Takashi Kurokawa, Seiji Fukushima, *NTT Optoelectronics Laboratories, Japan;* Hideo Suzuki, *NTT Communications & Information Processing Laboratories, Japan.* Dynamic parallel arithmetic processing is demonstrated for digital addition and subtraction. Real-time operation is achieved by synchronous control of logic gates and latche memories. (p. 356)

8:45 AM
WA3 Flexible-Structured Computation Based on Optical Array Logic, Jun Tanida, Masaki Fukui, Yoshiki Ichioka, *Osaka U., Japan.* Flexible-structured computation is considered with array logic. As examples of such a paradigm, a Turing machine and a systolic computing array are demonstrated. (p. 360)

SALON F

9:00 AM-10:00 AM
WB DIGITAL OPTICAL COMPUTING: 2
Ravindra A. Athale, *BDM Corporation, Presider*

9:00 AM
WB1 Reconfigurable Programmable Optical Digital Computer, P. S. Guilfoyle, F. F. Zeise, *OptiComp Corp.* Previous optical computing schemes offered analog or quasidigital accuracies with a single fixed primitive. This paper describes how programmable, arbitrary bit length, all-digital central processing unit computations are now possible. (p. 366)

9:15 AM
WB2 Programmable Logic Gate Array and Its Use in a Reconfigurable Network Based on Modified Sign Digit, Yoshiki Suzaki, Toyohiko Yatagai, *U. Tsukuba, Japan.* A ternary programmable parallel logic gate array is designed to make dynamic interconnection. We made a prototype electrooptic gate and simulated a MSD adder. (p. 370)

9:30 AM
WB3 Optical Programmable Binary Symmetric Logic Module, Yao Li, Berlin Ha, George Eichmann, *CUNY-City College.* An optical programmable binary symmetrical logic module (OPBSLM) is proposed. Diversified uses of the OPBSLM and an experimental demonstration are discussed. (p. 374)

9:45 AM
WB4 Optical Realization of Arithmetic Operations in a Ternary Number System, E. M. Dianov, A. A. Kuznetsov, S. M. Nefjodov, G. G. Voevodkin, *Academy of Sciences of the U. S. S. R.* Optical realization of addition and multiplication operations is possible in a ternary number system. Optically controlled LCLV bichromatic light source and optical feedback are used.(p. 378)

SALON D

10:00 AM-10:30 AM COFFEE BREAK

SALON F

10:30 AM-11:15 AM
WC DIGITAL OPTICAL COMPUTING: 3
Alexander A. Sawchuk, *University of Southern California, Presider*

10:30 AM (Invited Paper)
WC1 Business and Technological Issues for the Commercialization of Optical Computing, Henry Kressel, *E. M. Warburg, Pincus & Co.* The major technological elements encompassed by optical computing are discussed in terms of their applications. Comparisons with the successful commercial introduction of other optical technologies are made to highlight the elements contributing to their success. (p. 384)

11:00 AM
WC2 All-Optical Full-Adder Based on a Zinc Sulfide Optical Bistable Device, Ruibo Wang, Zizhong Zha, Lei Zhang, Chunfei Li, *Harbin Institute of Technology, China.* Operation of a single-gate full-adder with on-axis input has been demonstrated experimentally. A multibit full-adder circuit based on a single ZnS interference filter is proposed. (p. 385)

SALON F

11:15 AM-12:30 PM
WD MATRIX ALGEBRAIC PROCESSING
William T. Rhodes, *Georgia Institute of Technology, Presider*

11:15 AM
WD1 Electrooptic Architecture for Solving General Sparse Linear Systems, M. Mary Eshaghian, V. K. Prasanna Kumar, David W. Tang, *U. Southern California.* We present an efficient electrooptical implementation of the iterative solution of general sparse linear systems. Our design achieves an optimal time of O (log m) for each iteration, where m is the number of variables. (p. 390)

11:30 AM
WD2 Digital Optoelectronic Processor Array Architectures for Vector-Matrix Multiplication, Michael R. Feldman, Clark C. Guest, *UC-San Diego.* Interconnection networks for optically interconnected electronic processor arrays have been developed. These networks have area and time growth rates for vector-matrix multiplication close to fundamental lower bounds. (p. 394)

11:45 AM
WD3 Hybrid Optoelectronic Coprocessor Implementation Inside a Computer Workstation, Guy Lebreton, *U. Var, France;* Remy Frantz, *Compagnie Europeenne des Techniques de l'Ingenierie Assistee, France.* A loop with continuous relaxation is formed between sixteen parallel laser diodes and photodiodes, connected through an acousto-optic matrix (multichannel Bragg cell with multiplexed frequencies). (p. 398)

12:00 M
WD4 Theoretical Description of the Bimodal Optical Computer, A. V. Scholtz, E. van Rooyen, *U. Pretoria, South Africa.* A theoretical analysis of the analog feedback loop of the bimodal optical computer (including the multiple eigenvalue case) is discussed, with reference to convergence requirements. (p. 402)

12:15 PM
WD5 Sequential Logic Operation Using an Optical Parallel Processor Based on Polarization Encoding. Masashi Hashimoto, Ken-ichi Kitayama, *NTT Transmission Systems Laboratories, Japan;* Naohisa Mukohzaka, *Hamamatsu Photonics K. K., Japan.* Experimental sequential logic operation by an all-optical parallel processor using polarization encoding is shown. Optical latches and a spatial decoder in the optical feedback path are key elements. (p. 406)

12:30 PM-2:00 PM LUNCH BREAK

SALON F

2:00 PM-5:30 PM
WAA JOINT PHOTONIC SWITCHING AND OPTICAL COMPUTING PLENARY SESSION
Joseph W. Goodman, *Stanford University, Co-presider*
John E. Midwinter, *University College London, U.K., Co-presider*

2:00 PM (Plenary Paper)
WAA1 OEIC Technology for Photonic Switching, S. Yamakoshi, *Fujitsu Laboratories, Ltd., Japan.* OEIC technology promises to construct new optical systems such as photonic switching, routing and other optical processing operations. The state-of-the-art and future prospects of OEICs for photonic switching are discussed. (p. 412)

2:45 PM (Plenary Paper)
WAA2 Quantum Well Devices for Optical Computing and Switching. David A. B. Miller, *AT&T Bell Laboratories.* Quantum well electroabsorptive self-electrooptic-effect devices are attractive for 2-D arrays for switching and processing. Novel integrated configurations and progress toward arrays are summarized. (p. 413)

SALON D

3:30 PM-4:00 PM COFFEE BREAK

4:00 PM (Plenary Paper)
WAA3 Switching in an Optical Interconnect Environment, Joseph W. Goodman, *Stanford U.* The requirements placed on switching in an optical interconnect environment differ significantly from those present in long distance telecommunications, allowing new approaches to switch realization. (p. 416)

4:45 PM (Plenary Paper)
WAA4 Relationship Between Photonic Switching and Optical Computing, H. Scott Hinton, *AT&T Bell Laboratories.* An outline of the relationship between the hardware requirements of photonic switching and optical computing systems is presented. (p. 418)

5:30 PM-5:45 PM
CLOSING REMARKS
Alexander A. Sawchuk, *University of Southern California*
C. Lee Giles, *Air Force of Scientific Research*

SALON D

6:00 PM-7:30 PM CONFERENCE RECEPTION

MONDAY, FEBRUARY 27, 1989

SALON F

8:30 AM–9:30 AM

MA1–MA2

NEURAL SYSTEMS: 1

C. Lee Giles, Air Force Office of Scientific Research, *Presider*

Optical Implementations of Neural Computing

Ravindra A. Athale

The BDM Corp.

7915 Jones Branch Dr.

McLean, VA. 22102

BACKGROUND:

Achieving performance comparable to human beings in speech recognition, visual perception, motor control and knowledge acquisition, representation, and processing is one of the most difficult and exciting challenges facing the information processing research community. Recently neural net models of computation have been investigated as a novel approach for solving these problems. These proposed models are only loosely based on the known and postulated characteristics of biological systems and no claim is usually made for these models to be biologically accurate.

In spite of the diversity in the characteristics of neural net models reported in the literature, the following common features emerge:

1. A very large number of relatively simple processing elements (neurons) are arranged in densely interconnected layers,
2. the interconnections between the Processing Elements (PEs) have analog weights indicating the strength of the interconnections,
3. the interconnection weights evolve under the influence of external inputs without a central controller.

In different neural net models, the complexity of the operations performed within the PE varies from a simple sum-of-products to spatio-temporal differentiation of weighted inputs with some local storage. Similarly the details of the dynamics of the interconnection weights are also heavily model-dependent.

The research in neural nets can be divided into three main areas:

(1) theoretical investigation of new models,
(2) applications of existing models to interesting problems,
(3) hardware implementations (optical and electronic) of existing models.

In this paper I will primarily focus on the optical implementation of existing models.

OPTICAL IMPLEMENTATION OF NEURAL NET MODELS:

The simplest neural net models are based on a PE that evaluates a weighted sum of its input signals and applies a nonlinear transfer function to the resulting scalar value before transmitting it as its output signal to subsequent stages. This functionality can be realized in analog optics via systems that are largely identical to the correlators and matrix processors previously investigated in optical computing. The nonlinear transfer function can be achieved through the response of the active devices (all-optical or hybrid electrooptical). A more complex PE becomes harder to directly realize in optics except in the cases where the desired response maps onto the physics of the active device, e.g. exponentially weighted time-integration achieved in a device due to its finite response time. When such a match is not found, analog electronic components may be necessary in addition to the optical detectors and modulators/sources. Thus the PE implementation in optical neural nets will contain varying amounts of electronics depending its functional complexity.

The analog-weighted interconnects between the PEs can be implemented optically using several different technologies. If the neural net model does not require real-time adaptability of the weights, then film masks or holograms can be used to encode the analog weights, while conventional optical elements (spherical and cylindrical lenses) perform the signal distribution function. Spatial Light Modulators (SLMs) and real-time holographic materials (thermoplastics, photorefractive crystals) become necessary when the models demand real-time adaptability of the interconnection weights. For both cases, 2-D (planar) and 3-D (volume) analog storage techniques can be employed. The former has the advantage of relative technological maturity while the latter has a higher theoretical storage capacity.

Several optical neural net implementations have been proposed and demonstrated experimentally during the past four years. References 1 and 2 represent a comprehensive collection of papers describing this research. One class of optical neural net systems are based on optical matrix processing architectures where the PEs are represented by a 1-D vector. The second class of systems are primarily derived from Fourier plane correlators which use 2-D images to represent PE signals. Both of these systems utilize planar (2-D) storage media. The third class of optical systems uses volume holographic storage of analog interconnect weights. Finally a number of optical systems based on nonlinear resonators have been investigated. In these systems the similarity between the dynamics of the neural net model and the resonator is exploited.

The future directions of research in optical neural computing will be heavily influenced by developments in the studies of new models as well as the applications of these models. The ease of implementing simple PEs in analog optics, the high analog storage density of 2-D (10^6) and 3-D (10^9) optical systems, and the parallel access to these weights represent the primary advantages of optics as applied to neural net models. Hence it will be the ability of neural net models based on these features in attacking interesting problems that will determine the utility of optical neural net processors.

REFERENCES

1. Special issue of Applied Optics, 1 December 1987, Vol.26, No. 23.

2. "Neural Network Models for Optical Computing", Proceedings of S.P.I.E., Vol. 882, 1988.

Electronic vs. Optical Implementations of Neural Networks*

Jay P. Sage
Massachusetts Institute of Technology
Lincoln Laboratory
Lexington, MA 02173-0073

1 Introduction

This paper will address for the optical community the relative advantages and limitations of electronic neural network implementations in contrast to optical implementations. Its intent is by no means to be adversarial. It is hardly necessary to say that today electronics has an edge over optics in this field. The aim of the paper will be to help indicate the areas where electronic and optical implementations can each make their most important contributions and the areas in which advances in technology, particularly in optical technology, will be required.

We will begin with a discussion of the general criteria that determine the general suitability of a technology. Next, we will present a working definition of neural networks. Then we will look in very general terms at the suitability of electronics and optics in meeting the requirements of neural network implementations. Three areas will be considered: computation, memory, and connectivity. Finally, examples of electronic neural network implementations from Lincoln Laboratory will be described to illustrate the earlier points. These examples will not be covered in this summary, however.

2 Performance Criteria

The success of a technology in meeting practical application requirements depends on a number of factors, among them the following:

- speed
- density
- processing power
- power consumption
- capability
- maturity
- interface & control
- cost

The factors listed on the left are the technical factors that researchers readily appreciate. We can measure them easily, their significance is obvious, and I will say only a little about them. As for speed, some of the active optical devices required in neural networks are today quite slow and will quite likely never reach the speed of electronic circuits. Density, on the other hand, is likely to favor optics, especially where optics can take advantage of the third spatial dimension.

*This work was supported by the Department of the Air Force and the Defense Advanced Research Projects Agency.

Processing power involves a combination of speed and density, and, as the biological brain shows so dramatically, a technology can fall rather short in one area and still come out a winner! How much processing power is required is not known, but my view is that if the amount is small, special neural network hardware (and probably neural networks *period*) will not be needed. The amount of processing power required in each module of a complete neural network subsystem is a separate question. Power consumption is another area that is problematic for optical networks, since optical nonlinearities tend to occur only at rather high power levels.

Now we turn to the factors in the right column. Obviously, before a technology can be useful for an application, it must be able to perform the necessary tasks. Both optics and electronics offer great promise for neural network implementations, but it is probably fair to say that *neither* has yet demonstrated the capability completely and conclusively.

The time-frame within which practical solutions can be developed depends on the maturity of the technology. Electronics, because of the explosion in digital applications, is now a highly mature field. However, analog electronics, and especially nonprecision analog electronics of the type that may play an important role in neural networks, has received little attention. Thus technology development even in electronics is very badly needed. Nevertheless, electronic neural networks can benefit from the advances in fabrication technology and can take advantage of digital control and interface circuitry. Nonlinear optics, on the other hand, is relatively immature and will require significant technology development.

A neural network will be only a part of an information processing system, and the network will have to interface to and be controlled by other components in the system. Since the rest of the system is most likely to be electronic, optical neural networks face a handicap relative to electronic ones.

Being a scientist, I had originally left the last factor, cost, off my list, though it is in almost all cases the one with the *ultimate* importance!

3 Definition

As a point of departure, let us begin by attempting to define a neural network. This is not easy to do, and it took considerable reflection before participants in the DARPA Neural Network Study accepted, as a basis for discussion, the following three-part definition (slightly modified):

- A system composed of many simple processors, fully or sparsely connected, whose function is determined by the connection topology and strengths.
- This system is capable of a high level function such as adaptation or learning with or without supervision as well as lower level functions such as vision and speech preprocessing.
- The function of the simple processors and the structure of the connections are inspired by the study of biological nervous systems.

4 Computation

From the first section of the definition we see that a neural network comprises a large number of simple processors rather than the small number of complex processors found in conventional computers. We should note that "simple" does not necessarily have its obvious meaning here. The electrochemical processes in biological networks are, in fact, so complex that even supercomputers have trouble simulating them. Rather, the computational elements in a neural network are simple in the sense that they do not perform a multitude of different tasks as a CPU does; they perform a single, dedicated task, perhaps using a physical or chemical process.

Neural networks perform three different computational operations in two computational structures. The structural elements are (1) the neurons and (2) the synapses, each of which connects a pair of neurons. The first computation is performed by the synapse, which takes a signal from one neuron, operates on it in some (generally nonlinear) way that depends on the stored state of the synapse (often called the weight), and produces an output which is delivered to another neuron.

The second computation is performed in the neuron, which combines the inputs from all the synapses connected to it and produces an output that depends (again generally in a nonlinear way) on the inputs. In practice, the neural computation is greatly simplified by assuming that the individual synaptic signals are first combined by linear summation and that only this sum is operated on in a nonlinear way. For analog electronic implementations, this simplifying assumption is of critical importance. The same is probably true for optical implementations as well.

A third computational operation is required for the learning referred to in the second paragraph of the definition. Although learning could be accomplished by introducing structural changes in the network architecture (creating or removing neurons or synapses or changing the basic topology of their interconnection), learning is generally assumed to be limited to changes in the synaptic parameters that affect their transfer functions. For this, the synapses must perform a second type of computation, also nonlinear but different from the computation performed during network readout.

In general, electronics offers great flexibility in the kinds of computational operations it can perform, since high quality nonlinear devices are readily available. At one extreme, any and all the computations can be performed using digital logic. Thus, the issue is not whether the required network computations *can* be performed by an electronic network but whether they can be performed *efficiently*, that is, by compact, high speed circuits. It is this issue that drives implementers to analog electronics. For example, consider the addition operation required in each neuron. Parallel addition of 100 8-bit digital inputs from synapses would require an enormous ALU, but Kirchhoff's Law can add 100 analog currents effortlessly using a simple, tiny conductor. Optical detectors also perform this operation fast, accurately, and effortlessly.

5 Memory

A neural network requires a mechanism for storing the synaptic state parameters. For nonlearning networks, some kind of static storage, such as a mask-defined resistors or an

optical transparency, can be used. For adaptive networks, however, memory becomes a critical issue. Digital electronic memory is easily implemented, but it is not particularly compact, as one of our example electronic networks will illustrate. Neither electronics nor optics has yet demonstrated an ideal analog memory, one that retains its information faithfully yet is easy to update electrically or optically. Another electronic example will illustrate the technology we have been developing for this purpose.

6 Connectivity

The first section of the neural network definition points to two features that determine the function of a network. We have already discussed one, the strengths of the connections in the network. The other one is the topology or architecture of those connections. This is one of the key differences between neural networks and conventional computers, where a software program determines the function.

A corresponding implementation issue is the types of connectivity that a given technology offers. Biology can construct its networks in three dimensions, while electronics is essentially planar. Some network models, such as the Grossberg/Mingolla vision network, involve complex or irregular patterns of connectivity between neurons and probably cannot be implemented effectively using electronics.

Other networks map very well into a planar architecture. When only nearest neighbor connectivity is required, the neurons can be laid out efficiently in a two-dimensional array with the synapses in between. One serious problem remains, however: getting signals out of the network. External connections to a chip are limited in number and must generally be placed along the periphery of the chip. At the opposite extreme of connectivity — full interconnection between two sets of neurons — electronics is again efficient. In this case, the synapses greatly outnumber the neurons. The input neurons are arrayed in one line, while the output neurons are arrayed in a second, perpendicular line. The synapses then fill densely a two-dimensional rectangle. This arrangement is probably the most efficient one for electronics, because the neurons are at the edge of the chip where external connections can be made.

Connectivity is the one area where optics may have a clear advantage over electronics. There can be a very large number of parallel I/O channels to and from a network, and when synapses are packed into three dimensions the total density of processing elements can be much higher.

7 Final Remarks

All neural network technology development, and especially that in optics, must turn its attention from "toy" problems that illustrate only the particular strength of that technology and begin to address the whole range of implementation issues. These include the development of the necessary nonlinear devices not only for the network functions *per se* but also for the control and interface functions needed to implement hierarchical neural network systems.

MONDAY, FEBRUARY 27, 1989

SALON F

9:30 AM–10:00 AM

MB1–MB2

NEURAL SYSTEMS: 2

Henri Arsenault, Laval University, Canada, *Presider*

Implementation of Dynamic Hopfield-like Networks using Photorefractive Crystals

Jeff Wilde and Lambertus Hesselink

Department of Applied Physics
Stanford University, Stanford, CA. 94305

Abstract

We present an architecture for optically implementing a digital associative memory using a coherent optical system. Bipolar information is holographically phase encoded in a photorefractive crystal and upon readout is interferometrically phase decoded.

Summary

Numerous schemes for optical implementation of neural networks have recently been proposed and demonstrated [1-4]. From among the many possible approaches, those involving holographic storage of information appear very attractive due to the large potential memory capacity. In addition, if the recording medium is dynamic, it is possible to make "trainable" feed-forward networks as discussed by Psaltis *et al* [5]. In this paper we focus on the implementation of a readily reconfigurable Hopfield-like network in which a set of input/output vector pairs with bipolar bits (i.e. ± 1) are stored via a sum of weighted outer-products.

To optically encode a vector, we let each bipolar bit be represented by a phase-encoded plane wave having a unique direction of propagation. The fact that we have chosen a plane-wave representation is not essential, but merely simplifies the analysis. Assuming that each beam has unity amplitude, the general representation of the i'th element is simply $e^{i(\vec{k}_i \cdot \vec{r} - \omega t + \phi)}$, where ϕ assumes a value of either 0 or π. To record the outer-product between an input/output pair of vectors $(\vec{u}, \vec{v})$, all waves representing the two vectors overlap within the volume of a photorefractive crystal as shown in figure 1. The mutual interference between the two sets of beams produces the desired outer-product matrix $\mathbf{W} = \vec{u}\vec{v}^T$. Each matrix element W_{ij} is given by the time-averaged interference pattern formed between corresponding beams,

$$\begin{aligned} W_{ij} &= |e^{i(\vec{k}_{ui} \cdot \vec{r} + \phi_{ui})} + e^{i(\vec{k}_{vj} \cdot \vec{r} + \phi_{vj})}|^2 \\ &= = 2 + [e^{i(\vec{k}_{wij} \cdot \vec{r} + \phi_{ui} - \phi_{vj})} + \text{c.c.}], \end{aligned} \tag{1}$$

where the grating vector $\vec{k}_{wij} \equiv \vec{k}_{ui} - \vec{k}_{vj}$. Since the values of the phase factors are limited to 0 and π, we have $\Phi_{ij} = \phi_{ui} - \phi_{vj} = \phi_{ui} + \phi_{vj}$. Assuming that the modulation depth of any one grating is sufficiently reduced by the presence of other beams, the photorefractive response is linear and is given by.

$$\Delta n_{ij} = \kappa \, e^{i(\vec{k}_{wij} \cdot \vec{r} + \Phi_{ij} + \theta_o)}, \tag{2}$$

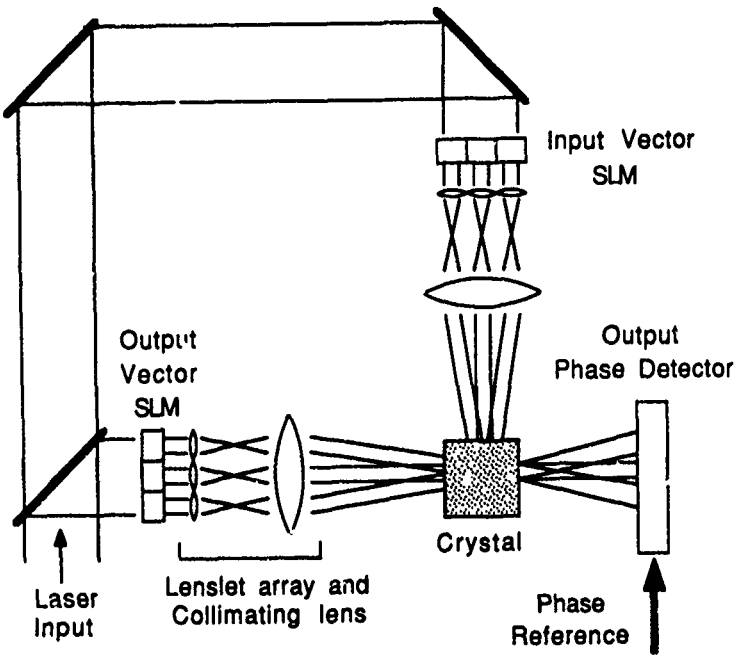


Figure 1: General architecture for a 2-layer photorefractive neural net.

where θ_o is a constant phase-shift between the intensity pattern and the resulting index modulation. Strictly speaking, the assumption of linear response requires that the multiplexed gratings add incoherently. This is not the case with simultaneous recording, but can be achieved with sequential recording [6]. Since parallel formation of the outer-product matrix is highly desirable, we are currently investigating the effects of simultaneous recording. For our discussion here, we will use the above equation as an approximate response, keeping in mind its limitations.

To store the information contained in a set of M vector pairs, a complete memory matrix $\mathbf{W^{net}}$ is generated by a linear combination of single-pair matrices $\mathbf{W^{(m)}}$. The time-averaging property of photorefractive crystals [7] makes them ideally suited for performing the required summation. Time-division multiplexing the exposures of the intermediate matrices allows the crystal to generate the average interconnection matrix. The exposure time τ_m of any one intermediate matrix $\mathbf{W}^{(m)}$ determines its associated weighting coefficient. Hence, repetitively illuminating the crystal with the complete set of vector pairs produces the following response,

$$\Delta n_{ij}^{net} = \frac{1}{T_o} \sum_{m=1}^{M} \tau_m \Delta n_{ij}^{(m)}. \tag{3}$$

The exposure period T_o for one complete M-step cycle must be no larger than the crystal time-constant which typically ranges between tens of milliseconds to tens of seconds for CW laser intensities. So Δn_{ij}^{net} represents W_{ij}^{net} and is simply a steady-state grating formed by the weighted average of M individual phase-encoded gratings.

Once the network has been programmed, memory recall is invoked through a multiplication of the storage matrix by an input vector. This operation takes place optically in parallel via the diffraction process. Presenting the crystal with an input vector generates the appropriate phase-encoded output vector. Each element of the output vector is a plane wave

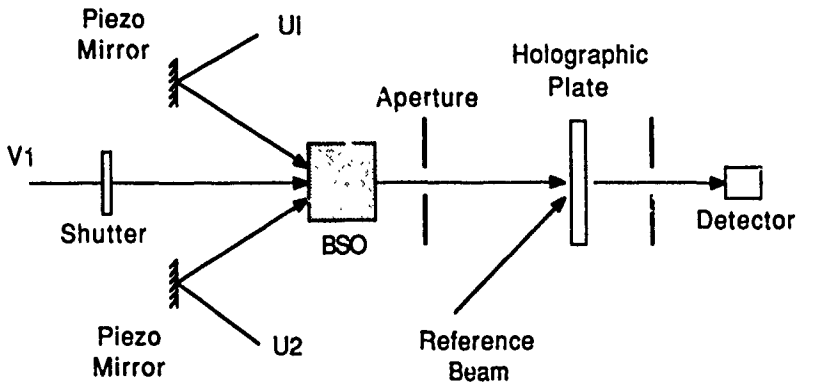


Figure 2: Experimental set-up for performing simple bipolar operations.

resulting from the superposition of many individual plane waves, the number of which is given by the dimension of the input vector. At this point, an optical thresholding operation that maintains the phase but normalizes the amplitude of each output beam will allow for feedback to the input or propagation into another network layer. Such a thresholding could possibly be implemented with 2 wave mixing techniques in ferroelectric photorefractives such as $BaTiO_3$ or SBN [8]. Under appropriate conditions, the output beams will be coherently amplified to approximately a constant amplitude with the additional benefit of providing the necessary gain to overcome diffraction losses.

We report the results of a simple experiment to test the feasibility of the phase-encoding method. The layout is shown in figure 2. Two input beams (U1 and U2) are reflected off piezo-mounted mirrors and interfere with a third output beam (V1). The intensity of all three beams was approximately 1 mW/cm^2. Two steady-state gratings form in a crystal of BSO that connect input to output. The output beam is then shuttered for 0.5 seconds, during which one or both piezo-mirrors phase shift their respective beams by π for two separate 100 ms durations. When only one of the two input beams is phase shifted, the two diffracted beams destructively interfere with each other as shown in figure 3a. The lower trace is the driving voltage to the mirror and the upper trace is the resulting diffraction intensity measured after the crystal (no phase detection). Note that the diffraction signal exhibits a slight overall decay due to grating erasure. We could consistently obtain a result similar to the one shown as long as high-frequency table vibrations were minimized. When both mirrors move simultaneously, the diffraction intensity remains relatively constant as shown in the lower trace of figure 3. However, the phase of the reconstructed beam alternates between 0 and π. To detect the phase, an additional grating recorded in a film plate is used to coherently combine an external reference beam of fixed phase with the diffracted output from the crystal. As seen in the upper trace of figure 4, the two different phase states are detectable as two intensity states. We should mention that the use of a fixed hologram to provide phase detection is not ideal since it cannot accommodate slow mechanical drift in the system. However, using a second crystal for phase-detection will overcome this difficulty. In conclusion, we feel these initial results indicate that the real-time volume holographic properties of photorefractive crystals can provide a feasible means for implementation of bipolar neural nets.

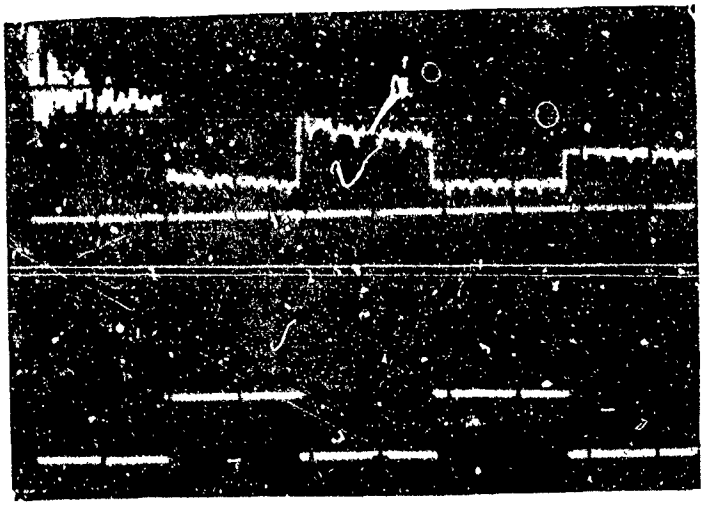


Figure 3: Results showing a) crystal output for one phase-shifting mirror (upper trace), and b) driving voltage (lower trace). Time sweep is 50 ms/div.

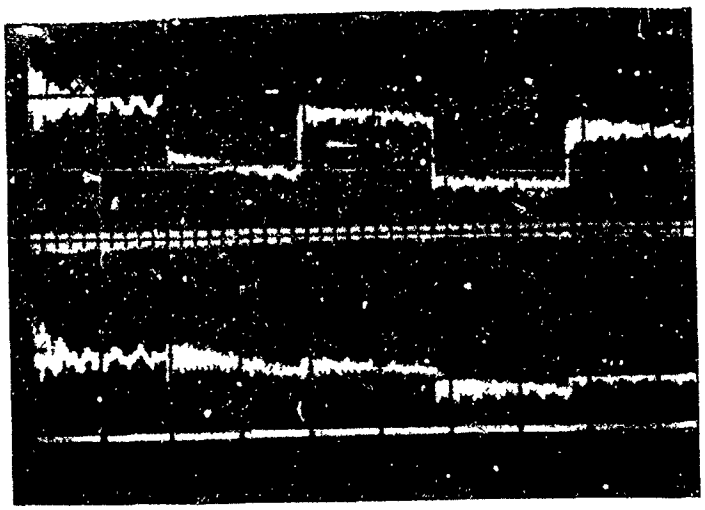


Figure 4: Results showing a) crystal output for both mirrors moving (lower trace), and b) same signal after holographic phase detection (upper trace).

References

1. Special issue on Neural Nets, Appl. Opt. **26**, (1 Dec. 1987).
2. N. Farhat, D. Psaltis, A. Prata, and E. Paek, Opt. Lett. **24**, 1469 (1985).
3. H. White and W. Wright, Appl. Opt. **27**, 331 (1988).
4. S. Song and S. Lee, Appl. Opt. **27**, 3149 (1988).
5. D. Psaltis, D. Brady, and K. Wagner, Appl. Opt. **27**, 1752 (1988).
6. L. Solymar and D. Cooke, *Volume Holography and Volume Gratings* (London, New York: Academic Press, 1981).
7. G. Hamel de Montchenault, B. Loiseaux, and J. Hiugnard, Appl. Phys. Lett. **50**, 1794 (1987).
8. P. Yeh, A. Chiou, and J. Hong, Appl. Opt. **27**, 2093 (1988).

Classification of Normal and Aberrant Chromosomes by an Optical Neural Network in Flow Cytometry

S. Noehte[1], R. Männer[1], M. Hausmann[2], H. Horner[3], C. Cremer[2]

[1)] Physical Institute, [2)] Institute for Applied Physics I, [3)] Institute for Theoretical Physics
University of Heidelberg, Heidelberg, West Germany
maen@dhdmpi5v.bitnet

Summary

This paper proposes to use an optical neural network for the real-time classification of normal and aberrant chromosomes in flow cytometry. It has been shown previously [1] that slit scan flow cytometry allows such an analysis for several problems of biomedical importance. Particularly, the rate of dicentric or translocation chromosomes can be used for a measurement of the radiation exposition of a patient in biological dosimetry.

To that, metaphase chromosomes stained by a DNA specific fluorochrome or by fluorescence hybridization are aligned in a flow cell by hydrodynamic focussing and pass a focussed laser beam [2]. The stain distribution along the chromosomes is recorded. These chromosome profiles allow a classification according to DNA content (centromeric index) or DNA sequences. Presently, chromosome profiles can be acquired at a rate of up to 100/s [3]. Using the reflection algorithm, profiles can currently be classified due to their centromeric index at a rate of 10/s [1]. To measure, e.g., radiation induced aberrations in the low dose range, it is possible to increase the data acquisition rate by a factor of 10. This requires to speed up the analysis by a factor of 100 which can be done by a multiprocessor system with 60 CPUs [4]. The classification of chromosomes by an optical neural network as outlined below would allow to increase the analysis speed by at least two further orders of magnitude which would make much lower dose ranges accessible for measurement.

The proposed classification method uses two profiles recorded for each chromosome. One corresponds to the total DNA content, the other one is specific to DNA sequences of a certain chromosome type. The classification is then done in two steps. The first profile allows to determine the chromosome type by means of the centromeric index. The second profile allows to identify translocations from the chosen chromosome into others, i.e. to find aberrant chromosomes. Both analysis steps can be done by an optical neural network operating as an associative memory in the following way.

Fig. 1 shows a typical example of the two chromosome profiles. Each one consists of 256 channels with a resolution of 6 bit. The total DNA signal (black) will be normalized in its intensity by the recording electronics. For each one of the 24 human chromosome types, the recorded length of a specific type may vary by a factor of up to 2. Because the hydrodynamic focussing does not align all chromosomes perfectly, the relative amplitudes of both peaks can also vary within the same range. To cover all possible variations of a single chromosome type, roughly 100*30 different patterns have to be stored. If the total amount of 24*3000 patterns with 256*6 bits each is stored in an associative memory, any recorded profile can be asserted and the most similar pattern will be retrieved. The storage medium is a hologram, providing the required storage capacity of 10^9 bits.

There are three basic shapes for the second profile that is specific for DNA sequences of a certain chromosome (gray in Fig. 1). Chromosomes that do not contain DNA sequences of the selected chromosome type simply create a profile with amplitude near zero. Chromosomes of the selected type generate a signal of the same shape as that of the first profile. Aberrant chromosomes, i.e. chromosomes which are not of the selected type, but contain some DNA sequence of it due to translocations, have an intensity peak at the corresponding position. Superimposed to all three profile types may be peaks of saturation intensity which are artefacts from the preparation of the chromosomes. There are roughly 3000 possible artefact patterns and 100 translocation patterns that have to be stored like for the first analysis step.

All patterns are highly correlated, so that the standard Hopfield learning algorithm can not be applied. Instead, pseudo invariant learning [5, 6] will be used. Training sets are generated from offline analyzed samples. The learning procedure itself will be done with a neural network simulated on a special purpose multiprocessor system [7]. This allows a high flexibility in selecting the optimal learning algorithm. The simulation result is a synthetic hologram, representing the synaptic weights of the neural network. Since learning is required only once, the time to compute this hologram is not a limiting factor. Important, however, is that the chromosome classification, which can be done by using the fixed set of synaptic weights stored in the hologram, is done fast.

A full optical neuronal network will be realized in two steps. The first step is a hybrid electro-optical setup similar to previously proposed systems [8, 9] (Fig. 2). The chromosome profile recorded by the cytometer is displayed on a transmission LCD as the input pattern under exploration. An expanded and mode cleaned beam of an argon-ion laser illuminates via the LCD and a cylindric lens the the hologram. The output is focused onto a photo transistor array which is read out by a local computer. The computer controls the amplification, the threshold or nonlinearity of the feedback function. The signal fed back represents now the new input pattern which is displayed on the LCD transmitter.

In a second step, it is planned to realize the feedback loop by an optical system as mentioned earlier [10] (Fig 3). In distinction to the first setup an argon-ion pumped dye laser is used for the realization of a laser beam amplification inside the loop. A Glan - Thompson prism in addition to a Kerr cell allows to direct the laser beam in a closed loop or to the output pattern detector. To compensate the loss of light a transversely pumped amplifier is placed in the loop. In this setup, one circulation in the optical loop requires <10 ns. With 512 neurons, the typical convergence time is 1000*10 ns, i.e. 10 µs, a factor of 100 faster than even an expensive multiprocessor solution.

References

[1] Lucas J., Gray J.W.: Centromeric Index Versus DNA Content Flow Karyotypes of Human Chromosomes Measured by means of Slit Scan Flow Cytometry; Cytometry, Vol. 8 (1987) 273-279

[2] Hausmann M., Dudin G., Aten J.A., Bühring H.J., Diaz E., Dölle J., Bier F., Cremer C.: Flow Cytometric Detection of Isolated Chromosomes Following Fluorescence Hybridization; *Biomed. Optics* (1988) (in press)

[3] Gray J.W., Langlois R.G.: Chromosome Classification and Purification Using Flow Cytometry and Sorting; *Ann. Rev. Biophys., Biophys. Chem.*, Vol. **15** (1986) 195-235

[4] Cremer C., Hausmann M., Zuse P., Aten J.A., Barths J., Bühring H.J.: Flow Cytometry of Chromosomes: Principles and Applications in Medicine and Molecular Biology; subm. to *Optik* (1988)

[5] Personaz L., Guyon I., Dreyfus G.: Information Storage and Retrieval in Spin-glass like Neural Networks; *J. Physique Lett.*, Vol. **46**, L359 (1985)

[6] Kohonen T.: Selforganization and Associative Memory; Springer-Verlag, Berlin (1988)

[7] Genthner A., Hauser R., Horner H., Lange R., Männer R.: NERV - A Simulation System for Neural Networks; will be publ. in: Proc. Int'l Symp. Connectionism in Perspective, Zürich (1988)

[8] Farhat N.H., Psaltis D., Prata A., Paek E.: Optical Implementation of the Hopfield Model; *Appl. Optics*, Vol. **24**, No. 10 (1985) 1469-1475

[9] Farhat N.H., Shae Z.Y.: Architectures and Methodologies for Self-Organization and Stochastic Learning in Opto-Electronic Analogs of Neural Nets; Proc. IEEE 1st Int'l Conf. on Neral Networks, San Diego, CA (1987) III565-III575

[10] Wagner K., Psaltis D.: Nonlinear Etalons in Competitive Optical Learning Networks; Proc. IEEE 1st Int'l Conf. on Neural Networks, San Diego, CA (1987) III585-III594

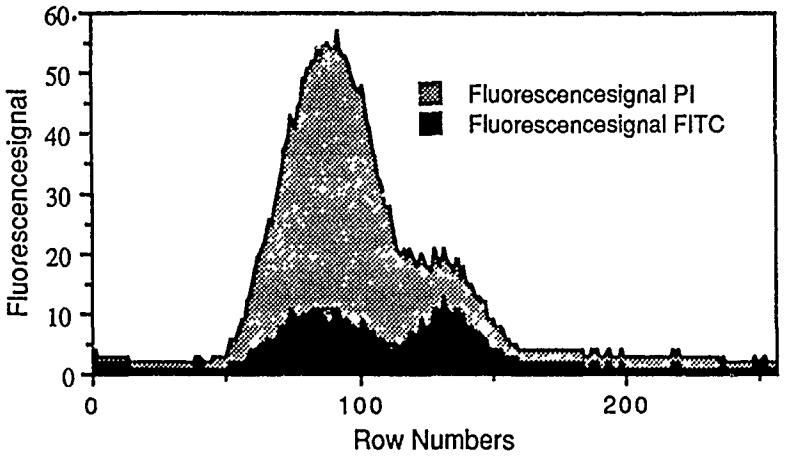


Fig. 1: Profile of an aberrant chromosome; gray - total DNA signal (PI colored chromosome), black - translocation labelled by fluorescence FITC after hybridization in suspension

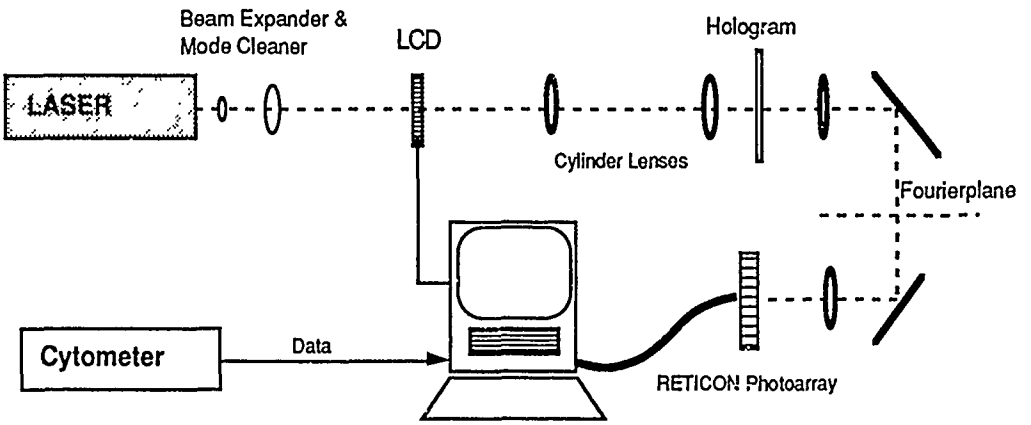


Fig. 2: Opto-electronical setup of the neural network

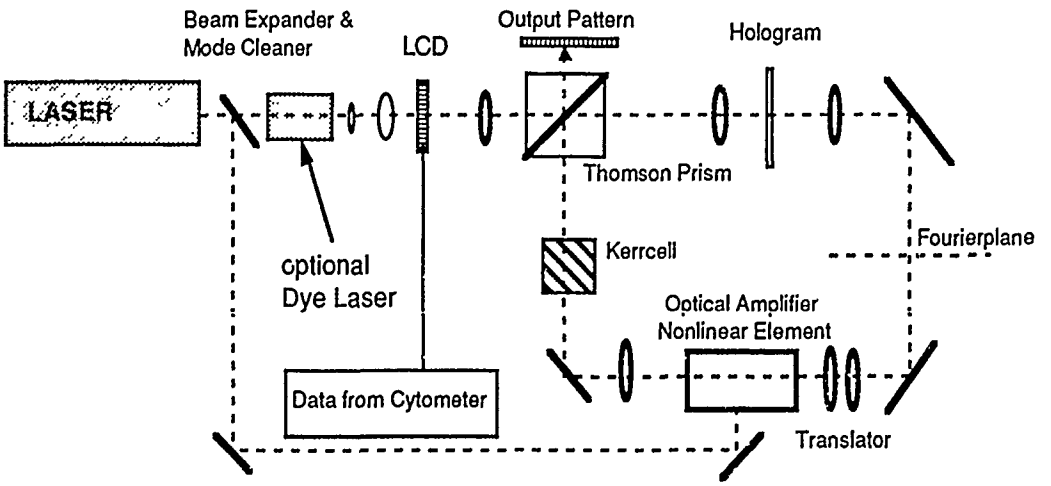


Fig. 3: Optical setup of the neural network

NOTES

MONDAY, FEBRUARY 27, 1989

SALON F

10:30 AM–11:30 AM

MC1–MC3

OPTICAL ARTIFICIAL INTELLIGENCE AND ADAPTIVE SYSTEMS

W. Thomas Cathey, University of Colorado, *Presider*

Optical Artificial Intelligence Based on Semantic Network Architecture

Toyohiko Yatagai

University of Tsukuba, Institute of Applied Physics
Tsukuba, Ibaraki 305, Japan

1. Introduction

In symbolic processing, associative network approaches show promise for solving difficult artificial intelligence problems.[1,2] Optical associative networks, including holographic[3,4] and matrix-vector multiplication [5] architectures, are one of the most attractive approaches toward large-scale associative processing. Optics provides both 2-D parallel interconnection ability between modules and parallel-computing mechanisms for parallel association algorithm. A hybrid optical inference architecture has been proposed.[6] Recently optical architectures for learning and self-organizing neural network are discussed.[7,8]

In this paper we are concerned with an architecture for optoelectronic implementation of a semantic network based on context-sensitive associations. The architecture proposed here is capable of computing the interconnectivity matrix for associations and of changing the weight matrix. Flexible data structures[2] used in artificial intelligence architectures are implemented in our system. This means that the concepts are represented by large patterns of activity and data-structures are stored by modifying the interconnections between these patterns.

2. Semantic Network Architecture

2.1 Relation formalism

There are many different ways of implementing semantic networks. As shown in Fig. 1, semantic nets are represented with nodes and directed arrows. This relational information is rewritten as

(NOBU FATHER TOYO)
(EMI FATHER TOYO)
(NOBU SISTER EMI)

One approach to implementation is to make each node in the semantic net correspond to a particular pattern of activity on a large assembly of units. Different semantic nodes may then be represented by different patterns of activity on the same set of

units. This semantic net formalism is seen as a crude description of the interactions between complex patterns of activity.

2.2 Semantic Net and State Vector

The information in any network consisting of nodes connected by labeled, directed arcs is equivalent to a set of triples, each of which consists of two nodes and an arc label. Cases in which the third component of a triple is not uniquely determined by the other two are particularly interesting.

The method that was used for implementing a semantic net in the associative architecture involved dividing the units into four groups or assemblies. The first three assemblies were called ROLE1, REL(short for relation), and ROLE2. The associative system could be queried about a particular triple by putting it into an initial state in which two of the first three assemblies had patterns of activity representing two components of the triple, and the remaining assemblies started off with all their units inactive. The associative system would complete the triple by setting down into a state in which the missing component of the triple was represented by the state of activity of the relevant assembly.

The fourth assembly was called PROP(short for proposition). For each particular triple stored by the associative memory system there was a corresponding particular state of the proposition assembly. Recall of triples from two of their components was achieved by making these states of the PROP assembly have the transition properties shown in Fig. 2.

Figure 3 shows the output of a computer program consisting of a simulator, which can simulate a parallel computer and a handler. The handler translates the first instruction into a set of operations that modify the weights of the associative memory system so as to store the three triples (NOBU FATHER TOYO), (EMI FATHER TOYO) and (NOBU SISTER EMI). The second instruction, (RECALL '(NOBU FATHER 0)), tells the handler to set up a particular initial binary state vector in the associative memory system and to print out a description of each subsequent binary state vector.

2.3 Context-Sensitive Association

The structure of the weight matrix for storing triples is shown in Fig. 4. The units in the first three assemblies have high thresholds but are also self-excitatory. Many matrices are null. For the three assemblies that code the constituents of a triple, the submatrices determining the effect of an assembly state on itself have all 0 weights except for the leading diagonal. This makes these assemblies retain whatever patterns they are given.

3. Optical Realization

Figure 5 shows an optical realization of the contextsensitive association system. The weight matrix is recorded in a micro-channel plate spatial light modulator (MSLM), and the recorded weight matrix and an input state vector displayed on a LED array are multiplied with an anamorphic optical system. The output vector detected with a photodetector array is transfered to a computer as shown in Fig. 6. The control computer reads the output vector and calculates the optimum weight matrix and decides a adequate level of thresholding to obtain a correct output of association. This procedure of modifying the weight matrix and determining the threshold level is realized by using a well known perceptron learning algorithm.[1]

4. Concluding Remarks

We present a contextsensitive associative memory system for artificial intelligence. A flexible data structure based on a semantic network is adopted. In the proposed optical implementation of the associative system a micro-channel plate spatial light modulator is used to store the weight matrix. Such a semantic model of contextsensitive association has the ability of storing and searching complex, flexible data-structures in highly parallel electro-optical hardware.

5. Acknowledgements

I would like to thank Yuzo Hirai for his helpful discussion, Yoshio Hayasaki and Nobuyuki Kasama for their experimental assistance. The constant support of Hiroyoshi Yajima, Satoshi Ishihara and Masahiko Mori of Electrotechnical Laboratory is highly appreciated. This work was supported in part by Grant in aid for scientific research by Ministry of Education, Culture and Science, Japan.

References

[1] T. Kohonen, Self-Organization and Associative Memory (Springer-Verlag, New York, 1984).
[2] G. E. Hinton and J. A. Anderson, Parallel Models of Associative Memory (Erlbaum, Hillsdale, NJ, 1981).
[3] D. Gabor,IBM J. Res. Dev., 13, 1256 (1969).
[4] E. G. Paek and D. Psaltis, Opt. Eng., 26, 428 (1987).
[5] D. Psaltis and N. Farhat, Opt. Lett., 10, 98 (1985)
[6] C. Warde and J. Kottas, Appl. Opt., 25, 940 (1986).
[7] A. D. Fisher, W. L. Lippincott and J. N. Lee,
[8] N. H. Farhat, Appl. Opt., 26, 5093 (1986).

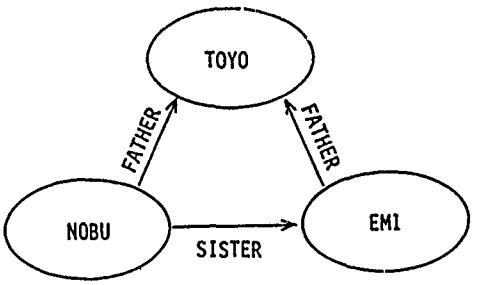


Fig. 1 Formalism for representing relational information.

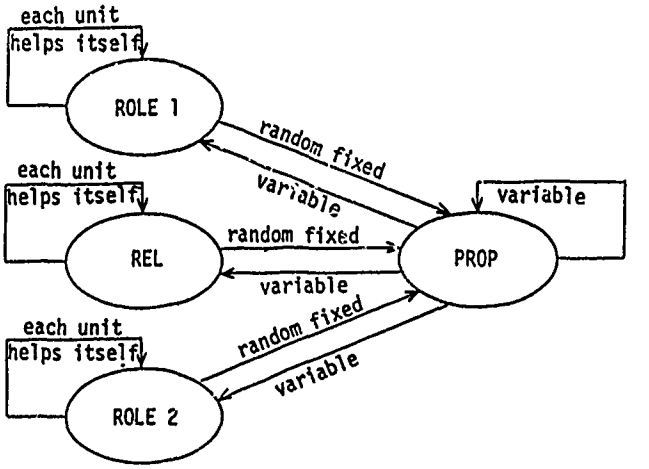


Fig. 2 Transition rule of assemblies for association.

```
*(STOREALL'((NOBU FATHER TOYO)
            (EMI FATHER TOYO)
            (NOBU SISTER EMI) ))

*(RECALL'(NOBU FATHER 0))
```

ROLE1	REL	ROLE2	PROP
NOBU	FATHER	0	0
NOBU	FATHER	0	NOBUFATHERTOYO4
NOBU	FATHER	TOYO	NOBUFATHERTOYO
NOBU	FATHER	TOYO	NOBUFATHERTOYO

Fig. 3 Output list of computer simulation of association.

INPUT VECTOR	OUTPUT VECTOR: ROLE1	REL	ROLE2	PROP
ROLE1	diagonal	0	0	random and fixed
REL	0	diagonal	0	random and fixed
ROLE2	0	0	diagonal	random and fixed
PROP	variable	variable	variable	variable

Fig. 4 Structure of weight matrix.

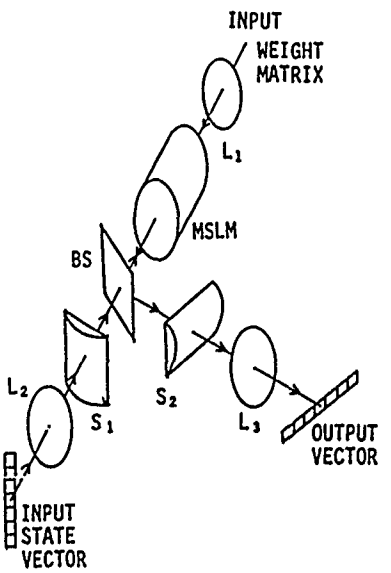


Fig. 5 Optical system for association based on multiplication.

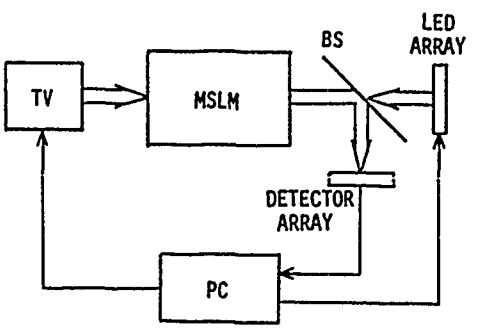


Fig. 6 Optoelectronic context sensitive association system.

Optical Matrix Encoding for Constraint Satisfaction

Gary C. Marsden, F. Kiamilev, S. Esener and Sing H. Lee

University of California, San Diego
Dept. of Electrical and Computer Engineering
La Jolla, California 92093

I. Introduction

Many Artificial Intelligence problems, such as theorem proving, computer vision, and expert systems, can be seen as constraint satisfaction problems[1,2,3]. In such problems, constraints are given on any possible solution. Most often, the problem is either explicitly stated in terms of allowed partial solutions or can be converted to such a representation. The objective is to find one or more solutions satisfying all constraints simultaneously, or the determination that no such solution exists. A simplistic approach to solving constraint satisfaction problems is to generate all possible solutions, then test each against the constraints to see if indeed they are satisfied. The process of backtracking provides a marginal improvement by testing increasingly larger partial solutions. Consistent Labelling is a more efficient procedure which eliminates allowed partial solutions that conflict with one another [2]. A problem is represented in a constraint network. Arc and path Consistent Labelling eliminate allowed partial solutions that are inconsistent over the smallest closed loops. The remaining allowed partial solutions can then be used in an efficient backtracking search. Using Consistent Labelling on larger loops, it is possible to obtain solutions to constraint satisfaction problems without backtracking[4].

Recently, a set of optically implemented algorithms were suggested for Arc and Path Consistent Labelling[3]. These algorithms operated on a matrix encoding traditionally associated with Consistent Labelling [1]. Arc consistency was obtained optically by focussing each matrix column to a point, while path consistency required matrix-matrix multiplication. Both of these require a large dynamic range for large problems. In this paper we describe a different method of matrix encoding constraint satisfaction problems that allows any degree of consistent labelling and therefore provides a complete solution. The algorithms do not differ greatly for the various degrees of consistency. Matrix encoding allows parallel optical execution. The dynamic range requirements on the optical system are minimal, since all matrix operations are executed locally at each element.

II. Matrix Encoding For Consistency Algorithms

Map coloring is an illustrative constraint satisfaction problem. Consider the network in Figure 1, which represents the problem of coloring the nodes with three colors (A, B, or C) so that no two connected nodes are the same color. As an additional constraint, node 2 may only be colored A or B. Also, suppose that relations between nodes 1 and 2 may not involve color C. At each node there are brackets showing the allowed labels for that node. Allowed relationships between nodes are represented within parentheses. Note that some relationships involving the labelling of node 2 with C are allowed even though it must be A or B. That is, the original problem did not directly forbid relationships between nodes 2 and 3 involving C as a label for node 2. Consistent Labelling identifies and removes these and other subtle inconsistencies among stated constraints.

Constraints on a network can be represented using vector and matrix encoding. The node vector in Figure 2A encodes the nodal constraints of Figure 1, while the matrix in Figure 2B encodes the relationship constraints. All node labels are explicitly allowed for each node except node 2, which may only be labelled A or B. There are more relationships represented in the matrix than in the network graph. Some of these represent the fact that a node cannot be multi-valued ; for example, (1A,1B) = 0 . Other allowed relations, such as (1A,4A), are set initially to one, since they are not listed explicitly in the constraints. Relationship constraints are assumed undirected, so that the encoding matrix is symmetric.

III. Consistent Labelling Algorithms

Before the relations matrix can be used, it must first be made compatible with the node vector. If a node label is disallowed, as is 2C in the present example, relations involving that labelling must be eliminated. A mask matrix is obtained from the outer product of the node vector with itself. This is multiplied (logical AND) element by element with the relations matrix. The resulting matrix is the revised relations matrix consistent with the node vector. This process is shown in Figure 3 for the encoding in Figure 2.

Arc consistency is obtained from the revised relations matrix by splitting the matrix into sections corresponding to each label value, as shown in Figure 4. These sections are added (logical OR) to produce a temporary matrix, which in turn is separated into rows. These rows are multiplied together to form the revised node vector. If the node vector changes by this process, the relations matrix is revised and the process repeated. Otherwise, the resulting relations matrix and node vector are arc consistent.

The path and higher order consistency algorithms require a core outer product routine, shown in Figure 5. In this routine, a matrix is formed from the outer product of a particular input vector with itself. Repeating this procedure for a set of input vectors, the resulting matrices are summed together. This sum is then multiplied with the

relations matrix to obtain a revised relations matrix. The particular vectors used in the core outer product routine depend on the algorithm. These vectors are obtained using columns from the relations matrix. In the path consistency algorithm, vectors presented to the routine are the columns corresponding to the labels for a single, fixed node. For higher order algorithms, several nodes are fixed. Columns corresponding to labellings of these nodes are multiplied together. The result is presented to the core outer product routine. Iterations are made on all labels for these nodes.

IV Optical Considerations

If the relation matrix is represented optically, operations on each element can be executed in parallel. The outer product required can also be executed in parallel. Since the logic operations are local to the elements, a dynamical range of two is required. The logical operations needed are OR, AND and XOR. The XOR operation is needed in comparing input to output matrices when terminating an algorithm. In addition, a memory buffer is needed for the summation matrices in the path/ higher order consistency algorithm.

V Conclusion

We have shown Consistent Labelling algorithms that use matrix encoding of constraints.,The nature of the matrix encoding allows optical implementation with efficiencies in parallelism and dynamic range. The algorithms are easily extended to arbitrary degrees of of consistency and therefore allow solutions to constraint satisfaction problems.

References:

1. Montanari, U., Networks of constraints: fundamental properties and applications to picture processing, *Inform. Sci.* 7 (1974), 95-132.

2. Mackworth, A., Consistency in networks of relations, *Artificial Intelligence* 8 (1977), 99-118.

3. Haney, M.W.,Athale,R.A., and Geesey,R.A., Optical techniques for increasing the efficiency of heuristic search, Special issue on optical computing, *Opt. Eng.* (March 1989).

4. Freuder, E.C., A sufficient condition for backtrack-free search. *J. ACM* 29 (1982), 24-32.

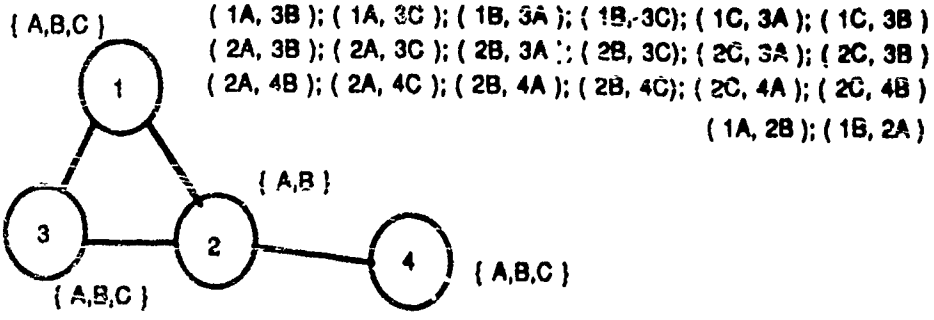


Figure 1: Map coloring is a typical constraint satisfaction problem. Allowed relationships between connected nodes are listed in parentheses. Brackets show allowed labels for each node.

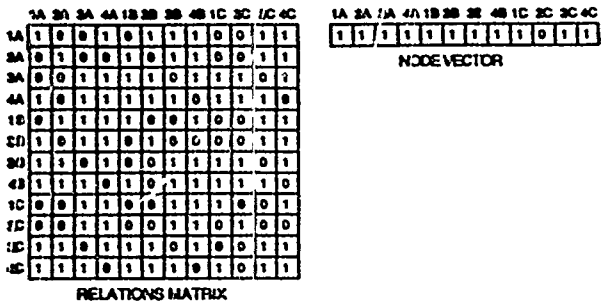


Figure 2: Explicitly allowed labels for each node are encoded with 1's in the node vector. Allowed relations between nodes with given labels are represented by 1's in the relations matrix.

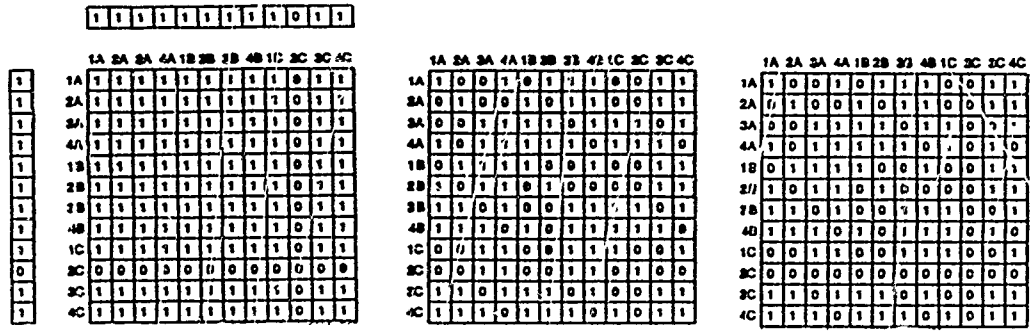


Figure 3: The outer product, (A), of the node vector is multiplied by the current relations matrix, (B), to yield a revised relations matrix, (C)

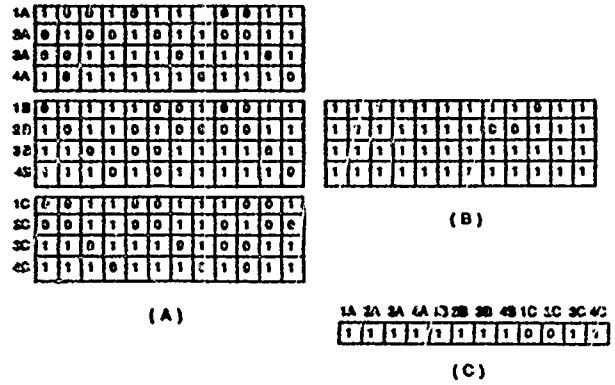


Figure 4: Arc consistency is obtained by separating the relations matrix into sections by label type (A), which are added to produce a temporary matrix, (B). The rows of this matrix are multiplied together to yield the revised node vector, (C).

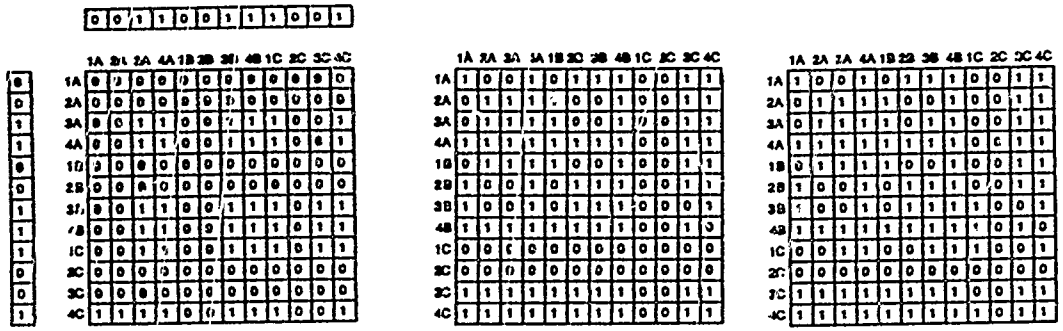


Figure 5: Path and higher order algorithms require a core outer product routine. An outer product matrix is produced from an input vector (a). This is added to the sum of previously generated matrices (b), to produce a new summation matrix, (c). The final summation is used as a mask to revise the relations matrix.

Adaptive Optical Filtering Architecture

Dr. Michael G. Price; Systeka Inc.
9525 Worrell Ave, Lanham/Seabrook, MD 20706

Dr. Alan E. Craig; Naval Research Laboratory
4555 Overlook Ave, Washington, D.C. 20375

Joseph C. Harsanyi; Systeka Inc.
9525 Worrell Ave, Lanham/Seabrook, MD 20706

Dr. John N. Lee; Naval Research Laboratory
4555 Overlook Ave, Washington, D.C. 20375

Abstract:

This paper discusses aspects of adaptive filtering applied to narrowband interference rejection for wideband receiver systems. A time/space integrating optical architecture using a spatial light modulator is described.

1.0 Introduction

Wideband signal acquisition or synchronization subsystems, even with processing gain, may be vulnerable to high-level jamming. Removal of strong CW interference can greatly improve the operation of these subsystems. The technique discussed here is adaptive cancellation, which consists of estimating magnitude and phase of the undesired signal, and using this information to cancel that signal.

An overall block diagram of an adaptive cancellation filter is shown in Figure 1. The estimator constructs an estimate $\hat{x}(t)$ of the evolving signal x(t). The filter output is the error or residual $e(t) = x(t) - \hat{x}(t)$. The filter takes advantage of the deterministic nature of narrowband signals; past values can be used to estimate present and future values. Noise and rapidly varying signals pass through the filter relatively unchanged.

Figure 2 shows the internal structure of the adaptive filter. This structure implements the algorithm defined by

$$a(iT) = \int_0^t [x(t) - \hat{x}(t)]\ x(t-iT)\ dt \qquad (1)$$

$$\hat{x}(t) = (1/N) \sum_{i=1}^{N} a(iT)\ x(t-iT) \qquad (2)$$

where a(iT) is the tap weight calculated at the ith tap, for N taps. This algorithm is known by many names, including LMS (for Least Mean Square) prediction; it is also called a Correlation Cancellation Loop (CCL).

Several variants of these equations exist. In particular, for delay time variable $\tau = iT$ and for large t, Equation (1) is seen as a correlation function $a(\tau)$ as tap spacing becomes small, Equation (2) becomes an integral convolution. This filter has been expressed as a time domain realization. An alternate form can be obtained by transforming the defining equations into the frequency domain [1].

2.0 Optical Architecture

Various optical methods have been proposed to implement this type of filter [1,2,3,4]. Optical approaches offer substantial bandwidth and time bandwidth product compared to digital approaches. Also, since this processor depends on feedback, some possible disadvantages of an optical approach, including low dynamic range and nonlinearities, may be reduced.

An alternate representation of the filter is shown in Figure 3. When the delay line is split into two parallel paths, the signal processing functions can be grouped into two distinct modules: a (time integrating) correlator and a transversal filter. The two modules communicate via the vector of tap weights $a(\tau)$ and the error $e(t)$.

An optical implementation corresponding to this representation is shown (in simplified form) in Figure 4. The tap weight calculation is performed with a time integrating correlator. That is, the tap weight information is obtained as a spatial optical distribution on a linear sensor. The convolver segment is implemented using an electrically addressed spatial light modulator (SLM) in a space integrating configuration. The output of the convolver is a time signal which represents the LMS estimate of the deterministic portion of the input.

This structure offers the potential for good CW suppression, since the time integrating module is expected to produce high quality estimates of the tap weights. In the convolver segment, the feedback loop is expected to

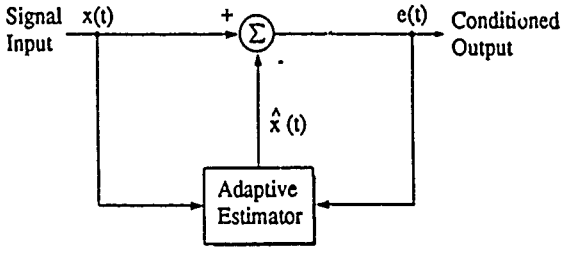


Figure 1. Adaptive Cancellation Overall Block Diagram

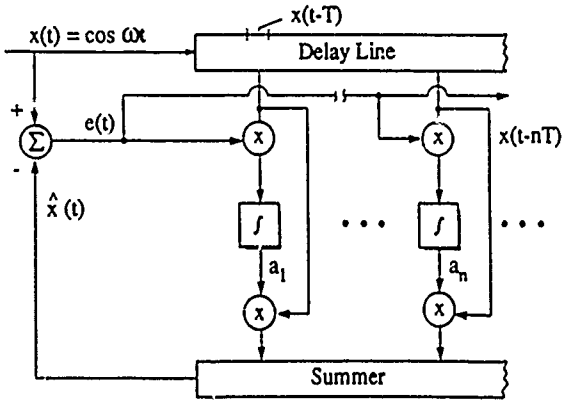


Figure 2. Adaptive Cancellation: Detail

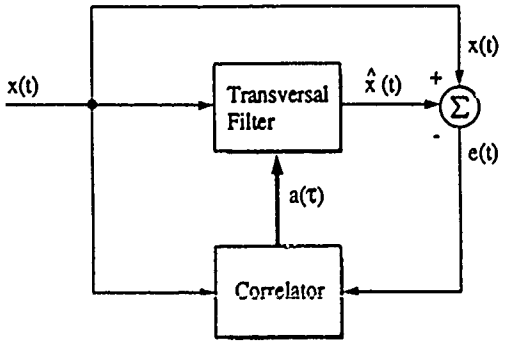


Figure 3. Adaptive Cancellation: Alternative Model

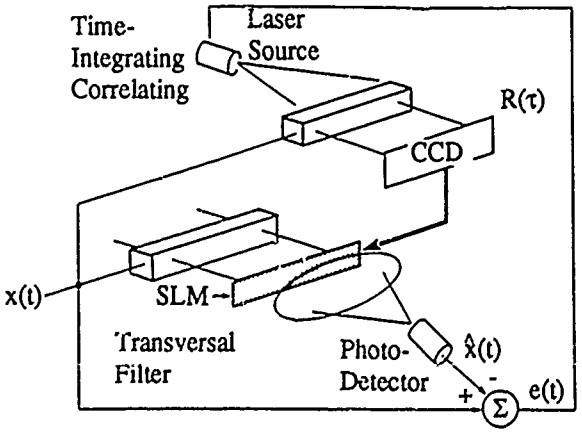


Figure 4. Adaptive Cancellation Architecture

reduce the SLM nonlinearities. In addition, initial simulation results suggest that the processor can be operated with unipolar (positive only) tap weights.

This processor is presently in integration and test at NRL. Figure 5 shows the optical layout. For both the correlator and convolver modules, the optical design is telecentric, unity magnification, with about 1 cm aperture. The 1 cm aperture corresponds to about 15 usec delay in slow shear TeO2 Bragg cells operating near center frequencies of 50 MHz. The laser sources are both HeNe.

The correlator module generates the adapted tap weights which are collected from the CCD array by a Data Translation 2861 frame grabber board installed in a Compaq 386/20 controller. The tap weights are then scaled, time averaged, and supplied to the convolver as a video signal. The convolver section uses a 128x128 pixel Texas Instruments deformable mirror device (DMD) as the SLM. The output of the convolver is the desired estimate.

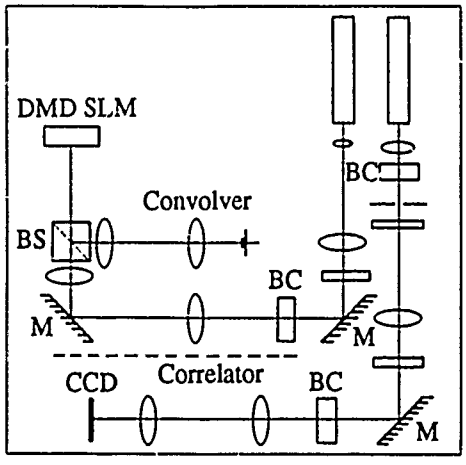


Figure 5. Optical Layout

3.0 Summary

An optical time integrating correlator and an optical space integrating convolver perform CW cancellation in an adaptive feedback architecture. Tests in process will provide an evaluation of system parameters and critical hardware components such as the SLM.

Acknowledgements:

The authors gratefully acknowledge the support of Ms. Paulette Rosner and the SAIC graphics publication staff in the preparation of this paper, and the many helpful technical discussions with Mr. John Colwell.

References:

(1) Rhodes, J. F. and Brown, D. E., "Adaptive Filtering with Correlation Cancellation Loops," SPIE, Vol. 341, Real Time Signal Processing V (1982), p. 140.

(2) VanderLugt, A., "Adaptive Optical Processor," Appl. Optics, 21, 1982, p. 4005.

(3) VanderLugt, A., "Optical Transversal Processor for Notch Filtering," Opt. Engr., 23, 1984, p. 312.

(4) Cohen, J. D., "Optical Adaptive Linear Predictors," Appl. Optics, 24, 1985, p. 4247.

MONDAY, FEBRUARY 27, 1989

SALON F

11:30 AM–12:30 PM

MD1–MD4

NEURAL SYSTEMS: 3

Demetri Psaltis, California Institute of Technology,
Presider

"Optical Associative Memory Utilizing Electrically and Optically Addressed Liquid Crystal Spatial Light Modulators"

I. Introduction

The goal of this research is to build optical connectionist machines (OCM's) capable of interconnecting many input units to output units via two-dimensional liquid crystal spatial light modulators (SLM's). Liquid crystals SLM's are chosen because of their low optical absorption and power dissipation, moderate switching speeds, potential for high extinction ratios and resolution. In addition, these materials are birefringent and can easily implement a polarization-based optical associative memory. Previously, we presented results on connecting five input units to five output units [1], and eight input units to eight output units [2] with the OCM. The number of interconnections was limited by the low extinction ratio of the two-dimensional SLM used as the OCM connection matrix (The Radio Shack Pocketvision Model 5 extinction is 5:1). In this paper we present results of using higher contrast computer controlled SEIKO liquid crystal pocket television model LVD.012 and optically addressable ferroelectric liquid crystal SLM's to hetero-associate pairs of 32 bit long input and output vectors. This association is performed using the least mean square algorithm (LMS), implemented with polarization encoding to represent both positive and negative weights.

II. Polarization-based OCM

Many popular connectionist architectures are based on a layered feedforward design, which can be easily implemented using the well known optical matrix-vector multiplier [3]. The simplest such system consists of two interconnected layers. One layer is used to represent the input activation units, and the other layer forms the output of the system. These networks are often used as associative memories [4]. The association between input and output can be learned by suitable modification of the connection strengths between the two layers. The Widrow-Hoff least mean square (LMS) algorithm is a rule for modifying the connection strengths, or weights to perform associative recall of a pattern from a partial or distorted input [5]. The OCM implements the LMS learning algorithm, as shown in Figure 1. Input patterns are encoded as intensities of vertically polarized light by the SEIKO LCTV1. Connection weights to the output layer are encoded by rotation of this light by a second SEIKO LCTV2. Vertically polarized light represents a +1, or an excitatory weight, horizontally polarized light represents a -1, or an inhibitory weight. Since the SEIKO LCTV's are twisted nematic displays, there are a range of analog weights between -1 and +1 that can be encoded , as shown in Figure 2. The input unit is imaged onto a column of the connection matrix, performing a rotation of polarization as the light is transmitted by the element. A polarizing beamsplitter separates the positive and negative components, as shown in Figure 3. These signed outputs are imaged onto linear detector arrays by cylindrical lenses, and subtracted in parallel electronics. The resulting output vector is compared to the desired vector which generates an error. This error prescribes a modification of the appropriate connection weight according to the LMS algorithm:

$$\Delta w_{ij} = \mu \varepsilon x_i , \tag{1}$$

where μ is a rate of convergence parameter, ε is the generated error, and x_i is the value of the ith input unit. After several passes through the system, a pattern association is learned. If part of the pattern, or a distorted version of the pattern is input to the system, the original pattern is retrieved, as shown in Figure 4. A photograph of the OCM is given in Figure 5.

III. Optical Outer Product Processing

The OCM described in section II requires electrically modifying the state of a weight in the connection matrix (LCTV2). This is performed by a serial computer, and therefore in the training of this matrix, the advantage of optics is lost. We can perform an outer product optical associative memory of one-dimensional input patterns by ANDing orthogonal projections of the input vector with optically addressed SLM's, as shown in Figure 6. We performed a proof-of-principle experiment with the input vector of Figure 7. Collimated light from an Argon Ion laser ($\lambda = 514.5$ nm) was transmitted by crossed one-dimensional SLM's and the resulting pattern illuminated a photoaddressed ferroelectric liquid crystal SLM,[6] forming the required outer product (Figure 8). Incorporating this component into the OCM will speed up both network training and convergence times. Further refinements of the OCM using the optical outer product processor will be discussed including multiple pattern storage and multilayer neural network architectures.

References

[1] M. Kranzdorf, K. M. Johnson, L. Cotter, L. Zhang, and B. J. Bigner, "A Polarization-based optical connectionist machine", Soc. Phot. Instr. Eng. vol. 963, (1988).

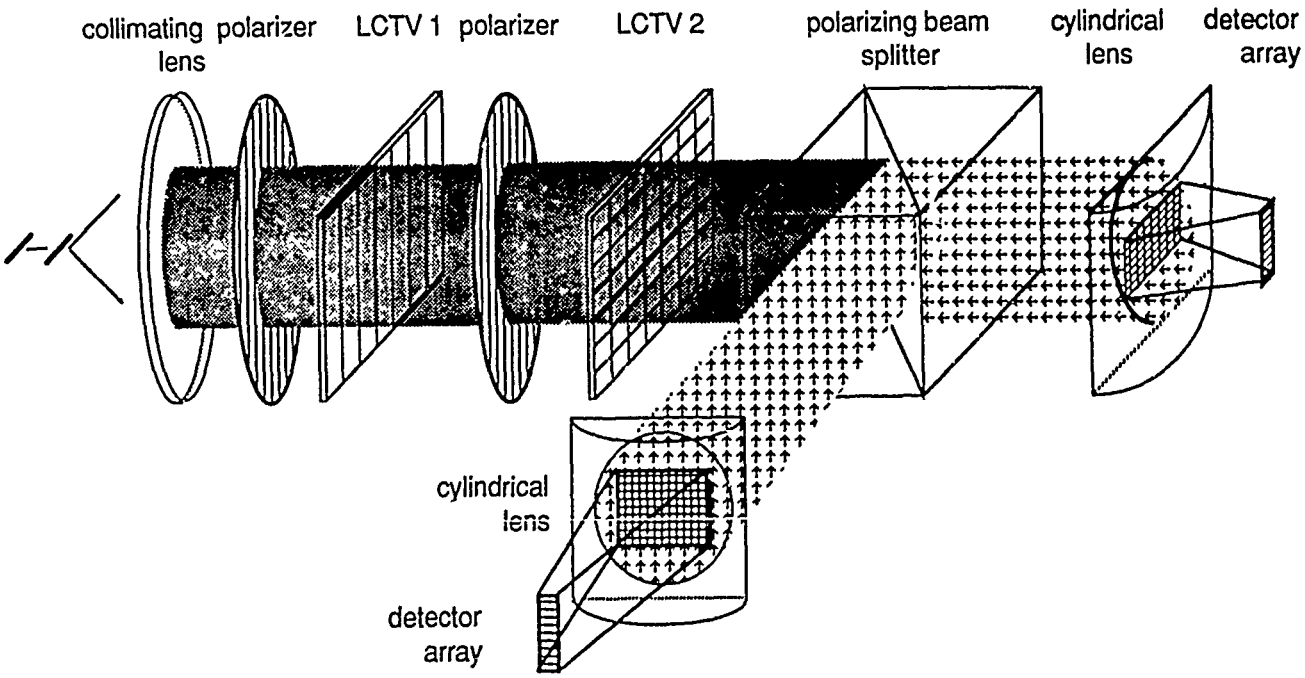
[2] K. M. Johnson, " A compact optical connectionist machine" OSA annual mtg. October, (1988).

[3] 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, (1978).

[4] See G. E. Hinton and J. A. Anderson, **Parallel Models of Associative Memory**, Lawrence Eribaum Associates, Hillsdale, NJ (1981).

[5] B. Widrow and M. E. Hoff, IRE WESCON. Rec. Part 4 96, (1960).

[6] G. Moddel, C.T. Kuo, K.M. Johnson and W. Li, "Optical Addressing of high-speed spatial light modulators with hydrogenated amorphous silicon" Mat. Res. Soc. Symp. Proc. 118 , 405 (1988).



The Optical Connectionist Machine

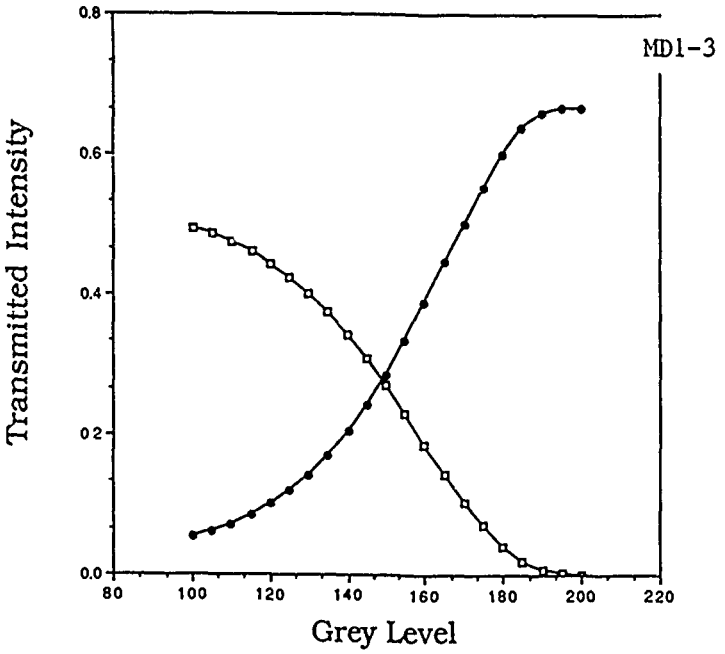


Fig. 2

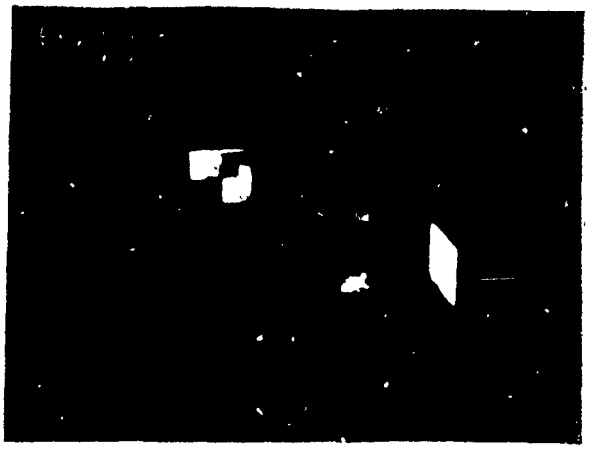
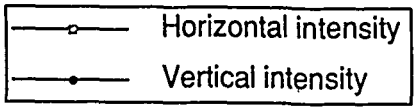


Fig. 3 Polarization Beamsplitter separates excitatory and inhibitory connection weights.



Seiko LCTV Polarization Modulation

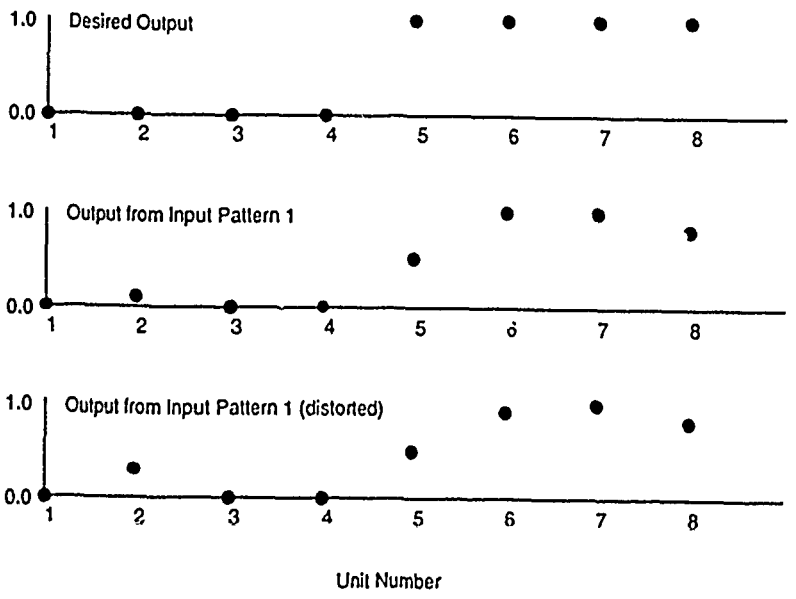


Figure 4 Results from two experiments are shown graphically. The top graph shows the desired result. For the second and third graphs, detector values are normalized and excitatory and inhibitory values are subtracted. Negative values are dropped. The middle graph is the output for a previously learned pattern displayed at the input. The bottom graph is the output for a distorted version of the same pattern displayed at the input.

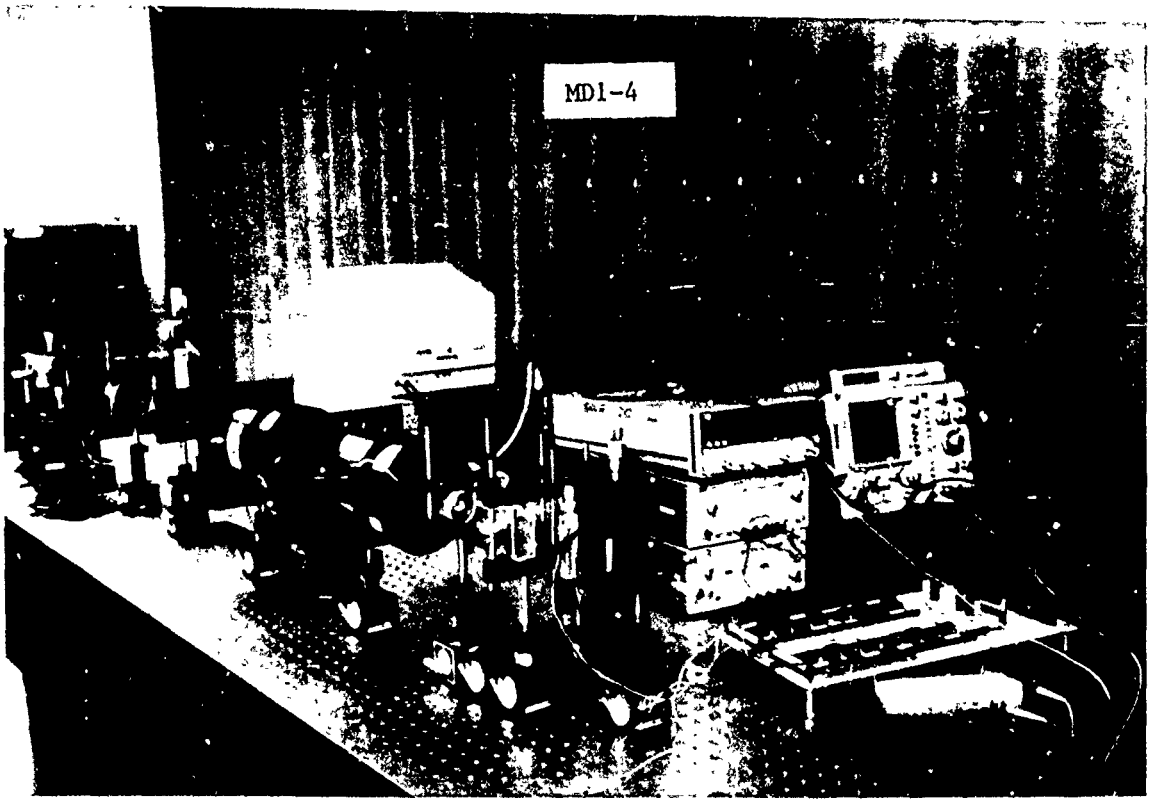


Fig. 5 Photograph of the Optical Connectionist Machine

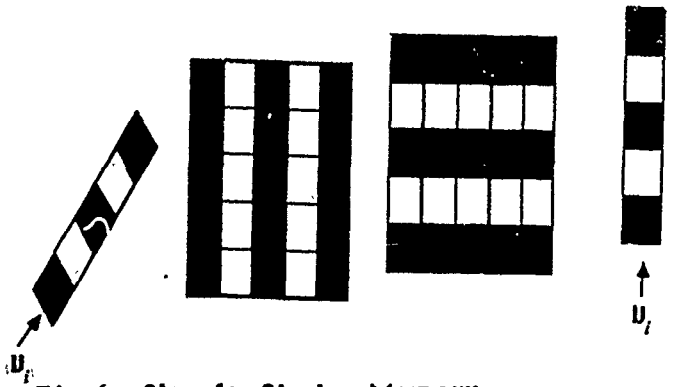


Fig.6 **Step 1: Storing Memory**

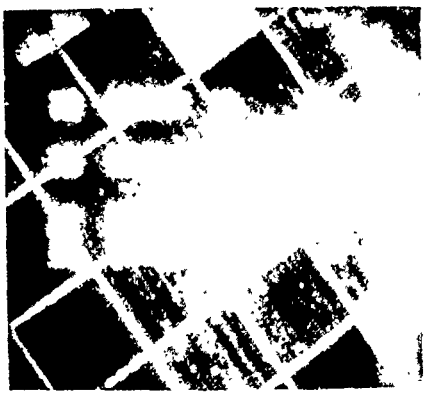


Fig. 8 Optical Outer Product

Fig. 7 Orthogonal Projections of an Input Vector

Competitively Inhibited Optical Neural Networks Using Two-step Holographic Materials

Michael Lemmon and B.V.K. Vijaya Kumar[1]
Dept. of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh PA 15213

abstract

Competitively inhibited networks automatically form nonparametric representations of density functions and therefore can be used as MAP predictors on a variety of problems. The dynamics of this class of networks is briefly reviewed and an optical implementation is proposed based on two-step holographic materials.

1. Competitively Inhibited Neural Networks

A competitively inhibited neural network is a specialized laterally inhibited network where the inhibition field has shrunk to a single neuron and the excitation field extends over all other neurons[Lemmon]. With n neurons in the network, let x_j be the short term memory (STM) state for the jth neuron and let z_{ij} be the ith long term memory state (LTM) for that neuron. The output of the jth neuron is denoted as $f(x_j)$ where the function is zero for $x_j < 0$ and monotonically increasing from zero to one for $x_j > 0$. The following network equations result.

$$\dot{x}_j = -Ax_j + B\sum_{i=1}^{m} z_{ij}y_i + C\sum_{i=1}^{n} D_{ij}f(x_i) \tag{1}$$

$$\dot{z}_{ij} = f(x_j)(y_i - z_{ij}) \tag{2}$$

where A,B,and C are positive network constants. Constants D_{ij} are chosen to induce competitive inhibition and $\bar{y}$ is the m-dimensional input stimulus vector.

We stimulate this network with a source which is generating stimuli once every T seconds. The stimulus vector, $\bar{y}$, lies in a set ,Ω, which we call the attribute space. Denote T as the presentation interval. Just prior to the presentation of a new stimulus we reinitialize the STM states to small equal negative values. The input stimuli in each presentation interval are drawn randomly according to the density function, $p(\bar{y})$. We are concerned with the convergence of STM and LTM states over a single presentation interval and the convergence over many presentation intervals. These two types of convergence are denoted as short run and long run convergence, respectively.

[1]We gratefully acknowledge the support of this research by the Internal Research and Development Funds of General Dynamics-Valley Systems

Short run convergent behaviour can be rigorously proven[Lemmon]. This behaviour consists of a given input ,$\bar{y}$, firing neurons whose LTM state vector,$\bar{z}_j$, are most closely correlated to the input. After firing, the LTM state vectors of these neurons begin converging to the applied stimulus vector.

Let $n(\bar{y})$ denote the fractional density of neural LTM states in the attribute space at vector $\bar{y}$. We can prove that long run convergent behaviour consists of the long run fractional neural density ,$n_s(\bar{y})$, satisfying the following relation, $p(\bar{y}) = g(n_s(\bar{y}))$, where $g()$ is a monotonically increasing function bounded between zero and one. This relationship means that neural LTM states will cluster about modes of the underlying source density, $p(\bar{y})$. This clustering behaviour is the hallmark of most self-organizing systems[Kohonen] [Amari].

The clustering capability of competitively inhibited networks is precisely the type of behaviour required by a MAP predictor. This means that such networks can be successfully applied to MAP pattern classification and MAP estimation. As the competitively inhibited network tends to form a nonparametric estimate of the source density function, we find that "neural MAP predictors" are applicable to a larger class of problems than can be currently handled by general purpose prediction methodologies. Such applications include ill-posed problems such as multitarget tracking and inverse scattering. Therefore, competitively inhibited neural networks represent an important neural network paradigm.

2. Optical Neural Predictor

The utility of a neural network paradigm rests on the ability to efficiently realize such networks. This section proposes an optical neural predictor based on two-step holographic materials. Figure 1 shows the basic architecture under consideration. The heart of this system is a holographic crystal which can only be rewritten in a two-step process. By selectively imaging a writing beam on specific portions of this crystal, we select neurons to rewrite their LTM states.

The input to the network is a vector imaged onto the face of a holographic crystal. A position on the crystal face will be denoted by the ordered pair (x, y). Therefore the equation of the object (input stimulus) on the crystal face is,

$$O(x,y) = \sum_{i=1,j=1}^{m,n} y_i \pi(i\Delta x, j\Delta y) \tag{3}$$

where y_i is the ith element of the input vector and Δy and Δx are a small spatial increment along the x and y coordinates, respectively. $\pi(i\Delta x, j\Delta y)$ is 1 for $(i-1)\Delta x < x < i\Delta x$ and $(j-1)\Delta y < y < j\Delta y$. It is zero elsewhere.

The holographic crystal has the LTM states for the neurons encoded on specific parts of its face. Therefore the transmittance of the crystal as a function of the x and y coordinates can be written as

$$\tau(x,y) = \sum_{i=1,j=1}^{n,m} z_{ij} \pi(i\Delta x, j\Delta y) \tag{4}$$

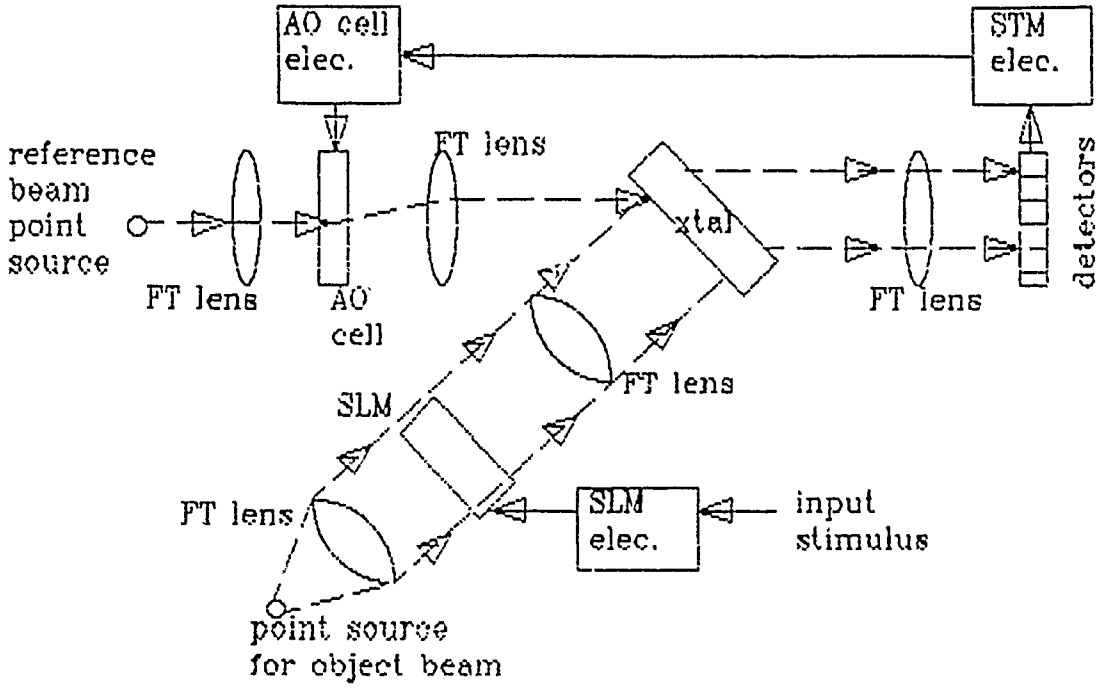


Figure 1: optical neural predictor

The output of the crystal is the element by element product of the LTM state vector and the input vectors. These outputs are summed along the x direction with a summing lens. The summed outputs image onto detectors. The output of these detectors is the inner product of the LTM state vector and input stimulus. There are n such outputs, one for each neuron.

Detector outputs ieed an analog VLSI circuit which emulates the STM dynamical equation 1. This block is easily realizable because of the reduced interconnection complexity associated with the competitively inhibited network. The outputs of the VLSI analog circuit then represent the outputs , $f(x_j)$, of the network.

The LTM equation 2 implies that rewriting LTM states (i.e., the holographic crystal) will be gated by the firing of that neuron. We therefore use a two-step holographic material so that rewriting only occurs when the object beam and an auxilliary writing beam are incident on that portion of the crystal's face encoding the LTM states for the fired neuron. To do this, we use the output of the VLSI neurons to drive an AO cell. The AO cell deflects the auxilliary writing beam so that its image on the crystal face is given by

$$f(x,y) = \sum_{i=1}^{m} f(x_j)\pi(j\Delta x, i\Delta y) \tag{5}$$

where $f(x_j)$ is the jth neuron's output given an STM state x_j.

The key component of the above proposed optical system is the two-step

optical material. There appear to be two materials that might be used for this crystal. One material is Lithium Niobate. However, because any illumination in the iron doped crystals will tend to redistribute charge, continual readout of the hologram will eventually erase the gratings. We can realize nondestructive readout if we can introduce long lived intermediate states between the conduction band and ground state through doping. In this case, absorption of photons of energy $h\nu_1$ excites electrons to the intermediate state. This illumination (i.e., the writing beam $f(x,y)$) "enables" rewriting on the medium. Then when the medium is illuminated with photons of enery $h\nu_2$, the electrons in the intermediate level are raised to the conduction band and the grating is recorded. This is precisely the two-step activation required by equation 2. Researchers using Chromium doped Lithium Niobate crystals were able to demonstrate this process[VonderLinde].

Another material capable of exhibiting nondestructive readout is ionic crystals with anisotropic color centers. In this case, the writing mechanism is through photobleaching or darkening of color centers in an ionic crystal. This results in an amplitude hologram with its associated poor diffraction efficiency. Researchers using NaF with M centers have been able to demonstrate two step nondestructive readout holograms [Casasent].

This paper has presented an important class of neural network architectures and has indicated how they might be realized using a hybrid optical architecture. The ultimate success of this optical implementation rests on the quality of two-step holographic recording materials.

References

[Lemmon] Lemmon and Vijaya Kumar, *First Annual INNS Proceedings*, Boston, Sept 1988

[Amari] Amari(ed.),*Competition and Cooperation in Neural Networks*, Springer Verlag, 1982

[Kohonen] Kohonen,*Associative Memory and Self-Organization*, Springer Verlag, 1982

[VonderLinde] Von der Linde,Glass, and Rogers, *Journal of Applied Physics*, Vol 47, No. 1,1976

[Casasent] Casasent and Caimi, *Applied Optics*, Vol 15,No. 3,1976

Adaptive two-dimensional quadratic associative memory using holographic lenslet arrays

Ju-Seog Jang, Sang-Yung Shin, and Soo-Young Lee

Department of Electrical Engineering
Korea Advanced Institute of Science and Technology
P.O. Box 150, Cheongryang, Seoul, Korea

1. Introduction

Optical implementation of two-dimensional (2-D) quadratic associative memory (QAM) that needs parallel N^6 weighted interconnections is described. We show that fully adaptive interconnections for the 2-D QAM are realizable by using two 2-D holographic lenslet arrays and two spatial light modulators (SLM's). Thus two extensions of our previous work[1] are proposed; they are learning capability and storage of 2-D images in QAM. We also show basic experimental results for the 2-D QAM.

2. QAM model of neural nets[2,3]

Extension of the 1-D QAM to the 2-D QAM is straightforward. Introducing an operator $\mathcal{T}$ that transforms 1-D vectors to 2-D matrices may be a solution[4]. A set of M binary matrices V_{ij}^s (s=1,2,...,M and i, j=1, 2,..., N) are stored in a sixth rank tensor,

$$W_{ijklmn} = \sum_{s=1}^{M} (2V_{ij}^s - 1)(2V_{kl}^s - 1)(2V_{mn}^s - 1) \tag{1}$$

where the unipolar binary [1, 0] is assumed. The tensor W_{ijklmn} is used for the retrieval of stored information from erroneous or partial input matrices. The tensor-matrix product

$$\hat{V}_{ij}^t \equiv \sum_{k,l}^{N} \sum_{m,n}^{N} W_{ijklmn} V_{kl}^t V_{mn}^t \tag{2}$$

with thresholding operation on $\hat{V}_{ij}^t$ yields an estimate of stored matrix that is most like the input V_{ij}^t. The thresholded estimate matrix is fed back to the input, and it converges to the correct stored image.

3. Optical implementation of the 2-D QAM using holographic lenslet arrays

In the optical implementation it is convenient to use unipolar W_{ijklmn}^* by adding a constant to every W_{ijklmn}. This is compensated by input-dependent thresholding operation.[1,4] Eq.(2) may be cast into two equations as follows:

$$T_{ijkl}^{t*} \equiv \sum_{m,n}^{N} W_{ijklmn}^* V_{mn}^t \tag{3a}$$

$$\hat{V}_{ij}^{t*} = \sum_{k,l}^{N} T_{ijkl}^{t*} V_{kl}^{t} \tag{3b}$$

where the terms marked by * means they have unipolar values. Then, V_{ij}^{t} is obtained by a proper input-dependent thresholding operation.[1]

Consider the implementation of Eqs.(3a) and (3b) with optics. We explain how each of the two tensor-matrix multiplication can be realized by using both a holographic lenslet array and an SLM. Each holographic lens plays the role of a lens for the first order diffracted beam when reference beam is illuminated. Thus it is possible to superpose the images of patterns positioned in front of the holographic lenses with the help of a lens as shown in Fig.1. Each holographic lens is made by exposing in a small area of the hologram a parallel reference beam with an object beam that is expanding from a focus. The array of such lenses are obtained by shifting the holographic plate and then exposing the two beams, repeatedly.

Since Eq.(3a) is a weighted sum, $\sum_{mn} W_{ijklmn}^{*} V_{mn}^{t}$, the holographic lenslet array is used to obtain T_{ijkl}^{t*}. First, we encode the tensor weight W_{ijklmn}^{*} into 2-D matrix pattern and position the pattern in front of the holographic lenslet array. Then, by illuminating an input beam through input $N \times N$ SLM that represent an input vector V_{mn}^{t}, Eq.(3a) is realized in the superposed image plane as shown in Fig.1. A coding rule of the sixth rank tensor into the 2-D SLM is shown in Fig.2 where each value of W_{ijklmn}^{*} is normalized and encoded by the degree of light transmission through the pixel of the SLM. Thus an $N^3 \times N^3$ SLM is required for W_{ijklmn}^{*}. Similarly, Eq.(3b) can be realized, after T_{ijkl}^{t*} is obtained. The total system is shown in Fig.3. Part A and B of Fig.3 is the implementation of Eqs.(3a) and (3b), respectively. Part B is similar to Part A except that the weight pattern T_{ijkl}^{t*} is the result of Part A.

Note that the weight pattern W_{ijklmn}^{*} is not recorded in the hologram array. It is only positioned in front of the holographic lenslet array. Thus an adaptive system can be realized by changing W_{ijklmn}^{*} value of the $N^3 \times N^3$ SLM.

4. Basic experiment

To show the feasibility of our system, Part A of Fig.3 is implemented since Part B is similar to Part A. Two binary images "L" and "T"

$$\mathrm{L} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \qquad \mathrm{T} \equiv \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{4}$$

are stored in 3×3 neurons. The $3^3 \times 3^3$ weight pattern W_{ijklmn}^{*} are calculated and the film mask for W_{ijklmn}^{*} is fabricated as explained in Fig.2. The element size of 3×3 holographic lenslet array we made for the demonstration of our ieda is 4mm×4mm, which may be made far smaller. Experimental results of Part A, T_{ijkl}^{t*}, are shown in Fig.4. The lumped 3×3 blocks in the photograph will be positioned in front of the 3×3 holographic lenslet array in Part B, respectively. Only the pixels of the input SLM that are switched on contribute corresponding block images to forming the output superposed image $\hat{V}_{ij}^{t}*$. Fig.4 shows that exact L and T can be obtained in the image plane of Part B if the T_{ijkl}^{t*}

is inserted in the $N^2 \times N^2$ 2-D SLM and the pixels of input SLM that stands for the input matrix are switched on.

5. Discussion

The storage capacity of this system M_q is proportional to N^4 (about $0.03N^4)^2$. The required maximum pixel number of 2-D SLM is $N^3 \times N^3$. If we want N^4 rate memory capacity with linear associative memory such as the Hopfield model memory, we require $N^2 \times N^2$ neurons and $N^4 \times N^4$ SLM for the weight pattern.

Pixel size of a few microns in the connection pattern can be imaged with our holographic lenses. Thus 10×10 neurons are easily implemented in our system with total holgraphic lenslet array size of 5cm×5cm. In this case M_q is about 300, which is remarkable storage capacity applicable to practical memory usages.

The implementation of the whole system and its experimental results will be discussed in more detail at the conference.

References

1. J.-S. Jang, S.-Y. Shin, and S.-Y. Lee, Opt. Lett. **13**, 693 (1988).
2. D. Psaltis and C. H. Park, AIP Conf. Proc. **151**, 370 (1986).
3. D. Psaltis and C. H. Park, and J. Hong, Neural networks **1**, 149 (1988).
4. J.-S. Jang, S.-W. Jung, S.-Y. Lee, and S.-Y. Shin, Opt. Lett. **13**, 248 (1988).

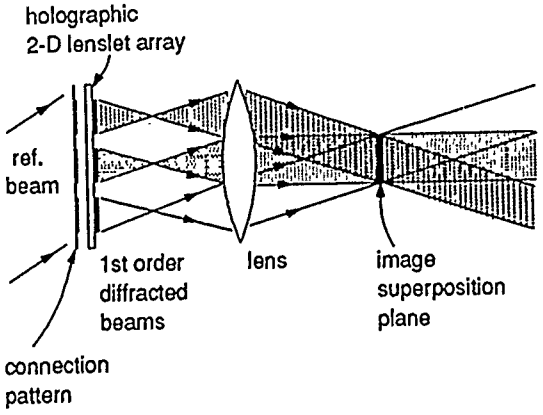


Fig.1. Superposition operation of holographic lenslet array. The image of pattern in front of each holographic array element is superposed in the image plane of the lens.

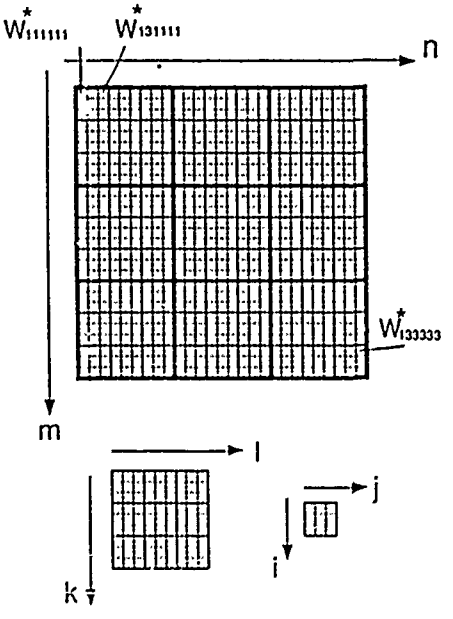


Fig.2. An example of the coding rule of the sixth rank tensor W^*_{ijklmn} into 2-D SLM when N=3.

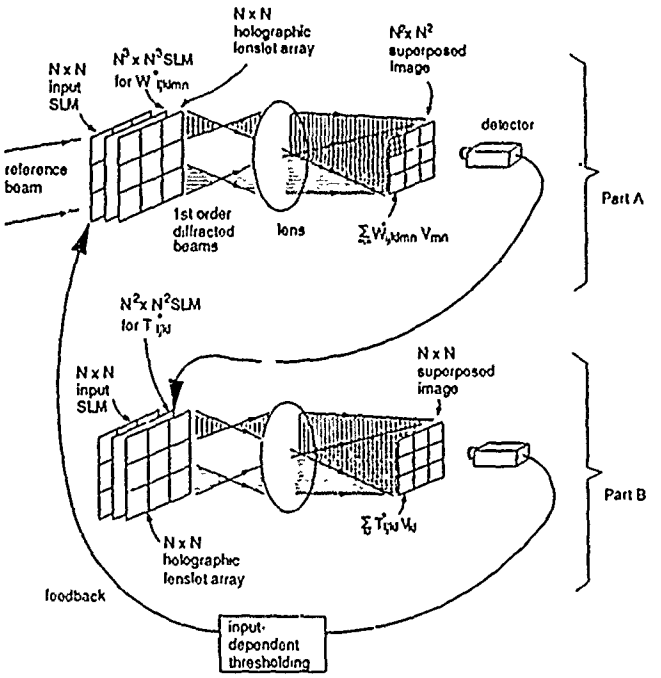


Fig.3. Total system setup.

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

(a)

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(b)

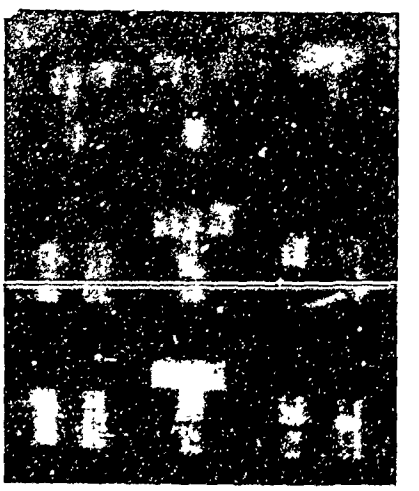


Fig.4. Experimental results: Photograph of output $T^{t*}_{ijk_l}$ in Part A of Fig.3. (a) The case when the input matrix is an erroneous L. (b) The case when input matrix is an erroneous T.

Self-Pumped Optical Neural Networks

Yuri Owechko
Hughes Research Laboratories
Malibu, California 90265

Neural network models for artificial intelligence offer an approach fundamentally different from conventional symbolic approaches, but the merits of the two paradigms cannot be fairly compared until neural network models with large numbers of "neurons" are implemented. Despite the attractiveness of neural networks for computing applications which involve adaptation and learning, most of the published demonstrations of neural network technology have involved relatively small numbers of "neurons". One reason for this is the poor match between conventional electronic serial or coarse-grained multiple-processor computers and the massive parallelism and communication requirements of neural network models. The self-pumped optical neural network (SPONN) described here is a fine-grained optical architecture which features massive parallelism and a much greater degree of interconnectivity than bus-oriented or hypercube electronic architectures. SPONN is potentially capable of implementing neural networks consisting of 10^5-10^6 neurons with 10^9-10^{10} interconnections. The mapping of neural network models onto the architecture occurs naturally without the need for multiplexing neurons or dealing with contention, routing, and communication bottleneck problems. This simplifies the programming involved compared to electronic implementations.

Previous optical holographic implementations of neural network models used a single grating in a photorefractive crystal to store a connection weight between two neurons (each pixel in the input/output planes corresponds to a single neuron). This approach relies on the Bragg condition for angularly selective diffraction from a grating to avoid cross-talk between neurons. However, because of the angular degeneracy of the Bragg condition, the neurons must be arranged in special patterns in the input/output planes to fully eliminate cross-talk. This results in sub-sampling of the spatial light modulators (SLMs) and incomplete utilization of the SLMs if the single grating per weight approach is used. Specifically, assuming the SLMs are capable of displaying NxN pixels, the single grating per weight method can store only $N^{3/2}$ neurons and N^3 interconnections.[1] I describe here an approach in which the Bragg degeneracy is broken by distributing each interconnection weight among a continuum of angularly and spatially distributed gratings. This eliminates cross-talk between neurons, making sub-sampling of the input/output planes unnecessary and allowing full utilization of the SLM space-bandwidth product. In other words, N^2 neurons can be fully interconnected provided the interconnection medium has sufficient degrees of freedom or space-bandwidth product to store the N^4 interconnection weights. By forcing signal beams to match the Bragg condition at many spatially distributed gratings, the signal-to-noise ratio should also be improved.

The continuum of gratings is generated by using a self-pumped phase conjugate mirror (SP-PCM) in conjunction with a SLM, CCD detector, frame grabber, and host computer. Several theories have been published for self-pumped phase conjugation in $BaTiO_3$ crystals, including internal resonators based on four-

wave mixing aided by Fresnel reflections and stimulated photorefractive backscattering. A common feature of these theories is that each pixel in the input plane writes gratings with and pumps all other pixels to form the phase conjugate wavefront. This results in a physical system which is massively interconnected and parallel, and which is a natural medium for implementation of neural network models. The distributed gratings in the crystal serve as the interconnection mechanism while the frame grabber in conjunction with the host computer implements programmable neuron activation functions. By spatially segregating the input/output plane, multiple layer neural network models can be implemented. This hybrid system combines the parallelism and interconnectivity of optics with the programmability of electronics.

A diagram of an experimental system used to demonstrate these concepts is shown in Fig. 1. The "object plane" corresponds to the plane of neurons represented by pixels on an LCLV (liquid crystal light valve). Activation patterns displayed on the LCLV are impressed on a light beam which is focused into the SP-PCM. Connections between the pixels are formed and the phase conjugate return is detected by a video camera. The return is processed on a point by point basis by the frame grabber/image processor before being displayed again on the LCLV. In neural network models such as backpropagation an error signal would be formed electronically and displayed on the LCLV to adjust the weights between neurons. The error signals are formed on a point-by-point basis (local operations) and so are not computational intensive.

An experimental demonstration of optical connectivity using the apparatus of Fig. 1 is shown in Fig. 2. Fig. 2a shows the phase conjugate return for an input consisting of a complete resolution pattern. The input was then switched to the region enclosed by the dashed ellipse in Fig. 2b. The return consisted of the complete resolution pattern, as shown in Fig. 2b, verifying that connection weights were formed globally among all the pixels. Cross-talk suppression is illustrated in Fig. 3. The input to the SP-PCM consisted of an array of dots on a rectangular grid (Fig. 3a). The conjugate return is shown in Fig. 3b. When the input was shifted even a slight amount, the return disappeared (Fig. 3c) which verified that pixels do not have to be arranged in special patterns on the SLM to avoid cross-talk. Finally, in Fig. 4 selective erasure of weights is demonstrated. The central neuron was deactivated in Fig. 4b by shifting the phase of that neuron on the LCLV. This shifts the phase of the gratings written by that neuron and selectively erases connections between it and the other neurons, demonstrating that learning using bipolar error signals is possible.

This work was supported in part by the Air Force Office of Scientific Research.

1. D. Psaltis, J. Yu, X. G. Gu, and H. Lee, "Optical Neural Nets Implemented with Volume Holograms," OSA Topical Meeting on Optical Computing, Incline Village, Nevada, 1987, Paper TuA3-1.

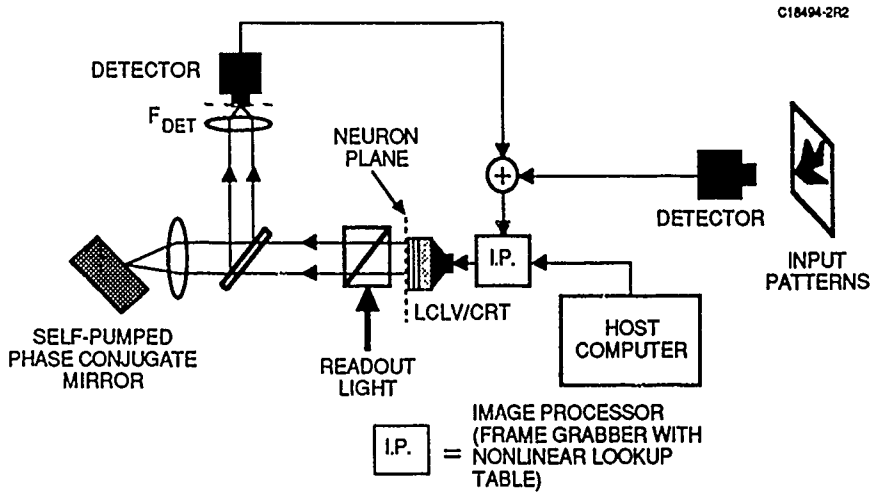
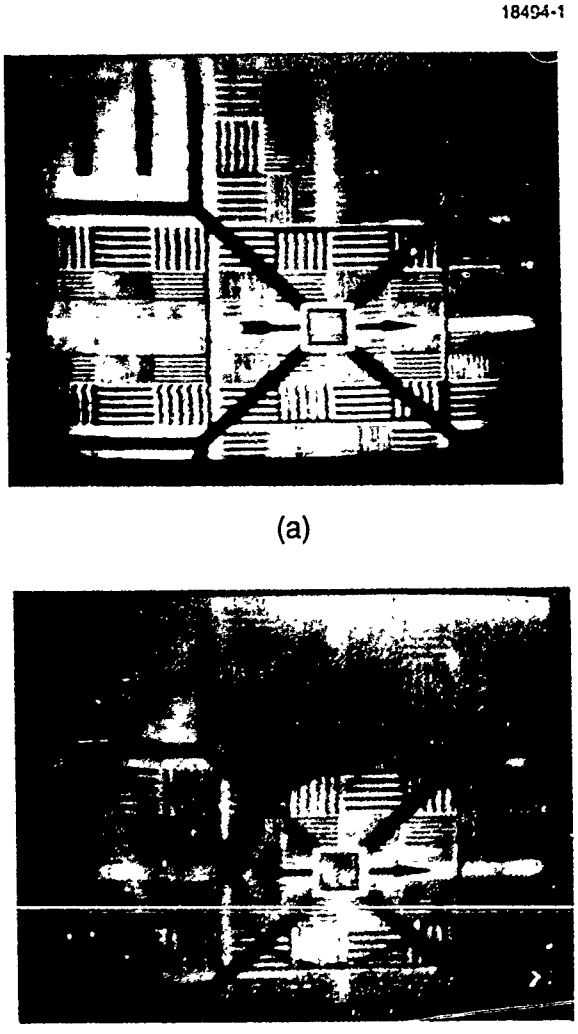


Figure 1. Schematic of self-pumped optical neural network apparatus

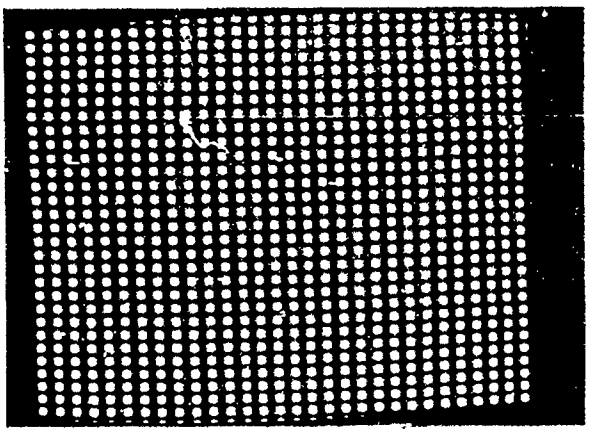


(a)

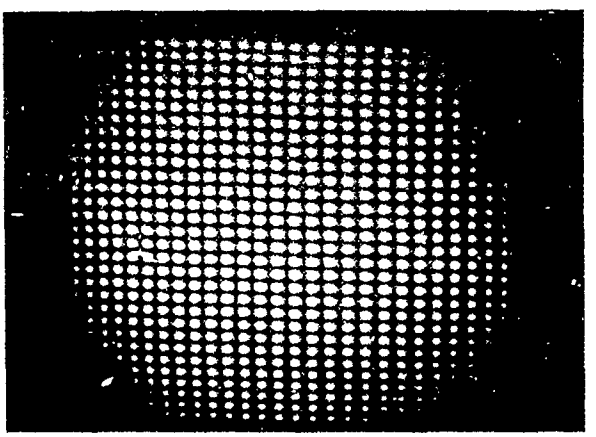
(b)

Figure 2. Demonstration of connectivity of self-pumped PCM

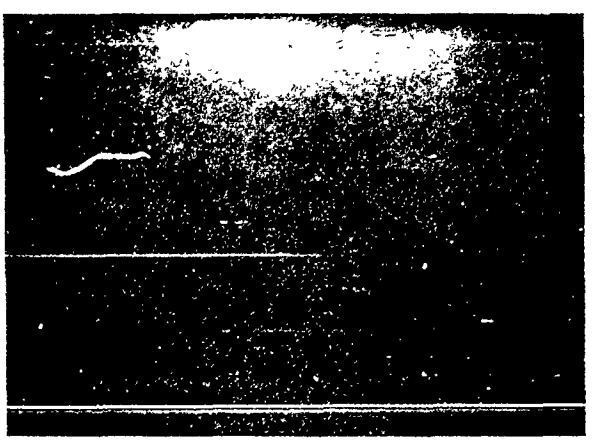
18414-1R1



(a)



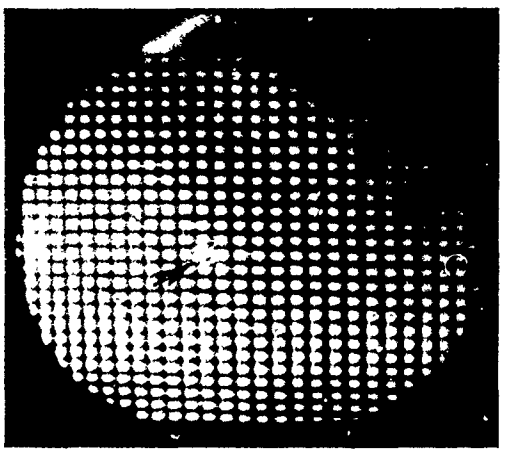
(b)



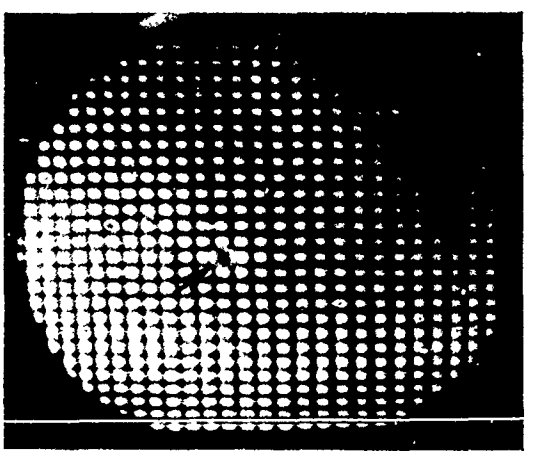
(c)

Figure 3. Demonstration of cross-talk suppression in self-pumped optical neural network

18441-1



(a)



(b)

Figure 4. Demonstration of selective weight erasure

NOTES

MONDAY, FEBRUARY 27, 1989

SALON F

2:00 PM–2:45 PM

ME1–ME3

SLMs AND OPTICAL DEVICES: 1

Arthur Fisher, U.S. Naval Research Laboratory,
Presider

64 ELEMENT HYBRID PLZT/SILICON SPATIAL LIGHT MODULATOR ARRAY

I Bennion, M J Goodwin, C J Groves-Kirkby and A D Parsons

Plessey Research Caswell Limited,
Allen Clark Research Centre,
Caswell, Towcester, Northants NN12 8EQ, England

Introduction

The high-throughput spatial light modulator (SLM) is a key component in an expanding range of optical signal processing applications and VLSI interconnect scenarios. With the continuing trend towards overall system integration, a number of approaches have been pursued in recent years, with the common aim of developing a silicon-compatible modulator system, preferably addressed via a fully integrated silicon back-plane with a 1:1 mapping between drive circuits and modulation elements. Potentially viable technologies investigated hitherto include liquid crystals [1,2,3], magneto optic materials [4] and micromechanical devices fabricated in silicon itself [5], all of which exhibit relatively limited modulation rates.

Electro-optic crystalline materials, with their intrinsically fast electronic response, offer considerable promise but, until recently, have lacked two major requirements for silicon compatibility, namely low drive voltage and mechanical interface capability. The advent of efficient quadratic electro-optic ceramics of the PLZT family, together with recent developments in hybridisation and thin-film deposition techniques, offers new possibilities in this field.

This paper describes a prototype high-speed two-dimensional SLM embodying recent advances in electro-optic materials and device processing technology. The device features electro-mechanical hybridisation of bulk PLZT electro-optic material to a monocrystalline silicon backplane, exploiting fully the inherent parallelism of the planar interface.

SLM Design

The hybrid PLZT/Silicon array is shown schematically in Figure 1. The structure comprises a two-dimensional monolithic electro-optic modulator array interfaced, via the flip-chip solder-bond technique [6], to a corresponding array of drive circuits. The device is operated in reflection at normal incidence, with optical input and output via a corresponding planar microlens array integrated and similarly aligned with the active elements using the solder-bond technique.

Despite its similarity to other recently reported structures based on polysilicon/PLZT [7,8,9], significant differences, both in physical construction and projected performance, exist. The principal novel feature of the device described here, and most significant in terms of fabrication and operation, is the physical separation of the electronic and electro-optic functions, enabling each to be developed and optimised separately and independently of materials processing constraints imposed by the other. In addition to utilising well established technologies, the hybrid approach offers the

significant major advantage of technological compatibility with a range of electro-optic material systems, permitting immediate incorporation of emerging materials as they become available without the need for redevelopment of drive arrangements. The monolithic approach, despite its apparent attractiveness, represents a technology compromise, with substantial effort necessary to establish fabrication techniques for each new electro-optic material.

Materials in the PLZT (lead lanthanum zirconate titanate) solid solution system are attractive for application as the hybridised modulator medium on account of their strong quadratic electro-optic coefficients, their availability in the form of large, high quality wafers and demonstrated fast switching response [10,11].

The modulator array is mechanically and electrically interfaced to the associated silicon circuit using the solder-bond flip-chip technique [6], c.f. Figure 2. This offers true metallic contact, photolithographically defined self-alignment with demonstrated sub-micron positioning accuracy, and clean, well-defined bonds. The present generation of devices uses a 16 x 8 solder-bond array (2 bonds per device) to provide electrical interface to the 64 element array, with further solder-bonds external to the active region for mechanical alignment and stability. This technique is compatible with arrays of up to 100 x 100 elements.

Projected Hybrid SLM Performance

Certain performance aspects of a 100 x 100 modulator array may be estimated on the assumption of an elemental switching response time of 50 ns [10]; significantly faster switching has been demonstrated in a waveguide format [11]. Calculations indicate a potential frame rate of 10^6 to 10^7 per second, with a corresponding throughput of the order of 10^{11} pixels per second. Thermal dissipation at these rates may be a critical, possibly limiting, factor in determining ultimate device performance. Using a modelled device capacitance of 0.6 pF, together with a 20 V drive requirement, 2.4 mW switching power per element is indicated. For 10^4 elements at 0.4 mm pitch, (625 cm^{-2}), this yields a total power dissipation of 1.5 $W.cm^{-2}$, within the capability of current heat-sinking technology.

Modulator Operation

Figure 3 shows a photograph of an assembled 8 x 8 element prototype PLZT/Silicon modulator array, packaged for individual element characterisation prior to incorporation of the input/output microlens array. Operation of devices of this kind has been demonstrated, drive voltages of the order of 20 V providing good contrast in 6 μm devices at 633 nm wavelength; full characterisation is currently in progress and detailed results will be reported at the meeting.

Acknowledgements

The authors wish to thank P. Burdett for his assistance in the fabrication of the hybrid PLZT/Silicon devices. This work has been supported by the Department of Trade and Industry (UK) under a Joint Optoelectronics Research Scheme (JOERS) programme.

References

[1] The silicon liquid crystal light valve: U Efron, J Grinberg, P O Braatz, M J Little, P G Re and R N Schwartz, J App Phys 57, 1356-1368 (1985)

[2] Low-cost LCD spatial light modulator: A M Tai, Appl Opt 25, 1380-1282 (1986)

[3] Status of the Hughes charge-coupled-device-addressed liquid crystal light valve: M S Welkowsky, U Efron, W Byles and N W Goodwin, Opt Eng 26, 414-417 (1987)

[4] Two-dimensional magneto-optic spatial light modulator for signal processing: W E Ross, D Psaltis and R H Anderson, Opt Eng, 22, 485-490 (1983)

[5] Micromechanical light modulators for data transfer and processing: R E Brooks, Proc SPIE 465, 46-54 (1984)

[6 Flip chip solder bonding for microelectronic applications: D J Pedder, Hybrid Circuits (ISHM UK), No.15, 4-7 (1988)

[7] Two dimensional silicon/PLZT spatial light modulators: design considerations and technology: S H Lee, S C Esener, M A Title and T J Drabik, Opt Eng 25, 250-260 (1986)

[8] One dimensional silicon/PLZT spatial light modulators: S C Esener, J H Wang, T J Drabik, M A Title and S H Lee, Opt Eng 26, 406-413 (1987)

[9] Two-dimensionally electrically addressed Si/PLZT spatial light modulator arrays: J H Wang, S C Esener, T H Lin, S Dasgupta and S H Lee ,Topical meeting on "Spatial light modulators and Applications", Lake Tahoe, Nevada June 1988, Paper ThC2

[10] Electro-optic switching response in cubic phase PLZT ceramic materials: C J Kirkby, Appl Opt 15, 828-830 (1976)

[11] Optical TIR switches using PLZT thin-film waveguides on sapphire: K Wasa, O Yamazaki, H Adachi, T Kawaguchi and K Setsune, J Lightwave Tech LT-2, 710-713 (1984)

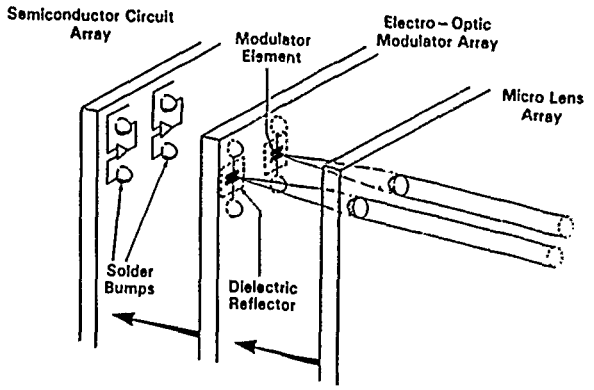


Figure 1:

Hybrid PLZT/Silicon Array Schematic

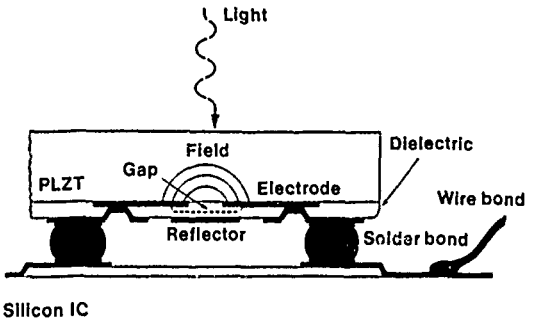


Figure 2:

Structure of Flip-Chip Solder Bond

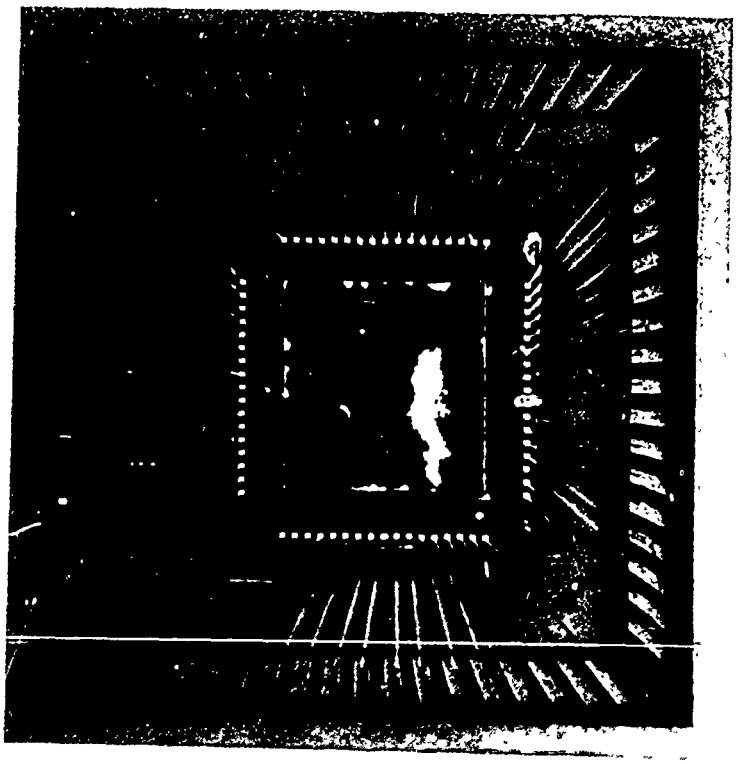


Figure 3: Assembled PLZT/Silicon Modulator Array

Dual Beam Recrystallization of Silicon on PLZT

Ali Ersen, Samhita Dasgupta, T. H. Lin, Sadik Esener and Sing H. Lee
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA 92093

INTRODUCTION

In the last few years extensive efforts have gone into the development of spatial light modulators (SLMs) for the realization of massively parallel optical processors [1-3]. The silicon-PLZT SLM approach is a promising approach and has the potential of combining the computational power of silicon with the communication power of optical interconnects [4]. Optical interconnects utilize the third dimension normal to the processing plane to provide the advantages of high speed parallel and global interconnections among simple silicon electronic circuits performing local computational operations. Fig. 1 shows the schematic of Si/PLZT SLM.

To fabricate a Si/PLZT SLM an Ar^+ laser has been used till now for the recrystallization of polysilicon on SiO_2 on PLZT [5,6]. In this paper we would like to report on the results of combining a CO_2 laser with an Ar^+ laser for the purpose of achieving better crystallization [7] of polysilicon by reducing the stress at the Si/SiO_2 interface. We shall also report on the improved performance of devices fabricated on the better quality recrystallized silicon.

MOTIVATION FOR DUAL BEAM RECRYSTALLIZATION OF Si/PLZT

Films recrystallized using only an Ar^+ beam exhibit residual stress that is induced during the laser crystallization process due to the coupled effect of the mismatch of thermal expansion coefficients between the polysilicon and the underlying silicon dioxide, and large thermal gradients that are present across the poly-Si/SiO_2 interface. The thermal gradient across the poly-Si/SiO_2 interface is present because the Ar^+ laser is strongly absorbed in the polysilicon layer ($\sim 2 \times 10^4$ cm^{-1}) whereas the SiO_2 layer is essentially transparent to the Ar^+ laser irradiation. Dual beam recrystallization reduces this temperature gradient since at 10.6 μm polysilicon absorbs weakly at room temperature (~ 1 cm^{-1}) whereas the SiO_2 layer beneath absorbs strongly($\sim 2 \times 10^3$ cm^{-1})[8]. The use of the two lasers in unison provides an independent source of local heating for both the layers and thus reducing the induced stress in the recrystallized film.

EXPERIMENTAL TECHNIQUES

A cross-section of the sample is illustrated in Fig. 2. The Si/PLZT samples are prepared by depositing a 3.5 μm layer of SiO_2 on the front surface of the PLZT substrate by PECVD at 250 °C. This layer is used for both thermal and electrical isolation. A 0.6 μm thick layer of polysilicon layer is then deposited on both sides of the composite by LPCVD at 640 °C. The front side polysilicon layer is later crystallized and used to host the silicon transistors while the backside layer is used as a masking layer protecting the PLZT substrate throughout the process. To define the location of grain growth and to obtain larger grain sizes anti reflection (AR) stripes have been used. During laser scanning the regions under the AR stripes get heated more than neighboring regions which results in a temperature profile. This temperature profile yields single crystalline regrowth in the region between the AR stripes and confines the grain boundaries to the

regions under the AR stripes. Doped regions for the fabrication of MOS transistors are defined by ion-implantation. The sample are then led through the usual device fabrication steps developed for PLZT substrates. [4]

The experimental setup for laser recrystallization is illustrated in Fig. 3. An 22W Ar^+ laser and a 50W CO_2 laser are focussed concentrically on the sample through an XY galvanometric scanner. The sample was positioned using a motorized XY translation stage. The stage movement, the scanner, the beam shutters were all computer controlled.

MATERIAL CHARACTERIZATION AND DEVICE TESTING

To evaluate the crystalline quality scanning electron microscopy (SEM) and Raman microprobe spectroscopy have been performed on the samples. To further characterize the new process a test chip has been designed. The chip is composed of several test structures such as CMOS inverters, NAND and NOR gates, MOS capacitors, ring oscillators and p-n junction diodes. The purpose of these test structures is to evaluate the electrical parameters as well as the uniformity of the fabricated devices across the wafer. We are in the process of measuring the devices and the results will be presented at the conference.

REFERENCES

1. A. R. Tanguay, "Material Requirements for Optical Processing and Computing Devices", Opt. Eng. 24 (1), January 1985, p.2-p.18.

2 C. Warde, A. D. Fisher, "Spatial Light Modulators: Applications and Functional Capabilities", in Optical Signal Processing, J. Horner, Ed., Academic Press, Inc., San Diego, (1987), p.477-518.

3 P. W. Smith, "Applications of All-Optical Switching and Logic", Phil. Trans. R. Soc. Lond., A313, 349-355, (1984).

4 S. H. Lee, S. Esener, M. Title and T. Drabik, "Two-dimensional silicon/PLZT spatial light modulators: design considerations and technology," Opt. Eng. 25 (2):250-260. February 1986.

5. J. H. Wang, S. C. Esener, T. H. Lin, S. Dasgupta and S. H. Lee, "Two-Dimensional Electrically-addressed Si/PLZT Spatial Light Modulator Arrays", Presented at the OSA meeting on Spatial Light Modulators at South Lake Tahoe, June 1988, (see 1988 Technical Digest Series, Volume 8:124-127).

6. T. H. Lin, J. Wang, S. Dasgupta, S. C. Esener, and S. H. Lee, "A 2-D Optically Addressed Silicon/PLZT Spatial Light Modulator Array", Presented at the OSA meeting on Spatial Light Modulators at South Lake Tahoe, June 1988, (see 1988 Technical Digest Series, Volume 8:128-131).

7. S. Dasgupta, H. E. Jackson, J. T. Boyd, "Two Beam Laser Recrystallization of Polysilicon on Insulator", J. of Appl. Phys. 64 (4),p. 2069 (1988).

8. H. R. Phillip, J. of Appl. Phys. 50, 1053 (1979).

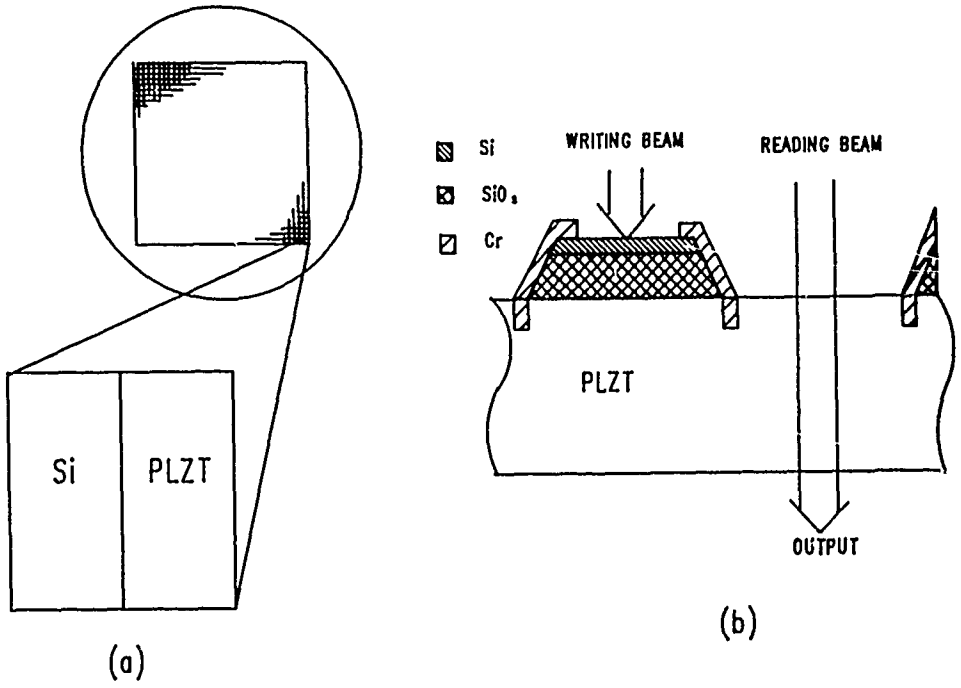


Fig. 1. Schematic of a typical Si/PLZT spatial light modulator array. Each pixel of the array consists of a photodetector fabricated on the Si which drives the electro-optic PLZT modulator.

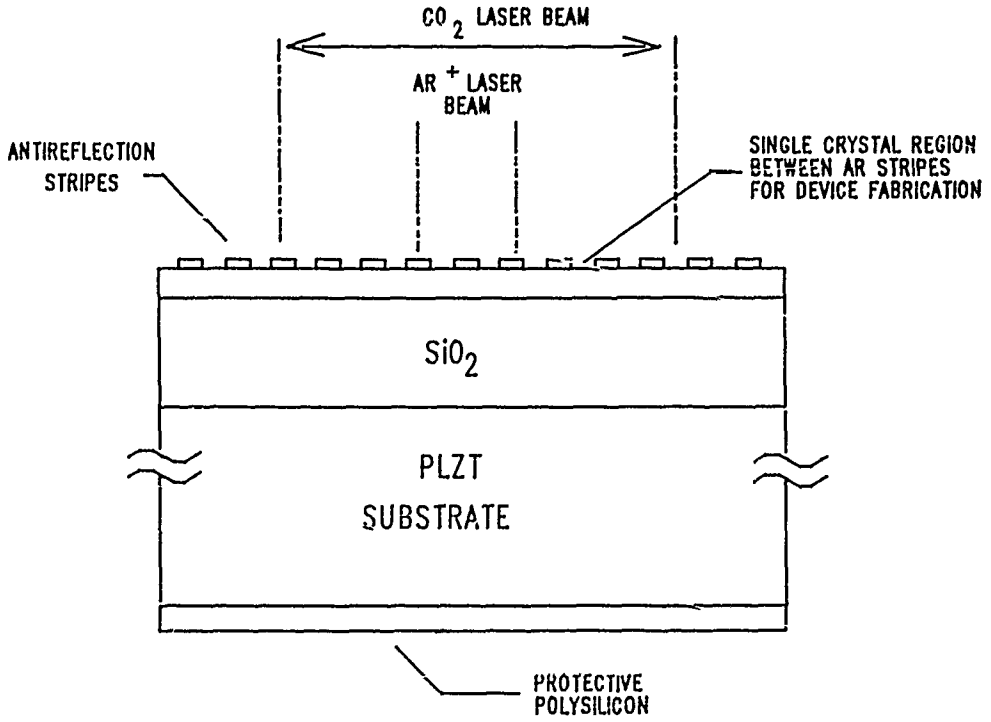


Fig. 2. Cross-section of the sample structure used to fabricate the modulator array. An isolation layer of SiO_2 is used between the PLZT and the polysilicon layer. Anti-reflection (AR) stripes are used to define the regions of single crystal grain growth during the recrystallization process.

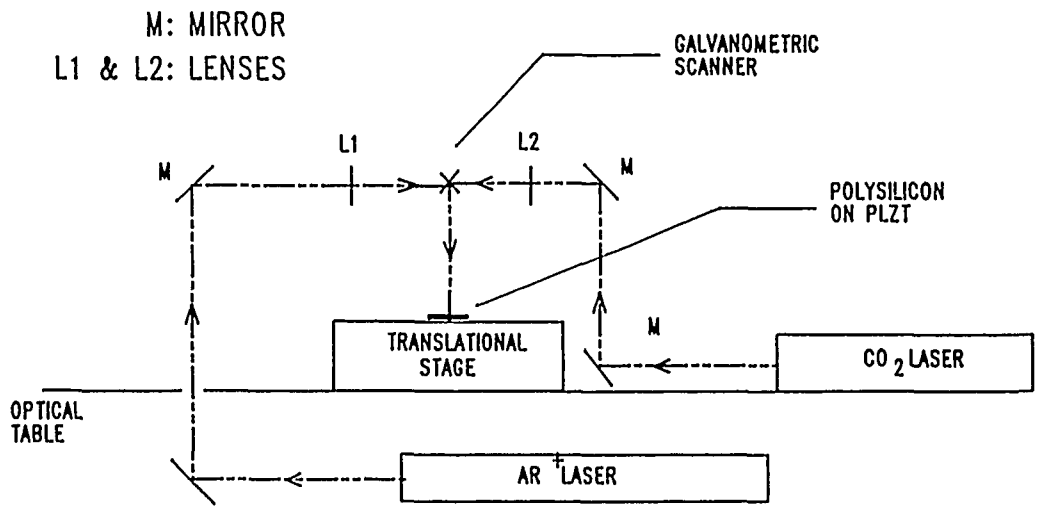


Fig. 3. Schematic of the dual beam recrystallization setup. The sample is placed on an XY translation stage. The two laser beams are focussed onto the sample through a scanner which has two mirrors mounted perpendiculaer to each other on the same shaft.

PARALLEL READOUT OF OPTICAL DISKS

Demetri Psaltis, Alan A. Yamamura, Mark A. Neifeld
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125

Seiji Kobayashi
SONY Corporation
Tokyo, 100-31, Japan

Optical memory disks have been developed in recent years as mass storage media for audio, video, and computer memory applications. Write-once systems are already widely used, and reprogrammable systems are now starting to become commercially available as well. In all the existing systems the information stored in the optical disk is recorded and readout serially by focusing a laser beam on a single pixel. With an optical memory however it is possible to illuminate the disk with an extended beam and readout (as well as record in principle) large amounts of data in parallel [1]. This distinction between serial and Parallel Readout Optical Disks (PROD) is schematically shown in Fig.1. If the potential of PRODs is realized in practice it can eliminate the bottleneck that currently exists between mass memory and the information processing portion of a computer and thus greatly impact the speed with which computers can execute memory intensive problems. There are three main issues that we will address in this paper: The suitability of commercially available disks for this applications including the experimental characterization of a prototype magnetooptic system from SONY, the limitations imposed on parallel access due to the optical system, and the types of problems and computer architectures that can make effective use of the PROD capability.

In order to readout in parallel information stored on separate tracks we need to be able to control the relative position of the recorded data across different tracks. For example, in order to record a 2-D image on a disk we must be able to align the lines of the image that are recorded on separate tracks. Commercial optical disk systems do not generally have this capability. Shown in Fig. 2a is a photograph of a standard disk, and it is obvious that there is no alignment of the data across tracks. The commercially available computer memory systems that we investigated do not have the capability to align the recorded bits along tracks. Recently, a group at SONY developed a disk drive that has the capability to record a bit in any specified location on the 2-D surface of the disk, and we are using a prototype of this system as part of a collaborative effort with SONY to look into PRODs. A photograph of a recording made on a write-once disk is shown in Fig.2b; excellent alignment across tracks is achieved. Each of the recorded spots has 1 μm diameter and the center-to-center spacing along track is $.5 - 1$ μm whereas the tracks are separated by 1.5 μm. At this recording density over 3.6×10^9 bits can be recorded on the 12cm diameter disk and the relative position between any two such bits can be controlled to within a μm. A photograph of the disk recorder apparatus is shown in Fig. 3. Most commercial disks have a plastic covering with coarse phase variations which makes their use in coherent processors very difficult and cumbersome even with incoherent light. Therefore, in the experiments we have used disks with optical quality glass covers specially made by SONY. Shown in Fig. 4 is a photograph of the diffraction pattern obtained by illuminating a 1mm diameter circular area on a write-once disk on which we had recorded a 2-D array of spots. The sharpness of the diffraction pattern is indicative of the accuracy of the positioning of the spots as well as the optical quality of the glass. The large spatial frequency ($\approx$500 cycles/mm) of the recorded information is convolved with the diffraction pattern due to the sectors in which the disk is divided in order to control the alignment. This remarkable quality of the diffraction pattern allows us for example to record and reconstruct computer generated holograms [2]. The complete characterization of the write-once disks will be presented at the conference. We have just started the characterization of the properties of the magnetooptic disks.

In general, the rate at which information can be retrieved is proportional to the optical power required. Therefore the speed-up that PRODs can provide has to come at the expense of power consumed and in many cases this consideration may pose the limitation on the degree of parallelism that can be achieved in a PROD system. Another limitation is the numerical aperture of the optical scanner that is used to select M out of N pixels stored on the disk and route them to the M parallel output channels. If we want to

arbitrarily select any M pixels then the total number of ways this can be done is

$$S = \binom{N}{M} \approx 2^{M(logN - logM)},$$

where the approximation is good for $M << N$. The optical scanning system must be capable of being set in S distinct states for this most general PROD system to function properly. For instance, if $N = 10^9$ and $M = 10^3$, then $S = 2^{6000log(10)}$ which is of course huge and impractical. This leads us to the conclusion that it is not possible to construct a *random access* memory with large parallelism. We must structure the memory and constrain the way in which the access to the stored data can be done. The most obvious constraint is to partition *a priori* the stored data in blocks that will be read out in parallel. In what follows we describe specific architectures and applications of PROD systems of this type.

The system shown in Fig.5 is designed to select any one of N/M stored blocks of data, each consisting of M bits, and present it to a fixed detector array with M elements. The entire memory to be scanned is presented at the input of the system. The output plane consists of a pinhole array that passes half the image (every other pixel). The input plane is imaged to the output through a beam deflector that shifts the image by one pixel only or not at all. Therefore depending on the setting of the deflector, one or the other half of the pixels from the disk will be selected by the pinhole array. The light that goes through the pinholes is demagnified by a factor of 2, fed back to the input, and processed in a similar fashion except that the deflector is reset for each iteration. An active element (e.g. 2-wave mixing amplifier) must be included in the loop to compensate for light loss and provide buffering. After $logN - logM$ such iterations the pattern at the input has been reduced to M bits. If the recording of the data on the disk is done by interleaving the M-bit blocks of data, then each distinct sequence of deflector settings will yield a different block of data. Notice that this architecture imposes minimum requirements on the deflection system since at any one time we only need to deflect by one pixel. This scheme does not involve mechanical motion and thus it can be very fast, limited by the response time of the optical amplifier. In the following two examples disk rotation is used as a convenient scanning method.

One of the areas where PRODs can be useful is parallel computation where contention problems that arise when different processing elements attempt to access data stored in memory simultaneously and/or in parallel are a major limitation of current machines. An example of a massively parallel architecture is a neural network. In this case the weights that are needed to specify the interconnection between the neurons in a large network require a very large memory capacity. The storage density of the optical disk can accommodate this memory requirement, and additionally the parallel access capability allows fast recall. For instance, suppose that we wish to implement a network that consists of 10^8 binary connections, or ten megabytes of memory. If these weights are stored in a serially readout disk it would require more than a minute for the disk to simply readout the weights. Thus a simple pass through the network would take a prohibitively long time. An implementation of a neural network using a PROD is shown in Fig. 6. The weights are recorded on the disk topologically matched to the network architecture itself. One or more silicon chips are used, each having photodetectors, one for each weight. The way an input signal propagates through the network is by first applying the external signal to the electronic chips, downloading in parallel the appropriate set of weights into the chip, evaluating on the chip the response of this particular section of the network, and sequencing through the complete network by rotating the disk so that the appropriate set of weights is aligned with the chips. This leads to tremendous speed up (perhaps down to several tens of milliseconds for the complete network). This approach assumes that the network can be decomposed into parts that can be sequentially executed. Fortunately, this is the case for almost all the networks that have been discussed in the literature, and most significantly multilayered feedforward networks.

If mechanical rotation of the disk is used for getting to the correct block of data, then parallel contiguous access *along* the track is not as crucial as parallel access *across* the track. This is due to the fact that a single detector on a given track can readout a string of bits recorded on this track in the same time it takes for the disk to rotate the string of bits opposite a 1-D, along track detector array. Therefore, across-track and non-contiguous (i.e. multiple heads) parallel access of information is more generically useful since it reduces the time required to get to information stored on the disk by a factor equal to the number of parallel access channels. One interesting possibility is to arrange the data in blocks recorded on the disk along radial lines that can be addressed and read out fully in parallel. The simplest method is to uniformly illuminate

a radial line and image the pixels onto a detector array. Spacing the detectors at 1μ approaches the limits of modern fabrication techniques. Alternatively the detector array can be broken into multiple segments with more widely spaced elements. In order to then read the entire vector at once, the detector segments are staggered in a fixed pattern, and the data is written on the disk in the same pattern. The system of Fig.7 is an example of radial block storage and it is aimed at a database management application. Each file is stored along a radial line on the disk and different portions of the disk are used to store different types of information (e.g. name, rank, etc. for the files of military personnel). The idea then is to be able to retrieve the entire file by entering one or more of the attributes or alternatively to enter an item (e.g. rank) and retrieve the names of all persons with the specified rank. The system shown in Fig.7 achieves this by having a long 1-D array of source/detector pairs along the radial direction. A typical disk has more than 10,000 tracks which implies that a source/detector array with an equal number of elements is required. At present it is feasible to monolithically fabricate such a device with roughly 1,000 elements. Ten or more such devices can then be combined or staggered on the disk (as mentioned earlier) to detect the entire file recorded on a radial line. The query for a name is coded as a modulation of the appropriate portion of the source array. The detector that is situated on the opposite side of the disk (one for each stored item) collects the light transmitted through the disk. If there is a correspondence between the pattern on the illuminating source array and the pattern stored on the disk, a strong signal will be sensed on the detector and the match between input and stored information will be detected. This then triggers the detector array on the top side to sample and hold the data of the file that produced the match. This information is then read out through parallel to serial conversion. In a system like this, it is possible to record more than 100,000 radial files, each with more than a kilobyte of available memory. The entire disk can be associatively searched and the data of a file retrieved within approximately 20 milliseconds (limited by the rotational rate of the disk). The strength of this approach is that it provides the very powerful capability to interrogate each file for a match in any one (or more) of the items stored in it during a single revolution of the disk. In almost all cases it would require an extremely long time to duplicate this capability with a serial read-out disk.

References

[1] Y. Abu-Mostafa and D. Psaltis, "Optical Neural Computers", Scientific American, March 1987.

[2] D. Psaltis, M. Neifeld, A. Yamamura, "Optical Disk Based Correlation Architectures", this conference.

Acknowledgements

This research is funded by the Army Research Office and the Defense Advanced Research Projects Agency. Alan Yamamura is supported in part by a fellowship from the Fannie and John Hertz Foundation.

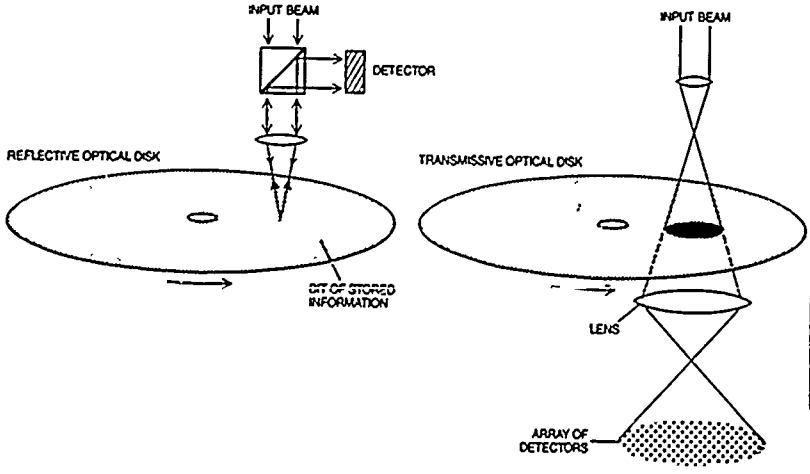


Figure 1

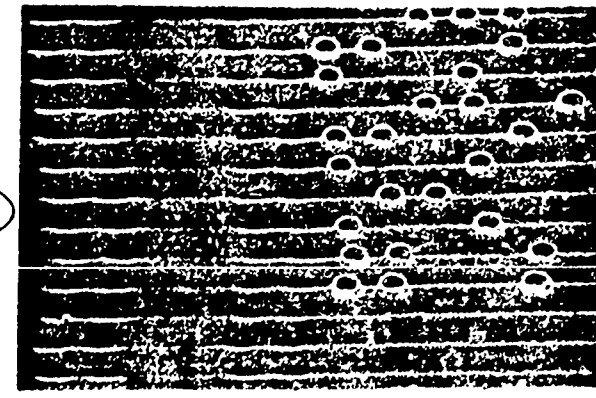


Figure 2a

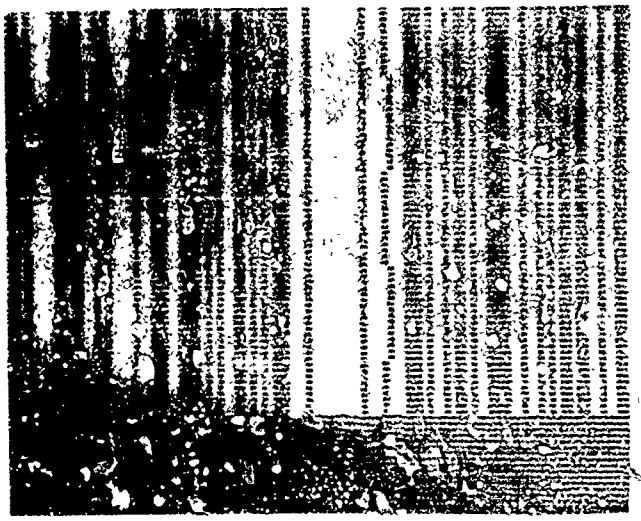


Figure 2b

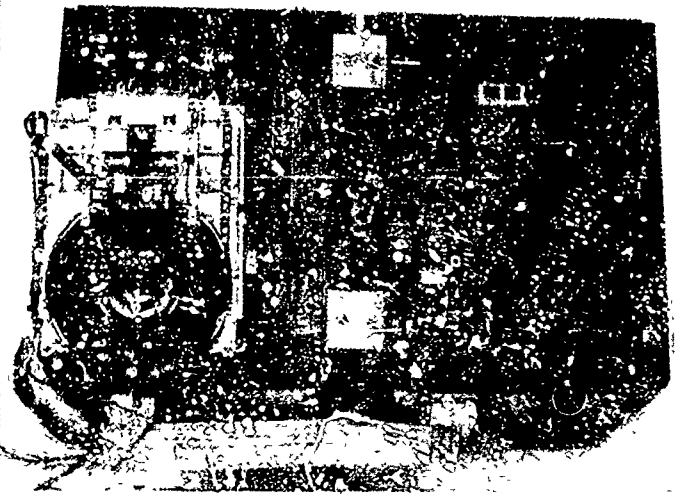


Figure 3

Figure 4

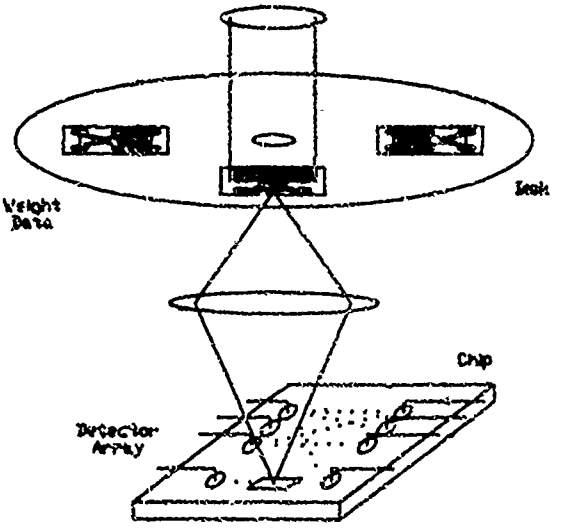


Figure 5

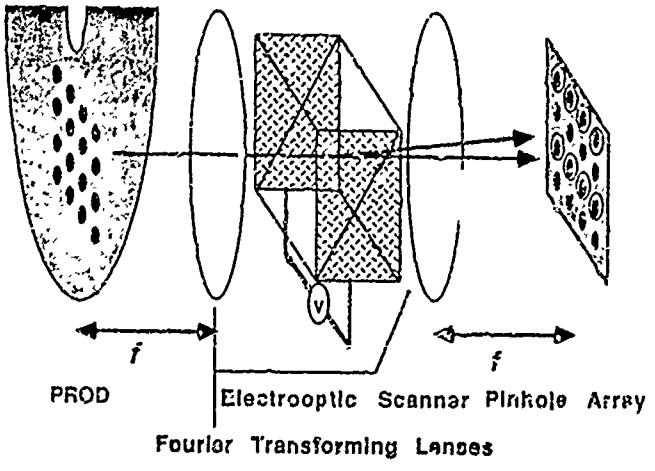


Figure 6

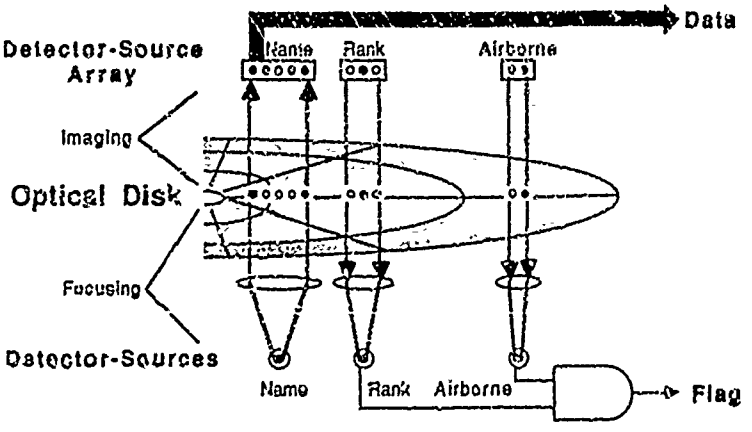


Figure 7

NOTES

MONDAY, FEBRUARY 27, 1989

SALON F

2:45 PM–3:30 PM

MF1–MF3

SLMs AND OPTICAL DEVICES: 2

Sing H. Lee, University of California–San Diego,
Presider

High-Speed Optically-Addressed Spatial Light Modulator for Optical Computing

R. A. Rice, W. Li, and G. Moddel

Electrical and Computer Engineering Department
and
Center for Optoelectronic Computing Systems
University of Colorado
Boulder, Colorado 80309-0425
(303) 492-1889

Optically-addressed spatial light modulators (OASLMs) provide a technique for processing two-dimensional optical data in parallel optical computing architectures [1, 2]. The OASLM presented uses a hydrogenated amorphous silicon (a-Si:H) photosensor and a ferroelectric liquid crystal (FLC) as the modulator.

The structure of the OASLM is shown in Figure 1. The a-Si:H p-i-n photodiode is deposited on a transparent conducting oxide (TCO)-coated glass substrate. The FLC is sandwiched between the a-Si:H and another TCO-coated glass substrate. Rubbed polymer on both the a-Si:H thin film and the second substrate induces the surface stabilized state of the FLC [3, 4].

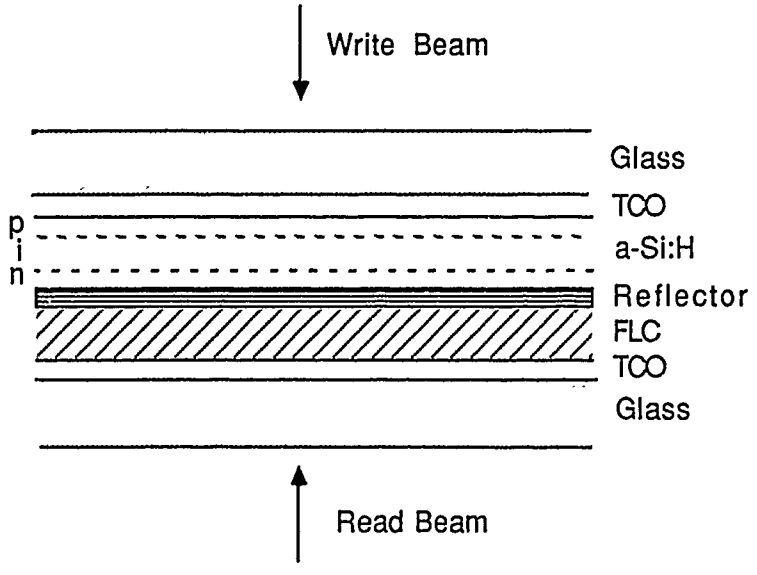


Figure 1. Optically-Addressed Spatial Light Modulator

An applied square-wave voltage drives the device. During forward bias, the a-Si:H photodiode conducts, and thus is relatively insensitive to any write-beam. The FLC is provided

with a uniform positive voltage and switches completely to a particular stable state, erasing any stored information. This is defined as the OFF state.

Under reverse bias and in the absence of a write light, the a-Si:H is highly resistive and sustains the majority of the voltage drop. Little voltage appears across the FLC, which remains in its OFF state. When the write light from an Ar laser (514 nm) is turned on, the photosensing a-Si:H thin film converts the optical image to a spatially varying electric field across the FLC. This allows the optic axis of the FLC to rotate into the ON condition. Ideally, the FLC switches ON only in those areas corresponding to illumination of the a-Si:H. The read light from a HeNe laser (633 nm) reflects off the a-Si:H/FLC interface, and therefore passes through the FLC twice, rotating the polarization of the light by 90°C where the FLC has been switched ON.

We have demonstrated an operational OASLM with a resolution of > 25 line pairs/mm, limited by the resolution of the measurement system. The response time of the device is shown in Figure 2. The overall cycle time is 0.3 msec, with a rise-delay time of 85 μsec between the onset of

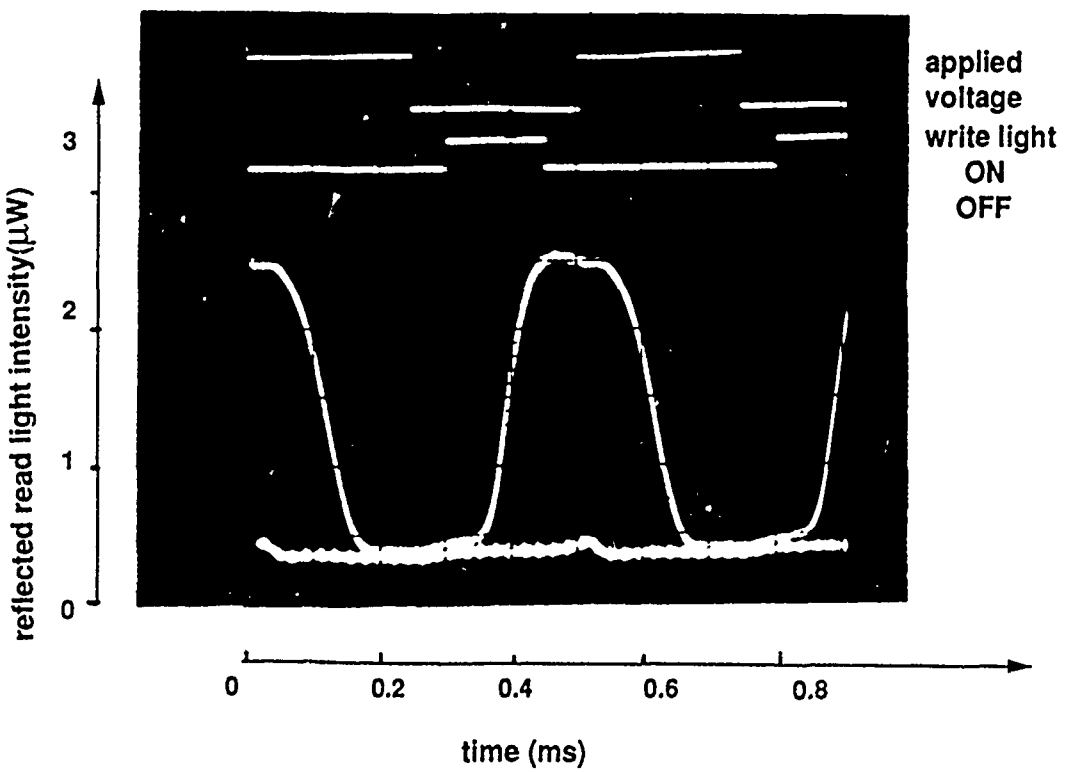


Figure 2. OASLM response to 2 KHz square-wave voltage and modulated write light (6 mW/cm^2).

the write-beam and the response, 10% - 90% rise time of 70 μsec., a fall-delay time of 50 μsec, and a 90% - 10% fall time of 80 μsec.

A circuit model of the OASLM, using the SPICE program, simulates the response time and modulation at various frame rates. Therefore the effect of characteristics such as the thickness and resistivity of the a-Si:H and FLC on the OASLM response may be modeled, and used to improve the performance of future devices.

The fabrication of the OASLM device was supported by the NSF/ERC Grant No. CDR 862236 and by the Colorado Advanced Technology Institute. The optical measurements were supported by the AFOSR under Contract No. AFOSR86-0819.

References

1. G. Moddel, K. M Johnson, and M. A. Handschy, Proc. SPIE **754**, 207 (1987).
2. K. M. Johnson, M. A. Handschy, and L. A. Pagano-Stauffer, Opt. Eng. **26**, 385 (1987).
3. N. A. Clark and S. T. Lagerwal, Ferroelectrics **59**, 25 (1984).
4. M. A. Handschy and N. A. Clark, Ferroelectrics **59**, 69 (1984).

Optical Nonlinear Neurons and Dynamic Interconnections Using the Field Shielding Nonlinearity in CdTe

William H. Steier and Mehrdad Ziari
Center for Photonic Technology
Department of Electrical Engineering
University of Southern California
Los Angeles, Calif. 90089-0483

An optical neuron and a high density dynamic optical interconnect are two of the critical elements in the realization of an optical neural net computer. The key to both elements is an optical nonlinear material with a sufficiently large nonlinearity at reasonable optical intensities which can be used at wavelengths compatible to integration with semiconductor sources and detectors. The field shielding nonlinearity in semiconductors such as CdTe is a charge transfer effect wherein an optical beam creates photocharge which then alters the electric field pattern in the material[1]. The change in electric field through the electrooptic effect can then be used to create a dynamic optical interconnect or a nonlinear neuron. In semiconductors this effect can have a formation time on the order of microseconds, has a relatively low required intensity on the order of tens of microwatts per cm^2, and can be used in the infrared. CdTe is a particularly promising material since its figure of merit, n^3r/ε, which is a measure of the index change per absorbed photon, is relatively large[2].

The optical neuron and its measured input/output response is shown in Figure 1. The response, measured at output 1 (analyzer axis), approaches that of the nonlinear sigmoid. The applied voltage is along the <111> axis, the beam propagates perpendicular to the applied field and is polarized at 45° to the field. The response was measured at 1.06μm but similar response can be expected over the wavelength band from 0.9μm to 1.4μm. Coherence is not required and LED sources could be used. The soft thresholding is caused by thermal reionization of the trapped charge which at low intensities prevents the field shielding from occurring. The measured response is for the steady state. The speed at which the neuron can change state decreases with increasing incident intensity ; it can change to the on state in 10 μsec with an input intensity of 10 W/cm^2.

By including mirrors on the sample, the neuron properties can be modified and tailored. If the etalon is tuned to resonance without an applied field, the application of the field will decrease the intensity inside the etalon since the electrooptic effect tunes the etalon for both principal polarizations; one up and the other down in frequency. However the effective birefringence is enhanced by the etalon effect. As the incident intensity is increased, the electric field seen by the intra-cavity beams is reduced by the field shielding effect, the etalon is tuned back to the incident wavelength, and the intracavity intensity increases nonlinearity. The etalon effect changes the shape of the neuron response, changes the threshold intensity, and reduces the required applied voltage. The effects of various mirror reflectivities, sample absorption, and applied fields on the neuron response will be presented.

The neuron can also be used with a polarization either parallel to or perpendicular to the applied field. The field shielding then effects the index and not the birefringence. The index change tunes the etalon response and changes the transmission as in the typical nonlinear etalon neuron. However, the index change is different for the parallel and the perpendicular polarization and hence the transmission vrs. intensity curves are different for the two polarizations. This type of response has application in backpropagating error driven learning networks[3]. The response of the CdTe neuron for these applications will be presented.

Optically controlled dynamic interconnections have application in several proposed optical computing systems and in optical networking. The field shielding effect can be used to produce a dynamic interconnect if the beam to be switched is at a wavelength which creates little photocharge (>1.4μm for CdTe) and the control beam is at a wavelength near the band edge (.85-.90μm for CdTe). The switched beam propagates and is polarized as in Figure 1 and the control beam is applied perpendicular to the switched beam or through a transparent electrode. In the absence of the control beam the electric field is approximately uniform throughout the sample; when the control beam is applied the field is present only in a small region near the negative electrode. Figure 2 shows measured electric field distributions inside the CdTe both with and without the 0.90μm control beam. An intensity of only 10 $\mu W/cm^2$ at 0.9μm was required to effect these substantial changes in the electric field pattern. These patterns were measured by observing the electrooptic effect at 1.5μm. A beam at 1.5μm placed just under the negative electrode can be efficiently switched between outputs 1 & 2 due to the electrooptic effect. Figure 3 shows the switching when the control laser diode is pulsed; it rapidly

switches to output 1 and, in the dark, remains latched for several milliseconds. The switching time decreases with increasing control beam intensity with a time of 1 μsec at 10 W/cm· It is not required that the control nor the signal beams be coherent and hence LED sources are possible.

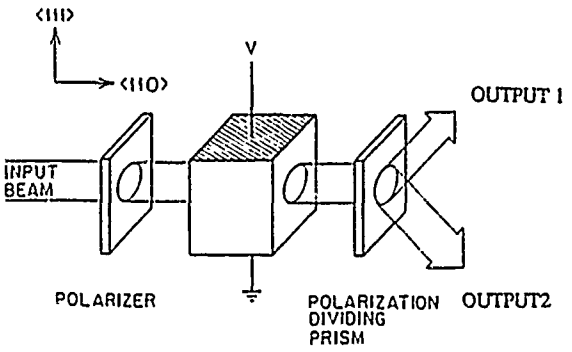
This redistribution of the electric field in high resistivity semiconductors and photorefractive dielectrics has been observed earlier[4,5] and is believed to be due to electrons excited into the conduction band by the control beam which drift toward the positive electrode but are trapped before reaching the electrode. The result is a positive space charge layer under the negative electrode and a high field region under the negative electrode.

This type of interconnect has the property that once placed into the switched state it tends to remain there for several milliseconds after the control beam is removed because of the long time required to thermally reionize the trapped electrons. Thus the interconnect can latch and hold into a particular state until erased. This can significantly reduce the energy dissipated by the control beams and the resulting thermal effects in a large array of interconnects.

These devices can be used for a 1-D interconnect array with the control beams entering through the transparent negative electrode as shown in Figure 4. Because the field occurs only over a small region when switching occurs, only modest applied voltages are required. The design and operating parameters of the array will be presented. The application of these arrays to a nonblocking crossbar by using polarization splitting optics will be reviewed.

An etalon can potentially be used on either the control or the signal beams or both to alter the switching characteristics and to reduce the required voltage. The advantages and difficulties of these approaches will be reviewed.

1. W. H. Steier, J. Kumar, and M. Ziari, Appl. Phys. Lett. **53**, 840 (1988).
2. A. M. Glass, MRS Bull. **16**, 36 (1988).
3. K. Wagner and D. Psaltis, Appl. Optics. **26**, 5061 (1987).
4. P. G. Kasherininr D. G. Matyukhin, and V. A. Sladkova, Sov. Phys. Semicond. **14**, 763 (:980).
5. V. V. Bryksin, L. I. Korovin, V. I. Marakhonov, and A. V. Khomenko, Sov. Phys. Solid State. **25**, 1244 (1982).



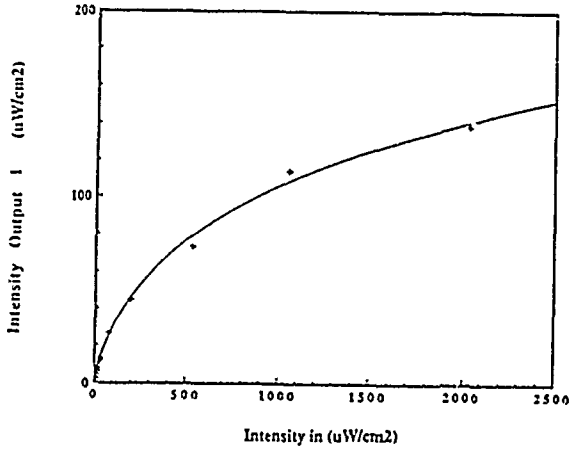


FIG. 1. Configuration of CdTe based optical non-linear neuron and its input/output response.

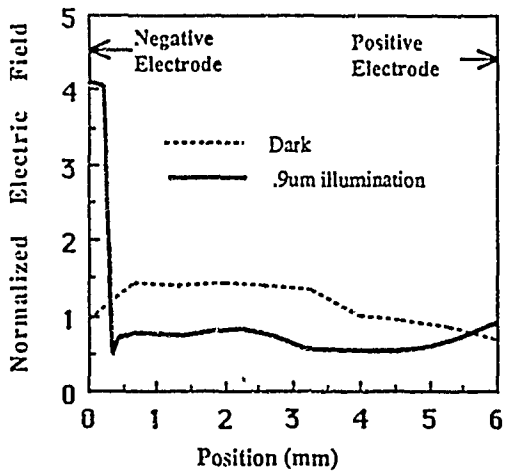


FIG. 2. Electric field distribution in the 6mm region between the electrodes. The illumination at .9 μm causes a region of high field at the negative electrode.

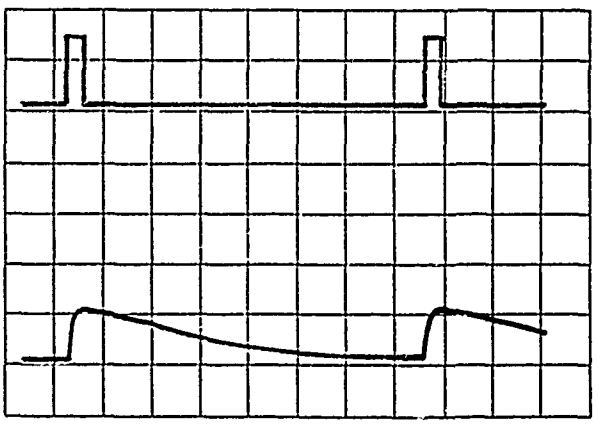


FIG. 3. Time response of CdTe to a pulsed control laser diode. The upper trace is the control laser output and the lower trace is the Intensity of output 1. The horizontal scale is 2 msec/div.

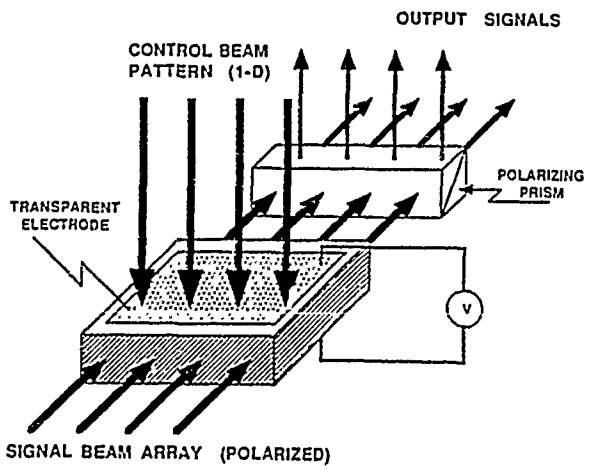


FIG. 4. One dimensional reconfigurable optical interconnects.

PHOTOREFRACTIVE NEURON BY TWO-WAVE MIXING

V. HORNUNG-LEQUEUX, Ph. LALANNE, J. TABOURY, G. ROOSEN

Institut d'Optique Théorique et Appliquée
U.A. CNRS, Bâtiment 503, B.P. 43
91406 ORSAY Cédex, FRANCE

=====

J.J. Hopfield has shown [1] that highly-interconnected neural networks have computational abilities for optimization problems and pattern recognition. The activity of a neuron can be modeled as shown in figure 1. The graded response to an excitatory or inhibitory force F is a non linear function of the input.

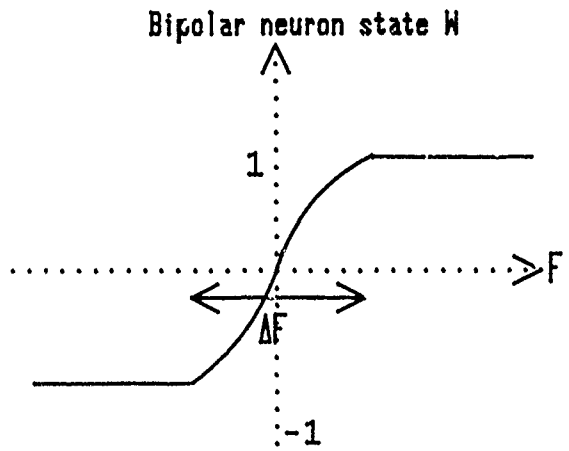


Figure 1 :
Input-output graded response of neurons.

This function can be realized by a component that preserves the wave phase and produces a saturation of the output intensity. Two wave mixing in photorefractive materials does'nt alter the phases of the interfering beams (as long as it involves a pure diffusion process of the photoinduced charge carriers) and provides amplification of one beam at the expense of the other [2, 3].

The amplified signal beam at the crystal output is :

$$I_1(L) = I_1(0)\,\frac{(r+1)\exp GL}{1 + r \exp GL}$$

where G is the photorefractive gain per unit length as defined in [2, 3], L is the interaction length of the two beams inside the crystal, $r = I_1(0)/I_2(0)$ the ratio of signal to pump intensities, at the crystal entrance. We note $\gamma = \exp GL$. For $r \ll 1$, and $\gamma r \ll 1$, the amplified intensity $I_1(L)$ is proportional to $I_1(0)$; the straight line slope is γ. When r is increased (but $r \ll 1$), γr becomes much larger than unity and $I_1(L)$ becomes quasi-constant and equal to the pump beam intensity : $I_1(L) = I_2(0)$, whatever $I_1(0)$. This condition that gives the level of the saturation is satisfied when, for example, $\gamma \geqslant 3000$ and r equal to some percents. If r is about unity, the amplified output intensity is $I_1(L) = I_1(0) + I_2(0)$. In conclusion, this short discussion shows that a saturation level is achievable for large variations of the input intensity. Moreover this saturation level can be modified by varying the pump beam intensity.

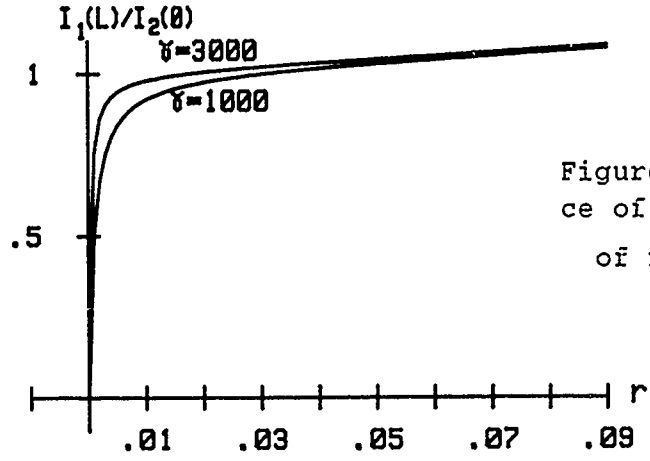


Figure 2 : Theoretical dependence of $I_1(L)/I_2(0)$ as a function of r for γ = 1000 and 3000.

In an optical implementation of neurons, losses due to the holographic synapses must be compensated at the saturation level. Moreover, this saturation level must remain quasi constant on a large range. Figure 2 illustrates, for γ = 1000 and 3000, that a large constant saturation level can be obtained as soon as $r > 0.03$ (i.e. the looses are smaller than 97 %).

In optimization problems, Hopfield [1] as shown that the graded part of the neuron response ($1/\gamma$) behaves as an effective temperature, permitting to escape from local minimum of the energy. With an incoherent third beam illuminating the crystal, an "annealing" situation can be provided by progressively decreasing the third beam intensity, thus increasing the photo-

refractive gain G (and therefore γ), corresponding to decreasing the effective temperature [3].

Experimental demonstration has been carried out using a $BaTiO_3$ crystal grown in Dijon (France). An illustration of the results is given figure 3.

The amplified signal $I_1(L)$ remains constant for a dynamic range of 10 of the incident signal intensity $I_1(0)$. The amplification gain given by $I_1(L)/I_1^W(L)$ ($I_1^W(L)$ = transmitted signal intensity without pump) varies from 10 to 100 for the r dynamic range giving a constant output signal $I_1(L)$.

Amplification with a thresholding nonlinearity has also been demonstrated on a non uniform object.

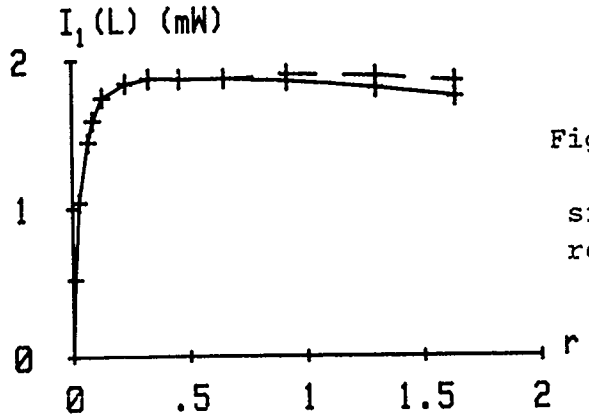


Figure 3 : Experimental dependence of the amplified signal as a function of r ; relative accuracy ≃ 1.5 %.

However our experimental arrangement gives high gains at the expense of important incident angles on the crystal. This considerably restricts the number of pixels that can be processed. A specifically cut crystal with its optic axis at an angle of the entrance face would eliminate this drawback [4].

(1) J.J. Hopfield, D.W. Tank, Biol. Cytern, 52, (1985), 141.
(2) N.V. Kukhtarev, V.B. Markov, S.G. Odulov, M.S. Soskin, V.L. Vinetskii, Ferroelect., 22, (1979), 949.
(3) V. Lequeux-Hornung, Ph. Lalanne, G. Roosen, submitted to Opt. Communications.
(4) J.E. Ford, Y. Fainman, S.H. Lee, Topical Meeting on Spatial, Light Modulators, Tech. Digest, 8, (1988), 40.

NOTES

MONDAY, FEBRUARY 27, 1989

SALON F

4:00 PM–5:00 PM

MG1–MG4

SLMs AND OPTICAL DEVICES: 3

Bernard Soffer, Hughes Research Laboratory,
Presider

MG1-1

PHOTOREFRACTIVE SPATIAL LIGHT MODULATION BY ELECTROCONTROLLED BEAM COUPLING IN SBN:Ce CRYSTALS

Jian Ma, Liren Liu, Shudong Wu, and Zhijiang Wang
(Shanghai Institute of Optics and Fine Mechanics, Academia Sinica)
(P.O.Box 8211, Shanghai, P.R.China)

The use of photoinduced dynamic refractive-index gratings for incoherent-to-coherent optical conversion or spatial light modulation has been reported, such as, four-wave mixing phase-conjugation in BSO crystals [1,2] and anisotropic self-diffraction in $KNbO_3$ crystals [3]. The coherent reconstruction in these experiments was a negative replica of the incoherent input image. We have found a new phenomenon in SBN:Ce that the coupling direction can be altered and coupling gain can be changed by an external electric field [4]. In accordance with this effect, in the present paper a new method for incoherent-to-coherent conversion and spatial light modulation using two-beam coupling in SBN:Ce is proposed, which is suited to control the replica contrast to be either negative or positive.

Our approach is based on such an idea that spatially overlapping an incoherent image to the said phase grating in a crystal will yield a modulation of coupling gain. Thus the spatial modulation can be transferred onto a coherent beam which either gains energy from pump (negative replica) or releases energy to pump (positive replica).

Consider the experimental configuration shown in Fig.1. For simplicity in calculation, we assume two coherent writing beams with intensities I_1 and I_2 being plane waves and an incoherent signal with intensity I_s being homogeneously distributed. The applicable coupled-wave equations are [5]

$$d\vec{A}_1/dz=(\gamma/I)\vec{A}_1 I_2, \qquad (1)$$

$$dA_2/dz=-(\gamma/I)A_2I_s, \tag{2}$$

where

$$I=I_1+I_2+I_s, \tag{3}$$

$$\gamma=\omega r_{eff}E/2nc\cos\alpha. \tag{4}$$

With the help of energy conservation law, Eqs.(1) and (2) are immediately integrable:

$$I_1=rI_c\exp(m\Gamma L)/[1+r\exp(m\Gamma L)], \tag{5}$$

$$I_2=I_c/[1+r\exp(m\Gamma L)], \tag{6}$$

with

$$r=I_1(0)/I_2(0), \tag{7}$$

$$\Gamma=2\mathrm{Re}(\gamma), \tag{8}$$

$$I_c=I_1(0)+I_2(0), \tag{9}$$

and the modulation factor

$$m=1/(1+I_s/I_c). \tag{10}$$

In this result, energy is coupled from beam 2 to beam 1. However, since the coupling direction is determined by the direction of c axis [6] which can be reversed by an applied electric field[4,7], we can reverse the coupling direction by reversing the applied voltage. Thus, it can be seen from Eqs.(5) and (6) that the intensity of the coherent beams are modulated by the incoherent signal, I_s. Either a negative or a positive replica will be resulted in the intensity distribution of the output (beam 1 as shown in Fig.1) with a positive or a negative voltage. It should be noted here that a pair of negative and positive coherent replicas can be simultaneously obtained by adding an optical branch for I_2, symmetrcal to I_1.

For a definite value of I_s/I_c, the optimum value of r for arriving a maximum slope of the curve (I_1/I_c versus I_s/I_c) is derived simply as

$$r_{opt}=\begin{cases}\exp(-m\Gamma L), & \text{for negative replica,}\\ \exp(m\Gamma L), & \text{for positive replica.}\end{cases} \tag{12}$$

This result is helpful for setting the experimental parameters, such as r, I_c and I_s.

During experiment, the crystal is placed at the near Fourier plane of both the incoherent image and beam 1 in order to decrease the response time and to match the dimension of the used crystal ($15 \times 10 mm^2$). A transverse electric field, E_o=5KV/cm, or E_o=-5KV/cm, is applied along c axis of the crystal to control the coupling direction.

Fig.2 shows the experimental photographs of negative and positive coherent replicas of a grey-level incoherent image respectively.

[1] Y. Shi, D. Psaltis, A. Marrakchi, and A.R. Tanguay,Jr., Appl. Optics, 22(1983)3665.

[2] A. Marrakchi, A.R. Tanguay,Jr., J. Yu, and D. Psaltis, Optics Eng., 23(1985)124.

[3] E. Voit, and P. Gunter, Optics Lett., 12(1987)769.

[4] J. Ma, L. Liu, S. Wu, Z. Wang, L. Xu, and B. Shu, Electrocontrolled beam coupling and bistable behaviour in SBN:Ce Crystals, Appl. Phys. Lett. (to be published soon).

[5] K.R. MacDonald, and J. Feinberg, J. Opt. Soc. Am. 73(1983)548.

[6] J. Feinberg, D. Heiman, A.R. Tanguay,Jr., and R. Hellwarth, J. Appl. Phys.. 51(1930)1297.

[7] D.M. Gookin, Optics Lett., 12(1987)196.

Fig.1. Experimental configuration for photorefractive incoherent-to-coherent optical conversion by using two-beam coupling. BS, beam splitter; Ms, mirrors; L_1-L_5, lenses; T, transparency.

Fig.2. Incoherent-to-coherent optical conversion of a grey-level transparency, (a) negative replica, (b) positive replica.

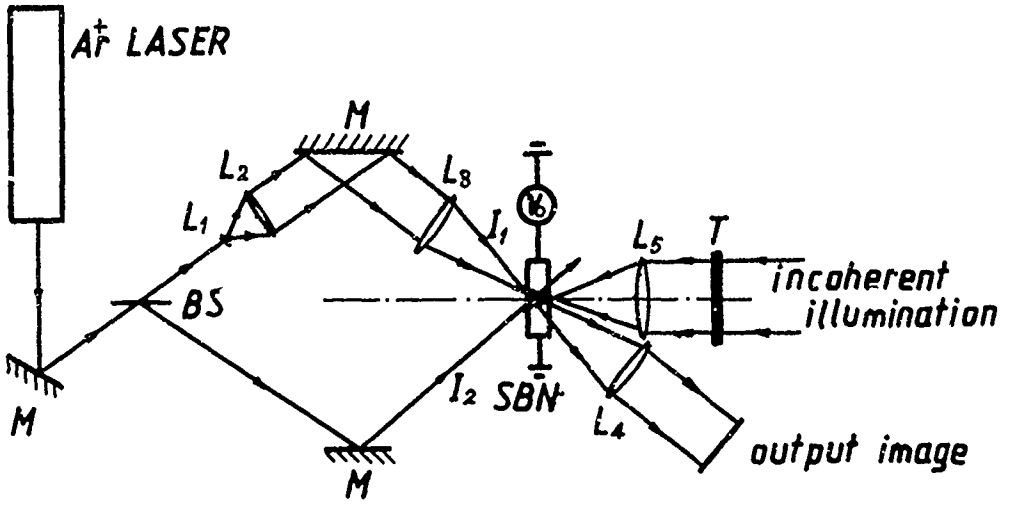

Ar+ LASER
M
L2
L1
L3
I1
V
L5
T
incoherent
illumination
BS
M
I2 SBN
L4
output image
M

Fig. 1

(a) (b)

Fig. 2

InP/InGaAs Based Charge-Coupled Devices for MQW Spatial Light Modulator Applications

K. Y. Han, R. Chang, C. W. Chen, J. H. Quigley
M. Hafich, G. Y. Robinson and D. L. Lile

NSF Engineering Research Center in Optical Computing
and Department of Electrical Engineering,
Colorado State University, Fort Collins, CO 80523

The use of optics for increasing the performance of signal handling systems beyond what can be achieved with electronics alone is becoming of more and more interest in a wide variety of areas. Optical communications via fiber links, interchip data routing via on chip emitters and detectors, integrated opto-electronics for a variety of circuit functions, and optical computing, all offer potential performance gains not only in speed and data handling capacity but also in ease of implementation. Associative memories and robot vision for example are applications seemingly naturally suited to optics.

One major component in a variety of these areas that does not yet appear to have achieved the level of performance desired by many systems architectures is the Spatial Light Modulator. Such devices certainly are available in a variety of formats, including those based on Ferroelectric Liquid Crystals, which offer outstanding contrast ratios and the potential for large array implementation. They do, however, suffer from the problem associated with any bulk phenomenon device of being quite slow. An alternative approach, demonstrated quite recently by Goodhue et. al,[1] involves using the Quantum Confined Stark Effect in MQWs to achieve electric field controlled light modulation[2,3] while a CCD provides the capability to spatially program the field across the area of the device. In this way potentially high speed Giga-bit data rate light modulation has been demonstrated using GaAs/GaAlAs with contrast ratios ~1.5 to 1[4].

One problem with GaAs based CCDs however, which has been well recognized in the literature[5], is the absence of a good dielectric and hence a lack of a two level gate technology. GaAs/GaAlAs quantum wells also operate at a wavelength ~0.85 μm, well below the bandgap of GaAs, thereby resulting in an absorbing substrate which must be removed if good transmittance in the on-state of the SLM is to be achieved. Although various techniques are available to minimize this problem it certainly is an additional factor further complicating what is to some extent an already complex device.

InP based CCD's, and the associated lattice matched InGaAs/InP MQW materials combination, suffer from neither of these difficulties. InP CCD's have been demonstrated to be fast, with f_{clock} as high as ~ 1 GHz, in even quite large geometry structures[6], to have a compatible dielectric which allows an overlapping two level gate technology analogous to that used in Si CCDs, and to have a transparent substrate over their ~1.3 to 1.6 μm operating range.

This paper will present results we have achieved on InP/InGaAs CCD performance, and on MQW field induced light absorption for high speed SLM applications. In particular we have fabricated a variety of 8 bit linear CCD arrays on InP and InGaAs epitaxial material grown by gas source MBE[7] and have evaluated their transfer efficiency, noise and linearity. We have also investigated the modulation capabilities of InP/InGaAs insulated gate MQW structures versus wavelength. The geometry of the devices we have been studying is shown in figure 1 and figure 2 shows their high frequency response[8].

Acknowledgement: The funding for this work was provided by the NSF, through their support of the ERC in Optical Computing, by the Colorado Advanced Technology Institute and by DARPA.

References

(1) W. D. Goodhue, B. E. Burke, B. F. Aul and K. B. Nichols, "Molecular-Beam Epitaxially Grown Spatial Light Modulators with Charge-Coupled-Device Addressing," J. Vac. Sci. Technol. A6, 2356 (1988).

(2) D.A.B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood and C. A. Burres, "Electric Field Dependence of Optical Absorption Near the Band Gap of Quantum-Well Structures," Phys. Rev. B32, 1043 (1985).

(3) H. Temkin, D. Gershani and M. B. Panish, "InGaAsP/InP Quantum Well Modulators Grown by Gas Source Molecular Beam Epitaxy," Appl. Phys. Lett. 50, 1776 (1987).

(4) K. B. Nichols, B. E. Burke, B. F. Aul, W. D. Goodhue, B. F. Gramstorff, C. D. Hoyt and A. Vera, Appl. Phys. Lett. 52, 1116 (1988).

(5) D. L. Lile and D. A. Collins, "Insulated Gate Inversion and Accumulation Mode Charge Coupled Devices on InP," Thin Solid Films, 103, 53 (1983).

(6) L. Messick, D. A. Collins and D. L. Lile, "A 64-bit 800-MHz Insulated-Gate CCD on InP," IEEE Electron Device Lett. EDL-7, 680 (1986).

(7) G. Y. Robinson, "Gas Source MBE Technology," First International Workshop on CBE/MOMBE/GSMBE," Sapporo, Japan, August 1988.

(8) Some of the data to be presented in this paper was obtained by one of the authors (DLL) while working at the Naval Ocean Systems Center in San Diego.

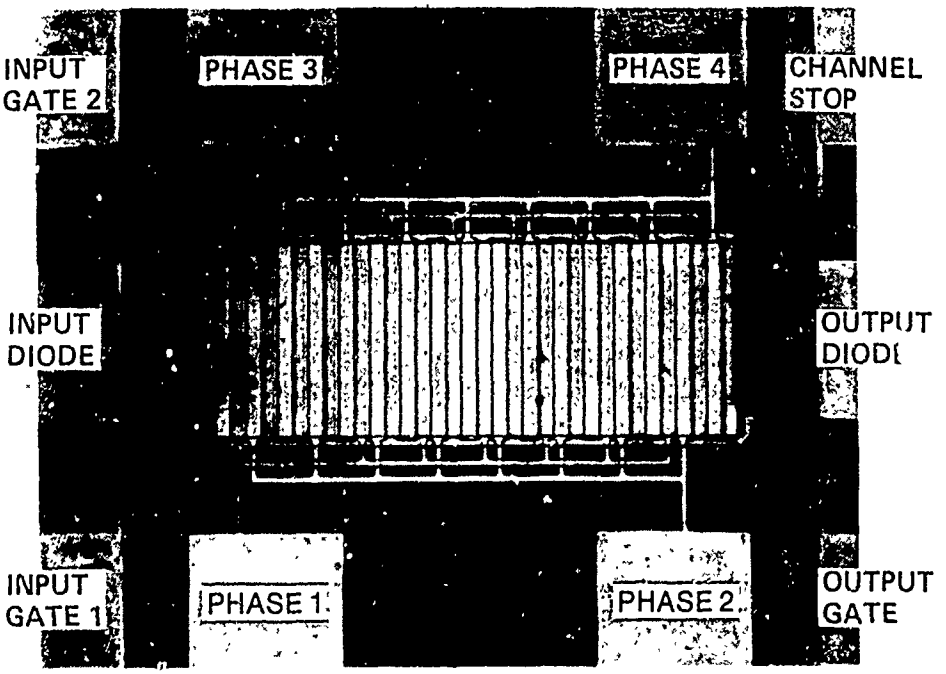
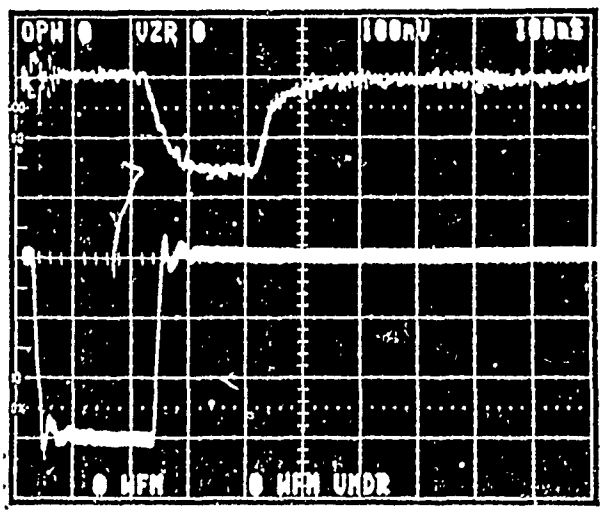
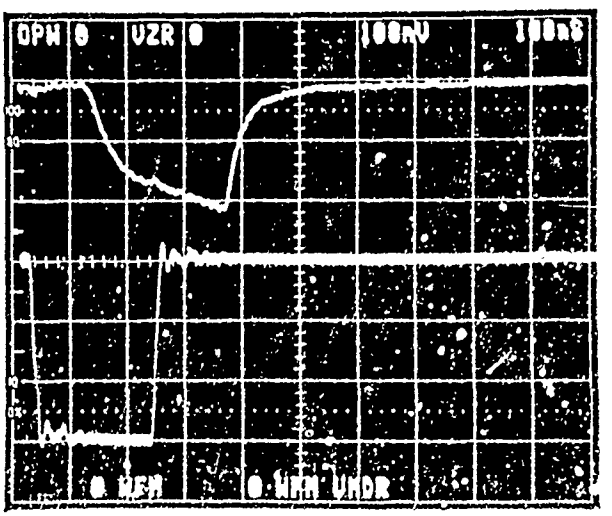


Figure 1. Geometry of the 8 bit InP/InGaAs MQW based CCD SLM. The gate geometry is ~10μm x 100μm and the 4ϕ structure is based on a two level insulated gate design.



400 MHz



800 MHz

Figure 2. InP CCD response at 400 and 800 Mz. The lower trace shows the input signal and the upper trace the resulting output.

Optical Space-variant Logic-gate Using a New Hybrid BSO Spatial Light Modulator

Zhang Ji Liu Weiwei Zhong Licheng Gou Yili

A novel method on the basis of spatial encoding technique has been advanced by T.Yatagai. In this method, the multiple instruction multiple datefluent (MIMD) is simply realizable in parallel by varing the decoding mask, but the method for encoding input pattern poses a problem in practical application. One of solution is to use the hybrid system, and encoding can be done with electronic computer. The other is using a new hybrid BSO SLM which can be used to encode input binary pattern with optical method. The hybrid BSO SLM can be used not only in encoding but also in neural logical process. In the further research, we will use it to enhance edge of pattern and to quantize pattern.

The hybrid BSO SLM and Itck's BSO PROM have identical principle in the operation. But they are different in composition. The diagram of hybrid BSO SLM was shown in Fig.1. Two BSO PROM are pressed close to both sides of a polarizer, which can keeps imput pattern on two BSO PROMs from influence each other when reading in pattern with polarized light. Optical aperture of the hybrid BSO SLM is 30x30 mm. Resolution is 1800x1800 resolution elements total. Contrast ratio $>$1500:1. One full opteration requires 50 msec. Half-wave voltage of BSO crystal is 3.9kv. But in practical operation optimum voltage is about 6kv for the highest cantrast ratio. The notable feature of the hybrid BSO SLM is that length and noise of optical process system can be reduced when it is used.

Arrangement of optical process system was shown in Fig.2. Input binary pattern A and its horizontal encoding mask pattern were written with high pressure Hg lamp on both sides of the SLM. Same were done for pattern B on the other SLM, but encoding mask pattern was vertical line. Decoding mask pattern, which has been memorized in electronic computer, can be written in the BSO PROM. The result of process was read out with 20mw He-Ne laser. All sixteen logical function have been realized in the MIMD logical way. In order to reprocess the result of the former opteration in next step, the wavelength eonversion device can be a Photo-DKDP SLM or waveguider SHG crystal device. Both of them are under research.

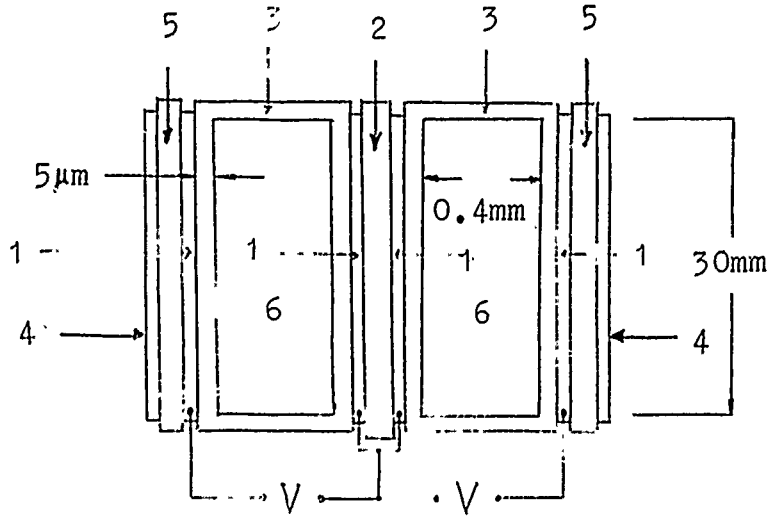


Fig.1 Composition and typical dimension of hybrid BSO SLM (1: Transparent electrodes 2: Polarizer 3: Insulator layer 4: AR coating 5: Glass plater 6: Photoconductive electro-optical crystal ($Bi_{12}SiO_{20}$))

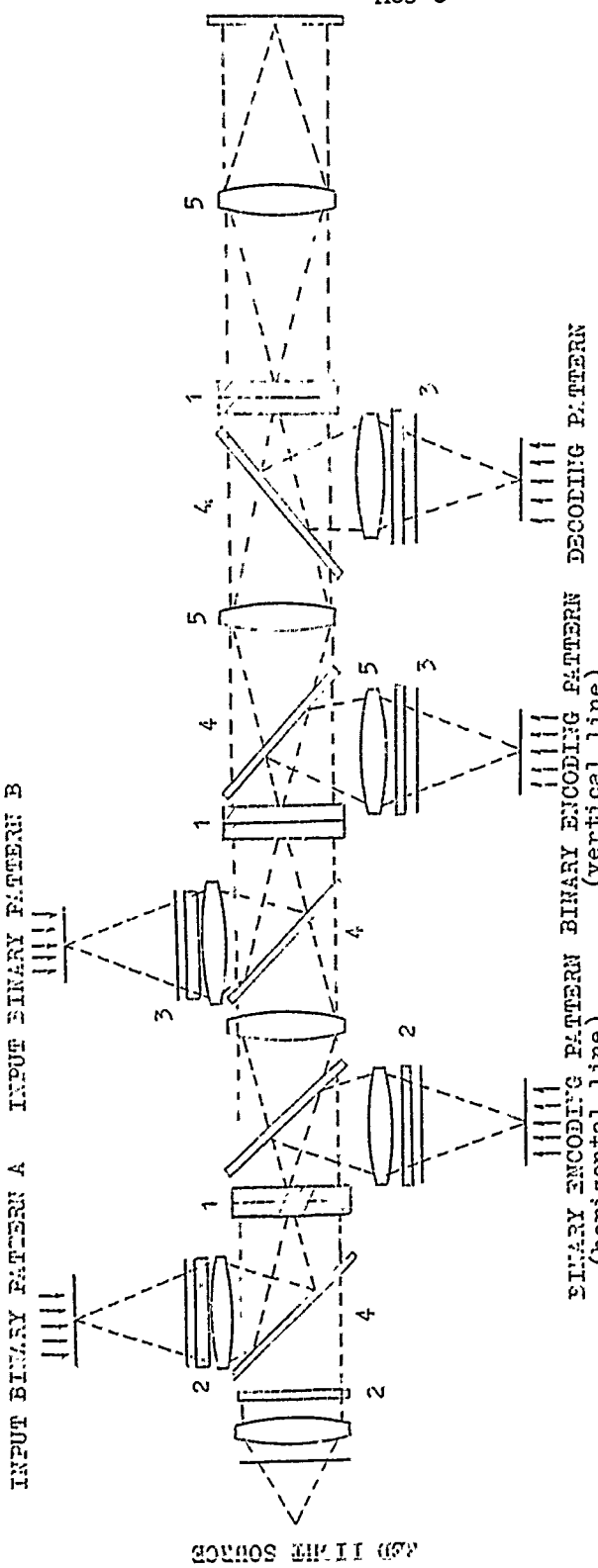


Fig.2 Arrangement of optical process system using hybrid BSO SLM
(1: Hybrid BSO SLM 2: Polarizer 3: Shutter
4: Dichroic mirror 5: Lense)

HIGH-SPEED PARALLEL OPTICAL PROCESSORS OF PHOTOREFRACTIVE GaAs

Li-Jen Cheng and Duncan T.H. Liu
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

It is known that a photorefractive crystal can act as a phase conjugate mirror via four-wave mixing or self-pumped phase conjugation. The use of a phase conjugate mirror in the interferometric system for parallel mathematic operations has been reported[1-6]. The photorefractive materials used are $BaTiO_3$[1-5] and $Bi_{12}SiO_{20}$[6]. The phase conjugation process can improve dynamic stability by reducing the sensitivity to beam path fluctuations and alignment. However, the slow responses of erials make operations not only slow, but also sensitive ntal fluctuations, such as air turbulence and v time constants shorter than those of the material resp

Recently, significant progress in the study of the feasibility of using photorefractive GaAs crystals as optical processing media has been achieved[7-9]. This paper reports the first demonstration of several basic optical processing operations using an interferometric technique with a phase conjugate mirror of photorefractive GaAs. The demonstrated processes include image subtraction, coherent and incoherent addition, inversion, parallel OR and exclusive OR(XOR) logic operations. The demonstration can be applied to other photorefractive semiconductors, such as InP and CdTe. The major advantage of using photorefractive semiconductors is the fast response

time[7,10]. This not only provides high speed operation, but also makes the system immune to low frequency vibration and fluctuation. The latter is due to the fact that the light-induced grating can follow the low frequency environment variation. The low frequency vibration and fluctuation are undesired environmental problems which commonly occur in interferometers, including those using phase conjugate mirrors of photorefractive oxides, such as $BaTiO_3$ and $Bi_{12}SiO_{20}$.

The sketch in Figure 1 briefly shows the configuration used for the experiment. A 1.06 micron light beam from a Nd:YAG laser entering from the left of the figure is split into two beams at a beam splitter(BS). Both reflected and transmitted beams, S1 and S2, are incident onto a GaAs crystal after passing through different optical paths and transparencies. Each beam creates an index grating with a coherent beam, PUMP 1, from the same laser. Each grating has a slight different orientation with respect to each other. Another beam, PUMP 2, also from the same YAG laser but incoherent with respect to the other beams travels against the direction of PUMP 1 and enters the crystal from the opposite surface. This beam is diffracted by the two gratings, resulting two phase conjugate beams, PC1 and PC2, which travel along the same path of S1 and S2 but at the opposite directions. Then, these two beams subsequently combine at BS and form an output beam which is imaged on the sensitive cathode of an infrared vidicon camera. The output image is the interference pattern of PC1 and PC2 and thus it depends on the relative orientation of the two gratings. Therefore, an adjustment of the alignment of S1 or S2 can change the relative phase between PC1 and PC2. As the result, different mathematical operations can be achieved.

Figure 1 gives a set of photographs and intensity scans illustrating that image subtraction, inversion, coherent and incoherent addition, parallel OR and exclusive OR (XOR) logic

operations can be obtained using the interferometric technique with GaAs phase conjugate mirrors. The result shows the potential of developing a fast versatile optical processor of GaAs capable of performing several basic computing operations.

It is worthwhile to note that the stable interference patterns are obtained under an experimental condition which are not usually suitable for ordinary interferometric experiment. This illustrates that a GaAs-based system has a high degree of immunity to the low frequency mechanical vibration and air turbulence, because the fast grating formation of GaAs can follow the disturbance.

This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the Defense Advanced Projects Agency and Strategic Defense Initiative Organization/Innovative Science and Technology through the agreement with the National Aeronautics and Space Administration.

References:

1. S.K. Kwong, G. Rakuljic, and A. Yariv, Appl. Phys. Lett., 48, 201 (1986).
2. S.K. Kwong, G. Rakuljic, V. Leyva, and A. Yariv,"Nonlinear Optics and Applications", SPIE vol. 613, p.36 (1986).
3. A.E. Chiou, P. Yeh and M. Khoshnevisan, "Nonlinear Optics and Applications", SPIE vol. 613, p. 201 (1986).
4. A.E.Chiou and P. Yeh, Opt. Lett. 11, 306 (1986).
5. P. Yeh, T.Y. Chang, P. H. Beckwith, Opt. Lett. 13, 586 (1988).
6. N.V. Vainos, J.K. Khoury, and R.W. Eason, Opt. Lett. 13, 503 (1988).
7. G. Gheen and L.J. Cheng, Appl. Phys. Lett. 51, 1481 (1987).
8. G. Gheen and L.J. Cheng, Appl. Opt., 27, 2756 (1988).

9. L.J. Cheng and G. Gheen, Appl. Opt. 27, 4236 (1988).
10. P. Yeh, Appl. Opt. 26, 602 (1987).

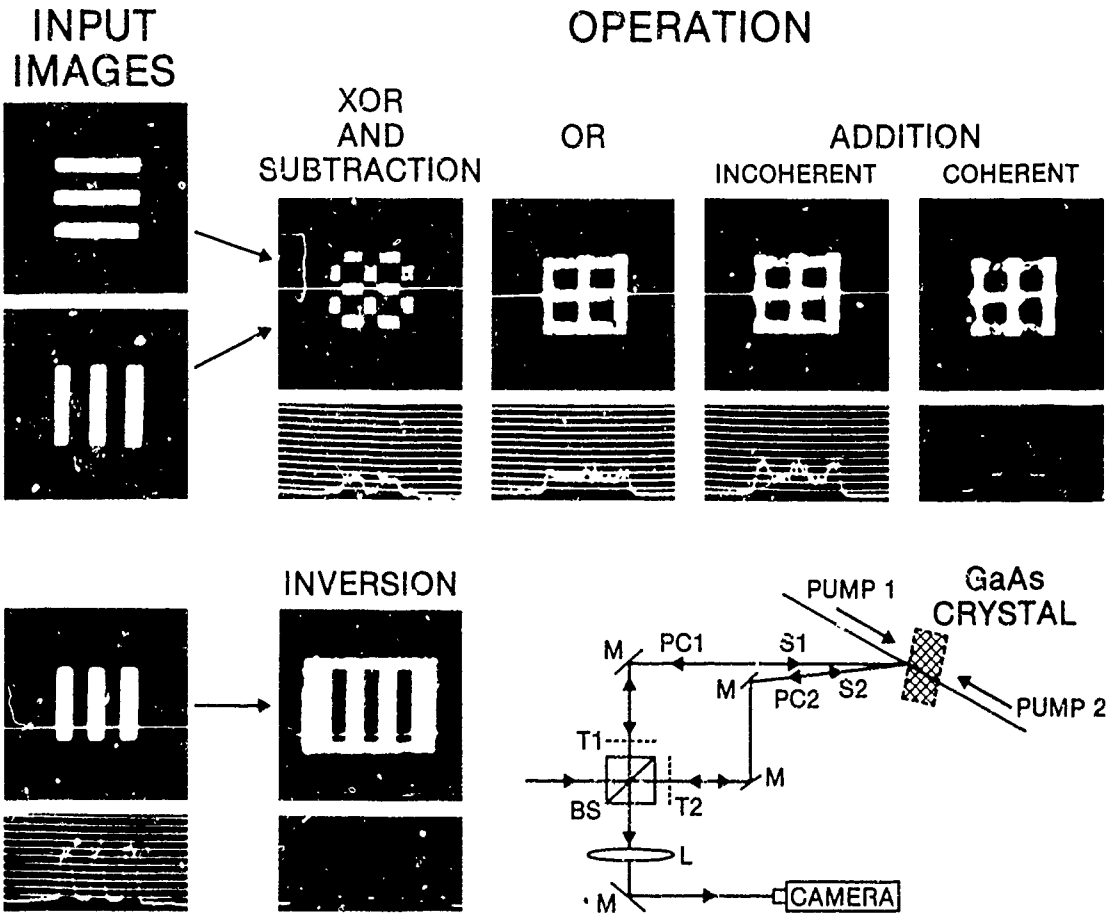


Figure 1. Obtained images and intensity scans illustrating that parallel mathematic operations can be obtained using an interferometric configuration with a GaAs phase conjugate mirror. A sketch of the experimental setup is shown at the low left part.

MONDAY, FEBRUARY 27, 1989

SALON F

5:00 PM–6:00 PM

MH1–MH4

SYMBOLIC SUBSTITUTION

Karl-Heinz Brenner, University of Erlangen-Nuremburg, Federal Republic of Germany, *Presider*

DESIGN OF A SYMBOLIC SUBSTITUTION BASED, OPTICAL RANDOM ACCESS MEMORY

M. J. Murdocca and B. Sugla
Room 4G-538, AT&T Bell Laboratories
Holmdel, NJ 07733

Abstract

Symbolic substitution is applied to the design of an optical random access memory. The design is near-optimal in gate count and circuit depth.

1 Introduction

Symbolic substitution[1] is a method of computing based on binary pattern replacement. A two-dimensional pattern is searched for in parallel in an array and is replaced with another pattern. An example of symbolic substitution is shown in Figure 1a. The pattern being searched for is called the *left hand side* (LHS) of the transformation rule and the pattern that replaces the LHS is called the *right hand side* (RHS) of the transformation rule. In Figure 1a the LHS of the rule is satisfied at two locations, so the RHS is written at those locations as shown in the transformed array. Cells that do not contribute to a LHS pattern disappear after the rule is applied. Transformation rules can be customized to perform specific operations such as addition,[1] Turing machines,[2] and sorting.[3] We have found that symbolic substitution provides rich connectivity for implementing complex functions when augmented with a log_2N interconnect such as a crossover[4] without introducing severe implementation constraints.

Consider the crossover interconnection scheme shown in Figure 1b. In this implementation of a crossover, a two-dimensional input image is passed through a beam-splitter where it is split into two identical images. One image is focused onto a mirror and is reflected back through the system to the output plane with no changes made to the spatial locations of data. The second image is passed to a grating where data is interchanged according to the period of the grating. Masks in the image planes customize the interconnect and an array of optical logic devices regenerates signals allowing for indefinite cascadability. The goal of this setup is to connect the output of every logic gate with the ouput of another gate according to the crossover pattern, except for connections that are masked out.

The crossover interconnect at any stage can be described by three symbolic substitution rules corresponding to the three angles of connections in a banyan interconnect as shown in Figure 2a. The logic gates in a banyan network can be rearranged into a crossover network by "rubber banding" the connections. This property is referred to as *isomorphism* and is characteristic of other interconnects such as the perfect shuffle.[5,6] The significance of this isomorphism is that we can design digital circuits with the banyan, which is conceptually simpler for applying symbolic substitution to gate level interconnects,[7] and map the design onto the crossover which is more efficient in terms of spatial bandwidth and conservation of light. The basic idea is to use three symbolic substitution rules for each stage of the banyan, where the rules reflect the angles of connections in each stage. Rules for the top stage of an eight wide banyan are shown in Figure 2a. Rules can be prevented from firing at specific sites by setting the corresponding locations on the masks in the image planes to be opaque. The use of the crossover allows properties of log_2N networks to be used in the design of optical digital circuits while maintaining the use of symbolic substitution at the foundation.

2 Design of the Memory

A computer memory is called *random access* if any word of the memory can be accessed in an equal amount of time, independent of the position of the word in the memory. Usually the time is logarithmic in the size of the memory. That is, if a random access memory (RAM) contains N words, then any element of the memory can be accessed in $C \cdot \lceil log_f N \rceil$ time, where f is the fan-out (here we assume a fan-out of two) and C is some constant. For a RAM of size N, $M = \lceil log_2 N \rceil$ address bits are needed to uniquely identify each word. Address bits are fed to the address decoder of the RAM which selects a word for reading or writing via an M-level deep decoder tree. Read and Write control lines determine whether the addressed location is to be read or written, and data lines provide a means for transferring a word to and from the memory. As the size of the memory grows, the length of the address grows logarithmically, so that one level of depth is added to the decoder tree each time the size of the memory doubles.

The model shown in Figure 2b illustrates the architecture of the optical RAM we propose. A two-dimensional input image contains an address and a new word of memory (when writing) and is passed

through five crossover stages of varying periods. The stored words of the memory travel through free space while the address is decoded, and then the decoded address and the memory are combined at Stage 5. Stage 6 is the final stage of writing into memory as described below. The new state of the memory is then fed back to the first stage. Regeneration and logic is performed by S-SEED devices[8] or any of a number of other suitable devices.

In order to write into the memory, the old word is erased and the new word is written in its place. There are three steps involved in writing into memory as shown in Figure 3. The first step is to find the addressed word via a decoder tree. The next step is to use that decoder tree to erase the old value. The last step is to use the decoder tree to enable the location to be written and write the new word into that location. In Figure 4, the old memory travels straight through on the right side of the diagram. It is interrupted at two locations, a NOR stage where the old word is erased and an OR stage where the new word is written. Since it is not known in advance where the new word is to be written, the new word is written to every leaf of an N-wide fan-out tree in the fan-out section. The output of the Decode section is NOR'ed with the output of the fan-out section so that one copy of the word to be written remains, in the proper location. The output of the Decode section is inverted and NOR'ed with the old memory so that every word in the old memory is enabled except for the word at the location to be written. The output is then OR'ed with the fan-out section to place the new word in the memory at the correct location.

The outputs of the Decode and Fan-out sections are superimposed to select one word at the addressed location as shown in Stage 4 of Figure 4. The Decode section is inverted and NOR'ed with the old memory to remove the old word at Stage 5, and the Fan-out memory is superimposed with the old memory at Stage 6 to yield the new memory. Reading from memory is performed in a similar manner, and is not detailed here for space considerations. The Read circuitry is included in Figure 4.

3 Discussion and Conclusion

Component count can be improved. Note that when the stored words of the memory travel alongside the decoding and memory collection trees that they do not contribute to any logic operation except where the flow is interrupted. For the levels where memory is simply flowing from one level to the next with no computation taking place on the memory itself, *no logic is needed.* Free-space propagation with appropriate delays provides the means for maintaining data on these levels, which improves the amount of logic devoted to storage to between one and two switching components per stored bit of information, the exact number depending on the size of the memory.

A design for a RAM based on symbolic substitution using planar arrays of optical logic gates interconnected in free space is proposed. Conventional serial readout is possible as well as parallel readout, which is an advantage over electronic integrated circuits. The latency between the time the address is presented and the outputs appear is $2\lceil log_2 N \rceil - 1$ gate delays for reading an N bit memory in serial. The latency is $\lceil log_2 N \rceil$ gate delays for parallel readout and $\lceil log_2 N \rceil + 3$ gate delays for serial writing. A parallel write requires only a single gate delay. The design shows that random access memory can be efficiently designed for an all optical free-space architecture.

The authors acknowledge Alan Huang for his many helpful comments on this work.

References

[1] Huang, A., "Parallel algorithms for optical digital computers," *IEEE 1983 10th International Optical Computing Conference*, 13, (1983).

[2] Brenner, K.-H., A. Huang and N. Streibl, "Digital optical computing with symbolic substitution," *Appl. Opt.*, 25, (18), 3054, (Sept. 15, 1986).

[3] Murdocca, M. J. and A. Huang, "Symbolic substitution methods for optical computing," Proceedings of the I.C.O. Meeting on Optical Computing, Toulon, France, (1988).

[4] Jahns, J. and M. J. Murdocca, "Crossover networks and their optical implementation," *Appl. Opt.*, 27, 3155, (Aug. 1, 1988).

[5] Lohmann, A. W., W. Stork, and G. Stucke, "Optical Perfect Shuffle," *Appl. Opt.*, 25, (10), 1530, (May 15, 1986).

[6] Brenner, K.-H. and A. Huang, "Optical implementation of the perfect shuffle interconnections," *Appl. Opt.*, 27, 135, (1988).

[7] Murdocca, M. J., A. Huang, J. Jahns, and N. Streibl, "Optical design of programmable logic arrays," *Appl. Opt.*, 27, (9), (May 1, 1988).

[8] Lentine, A. L., H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, "The symmetric self electro–optic effect device," in Conference on Lasers and Electro–optics, Technical

Digest Series 1987, vol. 14, (Optical Society of America, Washington, D.C., 1987, 249), postdeadline paper.

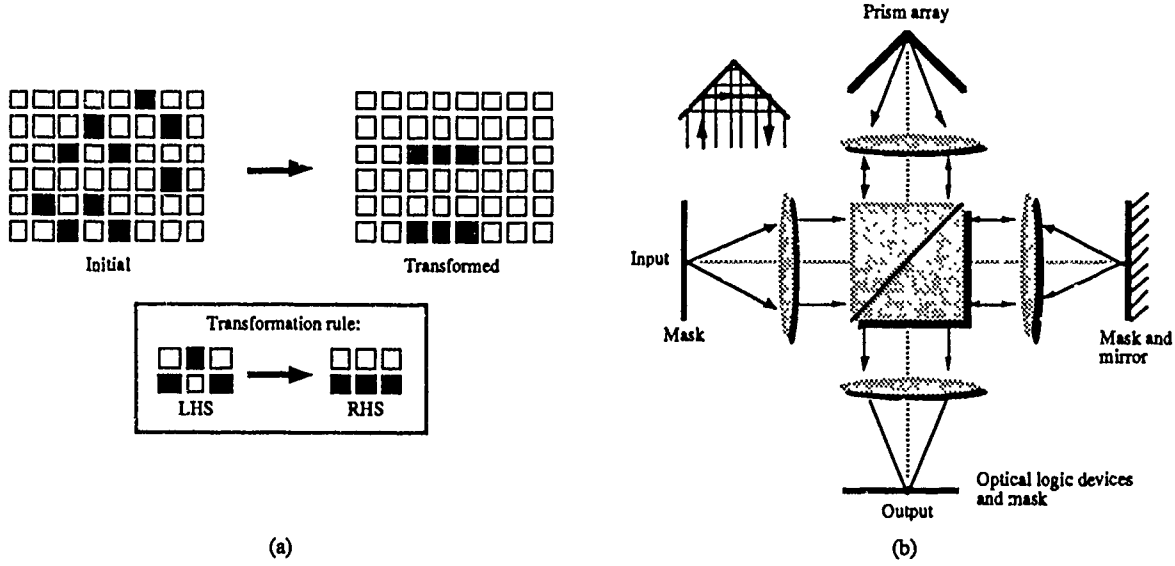


Figure 1: (a): Symbolic substitution. The transformation rule is applied to the initial array to produce the transformed array. (b): Optical implementation of crossover interconnect.

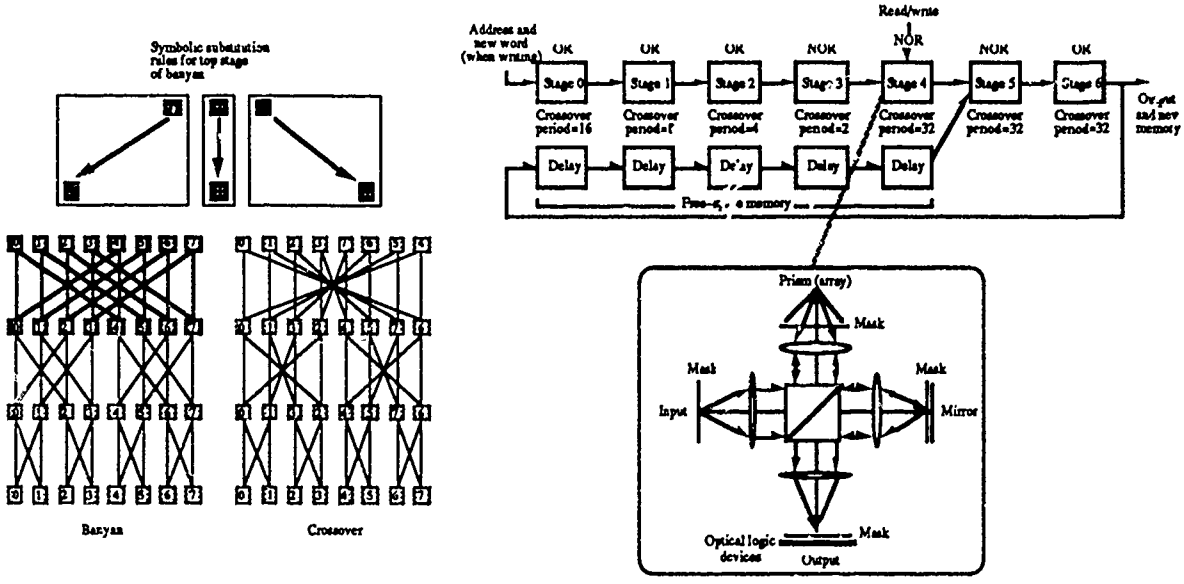


Figure 2: (a): Symbolic substitution rules for banyan interconnect, and topologically equivalent crossover network. Isomorphism between the banyan and crossover is shown in the logic gate numbering. (b): A two-dimensional input image contains an address and a new word (when writing into memory) and is passed through five crossover stages before being combined with the stored words of the memory that are propagating through free space. The output and new state of the memory are produced at Stage 6. A control image is used at Stage 4 to disable the Write logic when the desired operation is Read. Inset: A crossover stage (J. Jahns). The prism mask does not need to be in the image plane when both facets are masked.

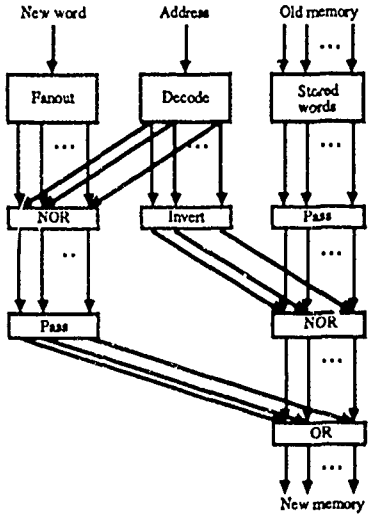


Figure 3: Block diagram of writing into RAM.

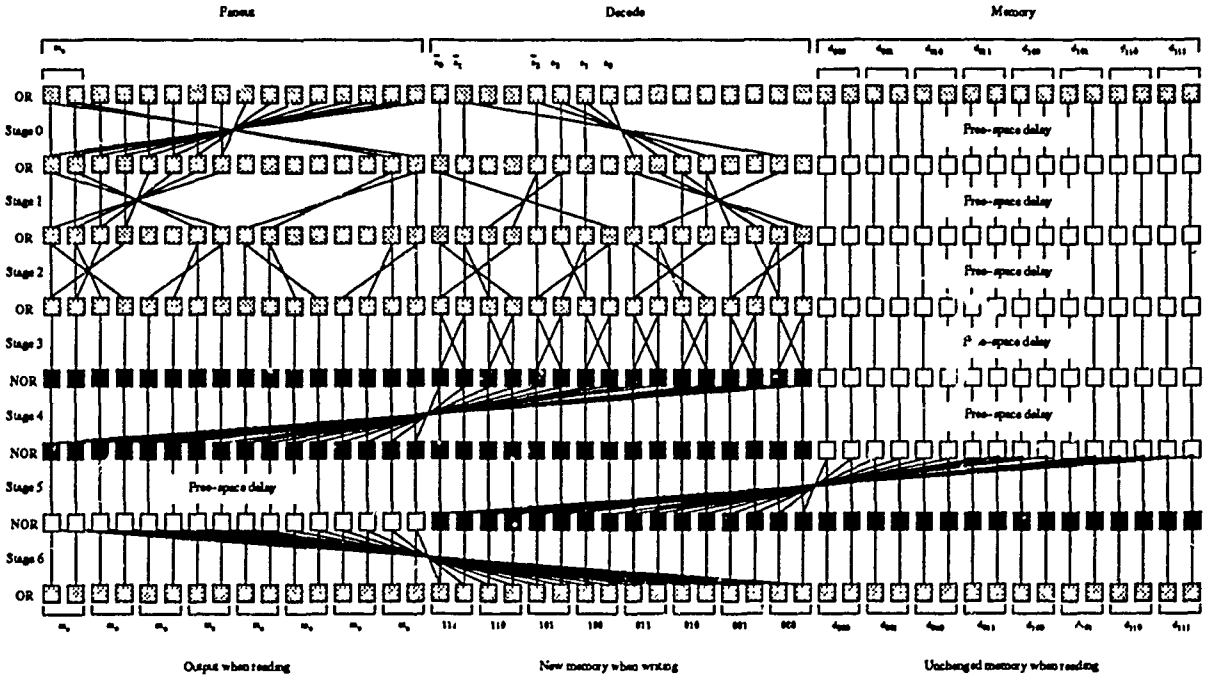


Figure 4: Memory expansion tree for writing into RAM. Fan-out tree is to the left, Decode section is in the middle, and the stored words of the memory are to the right. The Decode section is inverted and logically NOR'ed with the stored memory to remove the old word. The data to be written is then logically OR'ed with the stored memory. The Memory collection tree used for reading from memory is superimposed on stages 0-2 of the Fan-out section. Unshaded boxes indicate no logic operation due to free space propagation.

A Massively Parallel Optical Computer

Ahmed Louri
Department of Electrical and Computer Engineering
University of Arizona
ECE Building, Room 320J
Tucson, Arizona 85721
tel: (602) 621 - 2318

Abstract: In this paper, we present a new optical architecture for supporting massively parallel computations. The system processes two-dimensional arrays as basic data objects. The processing is based on the optical symbolic substitution (SS) logic. New SS rules are introduced.

1. Introduction: In this paper, we present a new optical computing architecture for implementing massively data-parallel computations. These applications exhibit a high degree of *data-parallelism* in which simple arithmetic and logic operations are simultaneously applied across large sets of data. Optical systems can simultaneously perform the same operation on all the entries of an image, hence are attractive for massively data-parallel processing. Explored in this paper is a parallel architecture that exploits optics advantages for efficiently implementing massively data-parallel algorithms, and a technique for mapping parallel algorithms onto the architecture.

2. The Parallel Optical Computing Model: Figure 1 depicts a block diagram of the basic components of the system. Unlike conventional computers that manipulate individual 0s and 1s as basic computational object, the optical architecture manipulates bit planes as basic computational objects. Up to three bit planes can be processed in parallel. For bit planes of $n \times n$ entries, it follows that up to $3 \times n^2$ operations are performed concurrently. The heart of the architecture is the processing unit. Locally, this unit can be viewed as a bit-serial processor, since it performs one logical operation on one, two or three single-bit operands. Globally, it is viewed as a plane-parallel processor, since it performs the same operation on large sets of data encoded as bit planes in parallel. This bit-serial plane-parallel processing combination allows flexible data formats and almost unlimited precision. Optical interconnects are used for data flow in the system. The architecture is conceived to be built with optical hardware that manipulates entire images simultaneously both at I/O and processing, so that the 2-D optics parallelism is sustained throughout various stages of the computation.

2.1 The Processing Unit: The processing unit operates in the SIMD (single instruction multiple data) mode, where the same operation is applied to all the data entries. In the proposed system, processing is based on the optical symbolic substitution logic[1]. Information is coded as spatial symbols in the input planes. Computation proceeds in transforming symbols into other symbols according to a set of substitution rules specifying how to replace every symbol. The processing unit is equipped with three fundamental operations *logical NOT* that inverts all the entries of an input plane, *logical AND* that performs the logical

AND operation on the overlapping bits of two input planes, and a *full Add* that performs the full addition of the overlapping bits of all the three input planes. These operations constitute a complete logic and arithmetic set, capable of computing any arithmetic or logic function. The optical implementation of this unit will be presented in the implementation section.

2.2 Input/Output Data Routing: The data represented as bit planes is fed to the processing unit either from the data memory or the outside world through three input planes, namely A-, B-, and C-plane as shown in Fig.1. Depending on the fundamental operation needed at a given computational step, the input combiner performs three data movement functions: for the logical NOT, it simply latches the relevant input plane to the processing unit. For the logical AND, the data movement required is called the *2-D perfect shuffle*. This function performs the shuffling of the row position of the data leaving the column position unchanged. The data movement function required for the full add operation is called the *2-D 3-shuffle*. This function performs a 3-way shuffling the rows of the three input planes.

The output router is responsible for directing the processed data to its appropriate destination. It also performs three data movement functions: *feeding back* to the input combiner, a partial result such as a carry bit plane resulting from a full add operation, *sending* a final result to the data memory for storage, and *shifting* the output either in the X or Y direction by a variable number of pixels. This shift enables communication between pixels in the plane.

3. Optical Implementation Considerations: In order to process information optically, we use light intensity and positional coding for the data representation. We encode the the binary bits 0 and 1 by dark-bright pixels and bright-dark pixels respectively as shown in Fig.2a. This encoding scheme has some implementation advantages[2].

Fig.2(b-d) depicts the symbolic substitution rules required to optically implement the fundamental operations: logical NOT, logical AND, and full Add. These SS rules are derived from the truth table specifications of these operations. The left-hand sides patterns (or search patterns) of the SS rules represent the input combinations and the right-hand sides (or replacement patterns) represent the table entries. The full add operation manipulates three bits which gives rise to eight combinations. If we put the bit symbols on the top of each other, we produce eight SS rules for the full Add. similary, the logical NOT, and AND give rise to two and four SS rules respectively. Note that for the logical AND and the full Add operations, each bit is provided by a separate bit plane. These bits have the same coordinates i, j in each plane. The grouping of bits into left-hand patterns is accomplished by the data movement functions described earlier. Optical implementation of the symbolic substitution has been suggested by several researchers[3,4,5]. The processing unit can be implemented in a modular fashion, where the rules are divided into functional modules : full Add module, NOT module, and the AND module. Each module comprises the SS rules corresponding to the function to be accomplished. An incoming plane (single plane for the NOT operation, two or three combined planes for the AND and full Add operations) is dynammically directed to the appropriate module depending on the operation required. Only one functional module is active at a time. Within each module, all the SS rules are

fired in parallel. Therefore, all the left-hand sides of the SS rules are searched and replaced by their corresponding right-hand sides in parallel. Details of the optical implementation of the input and output units and of the data memory will be presented at the conference.

4. Mapping Parallel Algorithms onto The Optical Architecture: We view the mapping process as a hierarchical structure as shown in Fig.3. At the highest level of the hierarchy is the application we wish to solve, i.e. signal and image processing, vision, radar application, etc. The next level identifies the various algorithms that can be used to compute these applications, i.e. matrix algebra, numerical transforms, solutions of PDEs, etc. A further analysis of these algorithms reveals that they share a common set of high-level operations. These high-level operations can in turn be decomposed into fundamental operations such the full Add, logical NOT and AND. The rationale behind the approach is that a lot of data-parallel algorithms share common features such as localized operations, intensive computations, matrix operations, and communications patterns. So the mapping process starts by identifying a set of high-level operations that captures these features. These high-level operations are then mapped onto the optical architecture. Next, parallel algorithms are constructed upon these high-level operations. This makes their mapping onto the architecture systematic and efficient. More details about this approach will be given through concrete examples during the conference.

5. Performance: If we assume input planes of size 1000 x 1000, and about 10 Mhz processing rate, then the proposed optical architecture is able to achieve 10^{13} bit operations per sec. This will represent a three orders of magnitude throughput improvement over existing array processors. More performance analysis will be given at the meeting.

References

[1] A. Huang, "Parallel algorithms for optical digital computers," in *Proceedings IEEE Tenth Int'l Optical Computing Conf.*, pp. 13–17, 1983.

[2] K. H. Brenner, A. Huang, and N. Streibl, "Digital optical computing with symbolic substitution," *Appl. Opt.*, vol. 25, 15 Sept 1986.

[3] E. Botha, D. Casasent, and E. Barnard, "Optical symbolic substitution using multi-channel correlators," *Appl. Opt.*, vol. 27, no. 5, 1 March 1988.

[4] J. N. Mait and K. H. brenner, "Optical symbolic substitution :system design using phase-only holograms," *Appl. Opt.*, vol. 27, pp. 1692 – 1700, 1 May 1988.

[5] K. Hwang and A. Louri, "Optical multiplication and division using modified signed-digit symbolic substitution," *Optical Engineering, Special Issue on Optical Computing*, To appear March 1989.

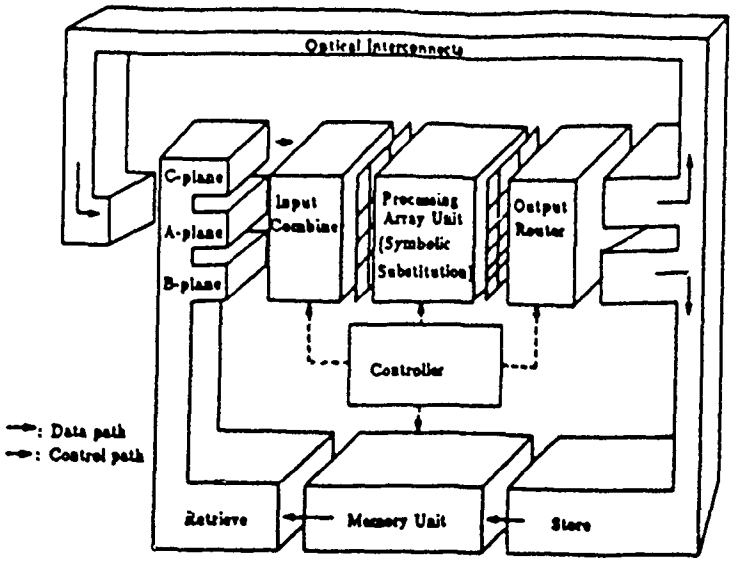
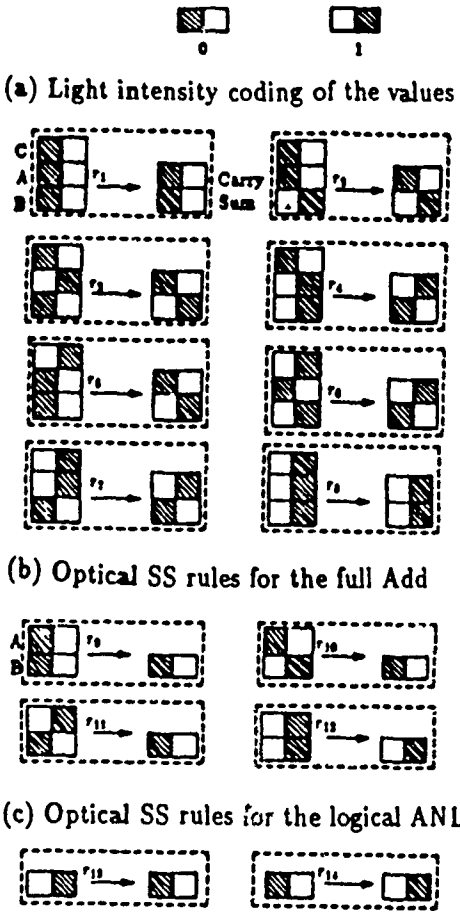


Figure 1. The block diagram of a massively parallel optical computer.



(a) Light intensity coding of the values 0 and 1.

(b) Optical SS rules for the full Add

(c) Optical SS rules for the logical AND

(d) Optical SS rules for the logical NOT

Figure 2. Optical symbolic substitution rules for the fundamental operations: full Add, logical NOT, and logical AND.

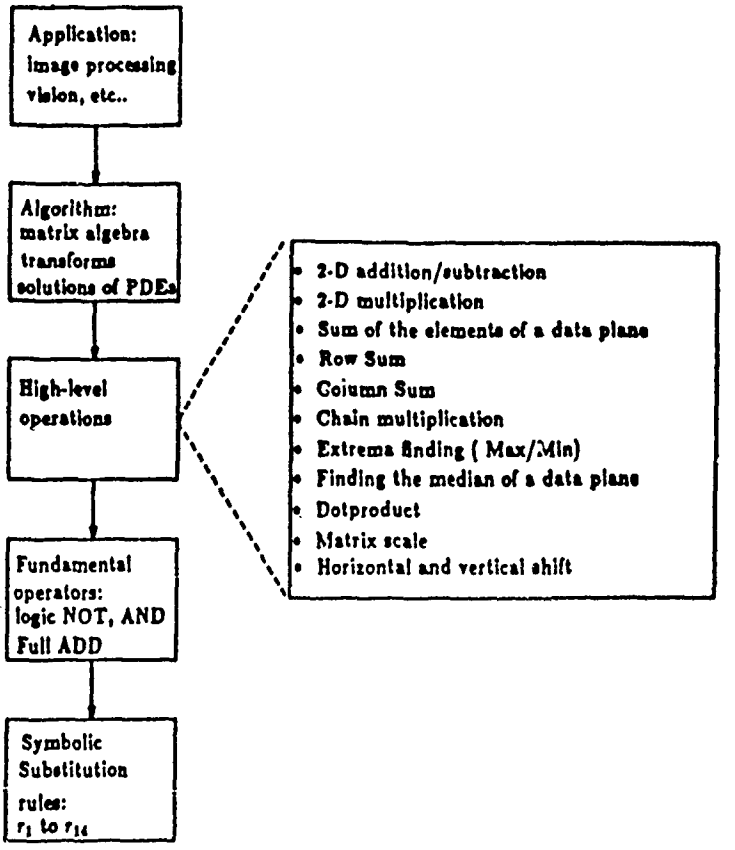


Figure 3. A hierarchical top-down approach to the mapping of parallel algorithms onto the optical architecture.

APPLICATIONS OF OPTICAL SYMBOLIC SUBSTITUTION TO IMAGE PROCESSING: MEDIAN FILTERS

Abdallah K. Cherri and Mohammad A. Karim

The University of Dayton; Department of Electrical Engineering; Dayton, Ohio 45469-0001

I. INTRODUCTION

Symbolic substitution (SS) based architectures [1] are actively sought for designing optical computing systems capable of processing binary data in parallel. The symbolic substitution is a two-dimensional parallel processing technique which maps a given pattern (referred to as search pattern) into a new pattern (referred to as scribe pattern).

A direct implementation of truth table (otherwise referred to as truth-table look-up processing) generally requires an insignificant execution time. A content-addressable memory (CAM) which is well known for its efficiency can be used for implementing a truth table. Using optical CAMs, SS based arithmetic operations such as addition and subtraction of modified signed-digit numbers are realized in either only two steps [2] or only three steps [3] irrespective of the number of bits present in the operands. In this paper, we demonstrate a particular image processing application of SS, namely, median filtering, which has been used to eliminate the noise present in an input image.

II. OPTICAL SYMBOLIC MEDIAN FILTERING

A. ONE DIMENSIONAL MEDIAN FILTERING

In one-dimensional median filtering, one takes the binary value of a pixel and replace it by the median of the binary values of this pixel and its neighbors either along a row or a column. For the case of a pixel having 2 neighbors, a window of size 3 pixels is taken in either the horizontal or the vertical direction. The binary value of the pixel position is replaced by the second largest value of the binary values present in the window.

For binary images, Table I shows all input combinations for 3-pixel window along with their expected median values. A, B, and C respectively represent either (i,j), $(i,j+1)$, and $(i,j+2)$ pixel positions of a 3-pixel horizontal window or (i,j), $(i+1,j)$, and $(i+2,j)$ pixel positions of a 3-pixel vertical window. For the 3-pixel window, only four of the input patterns (rows 5 through 8) having 1 as their median valuesneed to be recognized. Rows 1 through 4 are not considered since the medians of these patterns are 0. In general, for a one-dimensional window of size $(2 + 1)^n$ pixels a total of 2^n input patterns need to be recognized. It is obvious, therefore, that the increase in the number of to-be-recognized patterns makes the use of windows of size greater than 5 pixels relatively difficult.

The space-invariant mechanism described by Mait and Brenner [4] may be used to optically realize the SS-based median filtering. It has been shown that it is possible to construct optical systems for both recognition and substitution phases using classical elements and phase-only holographical elements. An alternative is to use an optical CAM based scheme like the one proposed by Mirsalehi and Gaylord [5]. In CAMs, holographic elements are used for storage while the system processing is based on truth-table look-up scheme. To use CAMs, the to-be-recognized patterns (which produces a 1 as an output) of Table I are subjected to a logical minimization. The resulting reduced minterms are either X11, or 1X1, or 11X where X is used to denote a don't care literal. These reduced minterms are used as references and stored in Fourier holograms. Consequently, for every output bit in the image, a total of 3 holograms will be required if one were to use a window size of 3 pixels.

B. TWO-DIMENSIONAL MEDIAN FILTERING

For example, in the case of a 3x3 two-dimensional neighborhood, the fifth largest value will be chosen as the median value. However, in practice, for a 3x3 window a total of 256

patterns will have to be recognized which makes the implementation of an optical SS questionable. Instead, one can take the median of the three values in each of the three rows and then take the median of these three in a column of three pixels. This procedure may not result in the true median, but it may be an acceptable approximation to the actual median. To realize median filtering, therefore, the one-dimensional scheme (as discussed in section A) for the 3-pixel window will have to be used twice - once along the row and then along the column.

Two-dimensional median filtering, as the one proposed herein, has its own problem. It eliminates thin lines as well as isolated points and it clips the corners. However, the horizontal and vertical lines as well as the corners can be preserved by using a 5-pixel cross-shaped window. In that case, the central pixel of the cross-shaped window takes up the third largest binary value from amongst the binary values of five pixels as its new value. Note that the cross-shaped window fails to recognize diagonally-oriented lines and corners. Consequently, the SS-based median filtering is best applied to only those images which are devoid of thin curves and sharp corners.

For the implementation of a 5-pixel cross-shaped window, the entries of Table I can be considered but with D, E, F, G, and H representing pixel positions (i,j), (i,j-1), (i,j+1), (i-1,j), and (i+1,j) respectively. Out of the 32 input combinations, only 16 are required to produce a 1 at the central pixel of the cross-shaped window. With the help of two don't care literals, the number of to-be-recognized patterns can be reduced to only 10. These patterns are shown in Fig. 1. Note that the patterns of Fig. 1 can be grouped into three classes. In the first class, for example, patterns A2, A3, and A4 can be realized from pattern A1, by respectively rotating it clockwise through 90° , 180° , and 270° . Similar rotation characteristic is also seen in the next pattern class. Note that only two patterns exist in the last class since the patterns are symmetric with respect to their centers. One can be realized from the other by a rotation of 90°. Consequently, instead of ten, only three patterns (A1, B1, and C1) are to be considered. On the other hand, with fixed A1, B1, and C1, the rotation of the input image can also be performed. Clockwise rotation of the input image is equivalent to a counterclockwise rotation of the to-be-recognized patterns. Ref. 6 describes an optical SS system that utilizes similar rotation of patterns to skeletonize a binary image Consequently, the optical system proposed in Ref. 6 can also be used to realize two-dimensional median filtering.

III. SIMULATION

For illustration purpose, a 64x64 binary image corrupted by "salt and paper noise" is considered for median filtering. Fig. 2 shows both the original as well as the corrupted image. The output of the SS-based system using a one-dimensional median filter is shown in Fig. 3 for the case of a 3-pixel window. One notices that depending on the locations of noise horizontal and vertical median filters eliminate the same and/or different noise pixels. Again, some of the noise values cannot be eliminated by either of the two filter directions. However, by applying two-dimensional median filters, most of the noises are eliminated, as hown in Fig. 4, but at the expense of loosing some of the corners of the original input. This is a small price that had to be paid in extracting the original image. By compering Figs. 4(a) and 4(b), not much of a difference is noticeable between the performances of the two two-dimensional filters.

IV. CONCLUSION

In this paper, we demonstrate the image enhancement of optical symbolic substitution based system. Median filtering is realized using SS architecture and it is shown that such an operation requires an acceptable number of substitution rules and reduced minterms.

Table I. Truth table for median filtering.

Row#	3-Pixel window ABC	Output pixel	Row#	5-Pixel window DEFGH	Output pixel
1	000	0	1	00111	1
2	001	0	2	01011	1
3	010	0	3	01101	1
4	100	0	4	01110	1
5	011	1	5	01111	1
6	101	1	6	10011	1
7	110	1	7	10101	1
8	111	1	8	10110	1
			9	10111	1
			10	11001	1
			11	11010	1
			12	11011	1
			13	11100	1
			14	11101	1
			15	11110	1
			16	11111	1
			17	00000	0
			18	00001	0
			.	.	.
			.	.	.
			.	.	.
			32	01100	0

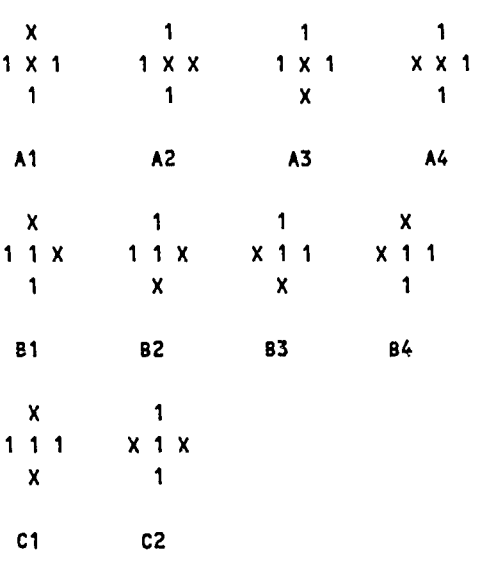


Fig. 1 Reduced minterms for the cross-shaped window.

REFERENCES

1. A. Huang, In Technical Digest, IEEE 10-th International Optical Conference, p. 13, 1983.
2. Y. Li, and G. Eichmann, Appl. Opt., Vol. 26, p. 2328, 1987.
3. A. K. Cherri, and M. A. Karim, Appl. Opt., Vol. 27, p. 3824, 1988.
4. J. N. Mait, and K. H. Brenner, Appl. Opt., Vol. 27, p. 1692, 1988.
5. M. M. Mirsalehi, and T. K. Gaylord, Appl. Opt., Vol. 25, p. 2277, 1986.
6. G. Eichmann, J. Zhu, and Y. Li, Appl. Opt., Vol. 27, p. 2905, 1988.

(a) (b)

Fig. 2 Images of (a) original input; and (b) corrupted input.

(a) (b)

Fig. 3 Median filtering output using a 3-pixel window: (a) horizontal; and (b) vertical.

(a) (b)

Fig. 4 Outputs of two-dimensional median filtering using: (a) two one-dimensional windows; and (b) a two-dimensional cross-shaped window.

PARALLEL ADDITION AND SUBTRACTION IN ONE COMPUTING CYCLE USING OPTICAL SYMBOLIC SUBSTITUTION

G. Pedrini, R. Thalmann, and K. J. Weible
Institute of Microtechnology, University of Neuchâtel
CH-2000 Neuchâtel, Switzerland

The mass parallelism offered by optical interconnection networks is exploited in the construction of optical parallel processors. Optical symbolic substitution is one form of parallel processing, which uses only space-invariant interconnections. It performs the search and replace for a set of specified spatial patterns in parallel upon an entire input matrix.[1] Symbolic substitution systems are being proposed for application in areas such as image processing or digital arithmetics, where it is desired to have a large data base operated on in parallel.

Different approaches for the realization of arithmetic processors using optical symbolic substitution have been proposed. In the implementation of a binary half adder four substitution rules are required.[1] Because of carry propagation, it requires n+1 computing cycles for the addition of two n-bit words. These cyclic iterations are time consuming and defeat the purpose of a parallel system. Other forms of number representation in the data encoding can be used to limit the number of cycles involved in the arithmetic operation. Using Modified Signed Digit (MSD) data encoding, the complete process can be carried out in three computing cycles, independent of the word length. Each of the cycles uses nine simple substitution rules involving a pair of ternary digits.[2,3]

In this paper, we present a simple technique for performing binary addition and subtraction in parallel, that is completed using only one computing cycle. The result is presented in MSD ternary form, which in our system could be reconverted, in parallel, to binary representation. The processor is based on the symbolic recognition of eight 2x2 binary symbols and the subsequent superposition of the results. Although, in our system we convert from binary input to MSD output, no ternary logic states are used throughout the process. In this paper, an optical implementation is described and experimental results are reported.

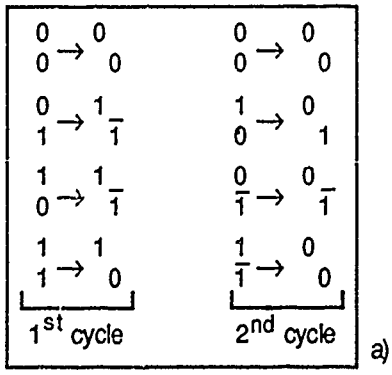
Restricted MSD addition

The MSD number representation is similar to the binary representation, except that a third digit value besides 0 and 1 is available, that is –1 (written as $\bar{1}$). The numbers are represented in the form:

$$a = \sum a_n 2^n \qquad \text{where } a_n = (1, 0, \text{or } \bar{1}).$$

As a consequence of MSD representation, each number has several possible representations. In Ref. 2, the rules for the addition of two MSD numbers using symbolic substitution are given. Due to the three possible states of each digit, nine substitution rules must be applied to each pair of digits being added. The addition process is complete after three such computing cycles.

If the input words are restricted to binary representation, the first cycle is reduced to the four rules shown in Fig.1 a). In the result of the first cycle, we find only 0's and $\bar{1}$'s in the lower digit and 0 and 1's in the upper (carry) digit. Therefore, for the second cycle, again only four rules are necessary. They all result with a 0 in the upper digit (Fig.1 a) and the addition is thus completed. Due to the fact, that this type of addition requires only two processing cycles, each digit in the result depends upon only two digit positions within the input words. It is therefore possible to



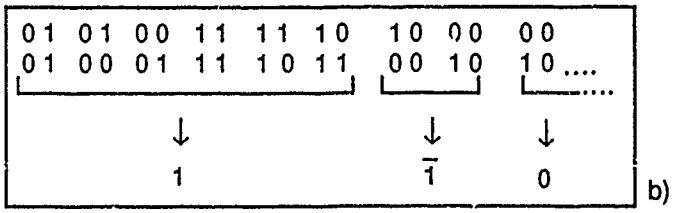


Fig.1a). Rules for MSD addition with binary input data.

Fig.1b) 2x2 bit rules for binary MSD addition, which complete the addition in one cycle.

perform the addition in one processing cycle, if 2x2 bit blocks are directly substituted. The resulting 16 rules are shown in Fig.1 b), where 6 of them result in a 1, 2 result in a $\bar{1}$, and the remaining 8 result in a 0. The use of the 8 rules for 1 and $\bar{1}$ is sufficient to decide whether the result is $\bar{1}$, 0, or 1. Note that the right and left border of the two input words have to be padded by zeros, 0's (see example in Fig.2 a), in order to accommodate for the two border columns as well.

The approach of completing the arithmetic operation in only one cycle, by the use of recognition rules involving more than one digit of the input words, has been proposed in Ref.4 for conventional MSD addition. In that case, 729 rules of 2x3 ternary blocks must be recognized and substituted. By logical minimization [4], the final number of patterns to be recognized can be reduced to 56, still impractical to be optically implemented in parallel.

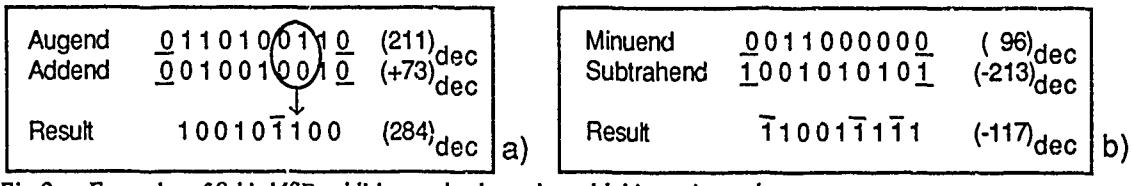


Fig.2. Examples of 8-bit MSD addition and subtraction with binary input data.

Subtraction

It is also possible, using the above computing scheme, to perform subtraction in parallel. The subtraction may be implemented by inverting the number to be subtracted, the subtrahend, and executing the same recognition rules as used for addition. The negation of a number is achieved by inverting all bit values (including the padded zeros, 0's). Thus the padded zeros, 0's, are transformed into padded ones, 1's, see Fig.2 b).

If the input data are polarization coded (see below), the inversion process (1's becoming 0's and vice-versa) may be implemented by using an appropriately oriented half-wave plate placed just behind the numbers to be negated. Ideally, the negation would be produced by an addressable spatial light modulator sandwiched with the input SLM. With such a system, the operation to be performed, i.e. addition or subtraction, becomes user selectable. Since the subtraction operation uses the same recognition rules as the addition, both operations may be performed in parallel upon a dataset within the same computing cycle. Thus, additions and subtractions may be performed in parallel and at the same time.

Optical implementation using symbolic substitution

The arrangement of the optical processor is shown schematically in Figure 3. The binary input data are coded on a spatial light modulator. The eight substitution rules are carried out using a multiple channel symbolic recognition unit. The NOR gate array restores the binary values after the recognition and yields a bright pixel at places where the search symbols have been recognized. Of course, a large number of word pairs may be entered in parallel if the words

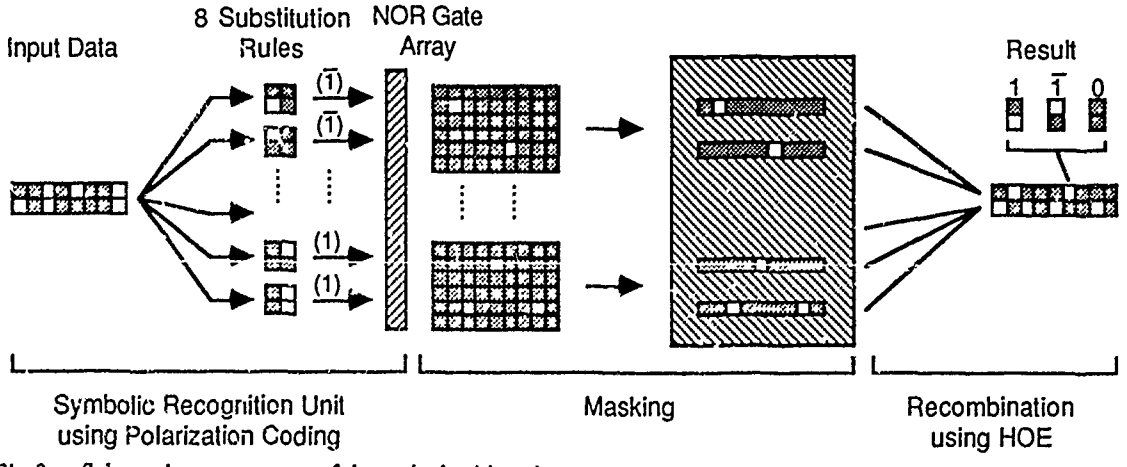


Fig.3. Schematic arrangement of the optical arithmetic processor.

are horizontally separated by the padded zeros. A mask is placed after the NOR gate array to block all but the relevant pixels of the symbolic recognition. The final stage consists of the recombination of the eight substitution rules, with the six patterns corresponding to a 1-recognition forming the lower row of the resulting word, and the two 1-patterns, forming the upper row. The three digit states of the ternary output word are thus encoded by two binary bits, dark/dark for 0, dark/bright for 1 and bright/dark for $\bar{1}$. The combination bright/bright is not a possible result.

Our experimental optical setup is sketched in Fig. 4. The recognition of the search symbols is performed in a 4f Fourier system using diffraction gratings and spatial filtering. [5] A detailed description of this setup can be found in the reference. The first 2-D grating splits the input pattern into four copies corresponding to the four pixels of the 2x2 symbols, the second grating produces the eight channels corresponding to the eight recognition rules. The input data are coded in two polarization states, [6] generated by a transmission liquid crystal display. It is desirable to encode the data using the polarization state of the propagating illumination for a couple of reasons. First, polarization encoding permits the recognition of both 1's and 0's and thus avoids dual-rail encoding necessary when using intensity. Second, the negation of an input binary word is easily obtained, using a half-wave plate, for the implementation of subtraction as described above.

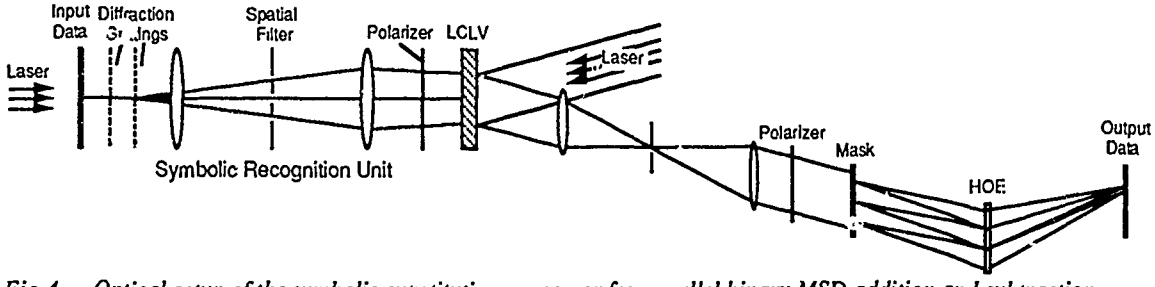


Fig.4. Optical setup of the symbolic substitution processor for parallel binary MSD addition and subtraction.

In the Fourier plane, the polarization states of some of the multiple copies of the input pattern (which appear as diffraction orders of the two gratings) are rotated by 90° according to the search symbols to be recognized in the different channels. [5] This is achieved by a half-wave plate, rotated at 45° with respect to the input polarizations, with holes in the plate passing the diffraction orders which do not need to change their polarization states (Fig. 5). The output of the 4f system is projected onto a liquid crystal light valve (LCLV) operating with a NOR gate input/output characteristic. The LCLV is read out with a plane polarized wave impinging from the opposite side of the device (Fig.4). The masking is carried out in an intermediate image plane after a telescopic imaging system. The final recombination of the different channels to form the resulting output vectors is performed by a holographic optical element (HOE) placed after the masking element. The HOE is composed of eight facets (holographic lenses), each of which produces an image of its corresponding channel appropriately shifted to achieve the superposition described above. The above superposition could also have been realized using prism elements instead of a HOE.

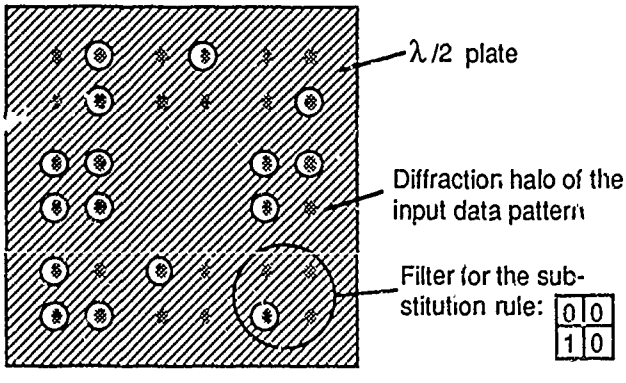


Fig.5. Half-wave plate placed in the Fourier plane of the symbolic recognition processor, which performs the filter function to realize the eight 2x2 bit rules.

Experimental results

The above described optical processor has been experimentally realized and tested. Figure 6 displays the results of both, the addition and subtraction of the examples presented in Fig. 2. In Fig. 6 a), the input data, which are polarization encoded by a transmission LCD, are displayed. The upper two rows correspond to the addition, while the lower rows correspond to the subtraction. The dark pixels in the padded pixel positions (first and last) for the bottom number indicate it as the negated value. Figure 6 b) shows the results at the system output after the holographic beam combiner. The photos are taken from a TV monitor used to observe the output data.

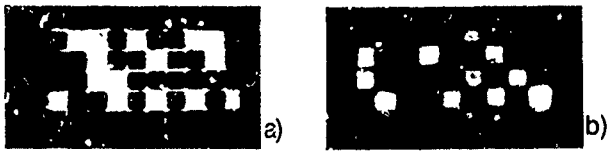


Fig.6. Experimental result of the addition and subtraction of the examples in Fig.2.

Discussion

A reduction of the space bandwidth product (SBWP) is the price which is paid for performing the digital addition and subtraction in only one computing cycle. The processor must carry out several operations at the same time, i.e. it contains multiple parallel channels. The generally precious space on the NOR gate array must be divided among the eight channels. Figure 5 illustrates the 36 channels (4 of which are not used) of the Fourier system, which must all be separated in the spatial frequency plane. Theoretically, the useful SBWP is 1/36 of the SBWP of the Fourier system, practically some safety factors need to be respected in order to avoid cross talk. Numerical example: In our setup, we used f = 38 cm, f/5 Fourier lenses. In such a system, approximately $5 \cdot 10^4$ bits could be processed, which corresponds to about 3000 8-bit additions in parallel.

The most serious problem of the proposed processor is the fact that it is not a cascadable system. The input must be binary and the output is ternary. It can however be shown, that the ternary result never contains two or more succeeding $\bar{1}$ digits. Therefore, the conversion to binary representation can be accomplished by converting all groups of the form $1[0..0]\bar{1}$ to the form $0[1..1]1$, where the brackets [..] contain an arbitrary number of 0's or 1's, respectively. After the transformation, the negative data will be represented in Two's Complement Binary form (TCB). One possible method of implementation in parallel is using a multi-channel symbolic substitution system. It is recognized by the authors, however, that this transformation technique is not very practical and a more elegant optical method to solve this problem is being sought. The symbolic substitution processor presented in this paper, although faced with the above limitation, would, of course, be well suited to prelude another system which accepts ternary input data.

The authors would like to thank R. Dändliker for many fruitful discussions. This work was supported by the Swiss National Science Foundation.

References

1. K.-H. Brenner, A. Huang, and N. Streibl, "Digital optical computing with symbolic substitution", Appl. Opt. 25, 3054-3060 (1986).
2. R.P. Bocker, B.L. Drake, M.E. Lasher, and T.B. Henderson, "Modified signed digit addition and subtraction using optical symbolic substitution", Appl. Opt. 25, 2456-2457 (1987).
3. P.A. Ramamoorthy and S. Antony, "Optical modified signed digit adder using polarization coded symbolic substitution", Opt. Eng. 26, 821-825 (1987).
4. M.M. Mirsalehi and T.K. Gaylord, "Logical minimization of multilevel coded functions", Appl. Opt. 25, 3078-3088 (1986).
5. R. Thalmann, G. Pedrini, B. Acklin, and R. Dändliker, "Optical symbolic substitution using diffraction gratings", ICO Topical Meeting **Optical Computing 88**, Proc. SPIE **963** (1988).
6. K.-H. Brenner, "New implementation of symbolic substitution logic", Appl. Opt. 25, 3061-3064 (1986).

NOTES

TUESDAY, FEBRUARY 28, 1989

SALON F

8:00 AM–9:00 AM

TuA1–TuA3

OPTICAL INTERCONNECTIONS: 1

H. John Caulfield, University of Alabama-Huntsville,
Presider

Optical Computing Research at MCC

Steve Redfield

Microelectronics and Computer Technology Corporation

MCC has been looking at the use of optics in computing systems as a means to overcome barriers which are inadequately addressed by electronics. The history, motivation, and successes of these efforts is presented.

Introduction

MCC was founded with the charter to seek revolutionary improvements in computer systems; that is, improvements giving several orders of magnitude more performance or capacity, or significantly new functionality; the focus being intelligence and parallelism. About four years ago, MCC began looking at optics with the hope that the different physics of light might be able to overcome some of the barriers encountered in trying to achieve these goals. The interest began when it was decided that the limiting constraint in designing a database machine was magnetic disk latency. One avenue of attack against this barrier was a search for alternative mass storage subsystems. Work was next expanded to see if the major inhibitor to massive parallel systems, the interconnection problem, could be successfully attacked by optics. More recently work has begun on optical neural nets.

Bobcat

In data intensive applications, it turns out that no matter how creative the system architecture, performance is always was limited by how fast data could be obtained from the disk. After looking at a number of things including holographic scanning of a stationary optical disk, we focused in on volume holographic storage in photorefractive media. This technology had been tried a couple of times in the past, but we thought advances in electro-optic devices might now make it possible. The one thing it excelled in was our very problem - latency.

A test bed, called Bobcat, was built. There were no surprises. Resolution was good enough for 10^5 to 10^6 bits in a material region with 1mm diameter surface spot. The number of recordings or pages that could be overlaid at different Bragg angles was order 10 limiting the 3-d aspect of the material. Read and write speed were very good with 10us for the read and 1ms for the write with reasonable power levels. Reads were partially destructive so after order thousand or so reads, a refresh was needed.

Work then began on overcoming the two biggest problems - capacity limits because of the small number of overlaid pages and stability due to the partially destructive read. Two significant advances have emerged from the effort. One is a novel non-destructive readout technique, based on a combination of applied electric field and the use of polarized light. It effects a highly asymmetric write/read cycle in photorefractive materials. The other is an

invention, called crystallytes, comprising a replacement for bulk photorefractive crystals. It permits much larger volumes for recording, selective control over regions in the volume, and the use of better non-linear materials.

The non destructive readout technique is a procedure for obtaining extended holographic readout in SBN. The procedure, in its optimum form, involves first recording at a spatial frequency of around 200 lines/mm for a particular length of time with a high applied electric field, around 6 Kv/cm, and ordinary polarized beams. The reconstruction is then done with the applied electric field reduced to around 1 Kv/cm and the polarization of the reconstruction beam rotated 90°. The reconstructed beam first drops in intensity, but subsequently grows in strength above the starting value, approaching 100% efficiency in some cases. The reconstruction is almost nondestructive with erasure times exceeding 3 hours of continuous readout. This equates to over 1 billion 10 us readouts with signal-to-noise ratios exceeding 20 dB due to the high efficiency.

This work was presented at the 1988 Optical Computing Conference in Toulon, France and has been described in a paper entitled "Enhanced Nondestructive Holographic Readout in SBN" by Steve Redfield of MCC and Lambertus Hesselink of Stanford University, in the October issue of *Optics Letters*.

The underlying crystallyte concept is to use a composite array of small, isolated photorefractive recording volumes in place of a bulk crystal of that material. These crystallytes are assembled in a matrix to synthesize a larger volume and may be touching or physically separated. Isolation may be achieved by refractive index differences, coatings on the sides, or the interposition of a substrate material.

The guiding of the light in fibers provides higher energy densities than are possible for free-space bulk material propagation. In holographic recording applications, longer interaction lengths give increased angular sensitivity and more dynamic range. These advantages are in addition to the ability to synthesize a much larger imaging cross-sectional area than is currently attainable using bulk materials.

Results of recording experiments suggest that an array of fibers might favorably replace bulk materials for certain computer and signal processing applications. This work was presented at the 1988 Optical Computing Conference in Toulon France and has been described in a paper entitled "Photorefractive Holographic Recording in SBN Fibers" by Steve Redfield of MCC and Lambertus Hesselink of Stanford University, in the October issue of *Optics Letters*.

Ox

In attempting to configure massively parallel processing systems, say order one thousand independent processing elements, the biggest hardware challenge is the interconnection of the processing elements. It was desired that these systems be extensible, that is, software which run on a system configured out of, say, ten nodes would also run on a system with one

thousand nodes, but appropriately faster. The ideal interconnect topology for doing this is a crossbar. The interconnect problem divides into two parts, wiring and arbitration. Optics is being looked at to address both of these problems in an effort called OX for Optical Crossbar.

Past approaches to optical crossbar design incorporate beam spreading/masking through the use of a SLM. These approaches have had problems with power dilution and the severe switching speed and contrast ratio requirements they make on the SLM. MCC has instead turned to a beam steering approach. A whole spectrum of configurations have been invented. The generic approach investigated is directed point-to-point free space connections using distributed arbitration logic with multiple access channel protocols.

These approaches use various deflectors (initially acousto-optic) in place of a masking device. A processor points its light beam via a beam deflector to the memory it wants to talk to. It appears that submillimeter size deflectors with nanosecond deflection times and 1000 resolvable spots are possible and will soon be available. It costs roughly 1 ns of deflection time per resolvable spot plus some overhead. A 2-d approach has been invented to allow deflection times on the order of 50ns for 1000 spots.

Any large-scale switch has its latency and throughput determined to a large extent by choice of protocols. Given a fixed latency budget, this often limits switch size before connectivity. To address this problem, distributed arbitration protocols, exploiting optical properties, were developed for the beam steering designs. In these designs each receiver does its own arbitration.

We are currently building a prototype of such a crossbar which is single sided. We expect this switch to be a liberator for parallel processing designs, allowing much larger numbers of processors to cooperate in the solution of non-localized problems (so-called "high flux" problems with significant communications loads). This work was presented at the 1988 Optical Computing Conference in Toulon France and has been described in a non proprietary MCC Technical Report entitled, "Ox-Design Sketches for Optical Crossbar Switches Intended for Large-Scale Parallel Processing Applications" by Al Hartmann and Steve Redfield of MCC, and also submitted to *Optical Engineering*.

Owl

One can make a rough partitioning of computer architecture into I/O, interconnect, and processing. In this last area, one of the directions in which work has been headed is what we call planar processing. Simply speaking, this is processing where the unit of information manipulated is not a string of bits or a word, but instead a 2-d plane of bits. We see a neural net as an instance of a planar processor. One of the major difficulties in implementing a large neural net is accommodating the weighted interconnects. To get experience in using optics to address this problem, we are constructing an electro-optic neural net. The effort is named Owl for *O*ptical *W*eighted *L*ogic. Initial plans were to follow a lot of the work which has been recently published using photorefractive materials to store the weighted interconnect matrix. We quickly

started to make some changes however. We wanted to use the Mean Field Theory learning algorithm which had been developed at MCC. This algorithm, which requires hidden units and multiple settling passes, naturally leads to changes. Also, based on our experiences with photorefractive recording in Bobcat, we had concerns about how many recordings or interconnect gratings could be recorded over the top of each other. A new approach, called direct projection, is being used for storing the links in the crystal. It has been simulated for a large network and found to work satisfactorily. In this optical neural network the input and output neurons are N^2 planes of pixels and the connection matrix is distributed spatially through the volume. This architecture is fundamentally different from that of other recent approaches since we use spatial rather than angular multiplexing of interconnections. We have discovered a way to correct for rescattering effects which might pose a problem by modifying the learning algorithm. We have confirmed by simulation that the changes can compensate for the crystal dynamics, beam depletion, and grating mutual rescattering. Construction of a working system is underway.

This novel optical neural network architecture using our photorefractive technology based on spatial rather than the more commonly used angular multiplexing of the interconnect gratings was presented at the 1988 Optical Computing Conference in Toulon France and has been described in a non proprietary MCC Technical Report entitled "Adaptive Learning with Hidden Units Using a Single Photorefractive Crystal" by Carsten Peterson and Steve Redfield of MCC.

Octopus

Recently, MCC has undertaken a study for DARPA on the injection of optics into existing or near future parallel processing systems. When DARPA proposed this study, individuals from the optical community suggested that MCC with its systems perspective might be ideal to lead it. We initially had some concerns about the potential for success, but eventually formulated a proposal which was promising. It was awarded to MCC and we have just started the work which we call Octopus, for *O*ptical *C*omponent *T*echnology for *P*arallel Comp*u*ter *S*ystems. The proposal divides the uses of optics into four categories: plug compatible where the optics is directly inserted into the system, interface modifying where the optics requires an interface change, system modifying where the optics requires a system organization change, and computational paradigm modifying where the optics utilizes to a new execution model.

The study will consist of three phases. Activities in each phase can be broadly categorized as (a) measurement/modeling, (b) opto-electronic technology application, and (c) architecture/systems design. The first phase will focus on plug compatible and interface modifying solutions, the second phase on system modifying changes and the third on new computational paradigms.

Conclusion

The emphasis at MCC in its optics work has been optics in computing, not optical computing. We seek new capabilities, tools, if you like, for our architects tool kit with which to build the computer systems of the future. We have strong hope for success.

Modified Brewster Telescopes

Adolf W. Lohmann, Wilhelm Stork
University of Erlangen, Physics
8520 Erlangen, Fed. Rep. of Germany

The telescope of Brewster (1781-1868) experiences today a renaissance. It is used for semiconductor lasers in order to change the beam shape from elliptical to nearly circular. It is also used for the temporal compression of laser pulses. In our application the Brewster telescope served to convert a square-shaped beam into a rectangular beam with an aspect ratio of 2:1. Such an anamorphotic process is needed in the context of perfect shuffling, which is an important link in many optical communication networks.

The particular version of the perfect shuffle, for which these modified Brewster telescopes are intended, requires two types of anamorphic changes of format /1/. The macro type will squeeze a quadratic array of pixels into an rectangular array, or vice versa. This change of format is needed if a 1D perfect shuffle is applied upon a 2D array.

The micro type consists of an array of micro Brewster telescopes, one for every data channel. A data channel may consist of a single pixel, or for example of 2x2 pixels, or more. In any case, the job of the array of micro Brewster telescopes is to squeeze every data channel by 2:1 such that there will be enough space to interlace another set of data channels among the array of squeezed channels. When operated in reverse, the Brewster arrays will blow up every data channel by 2:1 such that former gaps between channels are filled in.

For illustration we show in figure 1 two Brewster telescopes, with two or four prisms. Below it is indicated how the Brewster system fits as "afocal system" into an image forming setup.

Brewster telescopes will be more compact if the ordinary prisms are replaced by Amici prisms (fig. 2). Amici prisms are more costly since they consist of two or three kinds of glasses. A cheaper modification of the Brewster telescope emerges, if the concept of a Wadsworth prism is combined with it. It achieved straight view with only two prisms, and it is laterally more compact than the original design of Brewster.

References:

A.W. Lohmann, W. Stork, G.Stucke: Appl. Opt. 25 (1986) 1530
A.W. Lohmann: Appl. Opt. 25 (1986) 1543

BREWSTER-WADSWORTH TELESCOPE

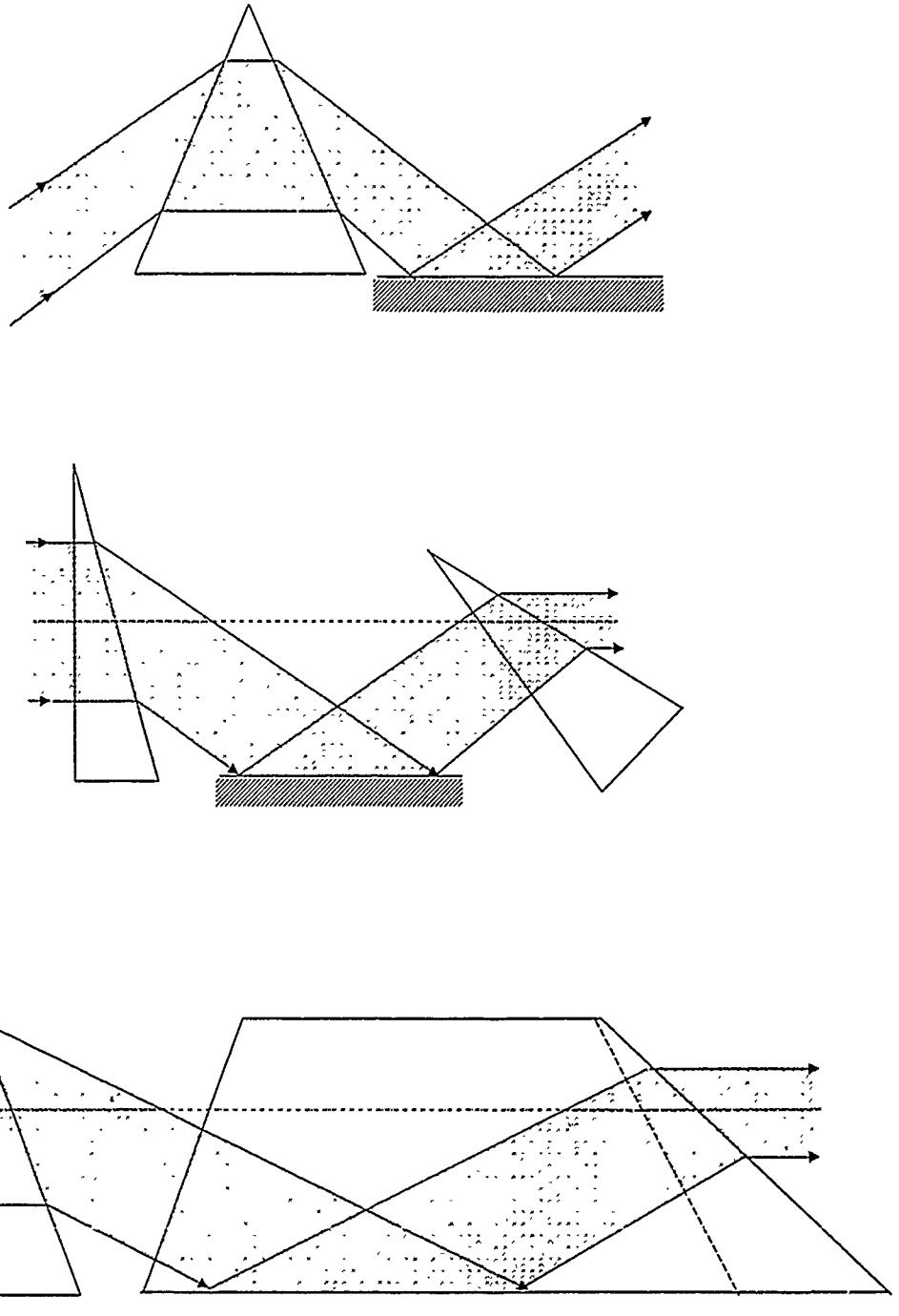
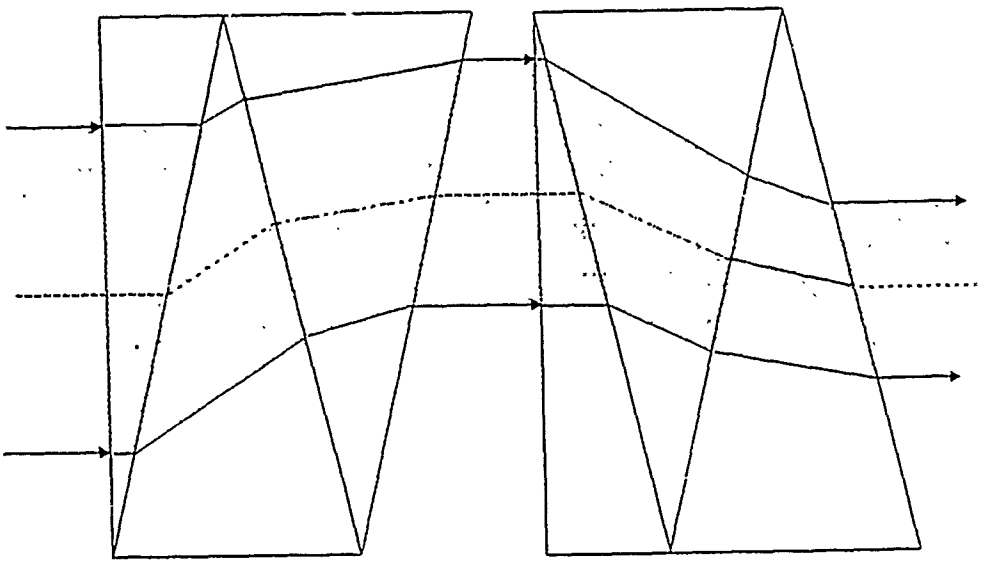


Figure 1

BREWSTER-AMICI TELESCOPE



BREWSTER TELESCOPES

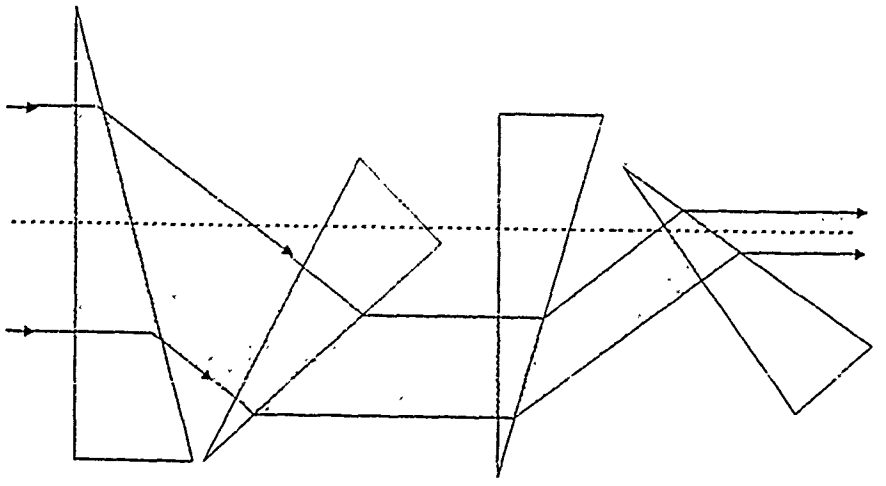


Figure 2

Optical Implementations of Interconnection Networks for Massively Parallel Architectures

Julian Bristow, Aloke Guha, Charles Sullivan, Anis Husain
Honeywell Sensors and Signal Processing Laboratory

Introduction: the I/O bottleneck

As semiconductor and electronic technologies approach fundamental physical limits in scaling and performance, the trend in high-performance computer design is to use large-scale to massively parallel architectures [1, 2, 3]. While the technology to design high-speed processing elements (PEs) has progressed significantly, the progress on designing high-performance interconnection network has not been adequate. Unfortunately, the bottleneck in performance of massively parallel architectures is typically the limited bandwidth of current interconnection networks. This is because while PEs can be densely packed on a printed wire board (PWB), there is never enough space on the board to provide all interconnection channels required for inter-PE communication at the maximum possible bandwidth. As a result, each PE is usually destined to communicate serially, often sharing communication channels with other PEs (e.g., 16 PEs in the Connection Machine share one serial line). This problem is particularly critical in fine-grained architectures where the processing time in relatively simple PEs is comparable to the communication overhead between PEs.

The board I/O requirement of interconnection networks as a function of scaling has been studied by the authors [6]. Figure 1 shows how the total number of I/O channels required by three interconnection topologies, the crossbar, the hypercube and the shuffle-exchange networks, scales with the total number of PEs in the architecture. It is assumed that a packet switching or message passing network is used [1, 2], with messages sent in parallel. We have also assumed that the switches of the interconnection networks are implemented electronically.

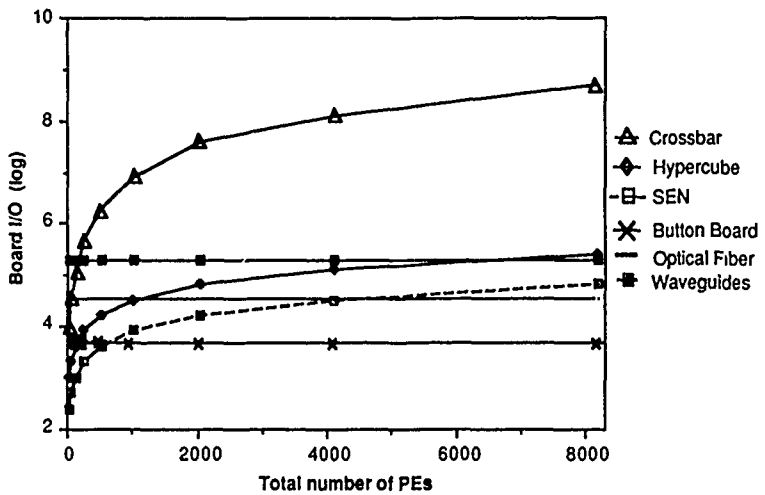


Figure 1. Board I/O versus number of PEs (message width =64 bits, number of boards =16, assumed board size: 15" x 18".)

Figure 1 also shows the I/O levels that could be supported by some high-density, board-level interconnect media based on current possible packing densities. These include state-of-the-art button boards [4], optical fibers, and polymer waveguides. Not surprisingly, large crossbars are infeasible. Of more significance is the fact that as the architecture scales up to 8K PEs, electronic button boards are grossly inadequate. Because of their much higher packing densities and bandwidth, optical interconnects hold much greater promise. Optics offers the possibility of alleviating the bottleneck associated with the interconnection in a parallel system in which the processing is performed electronically. While a fully parallel message transfer may not be possible in an optical hypercube connection for 8K PEs (Figure 1), an optical perfect shuffle connection of the same size could be supported by polymer waveguides. Our estimates indicate that when a single-stage shuffle-exchange network is used, button boards can support a network for only 400 PEs, optical fibers can support 3,000 PEs, and waveguides can support about 17,000 PEs.

The single-stage shuffle echange network makes efficient use of available interconnections when combined with the appropriate electronic processing. To enable a meaningful comparison to be made between the various implementations of the optical perfect shuffle, we consider a system consisting of a total of 1024 PEs on sixteen boards. For all boards except the first and last, all PEs communicate with PEs on other boards. In the worst case,

the communication is across eight boards. Parallel data transfer is considered, with 64 bits and therefore twice this number of unidirectional channels being associated with each PE. We assume that the boards of the system are separated by 2cm, and that to acheive the maximum possible density of interconnections, sources are used which emit light in a single spatial mode.

issues of concern in developing a practical system are reliability, bit error rate (which is affected by optical loss and crosstalk in the system) and the ease of mechanical assembly in situations where boards must be removed and re-inserted. These issues must be addressed whatever ointerconnection implementation is used.

Free-Space Network Implementations

Several implementations of the connectivity using free-space optics are possible. These include bulk and micro-optics, and holograms. [5] Holograms can be fabricated which serve the dual purpose of collimation and re-direction [7]. Ignoring aberrations present in real systems, the upper limit to interconnection density is determined by the diffraction limited spot-size.

One possible implementation consists of focussing all the sources from one board with high numerical aperture lenses, performing directional routing with a computer generated hologram. The highest density would be provided by arranging the sources in a square pattern. The signals would pass through transparent areas on each successive board until reaching a photodetector on the destination board as illustrated in figure 2(a). The highest required density is determined by the number of channels required in the vicinity of the two center boards, the number being one half of the total number of unidirectional channels, or 32K. This represents a square 181 x 181 channels. Assuming that the collimating lens has a numerical aperture of 0.5 and is given the same allocation of area as the detector array, consideration of data transfer across eight boards, indicates that the space occupied would be 1cm x 1cm Thus each channel is allocated an area 50um x 50um. In fact, the space allocated for each detector would be greater than this, since electrical connections must be made to the processing electronics. Using a multiple layer packaging technology, with a width of 25um, and ten layers would increase the area to be allocated to 4cm x 4cm. Assuming that the lateral deflection of the signal is 2cm in the worst case, it can be shown that the source wavelength must be maintained to within approximately 1nm. In addition, each of the 64K sources for the system must be fabricated with a wavelength within 1nm of the design wavelength. Each board must be aligned laterally to within 25um, with tolerances an order of magnitude tighter for the hologram alignment. An alternative implementation could use relay lenses between boards to reduce the diffraction limited spot size and relax some alignment tolerances The constraints on source wavelength however are unchanged.

Guided Wave Network Implementations

Fiber optics offers a relatively mature technology for interconnection. However, the diameter of fibers developed for telecommuniction systems renders them incompatible with the density required for board-edge connection, for which the technology is most developed. While special fibers could be developed, it would still not be possible for channels associated with different paths to occupy the same space. Thus the implementation would be bulky, in addition to being labour-intensive, and would scale poorly with increasing system complexity.

Polyimide waveguides fabricated on either rigid or flexible substrates [8] may be used to implement the required connectivity. This two-dimensional format dictates that the system consist of a set of boards interconnected using a backplane. The required planar packaging is compatible with established electronic manufacturing techniques and established device technology. The required routing may be performed on the boards, backplanes, or a combination of the two. Important parameters in the system are loss and crosstalk, which depend on the loss of the waveguides and components, and on the crosstalk of the waveguide crossovers. In this system, the worst number of crossovers sustained by a waveguide on the backplane is approximately half the number of channels in the complete system. An illustration of the system is shown in figure 2(b).

Preliminary results for the medium have already been presented [8] with propagation losses of 0.3dB/cm, losses of 0.4dB for right angle bends, and less than 0.0055dB for a right angle crossover. While these results suggest that the number of PEs which can be supported in a single-layer implementation is only 30, drastic improvements are to be expected from the use of several layers of polyimide in the backplane, or several independent flexible circuits. For

example, the use a number of layers less than the number of boards would eliminate crossovers, while optimization enables the number of layers to be reduced while still maintaining adequate optical performance. Another option, typically used in routing and layout of VLSI circuits and circuit boards, in reducing the total number of crossovers in the backplane would be in physically rearranging the PE layout in the board. This is quite feasible for the boards in the middle of the rack.

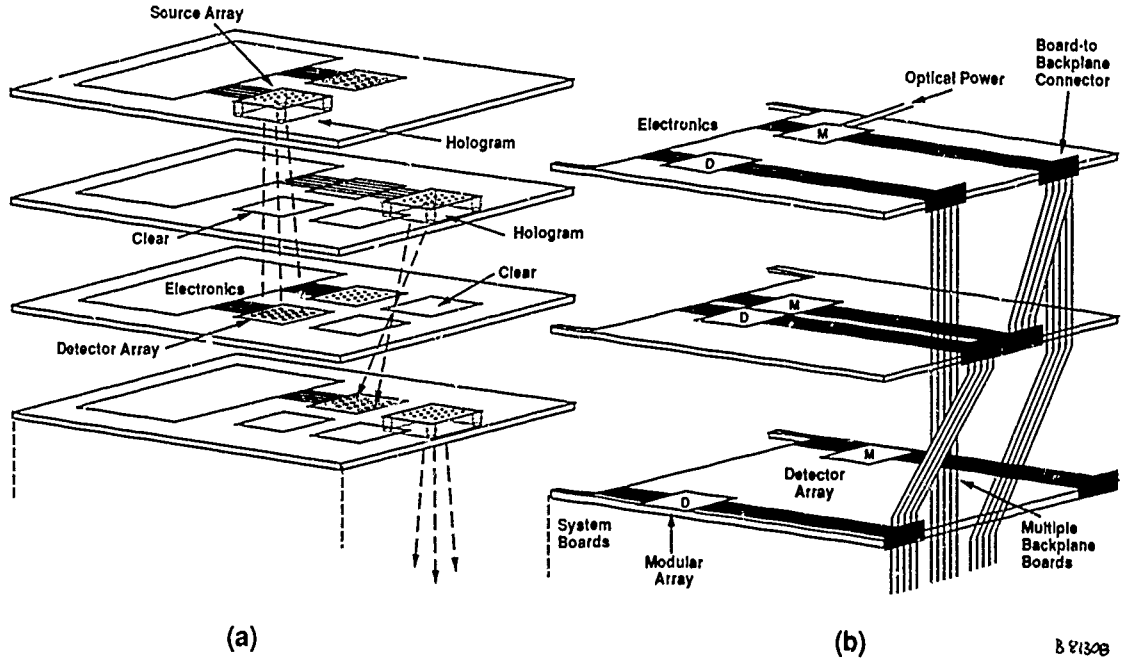


Figure 2. (a) Free-space and (b) guided wave interconnections for parallel, multi-board systems.

Practical multi-board systems require will require board-to-backplane connectors. While contacting connectors are feasible, giving alignment tolerances of the order of the waveguide dimensions, a more practical solution is found in the use of gradient index lenses. Properties of commercially available lenses indicate that one lens could be used with 128 channels with less than -40dB crosstalk. In a connector based on such a lens, angular alignments of a few milliradians are required. The waveguides in the source and image planes must be accurately located, while tolerances of tens of microns are permitted for location of the planes with respect to the lens. At the joint at which the boards would be demounted (at the right-angle prism), the tolerances are approximately 1mm. making the technique suitable for systems consisting of demountable boards. Multilayer polyimide technology would reduce the number required of connectors required to less than ten. The required width of the backplane would be 30cm.

Optoelectronic Interfaces

Independent of the choice of interconnection medium, optical power must be provided for each channel. Choices are LEDs and lasers. In both cases, one may have one channel associated with each source, or divide a given source between a number of channels. LEDs typically have large spatial extent, and are therefore incompatible with the aims of maximising the interconnect density. Lasers offer higher output powers and smaller spatial extent of the source. Consideration of the reliability of typical lasers at room temperature indicates that on average one of the 65536 sources would fail after six hours, if the system were operating at 50°C. Since failure of one channel is potentially as serious as failure of a much larger number, the system would be greatly improved by the use of a smaller number of lasers in controlled, remote environments feeding an array of external modulators. Redundancy of two would extend the mean time to first failure to 2.5 years. Operation of the lasers in a remote, environment at lower temperature would realize further improvements. The allowable fan-out will be determined by the performance of available receivers and the loss of the interconnection network.

To be immune to variations in operating environment, the modulators would be based on the electrooptic effect. Polarization based waveguide modulators with integral polarization filters offer the possibility of ease of fabrication, compatibility with standard planar processes, normally-off operation, high extinction ratios and low drive voltages. Logic compatible devices with multi GHz response have been reported in the literature.

Critical to the demonstration of a high-density interconnection medium is the development of high-density receiver arrays with low power dissipation compatible with standard packaging techniques.

Conclusions

The connectivity requirements of a massively parallel architecture have been examined. Optics enables a critical interconnection network bottleneck to be overcome. Free-space interconnects have potentially high interconection densities, but suffer from stringent source parameter requirements. A system consisting of several layers of polymer waveguides with micro-optical board to backplane connectors offers performance suitable for 1024 processors connected with a single-stage shuffle exchange network. For all parallel interconnection networks, electrooptic modulators in combination with external lasers appear to be the most attractive choice.

This research was supported by the Air Force Office of Scientific Research and the Advanced Research Projects Agency of the Department of Defense under Contract No. F49620-86-C-0082., and by the Defense Sciences Office at the Defense Advanced Research Projects Agency under contract number N66001-87-C-0205.

References

[1] W.D. Hillis, 'The Connection Machine,' MIT Press, 1985

[2] J. F. Palmer, 'The NCUBE family of parallel supercomputers,' Proceedings of the IEEE International Conference on Computer Design, 1986.

[3] C. L. Seitz, 'The Cosmic Cube,' Communications of the ACM, Vol. 28, No. 1, Jan. 1985, pp. 22 - 33.

[4] R. Smolley, 'Button Board: A New Technology Interconnect for 2 and 3 Dimensional Packaging,' Proceedings of the 1985 International Symposium on Microelectronics, pp. 326 - 333.

[5] A. R. Johnston, L.A. Bergman, and W.H. Wu, 'Optical Interconnection Techniques for Hypercube,' 1988.

[6] A. Guha and J. Bristow, Second Annual Technical Report for the Optical Symbolic Processor for Expert System Execution (AFOSR and DARPA), Honeywell Inc, 1988.

[7] K..H. Brenner, F. Sauer, "Diffractive-reflective optical interconnects", Applied Optics, Vol 27, No. 20, 1899 pp4251-4254

[8] C.T. Sullivan "Optical Waveguide Circuits for Printed Wire-board Interconnections", Presented at O-E/FIBER LASE, Boston, 6 September 1988.

NOTES

TUESDAY, FEBRUARY 28, 1989

SALON F

9:00 AM–10:00 AM

TuB1–TuB4

OPTICAL INTERCONNECTIONS: 2

John F. Walkup, Texas Tech University, *Presider*

Implementation of Dynamic Holographic Interconnects with Variable Weights in Photorefractive Crystals

A. Marrakchi and J. S. Patel

Bellcore

331 Newman Springs Road, Red Bank, NJ 07701-7020

Elementary holographic gratings can be used to implement interconnection links betweeen individual processing elements of two distinct planes in a multi-layered optical neural network. In such networks, one issue that has a direct impact on the development of learning machines is the capability of continuously modifying a given interconnection strength (or weight) without affecting the others, when the gratings share the same volume in the photorefractive crystal (i.e., frequency-multiplexed gratings). In the following, we extend the principle of coherent erasure by the double-exposure technique to the case of elementary gratings that implement real-time optical interconnections in photorefractive materials. The effect of continuously varying the phase shift between the two recorded gratings on the diffraction efficiency is quantified, and shown to be applicable to the simulation of synapses with programmable variable weights, as would be required in a learning neural network. Issues that relate to fan-in and fan-out capabilities, which ultimately determine the achievable level of parallelism and cascadability in such processing architecture, are also addressed. An experimental interconnection system based on two-dimensional liquid crystal phase and amplitude modulators and photorefractive recording is described. Finally, an extension of the double-exposure technique to time-average erasure is then discussed.

The proposed holographic interconnect system is shown in Fig. 1. The purpose is to connect a matrix of sources in plane P_{IN} with a matrix of detectors in plane P_{OUT}, with fan-in and fan-out capability. In order to write the respective gratings that will form the optical links, matrices of mutually coherent control sources are needed in the planes P_{C1} and P_{C2}. All the beams from P_{C1} represent the "object" wave in the traditional sense of holography, and each beam from P_{C2} represents the "reference" wave. Hence, in this coherent system, each interconnection set, defined by a specific configuration of P_{C1} and one reference beam from P_{C2}, has to be recorded separately from all the others in order to generate multiple optical links without appreciable crosstalk.

The double-exposure technique is a two-stage process in which a phase shift is induced on one of the control beams between recordings. The conventional ways of inducing this phase shift are either to electrooptically phase modulate the beam, or reflect it off a piezoelectrically driven mirror. In some of our experiments, we use a mirror mounted on a stack of piezoelectric ceramics, and in others a liquid crystal phase modulator. In the scheme proposed in Fig. 1, a matrix of such modulators would be placed in the control plane P_{C1}. In the "reference" arm P_{C2}, a matrix of amplitude modulators, such as ferroelectric liquid crystal gates, is used to control each set of interconnections. The recording medium is a single crystal of photorefractive bismuth silicon oxide ($Bi_{12}SiO_{20}$, or BSO). Holographic gratings are formed in the bulk of this material by interfering optical beams originating from an argon laser operated at a wavelength of 514 nm. To monitor the space-charge formation in the crystal, the composite grating is read out in real-time with a He-Ne laser incident at the Bragg angle, although for phase-matching over a wider spatial bandwidth it would be preferable to use the same wavelength as for writing.

The expected $\cos^2(\Phi)$ behavior of the diffracted intensity is illustrated in Fig. 2. Here, the normalized diffraction efficiency after erasure is plotted as a function of the phase shift between the two gratings. A variable and controllable interconnection strength is thus possible with this technique.

In a practical interconnection scheme, many gratings will share the same volume in the photorefractive crystal. In one experiment illustrating fan-out, two gratings with an angular separation of 4 mrad are recorded in the BSO crystal. Since the Bragg selectivity for the He-Ne beam is not critical with this small angular separation, two waves are diffracted. During recording, one of the writing beams is phase shifted while the efficiency is continuously monitored. The result in Fig. 3 shows oscilloscope traces of the diffracted intensity in each beam and the relative phase shift. In this particular experiment, erasure of one grating does not affect the other as would be necessary for a dynamically programmable system.

The proposed technique of coherent erasure has a response time that is practically independent of the phase between the two recorded gratings. Nevertheless, this system requires synchronization in order to stop the recording of the second grating when the efficiency reaches its minimum. To alleviate this problem, we extended the double-exposure technique to the case of time-average exposure. In such a configuration, the diffraction efficiency is a function of the average phase, and is stationary as long as there are no phase variations.

In summary, it is shown that the double-exposure technique with a variable phase shift between the two recorded gratings yields a continuously graded interconnection strength between two spatially separated planes. The non-linear relationship between the weight of this optical link and the phase shift is described by a $\cos^2(\Phi)$ function, as experimentally verified. When several holographic gratings are recorded to simulate fan-in and fan-out, it is possible to modify one interconnection weight without significantly affecting the others. Combined with the large storage capacity available with holographic recording, this double-exposure technique could be suitable for the optical implementation of learning neural networks with continuously variable weights. Extension to time-average holographic erasure simplifies the proposed interconnection system.

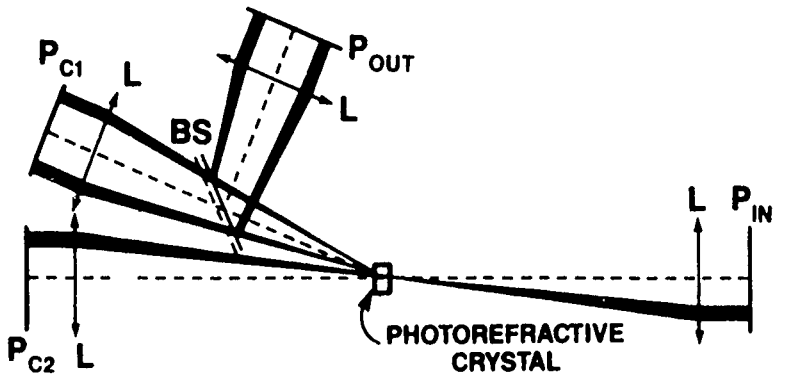


Fig. 1 Schematic of a holographic interconnect system.

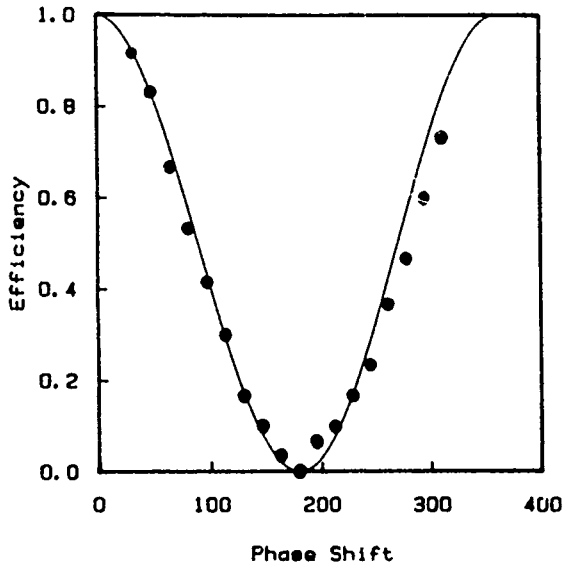


Fig. 2 Normalized diffraction efficiency as a function of the phase shift in units of degrees between the two gratings recorded with a double-exposure technique.

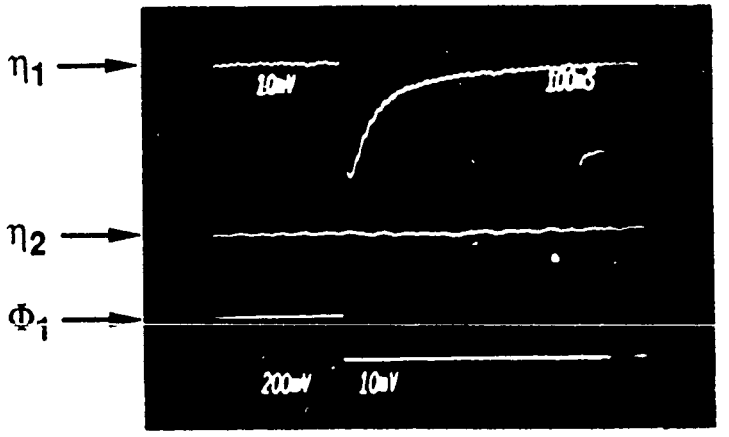


Fig. 3 Oscilloscope traces of the diffracted intensity in a fan-out situation and of the phase shift of one of the writing beams.

Energy Efficiency of Optical Interconnection Using Photorefractive Dynamic Holograms

Arthur Chiou and Pochi Yeh
Rockwell International Science Center
1049 Camino Dos Rios
Thousand Oaks, CA 91360

SUMMARY

Reconfigurable optical interconnection [1] linking laser arrays and detector arrays plays a key role in optical computing. A generalized crossbar switch [2] allowing arbitrary interconnection, including many-to-one and one-to-many (broadcasting), is the most desirable type of interconnection network for parallel processing. Such a switch can be implemented using optical matrix-vector inner product architecture [3] where a spatial light modulator (SLM) can be used as a binary matrix mask for configuring the interconnection pattern. For one-to-one interconnection (permutation) of a linear array of N-sources to a linear array of N-detectors (i.e., a normal crossbar), the upper limit of the energy efficiency of such an architecture is 1/N due to its fanout nature. Recently, we have proposed and demonstrated [4,5] that photorefractive dynamic holograms can be incorporated into this architecture to significantly improve the energy efficiency. In this paper, we report experimental results on the energy efficiency of such a reconfigurable interconnection using a $BaTiO_3$ crystal.

Referring to Fig. 1, we consider a scheme of two-wave mixing for the study of photorefractive energy transfer. An input optical beam (of power P_i) is split by a beam splitter into a pump beam and a signal beam which interact inside a photorefractive crystal. The crystal is oriented so that direction of energy transfer due to photorefractive two-beam coupling is from the pump to the signal beam. Let P_o be the optical power of the amplified signal beam. The energy efficiency (η) is defined as the ratio of the optical power of the amplified signal beam to that of the input beam (i.e., $\eta = P_o/P_i$). It is a function of the beam splitting ratio R, the photorefractive exponential gain ΓL (where Γ is the exponential gain constant and L is the interaction length), the beam overlap of the pump and the signal, and all the factors contributing to energy loss such as Fresnel reflection, absorption, and scattering. For a given photorefractive crystal, the optimal geometry and beam splitting ratio R that maximize the energy efficiency can be determined empirically. Using various samples of barium titanate crystals (of size ~ 5 mm × 5 mm × 5 mm, uncoated), we have achieved a maximum energy efficiency of 30%. When a neutral density (ND) filter is inserted into the signal arm, energy efficiency (η) decreases. The dependence of η on the transmittance (T) of the ND filter is investigated. For T = 0.1%, η as high as 10% can still be achieved.

For optical interconnection applications, the signal beam is expanded through a binary matrix mask (of dimension NxN) to carry the interconnection pattern. Depending on the experimental configuration, such a mask can be used to realize a 1 to NxN interconnection or a NxN crossbar switch. To achieve maximum energy efficiency, we need to match the beam profile spatially at the photorefractive crystal. Since both the signal beam (which carries the interconnection pattern) and the pump beam consist of array of beamlets of identical shape, Fourrier plane is an ideal place for maximum overlap. Near perfect overlap at the Fourrier plane is a result of the shift-invariance property of Fourier transformation. For 1 to NxN interconnection, spherical lenses are used to Fourrier transform the input spatial patterns of the signal mask and the "matched" pump mask (see the first row in Table I). For the NxN crossbar switch, we use astigmatic optics which image along the horizontal direction and Fourrier transform

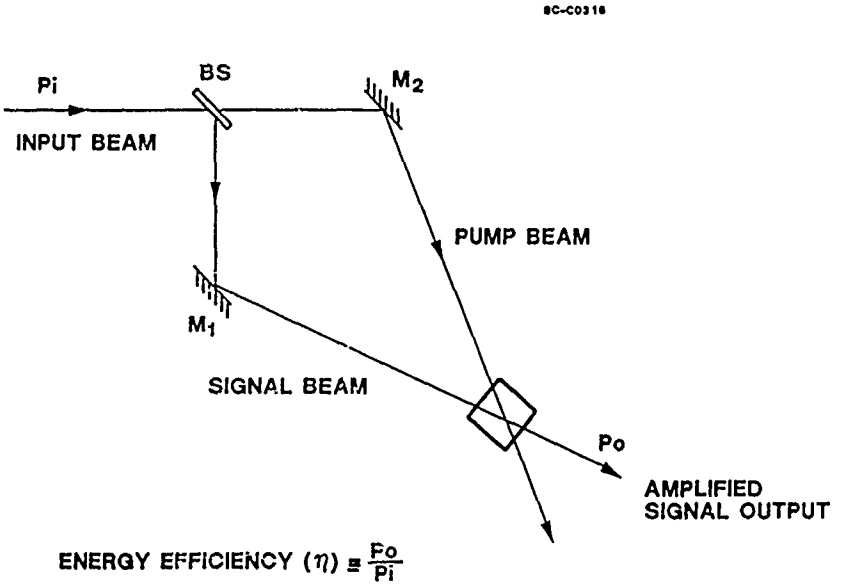


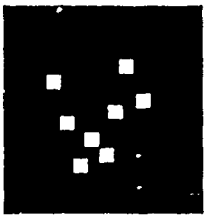
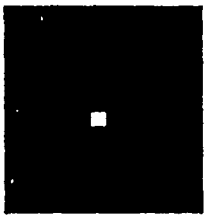
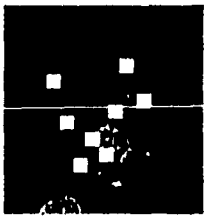
Fig. 1 The definition of energy efficiency of photorefractive dynamic holograms.

along the vertical direction of the input masks. This ensures that the i-th component of the pump beam array interact only with the part of signal beam that passes through the i-th column of the interconnection mask The appropriate mask for the pump is shown in the second row in Table I.

Table I

Comparison of the Configurations for 1 to NxN Interconnection and NxN Crossbar Switch (for N = 8)

SC-C0317

INTERCONNECTION SCHEME	MASK FOR SIGNAL BEAM	MASK FOR PUMP BEAM	OPTICS
1 TO NxN (a)			FOURRIER TRANSFORM LENS (SPHERICAL)
NxN CROSSBAR			ASTIGMATIC OPTICS IMAGING IN X-DIRECTION & FOURIER TRANSFORM IN Y-DIRECTION

(a) For the mask shown in the table, 8 out of the 8x8 channels are "on."

Using a 30°-cut $BaTiO_3$ crystal, we have measured the energy efficiency for the two interconnection schemes described above with N = 4, 8 and 16. The experimental configuration is shown in Fig. 2. An output beam from an argon laser (514.5 nm) is collimated by a lens (F.L. = 2 m). A variable beam splitter consisting of a half-wave plate and a polarizing beam splitter cube is used to vary the intensity ratio of the pump and the signal beams. The polarization of the reflected (signal) beam is rotated by 90° into the horizontal direction by another half-wave plate. After passing through a polarizer (to filter out the residual orthogonal polarization component), the signal beam is expanded to illuminate the interconnection mask. In the other arm, the transmitted (pump) beam illuminates a "matched" aperture mask (see Table I). Two spherical lenses, one in each arm, are used to Fourrier transform the two spatial patterns on to the crystal plane. A spherical lens is used to re-image the spatial pattern carried by the amplified signal beam on to the detector plane where the optical power in each channel is measured. For the NxN crossbar switch, the beam expander is placed at the upstream of the variable beam splitter so that both the pump and the signal beams are expanded through the same beam expander. The spherical lenses are also replaced by appropriate astigmatic optics as listed in Table I.

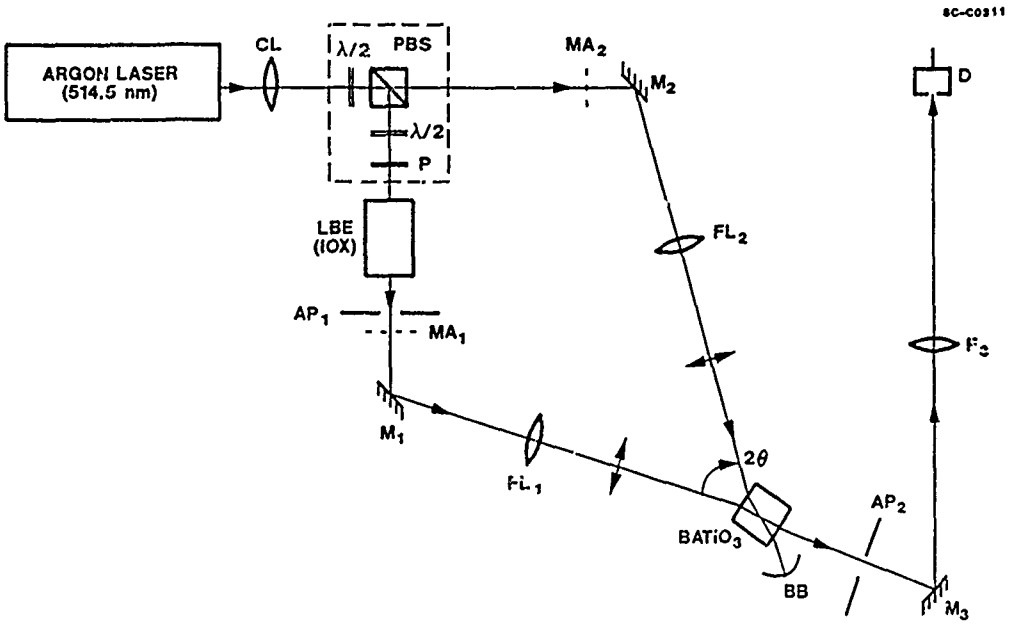


Fig. 2 Experimental configuration for the determination of energy efficiency of a photorefractive optical interconnection.

In summary, we have measured the energy efficiency of an 1 to NxN interconnection and an NxN crossbar switch using a 30°-cut $BaTiO_3$ crystal. The experimental results on non-uniform energy distribution among different channels, crosstalk, and the dependence of energy efficiency on N are presented and discussed.

ACKNOWLEDGEMENTS

This work is supported by DARPA/AFOSR under contract No. F49620-87-C-0015. We gratefully thank John Hong for helpful technical discussions.

REFERENCES

[1] See, for example, J.A. Neff, Ed., Optical Computing, Proc. SPIE **625**, 109 (1986).
[2] A.A. Sawchuk, B.K. Jenkins, C.S. Raghavendra, and A. Varma, "Optical crossbar networks," IEEE. Computer **20**, 50 (1987).
[3] R.A. Athale, "Optical matrix processors," in Optical and Hybrid Computing, SPIE **634**, 96 (1986).
[4] P. Yeh, A. Chiou, and J. Hong, "Optical interconnection using photorefractive dynamic holograms," Appl. Opt. **27**, 2093 (1988).
[5] A. Chiou, P. Yeh, and J. Hong, "Energy efficient optical interconnection using dynamic holograms in photorefractive media," paper FX1, OSA Annual Meeting/ OPTCON'88, Santa Clara, California (1988).

A Free Space Optical Interconnection Scheme

Alex Dickinson and Michael E. Prise
AT&T Bell Laboratories
Room 4E-514
Crawfords Corner Road
Holmdel, NJ 07733

Summary

Introduction. One of the limiting factors in the design of large scale digital systems is communication [1]. The use of conventional wiring for high speed communication entails high energy dissipation, crosstalk, ground loops and low interconnect density due to physical size constraints. Miller has shown [2] that optical interconnections have fundamentally lower energy requirements than electronic communication owing to the closer matching of impedances over all but the shortest inter-device distances, provided appropriate optoelectronic integration technology exists.

Free space interconnect technology involves using large arrays of optical beams imaged onto arrays of switching devices with lenses. These techniques provide a high degree of connectivity without a great deal of system complexity, and are being used for optical computing applications [3]. Integrated arrays of Self Electro-optic Effect Devices (SEEDs) have been constructed for this purpose [4].

In this paper we describe devices, circuits and optics for the optical interconnection of electronic subsystems (chips, wafers or boards). With this approach optical devices are used to provide dense high bandwidth communication between subsystems, thus alleviating conventional electronic communication problems. We also describe a particular interconnect architecture in order to illustrate the use of the components in a system context.

Input and Output Ports. An optical link requires a modulated light source (the equivalent of an electronic output pad on a chip) and an optical detector (the equivalent of an input pad). The output port is connected to an electrical signal that serves to modulate the beam. The beam carries the signal to an input port that detects the optical signal and converts it to an electrical equivalent. Devices for this application must meet several basic constraints:

1. The modulator should have an input capacitance of the same order as a typical transistor so that the energy advantages of the scheme over conventional wiring may be realized.
2. The detector should be capable of efficiently converting optical signals to electrical signals at digital logic levels. Inefficient detection would result in an energy/speed loss.
3. It should be possible to fabricate both classes of device in integrated arrays in order to achieve high interconnect density.

The following sections describe output and input devices that meet these criteria. The output pads are based on GaAs multpile quantum well(MQW) technology and the input pads on silicon technology. Currently this restricts the pads to fabrication on different substrates, thus restricting the techniques to board-level/multi-chip systems. The feasibility of fabricating appropraite GaAs devices on a Si substrate has however been demonstrated [5] [6], and will hopefully lead to integration of optical input and output ports on single chips, thus demonstrating an important extension of the SEED concept.

Output Ports. The output ports in this scheme are GaAs MQW modulators. These serve to modulate incoming 'power supply' beams generated by imaging a laser onto a Dammann [7] grating. To avoid referencing problems we use pairs of modulators to provide differential communication links. These can be driven in two configurations, either (a) by driving each modulator individually with complementary voltages or (b) by connecting the modulators in series and driving the voltage of the central node.

In either case if the modulators are illuminated with equal intensity light beams the output from the port (consisting of the two modulators) will consist of two spots each with a different intensity. Lentine [8] has demonstrated the second configuration using a single symmetric SEED device. He used a voltage of $V_{dd} = 15V$. This rather high voltage can be reduced significantly by using a reflecting substrate [9]. Using the reflective substrates also makes the processing easier.

Input Ports. Detection of the optical signals is an inherently less complex task than modulation: a simple reverse biased PIN diode acts as an admirable optical detector and may be fabricated in a standard silicon processes (CMOS, for instance). However in order to maximize the speed of the link and minimize optical energy requirements it is useful to introduce electrical gain into the detection process. The circuit shown in Figure 2(a) serves this purpose. The two spots representing the differentially encoded incoming optical signal are imaged onto two PIN diodes, $D_1 - D_2$. These are biased by devices Q_1 and Q_2 in order to produce voltages V_1 and V_2 related to the amount of optical power incident on each. The differential voltage $\Delta V = V_1 - V_2$ is amplified by the differential amplifier ($Q_3 - Q_7$) and the output fed to the inverter ($Q_8 - Q_9$) for restoration to CMOS digital levels.

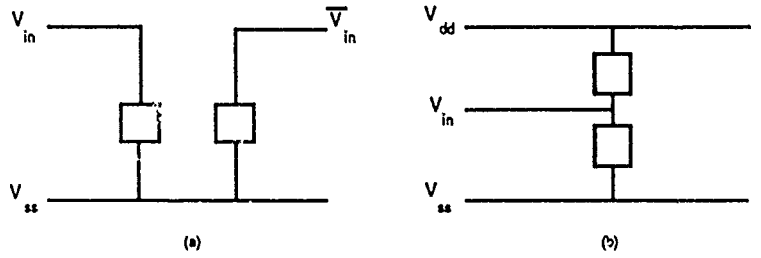


Figure 1: Electrical circuits for SEED devices configured for use as output ports.

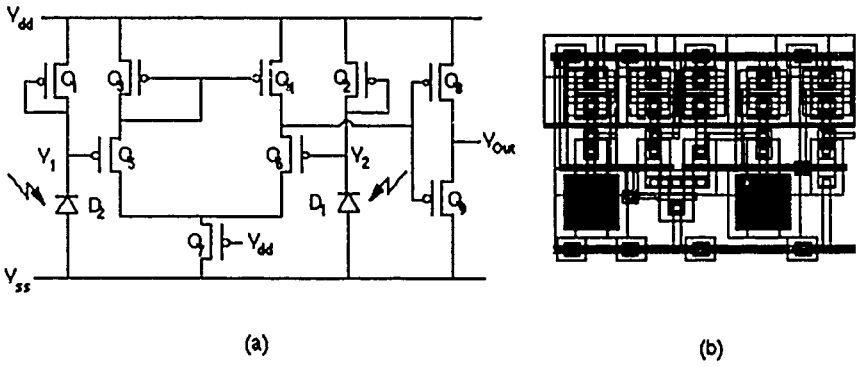


Figure 2: Differential amplifier based optical detection circuit (a) and a CMOS layout (b) with shaded regions indicating the photodiodes.

A number of variations of this circuit have been designed and are being fabricated in a $0.9\mu m$ CMOS process. The layout of the circuit shown in Figure 2(b) has been simulated at a rate of 10^8 bits/sec with an optical incident power of $20\mu W$ on each diode.

Optical Systems. One of the simplest interconnections we can implement is between two modules B1 and B2 one with an array of SEED modulators (OP1) as the outputs and another with an array of Silicon photodetectors (IP1) as the inputs. This is shown in Figure 3(a).

The necessary optical components are a polarization beam splitter (PB1), an array generator such as a Dammann grating (D1), a diode laser (LD1), a quarter wave plate LQ1, lenses (L1,L2 and L3) and a mirror M1. The light emitted from the laser diode is collimated at lens L1 and split into an array of equal beams by the Dammann grating. The collimated array of beams then passes through the polarization beam splitter PB1 and the waveplate W1. It is then focussed onto the array of SEED output ports OP1. The reflected signals which are the outputs are reflected off the polarization beam splitter (PB1) and reflected off the mirror (M1) and focussed by the lens L3 onto the Silicon input ports IP2. The system can be extended with another identical optical system.

This scheme is the easiest for us to implement since we only require discrete arrays of silicon input ports and GaAs output ports and will be our first working system. However to obtain a significant advantage over existing electronic systems it will be necessary to more fully integrate the optical input and output ports on one substrate. This interconnection is unidirectional and point to point. In the diagram to the left of the optical setup we have illustrated the functionality of the system.

The second system we show in Figure 3(b) describes a more flexible interconnection system that allows for *bidirectional* point to point interconnections. This requires that both the optical input and output ports be within the field of the input lens (L2 for B0). In order to minimize the optical system performance requirements both input and output ports should be located on the same substrate, as is also preferable when electronic system issues are considered.

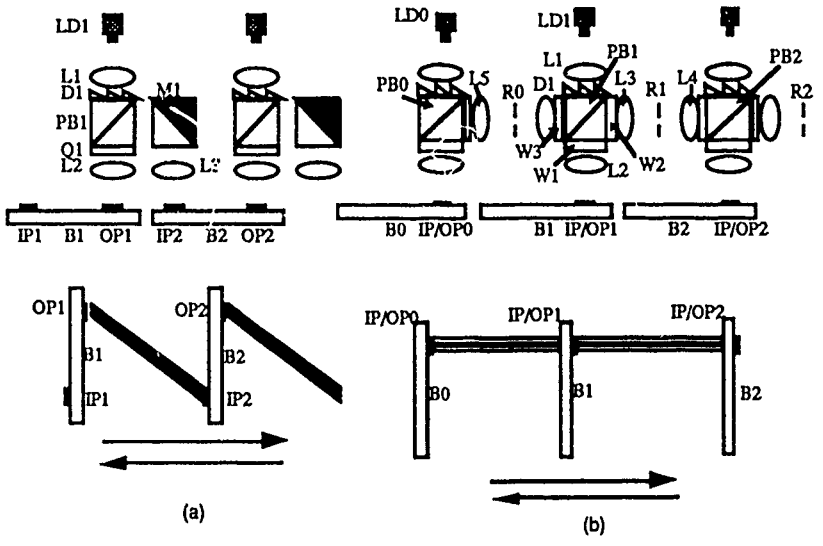


Figure 3: Inter-module, point-to-point, unidirectional and bidirectional interconnections.

This interconnection uses a very similar optical system to that described in [3] with space variant mirrors (R0, R1 and R2). This time the electronic modules B0, B1 and B2 (which could be boards, wafers or chips) have closely spaced optical input and output ports. Consider communications between B1 and the adjacent modules B0 and B2. Once again the laser diode LD1 is split by the array generator D1 and used to illuminate the output ports. We would initially have these ports arranged in a rectangular array as constrained by the Dammann grating method. The input ports can either be adjacent too, or interspersed with, the output ports. The reflected beams which contain the output information from the chip are collimated at lens L2, reflected at the beam splitter PB1 and then focussed down onto the reflector array R1. The reflector array can be used to configure the interconnect. By appropriately patterning the reflector we can reflect or transmit any of the signal we wish. The reflected signals go back through W2, PB1 and W3. They are then focussed down in the plane of reflector array R0. This reflector array is arranged so it is transparent at the appropriate points. The signals are then collimated by L1 and because their polarization has been rotated by half a wave, these signals are reflected at PB0 and then focussed down onto the appropriate input ports on B0.

Meanwhile the signals which have been transmitted at R1 are transmitted through PB2 and focussed down in the plane of R2. The patterned reflectors on R2 are arranged so that they are then reflected back onto PB2 where they are reflected down through the input lens and focussed onto the appropriately arranged input ports on B2. The arrangement can be continued indefinitely by simply replicating the optics.

The functionality of the system is indicated by the lower part of Figure 3(b).

An Optical Bus Architecture. In this section an architecture for the interconnection of N nodes (subsystems: chips, boards or wafers) will be described. The optical and electronic technology described above places two primary constraints on the design of an architecture:

1. The arrays of devices should all be identical to simplify design and fabrication.
2. The interconnect should be made up of point-to-point links (no fan-out).

The basic structure of this optical bus is illustrated for the four node case in Figure 4. The objective is to connect each of the N nodes to every other node. At each node an array of devices is used to generate, transmit, and receive optical signals. In the figure each array is represented in an abstract transparent form: in reality the arrays would be comprised of devices fabricated on a reflective substrate. The array is triangular and comprises the following component elements for an N element system:

1. A horizontal row of $N-1$ transmitting elements. Each of these is used to generate a modulated light beam that provides a connection to another node.
2. A diagonal row of $N-1$ receiving elements. Each of these is used to receive a modulated light beam carrying data from another node.
3. A pattern of $(N-1)^2/2 - 2(N-1)$ passive elements that allow the remainder of the beams to pass unhindered. Each of these beams is passing data between nodes not associated with this one.

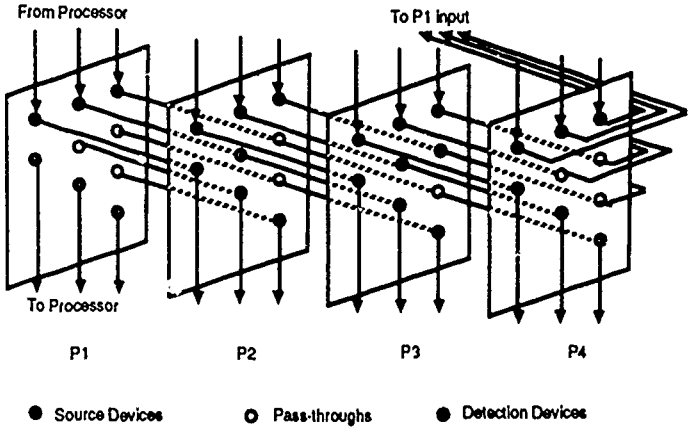


Figure 4: The optical interconnect for four nodes, P1 to P4.

The utility of the scheme is based around the observation that a simple *shift* of the beams 'downward' between each neighboring node is sufficient to route all of the beams to their correct destinations. The $N-1$ beams comprising the diagonal of the incoming bus are detected and represent the signals sent to the node from the $N-1$ other nodes. All other beams are simply shifted down one place (a physical dislocation will clearly produce this effect) and an additional $N-1$ beams are added in as the new top row — these are generated by modulating $N-1$ incident power supply beams with $N-1$ electrical signals from the *ith* node destined for each of the $N-1$ other nodes.

This architecture meets the requirements of regularity and fan-out, and additionally provides node-to-node routing by a simple relative shift of the arrays.

Conclusions. We have described an approach to optical interconnect based on the use of arrays of beams and integrated optical devices. With currently available technology it will be possible to use the scheme for board level interconnect, each board containing one CMOS array of input ports and one GaAs array of output ports. We are currently fabricating an array of CMOS differential detectors that we intend to use in conjunction with an array of GaAs SEED devices to form an simple point-to-point link. In the near future we intend to take advantage of the merging of the two fabrication technologies to provide chip-to-chip interconnection based on the mechanisms and architectures we have described.

References

[1] R. W. Keyes, "Fundamental limits in digital information processing", *Proc. IEEE*,69,(2),269-279, 1981

[2] D. A. B. Miller, "Optics for low-energy communications inside digital processors: quantum detectors, sources and modulators as efficient impedence convertors", *Optics Letters* (Jan. 1989)

[3] M. E. Prise, M. M. Downs, F. B. McCormick, S. J. Walker and N. Streibl, "Design of an Optical Digital Computer", in *Optical Bistability IV* edited by W. J. Firth et al., published by Les Editions de Physique 1988

[4] D. A. B. Miller, J. E. Henry, A. C. Gossard and J. H. English, " Integrated Quantum Well Self Electro-Optic Effect Device: 2 x 2Array of Optically Bistable Switches", *Applied Physics Letters*,49, 821-823, 1986

[5] W. Dobblelaere, D. Huang, M. S. Ünlü and H. Morkoç, "AlGaAs/GaAs multiple quantum well reflection modulators grown on Si substrates", *Applied Physics Letters*,55,94-96, 1988

[6] K. W. Goossen, G. D. Boyd, J. E. Cunningham, W. Y. Jan, D. A. B. Miller and D. S. Chemla (to be published)

[7] H. Dammann and K. Gortler, "High Efficiency in-line multiple imaging by means of multiple phase holograms", *Optics Communications*,3,312-315,1971

[8] A. Lentine et al. ," Photonic ring counter and differential logic gate using the symmetric self electro-optic effect device", *Conference on Lasers and Electro-optics 1988*, paper TUE4

[9] G. Boyd, D. A. B. Miller, D. S. Chemla, S. L. McCall, A.C Gossard and J. H. English, "Multiple quantum well reflection modulator", *Applied Physics Letters*,50,119, 1987

MICROOPTIC SYSTEMS: ESSENTIAL FOR OPTICAL COMPUTING

J.L. Jewell and S.L. McCall*

AT&T Bell Laboratories, Rm. 4G-520, Holmdel, NJ 07733 U.S.A.
**AT&T Bell Laboratories, Murray Hill, NJ 07974 U.S.A.*

INTRODUCTION

The history of information processing has seen a continuing trend toward ever-smaller devices and systems. Optics, well-suited for long distance communication, is currently the focus of many efforts to revolutionize computer technology, however many obstacles need to be dealt with realistically in order for it to have a chance of succeeding. Heat dissipation has long been thought to be a "fundamental" roadblock to optical computing. For any given architecture, minimization of propagation delay times is essential in the design of fast computers[1]. Architectural flexibility and manufacturability are additional issues which could potentially prevent optical computing systems from entering the marketplace. To overcome all of these hurdles it is necessary to develop microoptical systems far smaller[2] than those generally envisioned in the literature, either in the form of integrated optics[1] or microlenses and arrays. Here we address the latter approach and review some technological progress.

ARRAY SIZE SCALING

An array size of 1000×1000 is often considered a goal to be sought. Setting aside its glamourous appeal, let us compare the capabilities of a-small-number-of-large vs. a-large-number-of-small arrays by array size scaling. In this analysis we assume the devices in the array and their area density have already been optimized and therefore are unchanged; we only scale the size of the array itself. We take an initial N×N array and scale it down by a factor, s, in both dimensions, to N/s×N/s. Of course the number of devices/array is reduced to N^2/s^2, the power/array is reduced by s^2, and the intensity remains the same. It is also obvious that the lens diameters are reduced by s, but less well known is the fact that with this reduction the inherent lens aberrations are also reduced by a factor s. This is very significant because it means that the required lens performance level, e.g. diffraction-limited, can be achieved with a simpler system having fewer lens elements. Propagation delays (latency) must be minimized in a general purpose computing system. In our scaling, the latency will be reduced by s simply by the direct system scaling, and reduced even further by the fact that fewer lens elements need to be traversed. We can calculate the latency for a 32×32 array of GaAs microresonator[3] devices with 1-μm center-center spacing (1024 devices in a 32-μm square) and a $\sim$300-μm diameter lens. For the lossless crossover interconnection[4] and a reasonable lens design, the latency is about 30 ps. Since the array-based architectures are synchronous, the system clock cycle can be 30 ps or an integral fraction or multiple thereof. Banyan interconnections would be accomplished with 15 ps latency. It is certainly possible and it may be desireable to reduce the array size and latency still further. The latency and especially the clock cycle times compare very favorably with those projected for electronic technology in the next 10-15 years.

Thermal dissipation, often considered to have a "fundamental" upper limit of 100 W/cm^2 also improves as array size is scaled down, indicating that W/cm^2 is not a very fundamental unit for heat dissipation. The temperature rise accompanying a conductive dissipation of I W/cm^2 from an area A, conducting through a solid angle Ω over length L, can be compared to the rise in a scaled system of quantities I, A/s^2, Ω, and L/s. The rise ΔT becomes $\Delta T/s$. Experimentally we have pumped 1-2 μm devices at >100 kW/cm^2. They do not vaporize and even keep on working. The scaling behavior, $\Delta T \rightarrow \Delta T/s$, is also true for a high-performance convective cooling geometry[6]. Quantitative results of array scaling are summarized in Table 1.

Devices per Array	N^2	$\rightarrow$	N^2/s^2
Number of Arrays	M	$\rightarrow$	s^2M
System Volume	V	$\rightarrow$	$\sim V/s$
Power per Array	P	$\rightarrow$	P/s^2
Intensity	I	$\rightarrow$	I
Lens Diameter	D	$\rightarrow$	D/s
Lens Aberrations	A	$\rightarrow$	A/s
Latency	τ	$\rightarrow$	$<\tau/s$
Temperature Rise	ΔT	$\rightarrow$	$\Delta T/s$
System Flexibility		$\rightarrow$	more flexible
Manufacturability		$\rightarrow$	probably better

Table 1 - Approximate scaling behavior of various quantities under the assumptions given in the text.

Other relevant issues such as manufacturability and system flexibility do not lend themselves to quantitative comparisons, but we can attempt to judge them qualitatively. The system architecture is certainly more flexible in the scaled-down case. For example, if $s = 2$ we use 4 arrays for each array in the original case. In interconnecting them we are free to perform the same interconnection for all of them (which is equivalent to the original case), or to perform different interconnections for any or all of them. Thus the scaled-down design has increased flexibility. Manufacturability is more difficult to assess. Each array, packaged inside lens and cooling systems, is easier to manufacture in the scaled-down case, but there are more of them to assemble into the overall system. We expect that the manufacturing possibilities opened up by using small systems (e.g. making the lenses by non-labor-intensive techniques such as photo-electrochemical etching) will dominate over this issue, favoring smaller arrays. Assembly considerations will keep the optimum array size probably larger than 1×1.

TECHNOLOGICAL PROGRESS

Realization of microoptic systems will depend on nontraditional manufacturing techniques. To meet the requirements of small devices and the issues discussed above we propose the following

building block: a 32×32 array of InGaAs strained quantum well devices (~1-μm operating wavelength) on a transparent GaAs substrate. The substrate would have a 300-400-μm diameter lenslet formed by photo-assisted etching on its backside. Another lenslet element of GaAs can correct aberrations to extremely low values even with all surfaces spherical. A more attractive design utilizes a *single aspheric* surface on the substrate backside to accomplish the focusing still with low aberration. The lenslets, beamsplitters, waveplates, etc. would all be held in a solid piece by a transparent epoxy (Fig. 1). Importantly, as Fig. 1 shows, multiple arrays/lens elements can be left intact on a substrate, reducing the total number of pieces requiring assembly. All components must be small and cheap.

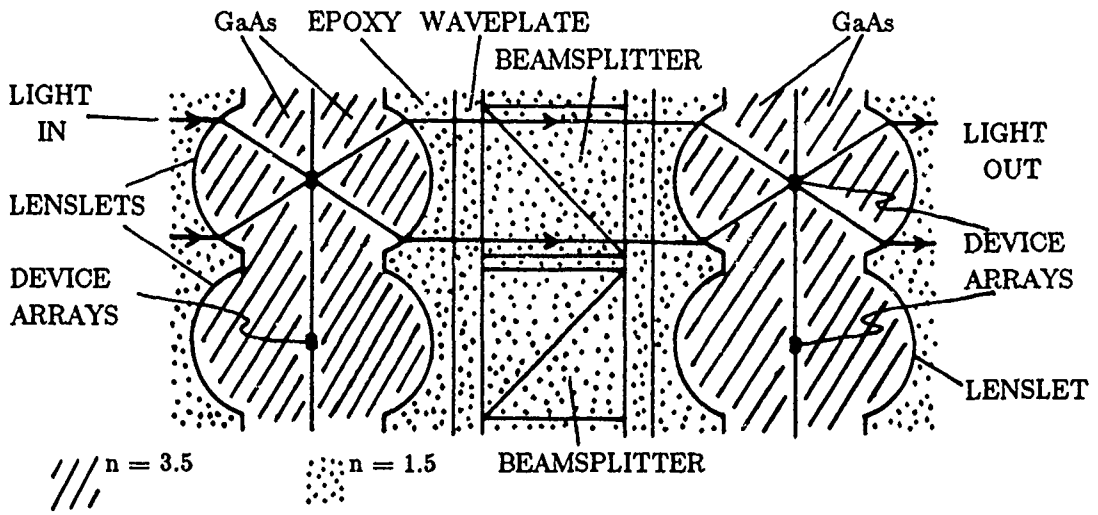


Fig. 1 Cross section of generic building blocks of a possible microoptic imaging system.

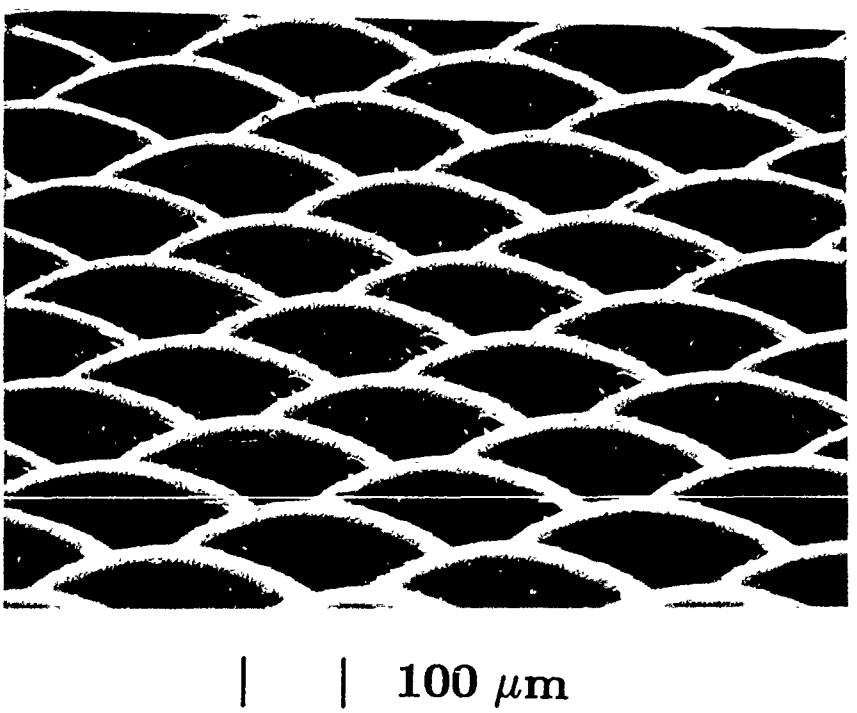


Fig. 2 Portion of an InP lenslet array.

Progress is underway in a variety of areas in microoptics including fabrication of refractive[6], diffractive[7] and gradient index[8] lenslets, and thin-film zero-order waveplates[9]. A portion of a lenslet array formed on InP by photo-electrochemical etching at AT&T Bell Laboratories is shown in Fig. 2. The lenses are designed to be part of a system having a focusing half-angle of 30°. This technology has previously yielded an *aspheric* lenslet having a surface within ±30 nm of the design[10]. The techniques used to etch 45° facets on diode lasers might be applied to the formation of microbeamsplitters. Although the state of these technologies is currently embryonic, they are the seeds of what must develop in order to make array-based photonic computing practical.

CONCLUSION

The performances of photonic information processing systems are dramatically improved by scaling down their sizes. Arrays of devices, if $\sim 32 \times 32$ in size instead of the usual 1000×1000 often sought, can dissipate many kW/cm^2 locally, have small propagation delays on the order of 10's of ps, and the systems can be manufactured by revolutionary microoptical techniques. In the history of microelectronic technology, large investments in well-chosen manufacturing techniques (e.g. clean rooms, photolithographic steppers, etc.) have been necessary and cost-effective for the achievement of high performance. We expect this to be true for microoptic technology also.

We acknowledge F.W. Ostermayer for fabrication of the InP lenslet array, and L.C. West for stimulating discussions.

REFERENCES

[1] L.C. West, IEEE Computer, pp. 34-46, December, 1987.

[2] J.L. Jewell, S.L. McCall, Y.H. Lee, A.Scherer, A.C. Gossard, and J.H. English, Journal de Physique, **C2**, 39-42, (1988); Proceedings of the Topical Meeting on Optical Bistability IV, Aussois, France, March 1988.

[3] J. L. Jewell, A. Scherer, S. L. McCall, A. C. Gossard, and J. H. English, Appl. Phys. Lett. **51**, 94-96 (1987).

[4] Jurgen Jahns and Miles J. Murdocca, Appl. Opt. **27**, 3155-3160 (1988).

[5] D.B. Tuckerman and R.F.W. Pease, IEEE Electron Dev. Lett. **EDL-2**, 126 (1981).

[6] Osamu Wada, Shigenobu Yamakoshi, Masayuki Abe, Yorimitsu Nishitani, and Teruo Sakurai, IEEE J. Quantum Electron. **QE-17**, 174-178 (1981).

[7] Teruhiro Shiono, Kentaro Setsune, Osamu Yamazaki, and Kiyotaka Wasa, Appl. Opt. **26**, 587-591 (1987).

[8] K. Iga, Y. Kokubun, and M. Oikawa, *Fundamentals of Microoptics*, Academic Press, 1984. Kenichi Iga, Proc. IEEE **813**; Proceedings of the 14th Congress of the International Commission for Optics, Quebec, Canada, Aug. 1987.

[9] Available from International Polarizer, Inc.

[10] F.W. Ostermayer, unpublished.

NOTES

TUESDAY, FEBRUARY 28, 1989

SALON F

10:30 AM–11:30 AM

TuC1–TuC4

OPTICAL INTERCONNECTIONS: 3

J. Shamir, University of Alabama-Huntsville, *Presider*

2D Optical Trimmed Inverse Augmented Data Manipulator Networks

T. J. Cloonan and M. J. Herron
AT&T Bell Laboratories, 200 Park Plaza, Naperville, IL 60566

1. Introduction

The network topologies that have been proposed for electronic switching systems must often be modified to satisfy the unique constraints placed on the system by the use of photonic technologies. This paper presents modifications to a class of O(N $\log_2$N) multistage interconnection networks known as Inverse Augmented Data Manipulator (IADM) networks[1] (Fig. 1) which are based on Plus-Minus-2^i connection patterns. The modified Inverse Augmented Data Manipulator network requires that some of these interconnections be *trimmed* from the edges of the network to simplify the optical components required in a photonic design, so the network is called the *Trimmed Inverse Augmented Data Manipulator* (TIADM) *network* (Fig. 2). The TIADM network can be modified (as in Fig. 3) to have three-dimensional interconnections between two-dimensional stages of nodes, and this network will be called a *two-dimensional (2D) TIADM network*. The large number of connections required in the 2D TIADM make the network attractive as a potential candidate for photonic switching architectures.

2. The 2D TIADM network

In one-dimensional networks, the nodes in a node-stage are normally arranged in linear (one-dimensional vector) fashion (Fig. 2). This forces the interconnections between node-stages to lie in a two-dimensional plane. In two-dimensional networks, the nodes in a node-stage are more easily arranged in planar (two-dimensional array) fashion (Fig. 3). This allows the interconnections between node-stages to take advantage of three-dimensional space. Many new implementations of two-dimensional optical networks have been proposed in the literature.[2] [3] [4] [5] [6] The 2D TIADM network is a new two-dimensional network that extends the 1D IADM connections into three dimensions while still using regular interconnects. The 2D TIADM network is illustrated in Fig. 3 for a system with N=64 input ports and N=64 output ports. For a system with N input ports and N output ports, the 2D TIADM network requires $(1/2)\log_2(N)+1$ node-stages. These node-stages are numbered from 0 to $(1/2)\log_2(N)$ from left to right. Between adjacent pairs of node-stages are link-stages. There are $(1/2)\log_2(N)$ link-stages, which are numbered from 0 to $(1/2)\log_2(N)-1$ from left to right. Each node-stage has N nodes arranged in linear fashion, which are numbered with ordered pairs (r,c) indicating each node's respective row number and column number (from (0,0) to ($\sqrt{N}$ - 1,$\sqrt{N}$ -1)). Each node is merely a 9-to-1 multiplexer with select signals gating the nine inputs.

Each link-stage of the 2D TIADM network provides $9N - 6\sqrt{N}\,(2^{i+1}) + 2^{2i+2}$ links between adjacent node-stages, because the single output from each node is fanned-out to form nine output links (but some of them are trimmed). Node (r,c) in node-stage (i) is directed the nine nodes described by (r+x,c+y), where x ε $\{2^i, -2^i, 0\}$, and y ε $\{2^i, -2^i, 0\}$. Since the resulting node number's row and column in node-stage (i+1) must be a value between 0 and $\sqrt{N}$ - 1, any connections to nodes outside of this range are connections that are trimmed from the edges of the switch.

There is always at least one path through an idle 2D TIADM network that allows data to be routed from any input source S to any output destination D.[7] [8] One method of determining a path through the 2D TIADM from input source (r_S,c_S) to output destination (r_D,c_D) requires the use of *vertical natural routing tags* and *horizontal natural routing tags*. The vertical natural routing tag is the signed difference (r_D-r_S) between the destination row number r_D and the source row number r_S, represented as a signed magnitude binary number. The horizontal natural routing tag is the signed difference (c_D-c_S) between the destination column number c_D and the source column number c_S, represented as a signed magnitude binary number. The vertical natural routing tag is used for vertical routing, and the horizontal natural routing tag is used for horizontal routing. Superposition of these two orthogonal paths yields the resultant three-dimensional path. The use of natural routing tags guarantees that the resulting natural path will never use the wrap-around connections (which were trimmed from the network).

4. Extra stages in TIADM networks

The 2D TIADM network is a blocking network, because it may not be capable of setting up a new path if active paths are already using some of the nodes required by the new path. In addition, the 2D TIADM provides no fault tolerance for many connections, because paths connecting input source S to output destination D do not have any alternate paths if S=D.[7] [8]

If the photonic network design requires decreased blocking probability and increased fault tolerance, then pairs of extra node-stages can be added to the 2D TIADM network to provide alternate paths. In Fig. 4, one of the two

node-stages in the pair is added before node-stage 0, and it is called *offset node-stage* 0. The link-stage that follows offset node-stage 0 is called *offset link-stage* 0, and it provides the connections from node (r,c) to node (r+x,c+y), where $x \in \{2^i, -2^i, 0\}$, and $y \in \{2^i, -2^i, 0\}$. The other node-stage in the pair is added after node-stage $(1/2)\log_2(N)$, and it is called *offset correction node-stage* 0. The link-stage that precedes offset correction node-stage 0 is called *offset correction link-stage* 0, and it also provides the connections from node (r,c) to node (r+x,c+y), where $x \in \{2^i, -2^i, 0\}$, and $y \in \{2^i, -2^i, 0\}$. If extra nodes are added to the top edge, bottom edge, left edge, and right edge of each node-stage, then every input source S and and output destination D has nine disjoint paths between them.

Multiple sets of offset stages and offset correction stages can be added to the 2D TIADM network to further increase the number of available paths. In general, if k offset node-stages are added at the input end of the network, and if k offset correction node-stages are added at the output end of the network, and if 2^k-1 extra nodes are added to the top edge, bottom edge, left edge, and right edge of the network, then there will exist $(2^{k+1}-1)^2$ paths between every input source and output destination.

5. Discussion

Blocking in the 2D TIADM network was studied via a computer simulation. Blocking probability can be plotted as a function of the offered call load, where the offered call load is described in terms of the percentage of the the maximum offered call load. Fig. 5 displays these plots for the different types of 2D TIADM networks with size N=64. It also displays the plots for the different types of 1D TIADM networks with size N=64, where 1D TIADM networks are like the network shown in Fig. 2. The plots show that blocking probability does increase as offered call load is increased, however the amount of increase is dependent on the network type. It is apparent that the extra stage TIADM networks offer better performance than the standard TIADM networks, because they provide lower blocking probabilities than the standard TIADM networks. In addition, it can be seen that 2D TIADM networks offer better performance than comparable 1D TIADM networks.

The entire 2D TIADM network can be implemented by appropriately connecting optical AND gate arrays and optical OR gate arrays. The optical gate arrays can be implemented with Symmetric-SEED devices[9] [10] or with OLE devices.[11] The beams propagating from a device array are oriented essentially perpendicular to the plane of the device array, and the beam-steering elements must redirect these beams to the appropriate spatial locations on the next device array. Different beam displacements are used to provide the connections in different link-stages.

One technique that can perform the beam-steering operations is based on multiple imaging techniques which employ computer-generated binary phase gratings.[12] An experimental implementation of the 2D TIADM interconnection was constructed to show system feasibility. The experimental set-up used a pair of crossed phase gratings whose Fourier transform produced a 3-by-3 array of spots. When the phase gratings were illuminated with the Fourier transform of three input spots (on a diagonal) and the output from the gratings was then inverse Fourier transfomed, the resulting output image contined three sets of 3-by-3 arrays of spots, as shown in Fig. 6.

7. Conclusion

An $O(N \log_2 N)$ multistage switching network (the 2D TIADM network) that is a modified version of the IADM has been described. An optical implementation using computer-generated binary phase gratings was also presented. It is shown that the 2D TIADM networks generally offer better performance than the 1D TIADM networks, because the 2D TIADM networks provide decreased blocking probabilities. These improvements in system performance seem to be related to the increased connectivity (pin-out) provided by the two-dimensional interconnections in the 2D TIADM network architecture.

8. Acknowledgements

The authors would like to thank Rick Morrison and Rick McCormick for useful discussions and for help in creating the computer-generated binary phase gratings.

REFERENCES

1. Siegel, H. J., *Interconnection Networks for Large-Scale Parallel Processing*, Lexington Books, Lexington, Massachusetts (1985).

2. A. W. Lohmann, "What classical optics can do for the digital optical computer," *Applied Optics* 25(10), 1543-1549 (1985).

3. T. Kumagai, and K. Ikegaya, "The two-dimensional inverse omega network," in 1985 *Int. Conf. on Parallel Processing*, D. Degroot, ed., IEEE, 325-327 (1985).

4. S.-H. Lin, T. F. Krile, and J. F. Walkup, "2-D optical multistage interconnection networks," in *Digital Optical Computing*, R. Arrathoon, ed., SPIE 752, 209-216 (1987).

5. S.-H. Lin, T. F. Krile, and J. F. Walkup, "Two-dimensional Clos optical interconnection network," in *Topical Meeting on Optical Computing*, A. Huang and A. W. Lohmann, eds., OSA 11, 98-101 (1987).

6. C. W. Stirk, R. A. Athale, and M. W. Haney, "Folded perfect shuffle optical processor," *Applied Optics* 27(2), 202 (1988).

7. D. S. Parker and C. S. Raghavendra, "The Gamma Network," *IEEE Trans. on Computers* C-33(4), 367-373 (1984).

8. K. Y. Lee and W. Hegazy, "The Extra Stage Gamma Network," in *13th Annual Int. Symp. on Computer Architecture*, IEEE, 175-182 (1986).

9. A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, "Symmetric self-electro-optic effect device: Optical set-reset latch," *Appl. Phys. Lett.* 52(17), 1419-1421 (1988).

10. A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, and J. E. Cunningham, "Photonic ring counter and differential logic gate using the symmetric self electro-optic effect device," in *Conf. on Lasers and Electro-optics*, OSA 7, TUE4 (1988).

11. J. L. Jewell, and Y. H. Lee, "Pulsed optical logic in GaAs etalons," in *Optical Bistability III*, H. M. Gibbs, P. Mandel, N. Peyghambarian, and S. D. Smith, eds., pp. 32-34, Springer-Verlag, Berlin (1986).

12. F. B. McCormick and R. L. Morrison, "Generation of Large Arrays from a Single Laser Beam via Multiple Imaging with Computer Generated Phase Gratings," to be presented at OSA Optcon '88 (1988).

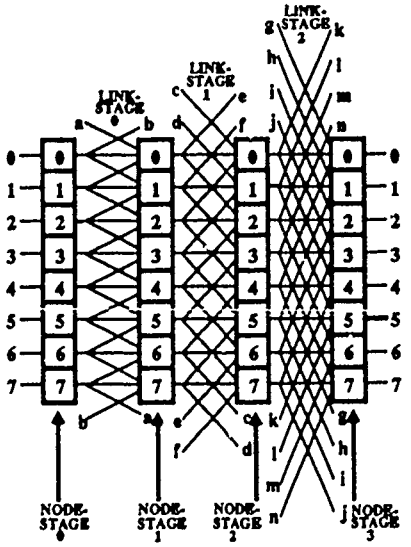


Fig. 1- IADM network of size N=8

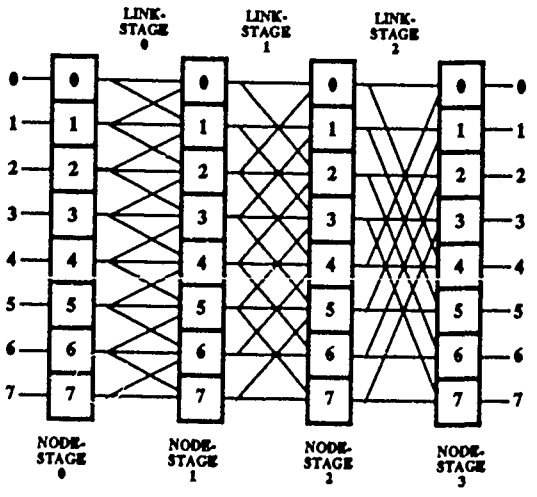


Fig. 2- 1D TIADM network of size N=8

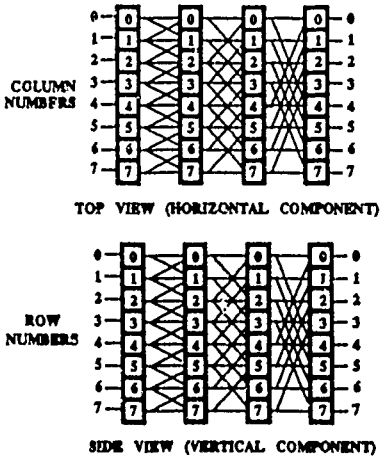


Fig. 3- 2D TIADM network of size N=64

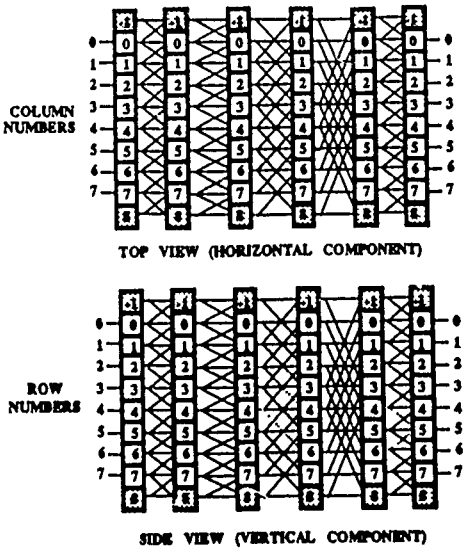


Fig. 4- 2D TIADM network with N=64 and 2k=2 extra node-stages

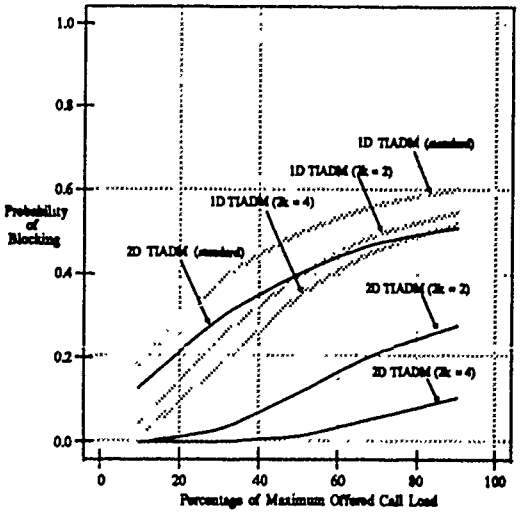


Fig. 5- Blocking probability in TIADM networks of size N=64 (2k = # of extra stages)

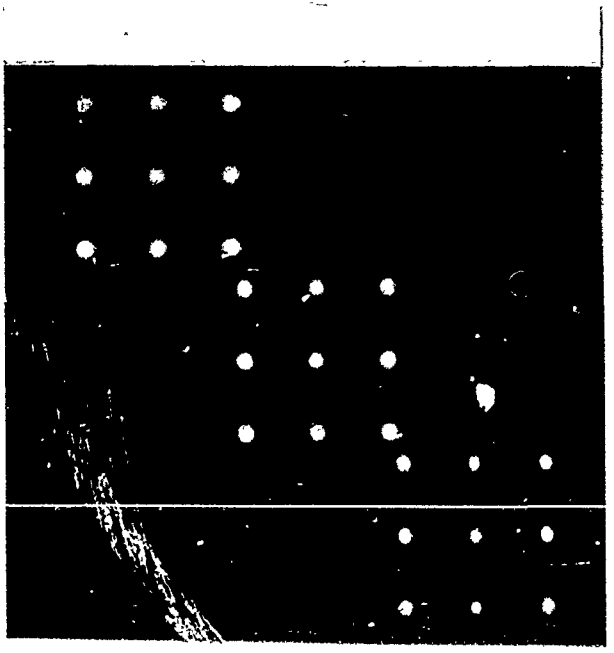


Fig. 6- Output from binary phase grating system (3 sets of 9 spots)

Alignment and Performance Tradeoffs for Free-Space Optical Interconnections*

Dean Z. Tsang

Lincoln Laboratory, Massachusetts Institute of Technology
Lexington, MA 02173-0073

Optical interconnections are of interest as a high-speed replacement for electrical interconnections in digital computers[1,2] for applications between mainframes, modules, boards, VLSI circuits,[3] and even between points within a VLSI circuit.[4] The effect of angular and positional alignment on the optical efficiency of free-space board-to-board optical interconnections is considered here for inexpensive lenses and shown to result in good system performance if reasonable care is taken in the design and assembly of the system. An experimental system was assembled and shown to operate at a rate of 1 Gb/s, a system efficiency of 18.8%, and an estimated aligned optical efficiency of 93%.

Prealigned transmitter modules, each with a diode laser and a collimating lens, and receiver modules, each with a focusing lens and a detector have been studied. The collimation and focusing lenses are assumed to be identical for optimal interconnect density. In order to maintain high optical efficiency, the lenses were assumed to be sufficiently large that optical interconnect separations are within the near field of the lens. The receiver field of view, transmitter beam angular misalignment relative to the receiver, and lateral misalignment of the two modules was considered separately based on simple thin lens approximations of the optics.

Without a lens in front of the detector, the receiver has a sensitive area proportional to the square of the detector diameter D. By placing a lens in front of the detector, we can tradeoff field of view for increased receiver aperture without increasing the capacitance of the detector. For a lens of focal length f the receiver field of view is $\theta = 2 \tan^{-1}(D/2f)$. The receiver field of view is the maximum angle allowed for light incident upon the detector lens before it no longer reaches the detector. Relative to the beam, this type of misalignment can be considered as receiver module misalignment. The field of view is plotted in Fig. 1 as a function of detector diameter for two inexpensive miniature lenses, a 1.7-mm-focal-length graded-index (GRIN) lens and a 3.9-mm-focal-length compact-disc (CD) lens . In order to maximize receiver field of view it is important to have a short-focal-length lens and a large detector. The detector can be as large as is consistent with the required speed of response. The 10-to-90% risetime is given on the top of Fig. 1 for a 50-ohm detector load impedance, a depletion width of 2.8 μm (achievable with a detector doping of 1×10^{15} cm^{-3} and a 5-V reverse bias), and a dielectric constant of 12.4. A larger detector diameter and better angular tolerance is achievable with no loss in speed if the detector is designed for biases of the order of 100 V but these voltages are not common in digital systems.

The second type of angular misalignment occurs when the beam is misaligned such that it is not fully collected by the detector lens. The output aperture of the transmitter is centered in the receiver's field of view but the angle of the transmitter module is misaligned. This can be considered as transmitter-module angular misalignment. The beam strikes the receiving lens at an angle, and the fraction of the beam collected is given by an overlap integral. For this calculation a uniform intensity across the beam is assumed. The uniform intensity assumption will result in a worst case calculation for small angles compared to a truncated Gaussian beam profile which better approximates an ideal diode laser beam. The alignment efficiency calculated by the overlap of the transmit beam and the receiver aperture (assuming the misalignment angle is within the receiver's field of view) is a function of the angle of misalignment and the separation between lenses. The allowable angle of misalignment as a function of the lens separation is shown in Fig. 2 for the

GRIN and CD lenses for an alignment efficiency of 80%. Clearly larger diameter lenses are less susceptible to angular misalignments of the source.

Positional or lateral misalignments are also determined by an overlap integral. The lateral misalignment calculation assumes that the axes of the transmitter and receiver are aligned, but that the positions of the modules are not. Again assuming a uniform intensity distribution, the overlap fraction is given in Fig. 3. Note, 80% of the light is received if the misalignment is limited to 0.3 of the lens radius, which is 0.27 mm for the GRIN lens and 0.77 mm for the CD lens. This type of misalignment is minimized with large lenses and is independent of the spacing between lenses until diffraction becomes significant. Additional positional misalignment tolerance is possible with no loss in signal if ones uses a transmitter lens that is somewhat smaller than the receiver lens, although the packing density may suffer if the lens sizes are minimized for best packing density.

The results of Figs. 1-3 can be combined to show the alignment tolerance and system performance possible with GRIN lenses for a board-to-board optical interconnect. The GRIN lens is 5 mm long with a 1.7 mm back focal length. Making modest allowances for the thicknesses of a module package and circuit board, one would expect the surface of the GRIN lens to be about 9 or 10 mm above the center of the circuit board. Thus for a 25 mm board-to-board separation there would be about a 6 mm separation between the surfaces of the transmitter and receiver lenses, which according to Fig. 2 would allow over 2 degree angular alignment tolerance with 80% optical efficiency. With the same 2-degree misalignment in the receiver module, detectors with risetimes of 50 ps should be possible.

The use of a 5.1-mm-diameter CD lens would relax the positional tolerance but the field of view for a 50-ps detector would be limited to about ± 0.7 degrees. These larger lenses are clearly better for optical interconnects over longer distances (e.g. between non-adjacent boards), although angular alignment is more critical. If the transmitter is aligned to 1 degree or better, 80% of the light can be collected at a lens separation of about 44 mm or a board-to-board separation of about 50 mm including the focal lengths of the two lenses. A 1-degree alignment tolerance on the receiver side corresponds to a 70 ps risetime. Figs. 1-3 show that the most demanding requirement for longer interconnections is determined by the transmitter module angular alignment. If very accurate transmitter module angular alignment can be maintained, these optics should be useful for even longer distance interconnections. An advantage of free-space optics for long distance interconnections is that the propagation velocity of free-space is larger than that of coaxial or fiber optic guides.

An experimental board-to-board optical interconnection was assembled to demonstrate the feasibility of efficient optical interconnections. A diode laser with 5-mA threshold current and 35%-per-facet differential efficiency was directly connected to the output of a GaAs code generator integrated circuit. A commercial GaAs D-type flip flop was directly connected to a 100-μm-diameter detector with no preamplifier. Two 5-mm-diameter 0.55-N.A. miniature lenses (attached to an optical bench) were used as collimating and focusing lenses, and the circuit boards were aligned with micropositioners. The output of the D type flip-flop at 1 Gb/s is shown in Fig. 4. Separate experiments with a 29% per facet differential efficiency laser show that the interconnection has an overall electrical current transfer efficiency (differential photocurrent out of detector / differential current into laser) as high as 18.8% with a lens separation of about 240 mm. The estimated optical efficiency is 93% at peak optical alignment.

These results demonstrate the feasibility of simple and efficient optical interconnections. Systems based on this technology and GRIN lenses for board-to-board interconnections could have card cage enclosures with flat surfaces and leaf springs to position optics between boards to

better than 0.010 inches and 2 degrees. Estimates show that optical efficiencies of 80% with 50 ps risetimes are possible if these tolerances can be met.

1. J. W. Goodman, F. J. Leonberger, S.-Y Kung, and R. A. Athale, Proc. IEEE 72, 850 (1984).
2. See for example Optical Engineering , October 1986.
3. R. K. Kostiuk, J. W. Goodman, and L. Hesselink, Appl. Opt. 24, 2851 (1985).
4. M. R. Feldman, S. C. Esener, C. C. Guest, and S. H. Lee, Appl. Opt. 27, 1742 (1988).

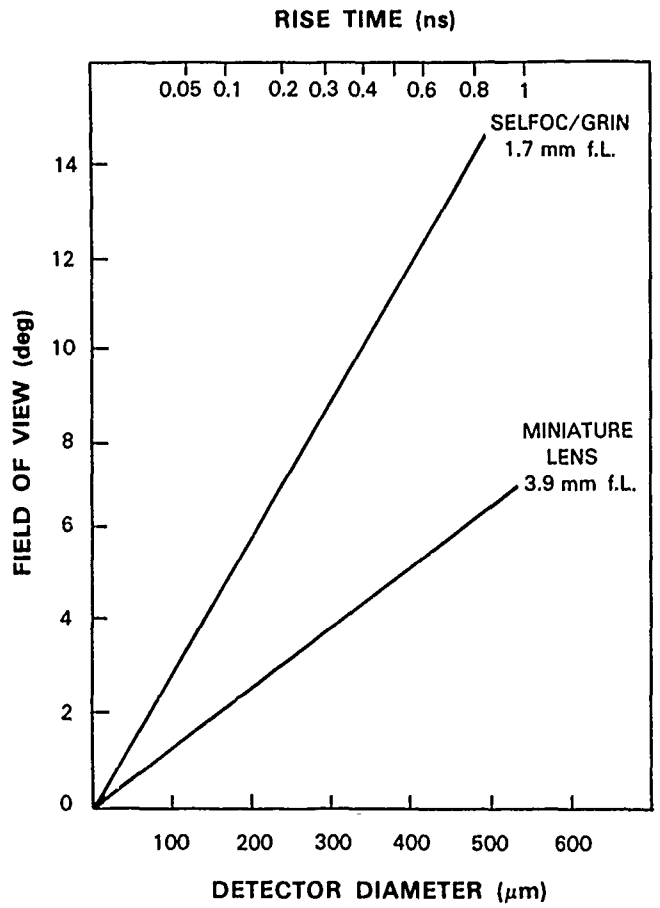


Figure 1. Detector module acceptance angle as a function of detector diameter for a 1.7 mm f.L. GRIN and a 3.9 mm f.L. miniature compact-disc (CD) lens. The upper scale shows the estimated system risetime for a reverse bias of 5 V and a 50 ohm detector impedance.

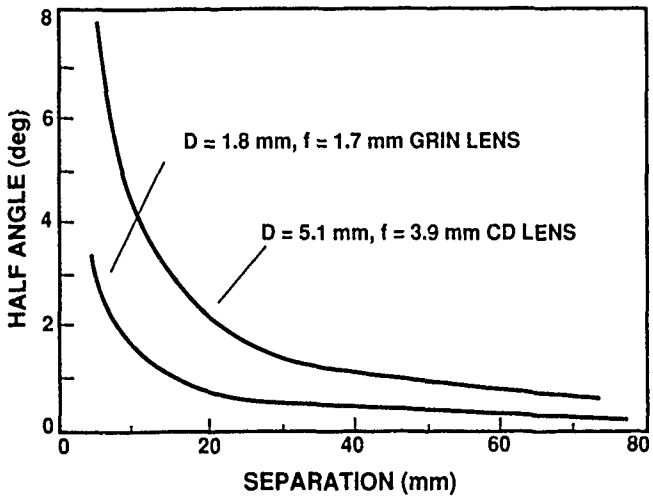


Figure 2. The source misalignment angle is plotted as a function of lens separation for both the GRIN lens and the CD lenses of Fig. 1 for the case of 80% of the peak aligned light.

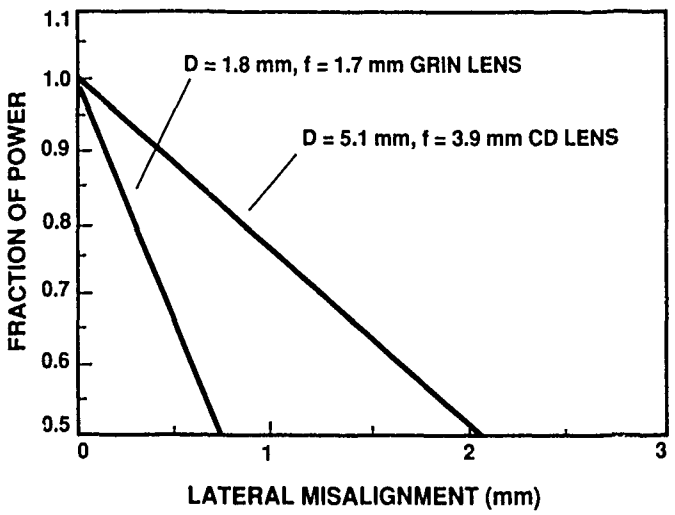


Figure 3. The effect of misaligned position is calculated for the GRIN and CD lens of Fig. 1.

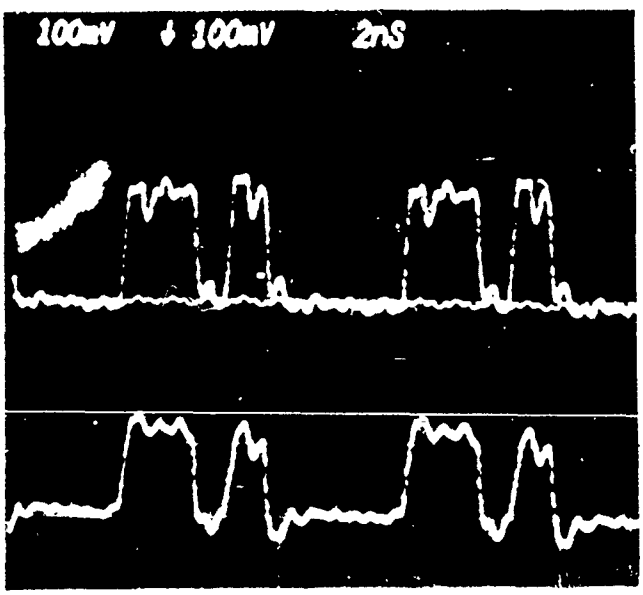


Figure 4. The top trace shows the output of the D-type flip flop on the receiver board. A superimposed trace shows the output of the flip flop with the light beam blocked. The lower trace shows the pattern of the input waveform on the transmitter board. The vertical scale is 1 V / division with 10 X probes while the horizontal scale is 2 ns / division.

Optical Holographic Interconnection Networks for Parallel and Distributed Processing

Freddie Lin

Physical Optics Corporation
2545 W. 237th Street
Torrance, CA 90505
(213) 530-1416

Introduction

With the rapid advances in technology, it is now feasible to build a system consisting of hundreds or thousands of processors [1-3]. Processors in such a parallel/distributed system may spend a considerable amount of time just communicating among themselves unless an efficient and fast interconnection network connects them. The first method for realizing optical interconnections that comes to mind is by means of optical fibers. However, optical fibers are not necessarily the ideal solution for large-scale multicomputer systems, since it requires a physical path for the interconnection between every two processors, and rather inflexible path at that.

Fortunately, if we observe most of the interconnection networks in use, we will find that the interconnections among processors have the nature of regularity. In other words, simply a hologram with the aid of several conventional optical components (such as lenses, mirror, etc.) can realize complex and massive interconnections among processors. Additionally, thick volume holograms provide very high diffraction efficiency and low crosstalk for the interconnections. For example, a single volume hologram with several stored gratings is capable of realizing numerous point-to-point interconnections simultaneously. For example, a 5-grating hologram can perform 2D mesh interconnection network of any size, and a 2^n-node binary hypercube interconnection network can be easily realized by a n-grating volume hologram.

Mesh Interconnection Network

As shown in Figure 1, a mesh interconnection network is used to communicate 4 nearest neighbors for any node in the array. Since a mesh network can be decomposed as 5 regular interconnection pattern (center, east, west, north and south), a hologram with 5 stored gratings can easily realize a mesh interconnection network of any size. Figure 2

shows the result of an optical mesh holographic interconnection network. Combining the technologies in holography and integrated optics, Figure 3 shown the planar design for mesh interconnection network.

Hypercube Interconnection Network

As shown in Figure 4, an 8-node hypercube network can be decomposed as 6 regular interconnect patterns (±1, ±2 and ±4 shifts), a hologram array with 3 stored gratings in each holographic element can, thus, implement this 8-node hypercube interconnection network. The photographic results are shown in Figure 5.

Multiplexed iInterconnection Network

The other advantage of holographic optical interconnections is that different interconnection networks are addressable for the same multicomputer system. Since volume hologram is able to multiplex stored gratings, the interconnections among processors can be changed through the addressing of various angles or wavelengths. Even a number of parallel/distributed systems can share the same hologram, while still communicating with different interconnection patterns within their own processors.

Acknowledgement

Applications of volume holography to various interconnection networks for large-scale parallel multicomputer systems are being sponsored by DOE under contract DE-AC03-88ER80673.

References:

1. J.C. Peterson, et al, "The Mark III hypercube-ensemble concurrent computer," Proc. Int. Conf. Parallel Process., Page 71 (1985).

2. W.D. Hillis, The Connection Machine, MIT Press, Cambridge (1985).

3. K.E. Batcher, "Design of a massively parallel processor," IEEE Trans. Comp. 29, 836(1980).

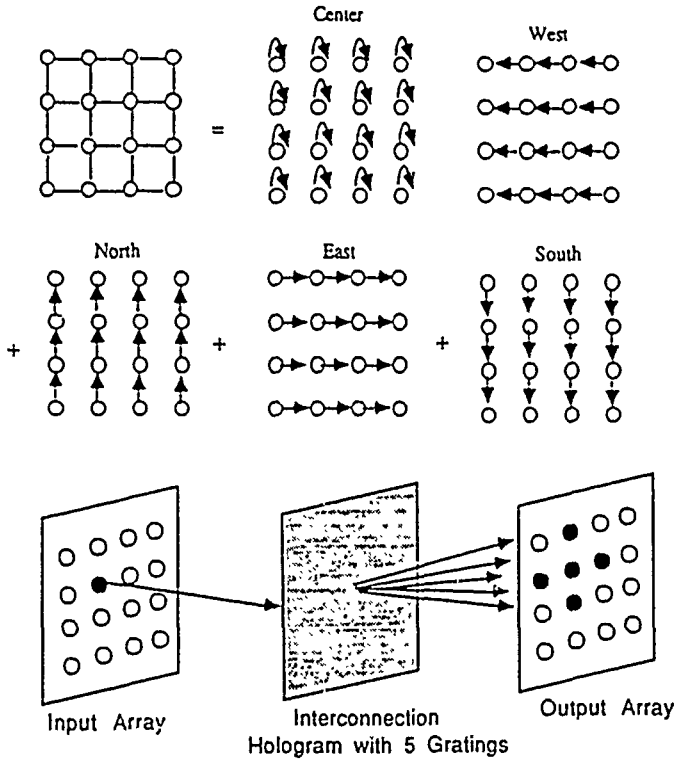


Figure 1 The decomposition of a 2D mesh interconnection network and holographic implementation.

Figure 2 Experimntal results of a mesh interconnection network for 5 points.

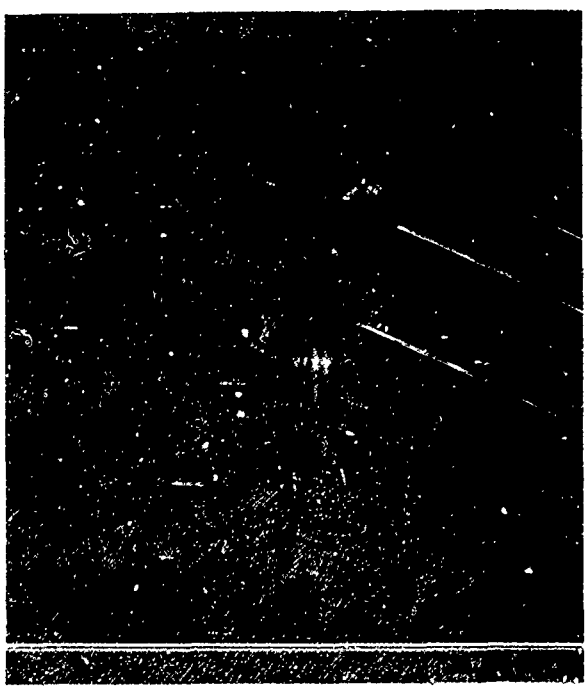


Figure 3 Experimental results of a planar mesh interconnection network.

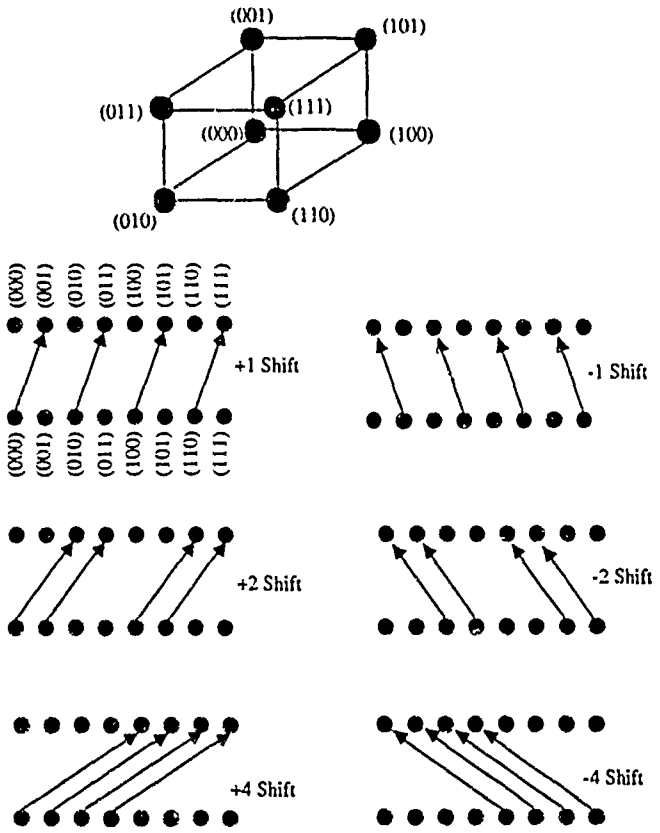


Figure 4 The decomposition of an 8-node binary hypercube interconnection network.

Figure 5 Experimental results of an 8-node binary hypercube interconnection network for node (001)

Light effective perfect shuffle using Fresnel mirrors

Yunlong Sheng and Henri H.Arsenault

Laboratoire de Recherche en Optique et Lasers
Universite Laval, Ste-Foy, Quebec, Canada G1K 7P4

Introduction

Three-dimensional optical multistage interconnection networks (MINs) can dynamically connect a 2-D array of NxN inputs to a 2-D array of NxN outputs. They exploit the 3-D nature of light propagation in free space for high density and high speed interconnections which are difficult to implement with planar electronic VLSI technology.

The 3-D multistage interconnection networks consist of a 2-D source array, a 2–D receptor array and a set of sandwiches of 2-D switch arrays and interconnect optics. The interconnect optics in each stage can be for arbitrary interconnections or for regular interconnections. An optical module for regular interconnections can easily be cascaded in multistage networks.

The most useful regular interconnection is the perfect shuffle (PS). It has been demonstrated that a limited number [O(logN)] of stages of PS followed by a switch array, or a so-called exchange box, can allow arbitrary permutations between input and output. The optical perfect shuffle was first introduced by Lohmann. Different optical PS's have been proposed using lenses, prisms, gratings, Michelson interferometers and spatial filters. We can classify those proposals into two categories: the optical PS's with a collimated input illumination and those with self-luminous sources. In most cases the optical signals to be interconnected come from LEDs, diode lasers or optical fibers. These are self-luminous sources. An optical PS with collimated input illumination requires collimating lenses for optical sources or spatial light modulators illuminated by a collimated beam to introduce the input signal.

An effective optical perfect shuffle in a practical enviroment of board-to-board and device-to-device interconnections should be: 1) able to operate in cascade; 2) have high light efficiency; 3) reliable, compact and use few optical components.

For self-luminous input sources, Lohmann proposed a perfect shuffle setup using four lenses and four prisms. Brenner and Huang proposed a Michelson arrangement for a 1-D perfect shuffle. Strik, Athale and Haney proposed a 2-D folded optical PS using four off-axis imaging lenses.

In this communication we propose to achieve a 2-D perfect shuffle and exchange box by using a Fresnel double mirror a spatial light modulator. This system has a high light efficiency. It is compact, reliable and easily cascaded for 3-D multistage architectures.

Perfect shuffle with Fresnel double mirror

An optical perfect shuffle with self-luminous inputs is simply an imaging system with a magnification of two. The system should yield two laterally shifted images of the input array. An optical mask placed in the input plane masks selectively half (in the 1-D case) or one-fourth (in the 2-D case) of each input element . The network proposed by Strik, Athale and Haney is a good achievement for such a system. Their limitation is only that the transverse shift of the imaging lenses from the optical axis must be equal to half the size of the input array. The size of a bundle with 64x64 optical fibers, for instance, can be less than 3x3 mm^2. Thus, the aperture of the imaging lenses could be limited to 3x3 mm^2, which could limit the light throughput of the PS system.

Figure 1 shows an alternative perfect shuffle for 1-D case. A Fresnel double mirror placed in front of an imaging lens yields two transversely shifted, slightly inclined images of the input array at the output plane. The optical mask placed in the input plane masks half of each input element. The perfect shuffled outputs have the same dimension as that of the input array. The next PS stage can thus easily cascaded. The inclination of the images will blur the output elements. To reduce the inclination, the double mirror should be as close as possible to the imaging lens. In this case the inclination is approximately $N\delta/3f$ where N is the number of the input elements,δ is the size of the elements and f is the focal length of the imaging lens. When N=64, δ=30 μm and f=10mm, the inclination is equal to 3.6^o.

The mirrors in the Fresnel double mirror can be replaced by two convex mirrors. The imaging lens can thus be removed, which makes the system very simple and compact. When the positions of the mirrors and the input array are fixed, the system needs no adustment. This optical PS module can easily be cascaded for multistage networks. A 2-D perfect shuffle can also implement by using two perpendicular Fresnel double mirrors.

Exchange box

The exchange box is placed in the output plane of a PS network, which is also the input plane for the next PS stage, for exchanging or not exchanging the data of adjacent elements and thus realize arbitrary permutations with the PS multistages. Lohmann proposed using Wollaston prism for the optical exchange box. For our proposed networks we can simply use a spatial light modulator in the place of the mask. All exchange functions can be implemented by reconfiguring the mask as shown in Fig.2.

The advantage of the above passive exchange box is the simplicity. It reduces the optical signal throughput by four times in each stage of a PS. The 4x4 optoelectronic switching modules with a 4-element detector and four output sources proposed by Sawchuk could be used in some stages inside a multstage network to amplify the optical signal.

Discussion

The light efficiency of the proposed optical perfect shuffle using the Fresnel double mirror is only limited by the numerical aperture of the imaging lens and is twice that of the PS using the Michelson arrangement. The proposed system needs no beam splitter and can implement a 2-D perfect shuffle.

References

1) J.Goodman, F.Leonberger, S.Kung and R.Athale, Proc. IEEE, 72, 850-865 (1984);

2) A.Sauchuk, Proc. ICO-14, Quebec Canada, 547-548 (1987);

3) A.Lohmann, Appl.Opt. 25, 1543-1549 (1986);

4)K.Brenner, A.Huang, Appl.Opt., 27, 135-137 (1988);

5) C.Strik, R.Athale and M.Haney, Appl.Opt. 27,202-203 (1988).

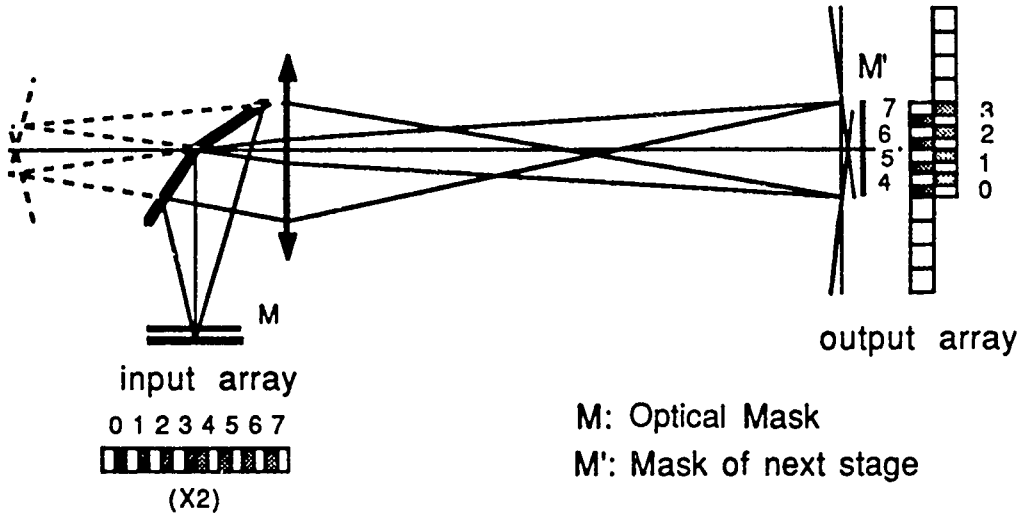


Figure 1: Optical perfect shuffle

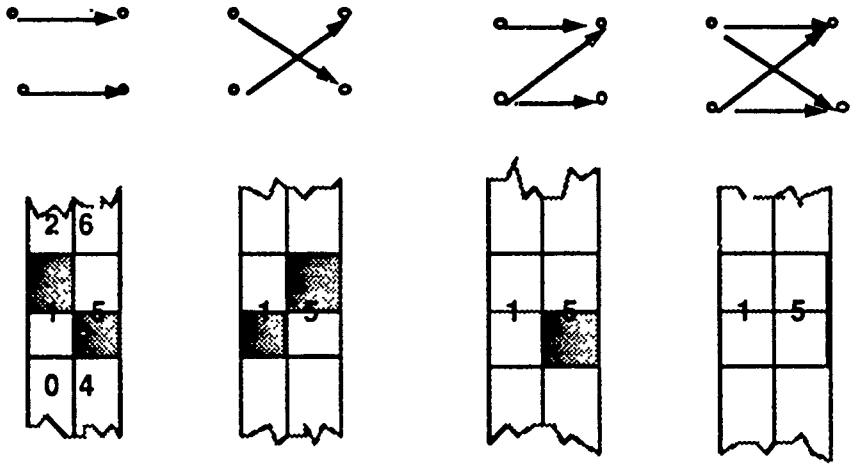


Figure 2: Exchanges by dynamic optical masks

NOTES

TUESDAY, FEBRUARY 28, 1989

SALON F

11:30 AM–12:30 PM

TuD1–TuD4

OPTICAL COMPUTING SYSTEMS AND COMPONENTS

Satoshi Ishihara, Optoelectronic Industry and Technology Development Association, Japan, *Presider*

TECHNIQUES FOR ARRAY ILLUMINATION

Norbert Streibl
Physikalisches Institut
Universität Erlangen-Nürnberg, FRG

Jürgen Jahns
AT&T Bell Laboratories
Holmdel, NJ, USA

Abstract

In an optical computing system comprising free-space interconnections the uniform illumination of two-dimensional arrays of nonlinear devices is a crucial task. Various techniques using Fraunhofer diffraction, Fresnel diffraction and spatial filtering are compared.

Summary

1. Introduction

Recently, optically bistable elements as well as opto-electronic devices which modulate the transmission of an incoming light beam have been demonstrated. These devices can be used as logic gates in an optical digital computer. In this case, the architecture would require to run two-dimensional arrays of devices in parallel.

An array generator is an optical system that splits one incoming beam of light into an array of many "beamlets". Such a system is necessary to distribute the light from one power supply laser to a one- or two-dimensional array of logic gates. The number of beams has to be adapted to the size of the array of devices. Array sizes from 16x16 up to 128x128 are desirable. In many practical cases the light efficiency is an important issue such that array illuminators comprising only nonabsorbing, i.e.phase components have to be designed.

2. Working Principles

A wide variety of optical systems for array illumination has been studied in the past [1]-[17]. These include the use of microlenses ("lenslets") [5,10,11], arrays of microtelescopes ("telescopelets") [13], computer generated diffraction grating (such as Dammann gratings [1-4,6-8] or number theoretic gratings[15]) in Fraunhofer diffraction systems, Fresnel diffraction techniques using regular phase gratings (fractional Talbot effect) [14], phase contrast imaging [10,12] and optical coordinate transformations [16]. Possibly, there are even more viable working principles.

All approaches to array generation have in common that a special component - in most cases a light efficient phase component - is used to achieve the light concentration into the generated spots. The array generators can be classified according to the position where the phase component is located within the optical system:

- Image plane array generators use a component situated in a plane conjugated to the array of devices [10,12].
- Fresnel plane array generators are characterized by a Fresnel diffraction step between the component and the devices [14,16].
- Fourier plane array generators based on Fraunhofer diffraction [1-9].
- Tandem or multicomponent configurations comprise several elements which can be situated in different successive planes [10,12,15].

3. Requirements

All approaches have to be compared with respect to different parameters. Here is a (possibly incomplete) list of important parameters:

- The splitting ratio is the number of beams generated out of one incoming beam. As mentioned before, sizes running up to 128 by 128 are of interest.
- The compression ratio is the quotient of the bright area of a spot and the whole area of the elementary cell of the spot pattern. It measures how well the beams of an array are separated from each other. Compression ratios of up to 100 are desireable for digital optical computers.
- The inhomogeneity is the difference in intensity of the brightest and the darkest spot within an array. The maximum allowable inhomogeneity depends on the tolerances of the devices. For many applications a 10% variation will be tolerable.
- Obviously, the manufacturability and cost are important issues. Lithographic techniques to produce the phase components would be favourable for ease of manufacturing.

4. An Experimental Example

As an example for an array generator, Figure 1 shows part of the binary phase structure of a Dammann grating which was manufactured by lithographic techniques and reactive ion etching [6]. The generated array is shown in Fig. 2.

5. Conclusion

A survey of array illumination techniques is presented. The different techniques can be classified depending on the position of the spot generating component. Criteria for comparisons between the different array generators are based on the requirements which are imposed by the logic devices used in the optical computer.

References

[1] H. Dammann, K. Görtler: Opt. Comm. 3 (1971) 312
[2] H. Dammann, E. Klotz: Opt. Acta 24 (1977) 505
[3] U. Killat, G. Rabe, W. Rave: Fiber and Int. Optics 4 (1982) 159
[4] W. B. Veldkamp, J. R. Leger, G. J. Swanson: Opt. Lett. 11 (1986) 303
[5] J. R. Leger, M.L. Scott, P. Bundman, M. P. Griswold: Proc. SPIE 884 (1987)
[6] J. Jahns, M. E. Prise, M. M. Downs, S. J. Walker, N. Streibl: OSA Ann. Meeting 1987, Rochester, NY
[7] F. B. McCormick: OSA Ann. Meeting 1988, Santa Clara
[8] M. E. Prise, M. M. Downs, F. B. McCormick, S. J. Walker, N. Streibl: Topical Meeting on Opt. Bistability IV, 1988, Aussois, France
[9] A. W. Lohmann, F. Sauer, N. Streibl, R. Völkel: subm. to HICSS-22, 1989, Hawaii
[10] A. W. Lohmann, J. Schwider, N. Streibl, J. A. Thomas: Appl. Opt. 27 (1988) 2915
[11] A. C. Walker, M. R. Taghizadeh, J. G. H. Mathew, I. Redmond, R. J. Campbell, S. D. Smith, J. Dempsey, G. Lebreton: Opt. Eng. 27 (1988) 38
[11] A. W. Lohmann, W. Lukosz, J. Schwider, N. Streibl, J.A. Thomas: to be publ. in Proc. SPIE 963
[12] A. W. Lohmann, F. Sauer: Appl. Opt. 27 (1988) 3003
[13] A. W. Lohmann: Optik 79 (1988) 41
[14] M. R. Schröder: Number theory in science and communications (2nd ed.), p. 296ff, Springer, Berlin, 1985
[15] W. J. Hossak, P. McOwan, R. E. Burge: Opt. Comm. 68 (1988) 97
[16] S. J. Walker, J. Jahns: OSA Ann. Meeting 1988, Santa Clara, CA

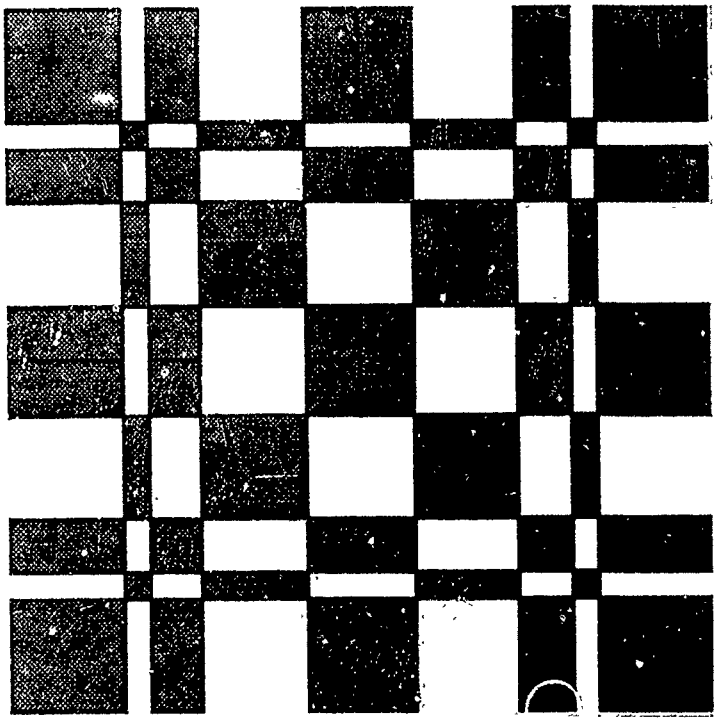


Figure 1:
Binary pattern for a grating used for array generation according to ref. [1]. Only one period is shown. Shaded areas correspondto areas which introduce a phase shift of pi.

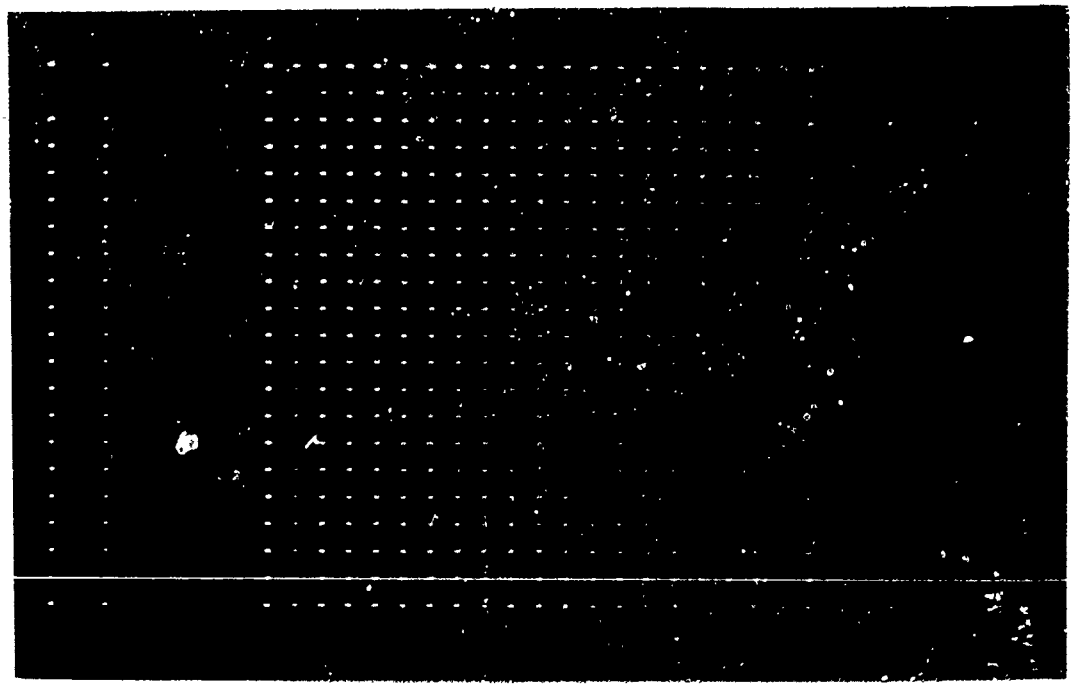


Figure 2:
Array of 17x17 spots generated with a binary Dammann grating.

ARRAY ILLUMINATOR USING GRATING COUPLERS

Mitsuo TAKEDA and Toshihiro KUBOTA †

University of Electro-communications
Department of Communications and Systems Engineering
1-5-1, Chofugaoka, Chofu, Tokyo, 182 JAPAN

† Kyoto Institute of Technology
Department of Electronics and Information Science
Matsugasaki, Sakyo-ku, Kyoto, 606 JAPAN

1. Introduction

Several schemes have been proposed for implementing array illuminators,[1-4] which are used to distribute optical power to an array of optical logic gates or bistable devices that require optical power supplies. The schemes reported so far are based on assembling bulk optical elements such as Damman gratings[1] or lenslet arrays[3]. In these schemes, relatively large spaces between optical elements (and between optical elements and devices, as well) constitute an essential part of the illuminating system because optical power is distributed by propagating light through these spaces. The system, therefore, tends to lack compactness, and alignments and stabilities of the total system may become a critical issue as the system grows complex. To solve these problems, we take an alternative approach based on integrated optics. In this paper, we propose to use grating couplers[5,6] to generate an array of many beamlets.

2. Optical Waveguide and Grating coupler

As shown in Fig.1, a film with thickness t and refractive index n_1 is sandwiched between air and a substrate whose indices n_0 and n_2 are lower than n_1.

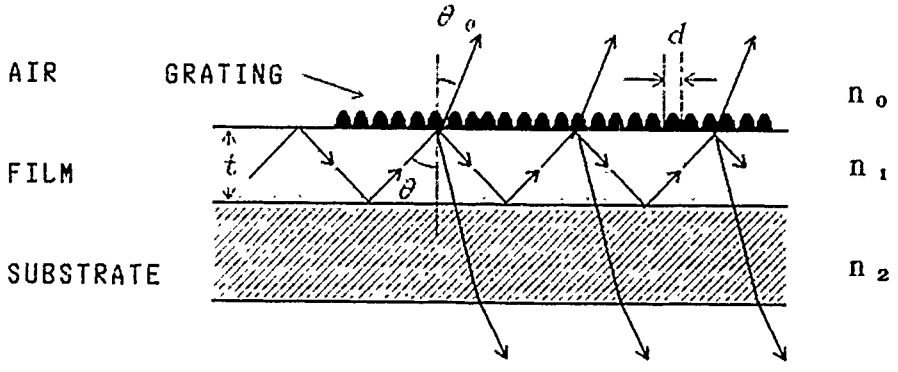


Fig.1 Geometrical configuration of grating coupler.

Because of total reflections at the two boundary surfaces, the film works as a lossless optical waveguide for those beams whose incident angle θ is greater than the critical angle. A phase grating of period d is fabricated on the top surface of the film, so that the evanescent waves are coupled with the grating, and radiate a part of their power out from the waveguide in the form of multiple beams exiting at regular points separated by $2t\tan\theta$. From the coupled-wave theory, the propagation constant β_0 of the radiated beams is given by

$$\beta_0 = \beta - 2m\pi/d, \tag{1}$$

where $\beta_0 = k_0\sin\theta_0$, $\beta = k_0 n_1 \sin\theta$ (λ_0 is a wavelength in the air), and m is an integer. Let us now consider a simplest case where the substrate is replaced with air, and suppose that the diameter of the guided beam is smaller than the beam separation $2t\tan\theta$. Then, we can obtain multiple beams running parallel to each other. These beams are obtained from both sides of the waveguide, so that the number of beams available is doubled. If we choose a grating period d which makes $\theta_0 = 0$, then we can obtain multiple beams exiting normal to the waveguide surfaces. A two-dimensional array illuminator can be made by combining two waveguides in such a manner that their grating lines run normal to each other, as shown in Fig.2. Prisms are used to introduce optical beams into the waveguides at an angle that causes total reflection.

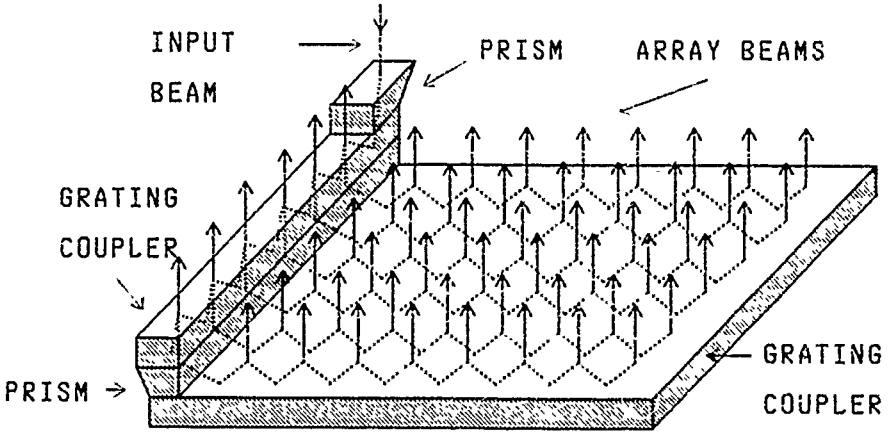


Fig.2 Two-dimensional array illuminator using grating couplers.

3. Experiments

Preliminary experiments have been conducted to examine the validity of the proposed principle. A waveguide was made of a 2.5mm thick glass plate with $n_1 = 1.51$, both surfaces being bounded by air so that $n_0 = n_2 = 1$. The top surface of the glass plate was coated with a photoresist film, in which a grating coupler was fabricated by holographic technique. The grating coupler was so designed that a collimated beam from a He-Ne laser source (λ_0 =632.8nm) incident on the grating at $\theta = 45°$ from inside the glass can generate multiple beams that propagate in the directions normal to the waveguide surfaces. The grating period that satisfies this requirement is $d = 593$nm. Figure 3 shows a picture demonstrating how the waveguide illuminator generates multiple parallel beams that propagate in the directions normal to its surfaces. A light beam is introduced from left

by a prism attached to the waveguide. Beams propagating upwards exhibit their trajectories through the cigarette smokes introduced for observation purpose. Beams radiated downwards do not exhibit their trajectories but their beam spots can be observed in the bottom of the picture. Figure 4 shows an array of beam spots generated by a two-dimensional array illuminator which was made according to the principle depicted in Fig.2.

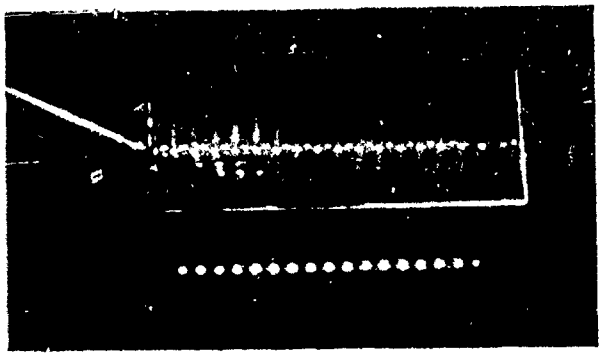


Fig.3 Mupliple beams generated by one-dimenionsal grating coupler.

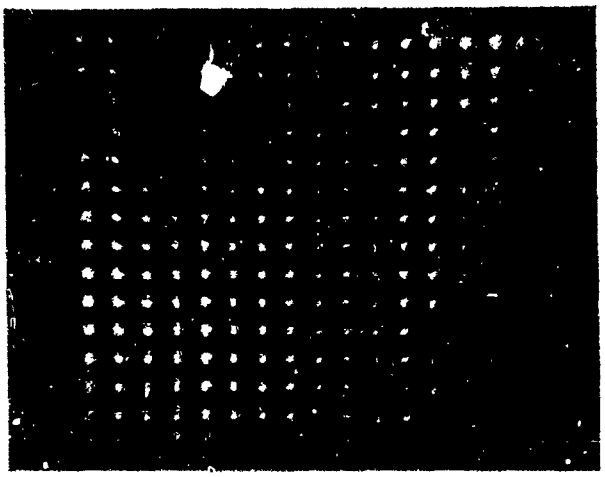


Fig.4 Beamspots generated by two-dimensional grating coupler.

4. Discussions

Since the grating fabricated for this preliminary experiment has a uniform coupling efficiency all over the waveguide surface, the power of the radiated beams decreases monotonically as the guided beam repeats internal reflections giving a part of its power to the outgoing beams. To correct this power nonuniformity, the coupling efficiency of the grating should be increased along the path of the guided wave. This can be accomplished by making a tapered grating that has an increasing groove depth. Since the light absorption in the waveguides and couplers can be made negligibly

small, a high power efficiency can be achieved. The array illuminator described above has a relatively large waveguide thickness. This thickness causes the diameters of the radiated beams to increase rather rapidly with the number of reflections experienced by the guided wave. This is because the beam departs from the its beam waist as it travels through the waveguide, and it will limit the number of separated beams available. One method to obtain a large number of beams with same diameters would be to use a very thin film like that used in conventional waveguides, and to fabricate many small islands of gratings only at locations where beams should be radiated from, as shown in Fig.5. Presumably, a 100×100 point array illuminator can be made by fabricating 10,000 grating islands, each having 250μm×250μm area and 250μm apart, on a 5cm×5cm waveguide which is capable of generating a total of 20,000 beams; a half of the beams exiting from the other surface of the waveguide propagate in the opposite direction. This type of array illuminator has various advantages. Locations and diameters of the exiting beams can be controlled easily in the fabrication process, and yet they are insensitive to the direction and the beam-waist position of the source beam introduced to the waveguide. This will give the possibility of making an illuminator which generates beams at arbitrary locations (without being restricted to be regularly spaced). Furthermore, we can envision stacking on this array illuminator various planar microoptics devices, like a monolithically integrated microlens array, which may be useful for interfacing. In addition, other waveguide-devices, such as light modulators and deflectors, could be fabricated on the same waveguide by choosing a proper material for the waveguide. This will add new functions and possibilities to the array illuminator.

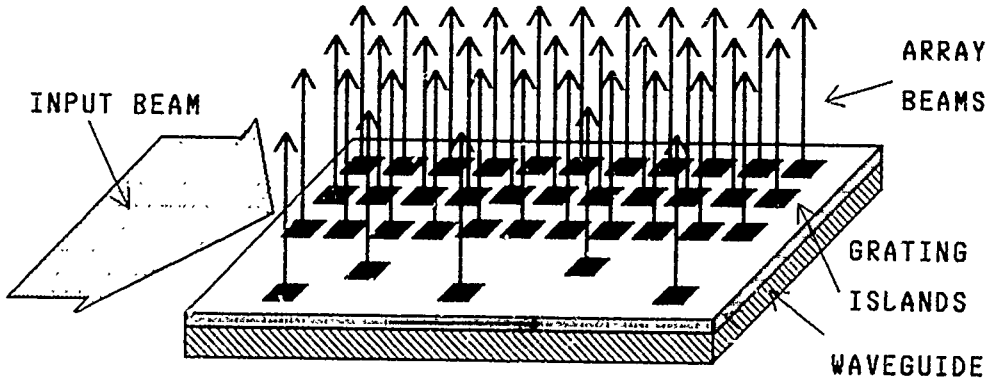


Fig.5 Array illuminator based on grating islands fabricated on waveguide.

A part of this work was supported by the Grant in Aid of Ministry of Education and Science, Japan (No.63460065).

References

1) J.Jahns, M.E.Prise, M.M.Downs, S.J.Walker, and N.Streibl, J.Opt.Soc.Am.(A), **4**, P69 (1987).
2) A.W.Lohmann, J.Schwider, N.Streibl, and J.Thomas, Appl.Opt., **27**, 2915 (1988).
3) N.Streibl, Angewandte Optik, Annual Rept. Univ. Erlangen, 37 (1987).
4) A.W.Lohmann, Optik, **79**, 41 (1988).
5) M.L.Dakss, L.Kuhn, P.F.Heidrich, and B.A.Scott, Appl.Phys.Lett., **16**, 523 (1970).
6) D.G.Dalgoutte, Opt.Communi., **8**, 124 (1973).

Substrate Mode Holograms for Optical Interconnects

Raymond K. Kostuk, Masayuki Kato, Yang-Tung Huang
Department of Electrical and Computer Engineering and
Optical Sciences Center
University of Arizona
Tucson, AZ 85721, USA
(602) 621-6172

Abstract:

The advantages and design considerations for free-space holographic interconnects are discussed. Substrate-mode holograms for this application are introduced and experimentally demonstrated.

Introduction:

Optical interconnects offer many potential advantages for both optical and electrical computer systems[1]. One of the principal advantages cited is 2-Dimensional transfer of information to a number of receiver locations for parallel processing applications. These systems require free-space optical elements to image light to the desired processing elements. In other connection applications such as board and wafer scale clock distribution, vertical and 90° crossover links are required which are difficult to implement by standard integrated optical approaches.

Our presentation begins with a discussion of relevent design criteria for a free-space holographic clock distribution system. Problems associated with alignment and chromatic aberrations, as well as multiplexing issues are analyzed and related to system performance. It will be shown that a holographic element used in conjunction with a substrate guided optical beam can greatly reduce alignment difficulties and improve overall system performance.

Planar multiplexed holograms have been suggested for coupling between totally internally reflected (TIR) beams propagating in a dielectric substrate[2]. Holograms have also been demonstrated which couple light from the evanescent field of a guided mode of a waveguide[3]. However a detailed description of the performance of these elements for optical interconnect applications has not been presented. In this paper we provide this description and demonstrate several types of substrate-mode holograms. Different arrangements for using these elements for optical clock and parallel processing systems will also be discussed.

Holographic Interconnect Requirements:

In a free-space optical interconnect system the separation of the hologram and receiver plane determines the magnitude of the shift in the image position as a function of misalignment. For efficient power transfer the detector must be large enough to encircle the

focused optical beam. When attached to an appropriate biasing circuit the frequency response of the detector can be evaluated and the performance of the system determined as a function of alignment. A set of curves indicating the required optical power for driving different silicon detectors to a supply voltage of 4.1 volts (2 μm CMOS process) are illustrated in Figure 1. As indicated, in a system with fixed illumination power the operating frequency must be reduced as the detector size increases. A larger detector however, can accomodate greater hologram misalignment. For example, a hologram-receiver plane separation of 1 cm and a 5 μm image shift corresponds to a misalignment tolerance of 0.03° which is a stringent engineering requirement. Increasing the detector diameter to 100 μm increases the alignment tolerance to 0.3°, but lowers the operating frequency by an order of magnitude.

A second consideration for holographic systems is the chromatic properties of the reconstruction source. A modulated laser diode is susceptible to mode hopping which causes the image to shift as a function of the wavelength change. Aging and fabrication variations have similar effects on the image position. One method of compensation is to use a pair of gratings with the -1 order from the first grating diffracted into the +1 order of the second.

Another consideration for holographic interconnects is the limit of detector separation on the receiver plane. It has previously been shown that the maximum separation length is restricted to the same order of magnitude as the hologram aperture[4]. Distances greater than this leads to severe image spreading from projection effects.

The principal advantage of using HOEs for interconnect applications is the ability to multiplex several images in a thin, light weight recording material. Using holograms to transmit information over appreciable distances however, is not desireable due to misalignment sensitivity and image projection effects.

Substrate Mode HOEs and System Configurations:

One possible solution to the above problems is the substrate-mode hologram illustrated in Figure 2. The basic component consists of a compound TIR holographic element which couples light from an optical source into a dielectric slab at an angle exceeding the critical angle, and couples light out at a predesignated location with a focusing HOE. The output location corresponds to a detector or spatial light modulator window. If the element is formed with the guiding beams it can be self-aligned to the receiver positions, thus reducing the detector size and increase the system operating frequency. Since a grating pair is used some chromatic compensation for laser diode wavelength variation is also possible. In addition, a multiple imaging element can be used instead of the single focus for parallel processing applications.

In our presentation we discuss several types of substrate-mode holographic components which are necessary to implement a complete interconnect system. These include high efficiency TIR transmission and reflection elements, focusing grating couplers, and quadrant beam splitters. Fabrication techniques and experimental results for individual components will also be presented.

Initially we plan to use substrate-mode holographic interconnects for a high speed optical clock distribution system at the chip connection level of an electronic system. One configuration for realizing this is illustrated in Figure 3. The incident beam is first coupled into the slab waveguide by a multiple grating which directs the beam to different quadrants of the chip. After propagating to a receiver, the individual beams are focused onto small area detectors.

In summary we present the effects of several important hologram characteristics on the performance of optical interconnect systems. Many of these difficulties can be overcome by using systems of substrate-mode (TIR) holographic optical elements. These components allow transverse displacement of optical signals, high coupling efficiency, and minimal sensitivity to misalignment.

References:

1. J.W. Goodman, F.J. Leonberger, S-Y Kung, and R.A. Athale, "Optical interconnections for VLSI Systems", *Proc. IEEE*, vol. 72, 850 (1984).
2. T. Jannson and S-H Lin, " Highly-Parallel Holographic Integrated Planar Interconnects", *Topical Meeting on Spatial Light Modulators and Applications, S. Lake Tahoe, NV.*, vol. 8, 56 (1988).
3. A. Wuthrich and W. Lukosz, "Holographic with Guided Optical Waves", *Appl. Phys.*. vol. 21, 55 (1980).
4. R.K. Kostuk, J.W. Goodman, and L. Hesselink, "Design considerations of holographic optical interconnects", *Appl. Opt.*, 26, 3947 (1987).

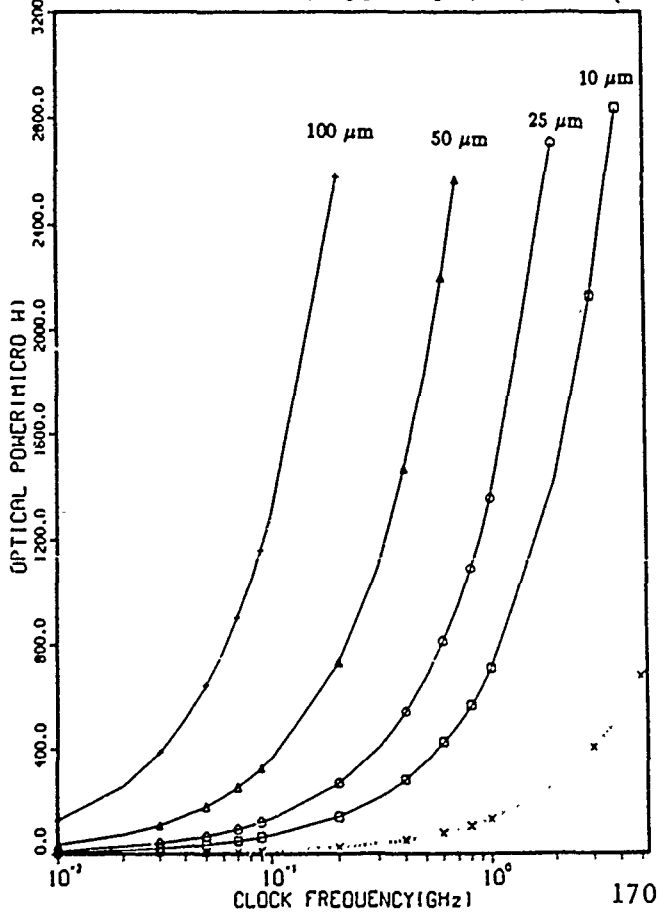


Fig. 1. Optical Power vs. Frequency for Different Diameter Detectors

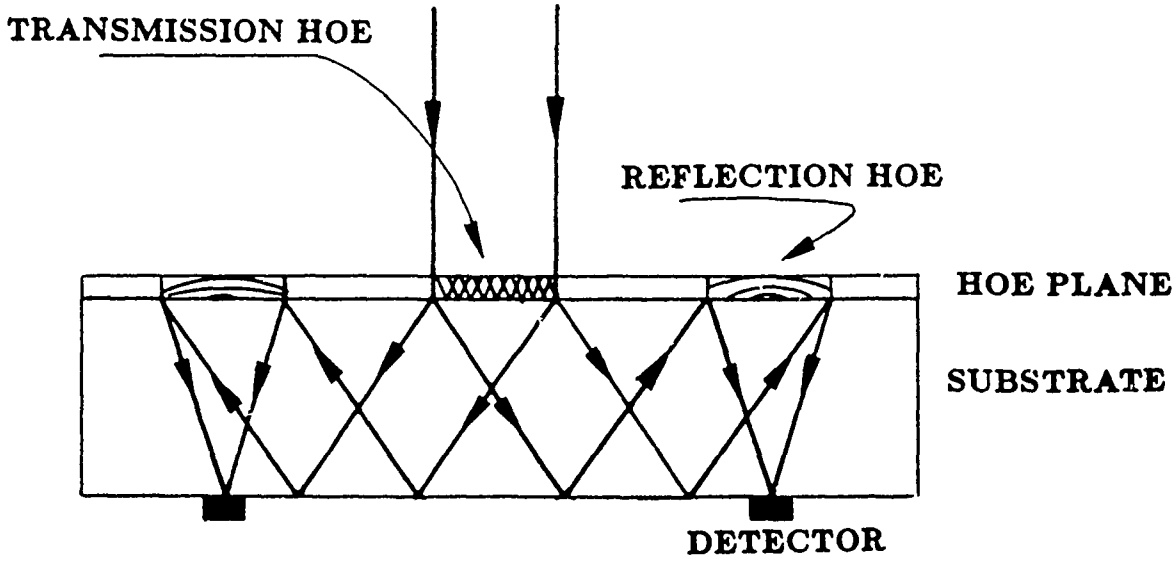


Fig. 2. Substrate-Mode Hologram

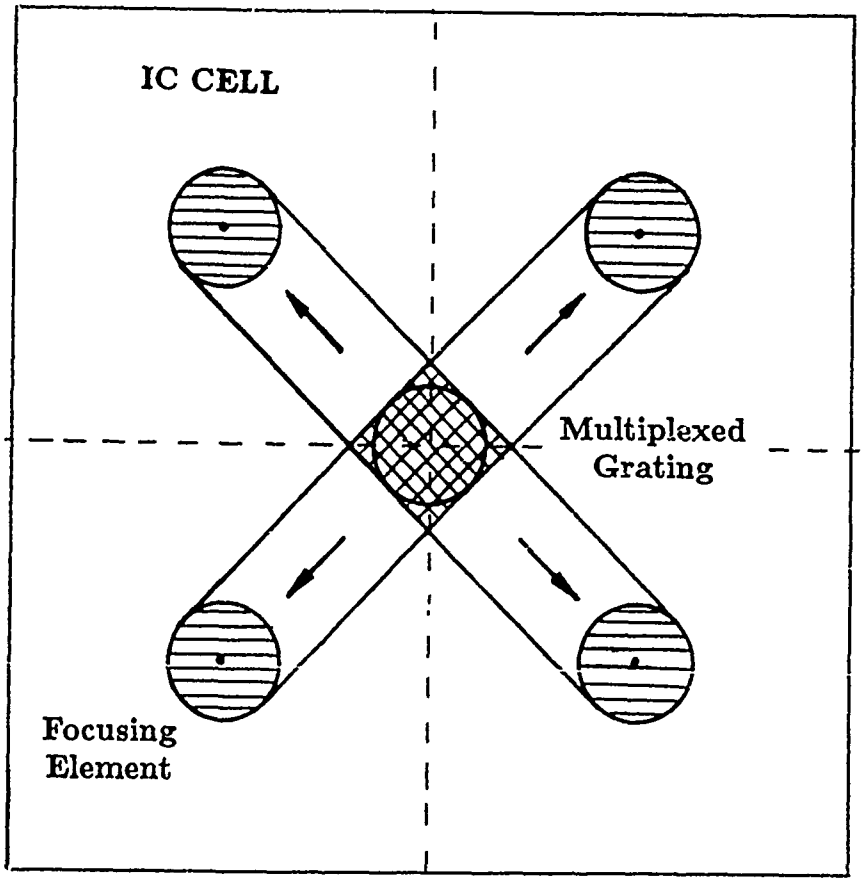


Fig. 3. Clock Distribution System

HYBRID ACOUSTOOPTIC SPECTRUM ANALYSER FOR RADIOASTRONOMY WITH SEMICONDUCTOR LASERS

N.N.Evtihiev and V.V.Perepelitsa
USSR, Moscow, Kashirscoe shosse 31,
Moscow Engineering Physics Institute
N.A.Esepkina, S.V.Pruss-Zhukovsky, O.N.Vlasov and
S.K.Kruglov
USSR,Leningrad, Politehnicheskay street 5,
Leningrad Polytechnical I stitute

Acoustooptic processors are mostly long-term devices for real time signal analysis and for modern optical parallel computers. Their merits are evident in the field of radioinformation, for instance in radioastronomy. But the realization of the whole cycle processing with the demands of high output SNR, high reliability and stability of the results of computation can be possible only in a hybrid optoelectronic processor with the controlled semiconductor laser. These lasers broaden the field of application of optical processors that includes the one pulse analysis of high rate processes, gives the opportunity to recieve compact and high stable devices.

The usage of semiconductor lasers in the optical processing systems was restrained by the lack of coherense, reliability and power. The last one is mostly important because of the high requirements defined by the sensitivity of a CCD pixel and the two-dimensional character of a CCD-camera. 1 W optical power or higher is necessary in a real time pulse processing. Characteristics of two types of semiconductor lasers are examined in our work in the scheme of optical spectrometers - one frequency stripe lasers with the optical power of 40 mW and phased arrays of Y-type diode lasers with the output power of 200 mW. The last one is very convinient for acoustooptic systems, because they form quAsionedimensional optical field in the far zone with high intensity. The characteristics of these lasers are discussed together with the optimized beam forming schemes.

The spectrometer for radioastronomy that is irvestigated in our report includes the acoustooptic modulator with the accumulation time-bandwith product 1000 in a high frequency bandwidth for heterodyned signals, linear CCD photodetector and a diode laser. The whole system is controlled by the computer that provides a low noise level. The elimination of the antenna noise is provided by it s subtraction from the signal, permanent calibration of the optical processing channel and by the adaption of the system to the changing conditions of sky observation. The feedback together with the programmable noise elimination provides an automatic processing of the device in a wide frequency band with high resolution.

The discussed spectrum analyser is creatrd for the RATAN-600 radiotescope and for other large antennas. Today it is the only device that can operate in a real time mode with high SNR and wide frequency band.

1.F.S.Guilfoyle - Optical Comp.-88, Toulon, France, p.191, (1988).
2.N.A.Esepkina,S.V.Pruss-Zhukovsky et al.- Izv.SAO, 14, p.144, (1981).
3.N.N.Evtihiev,N.A.Esepkina et al.- Kvantovaja Electr., 15, 4, p.847, (1988).

NOTES

TUESDAY, FEBRUARY 28, 1989

SALON F

2:00 PM–2:45 PM

TuE1–TuE2

OPTICAL COMPUTING SYSTEMS: 1

Steven C. Gustafson, University of Dayton, *Presider*

Perspectives in Optical Computing

John A. Neff
E. I. DuPont deNemours & Co., Inc.
Newark, DE 19714-6099

Optics is starting to get serious consideration in application areas previously dominated by electronics. One such application area is supercomputing where efforts to increase performance are beginning to turn toward the interconnection of many microprocessors rather than trying to attain a single super processor. The importance of communications in these new multiprocessing architectures has focused attention on optics to provide the necessary bandwidths. This overview of the opportunity for inserting optics into these parallel supercomputer architectures will begin with a brief description of parallel architecture, followed by a discussion of the critical optical device and materials needs.

Course-grained, Heterogeneous Multiprocessors

There are two categories of parallel machines worth discussing with respect to the benefits which can be derived from the use of optics. The first is referred to as a course-grained, heterogeneous architecture. This nomenclature represents the fact that there are a relatively few, complex processors that may differ from one another. These processors frequently interchange data via some sort of communications network. The reasoning behind such an architecture is that higher overall performance may be gained by using several different special purpose processing elements (PEs) in parallel rather than one super serial computer attempting to accomplish a diverse set of tasks. For example, such an architecture designed for image understanding might include such special purpose PEs as correlators, feature extractors, syntactic pattern recognizers, inference engines, and matchers. The PEs for a heterogeneous database machine would likely include sorters, joiners, and processors for performing union, intersection, and projection.

The possibilities for optics are twofold: to provide a high bandwidth communication network amongst the PEs, and to take advantage of the special purpose, high throughput nature of optical processors to function as some of the PEs themselves. An optical

correlator PE, for example, could prove valuable in a heterogeneous, image understanding multiprocessor. However, the most serious bottleneck to computing performance is the network. Since the course-grained architectures generally contain fewer than 100 PEs, a generalized crossbar switch would serve as the ideal network. Due to the gigahertz bandwidth requirements of such switches when used in computer applications, optics will likely prove to play a major role in switch implementation.

Fine-grained, Homogeneous Multiprocessors

The second category of parallel machines which optics is likely to impact is referred to as fine-grained, homogeneous and is composed of thousands of small microprocessors, all alike, each working on a small portion of the problem. An analogy is that many brick layers can build a wall faster than a single bricklayer. This assumes, of course, that the many bricklayers communicate so that their portion of the task interfaces well with those of the other bricklayers. This once again emphasizes the importance of communication in such an approach to computing architecture, and suggests why optics may have an important role to play. In fact, these architectures may be described as communication intensive as opposed to the switching intensive nature of other architectural classes. This communication intensive nature becomes very evident in one subcategory of these multiprocessors - neural networks. There is considerable sentiment that neural networks are so communication intensive that optics may be the only viable way to implement full scale systems.

Due to the large number of PEs in these fine-grained systems, crossbar networks are out of the question; however, optics may provide solutions to the interconnection problem via beam-steering, free-space channels which are reconfigurable in real-time. The most likely implementation of this beam-steering will be through either modifiable or multiple-selection holographic gratings. The former uses only one rewritable hologram per channel at any given instant of time, whereas the latter employs arrays of holograms per channel and the reconfiguration is done by activating the desired hologram.

The holographic interconnections will be done by planes of holograms interspersed between planes of switching and logic elements. Initial implementations will likely use optoelectronic

arrays for these planes, thereby accomplishing most of the switching actions electronically. As nonlinear material improvements permit the realization of practical optical logic elements, the electronics will be replaced by optics to avoid the losses associated with repeated conversions between photons and electrons.

Device and Material Needs

Attempts to demonstrate any of the above mentioned applications of optics have been largely frustrated by the lack of two-dimensional devices, the most critical of which is the logic or switching plane array. The type of device that is most urgently needed here is a spatial light modulator (SLM) which can combine electronic circuitry with optical detection and modulation elements at each array site. This will permit logic operations to be performed electronically while providing access to other logic or memory elements through high-bandwidth, non-interfering optical channels. An example of such a device is the Si/PLZT SLM at UC-San Diego. Not only can these devices be fabricated to function as logic planes, but also as two-dimensional detector planes with individually-addressable elements - another high priority need for optical computers.

Next in line of importance are the reconfigurable interconnect devices. For the course-grained systems, the need is for larger dimensional crossbar switches (32 x 32 and 64 x 64) and faster reconfiguration times (< 1 μsec). Reconfigurable interconnect devices are not available for the fine-grained systems. For the multiple selection approach mentioned above, there are currently two devices under investigation - one at Optron and the other at the University of Colorado. Since these approaches represent only a small sampling of what can probably be done in this area, additional ideas will be most welcomed. For the modifiable approach, the holograms could be formed in a photorefractive crystal; however, there are numerous problems at the material level, such as speed of response (1 ms or less needed at semiconductor laser power levels), control over the decay of recorded holograms as new holograms are written, and a reduction in beam fanning within the crystals. Research into other possibilities for storing and rewriting holograms in real-time could have large payoffs for reconfigurable interconnects.

With regard to material requirements, the most pressing needs are for better electro-optic materials for SLM applications (a high electro-optic coefficient with a low dielectric constant - e.g., $n^3r/c > 500$ pm/V), more sensitive photorefractive materials (mobility lifetime product > 10^{-5} cm^2/V), and material uniformity over an aperture size of at least 0.5 x 0.5 cm^2. The speed of response requirement for most envisioned applications is < 1 ms for photorefractive materials and < 1 μs for electro-optic materials for SLM applications. Organic polymers are being given serious consideration by several research groups for electro-optic applications.

Research is also needed at this time on further term goals of systems employing optical logic planes as discussed above. The search is underway for good nonlinear optical materials (third order) that could be used inside etalons to yield three terminal devices for logic plane applications. Such materials may also find application in wave-mixers; e.g., four-wave mixers for dynamic holography that could be used as a reconfigurable interconnect device. Promising materials here are organic polymers and semiconductor clusters in polymers and glasses (quantum dot effects). A yet unexplored area involves the search for the photorefractive effect in organic polymers.

THE ENERGETIC ADVANTAGE OF ANALOG OVER DIGITAL COMPUTING

H. J. Caulfield
Center for Applied Optics
The University of Alabama in Huntsville
Huntsville, AL 35899

I. INTRODUCTION

Our goal is to examine the energetics of computation for analog and digital computers. A careful analysis of digital computers has already been performed (1). That analysis suggests that each computation must have enough energy to overcome thermal noise. That is the energy per digital calculation, E_D', must satisfy

$$E_D' > kT ,$$

where k is Boltzmann's constant and T is the operating temperature. A second constraint is that the number, N_D, of events detected be large enough to give an acceptable probability that the (usually binary) decision have satisfactory certainty. Both constraints must be satisfied for both optics and electronics. For electronics, the energy per event is

$$E_E = eV,$$

where e is the electron charge and V is the accelerating voltage. For optics, the energy per event is

$$E_o = h\upsilon,$$

where h is Planck's constant and υ is the frequency. For their normally used values

$$eV << kT$$

and

$$h\upsilon \sim 10\ kT,$$

Thus, if we require M events per decision, electronics requires

$$\epsilon_E = ME_E = \text{MeV}$$

which can, in principle, approach kT. Current systems operate far away from this minimum, e.g.

$$\epsilon_E \sim 10^4 \text{ kT}.$$

For optics, we require

$$\epsilon_o = Mh\upsilon \sim 10MkT.$$

The best current optics also operate at

$$\epsilon_o \sim 10^4 kT.$$

The difference, however, is that digital optics has no room left for energetic improvement, while digital electronics can still improve by orders of magnitude in principle.

II. ANALOG COMPUTERS

We are concerned with a particular type of analog computer wherein a source of light or electrons illuminates a prepared passive apparatus which rearranges and "processes" the signal linearly. Thereafter the processed signals are subject to one or more nonlinear decision operations. In optics, some examples are Fourier optical pattern recognition [2] and massively interconnected optical neural networks (3).

Two observations must be made. First, the system consists of a complex linear operation followed by nonlinear decision making. Only the nonlinear decisions are subject to the analysis of Section I. Second, each detected event is a quantum mechanical measurement conditioned by the entire apparatus, however complicated.

Consider, for example, an optical neural network of this type which makes P weighted interconnections to each output position. Each photon event detected there was conditioned by the apparatus making all P

interconnections. In effect, each photon makes P calculations. The energy calculation, therefore, is

$$E'_0 \simeq Mh\upsilon/P.$$

For realistic cases

$$Mh\upsilon \sim 10^4 kT$$

and

$$P \sim 10^6.$$

Therefore

$$E'_0 \sim kT/100.$$

The important observation is that we can achieve

$$E'_0 << E'_E.$$

In other words, analog optics can have a significant energy advantage over digital electronics. Furthermore, there appears to be no fundamental thermodynamic minimum to the energy per calculation of this type of analog processor.

III. ANALYSIS

The only previous attempt we know of to design processors operating at less than kT per operation is the conservative logic (Fredkin) gate (4). It is interesting to note that it can be formulated in a way very similar to our analog processor. There are many "conservative logic gates" where no minimum energy need be expended followed by a nonlinear decision operation. Each decision is the result of many calculations.

It appears that it is important to distinguish between decisions (nonlinear) and calculations (which <u>can</u> be passive or conservative). In digital computers all computations are comprised of decisions. In some analog systems and in conservative logic systems, many calculations can be

made and subsequently probed by a single decision. The number of calculations, C, and the number of decisions, D, are related by

$$R = C/D.$$

For digital systems E=1 and

$$E \tilde{>} RkT = kT,$$

For some analog systems and some conservative logic systems, we have E << 1 and

$$E' \tilde{>} RkT << kT$$

and thus

$$E' << kT.$$

REFERENCES

dauer, "Irreversibility and Heat Generation in the Computing " IBM J. Res. Dev. 5, 183 (1961).

rsasent, "Scene Analysis Research: Optical Pattern Recognition cial Intelligence," Proc. SPIE 634, 439 (1986).

3. H. J. Caulfield, "Parallel N^4 Weighted Optical Interconnections," Appl. Opt. 26, 4039 (1987).

4. E Fredkin and T. Toffoli, "Conservative Logic," Int. J. Theoret. Phys. 21, 219 (1982).

NOTES

TUESDAY, FEBRUARY 28, 1989

SALON F

2:45 PM–3:30 PM

TuF1–TuF3

OPTICAL COMPUTING SYSTEMS: 2

George Eichmann, CUNY-City College, *Presider*

Tantalus and Optical Computing

W. Thomas Cathey
Center for Optoelectronic Computing Systems and
Department of Electrical and Computer Engineering
University of Colorado
Boulder, Colorado 80309-0525

The ancient Greek, Tantalus, had a problem that is similar to that of some optical computing and signal processor researchers. As a punishment by the Greek gods, Tantalus was placed in a lake with water up to his waist. Fruit was on branches just above his head. However, when he leaned over to drink, or reached up to eat, the water receded *just* beyond reach and the fruit evaded his grasp. Hence, he was doomed to never obtaining what seemed so close. I do not know what deed was done by the optics researchers that deserves similar punishment (I have some ideas.), but some of the anticipated results and applications of optics to computing seem to be always just barely beyond reach.

In this paper, I explore some of the promises of optics. Areas where optics seems to be able to make a contribution are examined, along with the barriers to success, and some potential solutions. The purpose is to help us to focus on the blocking technologies and to see where more effort is needed. First, some of the tantalizers or attractors are discussed. There are also drivers or clear needs in the areas of communications and computing. The problems or barriers, in some cases, are obvious, but some are not. Finally, potential solutions are delineated. Some may be tantalizing, but not achievable, but we can attempt to decide which are not in that category.

The tantalizers in the field of computing and communications include very short optical pulses (with the unreachable possibility of a high duty rate?) and high capacity fibers. These would, of course, be very useful in the construction of optical computers. Other tantalizers in computing include massively parallel computing and high-capacity optical storage of data. The ease of impedance matching leads to thoughts of high fan-out and fan-in of connections.

In addition to the tantalizing possibilities, some very real needs exist as drivers. Fiber-based

communications networks have electronic switching control systems. This demands a photon/electron/photon conversion at each switch. High-definition television and video teleconferencing require broadband switches. The need for higher capacity computing systems exists in the fields of high energy physics, weather prediction, fluid dynamics, and problems mimicking the reasoning capability of humans.

The barriers to achieving the goals of using femtosecond pulses, high capacity fibers, massive parallelism and high-capacity data storage in high-capacity computers and broad-band optical networks and switches fall both in the areas of devices and architectures. Some of the device limits relate to the speed of electronic detection and the limits on direct laser modulation rates. The conversion from photons to electrons and back greatly reduce the potential speeds of optical systems. Multiplexers can be used to interleave several optical signals to effectively use the fiber bandwidth, but effective ways of demultiplexing the signals without requiring synchronization must be found. If optical cross bars could be controlled optically, the need to do relatively slow photon / electron conversions would be reduced. For fan-in of several signals using spatial light modulators (SLM) where the number of inputs must be preserved, a very high contrast ratio is needed. If the output of the SLM is to go to an OR gate, there must be a tight tolerance on the variation of transmission levels in the on and off states. For fan-out, amplification or high source power levels are required. For systems requiring regeneration, high-speed optical amplifiers are necessary. Low power or very high speed devices are needed. In data storage, the demonstrated storage densities of holograms is several orders of magnitude lower than what is often quoted as the theoretical limit. Finally, the architectures that have been developed for electronic computers are not applicable to high-speed optical processors where the propagation time between gates is not negligible with respect to the gate delay. Nor are the old architectures applicable to massively parallel optically interconnected computers, be they electronic or optical computers.

There are several possible solutions and alternatives to these barriers. One is to use photon - electronic wavefunction interactions where possible, eliminating the need to do photon - moving electron conversions. This may be possible with semiconductors or organic materials. If orbital electrons are used to mediate the interactions rather than generating free electrons, the interactions can be faster. Devices and/or architectures for demultiplexers may be possible using tunable lasers and filters or with new ideas for separating time multiplexed signals. Research on photo-addressed optical crossbars

may lead to new ways of switching light without converting to electrons first. Research on the storage mechanisms of volume holograms may lead to ways to increase the number of non-planar holograms that can be stored in a medium for data storage or associative memory.

There is room for major research in the areas of architectures for optical or optoelectronic computers. These must be demonstrated, not simply proposed, architectures. The fact that one can not stop and store photonic signals as one can stop and store electronic signals must be taken into account in the architectures. Other algorithms and architectures, such as image algebra, may be useful for the design and analysis of massively parallel optical processors. In any architecture, either the short pulses and high speeds of optics, or the third dimension and parallelism must be incorporated. Ultimately, there will be a need for both.

The achievement of some of the earlier tantalizing possibilities are now recognized as not being very likely. At least, not likely with the approaches previously used. For example, the recognition of camouflaged vehicles in aerial reconnaissance photographs can not, after 25 years of work, be done using spatial filters. It may, however, be possible using other techniques such as optical neural networks trained to perform pattern recognition. Optical holographic storage of data still has not been practical. Two-dimensional optical disks, on the other hand, have been widely used. Optical data storage techniques using optical associative memory may prove practical or they may be one of the tantalizing, never reached, goals. Optical supercomputers remain on the far horizon, but simple processing of optical bit serial data is imminent. It is the job of device and systems researchers in optics to determine which of the goals are reachable and which are merely tantalizing.

The Mock Counter

Ann B. Yadlowsky and Harry F. Jordan
Optoelectronic Computing Systems Center
University of Colorado
Boulder, CO 80309-0525

Summary

The goal of the *Bit-Serial Optical Computer* project is to build the first all optical stored program digital computer. Because optical pulses may be extremely short and do not interfere, such a machine can potentially operate at very high speeds. The first step on the way to such a machine is a simple optical finite-state machine, the optical counter. This is an optical, bit-serial counter built using the technology to be employed in the Bit-Serial Optical Computer [1]. The *Mock Counter* [2] is a step on the way to the optical counter.

The current technology of optical devices resembles that of the devices used in early electronic computers of the late 1940's. There are few active optical logic devices and they are difficult to work with. The lithium niobate ($LiNbO_3$) directional coupler [3], logically represented in Fig. 1, has been chosen as the main logical element for this project because it is one of the few commercially available optical logical devices. This choice allows us to design and build interesting and technologically relevant optical architectures without waiting for the technology to produce a more optimal optical component.

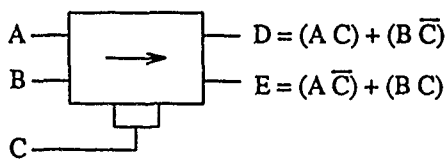


Figure 1: $LiNbO_3$ Switch as a Logic Element

This approach also means that we must work with what a digital electronic computer engineer would consider primitive analog devices. For example, because of the loss and crosstalk in each device, the optical signals must be digitized and regenerated after going through only a few switches. Also, the optical inputs to the directional coupler must be properly polarized to provide optimal switching. In addition, an electronic signal is required to switch the device. To perform logic on optical signals, this input must be converted to an electronic signal, shaped, and amplified to drive the device. These problems and high cost make it important to use as few switches as possible to perform the necessary logic.

Many switches are saved by operating bit-serially, as was done in early electronic computers [4]. In addition, expensive flip-flop memories are replaced, in this design, with a fiber delay-line memory. Three dB fiber couplers (devices which mix two incoming optical signals and then equally split them between the two outputs) save switches by providing fan-out and by acting as inexpensive optical OR gates.

The bit-serial counter is implemented logically as a half adder with synchronized feedback and an increment at the beginning of each K-bit word, as shown in Fig. 2. The sum bit is delayed by ΔK (K bit times) and the carry bit by Δ1 (one bit period) so that the carry from the previous bit is ORed with the increment bit (WCK) and then is added to the sum from the previous word.

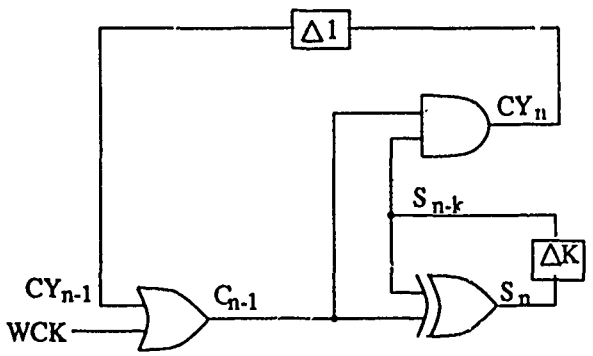


Figure 2: Bit-Serial Counter Logic

Fig. 3 shows the logic implementation using directional couplers, 3 dB couplers, and optical fiber memory. A 3 dB coupler implements the OR of WCK and CY_{n-1}, and the upper switch (SW4) implements the AND of C_{n-1} and S_{n-k}. SW3 provides regeneration for the carry loop. The carry is detected and amplified at input C. If the carry is high, a fresh clock pulse is switched into the loop at output D. SW3 also produces the complement of the carry at output E which is used by SW5 to generate the EXCLUSIVE-OR function. Fan-out is provided by 3 dB couplers, and fiber provides the necessary interconnections and the required one-bit and K-bit delays.

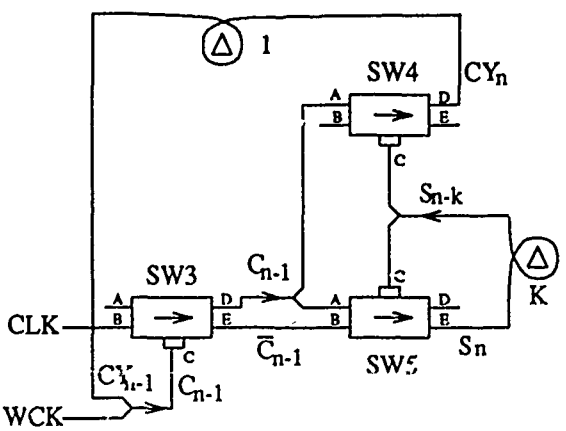


Figure 3: Bit-Serial Optical Counter Design

In Figs. 2 and 3, the delays are shown only where needed for logical delays or storage. In a real system, however, there are delays associated with every connection and every element.

Since pulses are not latched into flip-flops for synchronization, the system must be designed so that the inputs to each switch arrive at the same time. Such synchronization can be achieved by distributing the required delays for a path or loop among all of its components. Paths with no feedback may be lengthened as much as required, except that all parallel paths must be the same total length.

If the path lengths are not exactly right, one pulse may reach the switch early, while the other may be a little late. If these signals are ANDed together, the output is a shorter pulse. After this pulse travels around the loop many times, each time getting shorter, there is no signal left. If an early signal and the complement of a late signal are ANDed together, glitches are produced. To avoid pulse shortening and glitches, pulse stretching is used at input C. The fiber is cut so that the pulses arrive at C a small percent of a pulse time ahead of any early inputs. The electronics of the receiver-amplifier then stretch the pulse so that it remains after the exact pulse length is over to account for late inputs.

The challenges of building the bit-serial counter with $LiNbO_3$ directional couplers arise both from device problems and synchronization issues. The Mock Counter project started before we had the equipment or the experience to build the counter with directional couplers. "Mock" switches were built which receive the optical signals, convert them to digital logic levels, perform the proper logic, and transmit optical outputs. The mock switch acts like an ideal directional coupler switch. The electronic implementation has negligible crosstalk and regenerates the signal at each switch so that there is no attenuation. Also, all the inputs are detected by PIN diodes so that the input polarization does not matter. Building the Mock Counter allowed us to develop a procedure for building and synchronizing the counter system. It also gave us experience with optical fibers, connectors, optoelectronic components, and relatively high speed electronics.

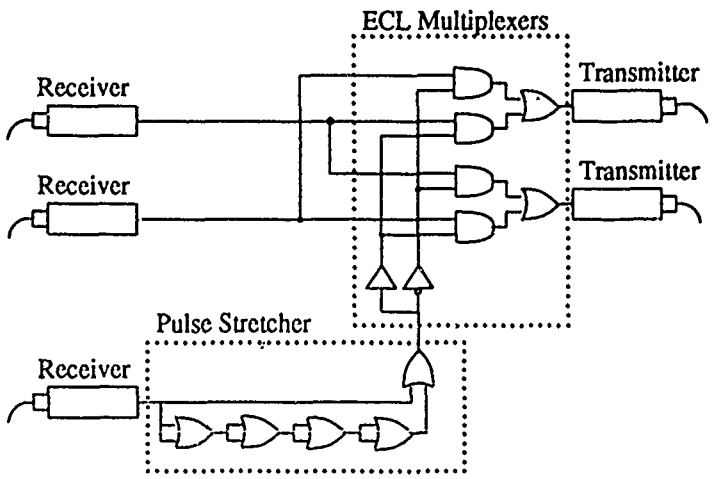


Figure 4: The Mock Switch

The details of the mock switch are shown in Fig. 4. The optical-electronic interface consists of AT&T ODL 200 Lightwave Data Link components. The logic is performed by two

ECL multiplexers. Pulse stretching is accomplished by ORing the C input pulse with a delayed version of itself. The clock for this system is a crystal oscillator driving an ODL 200 transmitter. The word clock for the incrementing is produced by dividing the clock by the word length electronically. In the directional coupler counter, the word clock will be produced optically using 2 $LiNbO_3$ directional couplers.

The first step in building and testing the Mock Counter was to measure the pulse stretch and the propagation delay of each of the mock switches. The switch parameters were used in a simulation program [5] to determine approximate fiber lengths for the system. The counter system was then put together one switch or delay loop at a time and synchronized at each step. The synchronization was done mainly by adjusting fiber lengths to provide the correct delay at each point in the system.

While this working counter serves as a good demonstration and proof of principle system, its main importance lies in the knowledge and experence gained in its construction. The procedure for assembling a bit-serial counter, developed for construction of the Mock Counter, will be extremely important in building the directional coupler counter. The advantage of the procedure is that it allows complete synchronization of each switch and feedback loop in the system without requiring intrusive observation of optical signals. Also, this procedure only requires input C to be held at a 1 or a 0 electronically; it does not require optical inputs to be held at high values. The only optical input is the clock.

Construction of the Mock Counter has successfully demonstrated a feedback state machine using fiber delay-lines for all information storage. It has established a step-by-step procedure for assembling and testing a $LiNbO_3$ directional coupler version of the counter that will operate at higher speeds. Work on constructing and testing the latter version of the counter, in which both switching and storage are optical, is currently in progress.

References

[1] V. P. Heuring, H. F. Jordan, and J. P. Pratt, "A Bit Serial Architecture for Optical Computing," *OCS Technical Report 88-01*, Optoelectronic Computing Systems Center, Univ. of Colo., Boulder, CO 80309-0525, March 1988.

[2] A. B. Yadlowsky, "The Mock Counter: A Hybrid Optical-Electronic Counter Using Fiber Delay Line Memory," *OCS Technical Report 88-04*, Optoelectronic Computing Systems Center, Univ. of Colo., Boulder, CO 80309-0525, September 1988.

[3] S. K. Korotky and R. C. Alferness, "Waveguide Electro-optic Devices for Optical Fiber Communication," in I. P. Kaminow and S. E. Miller, eds., *Optical Fiber Telecommunications*, Ch. 11, Academic Press, 2nd ed., to be published.

[4] S. P. Frankel, "The Logical Design of a Simple General Purpose Computer," *IRE Transactions on Electronic Computers*, March 1957, pp. 5-14.

[5] J. P. Pratt, private communication.

TuF3-1

CASCADE CONNECTIVE OPTICAL LOGIC PROCESSOR USING 2-DIMENSIONAL ELECTRO-PHOTONIC DEVICES

S.KAWAI, Y.TASHIRO, H.ICHINOSE, K.KASAHARA and K.KUBOTA

Opto-Electronics Research Laboratories, NEC Corporation
1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki-shi 213, Japan

Various approaches to achieve digital optical processing have been proposed[1]-[3]. In order to achieve general processing, it is necessary to execute various kinds of logic operations. Moreover, individual processors must be connected each other. Few processors, however, have been previously reported with data representation suitable for cascade connection. Most of the processors proposed so far require pre-coding process, which makes hardwares complicated. If optical output signals from a processor are to be used as optical input signals for a next stage processor, pre-coding process is not necessary, which results in the realization of cascade connectable processing systems.

The authors present a new logic algorithm for cascade connective processor, which is confirmed using 2-dimensional electro-photonic devices, called vertical to surface transmission electro-photonic devices (VSTEPs)[4],[5] and a ferro-electric liquid crystal spatial light modulator (FLC-SLM). The VSTEP has three functions, which are light emittion, photo-detection and thresholding, that is, the device emits light (LED mode) caused by incident light, when the incident beam intensity exceeds more than a specific level.

Figure 1 shows the proposed logic algorithm. Logic signals A,B and their complements $\overline{A}$,$\overline{B}$ are prepared for processing inputs. These four signals are multiplied individually by weighting factor W with three levels, according to logic category involved. Weighted signals are summed up and digitalized. Output signals C and their complements $\overline{C}$ are given by

$$C(A,B)=T(W_A*A+W_B*B+W_{\overline{A}}*\overline{A}+W_{\overline{B}}*\overline{B}-Th) \qquad (1)$$

and

$$\overline{C}(A,B)=T(\overline{W}_A*A+\overline{W}_B*B+\overline{W}_{\overline{A}}*\overline{A}+\overline{W}_{\overline{B}}*\overline{B}-Th), \qquad (2)$$

where $T(x)$ is a step function, and Th is a fixed thresholding level and is set at 0.75. W_A, W_B, $W_{\overline{A}}$, $W_{\overline{B}}$, $\overline{W}_A$, $\overline{W}_B$, $\overline{W}_{\overline{A}}$ and $\overline{W}_{\overline{B}}$ are weighting factors. For example, in the case of NOR logic operations, W_A and W_B are set at 0, and $W_{\overline{A}}$ and $W_{\overline{B}}$ are set at 0.5. If both A and B are 0 ,C(A,B) becomes 1. In other cases, C(A,B) becomes 0, then, producing correct answer for NOR logic operation. Complemental logic result $\overline{C}$ is obtained similarly using Eq.(2). In this algorithm, the data representation for the output signal C is equal to that for the input signals A or B. Thus, output signals can be used as input signals for next stage processor. Changing the W and $\overline{W}$ values, 14 kinds of logic operations are carried out within the processor, except EXOR and EXNOR. EXOR and EXNOR can be carried out by two stage processing using the two cascade connected processors.

The optics for the processor is shown in Fig.2. Outgoing light beams from light sources ($VSTEP_1$), A,B and $\overline{A}$,$\overline{B}$, are divided into two parts by optical branches, transmitted through operating masks and focused on optical thresholding devices ($VSTEP_2$). Lens $array_1$ collimates outgoing beams from light sources. Multiplying values W are changed by operating mask transmission. Lens $array_2$ performs the function of summing up transmitted light beams and of injecting then into $VSTEP_2$ as shown in Eq.(1) or (2). Digitalized optical outputs are obtained through $VSTEP_2$. The $VSTEP_2$ can be used as light sources for the next stage processor, allowing, thus, cascade connection.

Two-dimensional AlGaAs/GaAs VSTEP arrays (2x2) were used for $VSTEP_1$ and $VSTEP_2$. Individual element of VSTEP array were arranged at 0.25mm pitch. FLC-SLM array (4x4) was used for operating mask, whose pitch was as same as that of VSTEP arrays. Figure 3 shows the processing module (the size is 100mm x 50mm x 50mm). The module was constructed from basic elements in Fig.2,

except complimental logic processing. The SLM transmission was 75% at light emission wavelength of VSTEP, 870nm, when no voltage was apllied to the SLM. Fourteen levels of contrast were obtained. Planar micro lens array with 0.5mm focal length and 0.25mm pitch[6] was used as collimating lens array. Figure 4 indicates the operation mask patterns for NOR and OR operation. The upper half part of FLC-SLM mask transmission corresponds to W_A and $W_{\overline{A}}$, while the lower half part corresponds to W_B and $W_{\overline{B}}$. For NOR operation, $W_A=0$, $W_B=0$, $W_{\overline{A}}=0.5$ and $W_{\overline{B}}=0.5$. For OR operation, $W_A=1$, $W_B=1$, $W_{\overline{A}}=0$ and $W_{\overline{B}}=0$. Changing operating mask pattern like this, 14 variable kinds of logic operations are accomplished. Wollaston prism can be used as optical branches for performing complimental logic operations.

Experiments on logic operations were carried out. The FLC-SLM frequency characteristics is shown in Fig.5. Figure 6 shows the result on NOR operation. Fourteen kinds of logic operations were completely achieved. Operation changing speed is limited by the frequency response of the SLM. VSTEPs operates at the rate of 400Mbps and data can be transfered at this rate.

An optical processor with cascade connectability was proposed and successfully demonstrated. If these processors were connected in series, a pipe-line processor could be obtained, and image processors or data flow machines would be also realized.

The authors gratefully thank Drs. E.Okuda and M.Oikawa in Nippon Sheet Glass Co., Ltd. for supplying the planar micro lens array used in these experiments. They would also like to thank Drs. M.Sakaguchi, N.Nishida, R.Lang and K.Yanase for their suggestions and encouragement.

References

1)J.Tanida and Y.Ichioka, J.Opt.Soc.Am. 73, 800 (1983)
2)T.Yatagai, Opt.Lett. 11, 260 (1986)
3)K.H.Brenner et al., Appl.Opt. 25, 3054 (1986)
4)K.Kasahara et al., Appl.Phys.Lett. 52, 679 (1988)
5)Y.Tashiro et al., Appl.Phys.Lett., to be published
6)Y.Ikeda et al., Proc. 5th EFOC/LAN'87, p.103 (1987)

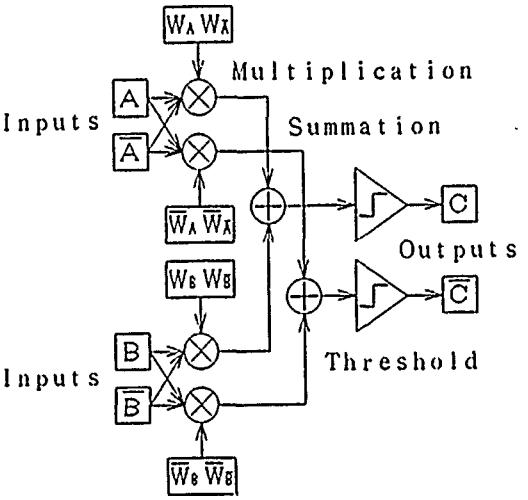


Fig. 1 Processing algorithm

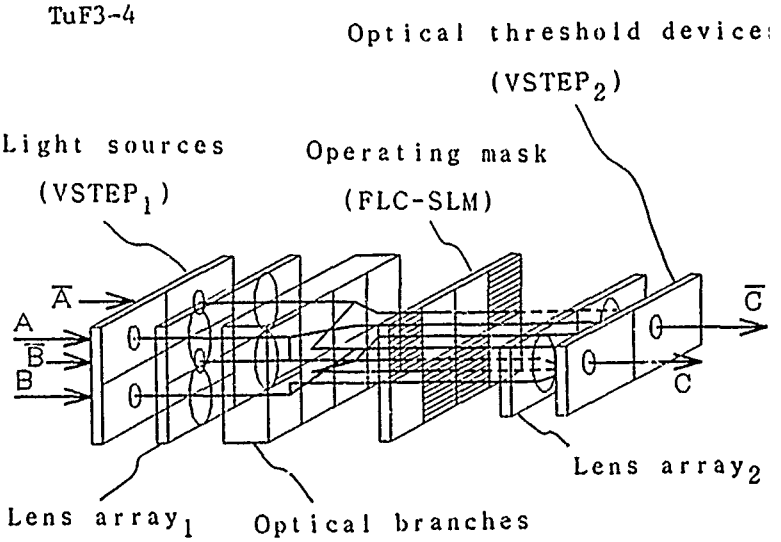


Fig. 2 Processor optics

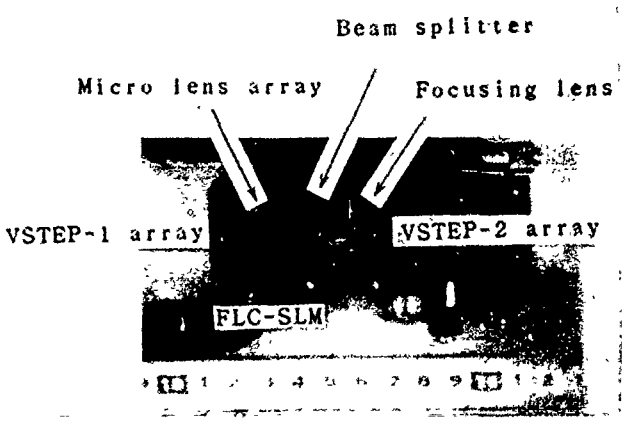


Fig. 3 Processing module

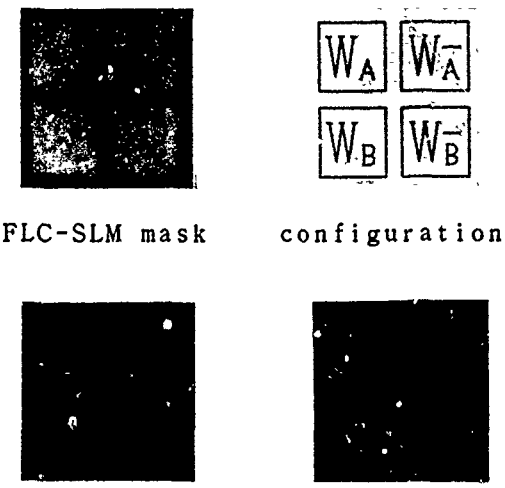


Fig. 4 Operating mask pattern

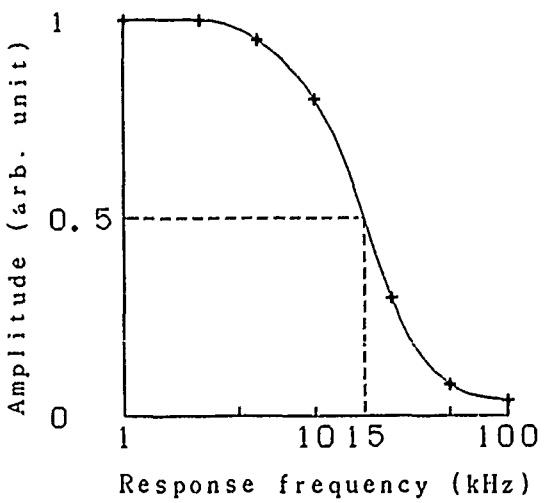


Fig. 5 FLC-SLM pulse response

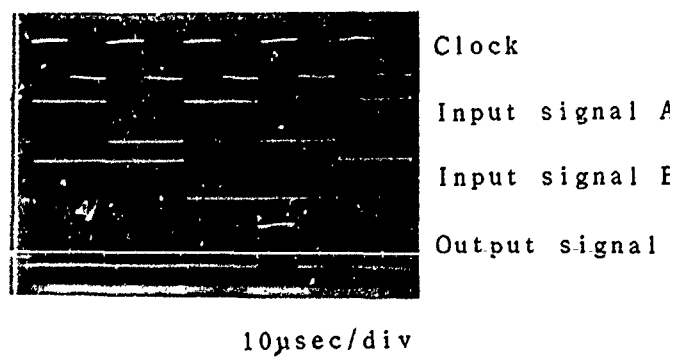


Fig. 6 NOR logic results

TUESDAY, FEBRUARY 28, 1989

SALON F

4:00 PM–5:00 PM

TuG1–TuG4

OPTICAL COMPUTING SYSTEMS: 3

William Miceli, U.S. Office of Naval Research,
Presider

Computational Origami – The Folding of Circuits and Systems

A. Huang
Room 4G-514, AT&T Bell Laboratories
Crawfords Corner Road
Holmdel, NJ 07733

Abstract

A technique which regularizes and folds circuits and systems to match the parallelism of optics is presented.

1 Computational Origami

Computational origami[1,2] involves the reformatting of computations. It takes a computation, regularizes[3] it, and then folds it into a format which is more suitable for processing. Any computer can be decomposable into state machines. A state machine can be partitioned into a combinatoric logic and latches. Any logic circuit can be recast into NOR gates with a fixed fanin and fanout. Such a circuit as shown in Figure 1 can be "regularized", and cast into a regular array as shown in Figure 2. Only four types of modules are needed to implement this or any circuit; a NOR gate, a crossover, a bypass, and a broadcast. Assume that instead of being fixed each module could be dynamically programmed to assume one of these functions. This circuit can then be folded down into the circuit shown in Figure 3.

Each Module has two inputs, two outputs, and a control input. The "A" function represents a NOR gate, "B" represents a crossover, "C" represents a broadcast, "D" represents a bypass, and "Z" represents a don't care condition which could be any of the other functions. The delay and multiplexer elements are also shown. For ease of explanation the propagation time through the modules is assumed to be instantaneous. During the first cycle the modules emulate the top two modules in the first column of Figure 2. The multiplexers connect two "z" (don't care) inputs to the inputs of the top module. The left output of the first module is fed to the bottom module while the right output is fed to a delay element which delays this signal for one unit of time and feeds it to the left input of the bottom module. The left output of the bottom unit is fed via a 5 cycle delay to the multiplexer attached to the left input of the top module. The right output of the lower module is fed via a 6 cycle delay to the multiplexer attached to the right input of the top module. During the second cycle the modules emulate the top two modules in the second column of Figure 2. The multiplexers then feed "a" and "b" input to the top module. The delayed right output of the top module fed to the left input of the bottom module. The outputs of the bottom module are then fed to the appropriate delay lines. This continues until after the sixth cycle at which time the multiplexers cut off the inputs and instead feed the outputs of the delay lines to the inputs of the top module.

Computational origami raster scans a hardware window across a circuit. The use of a CPU can also be viewed as passing a hardware window across a problem however in this case the CPU jumps in a directed manner over a problem rather than scanning it. The directed jumps seems to be more efficient in that it only visits the hot spots however this approach becomes less efficient the more parallelism there is in a problem. The directed jump approach suffers from the fact that it is difficult to introduce a second CPU since each CPU would have to worry about stepping on the other's tail whereas the scanning window can be made larger or shadowed by another hardware window.

2 The Relevance of Computational Origami to Optical Computing

Figure 4 shows arrays of optical logic gates interconnected via crossover networks to other arrays of optical logic gates. Techniques have been presented showing how it is possible to cut various interconnections via masks to implement various circuits.[4] This would provide a fixed circuit much like a printed circuit board does. The processor would be more powerful if we could dynamically alter the flow of information. Rather than altering the interconnections we could also tie some of the NOR gates high by injecting an input and forcing their output low. This would kill the flow of information along this path. With optics we have the capability of injecting such inputs, however which gates should we turn off. Computational origami answers this question. The hardware window corresponds to an array of optical PLA's (programmable logic array).

3 Computational Origami from a Computer Perspective

- The use of delay lines for memory is reminiscent of the old mercury delay line or magnetic drum based computers.

- The processor shown in Figure 3 can be viewed as the ultimate RISC (reduced instruction set computer) since it has only four basic instructions (A, B, C, and D).
- The processor can also be viewed as two linked Turing machines.
- The processor can also be viewed as a Boolean array processor.
- The processor is a MIMD (multiple instruction multiple data) machine.
- The technique can also be scaled up and applied to coordinate arrays of computers.
- This technique can provide a tunable amount of parallelism from bit serial to as parallel as the problem is.

4 Computational Origami from an Optics Perspective

- Provides an architecture for both a fast serial and a slow paraliel approach towards optical logic.
- The use of delay lines for memory is amenable to optics.
- It provides an initial design which requires very little hardware.
- The design is regular and expandable.
- It provides a means for harnessing the power of Symbolic Substitution.

5 Conclusion

Computational origami is a superset of several proposed optical digital computer architectures. It started as a means for efficiently programming Symbolic Substitution machines.[5] It provides a means of linking the Boolean equation machines proposed by Guilfoyle[6] or the electro-optical hybrid PLA processors proposed by Feldman.[7] It also provides a means of converting any processor into a bit serial implementation as proposed by Jordan.[8]

Computational origami will be initially used by the electronics community. The processors will eventually be scaled up to a point at which electronics will not be able to support the degree of parallelism.

References

[1] Huang, A., "Computational Origami," U.S. Patent Pending.

[2] Lu, H.-M., "Computational Origami: A Geometric Approach to Regular Multiprocessing," MIT Master's Thesis in Electrical Engineering, (May 1988).

[3] A. Huang "Architectural Considerations Involved in the Design of an Optical Digital Computer" *Proceedings of the IEEE* 72, No. 7, (July 1984).

[4] Murdocca, M. J., A. Huang, J. Jahns, and N. Streibl, "Optical design of programmable logic arrays " *Appl. Opt.*, 27, (9), (May 1, 1988).

[5] Murdocca, M. J. and A. Huang, "Symbolic substitution methods for optical computing," Proceedings of the I.C.O. Meeting on Optical Computing, Toulon, France, (1988).

[6] Guilfoyle, P. S. and W. J. Wiley, "Combinatorial logic based digital optical computing architectures," *Appl. Opt.*, 27, 1661, (May 1, 1988).

[7] Feldman, M. R., S. C. Esener, C. C. Guest, and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations," *Appl. Opt.*, 27, 1742, (May 1, 1988).

[8] V.P. Heuring, H.F. Jordan, and J.P. Pratt, "A Bit Serial Architecture for Optical Computing," *Proceedings of the I.C.O. Topical Meeting on Optical Computing 88* Toulon, France, (1988).

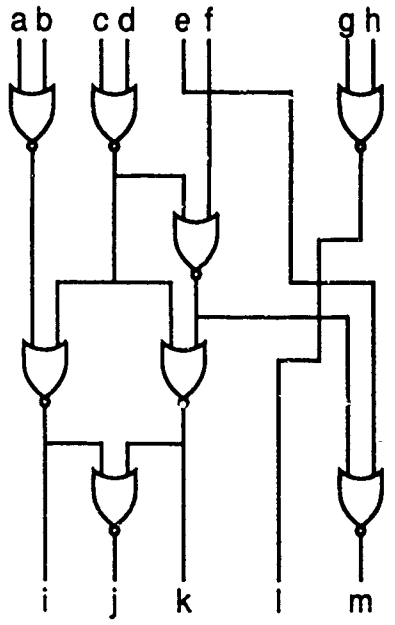


Figure 1: An arbitrary circuit with fixed fanin and fanout.

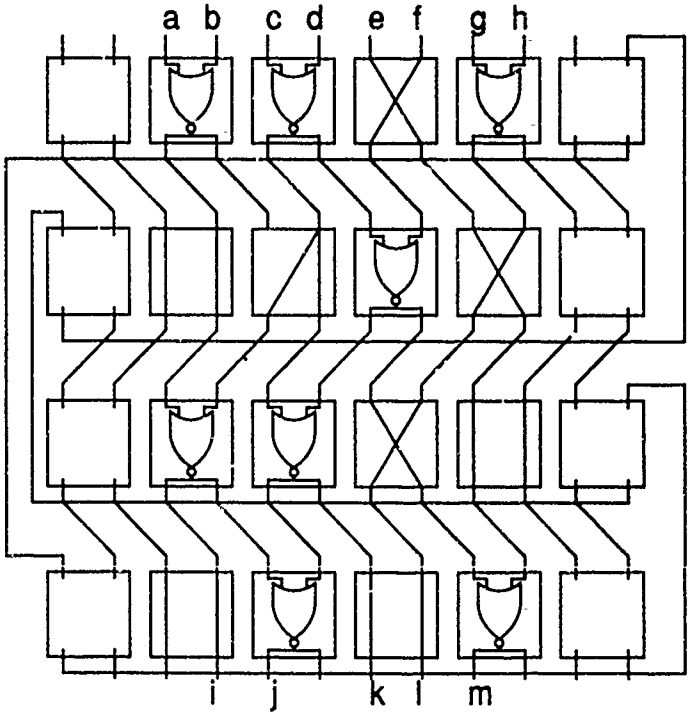


Figure 2: A regularized circuit.

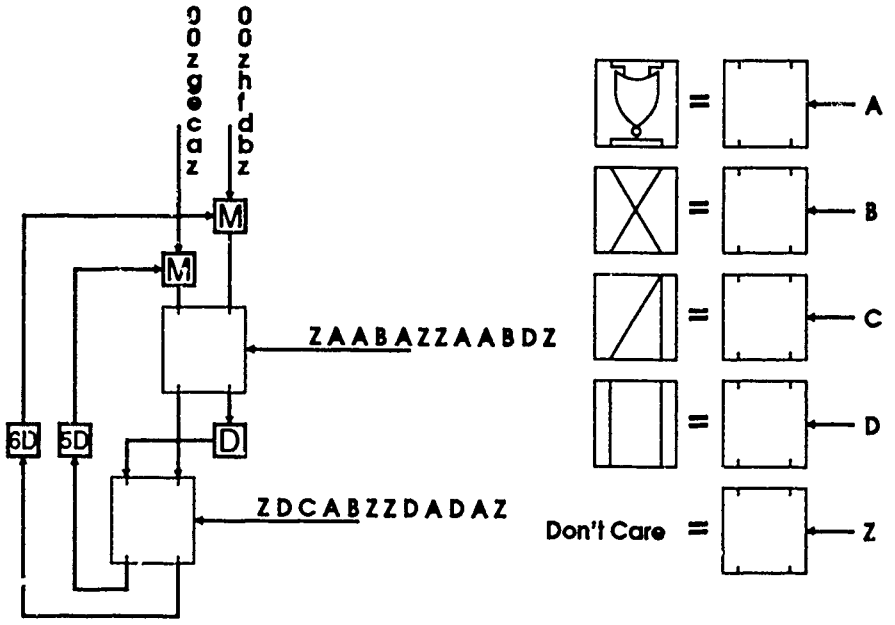


Figure 3: A folded circuit.

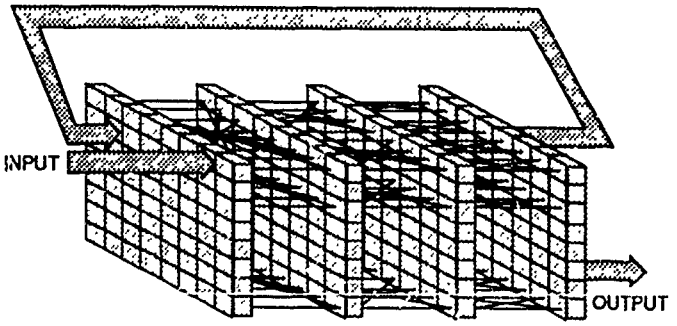


Figure 4: The architecture for an optical digital computer.

All-Optical "Game of Life" Computer

Lawrence H. Domash
Foster-Miller, Inc.
350 Second Ave.
Waltham, MA 02254

Mark Cronin-Golomb
Electro-Optics Technology Center
Tufts University
Medford, MA 02155

Introduction

Although a variety of optical computing architectures have been analyzed and many individual devices reported, to date few working optical computers have been demonstrated performing a complete, self-contained, recognizable computational function. A conceptually simple but significant computing architecture which appears accessible to complete all-optical implementation is the cellular automaton (CA). Murdocca discussed the CA in comparison with symbolic substitution (1). Taboury et al showed how general cellular processors could be realized using sets of holograms (2).

Cellular automata are significant architectures because some examples are known to possess computational universality. We report an analysis of a particular 2D CA known to have the property of computational universality, Conway's Game of Life, and propose a fully parallel implementation using a chain of cascaded photorefractive processing devices. This CA architecture seems particularly well adapted to utilize the natural strengths of optical methods by combining aspects of image processing, Fourier optics, and nonlinear dynamics in a massively parallel digital optical environment. Demonstration of a complete self-iterating system would be complex but appears to be within the state of the photorefractive art.

Cellular Automata

A simple 2D CA consists of a plane of multivalued cells. The cell values are updated in a sequence of discrete time steps, according to a rule which defines the value of a cell in the next time step as some function of its current value and that of its neighbors within a finite distance. Simple rules can lead to surprisingly rich and complex global behavior. Cellular automata have been of increasing interest in recent years because of their computational universality and also their importance for simulating scientific phenomena ranging from fluid flow to cell biology in which simple interactions among many similar elements lead to complex macroscopic behavior (3,4).

A specific CA which has been the subject of detailed study is the "Game of Life" invented by Conway (5). Life consists of a 2D pattern of binary cells (on/off) whose updating is defined by a three-part rule. **Birth:** an off cell turns on only in the case exactly three neighbors are on. **Survival:** an on cell remains on if has either two or three neighbors on. **Death:** an on cell dies (turns off) from isolation if it has fewer than two neighbors on or from overcrowding if it has more than three neighbors on.

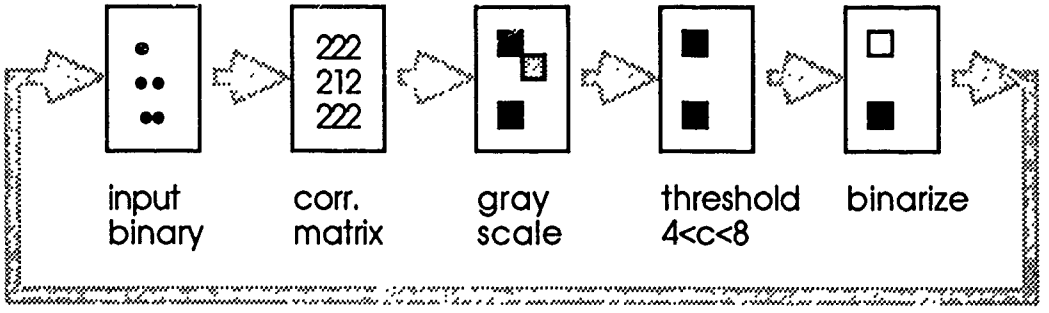
When applied to an initial input pattern to generate a sequence of updates, these rules give rise to a variety of intriguing "Life forms" including periodic structures, chaos, and emergent persistent correlations propagating in the plane. The figure shows a "glider gun" which, starting from an initial state of 26 cells, emits endless streams of "gliders," 5 cell propagating subpatterns able to play the role of clock pulses. Life has been proven to possess the property of computational universality (3). An all-optical Game of Life processor could serve in principle as a microcode basis for any higher level algorithm; programming would be formulated such that any computational problem is coded as an input Life pattern presented to the processor for evolution through a set number of updates into an output pattern.

Optical Formulation of the Game of Life

The Life rules may be recast in a form which makes their nonlinear image processing aspects more evident. The initial state lattice of Life cells Ln(i,j) is to be correlated with the 3X3 kernel matrix shown:

$$P(k,l) = \begin{matrix} 2\,2\,2 \\ 2\,1\,2 \\ 2\,2\,2 \end{matrix}$$

wherein the entries 2 in P count the nearest neighbors, and the central entry 1 determines if the cell in question is on or off in generation n. This operation yields a matrix of correlation values Cn(i,j) at each cell i,j in the generation n according to $Cn(i,j) = \Sigma\, P(k,l)\, Ln(i-k,j-l)$. The correlation function Cn can take integer grey-scale values 0 to 17 at each cell. Birth corresponds to C = 6, survival to C = 5 or 7, and death to C = 0-4 or 8-17. All three rules of Life may now be combined into one simple condition: Turn cell i,j on in the n+1 time step if and only if Cn(i,j) = 5, 6 or 7 units. In terms of optical processing steps, this formulation of Life calls for the application of four basic operations in sequence--**correlation**, interval **thresholding** to extract the desired levels, **binarization** to normalize the levels and form the next generation lattice, and **iteration** to continue the update cycle. The CA program is encoded in the P matrix entries and the threshold levels, and can be changed to express different rules.



...and iterate

OPTICAL STEPS IN THE GAME OF LIFE CA

Photorefractive Nonlinear Implementation

Photorefractive nonlinear optical image processing devices have been reported which perform each of the operations identified above. Real time correlation or convolution in a photorefractive four-wave mixing crystal was demonstrated by White and Yariv (6), and can be used to cross-correlate the input Life pattern with the kernel matrix P. Interval thresholding can be performed using the double phase conjugate mirror, a photorefractive device in which two beams or images are simultaneously phase conjugated, with the phase conjugate reflectivity of one beam (or pixel) controlled by the intensity of the other (7). Binarization of the thresholded image can be accomplished using the optical bistability properties of the semi-linear phase conjugator, which displays hysteresis and latching of the phase conjugate reflectivity of one beam with respect to the level of a second, seed beam (7). Interpixel crosstalk may limit the resolution of these operations with respect to images. Figure 3 illustrates a proposed processing chain formed by cascading all three devices to perform conversion of an input Life pattern to the correct updated pattern in the next time step. The storage of the updated pattern takes place in the third crystal by virtue of latching, requiring application of an erase laser beam as part of the update cycle. Figure 4 shows a complete self-iterating Game of Life computer formed by combining two identical processor chains, with a total of six photorefractive crystals operating as three distinct device types. Initially, timing can be accomplished by means of the shutters indicated together with periodic erase beams. Later high speed implementations can use all-optical timing in the form of pulses from a mode locked laser, with various optical delay paths to encode the processing sequence. Using very high speed photorefractive or other nonlinear materials currently under development and a compact optical design, loop processing update speeds of 1 ns may be possible. With a lattice of 1000X1000 Life cells, approximately 10exp[15] cell updates/sec might be achieved, 10exp[7] times more than a Cray XMP.

Discussion

In addition to demonstrating an optical approach to cellular automata as such, successful construction of an all-optical Game of Life computer may well provide the simplest test bed in which to study many of the practical problems generic to any complex optical computing architecture, including programming, storage, timing, thresholding and stability, interpixel crosstalk, binarization, iteration and noise/error propagation. Experimental realization of the proposed six-crystal system is likely to be complex, requiring the state of the photorefractive art to be extended from devices to systems. Because the optical Game of Life performs a recognizable computational function, its performance may be easily evaluated and the solutions to device problems may impact other architectures such as neural nets.

This research was supported at Foster-Miller by SDIO and managed by ONR under the SBIR program. The authors wish to acknowledge helpful discussions with W. Micelli and R. Barakat.

References

1. M. Murdocca, Appl. Opt. 26 682, 1987.
2. Taboury, Wang, Chavel, Devos and Garda, Appl. Opt. 27, 1643, 1988.
3. S. Wolfram, Theory and Applications of Cellular Automata, World Scientific, 1986.
4. T. Toffoli and N. Margolus, Cellular Automaton Machines, MIT Press, 1987.
5. Berlecamp, Conway and Guy, Winning Ways for Your Mathematical Plays, Academic Press, 1984, vol. 2.
6. J. White and A. Yariv, Appl. Phys. Lett., 37, 5, 1980.
7. M. Cronin-Golomb and A. Yariv, Proc. SPIE 700, 301, 1986.

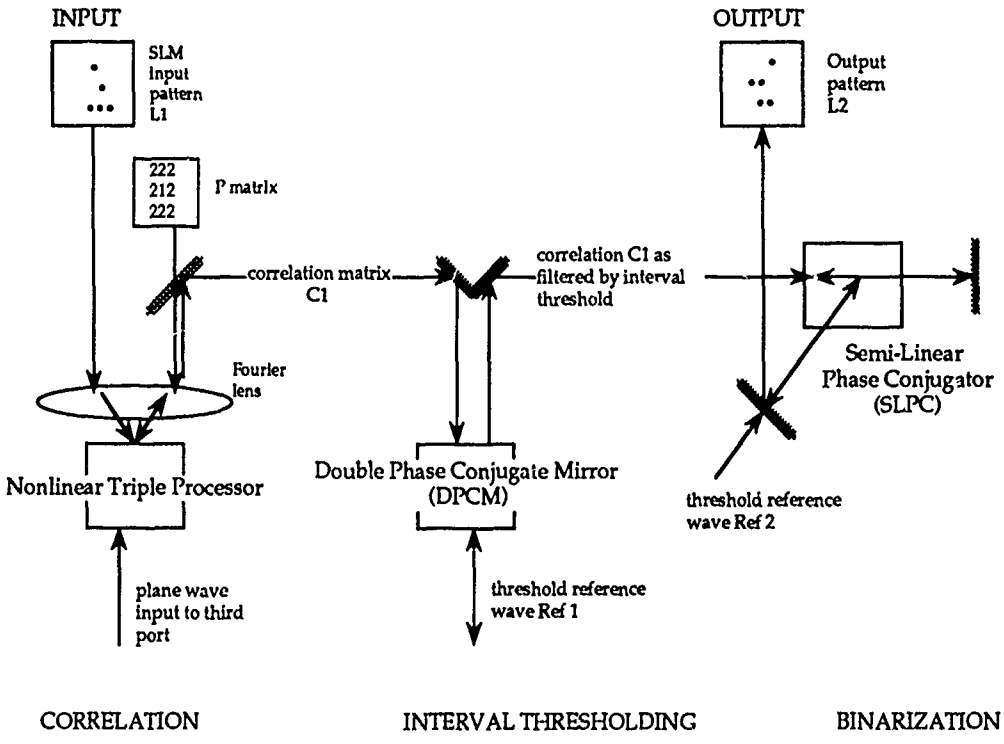


Figure 3. Main Processing Chain for the Optical Game of Life

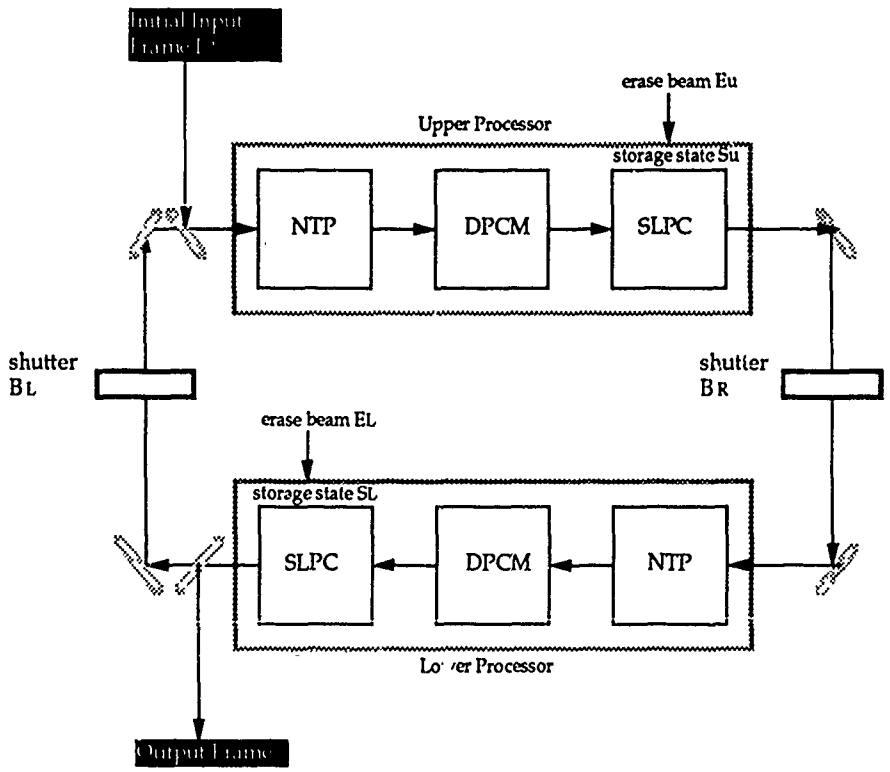


Figure 4. Self-Iterative Optical Game of Life Processor

OPTICAL DISK BASED CORRELATION ARCHITECTURES

Demetri Psaltis, Mark A. Neifeld and Alan Yamamura
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125

In this paper we describe and experimentally demonstrate optical image correlators that are implemented using optical memory disks. Optical correlation for pattern recognition [1] has long been considered a promising application for optical processing. One of the reasons such correlators have not been used in practical applications yet has been the lack of suitable spatial light modulators to be used as real time input devices. Recently, this limitation has to a large extent been removed through the development of a variety of 2-D SLM's [2] and concepts that allow the utilization of mature 1-D (acoustooptic) SLM's [3]. Attention has therefore shifted to the design of appropriate filters to perform reliable recognition [4]. In most practical applications a single filter is not suffiicient to produce reliable recognition, and the use of spatial [5] and temporal [3] multiplexing to search through a library of filters emerges as the most straightforward solution to the problem. The optical disk correlator architectures we describe in this paper provide an extremely efficient method for performing this task since they combine in a single device the huge memory required for storage of the library of reference images, the spatial light modulator needed to represent the reference in the optical correlator, and the scanning mechanism to temporally search through the library.

The first architecture we will describe is shown in Fig. 1. Each reference image is recorded as a 2-D computer generated Fourier transform hologram on the disk. The input image goes through the beamsplitter, it is Fourier transformed by the lens, and it illuminates the hologram on the disk. The reflected light contains a term proportional to the product of the transforms of the input and reference images. The same lens retransforms the reflected light and the correlation is produced. A principal issue of concern in this architecture is the suitability of commercially available disk systems for recording and reconstruction of holograms. We have identified a write-once disk system which is manufactured with glass (rather than plastic) covers of sufficient optical quality that has allowed us to reconstruct the recorded data using coherent light. We will report the results of this experiment at the conference. The rotation of the disk is used to perform a search through images centered at the same radial position on the disk. An auxillary scanning mechanism is needed in order to position the correlator "head" in the correct radial position. As the disk rotates the entire correlation pattern shifts in one dimension at the output as long as the reference hologram remains in the field of view. A time-delay-and-integrate (TDI) CCD sensor can be used to integrate this traveling correlation pattern in order to improve sensitivity. Alternatively a 1-D parallel read-out detector array can be used that sequentially produces slices of the 2-D correlation pattern as it travels past the detector array.

A straightforward modification of the system of Fig.1 is obtained by recording holograms that are Fourier transforms of the reference images only in the radial dimension since the rotation of the disks provides the necessary shift between the input and reference

along the tracks. The light reflected from such a hologram is Fourier transformed in the radial direction and integrated in the orthogonal dimension onto a 1-D parallel read-out array. The signal from the detector array is again the 2-D correlation presented as a sequence of 1-D slices. The advantage of this architecture compared to the previous one is that it has the same light efficiency as the TDI system without the relative complication of the TDI sensor. Therefore the experiments we will present are with this type of system.

The above architectures require storage of the reference images in the form of computer generated Fourier transform holograms. This provides the advantage of shift invariance which means that we do not need to be concerned with accurate positioning within a single track of the correlation head with respect to the data recorded on the disk. This is a very important practical consideration; the disadvantage however is an increase by a factor of 100 or more in the space bandwidth product required to record the hologram compared to the space bandwidth product of the image itself and an increased computational overhead to record the disk. In addition, the smaller size of the recording results in reduced phase uniformity requirement for the disk. In many cases it is only necessary to record the reference images as binary patterns [6] in which case they can be directly recorded on the disks. Gray scale images can be recorded using some form of area modulation as is done with video disks for example.

There are two types of architecture we will discuss that allow the reference images themselves to be stored on the disk rather than their Fourier transforms. The first is shown in Fig. 2. The input image goes through the beamsplitter and it is Fourier transformed by lens L_1. A Fourier transform hologram of the input is recorded in a photorefractive crystal using a reference beam that is incident from the right, as shown in the figure. Once the hologram is recorded the input is blocked and the the disk is illuminated. L_1 takes the Fourier transform of the reference image that is in the field of view of the illuminating beam and L_2 transforms the light diffracted by the hologram to produce the correlation at the output plane. The rotation of the disk is used to search through a library of images in the radial direction and a TDI detector can be used at the output to increase sensitivity as before. Multiple holograms could be multiplexed in the crystal to address different radial positions on the disk or the entire head can be scanned to address different radial positions as before. We have not yet completed the experimental demonstration of this system but we expect that at the conference we will present the experimental results from this system.

The final architecture we will discuss is shown in Fig. 3. The advantage of this architecture is that it operates on the light intensity and consequently the requirement for phase uniformity is greatly relaxed. As a result it is possible to implement this architecture with most existing disk systems. This correlator works as follows. The reference images are recorded on the disk and the input is imaged through a 1-D scanning device onto the disk. The scanner can be either acoustooptic (as shown in Fig. 3) or a rotating mirror. It provides the relative displacement in the radial direction between the input and reference images that is necessary to calculate the correlation function. The disk rotation provides the displacement in the orthogonal direction. The scanner translates the input image completely accross the stored reference image each time the disk rotates by a distance equal to a pixel of the reference. The intensity of the light reflected from the disk at any one time is proportional to the product between the input and a shifted

version of the reference. The reflected light is collected (integrated) on a single detector which produces as its output a temporal video signal of the 2-D correlation. This system was experimentally demonstrated with acoustooptic scanners. Two types of acoustooptic scanners can be used: A "flying spot" scanner in which a chirp signal propagates in the acoustooptic device acting as a traveling lens that scans the diffracted image at a rate equal to the acoustic velocity. This system completes a scan in a few μs, therefore a complete 2-D correlation takes approximately a few ms. The second scanner that we have demonstrated is a more conventional acoustooptic deflector that scans slowly but permits a higher space-bandwidth product of the input image. A sample of experimental results obtained with the system of Fig.3 is shown in Fig.4. Fig. 4a is a photograph of the pattern recorded on a write-once disk (the acronym CIT) and Fig. 4b is the 2-D correlation produced by the optical system of Fig. 3 and displayed by raster scanning the detector output on a 2-D monitor. Correlations can be produced with our experimental apparatus at rates up to 1000, 100X100 pixel images per second. The optically calculated correlation is in good agreement with the expected autocorrelation function of the CIT pattern. It should be pointed out that since this system operates on intensity we can only represent positive quantities. In order to represent bipolar input and/or reference images we need to add biases at the input stage and subtract it from the output [3], a technique that has been successfully used in a variety of incoherent architectures.

The number of bits that can be stored in the type of disk that we use for most of our work (a write-once, 12 cm diameter system from SONY) is more than 5 billion. The number of 100 × 100-pixel images that can be stored in such a disk is more than 5,000, assuming a generous factor of 100 for loss of spacebandwidth product due to representation (e.g. area modulation for gray scale representation). The rate at which all these images can be interrogated for a possible match with the input is limited by one or more of the following factors: The scanning speed of the disk (40Hz in our case), the speed of the radial scanning mechanism, and the sensitivity and the bandwidth of the output detectors and the electronics following them. As an example consider the system of Fig.2. At 40 Hz disk rotation rate, we obtain 1000 image correlations per 1/40th of a second (i.e. 40,000) image correlations per second), yielding a reasonable 4 MHz bandwidth per detector. It would be extremely difficult to duplicate this capability electronically and it can be achieved with *existing* optical technology. Moreover it is precisely such capability that is required for practical pattern recognition problems.

The research reported in this paper is supported by the Army Research Office.

References

[1] A. Vander Lugt, IEEE Trans. Inf. Theory IT-10, 2 (1964).
[2] OSA Topical Meeting on SLM Technology, Lake Tahoe, June 1988.
[3] D. Psaltis, Opt. Eng. vol 23,1 (1984).
[4] H.J. Caulfield, R. Haimes, and D. Casasent, Opt. Eng. 19(2), 152, (1980).
[5] B.D. Guenther, C.R. Chritensen, and J. Upatnieks, IEEE J.Q.E., vol. QE-15, pp.1348-1362, (1979).
[6] D. Psaltis, Paek, Venkatesh, Opt. Eng. vol. 23, pp.698-704, (1984).

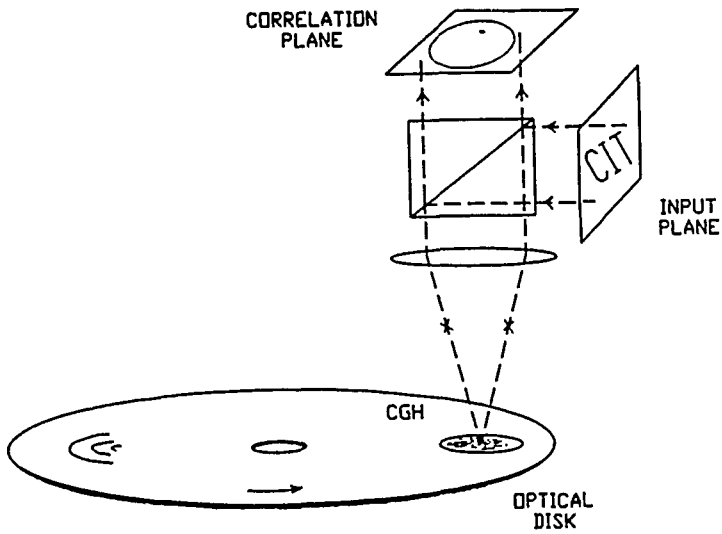


Figure 1

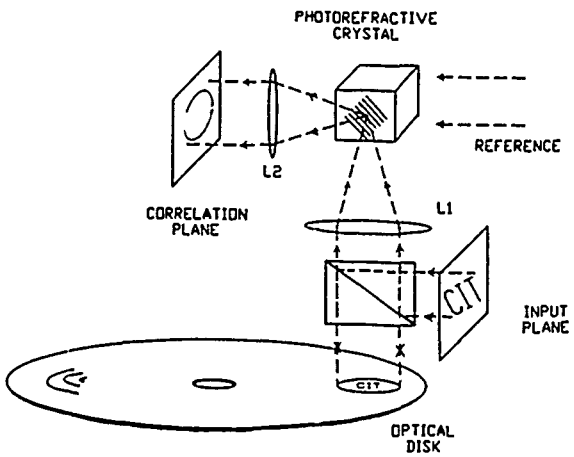


Figure 2

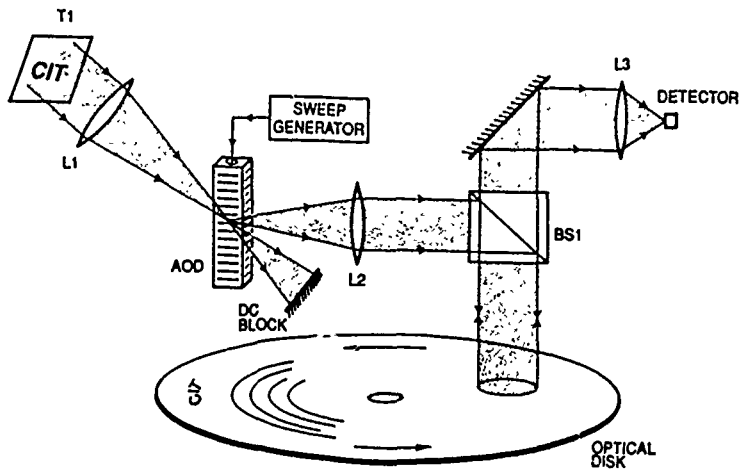


Figure 3

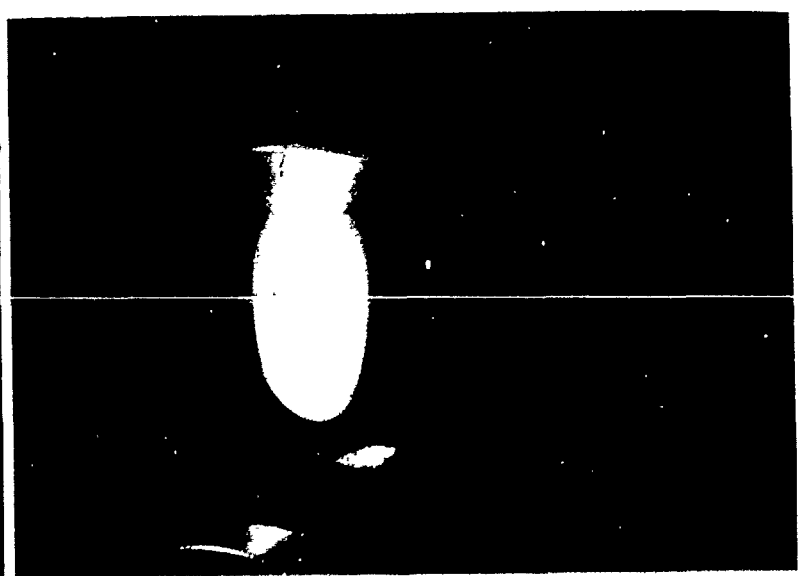


Figure 4a

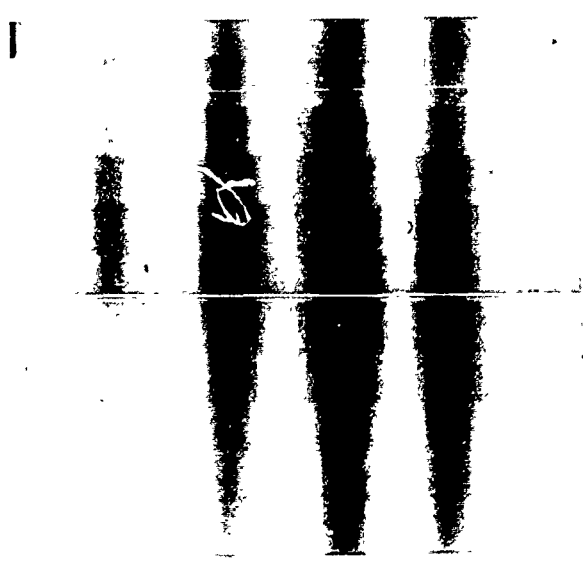


Figure 4b

PROPOSAL FOR AN OPTICAL CONTENT ADDRESSABLE MEMORY

Miles Murdocca
AT&T Bell Labs, Room 4G-538
Holmdel, New Jersey, 07733

John Hall, Saul Levy, Donald Smith
Department of Computer Science
Rutgers University
New Brunswick, New Jersey, 08903

Abstract

A content addressable memory (CAM) design that demands high throughput is proposed for arrays of optically nonlinear logic gates interconnected in free space.

1 Introduction

In a random access memory (RAM) each word of memory has a unique address. The physical position of a word in the memory is as significant as the value of the word. In a content addressable memory (CAM) a word is composed of *fields* that can be used as keys for indexing into the memory. The physical location of a CAM word is generally not as significant as the values contained in the fields of the word. Relationships between addresses, values and fields for RAM and CAM are shown in Figure 1. Values are stored in sequential locations in the RAM, with the address acting as the key to find the word. Four-byte address increments are used in this case. Values are stored in fields in the CAM, and in principle any field can be used to key on the rest of the word. If the CAM words were reordered, then the contents of the CAM would be virtually unchanged since physical location has no bearing on the interpretation of the fields. A reordering of the RAM may change the meanings of its values completely. This comparison suggests that CAM may be a preferred means for storing information when there is a significant cost with maintaining data in sorted order.

When a search is made through a RAM for a particular value, the entire memory may need to be searched one word at a time for that value when the memory is not sorted. When the RAM is maintained in sorted order, it may still require a number of accesses to either find the value being searched for or to determine the value is not stored in the memory. In a CAM, the value being searched for is broadcast to all of the words simultaneously, and a small processor at each word makes a field comparison for membership, and in two steps the answer is known. A few additional steps may be needed to collect the results but in general the time required to search a CAM is less than for a RAM in the same technology.

CAM's are not in common use largely due to the difficulty of implementing an efficient design with conventional technology. Consider the block diagram of a CAM shown in Figure 4a.[1] A Central Control unit sends a comparand to each of 4096 cells, where a comparison is made and the result is put in the Tag bits T_i which are collected by a Data Gathering Device and sent to the Central Control unit. When the Central Control unit loads the value to be searched into the comparand register, it sets up a mask to block out fields that are not part of the value. A small local processor at each cell makes a comparison between its local word and the broadcast value and reports the result of the comparison to the Data Gathering Device. A number of problems arise when an attempt is made to implement this architecture in a conventional technology such as very large scale integration (VLSI). The broadcast function that sends the comparand to the cells can be implemented with low latency if a tree structure is used. If the tree cannot be contained on a single chip, then connections must be made among a number of chips, which quickly limits chip density.[2] For example, a node of a tree that has a single four-bit input and two four-bit outputs needs 12 input/output (I/O) pins and three control pins if only one node is placed on a chip. A three node subtree needs 25 pins and a seven node subtree needs 45 pins. A 63 node subtree requires 325 pins not including power pins, and this outstrips most present day packaging technologies. A useful CAM would contain thousands of such nodes with wider data paths, so the I/O bandwidth limit is realized early in the design of the CAM. Compromises can be made by multiplexing data onto the limited number of I/O connections but this reduces effective speed and a major advantage of CAM over RAM.

Recent work on regular free-space optical interconnects such as perfect shuffles, banyans, and crossovers shows that comparable gate counts and circuit delays can be achieved when regular interconnects are used between optical logic gates rather than arbitrary interconnects.[3] Optical implementations exist for $log_2 N$ interconnects such as perfect shuffles[4,5] and crossovers[6] and progress has been made on optical logic devices[7,8] and systems[9] so that it is timely to design optical digital circuits such as a CAM for

a potential implementation. Our goal here is to show how a CAM can be efficiently implemented with optical-logic devices and regular, free-space optical interconnects.

2 Functional Layout of the CAM

The layout of the CAM we propose is shown in Figure 4b. A conventional RAM, a Register File, an Instruction Unit, and a Logic Unit make up a simple computer. A distribution and collection tree, the CAM words, and a Backing Store for the CAM words make up an extension to the simple computer. The Instruction Unit acts as the central control for the system. An instruction is sent from the Instruction Unit to the Logic Unit, where the instruction is decoded into microcode sequences which are distributed via a tree structure to all of the CAM cells in the tree. A Backing Store made up of serial memory extends the width of a CAM word without introducing significant cost into the processing element (PE) in each cell. The tree is used for collecting results which are then reported to the Instruction Unit via the Logic Unit and Register File.

An expanded view of a node in the CAM tree is shown in Figure 2. Two ALUs perform one of 16 operations such as in the 74181 four-bit arithmetic unit (ALU)[10] on the stored word and the input word when data is traveling down the tree. The ALUs send their outputs from the leaves of the node to the roots of the next nodes, or to the CAM words when there are no lower nodes. When data is collected up through the tree to the Logic Unit, an operation is performed on the two inputs from the lower nodes to decide which input is passed to the root of the tree and which input is stored in the local memory.

3 Regular Interconnect Design

The ALU's are similar in function and design to the Texas Instruments 74181 four-bit ALU chip. The 74181 is made up of four one-bit sections that operate in parallel for carryless operations and operate in parallel/serial mode for carry operations. The design for a one-bit section is shown in Figure 4. The target optical architecture for the ALU is made up of cascadable arrays of two-input, two-output NOR gates interconnected in free space with a regular interconnection scheme of the form shown in Figure 5. The restricted interconnection topology does not introduce a large cost in space or time as might be suspected, which supports previous results on regular free-space interconnects.[3]

The backing store can be implemented efficiently with four-bit wide serial delay lines since the four-bit ALUs can only process data in four-bit nibbles stored sequentially in this design. The linear distance between the output of the Read/Write mechanism and the input from the feedback path is great enough to store a large word (on the order of several hundred bits) in free space as it propagates around the feedback loop. Diagrams are omitted for space considerations. The tree connections can be implemented with a cascadable optical perfect shuffle (or a topologically equivalent crossover network).

4 Conclusion

An optical design for a content addressable memory is proposed. The design makes use of arrays of optically nonlinear logic gates interconnected in free space with simple components such as spherical lenses, mirrors, and gratings. The design is suitable for an optical implementation because the bandwidth requirements cannot be met with electronics technology in the foreseeable future, and because the optical interconnection scheme presented here is simpler than competing optical interconnect topologies such as fibers and one-to-many schemes based on holograms or magnification.

This project was funded in part by Air Force Office of Scientific Research grant AFOSR-86-0294.

References

[1] Foster, C. C., *Content Addressable Parallel Processors*, Van Nostrand Reinhold Company, (1976).

[2] Franklin, M. A., D. F. Wann, and W. J. Thomas, "Pin limitations and partitioning of VLSI interconnection networks," *IEEE Trans. Comp.*, C-31, 1109, (Nov. 1982).

[3] Murdocca, M. J., A. Huang, J. Jahns, and N. Streibl, "Optical design of programmable logic arrays," *Appl. Opt.*, 27, (May 1, 1988).

[4] A. W. Lohmann, W. Stork, and G. Stucke, "Optical Perfect Shuffle," *Appl. Opt.*, 25, 1530, (May 15, 1986).

[5] Brenner, K.-H. and A. Huang, "Optical implementation of the perfect shuffle interconnections," *Appl. Opt.*, 27, 135, (1988).

[6] Jahns, J., and M. J. Murdocca, "Crossover networks and their optical implementation," *Appl. Opt.*, **27**,3155, (August 1, 1988).

[7] Jewell J. L., A. Scherer, S. L. McCall, A. C. Gossard, and J. H. English, "GaAs–AlAs monolithic microresonator arrays," *Appl. Phys. Lett.*, **51**, (2), 94, (July 13, 1987).

[8] Lentine, A. L., H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, "The symmetric self electro–optic effect device," in Conference on Lasers and Electro–optics, Technical Digest Series 1987, vol. 14, (Optical Society of America, Washington, D.C., 1987, 249), postdeadline paper.

[9] Prise, M. E., N. Streibl, and M. M. Downs, "Computational properties of nonlinear optical devices," Topical Meeting on Photonic Switching, Technical Digest Series 1987, **11**, (Optical Society of America, Washington, D. C., 1987), 110, (1987).

[10] Texas Instruments, *The TTL Data Book for Design Engineers*, Texas Instruments, Inc., 2nd ed., (1976).

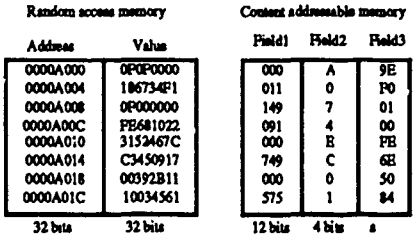


Figure 1: Relationships between random access memory and content addressable memory.

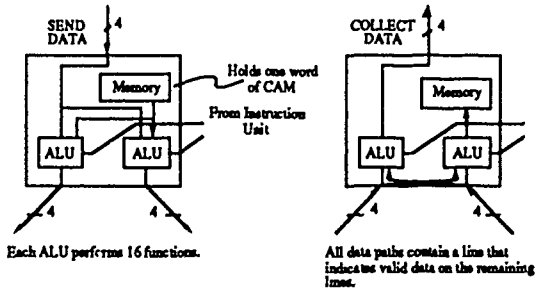


Figure 2: A node in the CAM tree contains two ALU's, enough memory to hold one CAM word, and bidirectional links for sending and collecting data.

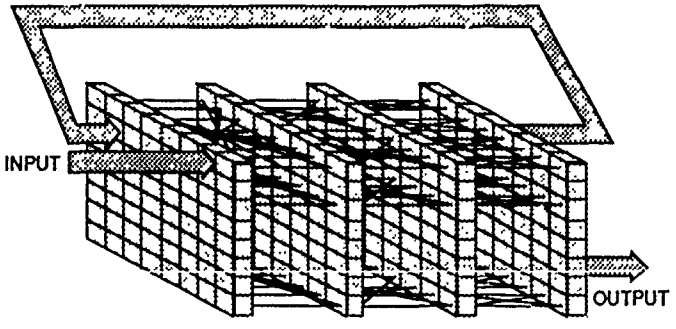


Figure 3: Arrays of optically nonlinear logic devices are interconnected in free space with regular interconnects. All logic gates perform the NOR function and have fan-in and fan-out of two.

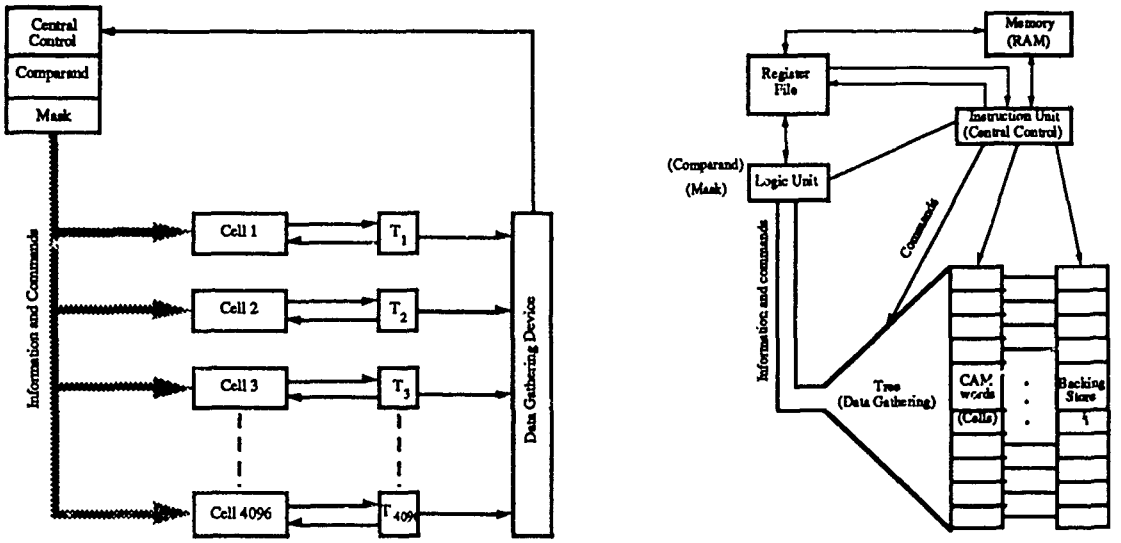


Figure 4: (a): Simplified CAM model. (b): Optical CAM model. The CAM words, the backing store, and the processing elements for the CAM words are designed for an optical implementation. The remaining components are suited for an electronic implementation.

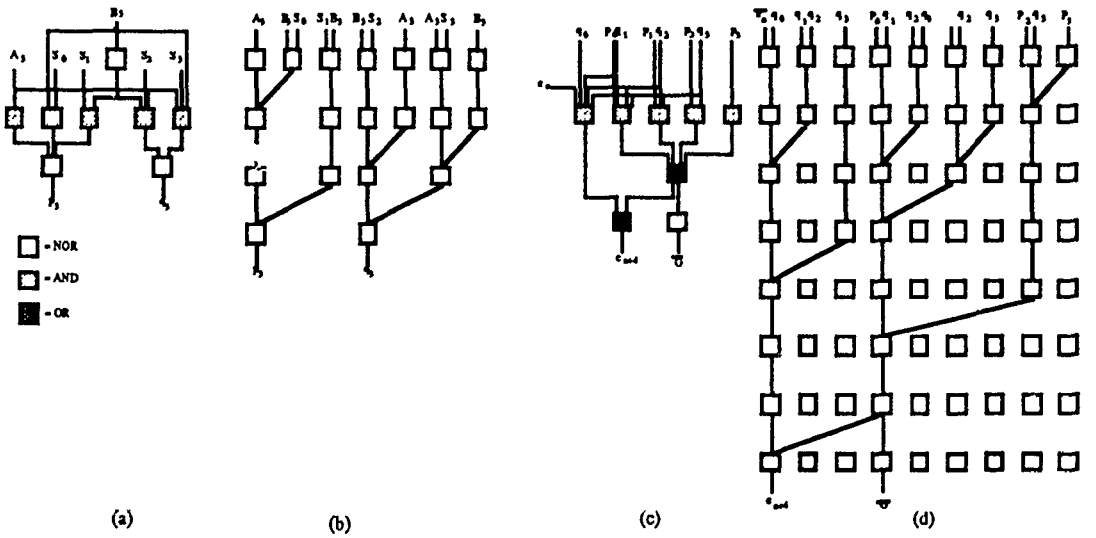


Figure 5: (a): Functional layout of top half of 74181 one-bit unit. (b): Two-input, two-output NOR equivalent of top half. (c) Functional layout of bottom half of 74181 one-bit unit. (d) Two-input, two-output NOR equivalent of bottom half.

NOTES

TUESDAY, FEBRUARY 28, 1989

SALON F

5:00 PM–6:00 PM

TuH1–TuH4

OPTICAL COMPUTING SYSTEMS: 4

Adolf W. Lohmann, University of Erlangen-Nuremberg, Federal Republic of Germany, *Presider*

Optical Outer Product Look-up Table Architectures for Residue Arithmetic

Mark L. Heinrich, Ravindra A. Athale, and Michael W. Haney
The BDM Corp.
7915 Jones Branch Drive
McLean, VA. 22102

BACKGROUND

The residue number system (RNS) allows high accuracy integer-valued arithmetic operations to be decomposed into independent (carry-free), low accuracy computations that can be performed in parallel. The RNS thus provides an attractive alternative to weighted number systems (e.g., binary or decimal) for high speed numerical computing [1]. The residue number representation is completely specified by a set of relatively prime moduli. The overall dynamic range is given by the product fo the moduli. Although this dynamic range can be arbitrarily high, the dynamic range required in any individual subcalculation is commensurate only with the associated modulus. The RNS also leads to a reduction in the growth of the total number of combinatorial logic elements required to perform a calculation via truth table approach. Specifically, the RNS exhibits additive growth in spatial complexity with input word size, contrasted by multiplicative (exponential) growth for weighted number systems.

Various optical encoding schemes have been proposed for residue based systems, including the polarization or phase state of a light beam where arithmetic is performed by a cyclic permutation of states [2], and binary-coding of residues (analogous to binary-coded decimals (BCD)) where arithmetic is performed via truth-table look-up [3]. In this paper we focus on spatial encoding for residue representation in which m spatial positions (for modulo m data) are available, but only one of which is active for any given operation. With position-coded inputs, residue arithmetic is realized by selecting the correct spatial "mapping" between input and output channels [4]. A bank of maps acts as a look-up table (LUT) by providing arbitrary permutations of input channels, thus realizing general processing. Position coded representation and LUTs for modulo 7 addition and multiplication are shown in Figure 1.

It can be seen from Figure 1 that the number of entries in the LUT grows quadratically with the size of the modulus. Recently, two approaches to overcome the quadratic spatial complexity of the position coded LUTs have been proposed based upon unique properties of the RNS. One approach is based on the Toeplitz form (constant along cross-diagonals) of the addition LUT matrix seen in Figure 1. This architecture requires only (2m-1) processing elements (PEs) to detect the proper cross-diagonal in a modulo m addition. Multiplication is more difficult and must be mapped into addition via logarithmic transformation [5]. The other approach uses second level factorization which is based upon eliminating multiplication by zero in the residue multiplication table and factoring the reduced residue set. However, addition is more complex since it does not possess the same property with respect to zero [6]. Both LUT architectures exhibit spatial complexities which are approximately linear in m, but in utilizing these unique features of residue addition and multiplication the generality of the LUT to perform arbitrary RNS arithmetic operations has been lost.

LUT processing with m^2 PEs is the most direct way to perform general residue arithmetic in a single gate delay. With binary position-coded inputs, all PEs can be realized with 2-input AND gates. An example of a direct LUT architecture is shown in Figure 2 for modulo 7. The table contains 49 (m^2) 2-input AND gates, of which only one will be active for a given table look-up. The interconnection requirements for such an architecture are as follows. First, each input must be broadcast over m row or column elements. Second, since the output of any modulo m residue arithmetic operation must also be modulo m, the architecture must provide an space-variant m^2-m mapping from the LUT processing elements to output channels. These direct LUTs can be cascaded since the mapping provides position coded data at the output. Since global and arbitrary interconnects are the main advantage of optics, it would appear that an architecture that implements all the necessary interconnects optically will be desirable.

APPROACH

In this paper we propose an optical outer-product LUT architecture which performs the required m^2 binary AND operations with only 2m active elements and retains the ability to compute general residue arithmetic functions by utilizing the 3-D interconnection capabilities of free-space optics. The outer product between an m-element column vector $\underline{a}$ and an m-element row vector $\underline{b}$ is an m x m matrix M whose

elements M(i,j) are defined by M(i,j) = a(i)*b(j). For binary input vectors, the outer product matrix contains the results of m^2 2-input AND operations as encountered in the LUT of Figure 2. One module of the optical outer product architecture is shown in Figure 3. Orthogonally-oriented input arrays (laser diodes and modulators) are cascaded in a multiplicative fashion so that an m^2-position (m x m) outer product matrix is formed in their common image plane. Each laser diode (LD) is broadcast in the vertical dimension over the m-element modulator array. Since position coded representation is employed, only one-of-m LDs and one-of-m modulators will be active for a given operation. Hence, the active LD channel provides the LUT column address and the active modulator channel provides the LUT row address. Only one of the m^2 LUT positions will contain light in the LUT plane. As with the direct approach, a passive space-variant m^2-m mapping will route light from the LUT plane to the proper output positions, thus realizing any integer-valued function of two independent variables by designing a suitable mapping.

The performance of the optical outer product LUT is measured in terms of system complexities and their dependence upon the modulus. We define temporal complexity (TC) as the number of sequential switching stages required to perform a LUT operation, spatial complexity (SC) as the number of PEs required to implement the LUT, and power complexity (PC) as the number of PEs which are active during table look-up multiplied by their respective power requirements. For one module of the outer product look-up table, TC is unit gate delay and thus independent of m. We have mentioned that SC has a linear dependence on m, namely m LDs and m modulators are required. Since only one laser diode and modulator is active in a given module regardless of the size of the modulus, it may appear that PC is independent of m. A closer inspection reveals, however, that the power requirement of each laser diode grows linearly with m due to the broadcast requirements. The PC of a module m is thus given by

$$PC = (m/\alpha)P_{ld} + P_{mod}, \qquad (1)$$

where P_{ld} and P_{mod} correspond to the power required to drive the respective active devices, m corresponds to the gain required to compensate for the broadcast loss of the LDs, and α (<1) incorporates losses due to inefficiencies in the optical mapping, LDs, and modulator transmittance.

Due to the mutual independence of the modules, total SC and PC exhibit additive growth with the number of modules. Table I shows the number of outer product modules required for 16-, 32-, and 64-bit dynamic range, the maximum and minimum required input array lengths, and the total SC of the processor. It can be seen from the table that the largest LUT required for 64-bit dynamic range is 47x47, which is quite modest.

SUMMARY

In this paper, we have described the design of an optical arithmetic unit performing high accuracy computation in the residue number system. The system is based on an optical outer-product architecture combined with an arbitrary m^2-m mapping that can be implemented holographically or with a fiber bundle. The architecture fully exploits the interconnection capability of optics while minimizing the number of active switching elements. The system is based on the rapidly advancing laser diode technology and can use multiple quantum well modulators with high switching speeds (<nanosecond) and low switching energies (<picojoule) [7]. The low contrast and small array size limitation of this technology will not adversely affect the performance of the system described in this paper. The residue number system does not allow general division or relative magnitude comparison, thus losing its attractiveness for general purpose computations. However, in high speed signal processing applications, which primarily require multiplication and addition, the residue-based optical processors potentially provides a distinct advantage over electronic systems based on weighted number system. The absence of clock skew and crosstalk in the optical interconnects combined with high speed optical switches makes the optical approach competitive to electronic approaches to residue computations. Conversion from analog/binary representation to residue and back is the next challenge that needs to be addressed with optoelectronic approaches.

REFERENCES

1. N.S. Szabo and R.I. Tanaka, Residue Arithmetic and its Applications to Computer Technology, McGraw Hill, New York, 1967.
2. S.A. Collins, Jr., "Numerical Optical Data Processor," Proc. SPIE, Vol. 128, pp. 313-319 (1977).

3. T.K. Gaylord, M.M. Mirsalehi, and C.C. Guest, "Optical Digital Truth-Table Look-Up Processing," Opt. Eng., Vol. 24, pp. 48-58 (1985).
4. A. Huang, Y. Tsunoda, J.W. Goodman, S. Ishihara, "Optical Computation Using Residue Arithmetic," Applied Optics, Vol. 18, pp. 149-162 (1979).
5. C.D. Capps, R.A. Falk, and T.L. Houk, "Optical Arithmetic/Logic Unit Based on Residue Arithmetic and Symbolic-Substitution," Applied Optics, Vol. 27, pp. 1682-1686 (1988).
6. A.P. Goutzoulis, E.C. Malarkey, D.K. Davies, J.C. Bradley, and P.R. Beaudet, "Optical Processing with Residue LED/LD Lookup Tables," Applied Optics, Vol. 27, pp. 1674-1681 (1988).
7. D.A.B. Miller, "Quantum Wells for Optical Information Processing", Opt. Eng., Vol. 26, p. 368 (1987).

Table I. Outer Product LUT Performance versus Dynamic Range

Dynamic Range	16-bits	32-bits	64-bits
# of modules	5	9	14
smallest linear array	7	5	7
largest linear array	13	23	47
Spatial Complexity	48 LDs + 48 mod	120 LDs + 120 mod	379 LDs + 379 mod

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
→ 3	3	4	5	6	(0)	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

x	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
→ 3	0	3	6	2	(5)	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

Figure 1. Modulo 7 Addition and Multiplication Look-up Tables

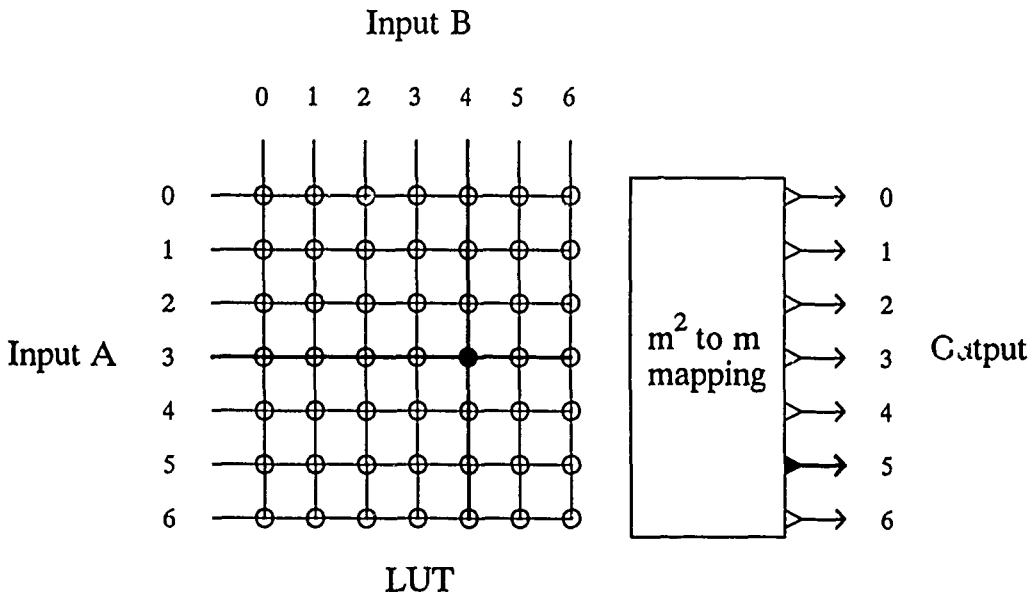


Figure 2. Direct Look-up Table Architecture

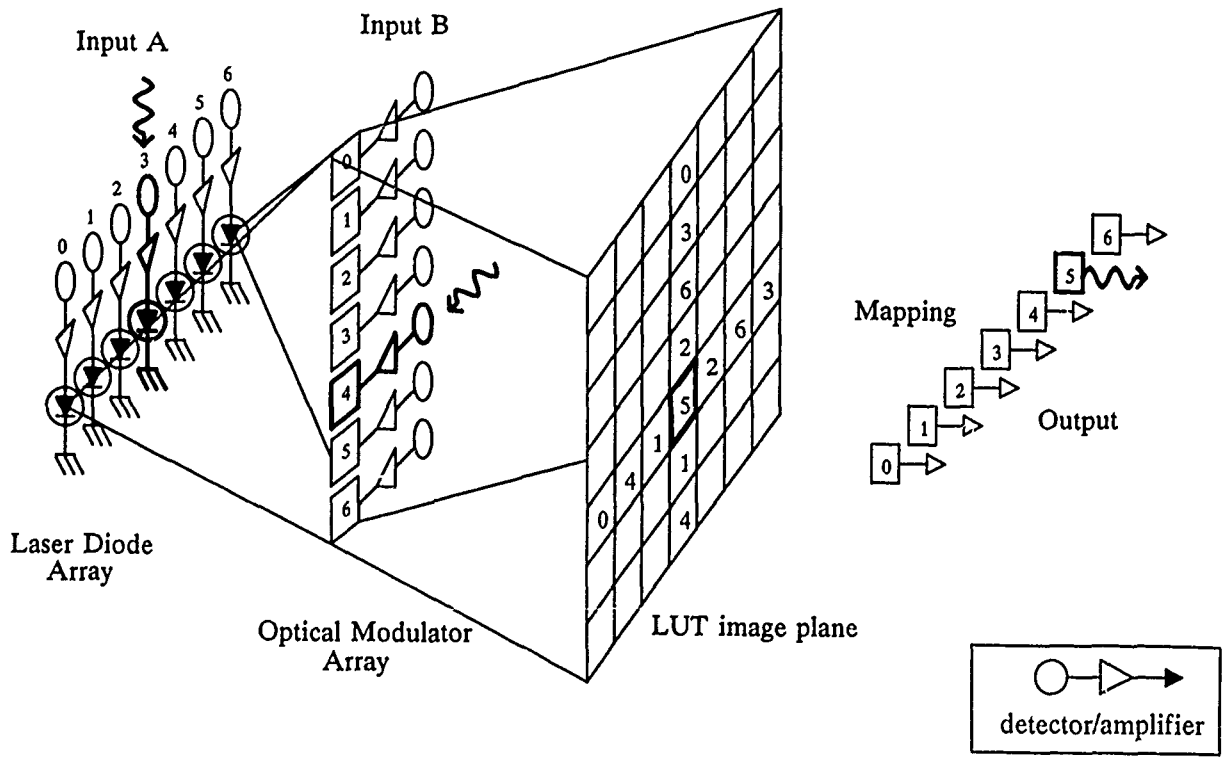


Figure 3. Optical Outer Product Look-up Table Architecture

Optical Transversal Filter With Variable Weights

Debra M. Gookin and Mark H. Berry
Naval Ocean Systems Center
San Diego, CA 92152

We describe an optical transversal filter constructed from off-the-shelf components. The filter is comprised of fiber optics and integrated optical devices. This filter differs significantly from previously constructed filters which all had fixed tap weights,[1-3] usually of value unity. An approach used to make fixed weight taps was to produce dielectric mirrors directly in the fiber.[3] All of these systems were limited by the inflexibility of their weighting scheme. This limitation adversely affects the filter performance in two ways. First, it is not possible to change the response function of the filter once the filter has been made. Second, it is not possible to correct for tap weight errors. When a fiber optic transversal filter is built, by any technique, there will be some error in the value of the loss and the delay between taps. Fixed weight systems have no technique for self correction.

Our variable weight system can be used to process large bandwidth (~10 GHz) analog electrical or optical signals. The sampling rate depends on the spacing between taps. A one centimeter tap spacing corresponds to a 20 GHz sampling rate. The tap weights are configurable at up to 10 GHz rates.[4] This system can be used to do one or several filtering functions. If the signal is stationary for periods greater than the reconfiguration time, then the tapped delay line can be programmed to perform a sequence of filter operations.

The tap weights in our transversal filter are integrated optic two-by-two directional couplers. The intensity at the outputs of the couplers, I_0 and II_0, depends on the input intensity, I_i, and on the applied voltage, V(t). (see figure 1) The second input of each device, II_i, is unused. The tap weights in our system are controlled using a 0-8 volt analog signal.

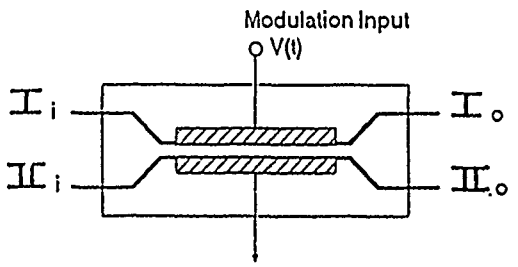


Figure 1. Integrated optical 2x2 directional coupler.

The 8 tap delay line transversal filter is shown in figure 2. The laser diode intensity is modulated by a high bandwidth electric signal. The light in the fiber is split by a 3dB coupler and two 1x4 trees. A fraction of the original light intensity is conveyed to each of the 8 integrated optic two-by-two couplers. The couplers serve as filter taps. The intensity of the beam passed by the coupler is controlled by the applied voltage. The light at output II_0 is collected by an asymmetric star coupler and another 3dB coupler. The signal at the detector is the incoherent sum of the optical intensities from each tap. The electrical output of the detector is the electrical input signal modified by the tapped delay line filter. The tap weights can be computed either in advance or during the filtering process (ie. for adaptive filtering and neural computing applications).

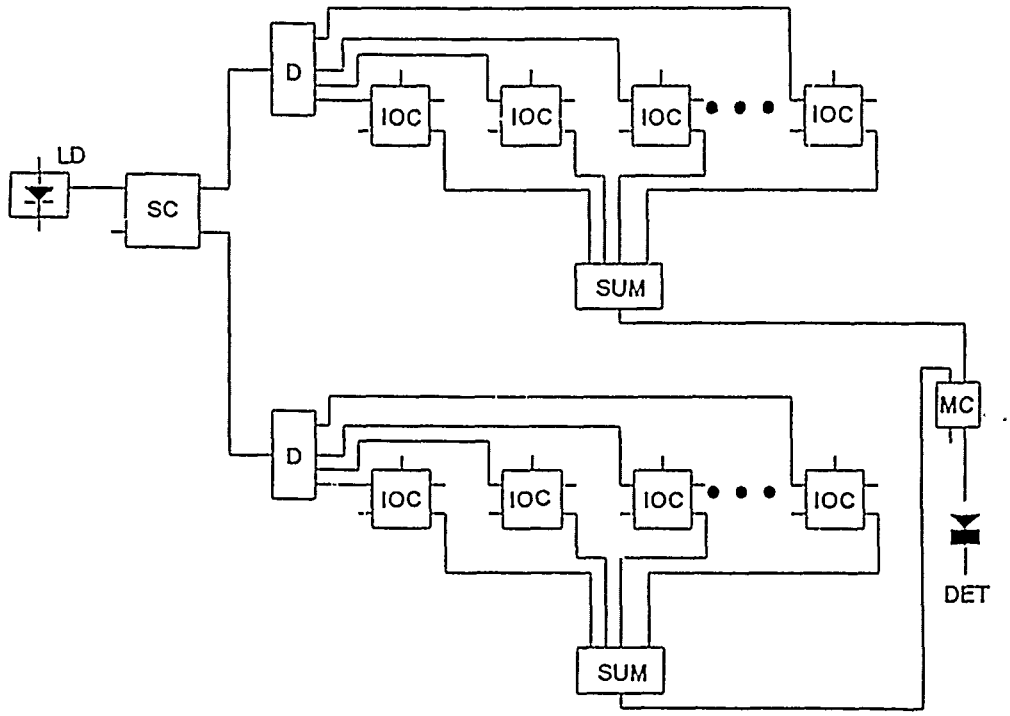


Figure 2. Eight tap variable weight transversal filter. (LD–laser diode, SC–single mode 3 dB coupler, D–1x4 divider, IOC–integrated optical 2x2 directional coupler, SUM–asymmetric star coupler, MC–multimode coupler, DET–detector)

There are other possible designs for the transversal filter. One possibility is to use a design similar to that used in previous fixed weight systems.[1] According to that design, the signal beam is not divided up and sent to each tap independently. Instead, the signal is sent along from one tap to the next serially. The reason we did not choose this approach is that the amplitude of the signal at each tap depends on the previous tap values. Therefore the fraction of the light signal you want to pass at each tap (and hence, the applied voltage) depends on the weight values at all the taps. The computation of the correct voltages to apply at each tap to get the desired filter response becomes complicated, especially when the effects of extraneous unknown losses are thrown in. The tapped delay line design in figure 2 simplifies the weight calculations. Using this design, the desired weight value is proportional to the applied voltage.

In order to get an estimate of the laser signal power necessary, we have calculated the losses in this eight tap system for a flat filter with maximum transmission at each coupler. We use the losses specified by the manufacturers for the off–the–shelf devices we are using: fiber optic splice – 0.15 dB (typical). 1x4 tree coupler – 0.6 dB (maximum), I.O. 2x2 coupler – 6.0 dB (maximum). (We are operating at $\lambda = 1.3$ μm.) With these assumptions for the losses at points A through K in figure 2, we predict the signal at the detector will be 11.8 dB down. Therefore, if we couple 2mW of power from the laser diode into the fiber, the optical signal at the detector will be about 140 μW.

The variable weight fiber optic tapped delay line we have described has the large signal bandwidth of other fiber optic transversal filters (0–15 GHz), with some of the versatility of an electronic system. The computational accuracy of this filter is limited by the extinction ratio of the integrated optic 2x2 directional couplers and by the linearity of the laser diode. Currently available off–the–shelf 2x2 directional couplers have extinction ratios specified to 20 dB,[4] and couplers are available by special order with extinction ratios specified to 30 dB.[4] Laser diodes are very linear for small modulation depths (small dynamic ranges), but in the tradeoff between linearity and large dynamic range, linearity is generally sacrificed. With currently available devices, this type of transversal filter does not have the high accuracy of electronic systems, however it can perform filter operations more quickly and on signals with larger bandwidths than electronic systems.

References

1 Stephen A. Pappert, Matthew N. McLandrich, and Ching–Ten Chang, J. Lightwave Tech. 3(2), 273(1985).

2 Stephen A. Pappert Ching–Ten Chang, and Matthew N. McLandrich, Fiber and Int. Optics 6(1), 63(1985).

3 C.–E. Lee, R.A. Atkins, and H.F. Taylor, Electron. Lett. 23(11), 596(1987).

4 Crystal Technology, private communication.

Integrated Optoelectronic Cellular Array For Fine-Grained Parallel Processing Systems

M. Hibbs-Brenner, S.D. Mukherjee, M.P. Bendett
Honeywell, Inc.
Sensors and Signal Processing Laboratory
10701 Lyndale Avenue South
Bloomington, MN 55420

A.R. Tanguay, Jr.
523 Seaver Science Center
University of Southern California
University Park MC-0483
Los Angeles, California 90089-0483

Introduction.

Due to the limitations of electrical interconnections, and the lack of compact optical logic elements, it may be advantageous to implement architectures requiring a high degree of parallelism and interconnectivity with a hybrid system which utilizes the massive interconnection potential of optics, while performing logic functions in electronics. This paper describes the development of an AlGaAs/GaAs based integrated optoelectronic cellular array (IOCA) in which each cell consists of an electronic processing element together with optoelectronic devices which perform the I/O function. Topics to be discussed in this paper include: potential applications of the IOCA, limits to scalability of the array, the choice of the optoelectronic components utilized in the chip, and the fabrication of integrable optoelectronic devices.

Integrated Optoelectronic Cellular Array Chip Description.

A schematic drawing of the Integrated Optoelectronic Cellular Array chip is shown in the upper left corner of Figure 1. The chip consists of a two dimensional array of cells, with each cell containing a vertically emitting AlGaAs/GaAs double heterojunction LED, an ion implanted GaAs photoconductive optical detector, and GaAs E/D MESFET based LED driver, amplifier, and logic. By connecting the cells together locally, data can be clocked into or out of the array electrically. Global interconnections can be achieved optically.

Potential Applications of the IOCA Chip.

Figure 1 also illustrates potential areas of application of the IOCA chip. The first is as an array of processing elements for fine-grained computing architectures requiring a high density of intra-chip connections. An example in which the cell logic function is very simple, but a high level of interconnectivity is required, is that of a neural network. The cell processing function could be as limited as a threshold, while the memory would reside in an interconnect element such as a volume hologram.[1,2] Another example of a fine-grained computing system which could be implemented with this chip is the Digital Optical Cellular Image Processor (DOCIP) which is well-suited to performing binary image algebra or pattern recognition functions.[3] Each cell in the array would correspond to a 54-logic gate DOCIP processor. The use of global optical interconnections would facilitate a hypercube interconnect architecture which would reduce the communication times between processors from O(N) to $O(\log_2 N)$ for an NxN array.

The high I/O density realizable with a 2-D array of emitters and detectors could also facilitate high density interchip interconnects. A polymeric waveguide can be used to perform the interconnect, with the coupling into or out of the polymeric waveguide accomplished with a grating or prism.

A third possible application is in the area of sensor signal processing. For instance, the IOCA chip could serve as a focal plane array of detectors with pre-processing for each array element performed on chip. The outputs from all elements would then be optically transmitted in parallel to a second processing chip.

System Scalability.

An important issue in determining the utility of the IOCA is the number of cells which can be fabricated on a single die. This is primarily limited by the power which can be dissipated on the chip. For GaAs substrates, power dissipation ranging from 1 to 10 W/cm^2 is reasonable, depending on heat sinking capability. A reasonable upper limit on chip size in the near term is 1 cm^2 based on achievable device uniformities. A typical operating condition for an LED is 10mA at 1.5V. If operated CW, the power dissipation due to the LEDs alone would limit the array size to 67. However, pulsing the LEDs with a 0.1% duty cycle would allow the array size to reach about 67000. Although the effective speed of the array is reduced, the increase in parallelism achieved by the increased array size can outweigh the effect of reduced speed.

Of course, since the electronics also consume power there is a trade-off between the complexity of processing which is carried out within each cell and the total number of cells. Figure 2 illustrates this trade-off. The assumptions in this calculation are outlined in the figure. Under these assumptions, the number of cells achievable for a neural network (e.g. Hopfield model with a simple threshold in each cell (number of gates is on the order of 1)) would be about 10^4. The DOCIP processor mentioned above would require 54 gates which would limit the number of cells on the chip to about 10^3.

Choice of components.

Figure 3 illustrates some of the considerations necessary in choosing the individual optoelectronic components which constitute the IOCA. The detector output current is shown as a function of the source CW power dissipation for four different combinations of LED, laser, PIN photodetector and photoconductive detector. Also shown are the noise floors for the two different kinds of detectors. Assumptions about source and interconnect efficiency are indicated in the diagram. It is clear that for applications where high signal-to-noise ratios are required, the laser/PIN detector combination is best. This is also the component combination which would be required for high speed operation (1 GHz). On the other hand, for lower speed applications (<200 MHz) the photoconductor has the advantage of providing a relatively high current output, reducing amplifier complexity and power requirements. The laser/photoconductor combination is the most promising for neural network type applications which are fault tolerant but require minimum power dissipation and maximum array size. In addition to the higher efficiency of the lasers, their advantage over the LEDs is also the potential for less cross-talk between cells of the IOCA due to their narrower spectral width. However, the development of surface-emitting lasers is currently at a preliminary stage.

Fabrication of individual components for the IOCA.

The greatest challenge in the development of the IOCA chip is the fabrication of an integrable LED structure. In order to be integrable with the standard E/D MESFET process the LED must be fabricated on a semi-insulating substrate and planarity of the wafer surface must be maintained. The LED structure chosen for this program is illustrated in Figure 4. An AlGaAs/GaAs double heterostructure is epitaxially grown in a well etched into a semi-insulating substrate using OMVPE. A thick p-GaAs layer is grown first in order to facilitate the formation of a contact to the p- side of the junction. The wafer surface outside the well is masked with silicon nitride causing polycrystalline material to grow there. This polycrystalline material is chemically removed after growth. The n-ohmic contact can be made to the top surface, whereas

a via must be etched to the bottom of the heterostructure in order to form the p-ohmic contact. The fabrication of the E/D MESFET I.C. begins after the removal of polycrystalline material and the nitride mask. We have fabricated LEDs using the full process described above. A photograph of the light output from one of these LEDs is shown in Figure 4.

Both photoconductors and PIN photodiodes have also been fabricated. They consist of interdigitated metal fingers deposited on an ion implanted substrate. Detector size is 40 x 60 μm^2 with the metal shadowing 35% of the detector area. A typical responsivity is 1000 A/W at 1 μW incident optical power for a photoconductor and 0.2-0.4 A/W at all incident power levels for a PIN photodiode.

The surface-emitting LED and photoconductor described above will be combined with a simple amplifier whose output will drive the LED. The transfer characteristic of the amplifier also acts as a thresholding function. The amplifier will be fabricated using a standard E/D MESFET process previously developed at Honeywell.[4]

Acknowledgements

This work was supported by the SDIO/Innovative Sciences and Technology Directorate via the Office of Naval Research, Contract No. N00014-86-C-0800. The authors gratefully acknowledge the technical assistance of E. Kalweit, M. Roth, R. Martin, C. Kelley and W. Goldberg.

References.

1. J.J. Hopfield, Proc. Natl. Acad. Sci. USA 79 (1982) 2551.
2. N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, Applied Optics 24 (1985) 1469.
3. K.S. Huang, A.A. Sawchuck, B.K. Jenkins, P.Chavel, J.M. Wang, A.G. Weber, C.H. Wang, I. Glaser, Proceedings of ICO Optical Computing 88, SPIE Proc. 963 (1988) 1.
4. H.K. Chung, G.Y. Lee, K.L. Tan, K. Betz and P.J. Vold, Proceedings of the GaAs IC Symposium (1986) 15.

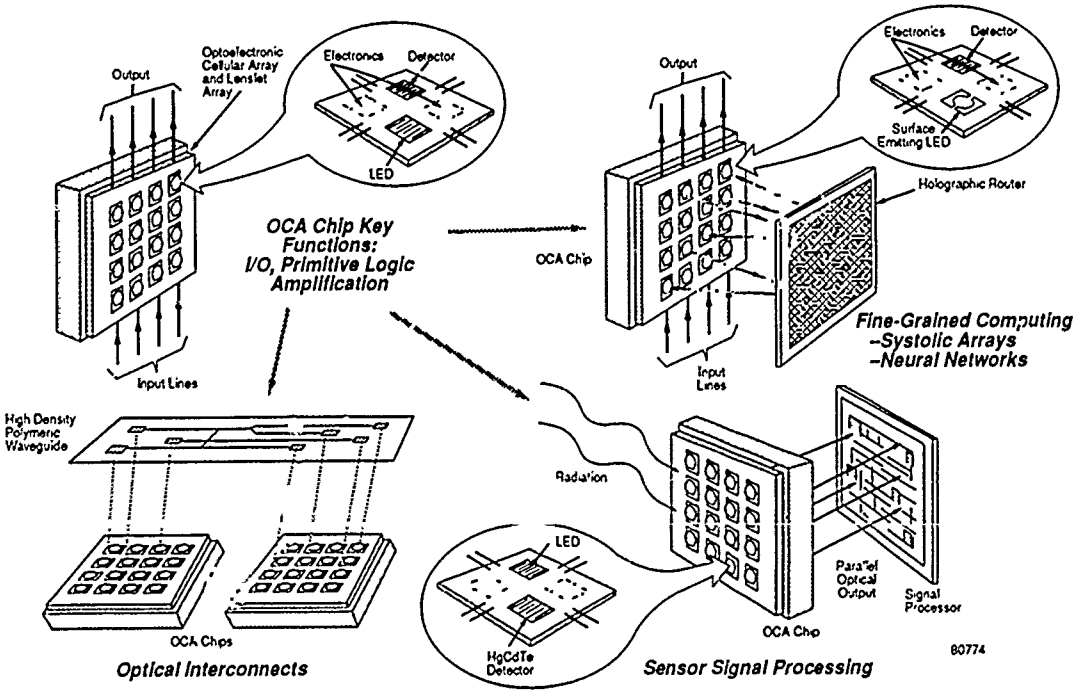
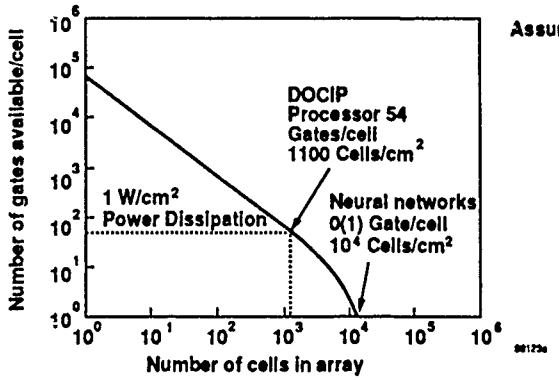


Figure 1. Applications of OCA Chip

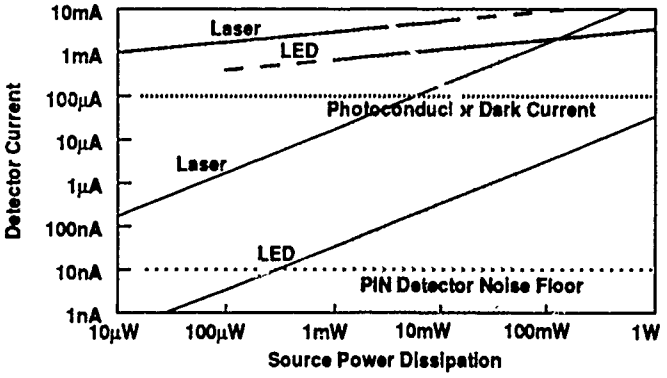


Assumptions: FETs run at 10-100KHz, power/gate=15 μW

Source: 1.5V, 10mA, duty cycle 0.1% (average power 15 μW)

Amplifier: 5V, 10mA, duty cycle 0.1% (average power 50 μW)

Figure 2. Number of FETs per Cell for Data Processing.

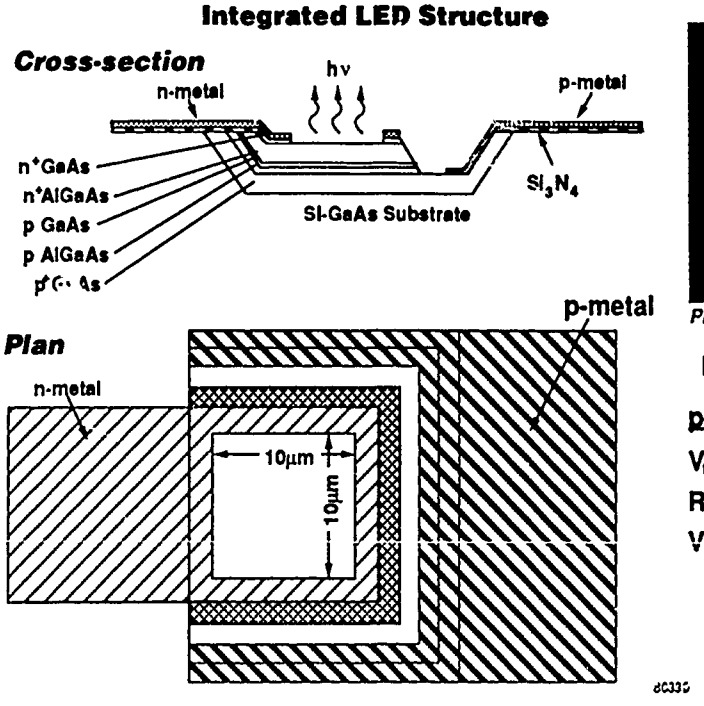


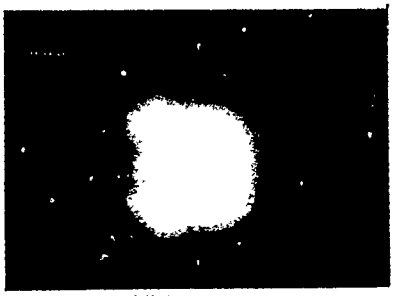
— Photoconductor responsivity varies from 7000 A/W at 0.1 μW to 10 A/W at 500 μW incident on detector

— PIN photodetector responsivity = 0.35 A/W

LED Efficiency = 0.001
Laser Efficiency = 0.5
Interconnect Efficiency = 0.1
Fanout = Fanin = n

Figure 3. Photodector Current vs. Source Power Dissipation





Photograph of light emitted from LED

Device Characteristics

p-n Diode:

$V_{turn\text{-}on} \simeq 1.2v$, n=1.55

$R_s \simeq 500$ ohms

$V_{br} \simeq 6V$

Figure 4. Integrated Optoelectronic Cellular Array

Optoelectronic Parallel Processing Arrays: System Architecture and Progress Toward a Prototype

Timothy J. Drabik and Thomas K. Gaylord
Georgia Institute of Technology
School of Electrical Engineering
Atlanta, GA 30332

I. Introduction

It is now generally acknowledged that incorporating the advantages of optics into digital multiprocessor systems can benefit performance. Considerable interest exists in the implementation of massively parallel systems tailored to problems having high inherent parallelism, that use optics to implement some or all of the tasks of interconnection, logic, and clocking.[1,2,3] We are exploring the design of digital systems based on two-dimensional arrays of electronic processing elements (PE's) produced at or near the wafer-scale level of integration, that are optically interconnected and synchronized by means of light modulators and detectors incorporated within the PE's and an external optical routing network. Provided a fast, reliable light modulator technology is developed that is compatible with silicon or GaAs VLSI circuit technology and signal levels, such systems will compete with all-electronic and all-optical computing systems in the regime of highly parallel and structured computation.

We describe below a system methodology that implements logic and local, intra-PE interconnections in electronic circuitry, and global timing and inter-PE interconnections optically, thus making optimal use of electronic and optical elements. This approach is contrasted with all-optical methodologies in terms of logic design, optical interconnect complexity, and physical limitations on size, density, and speed. Finally, a prototype shift-connected single-instruction, multiple-data (SIMD) array under construction at Georgia Tech is described, and the performance of fabricated light modulator arrays, silicon photodetectors, and silicon logic building blocks is discussed.

II. Optoelectronic Processing Array Methodology

We will discuss arrays of similar or identical processing elements based on silicon VLSI technology. The central idea is to use electronic and optical elements each to their greatest advantage. Logic is implemented with electronic devices. Interconnection is hierarchical since wire interconnects perform poorly over large distances, compared with optical interconnects,[4] but very well over short distances. Electrical interconnects are therefore used over short distances, and optical interconnects for global communication. This precept is observed in a natural way by realizing all interconnections within a PE electrically, and all communication among PE's optically. Optical interconnect complexity is minimized by adopting bit-serial communication among PE's, which admits the minimum of one modulator cell per PE. In turn, bit-serial computation is adopted within the PE both to match the serial optical data format and to decrease the PE complexity, and thereby increase the available parallelism, without reducing functionality.

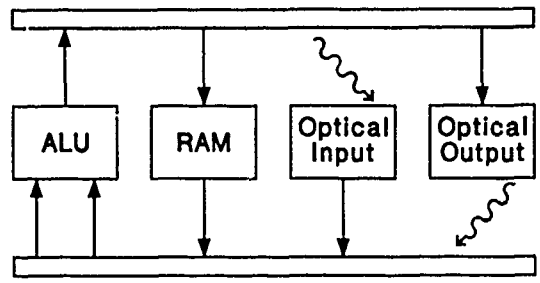


Figure 1. PE functional block diagram for shift-connected SIMD array implementation.

PE complexity is a design degree of freedom, but is limited by electrical interconnect performance and clock skew. Also, increased PE area reduces the number of PE's on a given wafer or chip and therefore the overall parallelism. It will be seen that a certain minimum PE complexity is required in very fast systems to avoid performance limitations arising from optical path latency.

A. Two examples. We describe two system scenarios to illustrate possible PE configurations. Figure 1 depicts a design suitable for an optoelectronic implementation of the shift-connected SIMD array, as described in [1, 5]. Each PE contains a 1-bit wide arithmetic-logic unit (ALU) that performs bit-serial addition, subtraction, multiplication and masking operations on data stored in its local random-access memory (RAM). Optical input and output allow inter-PE data movement. External optical interconnection of a square array of such PE's by programmable image shifts enables the shift-connected array to perform element-by-element arithmetic on arrays of numbers that can have an arbitrary relative shift, up to the extent of the array.[1] This architecture has been shown to admit efficient parallel algorithms for a wide variety of numerical problems.

Timing and control information must also be broadcast to the PE's. This can be accomodated by providing each PE with additional optical inputs for clock and control. However, the strictly feed-forward nature of clocking and control in this system renders latency in these signal paths harmless. It may therefore be beneficial to share optical inputs for timing and control among a local cluster of PE's. Layouts are possible that retain periodicity of optical data inputs and outputs and also allocate space for shared clock and control distribution. Here, a new optimal transition point from the electrical level of interconnection to the optical has been chosen that reflects the lesser importance of wire propagation delay for clocking

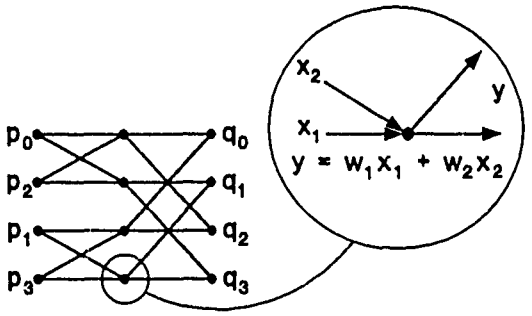


Figure 2. PE's as nodes in a radix-2 butterfly processor array.

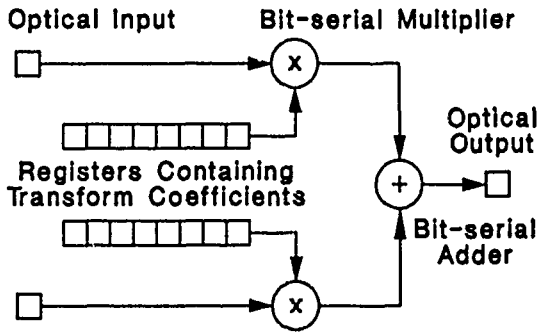


Figure 3. PE functional block diagram for pipelined radix-2 systolic array.

and control signals. The size of a local cluster is still constrained by timing skew.

The second scenario concerns an optically interconnected systolic array for the computation of fixed-point, radix-2 numerical transforms. If we associate one PE with each node of a radix-2 flow graph, as shown in Fig. 2, then each PE need only perform a computation of the form $y = w_1x_1 + w_2x_2$. A suitable PE design appears in Fig. 3. Operands enter the PE serially, least-significant bit first, and the result appears at the optical output after a constant delay, least significant bit first. Storage is required only for the transform weights w_i, as the result of the addition is passed on to the next stage bit by bit, as it is computed. A fixed optical interconnect distributes the result to two PE's in the next stage of the network. If the data representation uses k bits, a new transform is computed approximately every $2k$ bit times. Clocking and word synchronization information are distributed optically to each PE or cluster of PE's. The transform coefficients are downloaded over the data paths. This design extends in an obvious way to radix-3 transforms or to any transform of constant radix.

We can use this example to illustrate how optical interconnects can provide fault tolerance in wafer-scale systems: a wafer full of PE's is fabricated, complete with optical inputs and outputs. The PE's are probed optically and those identified as functional are interconnected with a holographic optical network tailored to that particular wafer. Power connections to defective PE's may be severed with laser surgery. A wafer need contain only a specified number of good PE's in order to be useable.

B. Technology requirements. The examples discussed above require light modulators that can be integrated in large arrays into VLSI circuitry, driven by low-voltage, high-impedance sources, and switched at a speed equal to that of the logic gates. A totally suitable technology is not yet available, but promising candidates are nonlinear electro-optic polymers,[6] multiple quantum well electroabsorption devices,[7] and ferroelectric liquid crystals (FLC's).[8]

III. Optoelectronic vs. All-Optical Systems

A. Disposition of optical interconnect capacity. All-optical computers must expend optical interconnect capacity on local as well as global connections. In all-optical systems with PE's comprising dozens of primitive gates, more optical interconnect capacity will be used to provide intra-PE than inter-PE interconnections. Optical imaging system requirements are further multiplied in the case of symbolic substitution methodologies that use two or more pixels per primitive gate.[3,9] Lower bounds have recently been derived for general[10] and specific[11] optical interconnect structures that relate the growth in physical volume required to implement an interconnect to the information content inherent in the data movement. The general result is that additional optical interconnects are achieved at the expense of additional system physical extent.

In contrast, the cost of electrical interconnects can be balanced against the cost of optical interconnects in the design of optoelectronic processors. For a given amount of functionality, optical interconnect requirements will be much lower than for an all-optical system.

B. Gate density. The physical density of circuitry in an optoelectronic system is free to track the technological state of the art and is decoupled from optical considerations. To be adequately resolved in the visible by an $f/2$ imaging system, pixels in an all-optical system must be spaced at least $5\mu m$ apart. As the number of gates and therefore the array diameter grows, so must the focal length and the physical size of the system, for constant $f/\#$.

C. The case of very fast logic. An advantage often claimed of all-optical systems is the potential for switching times in the ps or fs range. However, the effect of optical path delay on system performance in such cases must be examined. We hypothesize all-optical gates that switch in 10ps. Assuming an optical path length of 10cm for such systems as described in [3, 9], there will be a 300ps signal delay in all gate-to-gate interconnections. Extensive bit-level pipelining will be required to extract the full potential performance. We illustrate this with the example of a full adder used to perform serial addition on two bit streams, as in Fig. 4. If the delay in the carry path is equal to one bit period (Fig. 4a), the carry c_0 from the sum of a_0 and b_0 is available at the input when a_1 and b_1 arrive. If the feedback path delay is 30 bit periods, however (Fig. 4b), we must use the full adder to add the least significant bits of 29 additional pairs of numbers before the carry c_0 is available to be added in with a_1 and b_1. If this is not

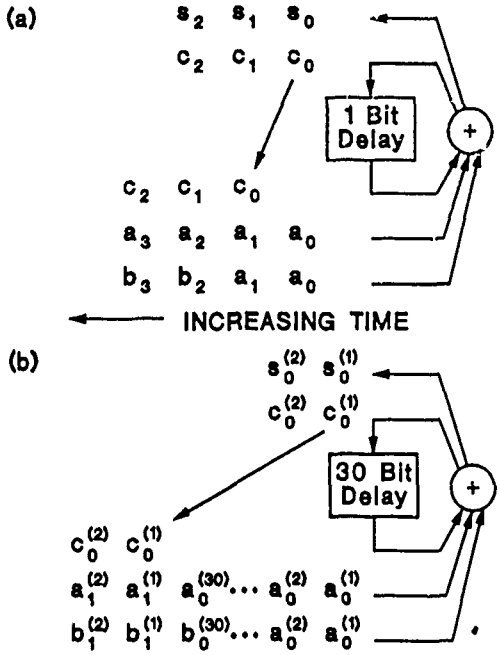


Figure 4. Efficient use of a full adder for bit-serial addition where carry feedback delay is (a) one clock period, and (b) 30 clock periods.

done, then the effective logic speed is equal to the path delay.

Logic design techniques exist for computer arithmetic that permit pipeline depth smaller than the ratio of path delay to switching time,[12] but the required degree of pipelining still grows with this ratio. The impact of pipelining requirements on systems where bit-level feedback is required, as switching times fall in the face of physically imposed interconnect delays, will be to diminish the marginal utility of faster logic.

If we construct a full adder as an optoelectronic PE with two optical inputs and one output, the carry feedback path is implemented electrically within the PE. Although electrical device characteristics remain an impediment to speed, the feedback delay is decoupled from optical considerations. Electrical implementation of critical feedback paths can be an important advantage of optoelectronic processors.

IV. System Prototype

A prototype shift-connected SIMD array based on 3μm bulk silicon CMOS technology is under construction. The PE, as shown in Fig. 1, consists of a 1-bit wide ALU capable of full addition and other logic operations, a 64-bit static RAM, an optical detector, and an optical modulator. In addition, a shift register cell is included that can be read or written by the RAM. Shift register cells of all PE's are concatenated to provide the array with electrical upload and download capability. Optical detection is implemented with silicon photodiodes or phototransistors. Optical outputs are reflective FLC light modulator cells integrated on to the silicon chip. The approximate PE layout is shown in Fig. 5. The 950μm dimension permits an 8 × 8 array

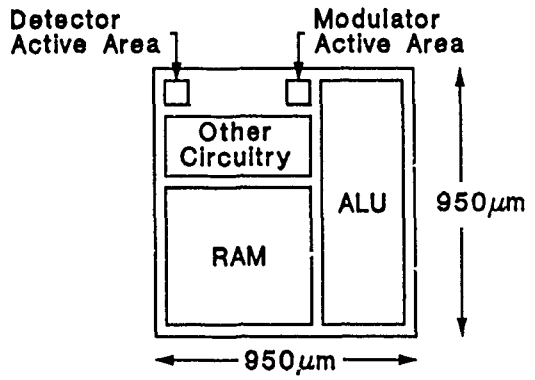


Figure 5. PE layout for prototype shift-connected SIMD array.

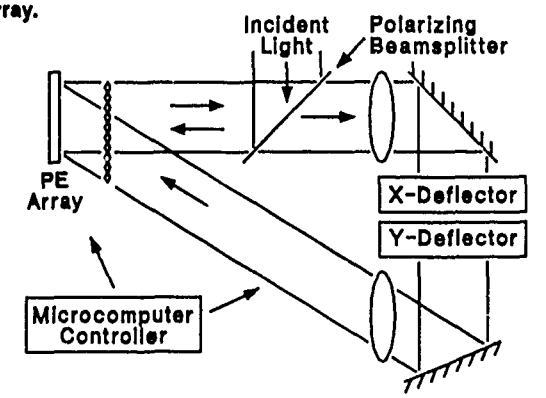


Figure 6. Prototype shift-connected SIMD system design.

of PE's to be fabricated on the largest standard die size provided by the MOS Implement on Service (MOSIS). Clock and control are electrically broadcast to all PE's and provided through the package pins. The overall system is illustrated schematically in Fig. 6. The modulator array is illuminated by a lenslet array provided by Corning Glass Works, which also collimates the beamlets exiting the modulators. The polarizing beamsplitter converts the polarization rotation of the FLC modulators to intensity modulation. Shifting of the output image is performed by programmable deflectors (galvanometers or acousto-optic devices) in the Fourier plane, and the data are then imaged onto the lenslets, which concentrate the light onto the detectors. An IBM PC XT effects control and data communication with the PE array and controls the deflectors by means of a plug-in interface board.

Several components of this design have been fabricated and tested. A chip carrying designs for a 64-bit static RAM and optical input elements based on junction phototransistors and photodiodes has been fabricated in 3μm CMOS technology and tested. All devices were found to be fully functional. The RAM has dimensions of 550μm × 650μm and an access time in the 100ns range. The detector configurations fabricated are shown in Fig. 7. Threshold sensitivities ranged from 25pW for the differential phototransistor design to 8.5μW for the single-ended photodiode design.

8 × 8 arrays of FLC light modulator cells have been fabricated and evaluated. The cell structure is shown

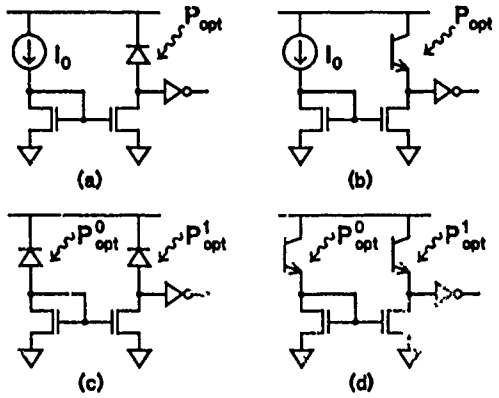


Figure 7. Fabricated detector elements with CMOS output signals: (a) single-ended photodiode, (b) single-ended phototransistor, (c) differential photodiode, and (d) differential phototransistor.

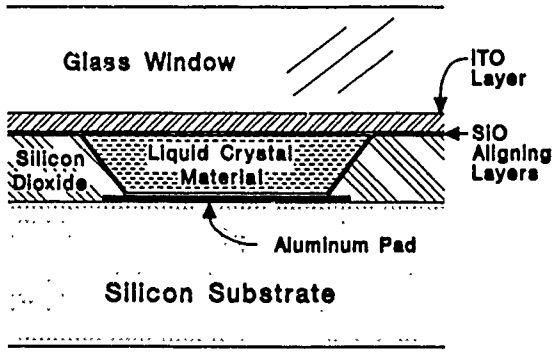


Figure 8. Fabricated FLC/silicon reflective light modulator structure.

in Fig. 8. A 1.3μm layer of SiO_2 was deposited over a 1200Å layer of aluminum evaporated onto a silicon wafer. 100μm × 100μm windows on 1mm centers were etched through the SiO_2, exposing the aluminum. A 250Å aligning layer of silicon monoxide was deposited on both the silicon substrate and on the transparent cover electrode of indium tin oxide on glass by evaporation at a 60° incident angle. The cell array was assembled and filled *in vacuo* with E. Merck ferroelectric smectic mixture ZLI-3489. Cells typically exhibited an intensity contrast ratio of 40:1.

In the eventual device, windows in the chip passivation above metal pads will be specified in the VLSI layout, and post-processing will involve only aligning layer deposition and cell assembly. The chief difficulty in fabricating such modulator arrays is the maintenance of constant cell thickness across the array.

Constructing and evaluating this prototype system will highlight the advantages of optoelectronic processor arrays as well as bring forward other important system issues and limitations.

V. Summary and Conclusion

We have presented a general design methodology for optically interconnected VLSI processor arrays based on the precept of employing optical and electronic elements where their respective characteristics are best exploited. Two examples were given to illustrate the importance of this central idea to specific computing systems. Optoelectronic processor arrays were contrasted with all-optical systems with regard to optical interconnect capacity requirements, gate density, and high-speed performance. Parallel systems combining the best features of optical and electronic elements will exhibit some advantages over all-electronic and all-optical multiprocessor systems. The required static RAM, optical detectors, and FLC optical modulators for a system prototype have been fabricated.

This work was supported by a grant from the Joint Services Electronics Program under contract no. DAAL-03-87-K-0059.

References

1. T. J. Drabik and S. H. Lee, "Shift-Connected SIMD Array Architectures for Digital Optical Computing Systems, with Algorithms for Numerical Transforms and Partial Differential Equations," Appl. Opt. **25**, 4053 (1986).
2. R. K. Kostuk, *et al.*, "Optical Imaging Applied to Microelectronic Chip-to-chip Interconnections," Appl. Opt. **24**, 2851 (1985).
3. K.-H. Brenner, "Programmable Optical Processor Based on Symbolic Substitution," Appl. Opt. **27**, 1687 (1988).
4. M. R. Feldman, *et al.*, "Comparison Between Optical and Electrical Interconnects Based on Power and Speed Considerations," Appl. Opt. **27**, 1742 (1988).
5. T. J. Drabik and S. H. Lee, "The Shift-Connected SIMD Array: a Digital Optical Computing Architecture Based on a Reconfigurable Data Passing Network," in *PROCEEDINGS OF COMPCON SPRING 1987* (IEEE, Washington, DC, 1987), pp. 258–262.
6. J. I. Thackara, *et al.*, "Advances in Organic Electro-Optic Devices," in *Nonlinear Optical Properties of Polymers* (A. J. Heeger, J. O. Orenstein, and D. R. Ulrich, eds.) (Materials Research Society, Pittsburgh, PA, 1988), pp.19–27.
7. G. D. Boyd, *et al.*, "Multiple Quantum Well Reflection Modulator," Appl. Phys. Lett. **50**, 1119 (1987).
8. N. A. Clark, *et al.*, "Ferroelectric Liquid Crystal Electro-Optics using the Surface-Stabilized Structure," Mol. Cryst. Liq. Cryst. **94**, 213 (1983).
9. T. J. Cloonan, "Performance Analysis of Optical Symbolic Substitution," Appl. Opt. **27**, 1701 (1988).
10. R. Barakat and J. Reif, "Lower Bounds on the Computational Efficiency of Optical Computing Systems," Appl. Opt. **26**, 1015 (1987).
11. M. R. Feldman and C. C. Guest, "Interconnect Density Capabilities of Computer Generated Holograms for Optical Interconnection of Very Large Scale Integrated Circuits," Appl. Opt. (in review).
12. H. S. Stone, *High Performance Computer Architecture* (Addison Wesley, Reading, MA, 1987), pp. 117–127.

TUESDAY, FEBRUARY 28, 1989

SALON D

7:30 PM–9:30 PM

TuI1–TuI32

POSTER SESSION

Programmable Emulation with the Optical Reconfigurable Logic Array

F.F. Zeise, P.S. Guilfoyle
OptiComp Corporation
P.O. Box 10779
310 Dorla Court, Suite 210
Zephyr Cove, Lake Tahoe, NV 89448

Abstract

An optically implemented reprogrammable logic array using control logic to compute ALU primitives for emulating a general purpose programmable computer.

Technical Summary

This paper discusses the implementation and programming of a high level ALU primitive instruction set from an optical PLA as described in the paper by P.S. Guilfoyle et al in this Proceedings[1]. A linear data source (25 - 65 channels) is diverged horizontally onto a SLM where independant columnar subsets are selected and converged onto a photodiode array (see fig 1a below). A detector threshold set to discriminate light from no light results in the effective calculation of boolean functions at each detector.

The SLM can be viewed as a matrix $S_{[m,n]}$ where light passage is a 'ONE' and blockage (down by > 1000) is a zero. The Detector outputs D_k are one when light is present. The data source vector pixels P_n are 'ONE' when light is emitted. Thus a high true interpretation gives the logical 'OR' of those data channels passed by the SLM. Thus D_k can be written for the high true interpretation as:

$$D_k^h = \sum_n P_n * S_{[k,n]} = \overline{D_k^l}$$

A low true interpretation (no light = 1) by DeMorgans Law gives:

$$D_k^l = \overline{D_k^h} = \prod_n \left[\overline{P_n * S_{[k,n]}} \right] = \prod_n \left[\overline{P_n} + \overline{S_{[k,n]}} \right]$$

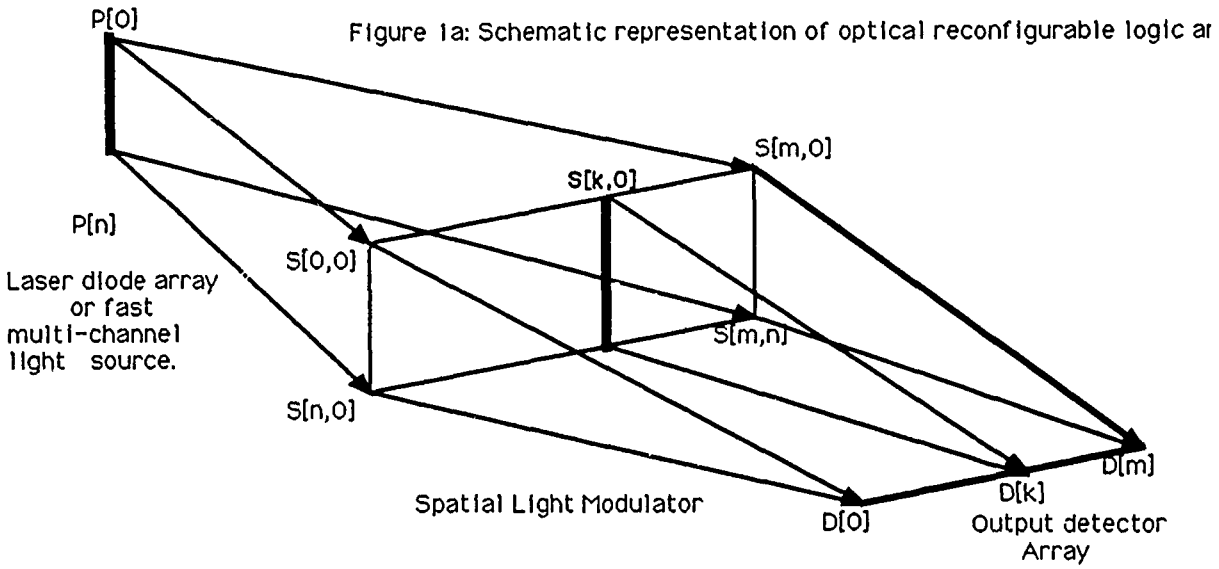
Thus for the low true interpretation, a product of arbitrarilly selected terms is obtainable at each detector from each individual column of the SLM. The data source pixels must be off for a 'ONE' for this interpretation. The effect of a ONE in the SLM is again to select the corresponding input signal for inclusion in the AND expression. Multiple AND expressions can be 'OR'ed together either in the detector electronics or by feeding the detector outputs back to the photodiode array and combining terms in a high true interpretation (logical OR) in a second pass through the ORLA. Thus arbitrary SUM of PRODUCTS expressions can be evaluated optically.

Logical Primitive Implementations

These examples will assume a data input array of length "N" and will discuss implementing 32 bit wide primitives for the emulation level. Data array inputs can be high true, low true or dual rail (both high true and low true) as needed for efficient SLM utilization. Otherwise either the SLM width or the calculation time would double for most primitives.

Bitwise OR: This is evaluated by combining N/2 A-inputs with N/2 B-inputs to generate N/2 bits of output data using high true signals. Thus with 32 data channels, 16 A-bits and 16 B-bits can be ORed at once. A typical 32 bit operation is done in two passes. See fig 1b below.

Bitwise AND: This is calculated using low true logic. Thus 16 bits of low true A are combined with 16 bits of low true B to provide 16 bits of low true AND output. A typical 32 bit AND operation is done in two passes.



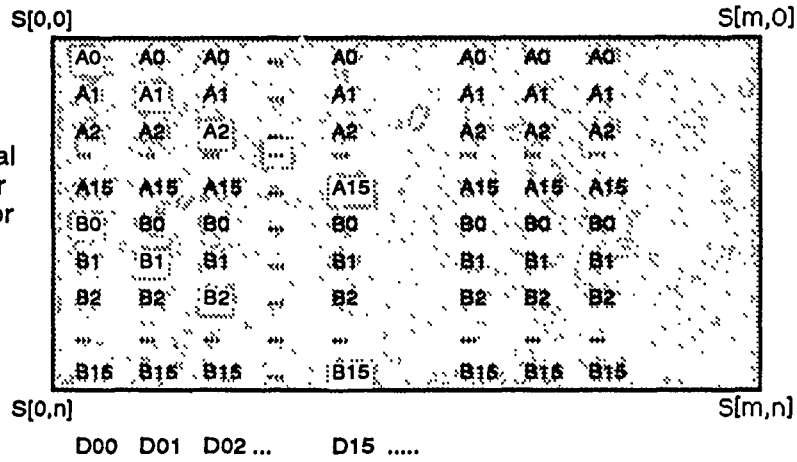


Fig. 1b: Spatial Light Modulator Programming for 16 bit AND-OR primitive

Bitwise Exclusive OR: This primitive for 2 inputs requires 2 terms for each output. Each input is required in both high true and low true polarities (A_n, $\overline{A}_n$, B_n, $\overline{B}_n$). $X_n = A_n \oplus B_n = A_n * \overline{B}_n + \overline{A}_n * B_n$ Thus with 32 data channels an 8-bit XOR is calculated using 16 detectors. The logical or of the output terms is calculated either by the detector electronics or by routing the data back through again. A 32 bit XOR operation is calculated in 4 passes and the 2-term OR's can be done in two passes.

ADD and SUBTRACT word: These primitives are very similar in that both calculate a 3-input XOR on the data bits A_n and B_n and the carry, borrow input C_{n-1}. $S_n = A_n \oplus B_n \oplus C_{n-1}$ [2] The only difference between add and subtract occurs in the carry / borrow calculation where in A - B the A_n appears complemented as $\overline{A}_n$ The carry in calculation can be done bit serially or via look-ahead for some number of terms. Each C_n is calculated in a carry lookahead by checking lower order terms for a carry generate (2 or more 1's) or a carry propagate (1 or more 1's) from a previous carry generate. Thus C_n can be expressed as the sum of the possibilities that a carry (or borrow) occured at each lower bit position and propagated to the bit position of interest. It is useful to define the intermediate functions Carry Generate and especially Carry Propagate in order to avoid exponential expansion of the number of terms required in a carry lookahead. These functions also enhance the intelligibility of the resulting equations shown below:

The generalized carry generate equation for n-bit addition carry look-ahead is:

$$C_{gn} = [A_n \cdot B_n]$$

The generalized carry propagate equation for n-bit addition carry look-ahead is:

$$C_{pn} = [A_n + B_n]$$

Thus the carry output from the lowest order bit can be written:

$$C_0 = C_{g0} + C_{p0} \cdot C_{in} = [A_0 \cdot B_0] + [A_0 + B_0] \cdot C_{in}$$

The carry output from the next higher bit can be set from either the current bit position or be propagated from a carry generate in previous bit positions as shown:

$$C_1 = C_{g1} + C_{p1} \cdot C_{g0} + C_{p1} \cdot C_{p0} \cdot C_{in}$$

This expands to:

$$C_1 = [A_1 \cdot B_1] + [A_1 + B_1] \cdot [A_0 \cdot B_0] + [A_1 + B_1] \cdot [A_0 + B_0] \cdot C_{in}$$

These equations are implemented in the SLM depicted in figure 2 below:

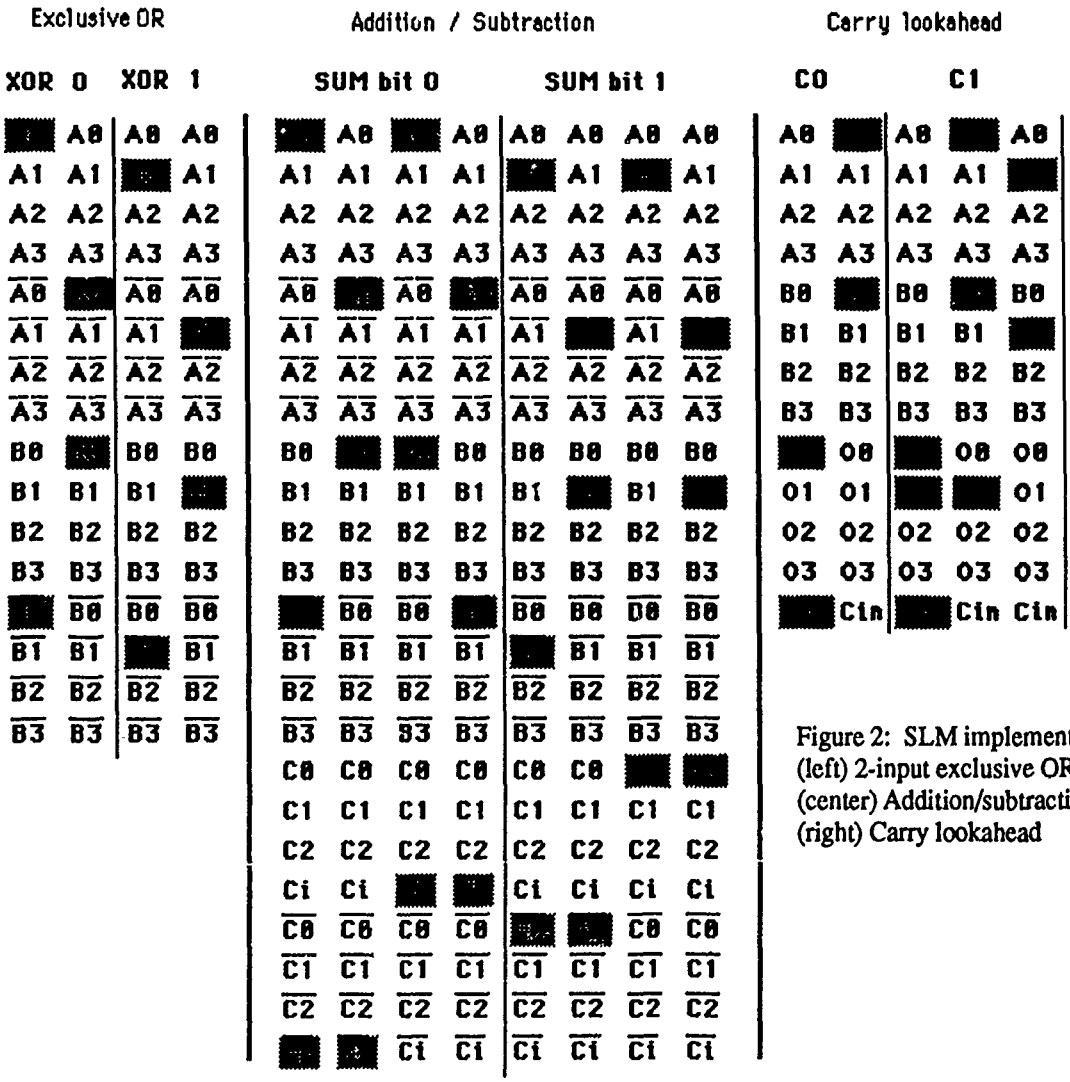


Figure 2: SLM implementations (left) 2-input exclusive OR, (center) Addition/subtraction, and (right) Carry lookahead

Thus for SLM calculation of carry lookahead 3 sets of input signals are used: A_n, B_n, $O_n = C_{pn} = A_n + B_n$, and C_{in}. For subtraction $\overline{A_n}$ is substituted for A_n in the above carry equations. The optical PLA makes calculation of multiple terms relatively easy. Routing data back to the inputs for use in the next calculation takes extra clock cycles so the optimal tradeoff tends towards increased width. The calculation of the nth bit of carry lookahead (1=lowest) takes n+2 terms. Thus to calculate carrys for w bits at a time takes $(w^2+3w)/2$ terms. The data inputs are A_n, B_n, $O_n = A_n+B_n$, and C_{in} for 3w +1 input channels. The best tradeoffs when data routing is condsidered are for widths of 8 and 16 for 25 or 49 input channels used. The SLM depth requirement increases rapidly with w : 44 terms or colums for 8 bits and but only 152 terms for a full 16-bit flash carry lookahead. These operations are repeated to calculate the full 32 bits of carry information. The sum output calculation evaluates the 3 input XOR as a sum of AND terms which requires 4 terms/output bit as given below:

$$S_n = C_{n-1} \text{ xor } A_n \text{ xor } B_n = C_{n-1} \cdot \left[A_n \cdot B_n + \overline{A_n} \cdot \overline{B_n} \right] + \overline{C_{n-1}} \cdot \left[A_n \cdot \overline{B_n} + \overline{A_n} \cdot B_n \right]$$

Thus high level logical primitives of an arbitrary width say 32 or 64 bits can be calculated with an SLM by sequentially stepping through the word width executing lower level primitives according to the available width of the SLM. Thus for a 32 channel (ch) light source and a 32(ch) x 128(d) SLM high level primitives of width P_w can be implemented as shown in table 1.

TABLE 1: PRIMITIVE IMPLEMENTATION TRADEOFFS

function	polarity	signals	Optical primitive width	W bits	SLM columns (d) / output	prim. width	for Pw bits	Number of sections	Pw=32 bits
OR	high	A, B	ch /2	(16)	1	ch/2	(16)	2Pw /ch	(2)
AND	low	A, B	ch /2	(16)	1	ch/2	(16)	2Pw /ch	(2)
XOR	low	$A,\overline{A}, B,\overline{B}$	ch /4	(8)	2	ch/2	(16)		
ADD Sum	low	$A,\overline{A}, B,\overline{B},C_n,\overline{C_n}$	ch /6	(4-5)	4	4*ch/6	(16-20)	6Pw /ch	(8-7)
Add Carry	low	A, B, C_{pg}, C_{in}	ch /3	(8-10)	w+2	$(w^2+3w)/2$	(48)	4Pw /ch	(4)

Thus with a 32 bit wide SLM a 32 bit 'OR' operation can be completed in 2 passes through the SLM in the following sequence. 1) calculate low order OR, 2) calculate high order OR.

A 32 bit XOR operation can be completed in 4 passes through the SLM in the following sequence.
1) Calculate low byte XOR 2 .. 4) Calculate successive output bytes.

A 32 bit ADD operation can be completed in 14 passes through the SLM in the following sequence.
1) Calculate Carry Propagate C_{pg} = A OR B 2 passes for 32 bits
2) Calculate carry look ahead C_n 8 bits at a time 4 passes for 32 bits
3) Calculate Sum outputs 4 bits at a time 8 passes for 32 bits

Additional operations can be built up such that an arbitrary width ALU - Arithmetic Logical Unit with arbitrary primitives can be implemented using the optical reprogrammable logic array. For example with some additional sequencing and electronic control circuitry a simple (Reduced Instruction Set Computer[3,4]) like a SUN SPARC (Scalable Processor ARChitecture) chip can be emulated. An emulation of only the user mode environment provides implements an expandable and versatile general purpose platform for further studies in optical computing architectures with an integrated software development environment.

REFERENCES

[1.] P. S. Guilfoyle, F .F. Zeise, "Reconfigurable Programmable Optical Digital Computer", 1989 Topical meeting on Optical Computing, (previous paper this session), OSA Feb 1989.

[2.] Swartzlander, Computer Arithmetic, Dowden, Hutchinson and Ross 1980.

[3.] M.G.H. Katevenis, Reduced Instruction Set Computer Architectures for VLSI, MIT Press 1986.

[4.] SUN Micro Systems, Mtn View Ca., The SPARC Architecture Manual, 1988

Optical Multiple-valued Logic Using Composite Bistable Laser Diode or Light Emitting Diode Circuit*

Shutian Liu Chunfei Li Jie Wu and Yudong Liu
Department of Physics, Harbin Institute of Technology
Harbin, People's Republic of China

SUMMARY

Multiple-valued logic offers many times the logic power and packing desity of binary logic, thus reducing the number of necessary logic gate [1]. Recently optical multiple-valued logic has been studied extensively using shadow casting method [2]. In this paper we describe a new configuration of optical multiple-valued logic gates using composite bistable laser diode or light emitting diode circuits (BILD or BILED). To our knowledge, this is the first time that optical multiple-valued logic have been obtained using optical bistable device. We have demonstrated four Post logic functions: Complement, Max(x,y), Min(x,y), and Suc(x) using composite BILED circuits. We only take Complement and Max(x,y) for example here.

Fig.1(a) shows the schematic diagram of the complement gate, two BILED circuits and a LED are connected in parallel. This configuration of BILED was first described by Y.Ogawa et.al.[3],which only consists of a phototransistor and a light emitting diode. In our experiment, we used the transistors as amplifiers. The resistors Rs and Ri(i=1,2,3) control the output power level; Rlj(j=1,2) control the switching-on power Pon1 and Pon2 of each BILEDs . The operation principle of ternary complement gate is simple .The current through Rs is tristable

due to the currents distribution between two parallel connected BILEDs with the BILEDs switching on at different input power.[4] The output power Po is downwards tristable vis a triangle wave input (Fig.1(b)). We take the output power Po of LED3 as the output of the Complement gate. The switch-on power Pon1 and Pon2 can satisfy the following condition by adjusting Rlj:

$$0 < Pon1 < 1,\ 1 < Pon2 < 2.$$

When the ternary input signal X=0, BILEDs are not switched on and all the current goes through LED3 and hence Po=2; if X=1, BILED1 is switched on and one unit of current goes through BILED1, then one has output Po=1; if X=2, both BILEDs are switched on and all the current goes through BILEDs, therefore Po=0. As a result of the operation above, the output Po and input ternary signal X have the following relation:

$$Po=\bar{X}=X-2,\ X=(012).$$

Fig.1(c) shows the input-output signal wave forms. In our experiment, the input and output light levels were indicated by the current through LEDs. For the level "1" and "2" the current were 9.5mA and 19mA, respectively. The current through LED3 was about 0.8mA for output Po=0.

Max(x,y) and Min(x,y) are very important logic functions in Post algebra for they are equivalent to OR and AND gates in binary logic. For ternary logic, Max(x,y) and Min(x,y) logic are defined as:

$$Max(x,y)=X,\ \text{if } X \geqslant Y,\ \text{with } X,Y=(012).$$

$$Min(x,y)=X,\ \text{if } X \leqslant Y,\ \text{with } X,Y=(012).$$

Fig.2(a) is the schematic circuit diagram of ternary Max(x,y) gate with four BILEDs connected in parallel. LED1 and LED2, LED3

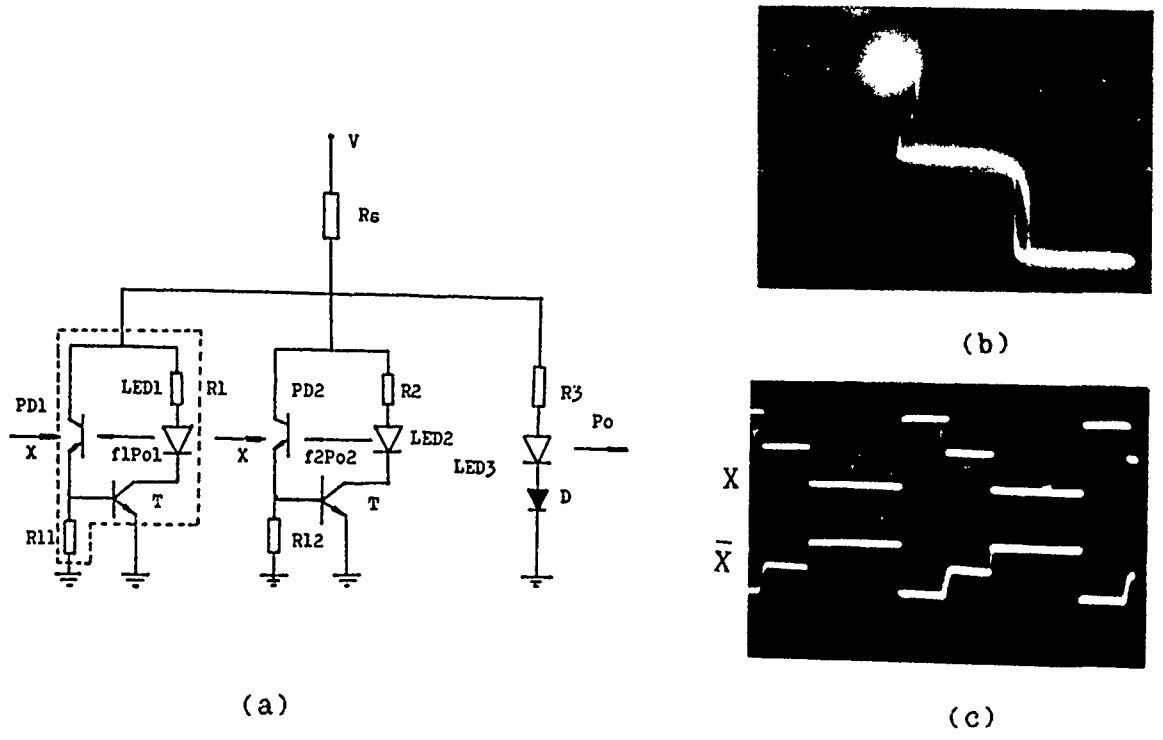


Fig.1.(a). Schematic diagram of ternary Complement gate.
(b). Outpul Po vis a triangle wave input.
(c). Input and output signals wave forms.

and LED4 are connected together, respectively. We take the output Po of LED5 as the output of Max(x,y) gate.To complete Max(x,y) function, BILEDi(i=1,2,3,4) must satisfy the following conditions:

$$0 < Pon1 = Pon2 < 1,\ 1 < Pon3 = Pon4 < 2.$$

In this working condition, BILED1 and BILED2 operate as one unit due to the existence of optical feedback. But they have two individual ternary input signals X and Y. Therefore they can be controlled either by X or Y. BILED3 and BILED4 operate in the same way as BILED1 and BILED2. If X = 1 or Y = 1, both BILED1 and BILED2 arc switched on, the output Po=1; if X=2 or Y=2, all the BILEDs are switched on, the output Po=2. Fig.2(b) is the experiment result of Max(x,y) gate. Using a four BILEDs combina-

tion circuit, one can also easily obtain Min(x,y) gate.

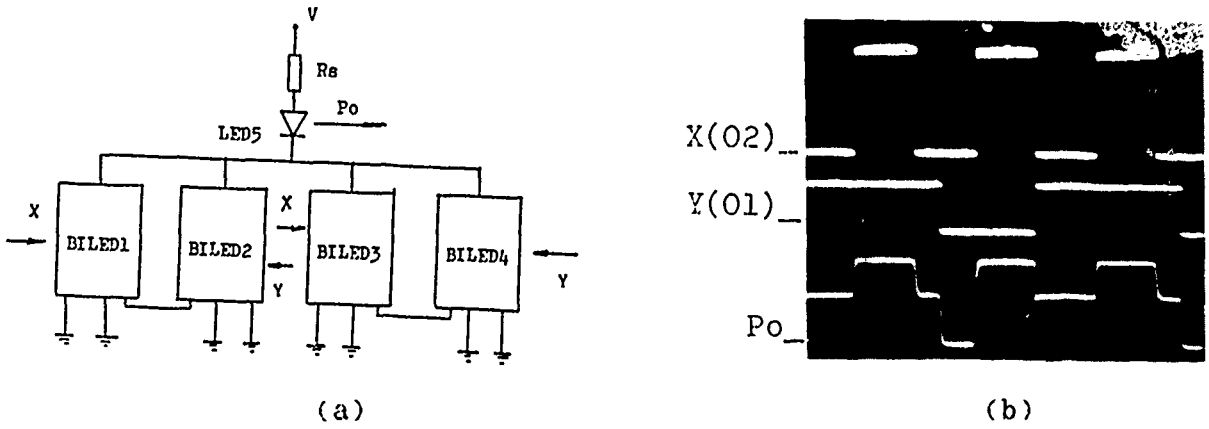


Fig.2.(a). Schematic diagram of ternary Max(x,y) gate.

(b). Input and output signals wave forms.

Composite BILD or BILED circuits are demonstrated to be promising for electro-optical hybrid digital computing, optical signal processing and optical telecommunication. These devices have many attractive features: low power and incoherent light operation; input amplification; and more available logic functions (binary and multiple-valued logic) which are necessary for digital optical computing. The Very Large Scale Integration fabrication technology will certainly make these devices more practical.

References:

1. T.T.Dao and D.M.Campbell, Opt.Eng.,25,014,1986.
2. R.Arrathoon and S.Kozaitis, Opt.Eng.,25,029,1986.
3. Y.Ogawa, H.Ito, and H.Inaba, Appl.Opt.,21,1878,1982.
4. Chunfei Li, Shutian Liu, and Jie Wu, Conference on Laser and Electro-optics, Baltimore, April 26-May 1, 1987.

An Optical-Holographic-Associative-Memory-Based Parallel Register Transfer Processor

George Eichmann, Andrew Kostrzewski, Dai Hyun Kim, and Yao Li

Department of Electrical Engineering,
The City College of the City University of New York,
New York, New York 10031.

A combinatorial logic circuit is an interconnected array of logic gates. However, for various arithmetic operations, iterative sequential computation is needed. To furnish feedback, memory elements, such as flip-flops or registers must be utilized. With this feedback, the overall logic circuit is a finite-state sequential logic machine. The use of optics to perform fast combinatorial logic processing was suggested[1-3]. However, for the various proposed combinatorial logic elements the efficient feedback generation is an active research area. To generate a sequential logic circuit, a viable hybrid approach is to use optics for both fast parallel logic and interconnect and high-speed bit-addressable electronics for storage and feedback [1]. In this paper, a specific hybrid sequential computing module, where optical array processors that perform the combinatorial logic and interconnect operations, are sandwiched between high-speed electronic parallelly-addressed storage registers, is described. This hybrid system can sustain various fast optical register transfer micro-operations (ORTMOs), operations that are the most primitive operations required for an optical digital computer. This new system will be referred to as an optical register transfer processor (ORTP).

For the design of a digital computer, the so-called register transfer language (RTL)[4] plays an important role. Based on an interconnected set of logic gates, registers, etc., RTL serves as the most primitive language that links a physical digital machine and its programmers. Any sophisticated iterative computation can be decomposed into a micro-sequence of logic and transfer operations. In Table I (II), some typical register transfer (logic) micro-operations[4] are listed. Here, the source and destination registers are denoted as A and O, respectively. It can be shown that for both a single bit and a full-length word, register parallel load, clear, rotate and shift as well as transfer operations are executable. These transfer operations together with a complete set of binary logic micro-operations can be combined for other more sophisticated arithmetic operations, such as addition, subtraction, and multiplication.

Table I. Inter-register transfer micro-operations.

Microoperation	Explanation
$C \leftarrow A$	Transfer A into C
$C \leftarrow srA$	Shift A right by 1-bit and transfer into C
$C \leftarrow slA$	Shift A left by 1-bit and transfer into C
$C \leftarrow rrA$	Rotate A right by 1-bit and transfer into C
$C \leftarrow rlA$	Rotate A left by 1-bit and transfer into C
$C_i \leftarrow A_j$	Transfer j^{th} bit of A into i^{th} bit of C

Table II. Register logic transfer micro-operations.

Binary logic	Microoperation	Explanation
$O_0 = 0$	$C \leftarrow 0$	Reset
$O_1 = 1$	$C \leftarrow 1$	Set
$O_2 = A$	$C \leftarrow A$	A
$O_3 = B$	$C \leftarrow B$	B
$O_4 = \overline{A}$	$C \leftarrow \overline{A}$	Complement
$O_5 = \overline{B}$	$C \leftarrow \overline{B}$	Complement
$O_6 = A \bullet B$	$C \leftarrow A \bullet B$	AND
$O_7 = A \bullet \overline{B}$	$C \leftarrow A \bullet \overline{B}$	AND
$O_8 = \overline{A} \bullet B$	$C \leftarrow \overline{A} \bullet B$	AND
$O_9 = A + B$	$C \leftarrow A + B$	OR
$O_{10} = A + \overline{B}$	$C \leftarrow A + \overline{B}$	OR
$O_{11} = \overline{A} + B$	$C \leftarrow \overline{A} + B$	OR
$O_{12} = A + B$	$C \leftarrow A + B$	XOR
$O_{13} = A + B$	$C \leftarrow A + B$	XNOR
$O_{14} = \overline{A \bullet B}$	$C \leftarrow \overline{A \bullet B}$	NAND
$O_{15} = \overline{A + B}$	$C \leftarrow \overline{A + B}$	NOR

For an optical register transfer micro-operation (ORTMO) implementation, the recently developed symbolic substitution technique can be used. In our approach, an optical holographic associative symbolic substitution (OHASS) technique, proposed by Yu et. al.[5] is employed. The two logic states 0 and 1 are encoded as two spatial orthogonal symbols. In the OHASS filter preparation stage, the interference pattern between the Fourier spectra of the input and the precalculated output symbols are recorded. To process either a single or two-variable RTMO, either two or four exposures at the correspondingly partitioned recording plane is affected. To generate an output, the input symbols are used as the reference beams. Details for the construction of an OHASS filter can be found in Ref. [4].

When a set of parallel micro-operations are required, by placing a duplicating grating past the input at the input plane, a parallel set of displaced Fourier spectra is used for the parallel holographic associative matching. In Fig.1, a schematic of a complete

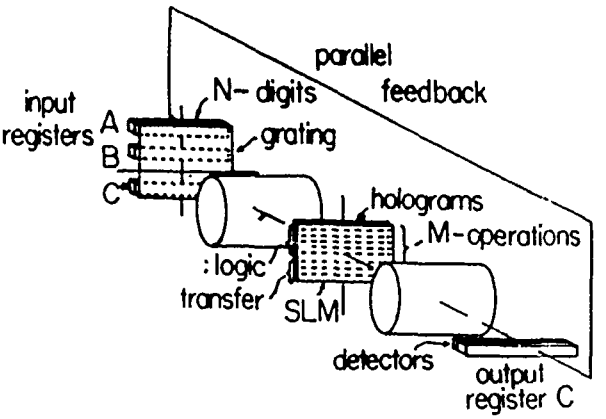


Fig.1 A schematic of a N-bit OHASS iterative processor. A, B and R, are three N-bit input registers driving channelized laser diodes; C, an N-bit output register storing the result of optical threshold detector array. In addition to the lenses, holograms, and an input duplication grating, a Fourier plane 2D SLM and a parallel electronic feedback are used.

optical register transfer processor (ORTP) is shown. The three N-bit inputs A, B, and R and the one N-bit output C electronic registers are employed. The input electronic registers are used to drive a parallel 1D array of fast laser diodes, while C stores the outputs of a fast optical detector array. For iterative computing, an one-to-one electronic feedback loop connecting A to C can be utilized. The register B acquires input data from an electronic system output port. In a learning phase to program the ORTP, register R is used. To actuate a microoperation, one of the M vertical Fourier spectrum replica is used. To ensure that all the two-variable input combinations are available, four OHASS bit-wise partitioned exposures are required. When all the M ORTMOs are encoded, the register R is deactivated. To control the ORTP's sequencing, located at the back of the hologram array, a 2D spatial light modulator (SLM) programmed to select, one at a time, one of the M horizontal slices, is employed. The thus selected result, after passing through the second cylindrical lens L_2, is detected, thresholded and then stored in the register C.

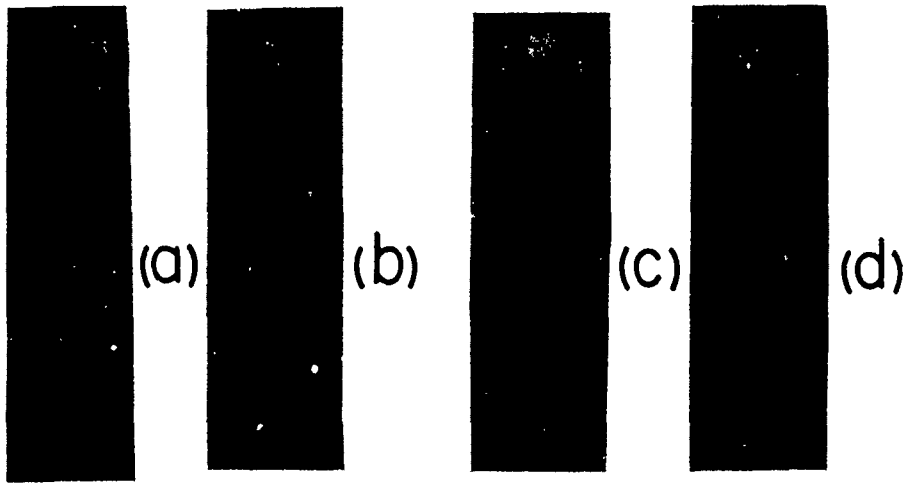


Fig.2 Results of a 1-bit OHASS interregister transfer micro-operation. (a) and (b), an associative transfer of a symbolic 1 and 0, respectively. (c) and (d), the associative complement of a symbolic logic 0 and 1, respectively. The top and bottom patterns are the input and output symbols, respectively.

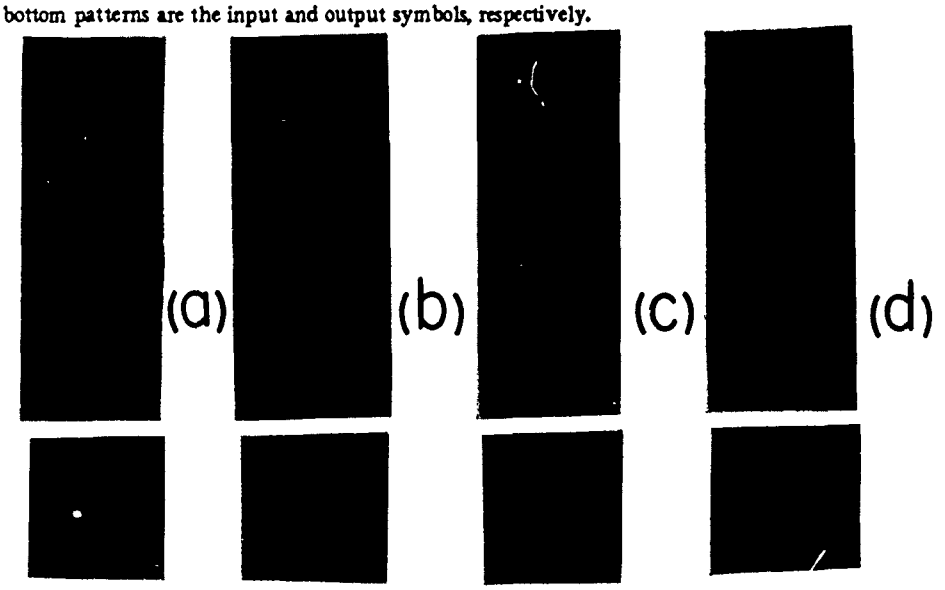


Fig.3 Results of a 1-bit OHASS logic AND micro-operation. (a)-(d), the associative AND operation results of the four input binary symbol pairs.

Because with an ORTP, using an OHASS, both logic and transfer operations are performed, the operation cycle time is equal to the free-space input-output beam propagation time. When the longitudinal dimension of the system is reduced to, say 1 cm, processing of N parallel bit pairs requires only about 33 picoseconds, independent of the word length. Since all the registers store the parallel data and the intermediate results for a short time, and because no serial intraregister operation is required, fast GaAs-based GHz electronic registers together with a fast system clock can be used. For the future all-optical ORTP, optical memory elements, such as the recently developed SEED[6] device, that can offer a dynamic storage for as long as 30 sec., may be used. In addition, because all microoperations require an identical processing circuitry, system synchronization is relatively easy. Finally, because of the ORTP speed is independent of the word length, by using a large space-bandwidth product optical system, an overall fast parallel ORT processing can be accomplished.

In our proof-of-principle ORTP experiment, parallel OHASS bit transfer, logic complement and AND operations were performed. For the duplication of

the Fourier spectrum into three laterally displaced spatial locations, two beam splitters and a mirror were used. For the bit transfer, negation and logic AND micro-operations, respectively, these three Fourier spectra were used. In Fig.2(a) and (b), the result of these register transfers are shown. To select a desired microoperation, at the Fourier plane, a binary mask was employed. On the left-most Fourier spectrum, an optical bit transfer micro-operation was performed. Since there are two bit transfer cases, i.e the transfer of either a 0 or a 1, the hologram associated with these two transfers was divided into two vertical parts. At each exposure, at the input plane, identical input symbols were used. To block one half of the spectrum, a sharp razor blade was used. The center Fourier spectrum was utilized for a bit complement logic operation. In this case, for each exposure, a pair of different binary symbols were used. In Fig.2(c) and (d), the two OHASS logic complement results are shown. To perform a two-variable logic AND operation, the right-most Fourier spectrum was used. In this case, a four quadrant composite hologram was constructed. For each exposure, the other three spectral quadrants were covered. Also, for each exposure, each of the four input pair symbols was inserted. In Fig.3, an experimental OHASS logic AND results are shown. Using this method, other Table I and II micro-operations can also be performed.

This research was supported in part by a grant from the Air Force Office of Scientific Research.

References

[1] R. Arrathoon, "Logic based spatial light modulators," Proc. SPIE, 881 (1988).

[2] P. S. Guilfoyle and W. J. Wiley, "Combinatorial logic based digital optical computing," Appl. Opt. 27, 1661 (1988).

[3] Y. Li, A. Kostrzewski, D. H. Kim, and G. Eichmann, "A compact folded path free-space optical programmable logic array," Opt. Lett. in print.

[4] M. M. Mano, *Computer System Architecture*, (Prentice-Hall, NJ, 1982) ch.4.

[5] F. T. S. Yu and S. Jutamulia, "Implementation of symbolic substitution logic using optical associative memories," Appl. Opt. 26, 2293 (1987).

[6] G. Livescu, D. A. B. Miller, J. E. Henry, A. C. Gossard, and J. H. English, "Spatial light modulator and optical dynamic memory using a 6 × 6 array of self electro-optic-effect devices," Opt. Lett. 13, 297 (1988).

Optical Network Design for a BIT-serial Parallel Processor

Adolf . Lohmann, Gregor Stucke
Uni· :sity of Erlangen, Physics
8520 ɛrlangen, Fed. Rep. of Germany

When designing a future optical parallel processor one might try to get inspiration from an existing architecture with the following features:

Software can be copied from the existing electronic system;
the processing elements are simple;
the control has SIMD character;
the connection network is fairly primitive.

The desirability of existing software is quite obvious. The simplicity of the processing elements allows for optical emulation in the foreseeable future. Hence, a hybrid system with electronic processors and with some parts of the interconnections in optical form makes sense as an intermediate goal. A SIMD-type control has a better chance for being implemented optically in the near future than would be the case for MIMD-type control. Finally, if the existing electronic communication network is quite primitive (such as four-neighbour connections) the hopes for noticable progress are justified, if a sophisticated optical network can be attached to the existing parallel processor.

We have studied the architecture of a specific bit-serial parallel processor which satisfied the criteria above (G. Stucke, accepted by Appl. Opt.). In that earlier publication we concluded that the following 6 communications commands would improve the performance of the system:

cyclic shifting in +/- x direction;
cyclic shifting in +/- y direction;
perfect shuffling in the x domain;
perfect shuffling in the y domain.

In the hybrid version of our systems design each processing element would possess one register which is outfitted with LEDs for emission and with photo diodes for detection of optical data. Upon a control signal all of these registers would emit simultaneously their status optically into an optical system which is able to perform anyone of the six operations or combinations thereof. Upon return from the optical system the optical signals will be detected and stored in the appropriate registers.

The body of this paper is dedicted to the design of this optical interconnection system. It consists of lenses, prisms, beam splitters, Wollastons and FLC switches. The concept is an extrapolation of earlier more specialised designs.

References:

A.W. Lohmann, W. Stork, G. Stucke: Appl. Opt. 25 (1986) 1530
A.W. Lohmann: Appl. Opt. 25 (1986) 1543
K.-H. Brenner, A.W. Lohmann: Appl. Opt. 27 (1988) 434

Symbolic Substitution Based Parallel Adder/Subtracter

S. Barua
Department of Electrical Engineering
California State University, Fullerton, CA 92634

1. Introduction

Fully parallel processors can be designed by employing a technology that is inherently parallel, a suitable number system, and an efficient encoding scheme for handling the data. Binary number system is accepted as the best suited in electronic computers. The delay due to carry propagation in binary arithmetic makes the binary number representation a very weak candidate for an optical processor that is inherently parallel. The modified signed-digit (MSD) number system[1] satisfies the requirements of fully parallel addition and subtraction by limiting the carry propagation to one position to the left. The design of an optical MSD adder capable of performing addition/subtraction in three stages has already been proposed.[2] The above design is based on polarization-coded symbolic substitution. A reduction in the number of stages can be achieved by exploiting some of the unique characteristics of MSD. The optical implementation of the MSD adder with reduced number of stages is discussed in this paper. The MSD digits are coded as three different polarization states of light. Polarization-coded symbolic substitution is used to implement the adder.

2. MSD Number System

The MSD representation is a subset of the signed-digit representation with radix $r = 2$ and is represented by three digits, 1, 0, or $\bar{1}$. For a precision of p bits, a given decimal number X can be represented in MSD as follows:

$$X = [1, 0, \bar{1}]\, 2^{(p-1)} + \ldots + [1, 0, \bar{1}]\, 2^{1} + [1, 0, \bar{1}]\, 2^{0} \qquad (1)$$

where one of the digits from the set $[1, 0, \bar{1}]$ is selected for each term to give the appropriate representation. A negative number is represented in MSD by taking the complement of the MSD positive number. To obtain the complement, simply replace 1 by $\bar{1}$ and vice versa and leave the zeros unchanged. Subtraction of two MSD numbers is accomplished by taking the complement of the subtrahend and adding it to the minuend.

3. Carry Propagation-free Addition

The addition/subtraction discussed in this paper is carried out in two stages by generating transfer and weight digits. These transfer and weight digits are

generated in parallel, as the operands flow through the MSD addermaking addition/subtraction of any length operands possible in a constant time equal to the time required for the addition/subtraction of two MSD digits.

The two-stage MSD adder/subtracter for two 4-digit MSD numbers X and Y ($X = x_3x_2x_1x_0$ and $Y = y_3y_2y_1y_0$) is shown in Fig. 1. The architecture consists of two stages, stage 1 being implemented by functional blocks A and stage 2 being implemented by functional blocks B. For the input digits x_i and y_i, block A generates the transfer digit t_i and the weight digit w_i. In stage 2, block B receives the transfer digit t_{i-1} and the weight digit w_i as inputs and generates the final output s_i. Addition/subtraction of two large numbers can be carried out by simply adding the required number of identical function blocks A to stage 1 and B to stage 2. The transfer and weight digits $[t_i\ w_i]$ for two MSD digits $[x_i\ y_i]$ are given in Table 1.

It can be seen from the table that when the input combination $[x_i\ y_i] = [1\ 1]$, $[1\ \bar{1}]$, $[0\ 0]$, $[\bar{1}\ 1]$, or $[\bar{1}\ \bar{1}]$, there is only one possible value for $[t_i\ w_i]$ and when $[x_i\ y_i] = [1\ 0]$, $[0\ 1]$, $[0\ \bar{1}]$, or $[\bar{1}\ 0]$ there are two possible values for $[t_i\ w_i]$. Since the result from the second stage corresponds to the final output, the addition of t_{i-1} and w_i by block B in the second stage should generate no carry. The generation of carry in the second stage can only be restricted by allowing both the inputs t_{i-1} and w_i to block B to be neither 1s nor 1s. The appropriate selection of $[t_i\ w_i]$ for all the possible combinations of $[x_i\ y_i]$ can be done as follows:

Case 1: $[x_i\ y_i] = [1\ 1]$, $[1\ \bar{1}]$, $[0\ 0]$, $[\bar{1}\ 1]$, and $[\bar{1}\ \bar{1}]$

For each value of $[x_i\ y_i]$ there is only one possible value for $[t_i\ w_i]$. Hence $[t_i\ w_i] = [1\ 0]$, $[0\ 0]$, $[0\ 0]$, $[0\ 0]$, and $[\bar{1}\ 0]$ respectively.

Case 11: $[x_i\ y_i] = [1\ 0]$ or $[0\ 1]$

Check the next lower order augend and addend digits $[x_{i-1}\ y_{i-1}]$. If both x_{i-1} and y_{i-1} are positive, then choose $[t_i\ w_i]$ as $[1\ \bar{1}]$. For all the remaining combinations of $[x_{i-1}\ y_{i-1}]$, choose $[t_i\ w_i]$ as $[0\ 1]$.

Case 111: $[x_i\ y_i] = [\bar{1}\ 0]$ or $[0\ \bar{1}]$

Check the next lower order augend and addend digits $[x_{i\ -1}\ y_{i-1}]$. If both x_{i-1} and y_{i-1} are positive, then choose $[t_i\ w_i]$ as $[0\ \bar{1}]$. For all the remaining

combinations of $[x_{i-1}\ \ y_{i-1}]$, choose $[t_i\ \ w_i]$ as $[\bar{1}\ 1]$.

The input/output relationship for blocks A and B for all input combinations are shown in Tables 2 (a) and 2(b) respectively. The augend and addend digits at the next lower order position to the least significant digits are assumed to be zeros. When $[t_i\ w_i]$ are chosen as per Table 2 (a) no transfer digit will be generated in the second stage. Thus, the addition of two numbers is accomplished in two stages.

4. Polarization-Coded Symbolic Substitution Logic (SSL)

The functional blocks A and B are implemented optically using polarization-coded SSL. The MSD digits are coded using three different states of polarization of light - 1 by vertically polarized light (denoted by vertical arrow), 0 by horizontally polarized light (denoted by a horizontal arrow), and $\bar{1}$ by light polarized at 45^0 (denoted by an arrow inclined at 45^0). Fig. 2 (a) shows the input/output relationship for block A, for the input combination $[x_i\ y_i] = [1\ 0]$. Fig. 2 (b) shows input/output relationship for block B for one of its input combination $[t_{i-1}\ w_i] = [1\ 0]$. The same for the remaining combinations of $[x_i\ yi]$ and $[t_{i-1}\ w_i]$ can be derived in a similar fashion using Table 2. The function blocks when implemented using SSL should be able to recognize the input patterns (search patterns) and substitute the recognized patterns by the output patterns (scribing patterns). A complete description of the recognition of the search patterns and the substitution by the scribing patterns can be found in Ref. 2.

It should be noted that there are eighty one possible search patterns for the first stage and seven for the second stage. All the pattern transformations take place in parallel at each stage.

5. Conclusion

The MSD adder presented here performs addition/subtraction of two MSD numbers in two stages regardless of the number of digits present in the two numbers. The architecture takes advantage of the parallelism offered by the MSD number system, symbolic substitution, and optics.

6. References

1. A . Avizienis, "Signed-digit number representation for parallel arithmetic," IRE Trans. Electronic Computers, EC-10, 389-400 (1961).
2. P. A. Ramamoorthy & S. Antony, "Optical modified signed-digit adder using polarization-coded symbolic substitution," Opt. Eng., vol.26, no.8, 821-825 (1987).

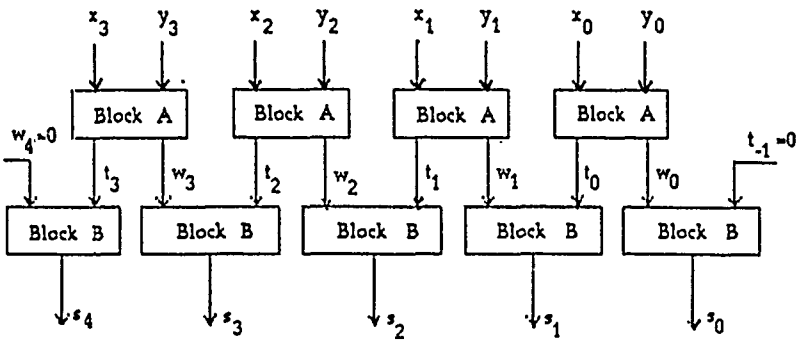


Fig. 1 Two-stage MSD adder/subtracter for two 4-digit numbers

x_i	y_i	t_i	w_i
1	1	1	0
1	0	0	1 or 1 $\bar{1}$
1	$\bar{1}$	0	0
0	1	0	1 or 1 $\bar{1}$
0	0	0	0
0	$\bar{1}$	0	$\bar{1}$ or $\bar{1}$ 1
$\bar{1}$	1	0	0
$\bar{1}$	0	0	$\bar{1}$ or $\bar{1}$ 1
$\bar{1}$	$\bar{1}$	$\bar{1}$	0

Table 1 Transfer and weight digits for two MSD digits

x_i	y_i	x_{i-1} y_{i-1}	t_i	w_i
1	1		1	0
1	0	both +ve	1	$\bar{1}$
0	1	otherwise	0	1
0	0			
1	$\bar{1}$		0	0
$\bar{1}$	1			
0	$\bar{1}$	both +ve	0	$\bar{1}$
$\bar{1}$	0	otherwise	$\bar{1}$	1
$\bar{1}$	$\bar{1}$		$\bar{1}$	0

Table 2 (a) Input/output relationship for block A.

t_{i-1}	w_i	s_i
1	0	1
1	$\bar{1}$	0
0	1	1
0	0	0
0	$\bar{1}$	$\bar{1}$
$\bar{1}$	1	0
$\bar{1}$	0	$\bar{1}$

Table 2 (b) input/output relationship for block B

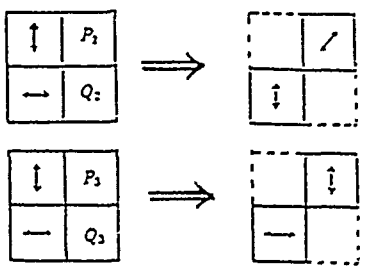


Fig. 2 (a) Input/output relationship for block A when $[x_i\ y_i] = [1\ 0]$

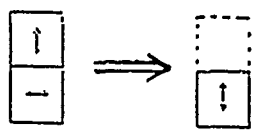


Fig. 2 (b) Input/output relationship for block B when $[t_{i-1}\ w_i] = [1\ 0]$

P_2 and $Q_2 = [1, 0]$
P_3 and $Q_3 = [1, 0, \bar{1}]$ such that both P_3 and Q_3 are not positive.

Using symbolic substitution method on optical matrix multiplication

Kuo-fan Chin, Minxian Wu, Shaomin Zhou
Dept. of Precision Instruments
Tsinghua University
Bejing 100084, China

INTRODUCTION The optical matrix computing often plays a very important role in optical digital processing, many transformation and information processings in optics can be converted into a group of basic matrix multiplications[1]. Now,many approaches to implement such calculations have been provided[2-4]. But to all the existed schemes and systems, the difficulty is the long time sequence and the lower accuracy obtained, as well as a rather complicated system is required. In this paper, we have tried to use optical symbolic substitution method[5] combined with the calculation of outer product of matrices to solve the multiplication of matrices. Preliminary experimental result has been obtained successively.

OPTICAL SYMBOLIC SUBSTITUTION AND OUTER PRODUCT OF MATRICES The method of symbolic substitution is based on pattern recognition of input transfered and the correlation of two 2-dimentional patterns. All these correlations can be substituted into different pattern according to assigned rules, as shown in Fig.1. While the computing with outer product of matrices is also a parallel calculation which fits for being realized by optical method. Assume A and B are two n*n element matrices, the product matrix C of which can be represented sum of n outer product matrices Ci is obtaind by ith colum of matrix A multiplies ith row of matrix B.

THE ENCODING OF A TWO-BIT MATRIX AND THE OPTOELECTRONIC SYMBOLIC SUBSTITUTION SYSTEM In order to perform a multi-bit (element) matrix multiplication, first, to encode each element of a matrix is needed. Although the binary encoding in more flexible, the disadvantage is that only the intermidiate results of those elements of the multiplied matrices can be obtained. Or, in other words, each result is often represented by a mixed binary number. If a pure binary system is wanted to be transformed, then a rather complicated system is required . As an example, using the symbolic substitution rules, the encoded pattern is made as shown in Fig.2 (a multiplication of two 2-bit (element) matrices), where, a and b represent the encoded input and output patterns respectively. As a 2-bit (element) input pattern consists of four possible values, 0,1,2 and 3, sixteen individual channels are needed for recognizing all the different combinations of two input patterns. Evidently, it makes the system too enormous. Therefore, we propose using technique of a multi-window mask to reduce the channels required. As shown in Fig.3, first, move the input pattern A relative to pattern B (the situation for different channel is different), and then recombine it. Put different multi-window decoding masks at the output, then several patterns of different combination can be recognized.

Thereby, sixteen channels are reduced into 6 ones. Of which, the output of the first channel is 0*, *0 (*: 0,1,2,3). They do not influence the results of substitution. Hence, only five channels are really needed.

Assume that the function of coordinate I(x) represents the position of the black spot and dx the distance between two grids, shown on figure 2(a), then, the encoding pattern of 0,1,2,3 may be expressed as follow:

$$I_k(x) = P(x_0 + kdx) \qquad (k=0,1,2,3) \qquad (1)$$

If the input pattern A does not move, only input pattern B moves one grid (dx) leftward, then:

$$I_{kB}(x) = P(x_0 + (k+1)dx) = I_{(k+1)A}(x) \qquad (k=0,1,2) \qquad (2)$$

When $I_{0B}(x)$ coincides with $I_{1A}(x)$, then $I_{1B}(x)$ and $I_{2B}(x)$ coincide with $I_{2A}(x)$ and $I_{3A}(x)$ respectively. This means, these three patterns can be recognized simultaneously. Similarly, this method can be used to discuss the situation of moving two grids (2dx) leftward and one or two grids (2dx) rightward.

In short, if N represents the number of possible value of the input pattern; P represents the number of channels required; and Δ the number of channels reduced. Then:

$$P = 2(N-2) + 1 \qquad (3)$$

$$= N^2 - P = (N-1)^2 + 2 \qquad (N \geqslant 2) \qquad (4)$$

Obviously, when N is large, the number of channels drops fierecly. A hybrid system has been used for implementation the symbolic substitution. Each optical pattern recognition channel is constructed by a Mach-Zehnder interferometer (Fig.4), Input A and B are placed on each arm of the interferometric system respectively, moving mirror M_1 and M_2, different displacements of A and B can produced before combining. The combined pattern are decoded by a NOR logic gates after which output are formed. All the recognized results are received by detectors masked by a multiple window and displayed by LEDs after amplification. The output C can be expressed as follow:

$$C = c_3\ c_2\ c_1\ c_0 \qquad (5)$$

where:

$$c_0 = d(11) + d(13) + d(31) + d(33) \qquad (6)$$

$$c_1 = d(12) + d(21) + d(13) + d(31) + d(23) + d(32) \qquad (7)$$

$$c_2 = d(22) + d(23) + d(32) \qquad (8)$$

$$c_3 = d(33) \qquad (9)$$

$$d(ij) = 0,1 \qquad (i,j = 0,1,2,3)$$

where, d(ij) represents the output of recognized channel of the ijth combination. Such an output is encoded just into a binary form (Fig.2(b)) which is very convienient for further computing. Fig.5 shows the logic combination of the pattern recognition and symbolic substitution of N = 4, P = 5.

MULTIPLICATION OF A TWO 2-BIT MATRICES Take two 2*2 element matrices for as an example, two outer product matrices C_1 and C_2 are used in solving C = A.B. Here, only the computing of C_1 is described. (the same is with C_2). First, expand the first colum of matrix A and the first row of matrix B. Then, we have A and B :

$$A_1 = \begin{pmatrix} a_{11} & a_{11} \\ a_{21} & a_{21} \end{pmatrix} \qquad B_1 = \begin{pmatrix} b_{11} & b_{12} \\ b_{11} & b_{12} \end{pmatrix}$$

For instance, suppose A,B as follow:

$$A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix}$$

Then:

$$A_1 = \begin{pmatrix} 0 & 0 \\ 2 & 2 \end{pmatrix} \qquad B_2 = \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}$$

Encode these two input and make patterns as shown in Fig.6. Fig.7 shows the experimental result of the six pattern recognition channels. The substitution is performed by electronics from the recognized pattern.

CONCLUSION The symbolic substitution of matrix multiplication is a pure digital processing, which is accurate and reliable, and it only P channels of optical NOR and electronic OR are needed. So the calculating speed can increase rapidly. Similarly, it may be extended to the matrix multiplication of more than 2-bits(elements). This work was supported by High-Tech fundation granted to Tsinghua University.

REFERENCE:

(1). J.M.Speiser and H.J.Whitehouse, Proc.1980 Government Microcircuits Applications Conf., Houston, Texas,(1980).
(2). D.Casasent, Proc. IEEE 72,831(1984).
(3). R.A.Athale and W.C.Collins, Appl. Opt., 21,2089(1982).
(4). Hidetoshi Nakano and Kazou Hotate, Appl. Opt., 26,917(1987).
(5). A.Huang, IEEE Tenth International Optical Computing Conference,13(1983).

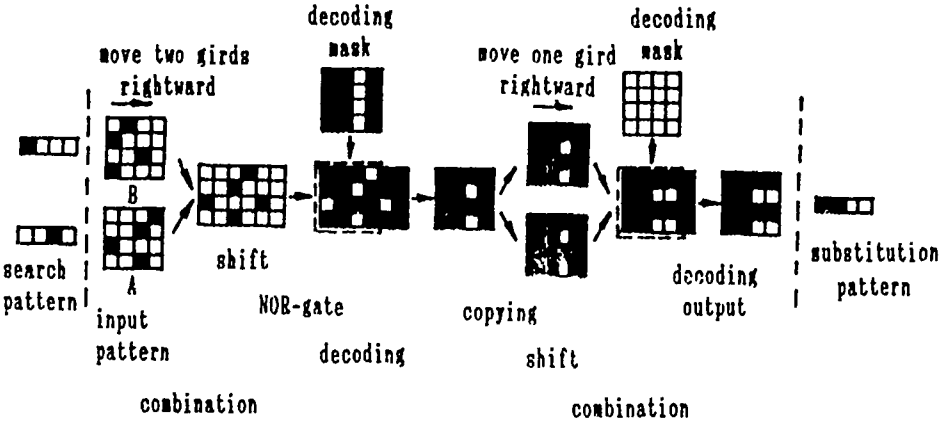


Fig.1 Optical symbolic substitution rule.

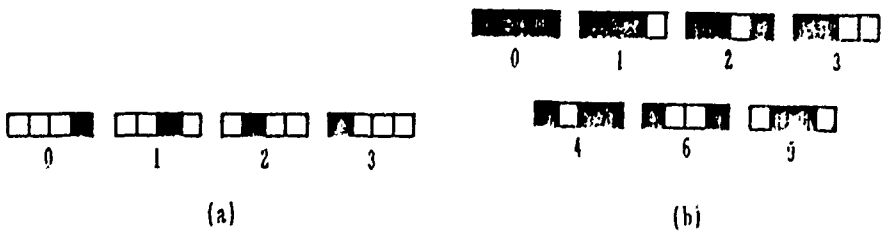


Fig.2 Encoding forms of input pattren and output result.

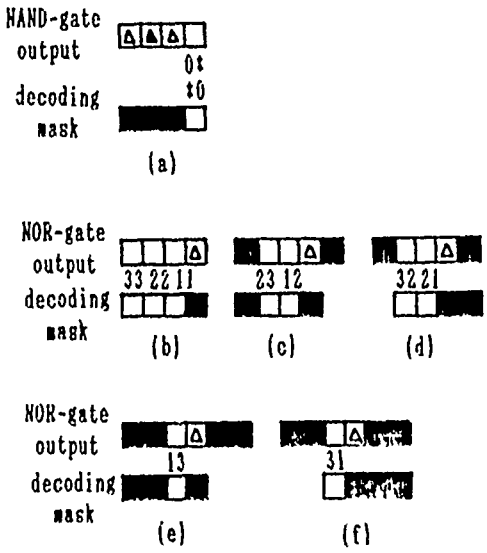


Fig.3 The required relative displacement of A and B input patterns for each channel. (a) and (b) A,B do not move. (c) B moves one gird leftward. (e) B moves two girds rightward. (f) B moves two girds leftward. "△" respresents 0 or 1.

Fig.6 The input mask after encoding.

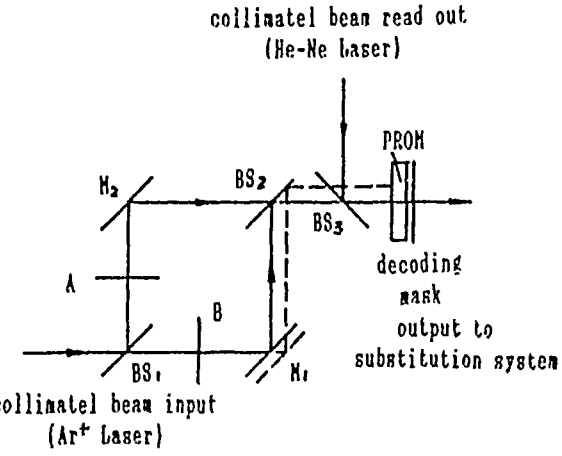


Fig.4 The optical symbolic recognition system (one channel).

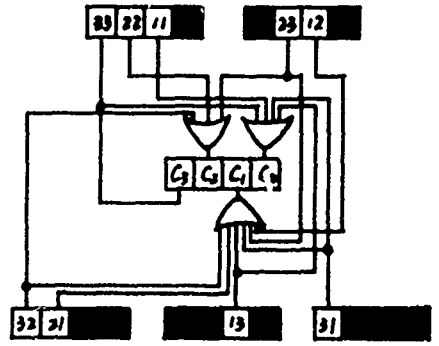


Fig.5 The logic relation of recognition and transformation parts.

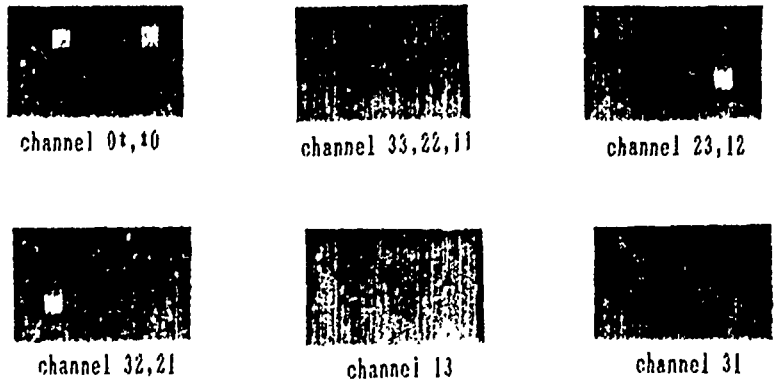


Fig.7 The experimental results of six optical pattern recognition channels.

CORRELATION ALGORITHM AND ARCHITECTURE FOR OPTICAL COMPLEX DISCRECTE FOURIER TRANSFORMATION

Hongxin Huang, Liren Liu and Zhijiang Wang

Shanghai Institute of Optics and Fine Mechanics, Academia Sinica. P.O.Box 8216,Shanghai,China.

1. INTRODUCTION

It is an important subject in optical computing to perform Fourier transform with innate characteristics of optics such as parallel, high-speed, wideband and no crosstalks. A lot of architectures and algorithms have been developed by using coherent and incoherent optical systems. This paper, we report a new correlator architecture and algorithm for performing complex DFT [1].

2. PRINCIPLES

2.1 Discussion of DFT

In mathematics the DFT operation may be regard as a matrix-vector multiplication. That is,

$$G_k=\sum_{l=0}^{N-1} M_{kl}\,g_l \qquad (k,l=0,1,\ldots,N-1) \tag{1}$$

Usually all M_{kl}, g_l, G_k are complexes. The difficulty in calculating Eq.(1) with incoherent optics is the complex could not be expressed directly by intensity. We adopt matrix-code-method of complex[2], that is, a complex is expressed by using a circle matrix with 3x3 nonnegative reals. Thus, substitute G_k, M_{kl}, g_l with encoded matrixs in Eq.1 respectively, we find,

$$\begin{pmatrix} G_K^{(0)} & G_K^{(2)} & G_K^{(1)} \\ G_K^{(1)} & G_K^{(0)} & G_K^{(2)} \\ G_K^{(2)} & G_K^{(1)} & G_K^{(0)} \end{pmatrix} = \sum_{L=0}^{N-1} \begin{pmatrix} M_{KL}^{(0)} & M_{KL}^{(2)} & M_{KL}^{(1)} \\ M_{KL}^{(1)} & M_{KL}^{(0)} & M_{KL}^{(2)} \\ M_{KL}^{(2)} & M_{KL}^{(1)} & M_{KL}^{(0)} \end{pmatrix} \begin{pmatrix} g_L^{(0)} & g_L^{(2)} & g_L^{(1)} \\ g_L^{(1)} & g_L^{(0)} & g_L^{(2)} \\ g_L^{(2)} & g_L^{(1)} & g_L^{(0)} \end{pmatrix} \tag{2}$$

In practice, only the first column elements of the left matrix are wanted to calculate. So Equ.(2) can be simplified into matrix-vector multiplication,

$$F_n = \sum_{m=0}^{3N-1} H_{nm} f_m \qquad (n=0,1,\ldots,3N-1) \tag{3}$$

Or expressed by the correlation form.

$$F'_{nm} = H_{nm} \ast\ast [f_m \delta(n)] \tag{4}$$

Here $\star$ denotes correlation. In this procedure, only addition and multiplication of two nonnegative reals are involved.

2.2 Encodes And Operations Of Nonnegative Real

For calculating Eq.(4), we use two binary masks to code those values contained in the right side of Eq.(4). (1) The maximal value,Max, is represented by a aperture with size dxd; (2) H_{nm} and f_m are respresented by apertures with sizes dx(H_{nm}/Max) and (f_m/Max)xd respectively (Fig.2a,2b); Thus, (3) The overlaping aperture of H_{nm} and f_m can represented product $H_{nm}f_m$ (Fig.2c); (4) The addition can be realized by using a lens to collelct all the relative lights.

3 OPTICAL ARCHITECTURE

A typical setup for calculating Eq.(4) is shown in Fig.2. In this scheme, LED arrary, lenses L_1,L_2 and PDs (photodetectors) are used. LEDs, which act as point-source-arrary and emit light with unit intensity, is imaged by lenses L_1 and L_2. The distance between the two masks is equal to $(d+\varepsilon)f_1/d_0$, where d_0 is the distance between adjacent LEDs and ε is the space between adjacent cells of mask.

4. EXPERIMENTS AND DISCUSSION

Experiments have been made to perform DFT of some input vectors. For simplificity, we suggested a 6-point DFT, the corresponding 18x18 matrix mask is binary. See Fig.3. The theoretical and experimental results with four suggested input functions are shown in Fig.4.

Fig.4 shows experiments conform to theories. The a few differences between them are caused mainly by errors in preparing the masks and the finite aperture of lens used. The design of the system described here depended on geometrical optics, thus the space bandwidth product is limited. The detail discussion of this problem can be found in Ref.3.

REFERENCES:

1. Hongxin Huang,etal. Accepted by Opt.Commun.
2. Hongxin Huang,etal. Submitted.
3. Hongxin Huang, Liren Liu. Submitted to Optica Acta.

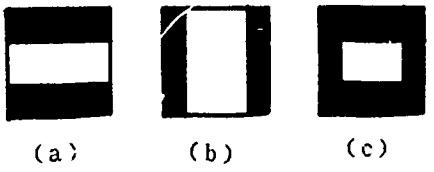


Fig.1 Codes of values H_{nm}, f_m and their product $H_{nm}f_m$. (a)H_{nm}, (b) f_m and (c) $H_{nm}f_m$.

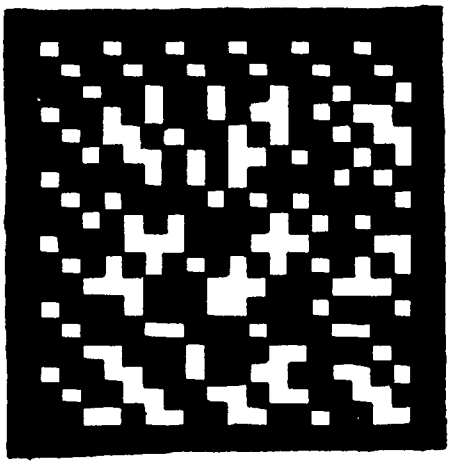


Fig.3 18x18 kernel matrix mask of 6-point DFT.

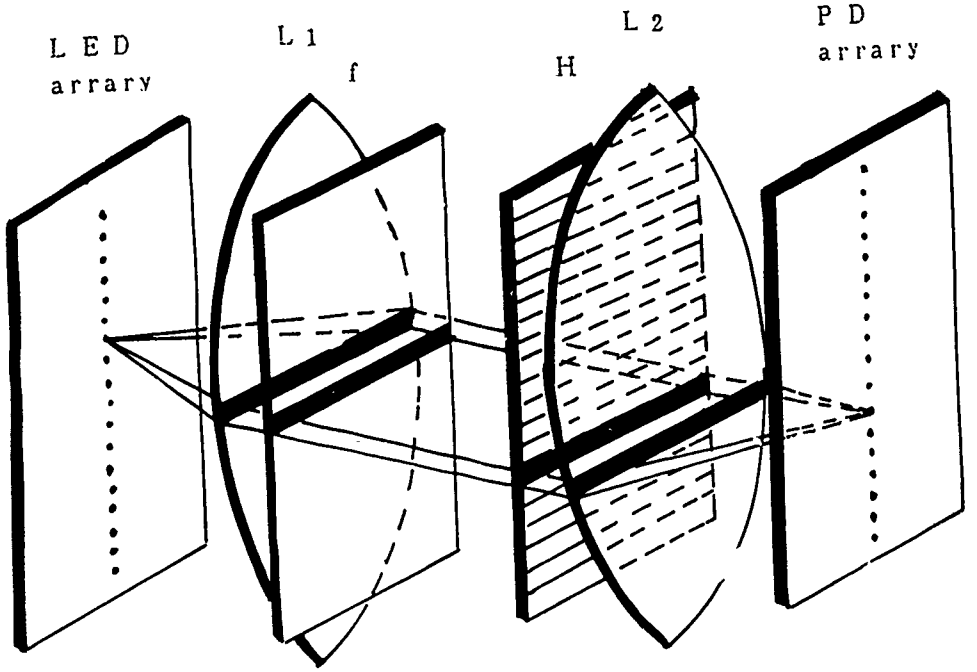


Fig.2 A typical multichannel correlator setup.

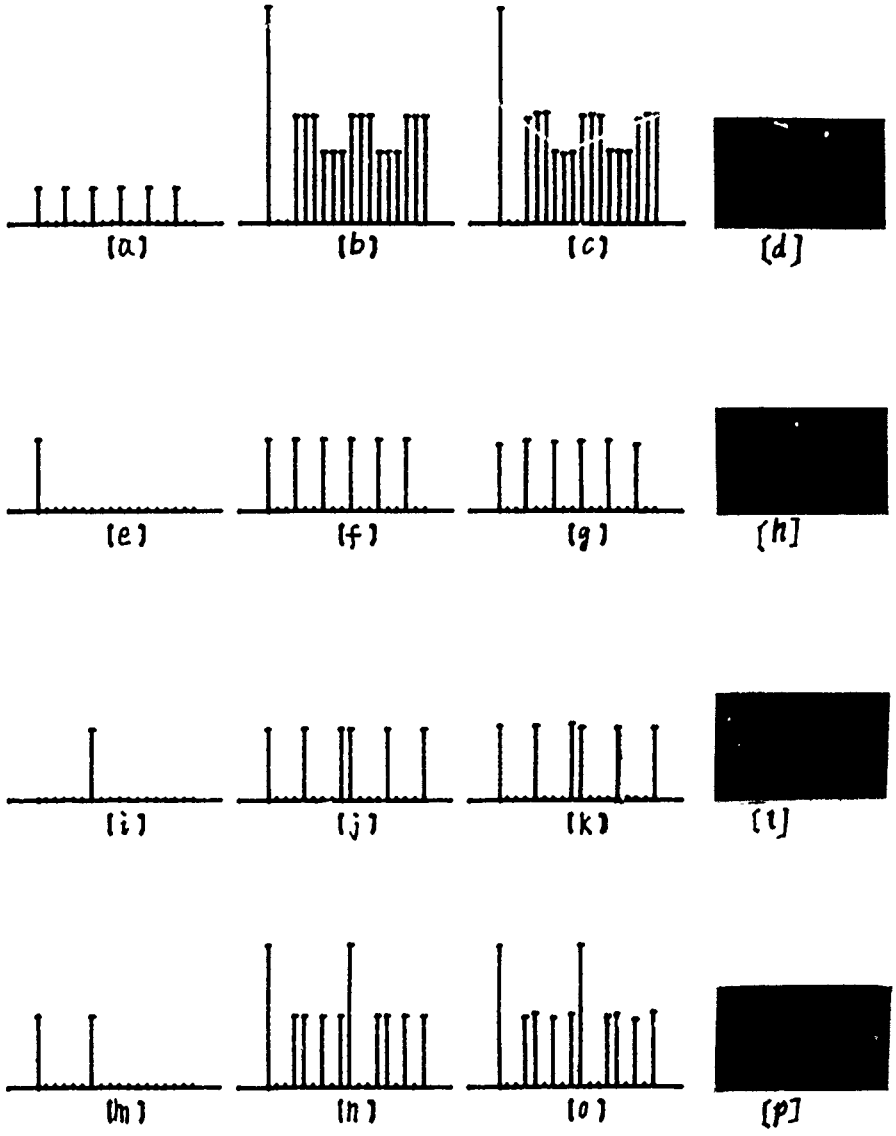


Fig.4 Theoretical and experimental DFT results. (a),(e),(i) and (m) for input distributions; (b), (f),(j) and (n) for theoretical distributions; (c),(g),(k) and (o) for experimental results obtaind with photodetector, and (d),(h),(l) and (p) for experimental photographs.

GaAs WAVEGUIDE MICROLENSES AND LENS ARRAYS WITH APPLICATIONS TO DATA PROCESSING AND COMPUTING*

T. Q. Vu and C. S. Tsai
Department of Electrical Engineering and Institute
for Surface and Interface Science
University of California, Irvine, CA 92717

SUMMARY

Waveguide lenses are among the essential components in construction of integrated optic modules or circuits for data processing and computing. For this purpose, various types of waveguide lens have been fabricated in $LiNbO_3$ substrate. These lens types include Luneburg, geodesic, index refraction via TIPE or two layers construction, chirp grating, and Fresnel. Some of these lens types have been utilized to construct RF spectrum analyzers, correlators,[1,2,3] and computers.[4,5,6] Despite the various successes with such $LiNbO_3$-based modules they have only been developed into hybrid integrated optic (IO) modules due to lack of technology for integration of lasers, detectors, and associated electronic circuits in the same substrate. In contrast, the GaAs-based substrate provides the capability for monolithic integration of all passive and active components. However, the material constraints such as a very high refractive index and high brittleness, and the relatively small reduction in refractive index in $Ga_{1-x}Al_xAs$ for a desirable fractional composition x have prevented any lens type from being fabricated in GaAs waveguide heretofore. We have most recently utilized ion-milling technique to fabricate waveguide lenses of high efficiency and diffraction-limited focal spot size in GaAs. In this paper, design, fabrication, and measured performances of single microlenses and microlens arrays of the analog Fresnel and chirp grating types as well as hybrid combination of the two are presented. IO modules that incorporate such waveguide lenses and acoustooptic and electrooptic Bragg modulators in channel-planar composite waveguides are also being constructed. The measured performances of such modules with applications to data processing and computing will also be reported.

In principle, both the analog Fresnel and chirp grating lenses may be formed using positive- or negative-index change phase zones. These two types of phase

---

* This work was supported in part by the AFOSR and the NSF.

zone require, respectively, deposition of a higher-index cladding material or reduction of the waveguide thickness. However, a high-quality higher-index cladding material for GaAs is yet to be grown. In the meantime, fabrication is greatly simplified by using the negative-index change technique as such negative-index change phase zones may be readily produced by forming grooves in the GaAs waveguide with ion milling. This approach eliminates the need for a higher-index cladding layer and also reduces the number of fabrication steps to a single masking, followed by a single etching of the waveguide with the ion mill. For example, Fig.1 shows the profile of the grooves to be formed by ion milling for an analog Fresnel lens (AFL). The AFL imposes a phase modulation on an incoming light to convert a planar 2-D wavefront into a converging 2-D wavefront. A symmetrical AFL with respect to the Z-axis (SAFL) that has the same phase modulation, but imposes lower resolution requirements can also be easily fabricated.

A limitation of the AFL described above is that the fingers of the lens come to sharp points, which are difficult to reproduce photo-lithographically as the finger period becomes small. The AFL has high efficiency in its center region where the pattern resolution requirements are low, but low efficiency in the smaller period zones where the lens pattern is distorted. A solution to this problem is to replace those smaller-period fingers with chirp gratings. Although the chirp grating period also decreases with the aperture of the lens, the individual fingers are of rectangular shape and therefore easier to reproduce. Thus, an optimum lens design would be to utilize analog Fresnel zones for the center section and chirp grating zones for the two outer sections (see Fig.2). The resulting hybrid lens should be high in throughput efficiency and near diffraction-limited in spot size. These desirable characteristics have been demonstrated in the hybrid lenses fabricated in this study.

In fabrication, the GaAs/$Ga_{1-x}Al_xAs$ waveguide samples with x = 0.07 and 0.15 were first coated with photoresist, exposed with the zone pattern, and milled with argon ion beam to form the lenses desired. The depth of the grooves obtained typically varied from 0.20 to 0.55μm. Single-element AFLs, multiple-element SAFL arrays, chirp grating lenses as well as hybrid lens arrays were readily fabricated. As the first example, three-element SAFL arrays, where each element lens had an aperture of .2 mm, a focal length of 2 mm, with a center to center separation of .25 mm were fabricated in the sample with 7% aluminum concentration. Fig.3 is a photograph of the ion-milled zone pattern for such lens array. The measured throughput efficiency for the element lens at the optical wavelength of 1.064 μm was 30%. The corresponding focal spot width

(defined at 1/e points) was 5.0μm which is slightly larger than the diffraction-limited spot size.

Next, consider a hybrid lens with 5.88mm in focal length, a total aperture of 1.2mm, consisting of an analog Fresnel section of 0.67 mm and two outer chirp grating sections. The GaAs/$Ga_{1-x}Al_xAs$ waveguide used for this lens had an aluminum concentration of 15%. The measured focal spot profile of the lens at the optical wavelength of 1.15μm is shown in Fig.4, indicating that a focal spot width as small as 3.1 μm was obtained. A throughput efficiency as high as 45% was also measured.

In summary, we have successfully fabricated, for the first time, planar waveguide microlenses and microlens arrays in GaAs by using ion milling. High throughput efficiencies and near diffraction-limited focal spot sizes were measured in the analog Fresnel and hybrid lenses of varying aperture and focal length. The fabrication process involved has been shown to be simple and versatile, requiring only patterning and ion milling to produce the phase zones with negative-index changes. We have also demonstrated the feasibility for extending the aperture of an analog Fresnel lens with a chirp grating. Both the ion milling process and hybrid lens combination are applicable to other waveguide substrates including other compound semiconductors. Similar to the titanium-indiffusion proton-exchange (TIPE) waveguide lenses[7] that make realization of $LiNbO_3$-based hybrid multichannel IO device modules possible,[4,5,6,8] such ion-milled microlenses and lens arrays should facilitate realization of GaAs-based monolithic multichannel IO device modules with applications to data processing and computing.

REFERENCES

1. C.S. Tsai, IEEE Trans. Circuits Syst.; CAS-26, 1072(1979)
2. C.M.Verber, R.P.Kenan, and J.R.Busch, Appl. Opt.; 20, 1626(1981).
3. K.Y.Liao, C.C.Lee, and C.S.Tsai, 1982 Topical Meeting on Integrated and Guided-Wave Optics, Pacific Grove, CA, Technical Digest; pp.WA4-1 to -4, Jan.6-8.
4. P. Le, D.Y. Zang, G.D. Xu, and C. S. Tsai, Proc. of 1987 IEEE Ultrasonic Symposium, PP.467-470, IEEE Cat. No. 87CH2492-7.
5. C.S. Tsai, D.Y. Zang, P. Le, Appl. Phys. Lett.; 47, 549(1985).
6. P.Le, D.Y.Zang, and C.S.Tsai, Appl. Opt.; 27, 1780(1988).
7. D.Y.Zang and C.S.Tsai, Appl. Phys. Lett.; 46, 703(1985).
8. C.S.Tsai, D.Y.Zang, and P.Le, 1988 Topical Meeting on Integrated and Guided-Wave Optics, Santa Fe, New Mexico, Technical Digest; pp.200-203, March 28-30.

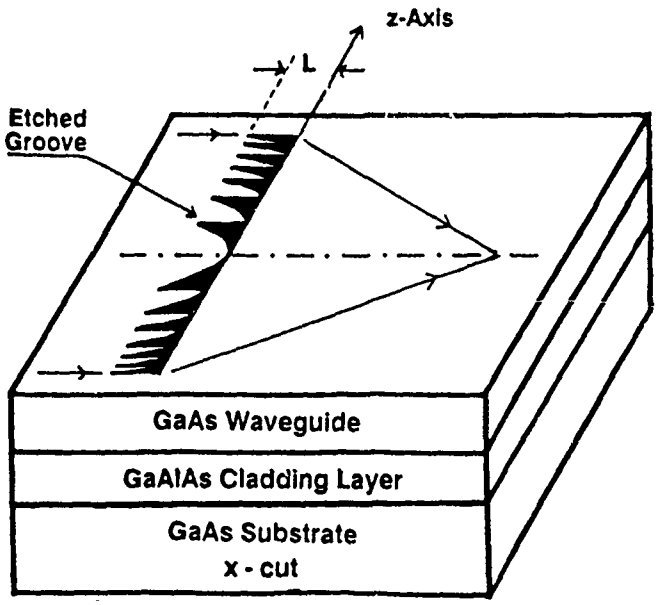


Fig.1
Negative Index-Change Analog Fresnel Lens with Etched Groove

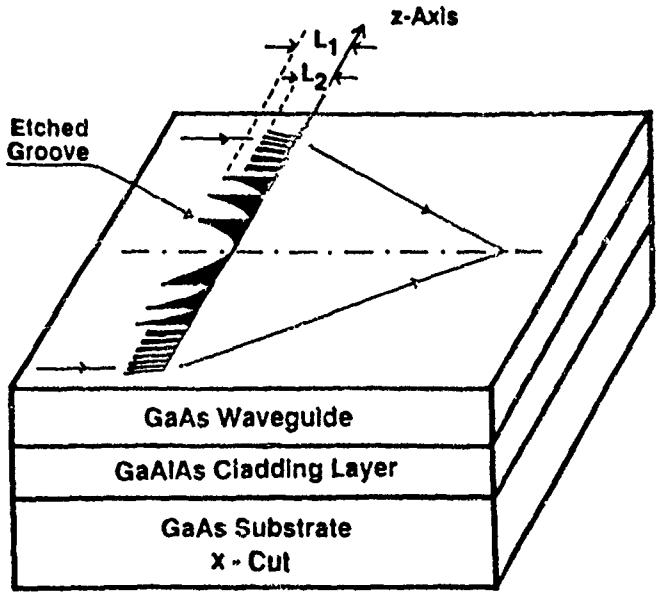


Fig. 2 Negative Index-Change Hybrid Lens with Etched Groove

Fig.3
Photograph of Z[illegible]e Pattern

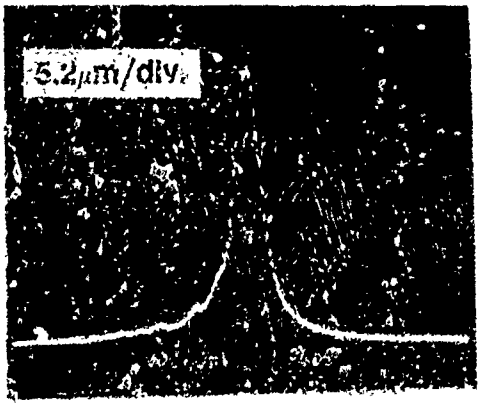


Fig.4
Focal Spot Profile of A Negative Index-Change Hybrid Lens in GaAs Waveguide

Intelligent Optical Processors

Anjan Ghosh

Department of Electrical and Computer Engineering
The University of Iowa
Iowa City, IA 52242.

The accuracy of matrix computations performed on high-speed analog optical associative processors (OAPs) is limited by the noise and spatial errors. There exist two different approaches for alleviating this limitation: (i) postprocessing with a bimodal system [1] and (ii) preprocessing with a preconditioner [2,3]. In this talk we show that these two approaches can be combined to develop an "intelligent" optical processor that can adapt the computational steps depending on the data and produce accurate solutions at a high speed.

The bimodal optical computing approach [1] is based on the idea of using a high speed analog optical processor coupled with a digital post-processor for obtaining the accuracy of digital computing while still retaining the speed and power advantages of analog optics. The analog processor is used to obtain an approximate solution to a problem and the digital processor is used to iteratively improve the accuracy of the final solution. The efficacy of the bimodal approach depend strongly on the condition number of the matrices involved; the smaller the condition number, the faster the bimodal system converges [1].

The preprocessing approach proposed by us is based on the fact that the accuracy of a computation depends on the data, the algorithm, and the hardware used. Thus the data used can be "improved" through a preprocessing step before the final solution such that the results of the computation are less susceptible to noise and spatial errors in the hardware. This preprocessing step requires high speed and is tolerant to computational inaccuracies to a great extent. Thus preprocessing is suitable for analog optical implementation. Apart from the robustness of the final solution, another advantage of such preprocessing is an improvement in the convergence rate of iterative algorithms. Matrix preconditioning is the preprocessing algorithm [2-4] we find most suitable for a large class of engineering problems involving solution of linear systems of equations.

Matrix preconditioning is a transformation for reducing the condition number C(A) of a matrix A [4]. The condition number of A is defined as $C(A) = \|A\| \cdot \|A^{-1}\|$, where $\|\cdot\|$ denotes any matrix norm. The rate of convergence of iterative algorithms depends on C(A); the smaller the condition number, the faster the convergence. Also, computational inaccuracies in problems involving matrix inversion or solution of linear algebraic equations are proportional to the condition number C(A) and the roundoff errors in the processor hardware. Thus preconditioning of matrix data prior to the final solution helps in reducing the time of computation and improving the accuracy of the solution.

Preconditioning of a linear system of equations, $\mathbf{Ax} = \mathbf{b}$, involves the computation of a preconditioning matrix M and the multiplication of both sides of the equations by M to obtain a modified system, $\mathbf{MAx} = \mathbf{Mb}$. The nonsingular matrix M is an approximation of $\mathbf{A}^{-1}$ such that MA has a small condition number $C(\mathbf{MA}) < C(\mathbf{A})$ [4]. This preconditioning process is a robust operation because the computational inaccuracies in matrix M only make it slightly less effective in preconditioning the original data matrix A and do not affect the solution $x = \mathbf{A}^{-1}\mathbf{b}$ obtained by the main algorithm. Thus the effects of errors and noise in optical processors on the result of preconditioning are not important.

There are several methods for calculating a suitable preconditioning matrix M [5]. The method we use is based on calculating M as a splitting matrix for linear gradient methods, such as the successive-overrelaxation algorithm [4]. This method is known as the polynomial preconditioning and is suitable for parallel processors. Moreover, this method allows us to control the degree of preconditioning by selecting a few parameters. The polynomial preconditioning method involves several matrix-matrix or matrix-vector operations, and thus, is a time-consuming process when performed on serial digital computers. These matrix multiplications can be carried out in order of n or n^2 steps on an optical processor [1]. This speed-up and the robustness of the preconditioning algorithm make the OAPs the ideal candidate for a parallel realization of matrix preconditioning.

Recently we developed a new and more efficient polynomial preconditioning algorithm called the split-step polynomial preconditioning (SSPPC) algorithm. In SSPPC the polynomial preconditioning steps are repeated several times in two interconnected loops. We split the basic polynomial preconditioning process with a large number of iterations, p, into several steps with smaller values of p in

each step, each time starting with an improved matrix. In SSPPC, the condition number C(**MA**) decreases at the rate $O(p^m)$ where m is the number of iterations in the outer loop. This rate is much faster than the O(p) rate of a standard polynomial preconditioning method. Our SSPPC algorithm is well-suited for realization on parallel machines, especially OAPs. The SSPPC is also versatile. With a few additional steps this algorithm can be used to calculate the inverse of a matrix or to estimate the condition number [6].

Using the SSPPC algorithm we can build an "intelligent" interconnected optical system for linear algebra processing as shown in Fig. 1. It is a pipelined system that can be adjusted for the given accuracy and amount of computational time to solve a problem of linear algebra. Simple analog OAPs are used throughout the system for their advantages in speed, parallelism, and interconnection capabilities. Since this is an adaptive system with an efficient software, costly high-accuracy hardware is not required.

In the first step this processor estimates the condition of the input problem to determine whether some preprocessing is required before giving the input data to the main processor. This initial decision making process involving an online estimation of the condition number based on the modified SSPPC algorithm can be realized on an OAP.

Then the condition number estimate is compared with a threshold value, C_{th}, representing the type of data that can be processed accurately in a given amount of time. If the estimated C(**A**) is greater than the threshold, the input data are preprocessed using the SSPPC method. Otherwise the input goes directly to the main processor for the final solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. C_{th} can be derived from the convergence characteristics of the algorithm implemented in the main processor. Thus, the second stage of the intelligent processor is an OAP realization of the SSPPC algorithm.

The final stage is a digital computer linked to the main processor in the fashion of a bimodal system. It is a postprocessor for improving the final accuracy. The amount of postprocessing to be used, i.e., the number of iterations in the bimodal loop can be estimated from the required accuracy and time for solution, and the estimated condition number. Research on the design and tuning of such an interconnected adaptive OAP that can adapt to a specified accuracy in a given amount of time is in progress in our laboratory.

This research is supported by the grant EET-8707863 from the NSF.

References:

[1] H. Caulfield et al, Appl. Opt. **25**, 3128 (1986).
[2] A. Ghosh and P. Paparao, Appl. Opt. **26**, 2734 (1987).
[3] A. Ghosh and P. Paparao, JOSA, A **5**, 39 (1988).
[4] G. Golub and C. VanLoan, *Matrix computations*, Johns Hopkins Pr., 1983.
[5] J. Ortega and R. Voigt, SIAM Rev. **27**, 149 (1985).
[6] A. Ghosh and P. Paparao, "High speed matrix preprocessing on analog optical associative processors", Tech. Report, Dept. of ECE, Univ of Iowa, 1988.

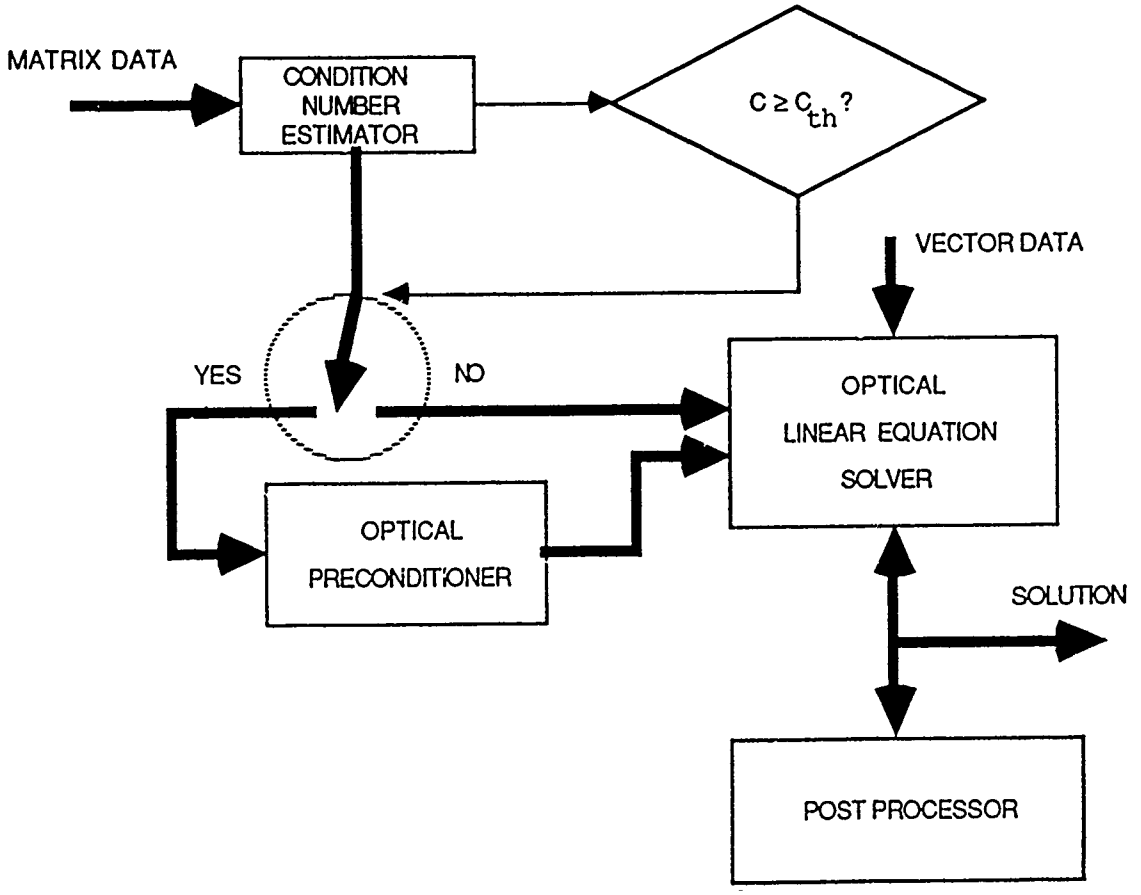


Fig. 1. An "intelligent" analog optical linear algebra processor with pre- and post processing units.

APPLICATIONS OF A POLARIZATION-BASED OPTICAL PROCESSOR

Abraham P. Ittycheriah, John F. Walkup, Thomas F. Krile
Texas Tech University, Dept. of Electrical Engineering
Lubbock, TX 79409

Introduction

Optical logic and transforms such as the Walsh and Haar are examples of optical binary arithmetic. A processor capable of evaluating polynomials[1] has been used to perform optical logic and Walsh and Haar transforms. This summary describes these applications.

Optical Logic

Boolean logic has been extensively used in electronic digital systems and has become the foundation of all digital systems that manipulate information. Figure 1 shows four basic logic functions of the two variables x_1 and x_2. Note that unlike digital binary systems, we have chosen to represent the binary values by -1 and 1 instead of the traditional 0 and 1 values. This choice is made to fully utilize the processor characteristics as described below. A method of encoding these values as images is shown in Fig. 2. This encoding (substitution) may be done efficiently by a simple look-up table. Although only two levels are used at the input, three levels are generated at the output of the processor, and the different logic functions may be realized by discriminating among the levels in different ways. A third state, grey, is the result of an interaction between a 1 and a -1. Other logic functions such as the Exclusive Or and Equivalence may also be implemented.

X_1	X_2	AND	NAND	OR	NOR
-1	-1	-1	1	-1	1
-1	1	-1	1	1	-1
1	-1	-1	1	1	-1
1	1	1	-1	1	-1

Fig. 1. Four Commonly used Boolean Functions.

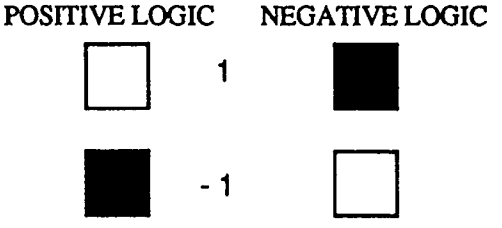


Fig. 2. Encoding Procedure.

Figure 3 shows a simple subsystem block of a logic processor. Note that this block is capable of implementing all logic functions whose variables are not dependent on the results of a future operation. These blocks may be cascaded to perform multi-layer operations. The processor works by ideally rotating the polarization 45 degrees if one of the input images is encoded to represent -1. If vertically polarized light passes through two such superimposed input images, the output light is polarized horizontally. Although in practice the LCTV will only rotate the polarization by about 22 degrees, operations may still be carried out by using a polarizer/analyzer pair to separate the polarization states. Use of polarization encoding will allow switching between positive and negative logic by using a half-wave plate to rotate the polarization by 90 degrees; in other words, this performs contrast reversal of the image. If both the positive and negative images of a logic operation are required at the output,

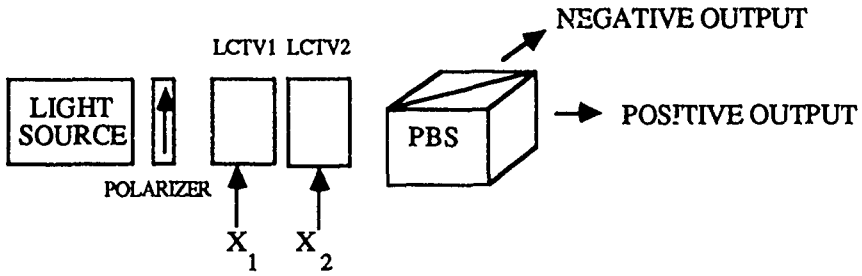


Fig. 3. Building block for digital optical processor (LCTV – liquid crystal television, PBS - polarizing beamsplitter).

then a polarizing beamsplitter may be used and this procedure very efficiently performs the duty of contrast reversal in one beam. If the output of a logic layer is required to be fanned out to several other gates, then one can use spatial replication or use a hologram to achieve one-to-many interconnections. The spatial replication technique, although reducing the space-bandwidth product, will allow for better resolving power and therefore higher accuracy at the present time. However, as the technology advances, active holograms (such as those obtainable using nonlinear media) will be able to provide restoration of power at intermediate steps. Polarization encoding has been discussed also by [2].

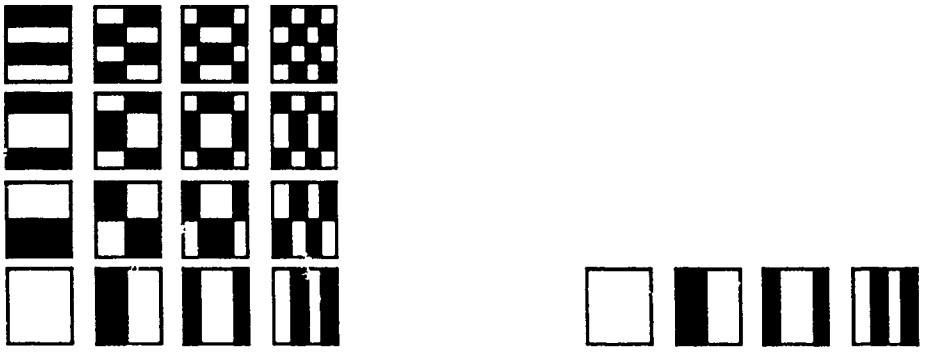
The Walsh and Haar Transforms

Functions of two variables are often encountered. Perhaps the most common of these applications is in image processing. In order to get some information on the frequency components of a two-dimensional function, F(x,y), one could use any of the binary or trinary valued transforms: Walsh, Haar, Hadamard, Her, etc.[3] While some of these transforms are globally symmetric, others are not and these provide local sensitivity to the input function F(x,y). The Walsh transform, shown in Fig. 4, is encoded using the procedure given in Fig. 2. The optical processor described above allows evaluation of these transforms at TV frame rates. In order to perform arithmetic with bipolar elements, a spatial separation into positive and negative number systems is employed. An architecture capable of performing this arithmetic is shown in Fig. 5. An additional feature of the architecture described above is that it is capable of generating the 2-D transform image presented in Fig. 4, by using only the 1-D basis functions. An outer product processor (LCTVs 1 and 2) allow the transform to be generated and then, using the generalized inner product algorithm for matrices, the function F(x,y) is transformed. This reduces the overall memory required to perform the transform. Although the space-bandwidth product of the architecture is limited by the finite size of the LCTV, time-multiplexing may be used to evaluate large transforms. For example, if performing one component at a time, a 64 point transform would require 64x64 TV frames or approximately 2.3 minutes. If performing m components at a time, a 64 point transform would take 2.3/m minutes. The savings in memory is substantial, since it would require approximately 16.7 megabits for storing the entire transform, but only 4096 bits for representing the 1-D basis functions. The image function F(x,y) may be decomposed into the Walsh series coefficients by using

$$a(k,m) = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} F(x,y)\,wal(k,x)\,wal(m,y)\,dx\,dy \qquad (1)$$

and reconstructed by the inverse series given by

$$F(x,y) = \sum_{k=0}^{N} \sum_{m=0}^{N} a(k,m)\ wal(k,x)\,wal(m,y) \qquad (2)$$



(a) (b)

Fig. 4. (a) The Walsh transform where white=1 and black=-1.
(b) The basis functions used to generate the Walsh transform image.

Figure 4 shows (a) the two-dimensional transform as an image and (b) the basis functions used to generate the transform. The transform image is generated by the outer product operation between the basis functions and may be formulated as

$$\text{transform image} = \begin{bmatrix} b_n^T \\ b_{n-1}^T \\ \vdots \\ b_1^T \\ 1 \end{bmatrix} \begin{bmatrix} 1 & b_1 & b_2 \cdots & b_n \end{bmatrix} \tag{3}$$

where b_i = i^{th} basis function and T=transpose operation.

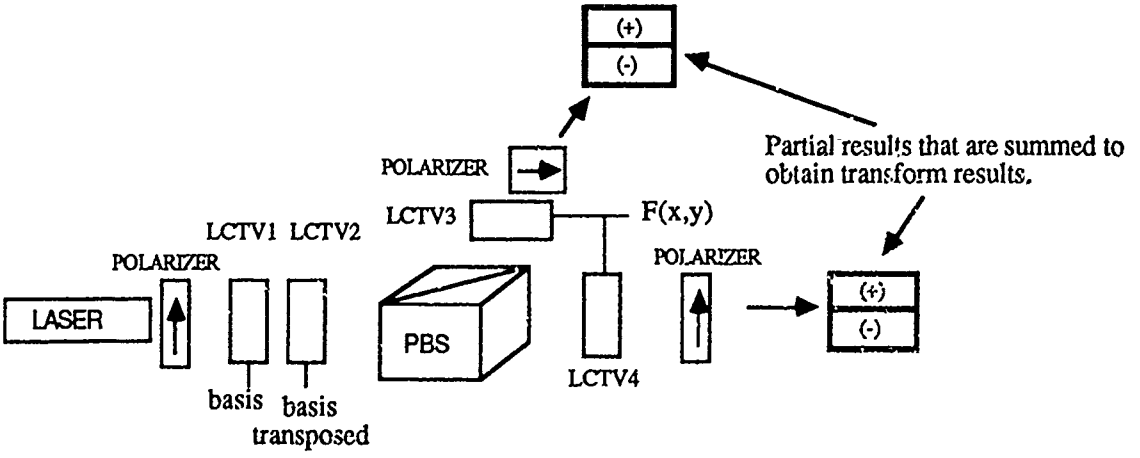


Fig. 5. Optical processor for Walsh and Haar transforms.

Similarly, the Haar transform may be used to decompose the image F(x,y). This transform has 3 elements: {-1 0 1} and the encoding is to use white squares to represent 1, black squares to represent -1, and grey squares to represent 0 as shown in Fig. 6. The zero term remains grey in both the reflected and transmitted images. The superposition of two black squares

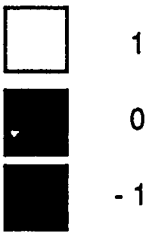


Fig. 6. Encoding procedure to include the zero element.

representing (-1)x(-1) generates a black square and if using an intensity detecting system, this requires contrast reversal of the image. The transform is generated by using the outer product formulation given by (3) and the transmitted image contains in the upper block white squares generated by products of positive numbers and the top block in the reflected image contains white squares to represent the product of two negative numbers. Similar reasoning applies to the bottom block to represent negative results. This transform is locally sensitive to the input, i.e., there is a lack of global symmetry. The transform is shown in Fig 7(a) and the basis function is shown in Fig. 7(b). The other transforms described in [3] may be similarly encoded.

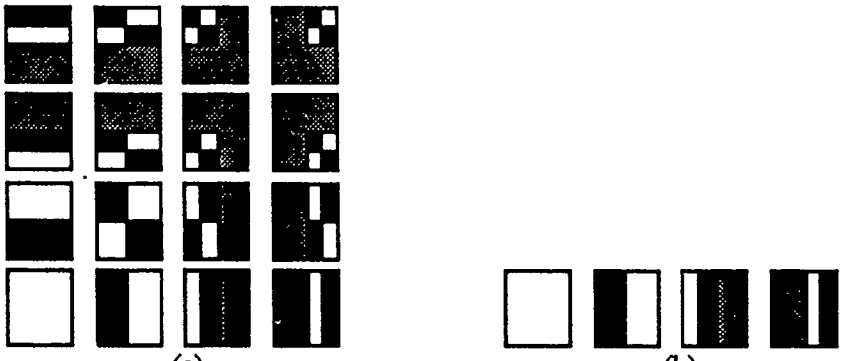


Fig. 7. (a) The Haar transform where white=1, black=-1, and grey=0.
(b) The basis function used to generate the Haar transform image.

Conclusion

An efficient polarization encoding procedure for performing optical logic as well as binary and trinary transforms has been described. Details of the architecture, experimental results and parameters affecting the speed of the processor will be presented in the talk.

Acknowledgements

The authors gratefully acknowledge the support of this research by the Air Force Office of Scientific Research under grant 88-0064.

References

1. A. P. Ittycheriah, J. F. Walkup, T.F. Krile. "An Outer Product Processor Using Polarization Encoding", submitted to Applied Optics, 1988.
2. K. H. Brenner. "New implementation of symbolic substitution logic". Applied Optics, Vol 25., No. 18. 1986.
3. H. Harmuth. Transmission of Information by Orthogonal Functions, New York: Springer-Verlag. 1972.

Guided Wave Vector–Matrix Multiplier

Mark H. Berry and Debra M. Gookin
Naval Ocean Systems Center
San Diego, CA 92152

Optical approaches to vector–matrix multiplication have been proposed for 25 years.[1] However the development of a versatile, accurate implementation has been lagging. The free space architecture used by Goodman, et al.[2] suffered the disadvantage of a fixed photographic matrix mask. Some of these disadvantages were eliminated through the use of acousto-optic[2] or electro-optic[3] devices. The accuracy of free space systems is critically dependent on the quality of large optics. Alignment, stray light, and temperature fluctuations are difficult to control. The use of fiber optics using a lattice structure approach for vector–matrix multiplication was introduced in 1985.[4] These operations were limited to Toeplitz matrices, and mechanical adjustments were necessary to change the matrix elements. Other approaches have included the use of photorefractive media[5] and optical implementations of special arithmetics.[6]

We propose the vector–matrix multiplier shown in figure 1. The laser diode intensities are controlled by a bias current. Each laser diode represents one element, X_i, of the input row vector. Each integrated optical two-by-two directional coupler represents one element, Y_{ji}, of the M by N array of matrix elements. The signal incident on the two-by-two coupler is divided into the two output channels. The ratio of the two outputs depends on the voltage applied to the two-by-two coupler. The light from one output channel of each of all the two-by-two couplers is combined together by a passive directional coupler and the

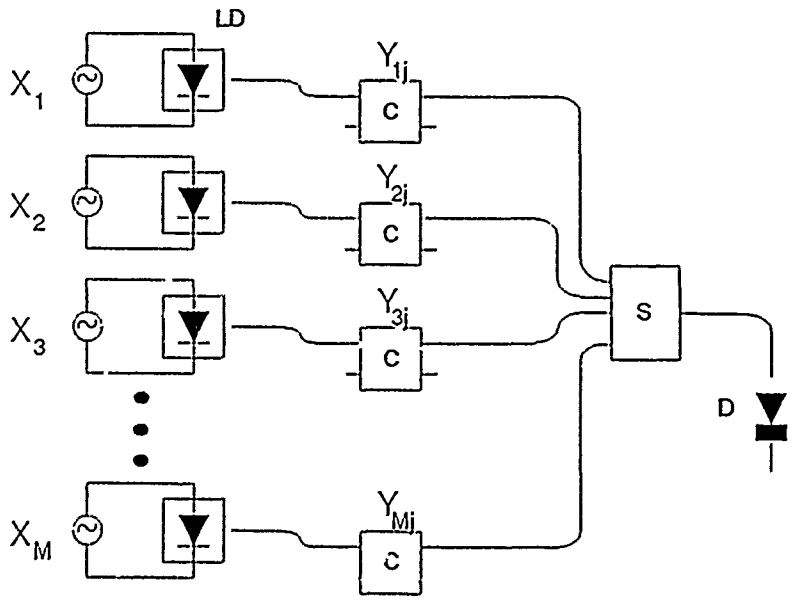


Figure 1. This vector–matrix multiplier performs the operation $Z_j = \Sigma X_i Y_{ij}$. (LD–laser diode, C–integrated optical 2x2 coupler, S–asymmetric star coupler, D–detector)

total light intensity is measured by a detector. In this way, the value of the product of the input vector multiplied by a single column of the matrix is computed. The output at the detector is Z_1 where $Z_1 = X_1Y_{11} + X_2Y_{21} + ...+X_MY_{M1}$.. The values of Z_1 to Z_N are calculated sequentially by changing the voltages on each of the two-by-two couplers to reflect the values of the elements of the next column of the matrix. The intensity of the light from the laser diodes must be kept constant during the calculation of all N of the vector-column products.

Since the matrix elements must be updated, the multiplication can not take place in real time. However the integrated two-by-two directional couplers can be adjusted at rates up to 10 GHz,[7] so the delay can be very small. The time it takes to form the vector-column product depends on the optical path length between the input laser diodes and the detector. For a 20 centimeter path length, we get a one nanosecond delay. If the column element values are updated every nanosecond, then the product of vector times an N column matrix will take:

$$\tau_{vm} = 1 \text{ ns} + N \text{ ns}$$

The duration of the calculation is independent of the number of vector elements.

This system can be extrapolated to a matrix-matrix multiplier by changing the values of the input vector elements following each vector-matrix multiplication. Assuming the laser diodes are modulated by an applied voltage, the time it will take to multiply an LxM matrix by an MxN matrix is:

$$\tau_{mm} = 1 \text{ ns} + LxN \text{ ns}$$

We have described a system for fast computation of vector-matrix products. This system accepts matrix inputs in a convenient computer compatible format. The accuracy of this system is determined by the linearity and stability of the laser diode and the detector. It is not necessary for the integrated optical two-by-two couplers to have a linear response to the applied voltage. As long as the characteristics of each device are well known, they can be taken into account in programming the corresponding voltage inputs. This structure requires only N couplers to form the product of an N element vector with an NxN matrix, whereas the lattice stucture approach[4] required 2N-1.

References

1 L.J. Cutrona, "Recent developments in coherent optical technology," in *Optical and Electro-optical Information Processing*, J.I. Tippett et al., Eds. Cambridge, MA: MIT Press, 1965.

2 J.W. Goodman, P. Kellman, and E.W. Hansen, Appl. Opt. 16, 733(1977).

4 Kenneth P. Jackson, Steven A. Newton, Behzad Moslehi, Moshe Tur, C. Chapin,Cutler, Joseph W. Goodman, and H.J. Shaw, IEEE Trans. MTT 33(3), 193(1985).

5 Pochi Yeh and Arthur E.T. Chiou, Opt. Lett. 12(2), 138(1987).

6 R.P. Bocker, M.E. Lasher, W.J. Miceli, R.H. Patterson, and B.L. Drake, SPIE Proc. 564, Real Time Signal Processing VIII, (1985).

7 Crystal Technology, private communication.

IMAGE PROCESSING USING POLARIZATION-ENCODED OPTICAL SHADOW CASTING II: EDGE DETECTION

A. A. S. Awwal and M. A. Karim
Department of Electrical Engineering, University of Dayton, Dayton, Ohio 45469.

I. INTRODUCTION

The polarization-encoded optical shadow-casting (POSC) system [1] has shown considerable promise in two dimensional digital image processing applications. Recently, POSC systems were utilized to perform gray level image processing [2]. It is very natural for the POSC systems to operate on images which are already in two dimensional form. The generalized POSC algorithm [1] has already been used to accomodate trinary designs [3], programmable logic arrays [4], associative content addressable memories [2], carry-free adder [5], and flip-flops [6].

In our current work, we use a novel spatial mask in the POSC system system to detect image edges in all directions. The POSC algorithm is utilized to determine the encodings for the to-be-processed input images. In the original work of Tanida and Ichioka [7] horizontal derivatives, for example, was accomplished by means of source pattern as operation kernel. However, as the number of the LED sources as well as the spread between the sources increases more and more difraction losses are introduced into the system decreasing the signal to noise ratio. Also no attempts were made in their works to reduce or modify the coding pattern. However, we show that it is actually possible to reduce the input pixel size to as small as one pixel sub-cell using polarization codes or to two pixel subcells using unpolarized codes.

II. POSC SYSTEM AND DESIGN CONSIDERATIONS

The lens-less optical shadow-casting system, as shown in Fig. 1, uses spatially encoded 2-D binary pixel patterns as its inputs. The coded input patterns are placed in perfect contact at the input plane that results in an input overlap pattern. The input overlap pixel is illuminated by a set of input LEDs. The overlapping of projected shadows results in an output overlap pattern at the output plane. A decoding mask, placed at the output plane, is used to spatially filter and detect the output. For the sake of encoding, each input pixel is subdivided into several pixel sub-cells.

In a binary image an edge is detected whenever a 0 --> 1 or a 1 --> 0 transition is encountered. For detecting a horizontal (vertical) line, as shown in Fig. 2(a-b), a vertical (horizontal) difference operator is necessary. Again either a 45° or 135° line, as shown in Fig. 2(c-d), and image corners, as shown in Fig. 2(e-h), can be detected using both horizontal and vertical difference operators. Considering a spatial 2x2 mask, the upper right corner acts like a don't care as it neither affects the horizontal nor the vertical edge calculation but redefines the edge as either a diagonal (having an orientation of either 45°or 135°) or a corner of an image. Consequently, an L-shaped mask (consisting of three pixels) is sufficient to extract edge information.

The edge detection scheme reduces to determining a 3-element window function. The three nontrivial pixels can accomodate up to eight different binary combinations. We form a truth-table whose output corresponds to the fact whether or not the input is a horizontal edge or a vertical edge or both. The truth table is shown in Table I, where A, B, and C are the three pixels ((1,1), (2,1), and (2,2) respectively) of the L-shaped mask.

To make an effective POSC system, it would be necessary to generate the coding patterns for the three inputs, A, B, and C, so that they satisfy the requirement of the truth-table function. Encoded image of B is to be overlapped with that of A (which is shifted downwards by one pixel location) and also with that of C (which is shifted leftwards by one pixel location)

before being introduced to the input overlap plane. The resulting overlap ensures that the three pixels of all L-shaped windows have overlapped with one another. An expanded L-shaped window for example, can be formed by shifting the image A and C by two pixels and then overlapping the three at the input overlap plane.

III. CODE DETERMINATION

From the truth-table one can see that there are only two minterms that need to be generated. The POSC overlap equation [1] correponding to these two minterms could be coupled as follows:

$$\bar{a} \wedge \bar{b} \wedge \bar{c} = V \qquad (1a)$$

$$a \wedge b \wedge c = H \qquad (1b)$$

where "$\wedge$" represents the overlap operation and where, for the sake of simultaneous solution of two equations, both V (vertically polarized) and H (horizontally polarized) codes are used to represent binary 1. Accordingly, $a = b = c = H$ and $\bar{a} = \bar{b} = \bar{c} = V$. If instead only unpolarized codes (transparent, T and opaque, F) were used two subcells ((1,1) and (2,1), for example, respectively representing conditions 1 and 8) would be necessary for each of the pixels to satisfy the coding requirement. Correspondingly, the POSC overlap equations are given by

$$\bar{a}_{11} \wedge \bar{b}_{11} \wedge \bar{c}_{11} = T \qquad (2a)$$

$$a_{21} \wedge b_{21} \wedge c_{21} = T \qquad (2b)$$

whose solutions are $a_{11} = b_{11} = c_{11} = F$ and $a_{21} = b_{21} = c_{21} = T$. The solutions also indicate that the three shifted versions of the image will have the same coding characteristics. For the first design involving polarized codes, only one LED is required. In the second design, however, two LEDs will be necessary.

IV. SIMULATION

For illustration purpose, a binary image, having edges in various orientations, is considered as an input to the POSC edge detection system. The binary image is encoded with unpolarized codes using the codes for A, B and C. The image encoded using the code for B, is overlapped with that for A but shifted one pixel down, and that for C shifted but one pixel to the left. This ensures that all possible L shaped window elements are overlapped with one another. The overlapped images are next illuminated by means of two unpolarized LEDs. The 2x1 LED pattern will generate a 3x1 overlapped shadow for each of the window at the output overlap plane. The center pixel subcell code of the overlapped pixels is then decoded as the output.

Fig. 3(a) shows the input object while its coded pattern is shown in Fig. 3(b). One may copy this coded version three times on a transparency and then overlap the three shifted coded images. The edge will become visible by doing so. The purpose of the LEDs and output mask will be to join the disjoint sets of edges. After processing it through the POSC edge filtering system the output is obtained as shown in Fig. 3(c). Fig. 3(d) shows the output when the same input is processed with an expanded L-shaped window. It may be noted that by doing so the edges

thickened.

V. CONCLUSION

The most recent shadow-casting based edge detector proposed by Tanida and Ichioka [8] corresponds to a multi-step sequential process and requires a total of 36 LEDs as well as huge memory. In comparison, the POSC edge detector developed here is not only a parallel processor but requires far less LEDs, two LEDs for unpolarized codes and one LED for polarized codes. Again since the proposed system operates in parallel, question of having a memory (for storing intermediate results) does not arise. In particular, two differnt L-shaped windows were used to detect the image edges. It should be mentioned here that the spike noise to which this filter may be very sensitive can be removed from the image by means of a low pass filtering or median filtering.

REFERENCES

1. M. A. Karim, A. A. S. Awwal, and A. K. Cherri, "Polarization-Encoded Optical Shadow-Casting Logic Units: Design", Appl. Opt., Vol. 26, 1987, p.2725.
2. A. A. S. Awwal, and M. A. Karim, "Associative polarization-encoded optical shadow-casting: gray-level image encoding for serial and parallel operations", Accepted for publication, Applied Optics(Special Issue on Optical Pattern Recognition and Associative Retrieval Process), 1989.
3. A. A. S. Awwal, M. A. Karim, and A. K. Cherri, "Polarization-Encoded Optical Shadow-Casting Scheme: Design of Multi Output Trinary Combinational Logic Units", Appl. Opt., Vol. 26, 1987, p. 4814.
4. A. A. S. Awwal and M. A. Karim, "Polarization Encoded Optical Shadow-Casting Programmable Logic Array: Simultaneous Generation of Multiple Outputs", Appl. Opt., Vol. 27, 1988, p. 932.
5. A. A. S. Awwal and M. A. Karim, "Polarization-Encoded Optical Shadow-Casting: Design of a Carry Free Adder", accepted for publication, Appl. Opt., Oct 1988.
6. A. A. S. Awwal and M. A. Karim, "Polarization-Encoded Optical Shadow-Casting: Design of a J-K Flip Flop", Appl. Opt., Vol. 27, 1988, p. 3719.
7. J. Tanida and Y. Ichioka,"Optical-Logic-Array Processor Using Shadowgrams II. Optical Parallel Digital Image Processing", J. Opt. Soc. Am. A, Vol. 2, 1985, p. 1237.
8. J. Tanida and Y. Ichioka, "Programming of Optical Array Logic. 1: Image Data Processing", Appl. Opt., Vol. 27, 1988, p. 2926.

Table I. Truth table for the POSC edge detector.

CONDITION	INPUTS			OUTPUT
Number	A	B	C	
1	0	0	0	1
2	0	0	1	0
3	0	1	0	0
4	0	1	1	0
5	1	0	0	0
6	1	0	1	0
7	1	1	0	0
8	1	1	1	1

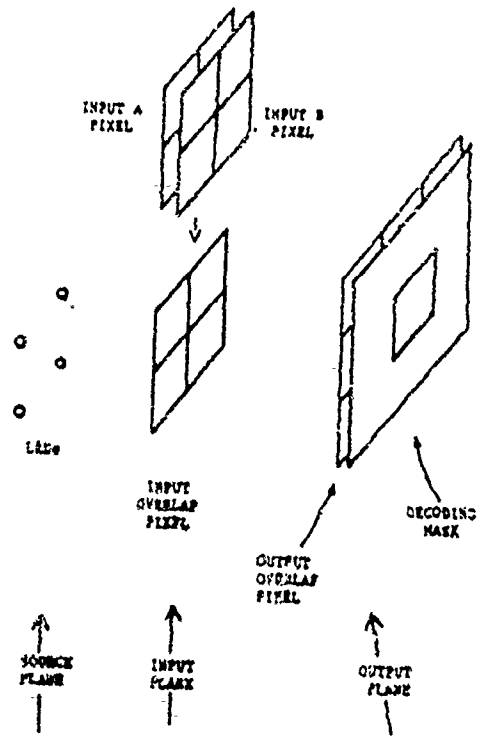


Figure 1. The POSC system.

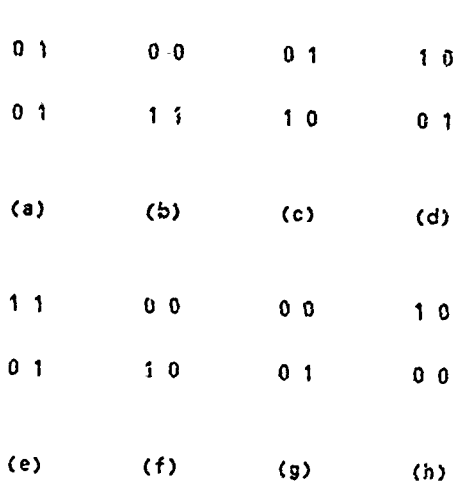


Figure 2. Edge Types: (a) vertical; (b) horizontal; (c-d) diagonal; and (e-h) image corners.

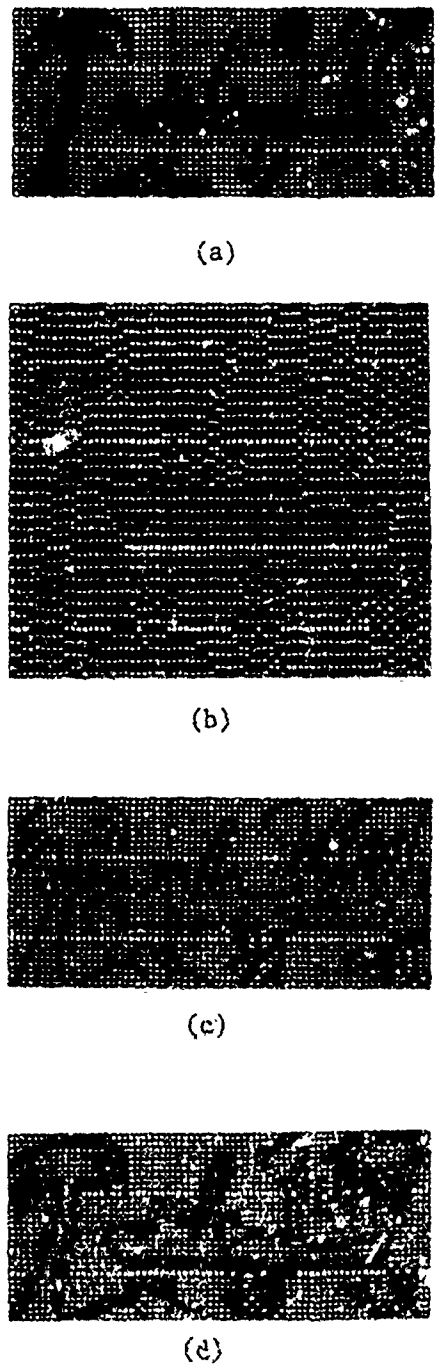


Figure 3. (a) Input image; (b) coded input; (c) output image obtained using a regular L-shaped window; and (d) output image obtained using an expanded L-shaped window.

Polarization-based optical computing using liquid crystals.

Johannes D. Roux
Dep. of Electronics and Computer Engineering,
University of Pretoria
Pretoria
0002
South Africa.

F. Wilhelm Leuschner
Dep. of Electronics and Computer Engineering,
University of Pretoria
Pretoria
0002
South Africa.

1 Polarization encoded logic.

Many architectures that perform digital optical logic have been proposed and built. Most of them use intensity-encoded logic where, for example, the presence of light would indicate logical true and the absence of light logical false. This way of encoding logic information has several disadvantages, e.g. light being irretrievably lost when switching from light ON to light OFF.

An alternative encoding scheme, as proposed by Lohmann et al [1,2,3] uses the inherent binary character of the polarization of light. Should the logic information be encoded using the two orthogonal states of linear polarization, several advantages become apparent. There is symmetry in the energy levels representing the two logic states and the inverse of any signal is produced easily by simply switching the polarization of all incident light by 90°.

Lohmann et al have implemented logic gates using polarization encoded logic. This system can produce any of the sixteen logic functions at will. This system may be classified as multiple data flow single instruction flow. While more than one independent input can be processed simultaneously, the same function is executed on all inputs. Their architecture is not a likely candidate for compact implementation as it makes use of Fourier plane filtering and a long optical axis.

The research group at Colorado University [4,5,6,7] have used ferroelectric liquid crystals to implement multiple data multiple instruction flow polarization based logic gates. They are however unable to implement other than the XOR and COINCIDENCE logic functions.

There is an obvious need for a compact, cascadable architecture using polarization based logic that can do multiple data multiple instruction flow operations, implementing a range of logic functions simultaneously.

2 Proposed architecture.

Recently [8] we have proposed an architecture that fulfills all the requirements spelled out above. Liquid crystal spatial light modulators (LC SLM's) are used to switch the polarization of the light. These devices have the characteristic that the polarization

of the incident light can either be rotated by 90° or passed unaltered. The decision to do either of these can be exercised by the application of the correct voltage across the plates of the LC SLM.

Fig.1 is a schematic diagram of the architecture. Vertical polarization is considered to represent logic true. Two sets of variables, 'A' and 'B' are to be processed. Collimated, vertically polarized light is incident on LC SLM 'A'. LC SLM 'A' encodes the information of the first set of logic variables 'A' in the polarization, switching the polarization of light passing through elements that correspond to a 'false' variable in logic set 'A' to horizontal. The light after LC SLM 'A' now consists of vertically polarized light at the places where elements of logic set 'A' were 'true' and horizontally polarized light at the locations where elements of logic set 'A' were false.

The light now enters a polarizing beamsplitter which deflects all horizontally polarized incident light downwards in the direction of LC SLM 'B1' and all vertically polarized incident light continues undeflected in the direction of LC SLM 'B2'.

The way in which the information contained in logic set 'B' will be encoded in the light depends on the logic function that will be implemented. Say for example that we wish to implement the logic funtion AND. The only elements that should be 'true' at the output are those elements corresponding to the case where both logic set 'A' and logic set 'B' are 'true'.

All the light corresponding to logic set 'A' being 'false' was deflected down to LC SLM 'B1'. All this light we wish to keep horizontally polarized and 'false' at the output. LC SLM 'B1' will therefore pass all the light incident on it unaltered and horizontally polarized. However, in order to implement logic function AND, those elements that correspond to logic set 'A' being 'true' and to logic set 'B' 'false' have to have their polarization switched t horizontal so that they will be 'false' at the output. This switch is done by LC SLM 'B2' through which all the light corresponding to logic set 'A' passes. The light emerging from LC SLM's 'B1' and 'B2' are now superposed on a combiner. In this way only the light corresponding to both logic sets 'A' and 'B' being 'true' has vertical polarization on the combiner. The funtion 'AND' has been implemented. Only information comming from logic set 'A' was used to switch LC SLM 'A' and only information comming from logic set 'B' was used to switch LC SLM's 'B1' and 'B2'.

Using similar reasoning as above, any of the sixteen possible Boolean functions may be implemented. Any combination may also be implemented side-by-side, giving multiple instruction flow operation.

3 Implementation.

In order to test the principle of the architecture, seven-segment twisted nematic liquid crystal displays were stripped of polarizers and backing to be used as LC SLM's. This makes for rather inexpensive laboratory equipment.

Results have been very encouraging. All logic functions have been implemented. Shown in fig.2 are drawings representing the logic sets 'A' and 'B' as well as a photograph of the output after the implementation of the function AND. An analizer-polarizer was used to make the output visible to the human eye, giving a light-true, dark-false output. The drawings of logic sets 'A' and 'B' are also shown light-true, dark-faise in order to make the result obvious.

4 Conclusion.

A compact architecture has been demonstrated that will implement any combination of logic functions in parallel on a matrix of logic variables. The size of the matrix is only limited by the physical constraints of the components. The minimum size of the LC SLM elements will be determined by diffracton effects.

5 References

[1] A.W.Lohmann,"Polarization and optical logic," Applied Optics **25**,1594 (1986).
[2] A.W.Lohmann and J.Weigelt,"Spatial filtering logic based on polarization," Applied Optics **26**,131 (1987).
[3] J.Weigelt,"Binary logic by spatial filtering, Optical Engineering **26**,028 (1987).
[4] L.A.Pagano-Stauffer, K.M.Johnson, N.Clark and M.Handschy, "Optical logic gates using ferroelectric liquid crystals," SPIE **684**,88 (1986).
[5] L.A.Pagano-Stauffer, K.M.Johnson, D.A.Heimmermann and M.A.Handschy, "Digital optical computing with ferroelectric liquid crystals," SPIE **825**,141 (1987).
[6] K.M.Johnson, M.A.Handschy and L.A.Pagano-Stauffer, "Optical computing and image processing with ferroelectric liquid crystals," Optical Engineering, **26**,385 (1987).
[7] M.A.Handschy, K.M.Johnson, W.T.Cathey and L.A.Pagano-Stauffer, "Polarization-based optical parallel logic gate utilizing ferroelectric liquid crystals," Optics Letters **12**,611 (1987).
[8] J.D.Roux and F.W.Leuschner, "Polarization encoded optical logic device," Applied Optics, presented for publication (1988).

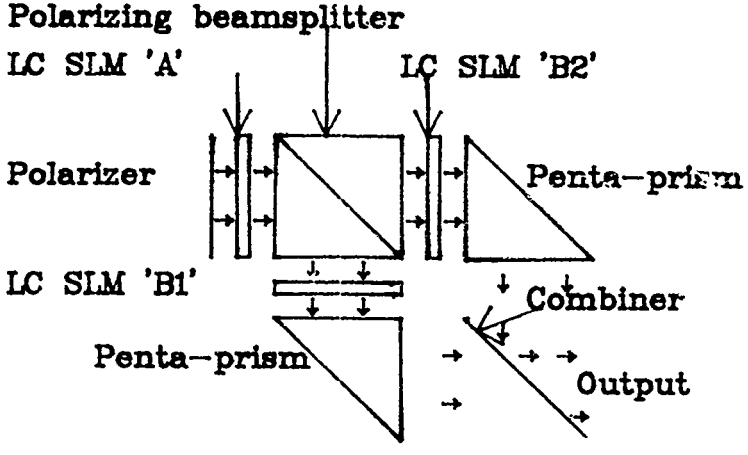


Fig.1. Setup for performing Boolean logic with data encoded by the state of polarization using LC SLM's. Small arrows indicate direction of light propagation

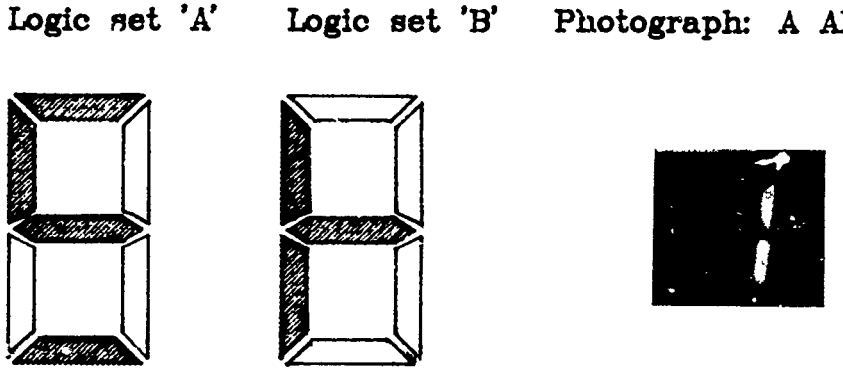


Fig.2. Logic function AND performed on logic sets 'A' and 'B'. Actual logic is encoded in polarization. Light-true, dark-false logic shown here for clarity.

Single Element 2-D Bragg Cells for Optical Computing

Jolanta I. Soos, Douglas C. Leepa, and Ronald G. Rosemeier
Brimrose Corporation of America
7720 Belair Road
Baltimore, Maryland 21236
(301)668-5800

Summary

A single element 2-D Bragg cell can be used in the following subsystems:

1. Vector-vector multiplications.
2. Words equality detection.
3. Half adder optical system.
4. Global interconnection capabilities.

The Brimrose 2-D device can be applied to the multiplication of two binary-digit vectors. The outer product matrix of two vectors $\underline{a}$ and $\underline{b}$ in the binary matrix form can be obtained as

$$C = [a]^T[b] = \begin{bmatrix}1\\0\\1\end{bmatrix}\begin{bmatrix}0 & 1 & 1\end{bmatrix} = \begin{bmatrix}0 & 1 & 1\\0 & 0 & 0\\0 & 1 & 1\end{bmatrix}$$

Multiplication between two single bits is equivalent to a logic operation in which the outer product operation can be carried out with a 2-D device. In other words, the row and column of a 2-D device operating at differenct frequencies can be addressed with two binary vectors $\underline{a}$ and $\underline{b}$, and the outer product C can be directly evaluated by the 2-D matrix.

The paper will present algorithms and equivalent optical circuits for the above-subsystems as well as some fabricated hardware.

Multiplexed Waveguide Hologram for Optical Processing and Computing

Freddie Lin

Physical Optics Corporation
2545 W. 237th St., Suite B
Torrance, CA 90505
(213)530-1416

Introduction

The concept of using holograms in optical processing and computing has been found in a variety of areas such as optical interconnects [1], optical associative memory [2] and optical computing systems [3]. In conventional holography, holographic optical elements (HOEs) are generally located in three-dimensional free space (see Figure 1(a)). This type of 3-D free-space holograms has important limitations. First, alignment problems are critical if the sources and detectors are not in exact 3-D alignment with the holographic elements, performance suffers possibly to the point where the system becomes inoperative. Second, and more importantly, conventional 3-D Bragg holographic elements have low angular and wavelength selectivity (i.e., multiplexing) due to the limited hologram thickness (t ~ 20μm, for dichromated gelatin). Although multifacet holograms [4] are proposed, to improve performance, they suffer from diffraction-limited because of the small facet apertures.

If we place holograms in the same plane as the incident waves (see Figure 1(b)), (i.e., light enters and leaves the hologram in the plane of the 2-D fringe patterns), the thickness of hologram (T) is, thus, not limited by the hologram coating thickness (t), where T>>t. This class of volume hologram can be operated with guided waves in the monolithic integrated optic substrates and the Bragg selectivity of a waveguide hologram can be improved by several orders of magnitude (as compared with conventional 3-D holograms).

Storage Capacity and Bragg Selectivity

Based on Kogelnik's coupled wave theory[5], the required index modulation, Δn, for obtaining very high efficientcy gratings is inversely proportional to hologram thickness. For example, if λ=0.5μm, T = 1mm (i.e., 1-mm waveguide hologram) and Δn_{max}=0.05 (typical for dichromated gelatin), theoretically this waveguide hologram can store up to 10^4 arbitrary gratings.

Furthermore, the angular selectivity of Bragg holograms is given by [4] for unslanted transmission hologram

$$\Delta\theta = \frac{d}{T} \tag{1}$$

where d is the grating spacing and T is the waveguide hologram "thickness."

Assuming T = 1mm and d = 0.5μm, we obtain $\Delta\theta = 0.06$ degree. In other words, there are 1,000 multiplexing channels in a 30°-total scanning angle.

Applications

Since waveguide holograms have superior storage capacity and multiplexing channels, optical processing and computing systems based on waveguide holograms can find many applications. Recofigurable optical interconnection based on turable laser diodes and wavelength-multiplexed waveguide holograms is illustrated in Figure 2. A matrix-vector multiplexer based on waveguide implementation (see Figure 3) has been demonstrated [5]. Addressable Vander Lugh filter using a surface acoustic wave device and a angular-multiplexed waveguide hologram is depicted in Figure 4. Figure 5 shows a preliminary experimental result of one grating waveguide hologram (T~1mm). The diffraction efficiency of 50%, can be greatly improved probably to 80%, and $\Delta\theta$ is less than 0.2°.

References

1. J.W. Goodman, et al, "Optical Interconnections for VLSI," Proc. IEEE, 72, 850 (1984)

2. H.J. White, N.B. Aldridge and I. Lindsay, "Digital and Analogue Holographic Associative Memory," Opt. Engr. 27, 30 (1988)

3. C.C. Guest and T.K. Gaylord, "Parallel Truth-Table Look-Up Digital Holographic Processing," SPIE 232, 110 (1980)

4. H. Kogelnik, "Coupled Wave Theory for Thick Hologram Gratings." Bell Syst. Tech. J. 48, 2909 (1969)

5. D. Psaltis and D. Brady, "Photorefractive Integrated Optical Vector Matrix Multiplier," SPIE 825, 106 (1987)

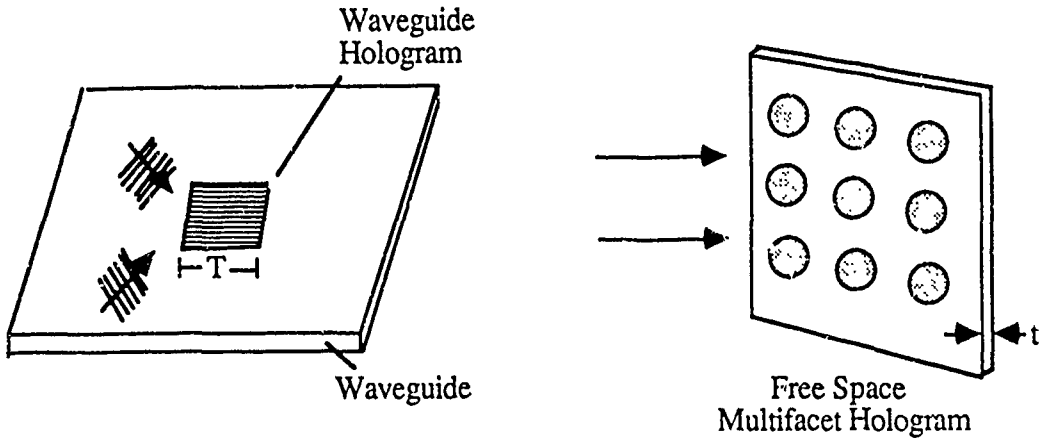


Figure 1 Two types of volume holograms in different configuration.

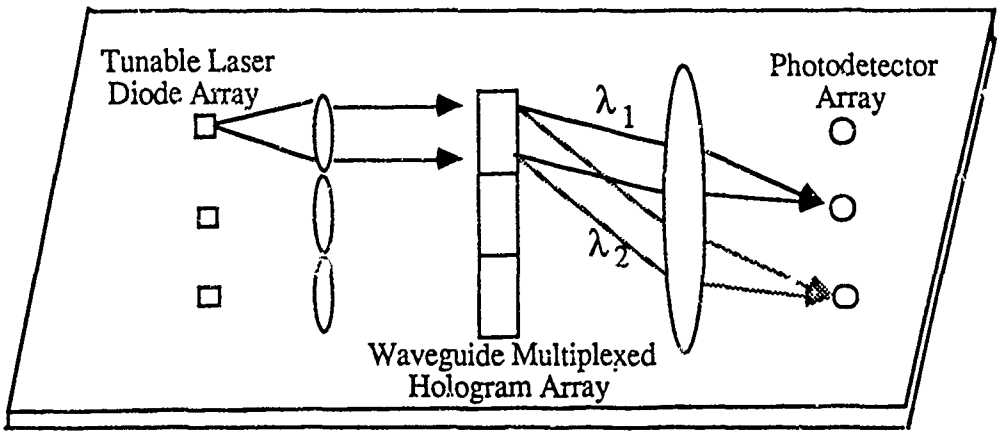


Figure 2 Reconfigurable optical interconnections based on wavelength-multiplexed waveguide holograms.

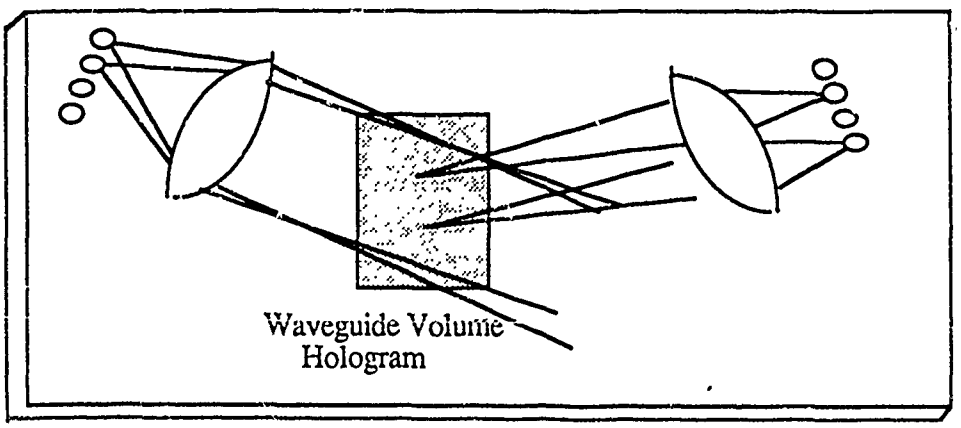


Figure 3 Optical waveguide matrix-vector multiplier (after Ref. 5)

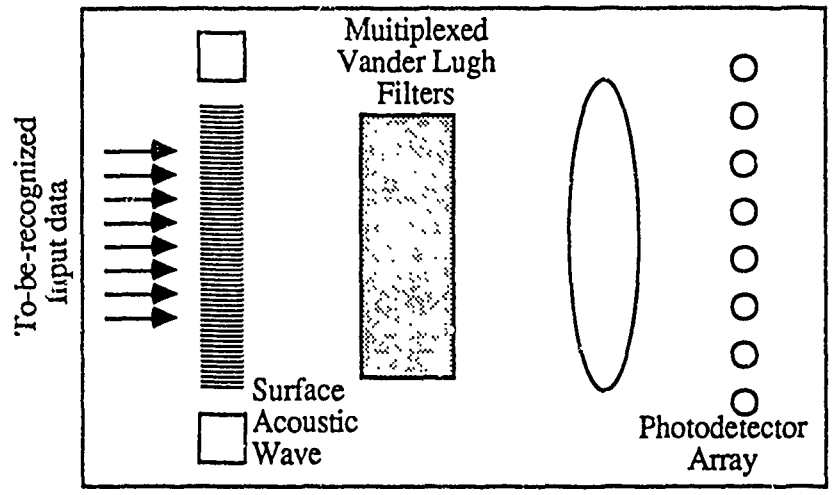


Figure 4 Integrated optical addressable Vander Lugh filters.

Figure 5 The preliminary experimental result of a waveguide hologram based on dichromated

Dynamic Optical Interconnection for Reconfigurable Neural Networks

Bradley D. Clymer and Qiu-Shi Ren
The Ohio State University, Department of Electrical Engineering, 2015 Neil Avenue, Columbus, OH 43210

1 Introduction

Neural networks have become attractive resources for the solution of a large class of problems which are not easily handled by traditional computing tools and strategies[1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17] [18]. The massive sharing of information between nodal processors in neural network configurations allows both redundancy for fault tolerance and a high level of association of data for classification problems such as sorting and character or pattern recognition. It is this massive data distribution that is at once the advantage of neural computing in that it facilitates solutions to difficult information processing problems and a disadvantage in that the enormous number of required interconnection paths can severely limit the size or capability of the network if wire paths are used.

Many optical implementations of neural networks designed to reduce or eliminate the disadvantages of massive global interconnection have been presented in the literature[7] [8] [9] [10] [11]. Because optical signals can pass through each other without interference, a more efficient use of space is possible in optical implementations. Furthermore, optical systems can be arranged in geometries in which many data operations are concurrently handled by shared resources such as lenses, mirrors, holograms and spatial light modulators. Most of the optical processors which have been presented in the literature also take advantage of the ability to store several holograms in a single material and have been developed as associative memory systems. Many also take advantage of materials in which light modulation patterns can be stored and erased or modified to implement learning algorithms in which the weights of interconnections are determined by the properties of the patterns stored in the optical material.

The optical system that we present is significantly different from the optical neural network implementations that have previously been presented in that we use the optics and holography primarily to map fixed interconnection patterns between processors. The details of weighting data passed through specific interconnection paths and redistribution of weights for the network are not a part of this system at this time, and may be handled either electronically or by a separate optical mechanism. The strength of this system is that a large collection of fixed patterns is predetermined to match the requirements of the specific task at hand, and changing between fixed patterns can easily accomplish many tasks that are difficult for both neural and traditional computers. To illustrate how this system might work, consider as an analog the complicated series of events that occur when a human responds to visual stimulations. If an interesting pattern is detected, the eye adjusts to position the object so that it is centered in the field of view. Further visual processing is then applied to heighten awareness of the centered object so that the details of its shape, color and shading can be discerned. In many cases, the object can be mentally rotated so that an internal representation of the image can be matched to a more familiar orientation and classification results. All of these processes are very difficult to implement with either electronic neural networks or traditional computers. Furthermore, an associative memory with a large enough capacity to store the number of records required to recognize different scales and rotations for each object is not feasible. If, however, a single orientation and size for each object can be stored and dynamic interconnection of the network is employed to perform magnification, translation and rotation in predetermined quantum units, an entire system with capabilities which begin to approach human vision is possible.

In the remainder of this paper, we present a description of the general operation and components of the two-level dynamic optical interconnection system.

2 General System Description

The dynamic interconnection system is composed of two levels of holographic reconstruction. An example geometry is shown in Fig. 1. In this architecture, a holographic look-up table contains all of the interconnection patterns required for system operation, much in the same manner as presented by Healey and Smith[19], with the exception that the reconstruction of information stored here is removed by one level from the actual

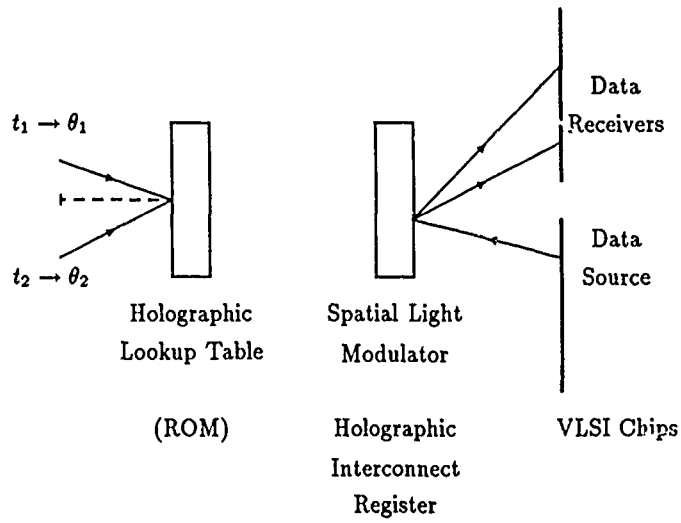


Figure 1: A two-level approach for dynamic holographic interconnection.

interconnection. In this thick multiple hologram, the stored images are fringe patterns which are used when reconstructed to form the interconnection hologram for the second level. Rather than using the data source beam location as the information allowing Bragg reconstruction of the desired pattern directly, a sequential reconstruction of current-state interconnection patterns is used to recallrecallrecall a set of image fringes which have previously been encoded into the look-up table hologram. This method allows an optical data source to be used for any of several interconnection patterns allowable in the system, and is the analog of a permanent memory which contains program information for electronic computers. A given fringe pattern image, representing the current-state holographic mapping information for the system, is reconstructed by addressing the permanent holographic memory at the appropriate Bragg angle and projected onto a temporary storage medium. The temporary storage medium converts the intensity information of the fringe pattern into either a refractive index variation pattern or an absorptive fringe pattern, which is then used for redirection of the data beam to the desired detector locations. The permanent multiple hologram represents the first level of the interconnection system and is roughly equivalent to a read-only memory (ROM) for program storage, and the temporary hologram is the second level and is analogous to an instruction word register which holds an instruction once it has been fetched from the program storage area while a processor decodes the bit pattern to determine which operations are dictated by the instruction.

The temporary holographic mapping register in the system writes the fringe pattern as a hologram that is used for the diffraction of light from data sources to data receivers. Since this fringe pattern is originally one of the images stored in the first-level permanent holographic memory crystal, the temporary storage device must be able to convert an image intensity pattern of fringes into an absorptive or phase modulating pattern. Since the upper limit diffraction efficiency is inherently higher for phase holograms, materials in which the phase of a reading beam can be altered as a function of the intensity pattern of a writing beam are most interesting for this application. The two-level nature of the proposed system allows the reading and writing beams to be of different wavelengths so that the write beam can be chosen to have a wavelength of maximum sensitivity for refractive index changes, and the reading beam can be chosen to have a wavelength for which the material is passive. Because the fringe patterns in this system are formed by reconstructing an image with a single beam rather than by interference of two waves as is typically done to form holograms, the reduced diffraction efficiency and distortions normally caused by writing volume holograms with one wavelength and reconstruction with another are avoided. In addition to the writing and reconstruction capabilities of the material used for the temporary mapping register, the pattern which is stored should be easily erased in the period in which the system control determines the next pattern is to be written.

Since the images reconstructed by the first-level hologram which are to be used as writing inputs for

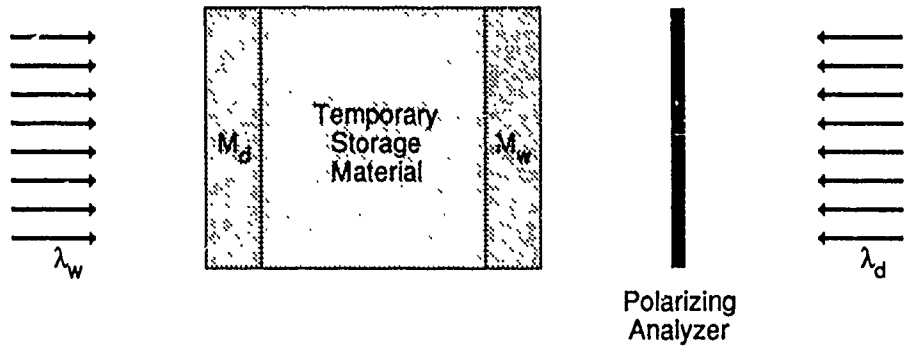


Figure 2: Holographic Interconnect Register

the second-level device are expected to be planar intensity patterns, controlled depth variation of refractive index in the proposed system is not possible. Since reflection holograms require depth variation of refractive index as well as planar variation, a novel geometry is employed for the mapping register to allow reflection of waves with a transmission hologram. The geometry is shown in Fig. 2 will achieve this. In the figure, two dielectric mirrors are deposited on opposite surfaces of the storage medium. The writing information with wavelength λ_w passes the mirror labeled M_d, which reflects only light of wavelength λ_d. This writing information passes through the storage material and is reflected by the mirror labeled M_w, which reflects only light of wavelength λ_w. For the writing process, the mirror M_w allows two passes of the information through the material to improve the writing sensitivity while keeping the thickness of the material small. This can be important, because in many photorefractive crystals the sensitivity and writing/erasure speed can be improved with the application of an external electric field, and a larger field can be generated for a given voltage if the material is thinner. In addition to assisting the writing sensitivity and speed, this mirror blocks unwanted light from reaching the plane of data receivers, and reduces background noise. The data beam passes through mirror M_w, travelling right to left; is diffracted by the refractive index variation present in the material; is reflected by mirror M_d; and passes through the material a second time, increasing the diffraction efficiency shift by effectively increasing the interaction length. The choice of writing wavelength λ_w depends on writing sensitivity of crystal. The data wavelength λ_d depends on available optical sources and detectors on the chips or circuit boards. The data wavelength should not be a sensitive writing wavelength for the material to avoid unwanted grating erasure during data transmission.

Undiffracted light can represent a considerable cross talk term if a great deal of care is not taken in system layout to ensure that none of this light is incident on data photodetectors. This is a problem inherent in many systems which use simple transmission holograms for interconnection, yet might be avoidable by choosing the correct material for the temporary holographic storage medium. In the system we propose, a birefringent storage material can be used to rotate the polarization of the undiffracted beam. A polarizing analyzer can then be used to filter the undiffracted light so that it it does not reach the detectors.

3 Concluding Remarks

The work on this project is just now beginning a phase in which computer simulations are performed on the diffraction of light through various crystal and grating geometries to determine an estimated diffraciton efficiency for the changeable mapping hologram. Simultaneously, laboratory experiments are beginning to study the feasibility of reconstructing fringe pattern images from a long term storage medium for use as the writing pattern of the mapping hologram.

Support for this project has been provided by a University Seed Grant from The Ohio State University and the Department of Electrical Engineering.

References

[1] M. A. Sivilotti, M. R. Emerling, and C. A. Mead, "VLSI architectures for implementation of neural networks," in *Am. Inst. Phys. Conf. Proc. on Neural Networks for Computing*, pp. 408–413, 1986.

[2] H. P. Graf, L. D. Jackel, and W. E. Hubbard, "VLSI implementation of a neural network model," *IEEE Trans. Computers*, vol. 21, no. 3, pp. 41–49, 1988.

[3] J. Hutchinson, C. Koch, J. Luo, and C. Mead, "Computing motion using analog and binary resistive networks," *IEEE Trans. Computers*, vol. 21, no. 3, pp. 41–49, 1988.

[4] L. D. Jackel, H. P. Graf, and R. E. Howard, "Electronic neural network chips," *Appl. Opt.*, vol. 26, no. 23, pp. 5077–5080, 1987.

[5] A. P. Thakoor, A. Moopenn, J. Lambe, and S. K. Khanna, "Electronic hardware implementations of neural networks," *Appl. Opt.*, vol. 26, no. 23, pp. 5085–5092, 1987.

[6] D. Z. Anderson and D. M. Lininger, "Dynamic optical interconnects: Volume holograms as optical two-ports operators," *Appl. Opt.*, vol. 26, no. 23, pp. 5031–6038, 1987.

[7] A. D. Fisher, W. L. Lippincott, and J. L. Lee, "Optical implementations of associative networks with versatile adaptive learning capabilities," *Appl. Opt.*, vol. 26, no. 23, pp. 5039–5055, 1987.

[8] C. C. Guest and R. T. Kolste, "Designs and devices for optical bidirectional associative memories," *Appl. Opt.*, vol. 26, no. 23, pp. 5055–5060, 1987.

[9] K. Wagner and D. Psaltis, "Multilayer optical learning networks," *Appl. Opt.*, vol. 26, no. 23, pp. 5061–5076, 1987.

[10] N. Farhat, "Optoelectronic analogs of self-programming neural nets: Architecture and methodologies for implementing fast stochiatic learning by simulated annealing," *Appl. Opt.*, vol. 26, no. 23, pp. 5093–5103, 1987.

[11] Y. Owechko, "Optoelectronic resonator neural networks learning networks," *Appl. Opt.*, vol. 26, no. 23, pp. 5104–5111, 1987.

[12] T. Kohonen, "Adaptive, associative, and self-organizing functions in neural networks," *Appl. Opt.*, vol. 26, no. 23, pp. 4910–4918, 1987.

[13] G. A. Carpenter and S. Grossberg, "ART 2: Self-organization of stable category recognition codes for analog input patterns," *Appl. Opt.*, vol. 26, no. 23, pp. 4919–4930, 1987.

[14] C. L. Giles and T. Maxwell, "Learning, invariance, and generalization in high-order neural networks," *Appl. Opt.*, vol. 26, no. 23, pp. 4972–4978, 1987.

[15] M. H. Brill, D. W. Bergeron, and W. W. Stoner, "Retinal model for adaptive contrast sensitivity and resolution," *Appl. Opt.*, vol. 26, no. 23, pp. 4993–4998, 1987.

[16] D. Kersten, A. J. O'Toole, M. E. Sereno, D. C. Knill, and J. A. Anderson, "Associative learning of scene parameters from images," *Appl. Opt.*, vol. 26, no. 23, pp. 4999–5006, 1987.

[17] J. Barhen, N. Toomarian, and V. Protopopescu, "Optimization of the computational load of a hypercube supercomputer onboard a mobile robot," *Appl. Opt.*, vol. 26, no. 23, pp. 5007–5014, 1987.

[18] S. Grossberg and D. S. Levine, "Neural dynamics of attentionally modulated Pavlovian conditioning: Blocking, interstimulus interval, and secondary reinforcement," *Appl. Opt.*, vol. 26, no. 23, pp. 5015–5030, 1987.

[19] P. Healey and D. W. Smith, "Holographic associative memory switching system," in *Optical Society of America Annual Meeting, Technical Digest*, pp. 119–120, Rochester, New York, October 1987.

ENTROPY-OPTIMIZED FILTER FOR PATTERN RECOGNITION

Uri Mahlab, Michael Fleisher and Joseph Shamir
Department of Electrical Engineering
Technion, Israel Institute of Technology
Haifa 32000, Israel

One of the procedures in the generation of spatial filters for pattern recognition starts from the correlation plane and defines a desired output pattern such as the Synthetic Discriminant Function (SDF) and the desired filter is generated to yield that output for a given input.
We follow the idea that led to the definition of an SDF but instead of requiring a well defined output function we are interested only in its general behaviour that may be described by introducing the concept of entropy.

Using a simple 4f optical correlator we define a filter function $h(x,y)$ that leads to a complex amplitude distribution

$$C(\lambda_x, \lambda_y) = \int_{-\infty}^{\infty}\int f(x,y)\, h^*(x+\lambda_x, y+\lambda_y)\, dx\, dy \qquad (1)$$

over the output plane for an input function $f(x,y)$.

Assuming a set of input patterns $\{f_n(x,y)\}$ we define our goal to be the detection of the presence of patterns out of a subset $\{f_n^D(x,y)\}$ and the reject of all other patterns, denoted by $\{f_n^R(x,y)\}$. A reasonable criterion for detection is the appearance of a strong and narrow peak as contrasted with a uniform distribution for a pattern to be rejected. To quantify this criterion we normalize the energy distribution over the output plane and define a normalized distribution function (DF) by the relation

$$\phi(\lambda_x, \lambda_y) = \frac{|C(\lambda_x, \lambda_y)|^2}{\int_{-\infty}^{\infty}\int |C(\lambda_x, \lambda_y)|^2\, d\lambda_x\, d\lambda_y} \qquad (2)$$

Our detection criterion states that for a rejected pattern we would like to obtain a uniform DF over the whole output plane, while for a desired pattern the result should be a strong and narrow peak.

For a given input pattern the entropy function is defined by

$$S = -\int\int \phi(\lambda_x, \lambda_y) \log \phi(\lambda_x, \lambda_y)\, d\lambda_x\, d\lambda_y \tag{3}$$

that should be maximized for the rejected patterns and minimized for the desired patterns.

Defining a cost function M by the relation

$$M = \sum_{\{f_n^D\}} S_n - \sum_{\{f_n^R\}} S_n \tag{4}$$

we search for

$$M_{\min} = M[h_{\mathrm{EOF}}(x, y)] \tag{5}$$

Several simulation experiments were performed to investigate the performance of the novel entropy optimized filters with a sample shown in the figures.

An EOF was designed to detect the letter P and reject the letter F in the training set of Fig. 1. Fig. 2 illustrates the amplitude distribution of the EOF and Fig. 3 shows the distribution over the output correlation plane. Fig. 4 illustrates a training set of four patterns, and an EOF was designed to detect G and P while rejecting O and F. Excellent detection levels were obtained as shown in the output plane distribution of Fig. 5.

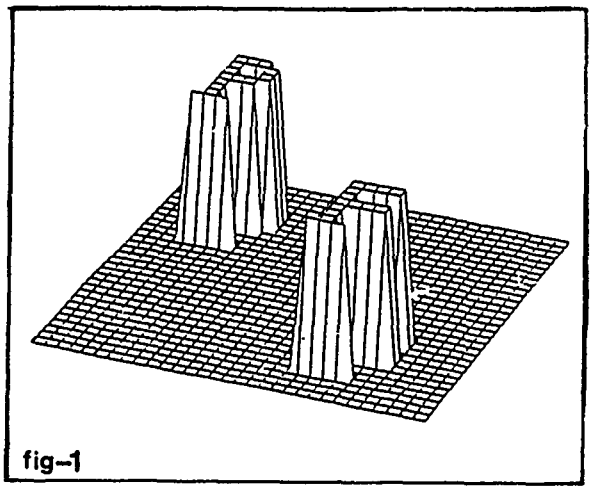


fig-1

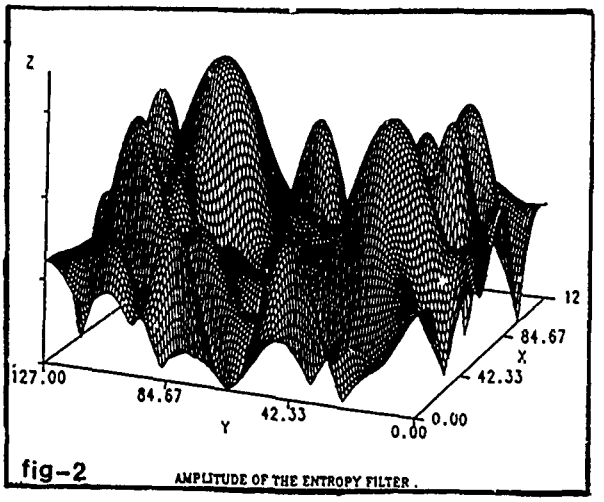


fig-2 AMPLITUDE OF THE ENTROPY FILTER.

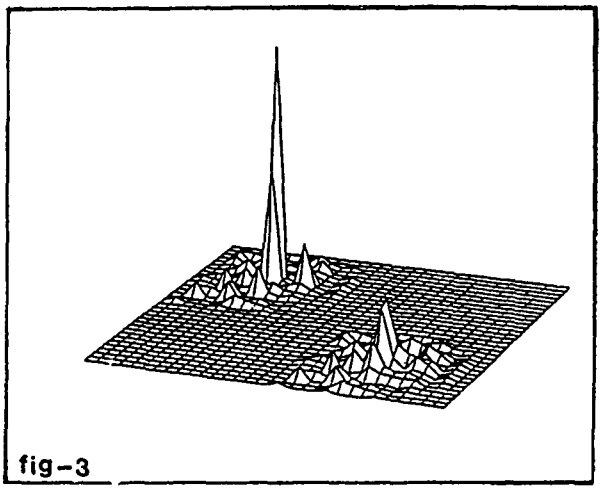


fig-3

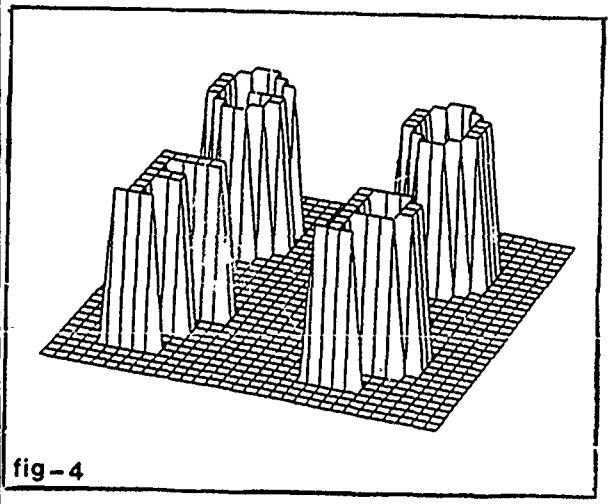


fig-4

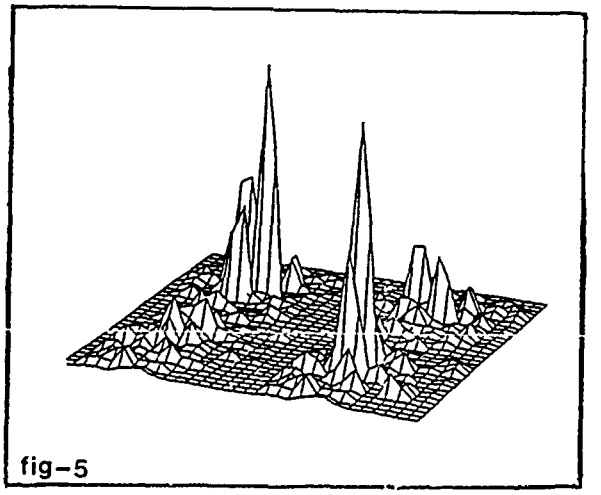


fig-5

SYNCHRONOUS DISCRETE NEURAL NETWORKS FOR MINIMIZATION

Hyuk Lee
Department of Electrical Engineering
Polytechnic Institute of New York
333 Jay Street
Brooklyn, New York

ABSTRACT

A general neural minimization algorithm which can be applied to arbitrary types of polynomial energy functions is presented. The algorithm can be operated in a synchronous as well as asynchronous way. The synchronous algorithm can be implemented by highly parallel optical systems.

Highly parallel neural computing algorithms have been investigated extensively. The Hopfield model has been successfully applied for solving combinatorial minimization problems [1,2]. However, the energy function in the Hopfield is restricted to a symmetric quadratic form having all the diagonal elements zero. Higher-order Hopfield model [3,4] has also been considered. In this case, the energy function is a polynomial of the state variables, and it is assumed to have special symmetry properties. Furthermore, the updating rules of such algorithms are based on the asynchronous operation. At each step, a variable is selected randomly and minimization is carried out by updating only the selected state variable and leaving all the other variables unchanged. Therefore, the algorithm can be operated in a totally asynchronous way but it can not be operated in a synchronous way. Optical implementations of the asynchronous algorithms have been studied [5,6]. However, synchronous neural algorithms are necessary to exploit the full parallelism of optics. In this paper, a general neural algorithm applicable to arbitrary types of polynomial energy functions is developed. It can be operated in a synchronous as well as asynchronous way.

The energy function is assumed to be an arbitrary type of polynomial function of the state variables. Real binary variables having values -1 and 1 are considered as state variables. The energy function can be described as

$$E = E(\{B_1,B_2,...,B_N\}), \tag{1}$$

where B's are the state binary variables and the total number of the state variables, i.e., the neurons is N. Partially synchronous minimization is considered for the most general case. Synchronous or asynchronous minimization algorithms are specific examples of the general case. Assume that at each step, M state variables are selected randomly and minimization is carried out by updating the M selected state variables simultaneously and leaving all the other state variables unchanged. M can be any integer from 1 to N, and the minimization algorithm becomes synchronous or asynchronous if M is equal to 1 or N. A set P is defined to consist of the indices of the selected state variables. Another set P' = {1,2,...,N} - P is defined to represent the indices of the state variables which are not selected. The updated state variables B'_i and the change of state variables ΔB_i satisfy the relation

$$B'_i = B_i + \Delta B_i, \tag{2}$$

where $i \varepsilon P$. The updated state variables are also binary variables having values -1 and 1. Therefore, the possible values of ΔB_i are -2, 0, and 2.

The incremental energy change ΔE due to the updated state variables given by Eq. (2) is considered to develop an algorithm which minimizes the energy function described by Eq. (1). ΔE is defined as

$$\Delta E = E(\{B'_i,B_j\}) - E(\{B_i,B_j\}), \tag{3}$$

where $i \varepsilon P$ and $j \varepsilon P'$, and utilizing Eq. (2), it becomes

$$\Delta E = E(\{B_i+\Delta B_i,B_j\}) - E(\{B_i,B_j\}), \tag{4}$$

The first term in the right hand side of Eq. (4) can be expanded as a Taylor's series in several variables because E is a polynomial of the state variables. Therefore, the incremental energy change ΔE becomes [7]

$$\Delta E = \sum_{m=1} \sum_{i1 \varepsilon P} \cdots \sum_{im \varepsilon P} (1/m!)\, [\Delta B_{i1} \cdots \Delta B_{im}]\, D[B_{i1} \cdots B_{im}]E, \tag{5}$$

where $D[B_{i1} \cdots B_{im}]E$ are the partial derivatives with respect to the state variables $B_{i1}, \ldots, B_{im}$ at $\Delta B_i = 0$ for all $i \varepsilon P$. The total number of terms in the summation of Eq. (5) is finite because E is a polynomial.

It can be shown that the products of changes in Eq. (5) satisfy the following relation

$$A \Delta B_{i1} \cdots \Delta B_{im} \leq |A| I(m) \sum_{k=1}^{m} B_{ik} \Delta B_{ik}, \tag{6}$$

where A is an arbitrary function of the state variables, and

$$I(m) = -(1/m)[\{ \sum_{l=1}^{m-1} 2^{(2^l - l - 1)} \} + 2^{(2^{m-1} - m)}], \tag{7}$$

If Eqs. (6) and (7) are used in Eq. (5), the incremental energy change is given by

$$\Delta E \leq - \sum_{i \varepsilon P} G_i B_i, \tag{8}$$

where

$$G_i = - \{ \sum_{m=1} \sum_{i1 \varepsilon P} \cdots \sum_{im \varepsilon P} (1/m!)\, I(m+1)\, |D[B_i B_{i1} \cdots B_{im}]E| + D[B_i]E \}. \tag{9}$$

From Eq. (8), it is clear that if

$$G_i B_i \geq 0 \text{ for all } i \varepsilon P, \tag{10}$$

the incremental energy change is always negative or zero. Therefore, the conditions described by Eq. (10) update the state variables in such a way

conditions described by Eq. (10) update the state variables in such a way that it minimizes the energy function in the limit of iteration. Equation (10) can be interpreted and transformed as follows. For positive G_i, Eq. (10) is satisfied if ΔB_i is positive or zero. If B_i is equal to 1, B'_i should be equal to 1 because $\Delta B_i=0$ is the only solution for this case. However, if B_i is equal to -1, B'_i should be 1 because this makes the incremental energy change more negative than using the condition $\Delta B_i=0$. Therefore, if G_i is positive, the updated value becomes 1 which is the same as the sign of the value G_i. If G_i is negative, the above argument can be applied to show that the updated value for B'_i becomes -1. This is the same as the sign of G_i. If G_i is zero, B'_i can have any value. Summarizing the above result, the updating rule which satisfies Eq. (10) can be written as

$$B'_i = T(G_i) \text{ for all } i \varepsilon P, \tag{11}$$

where T is a unit step function defined by $T(x)=1$ if $x \geq 0$ and $T(x)=-1$ if $x<0$. Equation (11) is the general neural algorithm which minimizes energy functions consisting of arbitrary types of polynomials in a partially synchronous way.

This work was supported by the National Science Foundation Grant No. EET-8810288 and the Center for Advanced Technology in Telecommunications.

REFERENCES

[1]. J.J. Hopfield, Proc. Natl. Acad. Sci. USA. **79**, 2554, 1982.
[2]. M. Takeda, and J.W. Goodman, Appl. Opt. **25**, 3033, 1986.
[3]. D. Psaltis, and C. Park, APS Conf. Proc. **151**, 370, 1986.
[4]. T. Maxwell, C.L. Giles, Y.C. Lee, and H.H. Chen, APS Conf. Proc. **151**, 299, 1986.
[5]. D. Psaltis, and N. Farhat, Opt. Lett. **10**, 98, 1985.
[6]. J.S. Jang, S.Y. Shin, and S.Y. Lee, Opt. Lett. **13**, 693, 1988.
[7]. L.V. Toralballa, *Calculus* (Academic Press, New York, 1967).

Optical Error Correction by Adaptive Thresholding

David Kagan[1] and Harry Friedmann[2]

[1]California Institue of Technology

Department of Electrical Engineering

Pasadena, California 91125

[2]Bar-Ilan University

Department of Chemistry

Ramat-Gan, 52100, Israel

Summary

During the last few years we have suggested[1-2] the use of a four-level system irradiated by a bichromatic beam to obtain a tunable optical bistable device. By tunable we mean that the switching threshold can be altered without changing the hardware characteristics, the operating temperature, or the frequencies of the inputs to the device. The basic principle of our suggestion is that by irradiating the non-linear optical material by two beams of appropriately chosen frequencies in an optical cavity, one can act as a signal and the other as tuning control beam. The advantage of the four level system lay in the possibility of using low power control beams which could tune the characteristic optically bistable output curves.

An interesting application of this tuning capability, could be as a self-correcting neuron. The parameters of the cavity can be chosen such that the output curve of the signal

is not bistable but rather exhibits soft thresholding. By using a control beam the soft thresholding can be shifted to the right as shown in Figure 1, where the dashed part of the curves indicate the use of a limiter. By detecting the difference of the output from a desired target an error signal can be formed. The control beam is then either increased or decreased in proportion to this error signal and reinput to the device. Since the new control beam shifts the threshold, the output signal will move towards the desired target value. The scheme is illustrated in Figure 2. Parametric representation of the output curves can be used to prove the convergence to the desired target, and the neuron has the desirable property[3] of fast convergence in the high gain region (undecided) and slow convergence in the low gain regions, where a decision has been reached.

It would be interesting to set up a a variable threshold network interconnecting many such devices. It has been suggested[4] that the variable threshold decision element is a realistic model of the brain neurons, where the overall potential of small sections of the brain may be changed selectively (either chemically or through propagation of "brain waves") as a method of information transfer. It should be noted that the tuning can be achieved as quickly as the signal switching so that a network could be speedily reconfigured using control beams. More research is required to develop useful learning algorithms for a variable threshold network.

References

1. D. Kagan and H. Friedmann, "Adaptive optical logic and switching devices," Proc. Soc. Photo-Opt. Instrum. Eng., Vol. 835, pp. 359-361, (1987).

2. D. Kagan and H. Friedmann, "Tunable optical bistabilty of bichromatic fields interacting

with two band systems," submitted for publication IEEE J. of QE.

3. K. Wagner and D. Psaltis, " Multilayer optical learning networks," Applied Optics, Vol-26, pp. 5061-5076, (1987).

4. W. Meisel, "Variable-threshold threshold elements," IEEE Trans. on Computers, Vol C-17, pp. 656-667, (1968).

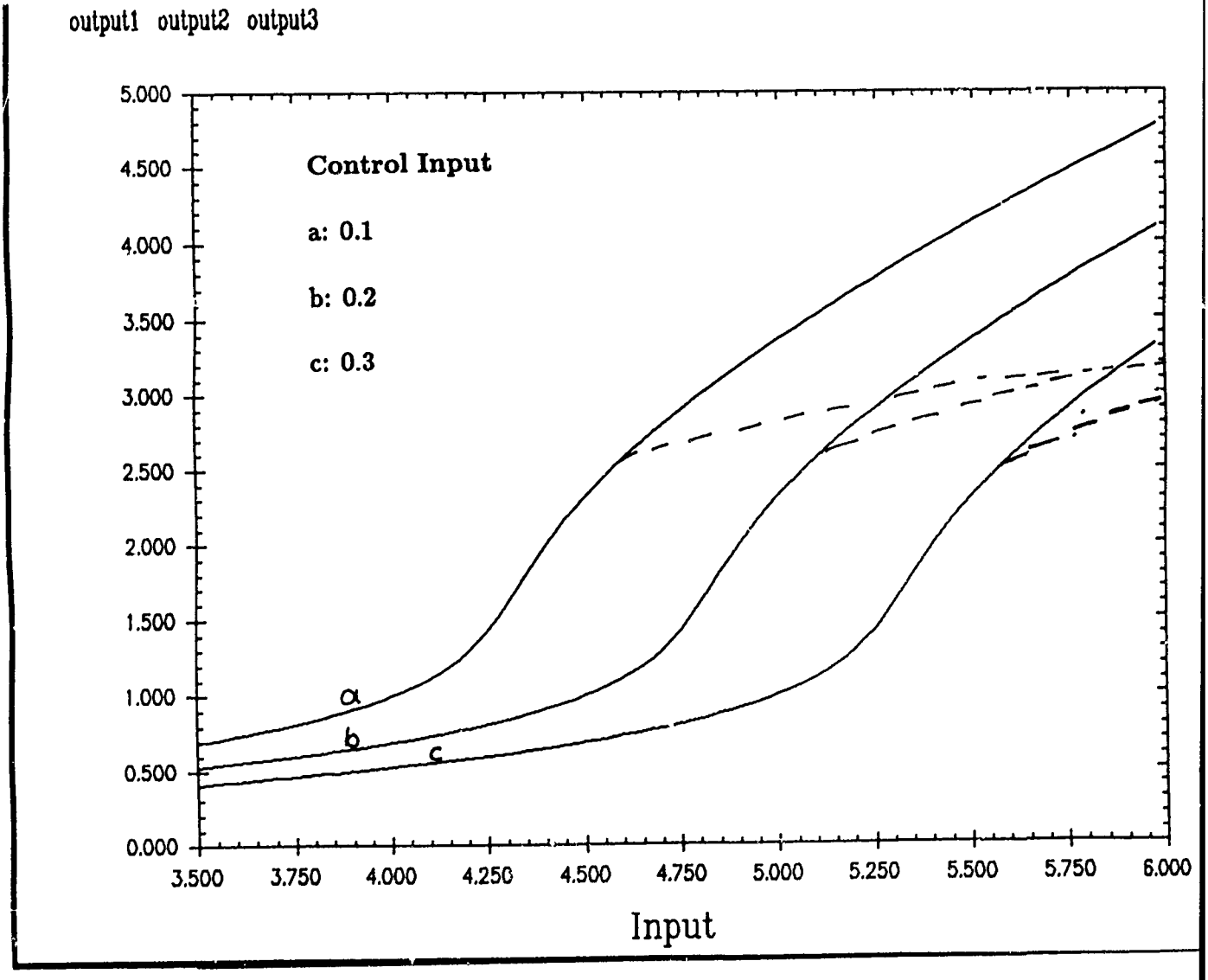


Figure 1: Three curves representing the signal output with three different control beam values. As the control beam intensity increases the threshold shifts to the right. The dashed lines indicate the use of a limiter.

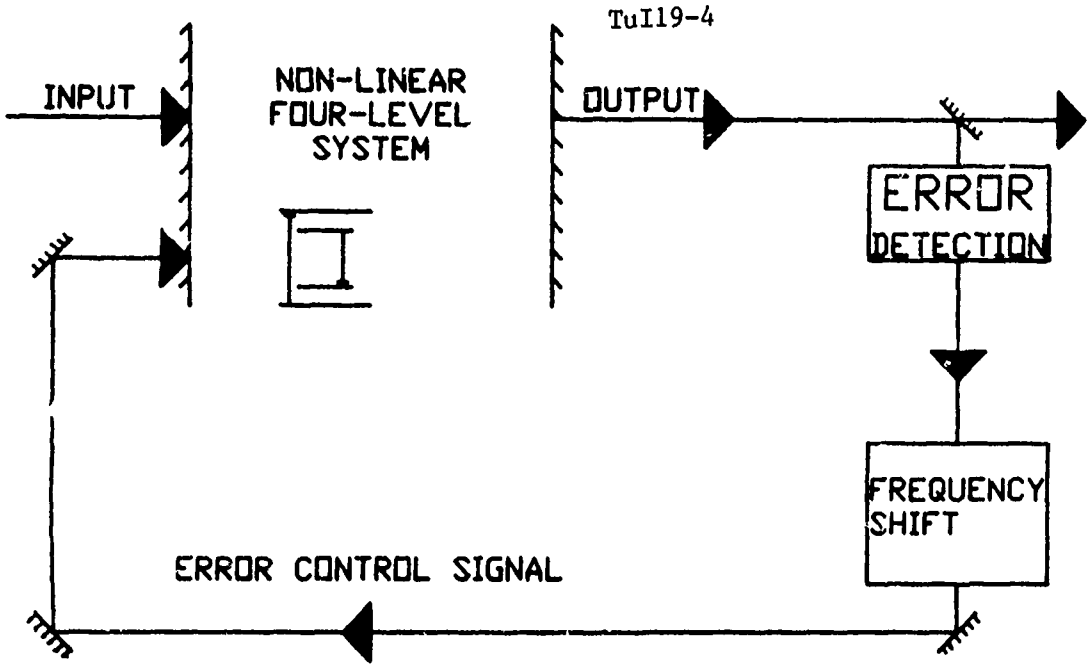


Figure 2: Self-correcting neuron from four level non-linear optical element. The ontrol signal is proportional to the error and it is fed back to shift the output curves till he desired response is achieved.

NEURAL NETWORK MODELS BASED ON OPTICAL RESONATOR DESIGNS

Steven C. Gustafson and Gordon R. Little
University of Dayton Research Institute
Dayton, Ohio 45469

This paper specifies a class of neural network models in a form suitable for performing computer simulation experiments and assessing possible optical implementations. These models are consistent with optical resonator designs that may include dynamic holograms and thresholded phase conjugate mirrors. The specification obtained here could be of near-term value in the development of new pattern recognition algorithms.

In many all-optical computing architectures, holograms are envisioned for interconnection and storage functions. Also nonlinear components, such as arrays of bistable optical devices or thresholded phase conjugate mirrors, are envisioned for decision operations. The necessary adaptation and feedback interactions between the interconnection and decision components are often achieved by incorporating these components in linear or ring resonators [1-6].

A simple and general mathematical formulation of a neural network model consistent with such optical resonator designs may be obtained by well-known methods in which plane wave amplitudes and phases are specified at discrete times separated by the resonator period. The model inputs and outputs are complex-element vectors. A state vector and a hologram matrix evolve in time according to a set of coupled nonlinear difference equations that represent, in general, a high-order threshold logic [7]. The hologram matrix is a function of the outer product matrix of the evolving state vector and has a form that depends on the hologram and resonator geometry.

A diagram of the model and equations for the model are given in Figure 1. Note the term in the hologram matrix equation proportional to the outer product matrix of the state vector with the diagonal elements replaced by the trace. This term may be readily derived for state vector elements interpreted as plane waves with pairwise-unequally-spaced propagation directions. (For the special case of equally-spaced propagation directions, the elements on each diagonal of the outer product matrix are replaced by their sum.) The hologram matrix is required to be stable and therefore attains a near-constant value after a sufficiently long time, and it is self-referenced in that no separate reference beams (e.g., at different angles for different recordings) are involved. Note also that the nonlinear operator performs no interconnection operations because it independently replaces each complex element of its argument by another complex element.

The hologram matrix could at least approximately represent many forms of diffracting structures: thin or thick, amplitude or phase, static or dynamic, reflection or transmission. The nonlinear operator could also approximate many types of components, including arrays of bistable optical devices and phase conjugate mirrors with thresholding and gain. Note that gain or some mechanism (such as phase conjugation) to compensate for wide-angle scattering from the hologram is necessary. Finally, note that the input and output

matrices A and B represent input and output devices such as beam splitters or phase shifters.

As a simple example of the model specified in Figure 1, suppose that input u, state v(t), and output w(t) are single-element vectors, that the hologram matrix H(t) is a complex constant $a\exp(i\phi)$, that the nonlinear operator is N = 1, and that A = B = b = 1, c = 0. Then the squared magnitude output of the model as a function of time is

$$|w(t)|^2 = \frac{a^{2(t+1)} - 2a^{t+1}\cos\phi(t+1) + 1}{1 - 2a\cos\phi + a^2}$$

which is a damped sinusoid modified by an additional exponential decay term. This example indicates that even a static-hologram, no-nonlinearity specification can lead to relatively complex but stable behavior in time.

One possible test of the model as a neural network pattern classifier is as follows. Set H(0) = I. Input a vector u of features from a training pattern. Find the stable matrix $H(\infty)$ and the stable vector $w(\infty)$. Repeat this process for one training pattern from each pattern class. Set H(0) equal to the sum of the $H(\infty)$ for all training patterns. Input a vector u of pattern features for a pattern that may or may not be a training pattern. Test the stable vector output $w(\infty)$ to determine if its squared absolute difference with the output $w(\infty)$ for the training pattern of the correct class is smaller than for the training pattern of any other class. If this test is successful, a pattern classification algorithm based on the model could be developed. Assuming that suitable optical materials and components become available, a long-term consequence could be the development of neural network pattern recognizers based on optical resonator designs.

1. S. C. Gustafson and G. R. Little, "Optical Classification of Binary Patterns," Proc. SPIE 882, 83, Los Angeles, CA, January 1988.

2. K. Wagner and D. Psaltis, "Multilayer Optical Learning Networks," Appl. Opt. 26, 5061 (1987).

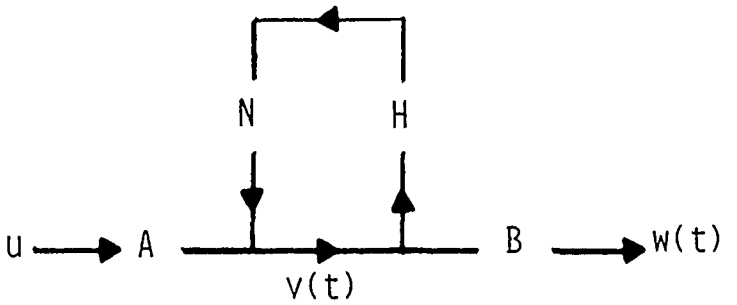
3. D. Z. Anderson and M. J. O'Callahan, "Competitive Learning, Unlearning, and Forgetting in Optical Resonators," Proc. SPIE 882, 181, Los Angeles, CA, January 1988.

4. Y. Owechko, "Optoelectronic Resonator Neural Networks," Appl. Opt. 26, 5104 (1987).

5. M. S. Cohen, "Coupled Mode Resonators for Optical Computing," Proc. SPIE paper 882, 122, Los Angeles, CA, January 1988.

6. D. Psaltis and N. Farhat, "Optical Information Processing Based on an Associative Memory of Neural Nets with Thresholding and Feedback," Opt. Lett. 10, 98 (1985).

7. C. L. Giles and T. Maxwell, "Learning, Invariance, and Generalization in High-Order Neural Networks," Appl. Opt. 26, 4972 (1987).



$$v(t+1) = N[H(t)\ v(t)] + Au$$

$$v(o) = A\ u$$

$$w(t) = B\ v(t)$$

$$H(t+1) = b\ H(t) + c\ M[v(t)\ v(:)^{\dagger}]$$

u	Input vector of m complex elements.
v(t)	Internal vector of n complex elements.
w(t)	Output vector of p complex elements.
H(t)	Hologram matrix of n x n complex elements, generally near-diagonal and stable.
M[A]	Matrix equal to n x n matrix A with diagonal elements replaced by trace (A).
N	Finite operator, generally nonlinear.
A, B, b, c	Matrix and single-element complex constants.
I, †	Identity matrix and complex transpose.
t	Discrete time t = 0, 1, 2, ...

Figure 1. Neural Network Model Based on Optical Resonator.

Implementation of NETL Knowledge-Base System with Programmable Opto-Electronic Multiprocessor architecture.

Fouad Kiamilev, Sadik Esener, Dept. of Electrical Engineering, University of California - San Diego, Mail Code R-007, La Jolla, CA 92093

The field of Artificial Intelligence has reached a critical stage. A good Artificial Intelligence system, must include a knowledge-base with abilities comparable to those possessed by humans. To date, we have not been able to achieve this goal. Even in a restricted problem domain, current sequential search techniques are much too slow to handle a knowledge-base of sufficient size to produce a human-like intelligence, since the search time increases linearly with the increasing size of the knowledge-base.

On the other hand, theoretical work by S.E. Fahlman [1] has shown that storing knowledge as a pattern of interconnections between many very simple processing elements allows searches to be performed very quickly. The basic idea is to store the knowledge-base as a graph where individual concepts are assigned to processing elements and the interconnections between the processing elements represent the relations between the concepts [see fig. 1]. Search operations are performed by marking specific node processors and then propagating these markers in parallel through the network. The set of conventions and processing algorithms for representing the knowledge in such a parallel network is called NETL. Fahlman has shown that NETL is capable of performing search operations on the knowledge base, simple deductions, learning, consistency checks, matching and symbolic recognition tasks. Fahlman's theoretical work paves the way to creation of human-like knowledge bases. The important and unique feature of the NETL system is that the time required to perform a search is essentially a constant independent of the size of the knowledge-base.

The importance of Fahlman's work was recognized by W. D. Hillis, who originally designed the Connection Machine parallel computer to implement NETL [2]. The Connection Machine consists of 64K simple 1-bit processing elements interconnected in a hypercube network topology. However, the implementation of NETL on the Connection Machine has not been effective. First, a large number of processing elements (>64K) is required to implement realistic problems. Second, the NETL interconnection can not be directly mapped onto the hypercube. Instead, routing is used to communicate between processing elements that are connected in the NETL graph. Since marker propagation is performed in parallel, there is a large overhead due to the latency and contention introduced by the fixed hypercube interconnect.

Fortunately, recent advances in opto-electronic computing give us hope that a hardware implementation of NETL will soon be possible. The hardware implementation will be based on Programmable Opto-Electronic Multiprocessor (POEM) systems. Each POEM system consists of wafer scale integration of simple processing elements with programmable optical interconnects [3]. The NETL system can be directly mapped onto the POEM hardware by programming the optical interconnects.

The NETL System for POEM.

As described earlier, the NETL system is basically a massively-parallel semantic network, consisting of nodes representing the concepts, and links representing relationships between these concepts. For example, in figure 1, GRAY is a concept and HAS-COLOR is a link. Concepts and links are stored in simple processing elements with a few marker bits of storage. Commands are send to the processing elements by the system controller for SIMD style execution. The processing elements can conditionally execute the commands sent to them based on the state of their marker bits. For

example, the system controller can order all nodes whose markers 1 and 2 are turned on, to set their marker 3 on.

On the other hand, all of the links in the system might be instructed to sense whether one of the nodes they are connected to has marker 1 on and, if so, to set marker 1 in the other attached node on. This has the effect of propagating the markers through the network and gives NETL the capability to perform fast knowledge-base searches. For example, in figure 1, to find the common properties of the CARROT and CLYDE concepts, the system controller turns on marker 1 in the CLYDE element and marker 2 in the CARROT element. Next, the markers are propagated in parallel through the network. Finally, the system controller can direct all processing elements who have both markers 1 and 2 turned on, to report to the controller. In this example, in a few steps we can determine that CARROT and CLYDE are both LIVING-THINGS and PHYSICAL-OBJECTS.

The fine-grain POEM machine seems to be ideal for implementing NETL because of its large number of processing elements (as large as 500K) and the programmable optical interconnects. In POEM, we should be able to directly map the NETL network onto the hardware of the machine by programming the optical interconnections. The node and link elements in the NETL graph are represented by the processing elements and the interconnections in the POEM machine, the system level controller in NETL by the system controller described in POEM, and the marker bits can be implemented in the 64 bits of local memory in each of the POEM processing elements.

In NETL graphs, however, the number of edges that are incident onto a node or that emerge from a node is not limited. Such arbitrary fan-out and fan-in capability cannot be implemented in a straight-forward fashion in POEM, because the hardware complexity of the processing element as well as the time required to perform the search would linearly increase with increasing node fan-in and fan-out. The solution to this problem is to use fan-in and fan-out units [2]. Fan-out from a node can be accomplished by allocating additional processing elements, called fan-out units, and appropriately programming the interconnections to attach the fan-out units to the node in a tree topology, with the node being at the root. In such fashion, a fan-out of N is accomplished in O(log(N)) time steps, with O(N) fan-out processing elements. Arbitrary fan-in can be accomplished in similar manner. Figure 2 shows the semantic network from figure 1 with added fan-in and fan-out units. This network can be directly mapped onto the POEM machine. In POEM the same processing element is used for concept and link nodes as well as fan-in and fan-out units. It should be noted that fan-in and fan-out units accomplish funnelling and broadcasting of information in the minimum possible amount of time, and therefore they do not introduce latency into the computation.

The POEM NETL system is capable of learning. For instance, suppose that we wish to add a new piece of information that CABBAGE is a PLANT to the network shown in figure 2. To perform this, the system level controller selects an unused processing element and assigns the concept of CABBAGE to it. An IS-A link and a fan-out processing elements are also allocated by the system controller. The interconnection pattern is reprogrammed to connect the CABBAGE processing element to the PLANT node, via the IS-A link and the fan-out processing elements. The IS-A link provides a way for a concept to inherit a large amount of information without actually allocating space to store the inherent information again, but rather by pointing to the node that defines the class for this information.

It is noted that the POEM NETL system can be enhanced by adding a numerical weight to each connection [4], such that conclusions carry a certainty factor instead of a boolean truth value. In such a system, each observed feature could vote for the hypothesis that it supports, with various strength for each vote. With this modification the NETL system becomes very similar to a neural network.

As a demonstration of POEM NETL usage, we can envision a POEM NETL knowledge-base storing information for a medical diagnosis expert system. The system controller implements the expert system which queries the user about the symptoms of the disease and uses the knowledge base to match the symptoms against the diseases. For example, if it is known that the patient has high temperature, the POEM NETL system can look up some other feature that is related to the high temperature concept, such a feature could be coughing. Next the expert system can query the user to verify that the newly uncovered features are indeed present. Once this is done, we have more specific information about the disease. Now we can repeat this process again to identify still more features of the disease. In such a way, the disease and its features emerge together, little by little.

Conclusion

In this paper, we have shown that POEM is well suited for implementing massively-parallel knowledge-bases. Basically, this is a result of the fact that we are able to map the data structure of the symbolic application directly onto the opto-electronic multiprocessor system. In the POEM NETL system, the number of processing elements increases linearly as new concepts and relations are added to the knowledge-base, while the search time remains essentially constant. The POEM NETL system uses space very efficiently, by storing only the required information.

Although the technology for full-scale POEM has not yet been fully developed, there should be much motivation for doing so. The wafer scale POEM machine can implement a NETL system in a very compact structure and with low energy requirement. Therefore, this system has potentially a low cost and high reliabilty. These attributes will not be attainable with the current VLSI technology no matter what the advancement there will be, because these capabilities stem from the unique abilities offered to us by integrating the signal processing capability of silicon and the connection capability of optics.

Acknowledgment

This work has been funded by DARPA under the management of AFOSR by contract 88-0022, and by ONR by contract N00014-86-K-0697. Additional funding has been provided through the Office of Naval Research Graduate Fellowship program.

References:

1) S.E. Fahlman, "*NETL: A System for Representing and Using Real-World Knowledge*," Cambridge, MA: MIT Press, 1979

2) W.D. Hillis, "*The Connection Machine*," Cambridge, MA: MIT Press, 1985

3) F.Kiamilev, S.Esener, R.Paturi, Y.Fainman, P.Mercier, C.Guest, S.H.Lee, "*A Comparison of Programmable Opto-Electronic Multiprocessors and Symbolic Substitution for Digital Optical Computing*,", to be published in Opt. Eng. Jan 1989

4) J.A. Feldman and D.H. Ballard, "*Connectionist Models and their Properties*," Cognitive Science 6:205-254, 1982

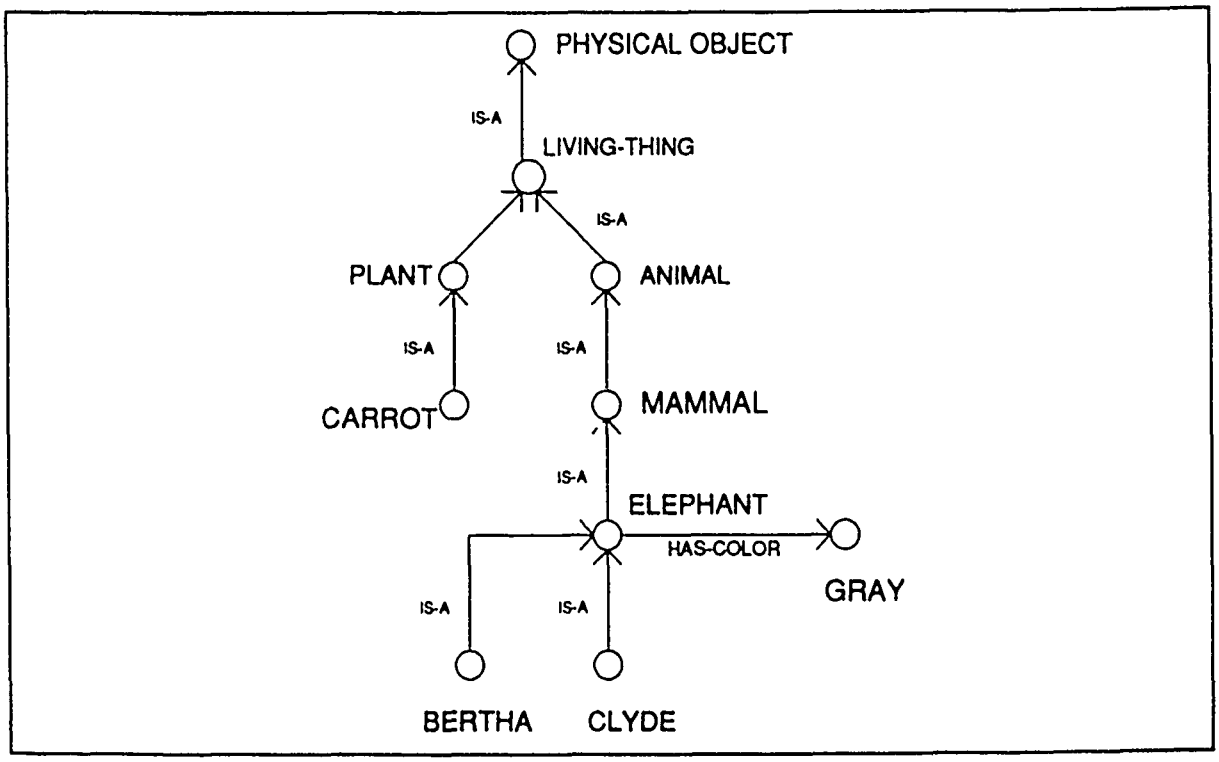


Figure 1: An example of a Semantic Network

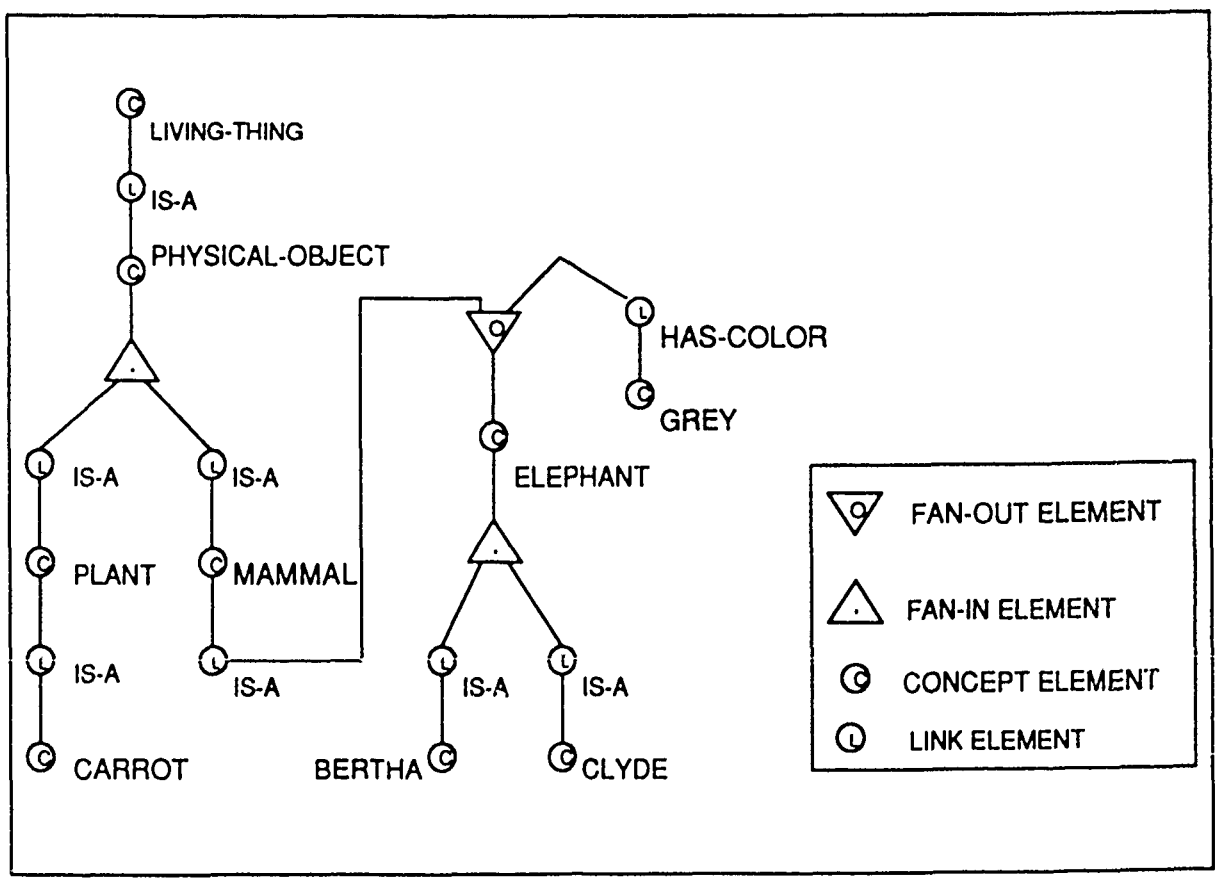


Figure 2: Semantic Network with fan-out/fan-in units

Trainable Optical Network for Pattern Recognition

John H. Hong and Pochi Yeh
Rockwell Science Center
1049 Camino dos Rios, A25A
Thousand Oaks, CA 91360

Abstract

An optical implementation of a single layer network for pattern recognition is described in which both subtractive and additive changes of the weights can be made.

Summary

The processing of information in neural networks differs from conventional approaches in that the interconnections play the dominant role rather than acting as mere communication pathways. The fact that this interconnection intensive computation can be achieved using optical techniques was realized by many, the use of volume holograms being one such example [1-3]. Although the interconnections can be computed and fixed for prescribed tasks in which the problem parameters do not change, the idea of a neural network which can be adapted on-line to solve problems is especially appealing.

Shown in Fig. 1 is a diagram depicting the most basic one-layer network with N input elements and one output. The weighted sum of the input pattern elements is thresholded to yield the output

$$y = g(\sum_{i=1}^{N} w_i x_i), \qquad g(z) = \begin{cases} 1 & \text{if } z>0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where g(.) is the thresholding nonlinearlity, w_i is the i^{th} weight, and x_i is the i^{th} element of the input pattern. Such a system can be used to dichotomize a set of patterns, and multiple layered networks can be built up using this as the basic building block. Learning algorithms are simple and can be characterized by the update equation

$$w_i(p+1) = w_i(p) + \alpha(p)\, x_i(p), \tag{2}$$

where $w_i(p)$ is the i^{th} weight at time p, $x_i(p)$ is the i^{th} element of the pattern shown at time p, and $\alpha(p)$ is a multiplier which depends on the particular learning algorithm. For perceptron learning [4],

$$\alpha(p) = \begin{cases} 0 & \text{if output } y(p) \text{ was correct} \\ 1 & \text{if } y(p)=0 \text{ but should have been 1} \\ -1 & \text{if } y(p)=1 \text{ but should have been 0.} \end{cases} \tag{3}$$

The threshold bias can be absorbed into the patterns by choosing one element of each pattern to be always equal to 1. Note that both additive and subtractive changes must be made to implement the algorithm directly. Extension to multi-category pattern classification can be achieved simply by having a matrix of weights and a multiplicity of output units.

The basic components to implement the network described above are an input device to convert the patterns into the appropriate format (e.g., electrical to optical, incoherent to coherent optical), interconnection device, and a thresholding nonlinear device for the output unit. The function of the interconnections in this context is to simply compute the inner product between the input pattern x_i and the weights w_i. Volume holograms can be used to implement such functions [5] in a way that is extendable to the multiple category case (i.e., multiple inner products). Consider the arrangement shown in Fig. 2 where a holographic medium is positioned at the Fourier plane of lens L1. The input pattern is displayed in the spatial light modulator (SLM) which is positioned at the front focal plane of the same lens. A hologram is exposed with a pattern w(x,y) in the SLM and a reference plane wave as shown. After development, another pattern f(x,y) is loaded into the SLM. The light passing through the SLM is diffracted by the hologram and the diffracted amplitude is the inner product between the two patterns w and f. Clearly, this is an overkill since the same function could have been achieved with a

planar hologram. However, in the multiple category case where a number of different inner products need to be computed simultaneously, the added dimension afforded by the volume hologram is necessary unless one resorts to spatial multiplexing of the planar hologram [6]. Multiple category classification is achieved by what is essentially an angular multiplexing of the volume hologram. This is shown in Fig. 3 where multiple holograms are written using the various reference plane waves. For simplicity, we focus on the single output case for our discussions.

By virtue of its dynamic nature, photorefractive crystals are ideal candidates for the holographic medium. In addition, crystals such as $LiNbO_3$, $BaTiO_3$ and SBN are by far the most efficient holographic using relatively low optical intensity levels (e.g., $1W/cm^2$). Volume index gratings which realize the interconnective weights w_i can be recorded or updated using holographic interference.

As the basic unit of the overall system, consider Fig. 4, in which an arrangement to allow for multiple exposures of a photorefractive hologram is shown. This setup exploits the Stoke's principle of reversibility for light to allow for $(0,\pi)$ phase control of the exposed gratings [7]. The two light sources can be mutually incoherent as long as their nominal wavelengths are the same. If the SLM contains a picture whose amplitude distribution is given by a(x,y), then the grating written in the crystal due to source 1 can be described by

$$g_1(x,y) = K\,(1- \exp(-t_1/\tau))\,\frac{a(x,y)}{1+|a(x,y)|^2}, \qquad (4)$$

where τ is the time constant of the medium (assuming intensity is kept constant for all exposures), t_1 is the exposure time, and K is a constant determined by the characteristics of the particular crystal. If, without changing the picture, source 1 is turned off and source 2 is turned on, the new grating can be shown to be proportional to the first with the opposite sign. In particular, if the second exposure time duration is t_2, then

$$g(x,y) = K\,\left\{(1-\exp(-t_1/\tau))\exp(-t_2/\tau) - (1-\exp(-t_2/\tau))\right\}\,\frac{a(x,y)}{1+|a(x,y)|^2}. \qquad (5)$$

True subtractive weight changes are thus possible without the use of external phase shifters (e.g., pzt mirrors, eo modulators). The network implementation is completed as shown in Fig. 5 by adding a photodetector with subsequent electronic thresholding and shutters controlled by the error signal as shown.

Reference

1. D. Anderson, D. Lininger, "Dynamic optical interconnects: volume holograms as optical two-port operators," Appl. Opt. **26**, p.5031 (1987).
2. D. Psaltis, D. Brady, and K. Wagner, "Adaptive optical neural computers," Appl. Opt. **27**, p.1752 (1988).
3. A. Yariv and S. Kwong, "Associative memories based on message-bearing optical modes in phase conjugate resonators," Opt. Lett. **11**, #118 (1986).
4. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis* (John Wiley & Sons, New York, 1973).
5. D. Psaltis, C. H. Park, and J. Hong, "Higher order associative memories and their optical implementations," Neural Networks **1**, p.149 (1988).
6. H. J. Caulfield, "Parallel N^4 weighted optical interconnections," Appl. Opt. **26**, p.4039 (1987).
7. P. Yeh, T. Y. Chang, and P. H. Beckwith, "Real-time optical image subtraction using dynamic holographic interference in photorefractive media," Opt. Lett. **13**, p.586 (1988).

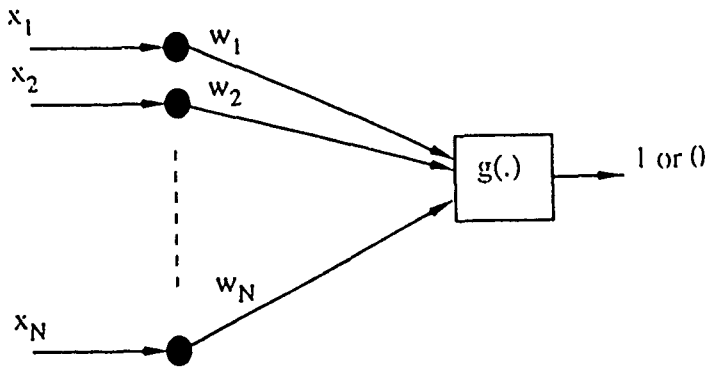


Fig. 1 One Layer Network

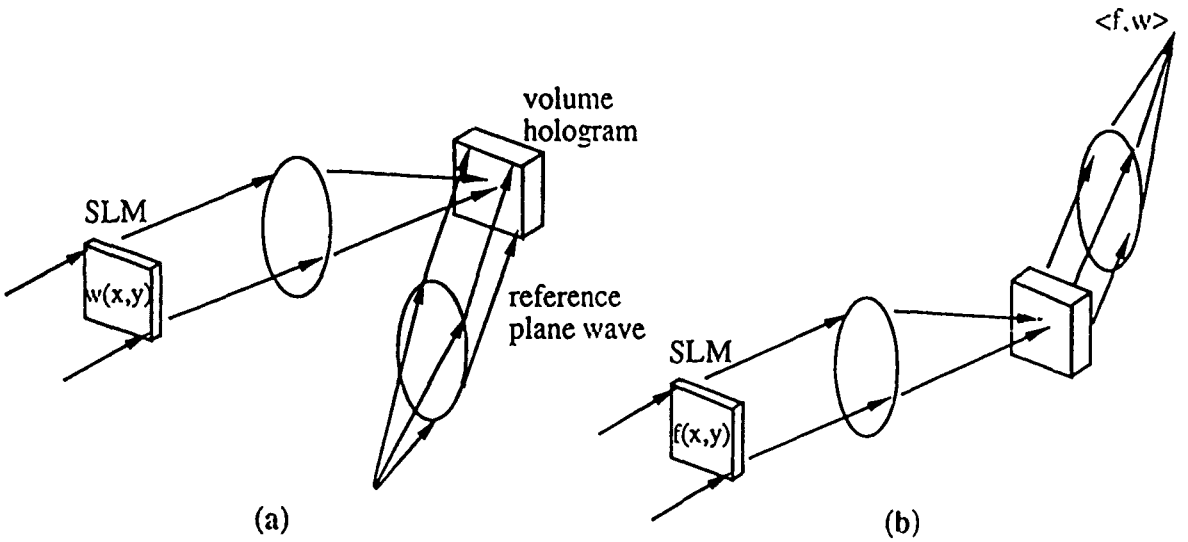


Fig. 2 Holographic Inner Products
a) recording b) read out

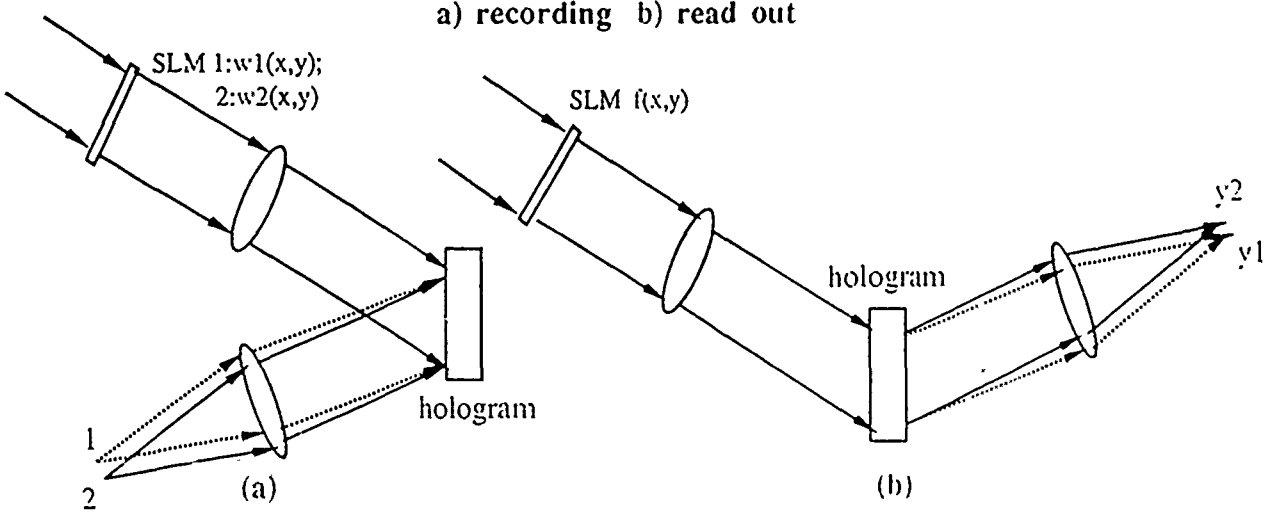


Fig. 3 Angular Multiplexing for Multiple Category Classification
a) recording [pt. source 1 is activated when w1(x,y) is in SLM and pt. source 2 is activated when w2(x,y) is in SLM; b) read out [y1=<w1,f> and y2=<w2,f>].

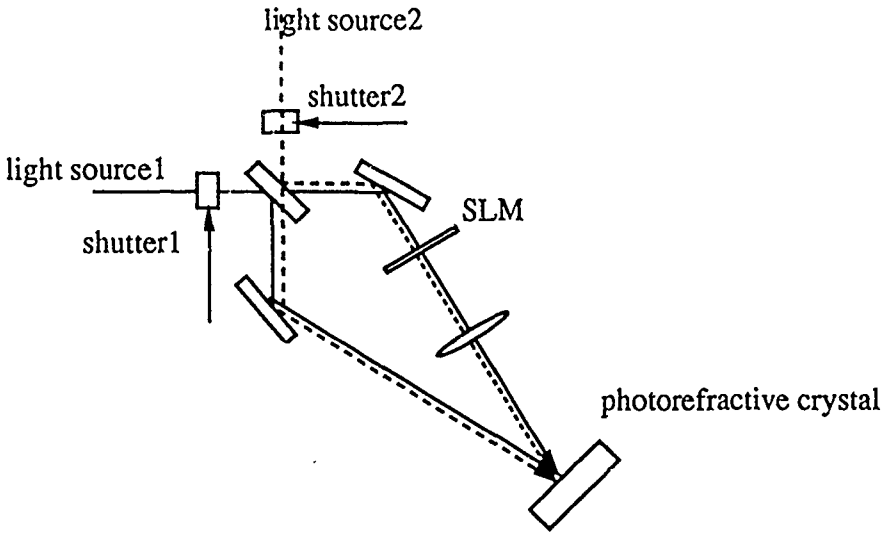


Fig. 4 (0,π) phase control using Stoke's Principle [the grating written by source 1 is 180 degrees out of phase with respect to that written by source 2].

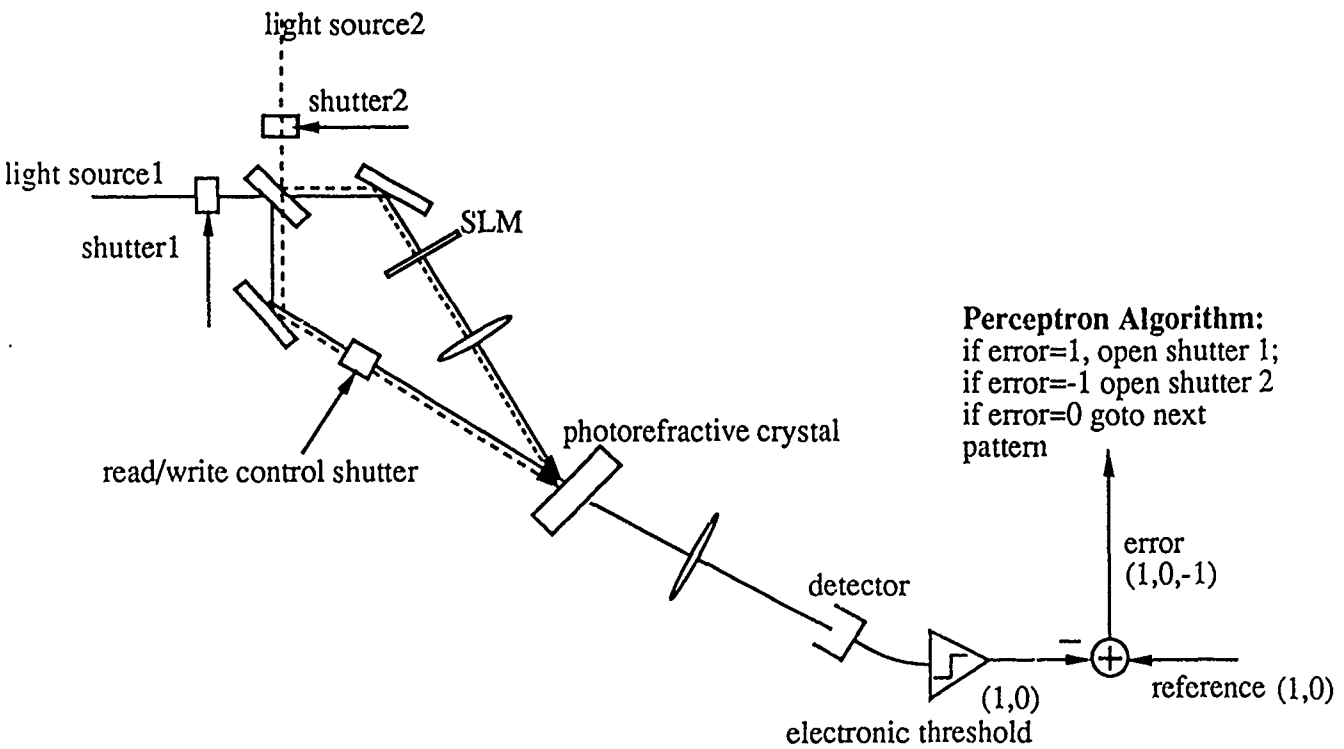


Fig. 5 Optical Network with subtractive weight change capability read/write control shutter (open for weight changes and closed for interrogation); [operation of shutters should be mutually exclusive: i.e., if 1(2)is on, 2 (1) is off]; error signal is 1 if output=0 and reference=1, 0 if output=reference, and -1 if output=1 and reference=0.

OPTICAL IMPLEMENTATION OF ASSOCIATION AND LEARNING BASED ON PRIMO/LIGHT VALVE DEVICES

U. Efron and Y. Owechko
Hughes Research Laboratories
3011 Malibu Canyon Road
Malibu, California 90265

Summary

The effort proposed here is aimed at demonstrating the use of existing optical and electrooptical components in implementing adaptive neural network systems. Specifically, we propose a system which implements the outer product model[1,2] of auto- or hetero-association. Both learning and association operation can be executed using this system which is based on the use of 1-D striped-electrode fast input modulators (based on the PRIMO technology[3]) coupled with a time integrating photoactivated liquid crystal light valve.

The purpose of this concept is (1) to demonstrate the potential of adaptive optical systems for use as efficient parallel-addressed neural net systems, and (2) to study their capabilities and evaluate the ultimate performance expected in these implementations. The system is essentially based on three electrooptic components (Figure 1): (a) two 1-D PLZT modulators, (b) a liquid crystal light valve, and (c) a 1-D imaging detector. Linear array detectors are basically available as off-the-shelf items. As for the 1-D PLZT modulators, such devices have been under development for an optical computing PRIMO system.[3] Operation of a 64-element modulator at ~1 μSec response time was recently demonstrated. For the integrating, photoactivated liquid crystal light valve either the CdS[4] or the silicon-based devices[5] can be used.

The main approach is shown in Figure 1. The system consists of two 1-D modulators (MOD1, MOD2) which will be based at this point on PLZT technology. These two layers will be used to construct the interconnect matrix, T_{ij}. Each of the 1-D modulators consists of striped-electrode patterns on PLZT. The two layers are oriented so that their electrodes are crossed. Thus by modulating one with a set of m vectors $U_i^{(m)}$ and the other with a set of m vectors $V_j^{(m)}$ supplied by the microprocessor, one optically forms the T_{ij} matrix as an outer-product[6,3] where:

$$T_{ij} = \sum_m U_i^{(m)} V_j^{(m)}$$

The vector elements are assumed to be ±1. The vectors $U_i^{(m)}$ and $V_j^{(m)}$ will be supplied at a relatively fast rate (≅10 μSec/vector) by the PLZT modulators. Since the LCLV has a response time of ≅10 msec, one will be able to integrate up to a few hundred outer products or vectors in this LCLV-based T_{ij} matrix. Bipolar analog T_{ij} values can be represented in PRIMO using temporal or spatial multiplexing.(3) Having completed the learning phase, the liquid crystal will be modulated with the T_{ij} information for a duration of ≅ 10 msec. During this period one can proceed with the interrogation or the association operation. A third 1-D PLZT layer (MOD3) will then input the vector $V_j^{(o)}$ to be associated. This 1-D vector, whose components are spread in the vertical dimension, will be optically multiplied by the T_{ij} (LCLV) matrix by illuminating the MOD3 modulator using the polarizing beam splitter as shown. Thus in each line, i, of the T_{ij} matrix the columns (running j) are multiplied by the (same) $V_j^{(o)}$ information. By using a cylindrical lens at the output of the beam splitter, as shown, we effectively sum:

$$\hat{V}_i = \sum_j T_{ij} V_j^{(o)}$$

for each i-line. Thus each of the i-pixels formed will correspond to the desired i^{th} component $\hat{V}_i$ of the matrix-vector product to be compared against a threshold level according to the outer-product model. This operation will be carried out by detecting the resultant vector $\hat{V}_i$ using the linear detector and an electronic thresholder controlled by the microprocessor. To complete the association, the thresholded $\hat{V}_i$ is fed back into MOD3 and the matrix vector multiplication operation is repeated. The resultant sequence of $\hat{V}_i^{(n)}$ ($V_i^{(1)}, V_i^{(2)} \ldots V_i^{(n)}$) will be tested for convergence which, once reached, will yield the closest association with the interrogating input vector, $\hat{V}_i^{(o)}$. One type of learning that this system can perform is a statistical learning as suggested by Anderson.(1) He showed that for a neuron system coupled in an auto association scheme the multiplication of the weight matrix W_{ij} by the interrogating vector $\hat{V}_i^{(o)}$ will

result in the output vector being one of the stable state, with a weight which is proportional to the frequency in which this vector appeared during the learning phase.

The system can therefore learn to enhance common features which appear in different patterns during the teaching (learning) phase. Thus when a vector appearing during the interrogation phase has a feature which had appeared as a vector with a high frequency of repetition during the learning phase, the output of the system will tend to be that particular feature. The statistical learning capability is strictly true only for orthonormal state vectors. We do expect, however, that the enhancement of the T_{ij} weights associated with this effect will also occur to some extent for non-orthonormal vectors. It should be emphasized, however, that even without this interesting feature, the proposed system offers adaptive learning in the sense of learning the weights corresponding to the association of vectors U_i, V_j -- in other words, a modifiable-weight T_{ij} matrix.

Another interesting subject to be studied under this program is the possibility of using the vectors $\hat{V}_i$ obtained during the association (interrogation) phase as inputs for a new, modified T_{ij}. This opens up the possibility of demonstrating a system that would adapt itself to new state vectors (environment). this can be implemented if the interrogating vectors (which are input to MOD3 of Figure 1) are made to represent external vectors supplied by the environment which we wish to learn and recognize.

Finally, we wish to point out that an electro-optical implementation of the Hopfield-Anderson model was previously demonstrated.[7] The use of acoustooptic cells in conjunction with a 2-D spatial light modulator for similar implementation was recently suggested.[8]

References

1. J. A. Anderson et al, Psych. Rev. 84, 413 (1977).
2. J. J. Hopfield, Proc. Natl. Acad. Sci. 79, 2559 (1982).
3. B. H. Soffer et al, Appl. Opt. 25, 2295 (1986).
4. J. Grinberg et al, Opt. Eng. 14, 217 (1975).
5. U. Efron et al, J. Appl. Phys. 57, 1356 (1985).
6. R. Athale and W. C. Collins, Appl. Opt. 21, 2089 (1982).
7. D. Psaltis and N. Farhat, Opt. Lett. 10, 98 (1985).
8. A. D. Fisher et al, Appl. Opt. 26, 5039 (1987).

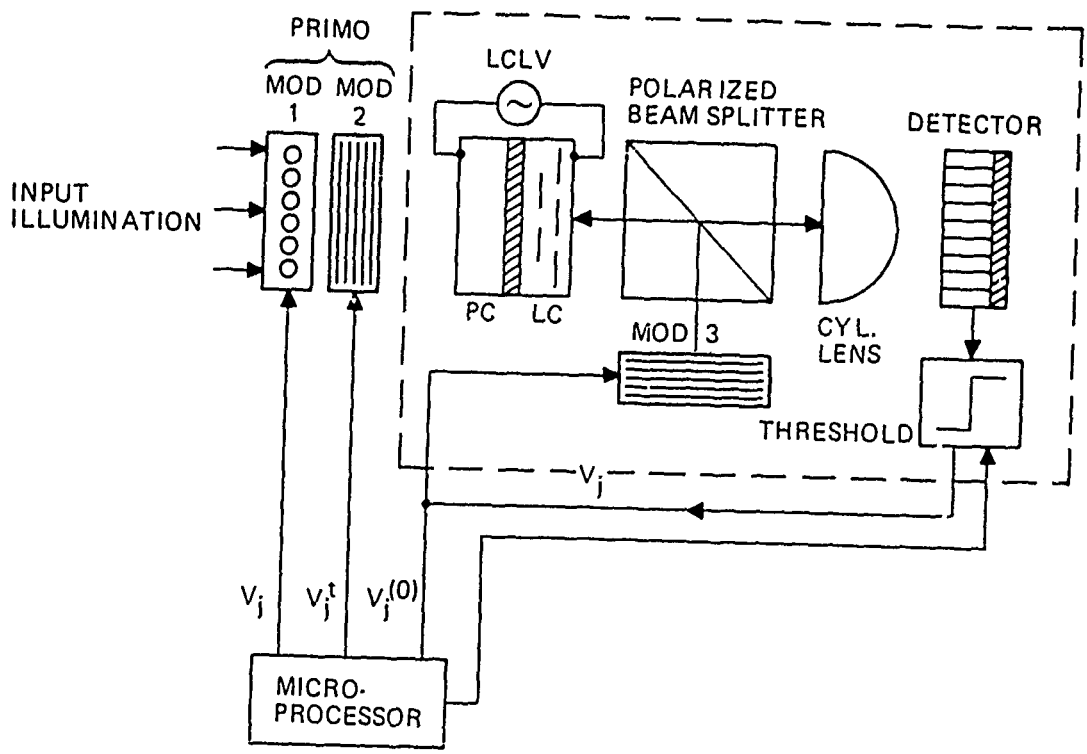


Figure 1. An optical associative memory system with learning capability using PRIMO/LCLV.

The Thermal Nonlinear Microcavity and Optical Computing

C. Godsalve and E. Abraham

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

Thin film Fabry-Perot etalons which have a temperature dependent refractive index exhibit bistability or gain at room temperature (e.g. ZnSe) and at optical frequencies. These features make them candidates for digital optical computing. An N x N array of elements can be generated in a single filter by an array of laser beams. As a result, thermal crosstalk develops which is long range and only a few elements per cm^2 on such a filter can operate independently [1,2]. However if each filter is mounted on its own separate 'turret', crosstalk can be reduced to the extent that 10^4 microcavities (or pixels) can operate independently per cm^2 [3,4].

We consider a cylindrical pixel mounted on some substrate as in Fig. 1, illuminated by a Gaussian beam of spot radius s, the pixel dimensions are shown in Fig. 2, and the symbols κ_p and κ_s are used for the pixel and substrate thermal conductivities.

We solve the heat equations for pixel and substrate by making the approximation that the NLIF is thin compared with the pixel and include the nonlinearity by using the temperature dependence of the absorptance 'a' [5]

$$a(T_f) = a_0/(1 + g(\bar{T}-T_f)^2)$$

where T_f is the film temperature, (g is related to the finesse and T the detuning) and include the absorption of the laser beam via the heat source Q(r,t)

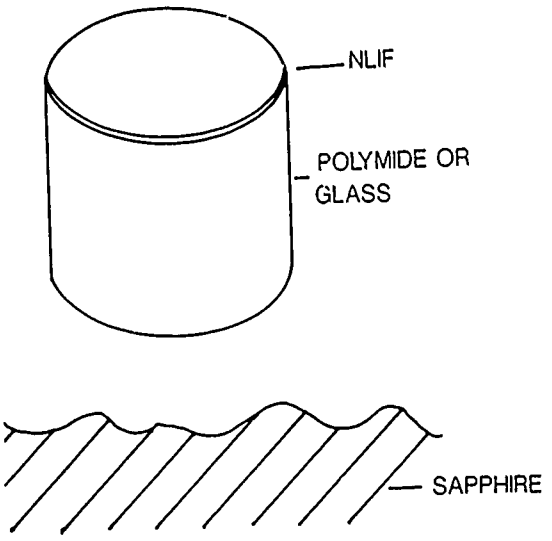


Fig. 1. The nonlinear microcavity or pixellated NLIF.

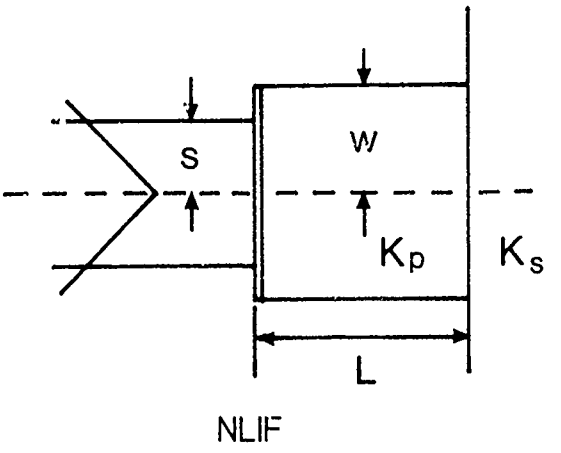


Fig. 2. Pixel dimensions and thermal conductivities.

$$Q(r,t) = I_o e^{-r^2/s^2}\ \delta(z+l)\ \alpha(T_f(r))$$

By making the approximation that the NLIF is thin compared with the pixel we solve the heat equation analytically (this is not included here for the sake of brevity). A promising combination is polyimide pixels on a sapphire substrate. Here we include results for the switching powers for both glass and polyimide pixels as a function of spot radius and pixel height for a 10 μm radius pixel.

We see that milliwatt switching powers are possible for suitable pixel dimensions and thermal conductivities. The thermal time constant of the pixel scales with $s^2\rho_p C_p/\kappa_p$ where ρ and c are the pixel densities and specific heat capacities for s = 5 μm, w = 10 μm, l = 10 μm. This gives switching times for a 5 μm spot size of the order of 5 ms for glass and 35 ms for polyimide using packing densities of 5×10^3 cm^{-3} for glass pixels, and

10^4 for polyimide, this yields the number of operations per second per cm^2 as $\sim 10^6$ for glass and $\sim 1.4 \times 10^6$ for polyimide.

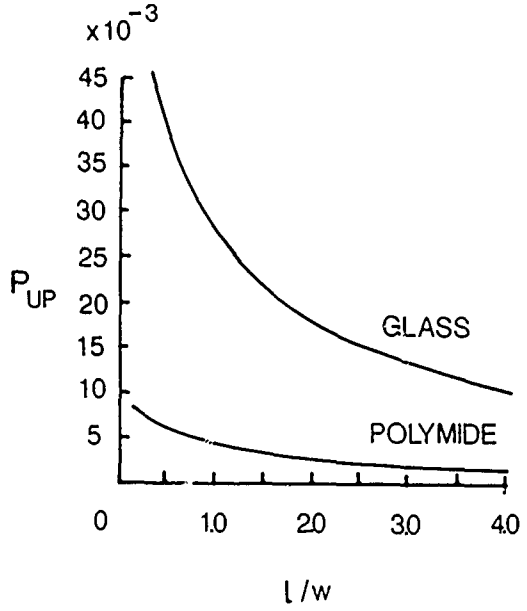


Fig. 3. Switch up power as a function of l/w for w = 10 μm.

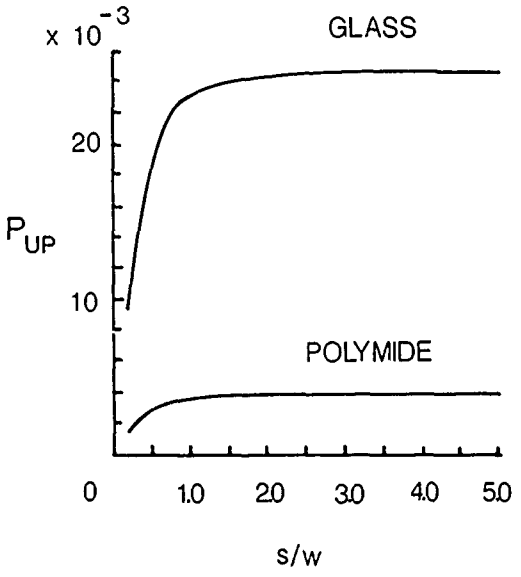


Fig. 4. Switch up power as a function of l/w for w = l = 10 μm.

References

[1] Abraham, E. and Rae, C., 1987, J. Opt. Soc. Am., 4, p. 490.

[2] Abraham, E., 1986, Opt. Lett., 11, p. 689.

[3] Abraham, E., Godsalve, C., Wherrett, B.S., 1988, J. Appl. Phys., 64(1), p. 21.

[4] Godsalve, C., Abraham, E., 1988, J. Phys. D: Appl. Phys., 21, p. S121.

[5] Janossy, I., Taghizadeh, M.R., Mathew, J.G.H., Smith, S.D., 1985, IEEE J. Quantum Electron., QE-21, p. 1447.

TWO BEAM COUPLING POLARIZATION PROPERTIES IN BSO USING ALTERNATING ELECTRIC FIELDS

G. Pauliat, G. Roosen

Institut d'Optique Théorique et Appliquée,
Unité Associée au CNRS,
Centre Universitaire d'Orsay, Bâtiment 503, B.P. 43
91406 ORSAY Cédex FRANCE

=====

Because of the inherent loss of optical systems, image amplification can be required in massively parallel optical architectures. This amplification can be obtained through the two wave mixing interaction in photorefractive crystals [1]. However, for sensitive material such as sillenite crystals one has to compensate for the relatively low electrooptic coefficient by an appropriate enhancement recording technique. One method to enlarge the photorefractive gain is to apply an alternating field to the crystal [2]. Consequently all the properties of the amplified beam (phase, polarization and intensity) will oscillate at the electric field frequency. Therefore, special care should be taken when inserting such an amplifier in an optical system so that the time dependent characteristics of the amplified beam are not disturbing the operation of any next non linear device. For this reason, using the coupled wave formalism [3], we study in the following all the features of the amplified beam under an alternating electric field.

Two beam coupling enhancement under an AC field in sillenite crystals is achieved using the following configuration. The two coherent optical waves, the pump beam $\vec{R}$ and the probe beam $\vec{S}$ lying in the (110) plane are incident on the $(\bar{1}10)$ face while the AC field is applied along the [001] crystallographic axis. The induced photorefractive grating is $\pi/2$ phase shifted relative to the interference pattern and gives rise to an energy redistribution. The probe beam can thus be amplified (or depleted) at the expense of the pump beam. Hereafter we assume that this index grating is sinusoïdal and proportional to the space charge electric field the expression of which was previously derived [2].

In order to simplify the notations in the derivation of the coupled wave equations, two legitimate approximations are made. First, because this enhancement technique is only efficient to record gratings with large fringe spacings, we approximate the eigen waves of the two beams by the eigen waves of a beam propagating along the $[\bar{1}10]$ axis. Second, because the

linear birefringence of sillenite crystals is very low, the longitudinal component of the electric field is neglected. Within these two approximations and taking into account the large optical activity of our materials, the pump and probe beam electric fields $\vec{R}$ and $\vec{S}$ versus time t and coordinate z along the [$\bar{1}$10] axis can be written [4] as :

$$\begin{cases} \vec{R}(z,t) = R_+(z,t)\, e^{i k_+(t) z}\, \hat{e}_+(t) + R_-(z,t)\, e^{i k_-(t) z}\, \hat{e}_-(t) \\ \vec{S}(z,t) = S_+(z,t)\, e^{i k_+(t) z}\, \hat{e}_+(t) + S_-(z,t)\, e^{i k_-(t) z}\, \hat{e}_-(t) \end{cases} \quad (1)$$

with the eigen polarization vectors defined in the fixed $(\vec{i}, \vec{j})$ frame of the optical axes by :

$$\hat{e}_+(t) = \frac{1}{\sqrt{1 + r(t)^2}} \begin{vmatrix} 1 \\ i\, r(t) \end{vmatrix} \quad ; \quad \hat{e}_-(t) = \frac{1}{\sqrt{1 + r(t)^2}} \begin{vmatrix} r(t) \\ -i \end{vmatrix} \quad (2)$$

where r(t) is the ellipticity of the eigen waves :

$$r(t) = \left(\sqrt{\delta n(t)^2 + \Gamma^2} - \delta n(t) \right) / \Gamma \quad (3)$$

Γ is the circular birefringence and $\delta n(t)$ the linear birefringence induced by the applied electric field E(t) and related to the electrooptic coefficient r_{41} and to the refractive index n by :

$$\delta n(t) = \left(n^3 r_{41} E(t) \right) / 2 \quad (4)$$

The two eigen wave vectors $k_\pm(t)$ corresponding to the two polarization states are :

$$k_\pm(t) = \frac{2\pi}{\lambda} n_\pm(t) \quad \text{with} \quad n_\pm(t) = \left(n + \frac{\delta n(t)}{2} \right) \pm \frac{1}{2} \sqrt{\delta n(t)^2 + \Gamma^2} \quad (5)$$

To get an optimum enhancement of the two wave mixing interaction, the period T of the AC electric field E(t) is chosen to be much shorter than the characteristic time τ_g of the hologram buildup [5] . Consequently, for the steady state, the two wave mixing process can be seen as the diffraction of two time dependent beams onto a stationnary grating, this grating being previously written by the time average of the interference pattern. Within the slowly varying envelope approximation and following the usual procedure we derive the coupled wave equations [2] . Because of the optical activity the anisotropic diffraction processes are neglected. We get for R_+, R_-, S_+ and S_- :

$$\begin{cases} 2 \dfrac{\partial}{\partial z} R_\pm(z,t) = - G_\pm(t)\, S_\pm(z,t)\, m^*(z) - \alpha\, R_\pm(z,t) \\ 2 \dfrac{\partial}{\partial z} S_\pm(z,t) = G_\pm(t)\, R_\pm(z,t)\, m(z) - \alpha\, S_\pm(z,t) \end{cases} \quad (6)$$

in which the asterix denotes complex conjugation, α is the opti-

cal absorption and m(z) is the time average of the modulation index :

$$m(z) = 2\int_0^T \frac{R_+^*(z,t)\,S_+(z,t) + R_-^*(z,t)\,S_-(z,t)}{\|\vec{R}(z,t)\|^2 + \|\vec{S}(z,t)\|^2}\,\frac{dt}{T} \qquad (7)$$

Taking into account the crystal symmetry [6] it is straightforward to derive the two coupling coefficients :

$$G_-(t) = g\,\frac{r(t)^2}{1 + r(t)^2} \quad \text{and} \quad G_+(t) = g\,\frac{1}{1 + r(t)^2} \qquad (8)$$

with the new parameter : $g = \frac{\pi}{\lambda} n^3 r_{41} E_{sc}$

E_{sc} is the enhanced amplitude of the photoinduced space charge electric field the expression of which is for instance given in Ref. [2,5].

This set of coupled wave equations can be solved numerically. However, the expression for the space charge electric field we have [2] is only valid when the modulation index m(z) remains small compare to unity : e.g. when the coupling strength is not large enough to deplet the pump beam. Therefore we only need to solve system (6) in the undepleted pump beam approximation. We also assume that the two incoming beams have the same polarization state. Within this limit an analytic expression can be derived in the following way : first, taking into account the absorption, the expressions for $R_\pm(z,t)$ are obtained ; second, combining all the equations (6) together we derived a first differential equation for the modulation index which is then solved ; third, using the previous expression for m(z), the derivative equations (6) for $S_+(z,t)$ and $S_-(z,t)$ are solved ; fourthly, inserting the amplitudes $S_\pm(z,t)$ into equation (1) we get the final expression for the transmitted beam. We have :

$$\vec{S}(z,t) = e^{i\Phi(z,t) - \alpha z}\left[M_0(z,t) + \frac{e^{\eta\delta z} - 1}{\delta}\,M_1(z,t)\right]\vec{S}(z=0) \qquad (9)$$

where $\Phi(z,t)$ is the average phase : $\Phi(z,t) = (k_+(t)+k_-(t))z/2$.

δ is a constant factor depending on the incident polarization :

$$\delta = \frac{1}{\|\vec{R}(z=0)\|^2 + \|\vec{S}(z=0)\|^2}\int_0^T \frac{\|R_+(0,t)\|^2 + r(t)^2\,\|R_-(0,t)\|^2}{(1 + r(t)^2)}\,\frac{dt}{T} \qquad (10)$$

The two matrices $M_0(z,t)$ and $M_1(z,t)$ can be expressed in the fixed $(\vec{i}, \vec{j})$ frame as a function of the phase difference $\varphi(z,t)$ between the two eigen waves :

$$\varphi(z,t) = (k_+(t) - k_-(t))z/2 \qquad (11)$$

$M_0(t)$ is the usual transfer matrix [4] when there is no amplification :

$$M_0(z,t)=\begin{pmatrix} \cos\varphi+i\,\dfrac{1-r^2}{1+r^2}\sin\varphi & \dfrac{2r}{1+r^2}\sin\varphi \\ \dfrac{-2r}{1+r^2}\sin\varphi & \cos\varphi-i\,\dfrac{1-r^2}{1+r^2}\sin\varphi \end{pmatrix} \qquad (12)$$

$M_1(z,t)$ represents the amplified part of the probe beam :

$$M_1(z,t)=\frac{1}{[1+r^2]^2}\begin{pmatrix} (1+r^4)\cos\varphi+i(1-r^4)\sin\varphi & i(r^3-r)\cos\varphi+(r+r^3)\sin\varphi \\ -i(r^3-r)\cos\varphi-(r^3+r)\sin\varphi & 2\,r^2\cos\varphi \end{pmatrix} \qquad (13)$$

The expression (9) for $\vec{S}(z,t)$ is valid whatever the temporal shape of the applied electric field and whatever the incident polarization. We can note that usually, all the properties of the transmitted beam are oscillating at the electric field frequency. However using a specific experimental arrangement it is possible to make the transmitted beam time independent. Such an amplifier can thus be inserted inside a larger optical system without any disturbance. This possibility and the experimental verification of equation (9) will be discussed in more details during the lecture.

[1] P. Günter
Physics Report, Vol. 93, n° 4, pp. 199-299 (1982).
[2] S.I. Stepanov, M.P. Petrov
Opt. Commun., Vol. 53, n° 5, pp. 292-295 (1985).
[3] H. Kogelnik
Bell Syst. Tech. J., Vol. 48, n° 9, pp. 2909-2947 (1969).
[4] R.C. Jones
J. Opt. Soc. Am., Vol. 38, n°8, pp. 671-685 (1948).
[5] C. Besson, J.M.C. Jonathan, A. Villing, G. Roosen
The International Congress on Optical Science and Engineering (Hamburg 1988), SPIE Proceedings, vol. 1017, to be published.
[6] A. Yariv, P. Yeh
"Optical waves in Crystals" (Wiley, New-York, 1984).

Band-tunable multichannel scale invariant pattern recognition system with zone plates

Liang Minhua,Wu Shudong,Liu Liren,Wang Zhijiang

Shanghai Institute of Optics and Fine Mechanics
Academia Sinica

P. O. Box 8216, Shanghai, China

1. Introduction

Distortion invariant pattern recognition is wanted very much in real circumstances, such as robot vision, target tracking, pattern recognition and so on. Matched filter, also called the Vander Lugt filter, suffering from the requirement for the input object with the same scale size and orientation as the reference pattern, is hard to meet those requirements in the real circumstances. Many improvements, however, have been made on the technique of matched filtering recognition. The factor of scale is the one to which a system of matched filtering recognition is expected to immune.

We report here a multichannel scale invariant pattern recognition system which has one advantage over those previous multichannel recognition systems that have been proposed and demonstrated[1,2]: it is band-tunable and band-movable so that the recognition can be performed in the real circumstances, just like eyes. If a circular harmonic expansion term of the reference pattern is utilized,shift-, rotation- and scale-invariant recognition will be achieved.

2. A new scale invariant pattern recognition system

Such a recognition system is based on a novel Fourier transforming system which takes advantage of high dispersion of the same two zone plates to constitute an achromatic imaging system[3,4]. See figure 1. Two converging lenses with different focuses and the same two zone plates are connected in the way illustrated in Fig. 1 . A spatial coherent white-light point source is used. Putting an object between the lens L_1 and the second zone plate L_2, we will get a multi-scale Fourier spectra on the image plane of the point source via the whole optical system. The scale factor of the Fourier spectra is

$$R_1(\lambda)=\lambda[xF_1+dx\,F_1P_1(\lambda)+dF_1-dx]/(F_1-x)$$
$$=\lambda F_1[\alpha+\beta-\alpha\beta+\alpha\beta F_1P_1(\lambda)]/(1-\alpha)\ ; \qquad (1)$$

and

$$\alpha=x/F_1\ ; \qquad (2)$$

$$\beta = d/F_1 \quad ; \tag{3}$$

where λ denotes wavelength, x is the distance between the image of the source via lens L_C and the first zone plate L_{P_1}, d is the distance between the input object and the second zone plate L_{P_2}, F_1 is the focal length of lens L_1 and $P_1(\lambda)$ represents the dispersion function of dioptric power of L_{P_1}. The distance between the final image of the source and the second zone plate is

$$y = F_1 x/(F_1 - x) \quad . \tag{4}$$

From Eq.(1) we see that with α and β changed the value of the scale factor can be tuned and moved to the state we need. Achromatic imaging system can be obtained only when

$$P_1(\lambda) + P_2(\lambda) = K \quad , \tag{5}$$

where K is a constant irrelevant to wavelength. When K is equal to zero, there will be Eq.(1).

Combining two novel Fourier transforming systems together in the way illustrated in Fig. 2, we will get a band-tunable multichannel scale invariiant pattern recognition system. In the system in Fig. 2, the dioptric powers of P and P are the same while the focuses of L and L may not be the same. On the output plane in Fig. 2, the correlation output of intensity is

$$I(x,y) = I_0 R_1^2 h^*(-x/R_2, -y/R_2) * f(xR_1/R_2, yR_1/R_2) * \delta(x, y-y_0) \quad , \tag{6}$$

and

$$R_2 = \frac{-ZF_{C_2}[\alpha+\beta-\alpha\beta+\alpha\beta F_1 P_1(\lambda)]}{(Z+D-F_{C_2})(1-\alpha)\beta} \quad , \tag{7}$$

so

$$R_1/R_2 = -d(Z+D-F_{C_2})/ZF_{C_2} \quad , \tag{8}$$

where Z is the distance between the output plane and P_3, D is the distance between P_3 and L_{C_2} and F_{C_2} is the focal length of ler L_{C_2}, $h^*(*)$ denotes the conjugation of the reference pattern, $f(*)$ denotes the input object, y_0 is a dispersion term and can be eliminated by means of the technique proposed by K. Mersereau and G.M. Morris[5]. Those results are just what we want.

3. Discussion

We use zone plates to construct those systems because a zone plate has a linear and large dispersion of dioptric power. Surely, there will be two groups of dioptric power combination, which will impose a quasi-white-noise back-

ground on the output plane. Applying two dispersion lenses which have the opposite dioptric powers to each other, we can achieve an output without the noise background. There exists a possibility to construct a large dispersion thick lens by combining two thin lenses with different refractive indices.

References

1. A. Vander Lugt, Proc. IEEE 54, 1055(1966).
2. J.D. Gaskill, Linear System, Fourier transforms, and Fourier Optics, p.410(Jhon Wiley&Sons, New York, 1978).
3. Liang Minhua et el, submitted to Opt. Comm. .
4. Liang Minhua et al, submitted to J. Mod. Opt. .
5. K. Mersereau and G.M. Morris, Appl. Opt. 25, 2338(1986).

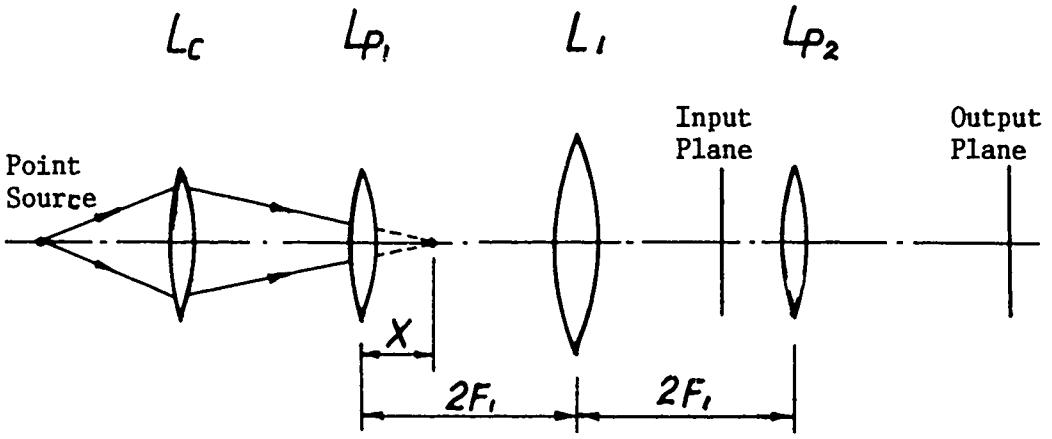


Figure 1 Scale range tunable Fourier transforming system.

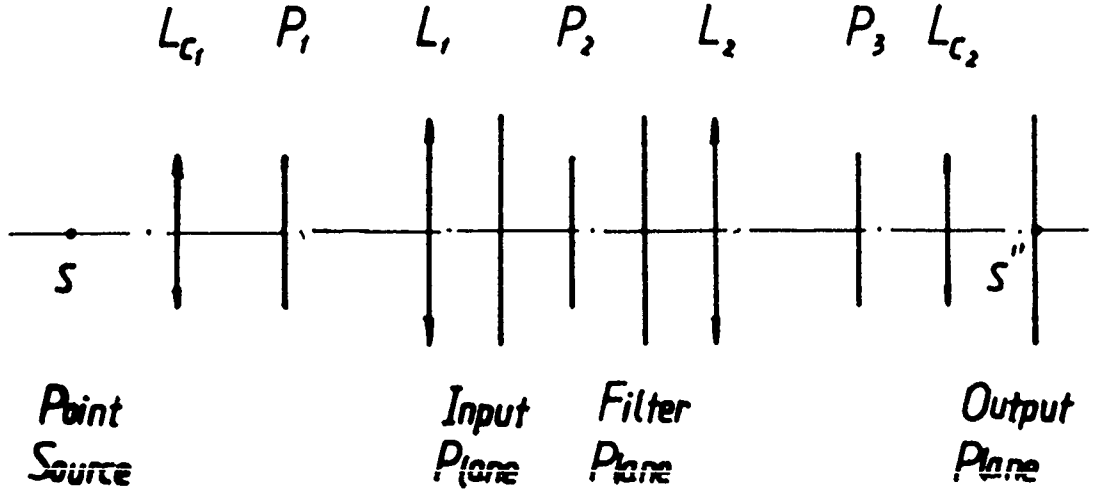


Figure 2 Multichannel scale invariant pattern recognition system with zone plates. P_1, P_2, P_3 represent three zone plaes, respectively.

Hybrid Optical Processing for Measuring the Refractive Index Profile in Single-Mode Fibers

T.Nobuyoshi

Department of Electronics, Okayama-Rika University
Ridai-cho 1-1, Okayama-shi, 700 Japan

1. INTRODUCTION

In estimation of propagation characteristics for single-mode fibers, the refractive index profile plays an important role. Several measurement methods such as the exit radiation pattern method (ERP)[1] and the near field intensity method (NFI)[2] have been recently developed with a relatively high spatial resolution. Both methods, however, require complicated features: cumbersome computation, with inverse Hankel transformation and correction of dynamic range in the case of ERP; numerical derivations and smoothing for noise reduction in the case of NFI. In this study, we present a new processing technique for calculation of the refractive index profile in single-mode fibers. This technique makes use of hybrid optical processing with a spatial filter (a radially inhomogeneous transparent filter) and two lenses, creating the Hankel transformation and the inverse transformation. This method allows real time monitoring of the profile and permits significantly less computational time as compared to that of conventional computer methods.

2. THEORY

2.1 Near-Field Pattern and Refractive Index Profile

For a single-mode fiber, the processing procedure is based on the fact that the near-field pattern R(r) fulfills the following scalar wave equation:

$$(\Delta+(k^2n^2(r)-\beta^2))R(r)=0 \qquad (1)$$

where k and β are the free space wave number and the longitudinal propagation constant, respectively, and Δ is the Laplacian operator in the cylindrical coordinate. From modification of eq.(1), the refractive index profile n(r) is represented by

$$n^2(r)=(\beta^2-\Delta R(r)/R(r))/k^2 \qquad (2)$$

2.2 Far-Field Pattern by a Convex Lens

Recently it has been shown [1] that the far field pattern (FFP) F(ρ) (ρ-spatial frequency) of the single-mode fiber has a simple relationship to the NFP through Kirchhoff's diffraction integral theorem. If the NFP is axially symmetric, then the FFP is proportional to the Hankel transform H[] of R(r), that is, $F(\rho) \propto H[R(r)]$. In Fourier optical processing [3][4], we get the FFP on the back focal plane of a convex lens as the image of the Hankel transform of R(r) on the fiber end surface located on the front focal plane. Conversely, the FFP on the front focal plane is converted to R(r) on the back focal plane using the inverse Hankel transformation $H^{-1}[]$. The back focal plane is known as the spatial frequency plane 'ρ-plane). At the ρ-plane, it is satisfied that $\rho = fk\tan\theta$, where f and θ denote the focal length of the lens and the exit radiation angle of the fiber, respectively.

2.3 Optical Hybrid Processing by Spatial Filtering

In the cylindrical coordinate the Hankel transform of ΔR(r) may be expressed in the form [3]

$$H[\Delta R(r)] = \rho^2 F(\rho) \tag{3}$$

This leads to

$$\Delta R(r) = H^{-1}[\rho^2 F(\rho)] \tag{4}$$

Substituting eq.(4) into eq.(2),

$$n^2 = (\beta^2 - H^{-1}[\rho^2 F]/H^{-1}[F])/k^2 \tag{5}$$

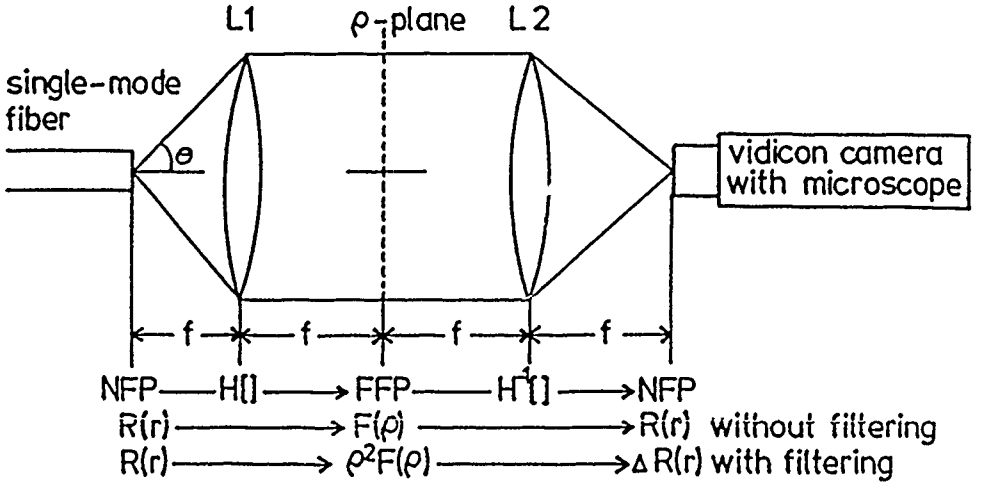


Fig.1 Optical processor for calculating the refractive index profile in single-mode fibers

Using the Hankel transform expression for F, eq.(5) becomes

$$n^2=(\beta^2-H^{-1}[\rho^2H[R]]/H^{-1}[H[R]])/k^2 \qquad (6)$$

The second term on the right side of eq.(6) represents the Hankel transformation and the inverse transformation of R, consecutively. As shown in Fig.1, these operations are achieved by the typical 2 lens optical processing system. The first lens L1 performs the Hankel transformation and the second lens L2 the inverse transformation. In the calculation of the numerator, $H^{-1}[\rho^2H[R]]$, we use the spatial filter whose transparency changes parabolically to the radial direction at the ρ-plane. Sampling two images, $H^{-1}[H[R]]$ and $H^{-1}[\rho^2H[R]]$, the refractive index profile is obtained from the ratio of the two images. $H^{-1}[H[]]$ means no transformation; therefore, it may be possible to measure R directly instead of through a process $H^{-1}[H[R]]$.

3. OPTICAL MEASUREMENT SYSTEM

Fig.2 shows our experimental design. L1 and L2 are the convex lenses with a 35mm diameter. As an image sampler we used a vidicon camera system (C1000:Hamamatsu Photonics, Inc.) with a 2 field input microscope. At the same time, two images, $H^{-1}[\rho^2H[R]]$ and $H^{-1}[H[R]]$, were treated as 8-bit sampling data by the A/D converter which followed this image sampling system. After calculation of the ratio of these two images and correction of the term β^2/k^2 in eq.(6) by the microprocessor unit(MPU), the refractive index profile was plotted out. We measured the step-index single-mode fiber fabricated by the rod-in-tube method (3μm: core radius) and used an He-Ne laser (λ=633nm) as an optical source.

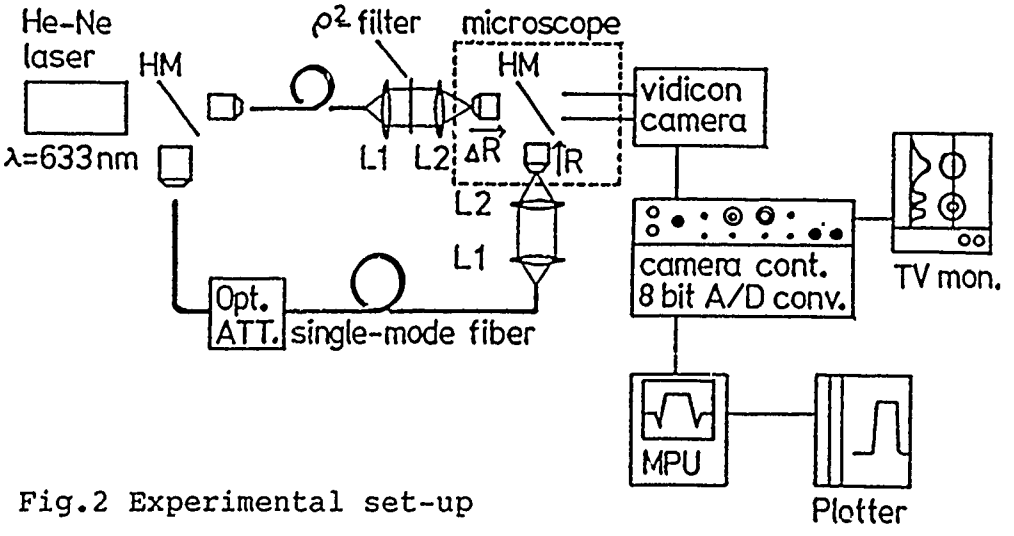


Fig.2 Experimental set-up

4. RESULTS and DISCUSSION

In Fig.3 the solid line shows the measured refractive index profile and the dot-dash line depicts a numerical simulation of eq.(5). By the limitation of numerical aperture (NA) in the transformation lenses, the integral region of eq.(5) is limited to under 27° which is the maximum radiation angle The reason for differences in the two lines may be due to error of parabolic characteristics in the spatial filter and/or astigmatism of the lenses. We used the ND filter; its transparency was approximated by a broken line in 4 sections instead of parabolically. The MPU was used only in the operation of rooting and dividing two images. These operations are readily changed by unprogrammed analog IC and it is possible to monitor the refractive index profile in real time.

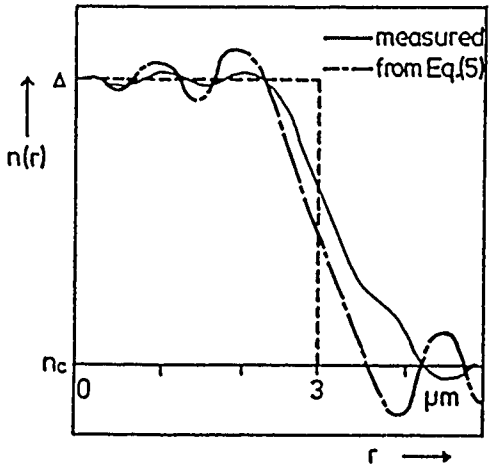


Fig.3 Simulated and measured refractive index profile of a step-index single-mode fiber

5. CONCLUSION

In conclusion, we have demonstrated that by making use of hybrid optical processing, convienient refractive index profile measurements are feasible in single-mode fibers. In view of various improvements that should enhance the accuracy of parabolic transparency in spatial filter and that should allow fabrication a single, large NA lens with a large diameter, this simple technique seems very promising, and could soon become a real time monitoring technique for the refractive index in single-mode fibers.

References:

1. K.Hotate and T.Okoshi: Appl.Opt., vol.18, No.19,p163, 1979
2. G.Coppa et al.: Electron.Lett., vol19, No.8, p293, 1983
3. K.Iizuka: *Engineering Optics*, Ch.1,5, (Springer-Verlag 1983)
4. J.D.Gaskill: *Linear Systems, Fourier Transforms, and Optics*, p323 (John Willy & Sons 1978)

On the Hardware Requirement for 2-D Image Convolution in Electro-Optical Systems[1]

M. Mary Eshaghian, D. K. Panda and V. K. Prasanna Kumar

SAL 344, Department of EE-Systems
University of Southern California
Los Angeles, CA 90089-0781, USA

1 Introduction

The computational throughput of a parallel architecture is largely dependent on it's communication bandwidth. The communication bandwidth available in the present day parallel systems is reaching saturation due to the inherent limitations of transmitting electronic signals. To overcome this problem, researchers have considered contention-free optical beams as an efficient means of interconnection [GLKA84,JS87]. Unlike the electronic signal, the optical interconnection offers dual advantages of larger bandwidth and fan-out. With the availability of such interconnection, the field of Optical computing is being diversified from the development of analog optical processors [CRFH81] to digital optical and electro-optical computers [Hua84,EK88]. Such electro-optical systems, capable of exploiting the speed and parallelism of optical systems together with the programmability and accuracy of electronic computers, promise tremendous computational power.

In this paper, we consider an electro-optical system from a computational perspective and study some inherent limitations of such a system in parallel computation. As an example, we study the electro-optical resource requirements for solving a fundamental, computationally intensive operation such as 2-D image convolution. The 2-D image convolution is extensively used in signal and image processing. A lower bound on the storage(memory) requirement of an electro-optical chip to solve a problem reflects the hardware requirement for fabricating such a system. We present a lower bound of $\Omega(nw)$ on the volume requirement of an electro-optical chip for computing image convolution. Irrespective of the I/O scheme and the order of computation, we show that any image convolution design must satisfy this bound for convolving a $w \times w$ kernel with a $n \times n$ image, as long as the input bits are given to the system once only. Most of the VLSI designs for 2-D image convolution use the input image only once. All these designs satisfy this bound.

The rest of the paper is organized as follows. In the next section, we describe an optical model of computation, the relationship between the minimum volume requirement in this model and information transfer. In section 3, we derive a lower bound on the information transfer for image convolution under several input formats used in practice. In the last section, we compare the memory requirements of the known VLSI designs for image convolution.

2 An Optical Model and Information Transfer

In this section, we define an Optical model of computation which is an abstraction of currently implementable optical and electro-optical computers [EK88]. Similar to the VLSI model of computation, this model can be used to understand the limits on computational efficiency in using optical technology. We show that minimum volume requirement of an optical model of computation is same as the minimum VLSI area in the VLSI model. Using information transfer argument, we also show a methodology to determine the minimum volume requirement of an electro-optical system for solving a problem.

[1]This research was supported in part by the National Science Foundation under grant IRI-8710836.

[2]A function $f(n)$ is said to be $\Omega(g(n))$ if there exist positive constants c and n_0 such that for all $n > n_0$, $f(n) \geq c \cdot g(n)$. A function $f(n)$ is said to be $O(g(n))$ if there exist positive constants c and n_0 such that $f(n) \leq c \cdot g(n)$ for all $n \geq n_0$.

2.1 An Optical Model

An optical model [EK88] is shown in figure 1. More formally this model is defined as follows:

Definition 1 *An optical model of computation represents a network of processors each associated with a deflecting unit and a receiving unit capable of establishing direct optical connection to another processor or a set of processors.*

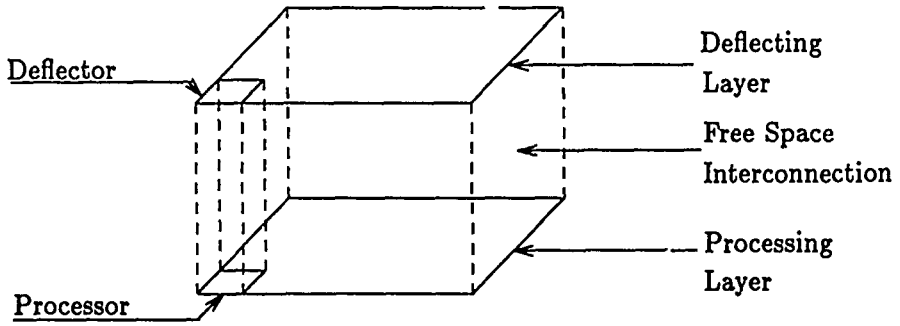


Figure 1: The optical model of computation

We make the following assumptions to capture the real life optical designs:

(a) The processing layer consists of processors and memory elements. In one unit of time, a processor can compute a simple arithmetic/logic operation and a deflector can redirect an incident beam.

(b) The intercommunication is done through free-space optical beams. An optical beam carries a constant amount of information in one unit of time, independent of the distance to be covered. We make this assumption for deriving lower bounds.

(c) I/O is performed at I/O pads. Each I/O pad occupies one unit of volume. Similarly, one bit of memory contributes to at least one unit of volume. The volume occupied by processors, deflecting elements, memory, and I/O pads together determine the total volume of the system.

(d) The time, T for computation is the time between the arrival of the first input to the departure of the last output.

(e) The input and output are performed according to a pre-determined sequence of time instants at pre-specified locations which depends entirely on the circuit design, not on the data.

(f) Each input bit is read by the chip exactly once.

2.2 Optical Volume, VLSI Area and 1-way Information Transfer

The minimum VLSI area requirement for computing a problem is related to the *lower bound* on the *1-way information transfer* [Kum83,Yao79]. The following abstract setting has been shown to be useful in estimating the minimum VLSI chip area [Kum83,Yao79]. Two sets of processors $P1$ and $P2$ each receive $\frac{n}{2}$ bits of an n input function f to be computed. The input partition can be denoted by (π_1^i, π_2^i) where $\pi_1^i(\pi_2^i)$ are the inputs known to $P1(P2)$. The minimum *information transfer* from $P1$ to $P2$ to compute f over all possible input partitions is denoted by $I_1(f)$. The area requirement A of any VLSI chip computing f is $\Omega(I_1(f))$ [Yao79,Kum83].

The above 1-way protocol for single output function can be easily extended to *multiple output functions* by introducing a suitable output partition (π_1^o, π_2^o) over the set of output functions $F = \{f_1, f_2, \ldots, f_l\}$. Both processors $P1$ and $P2$ are allowed to compute the output functions belonging to their respective subsets based on the input bits available to them. Since the 1-way communication link is from $P1$ to $P2$, any requirement for tranfering data from $P2$ to $P1$ to

compute a function belonging to π_1^o leads to an *infeasible* output partition. Hence, $I_1(F)$, the 1-way complexity of computing a set of output functions F is defined as follows:

$$I_1(F) = \begin{array}{c} Min \\ feasible\ output \\ partition \end{array} \left\{ \begin{array}{c} Min \\ input\ partition \end{array} \left\{ \begin{array}{c} worst\ case \\ information\ transfer \\ from\ P1\ to\ P2 \end{array} \right\} \right\}$$

It is interesting to observe that this 1-way information transfer $I_1(F)$, is related to the optical volume requirement of an electro-optical system in solving a problem. In the full version of the paper [EPK88], we show:

Theorem 1 *The volume V_o of any electro-optical system computing F satisfies $V_o = \Omega(I_1(F))$.*

The *computational rectangle* corresponding to the input partition for single output function can be extended to the third dimension resulting a *computational parallelepiped P* which represents the computation of F over an input and output partition. For a fixed value of input bits, the output functions in set F are represented by a vector of length l in the third dimension of P. Based on the concept of *distinct planes* in P, the information transfer, $I_1(F)$, can be estimated [EPK88] in a similar spirit as in $I_1(f)$ [Y 79]. This leads to:

Proposition 1 *For a fixed input partition (π_1^i, π_2^i) and a fixed output partition (π_1^o, π_2^o) , the minimum number of bits of information transfer from P1 to P2 to compute F is* log $d(F)$, *where $d(F)$ is the number of distinct planes in P. Also, the 1-way complexity $I_1(F)$ is equal to* log d, *where d is the minimum $d(F)$ over all possible input and feasible output partitions.*

3 Optical Volume for Computing Image Convolution

In this section, we derive a lower bound on the information transfer, $I_1(F)$, for convolving a $w \times w$ kernel with an $n \times n$ image using the technique of last section. These bounds are translated to lower bounds on optical volume for computing image convolution.

3.1 Lower Bounds on Optical Volume

Let the pixels of the input image be arbitrarily colored as R(Red) or B(Blue) such that equal number of R and B pixels exist. Figure 2 (a) demonstrates one such coloring. The $\frac{n^2}{2}$ R(B) pixels correspond to $\pi_1^i(\pi_2^i)$. This arbitrary pixel-coloring can be reduced to an instance of arbitrary row(column)-coloring of the image by defining the color of a row(column) to be the majority of the color of the pixels available in that row(column). Thus, a row or column is R(B) it at least $(\frac{n}{2}+1)$ pixels in that row or column are R(B). A set of w consecutive rows starting at row(x), $1 \le x \le n-w+1$ can be defined as a window $wd(x)$. Irrespective of the arbitrary pixel-coloring, the following claim can be verified by combinatorial analysis [EPK88].

Claim 1 *For $n \ge w(\frac{w}{2}+1)$, there exists a window $wd(x)$, $1 \le x \le n-2w+2$, consisting of $R(B)$ rows with indices $i_1, i_2, \ldots, i_l$, $l \ge \frac{w}{20}$, and an integer k, $1 \le k \le w-1$, such that the rows with indices $i_j + k$, $1 \le j \le l$, are $B(R)$.*

The above window WD can be used to compute a lower bound on volume requirement. With the value of k as defined in claim 1, choose a kernel K as shown in figure 2 (b). With the given kernel K, the computation of at least $(\frac{n}{2}-w+1)$ bits in each of the rows with indices $i_1, i_2, \ldots, i_l$ belongs to π_2^o. Let the set X_1 represents these identified bits. According to the above claim, $|X_1| \ge \frac{w}{20}(\frac{n}{2}-w+1)$.

The number of *distinct planes* in the computational parallelepiped P is at least equal to the number of distinct rows in any vertical plane. Consider the vertical plane corresponding to value 0 for all input bits in π_2^i. The output of the convolution operation for bits in X_1 are identical

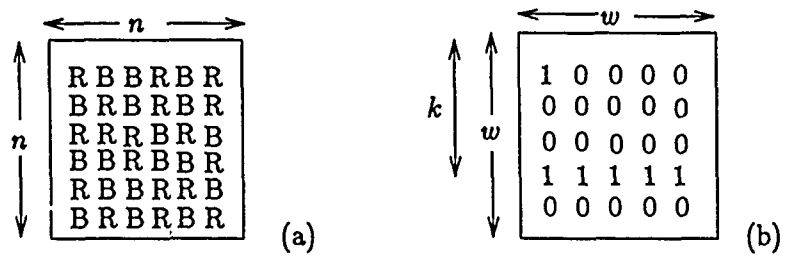


Figure 2: (a)Arbitrary coloring of pixels. (b) A kernel for general case

to the respective input values. These bits in X_1 can take $2^{|X_1|}$ distinct values, resulting in $2^{|X_1|}$ distinct rows in the vertical plane under consideration. This leads to, $d \geq 2^{|X_1|}$. By proposition 1, $I_1(F) \geq |X_1|$ or $I_1(F) \geq \frac{w}{20}(\frac{n}{2} - w + 1)$. Applying theorem 1, we get $V_o = \Omega(nw)$. This leads to the following result:

Theorem 2 *The volume V_o of any electro-optical design for convolving a $w \times w$ kernel with a $n \times n$ image satisfies $V_o = \Omega(nw)$.*

4 Comparisons and Conclusions

In this paper, we considered an electro-optical system from a computational perspective. As an example, we studied the electro-optical resource requirements of a computationally intensive operation such as 2-D image convolution. We showed that any electro-optical system, regardless of implementation, must have $\Omega(nw)$ volume for convolving a $w \times w$ kernel with a $n \times n$ image, as long as the input pixels are given to the system only once. This lower bound on the volume requirement of an electro-optical system is also same as the minimum VLSI area requirement of a VLSI chip to carry out such a computation. Our scheme can be used to study the hardware requirement for solving several other problems in signal and image processing using electro-optical systems.

References

[CRFH81] H. J. Caulfield, W. T. Rhodes, M. J. Foster, and S. Horvitz. Optical implementation of systolic array processors. *Optical Communication*, 40:86–90, 1981.

[EK88] M. M. Eshaghian and V. K. Prasanna Kumar. Massively parallel architectures with optical interconnection network. In *Topical Meeting on Optical Computing*, Toulon,France, 1988.

[EPK88] M. M. Eshaghian, D. K. Panda, and V. K. Prasanna Kumar. *On the Hardware requirement for 2-D Image Convolution in Electro-Optical Systems.* Technical Report 129, Signal and Image Processing Institute, University of Southern California, Aug 1988.

[GLKA84] J. W. Goodman, F. J. Leonberger, S. Y. Kung, and R. A. Athale. Optical interconnections for VLSI systems. In *Proc. IEEE*, pages 850–866, 1984.

[Hua84] A. Huang. Architectural considerations involved in the design of an optical digital computer. In *Proc. IEEE*, pages 780–786, 1984.

[JS87] H I Jeon and A. A. Sawchuk. Optical crossbar interconnections using variable grating mode devices. *Applied Optics*, 26:261, Jan 1987.

[Kum83] V. K. Prasanna Kumar. *Communication Complexity of various VLSI Models.* PhD thesis, Pennsylvania State University, 1983.

[Yao79] A. C. Yao. Some complexity questions related to distributed computing. In *Proc. 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979. Atlanta.

Optical Systems Tolerances for Symmetric Self Electro-optic effect Devices in Optical Computers

Nick C. Craft
Heriot-Watt University
Riccarton
Edinburgh, EH14 4AS
Scotland

Michael E. Prise
AT&T Bell Laboratories
Crawfords Corner Road
Holmdel, NJ 07739

Summary

Introduction The s-SEED as a logic gate has been demonstrated [1] and we are building optical systems using them [2] to implement the architectures proposed by Murdocca [3]. Here we investigate the systems requirements we must meet in order to do this. We show that the s-SEED is a significantly easier optical logic device with which to build optical digital computers, than previously demonstrated devices. We demonstrate two methods of fully utilizing its advantages in the systems we are building. It is also the purpose of this paper to stimulate the design of other devices which will have the advantages of s-SEEDs.

Device Operation We are attempting to use the s-SEED's as 2 input NOR gates as shown in Figure 1.

In the diagram the output of the SEED is shown as the transmitted power supply whereas in an actual device the output is the reflected power supply. The stages of the temporal cycle in order are

1. The preset beam : This is only incident on modulator R. The purpose of this cycle is to set the device in the LO state. $\lambda = 780nm$ is a convenient wavelength since good semiconductor lasers are available at this wavelength and this beam can be brought in to the system using a dichroic beam-splitter.
2. Two signal beams : These are the outputs from the two previous s-SEEDs - the logic signals. They are are the reflected power supply beams from these devices. Notice that only half this power is available since the other half is necessary to drive another logic gate. It is during this cycle that the logic operation occurs.
3. Power supply : This consists of two equal beams one incident upon each modulator. During this cycle the logic state of this s-SEED device is transferred to the input of the next devices.

We have also introduced attenuators on the outputs of the R modulators. The transmission of these attenuators is denoted by α and we later show that by selecting the appropriate value of α the operating tolerances of the system can be maximised. These attenuators can be interlaced with the patterned reflectors described in [2]. In this mode of operation the upper branch of the hysterisis loop is unused. We therefore describe the relevant optical characteristics of the device as shown in Figure 2.

This is the characteristic of a general two input differential logic device which switches from HI to LO at a differential power ratio of S and has a contrast $C_{SW} = \frac{R_{HI}}{R_{LO}}$. For clarity the outputs have been shown to be transmitted rather than reflected.

Here we discuss only use of the device as a NOR gate. It can be shown that for device operation,

$$C_{SW} > \frac{\alpha}{S} > 1 \qquad (1)$$

The relationship 1 is shown in Figure 3. There is no dependence on P_{SUPPLY} therefore, only the ratio of the power supply beams on the two input modulators which determines the state of the gates. No holding beam is necessary as is the case for conventional bistable devices [4] so, for architectures based on sequential arrays of logic devices , there is no relationship necessary between the power supplies for different device arrays. *Critical biasing* which has made building anything with conventional bistable devices very difficult [4] is avoided.

Notice that the quantity $\frac{\alpha}{S}$ is a critical parameter. In the case of the s-SEED device S is a device parameter and α can be chosen ($\alpha > 1$ can be implemented by putting attenuators $\frac{1}{\alpha}$ in front of the S- modulators). Looking at the figure one can see that the devices can be made to operate with any contrast at all and for given C_{SW} , α

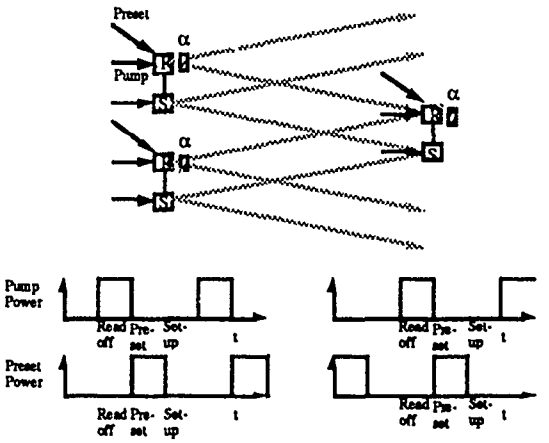


Figure 1: Operation of an s-SEED as a two input NOR gate using the preset method.

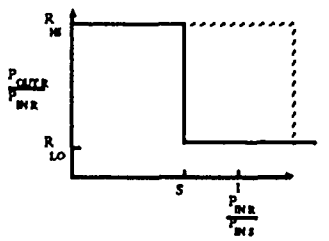


Figure 2: Generalized differential logic device characteristics (dotted line shows upper branch of the hysterisis loop for s-SEED only).

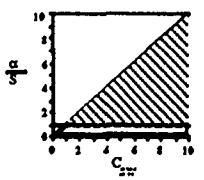


Figure 3: Relationship between C_{SW} and $\frac{\alpha}{S}$ necessary for operation of an s-SEED as a two input NOR gate and the shaded area shows the range of allowable parameters.

can be selected to allow for the maximum range of S so allowing operation of an array of devices which is not completely uniform.

Array Generation It is important to consider the behaviour of arrays of these devices when there is variation in the power supply beams to each modulator in the array. This is a problem since one of the difficulties we have is generating a completly uniform array of power supply beams [6]. Here we introduce a special method of array generation which is suitable for such differential logic devices.

A splitter and an array generator : We use an element to split the beam from the power supply laser in two, such that one beam powers the R modulator of a single s-SEED and the other beam powers the S modulator of the same device (the *splitter*). These two beams are then passed through another array generator which distributes them to all the devices. In our experiments so far we have used a binary phase grating consisting of equally spaced strips as the two beam *splitter*. The required period of the grating p is determined by the spacing between the two modulators s and the focal length of the input lens F and the wavelength λ, by

$$d = \frac{2F\lambda}{s} \tag{2}$$

So the strip width would be $\frac{d}{2}$. The strips have a phase depth of $\frac{\pi}{2}$. 86% of the power is in the correct orders ($+1$ and -1) and the other orders of the grating can be filtered out (the 0 order has no power in it). It may be possible to use a holgrahic grating her to reduce the power loss. This splitting can be done very accurately since the fabrication involved is much less complex, than the main array generator. In fact experimentally we have manufactured binary phase grating splitters with a period of $512\mu m$ which produce two beams with less than 1% difference in power.

The second array generator in our initial experiments is a Dammann grating with a period in one direction twice that in the orthogonal direction. This will the distribute the two beams from the splitter to all the devices. This element is in general harder to fabricate than the splitter since it is more complex and in the case of binary phase gratings its uniformity will be much more sensitive to fabrication limits [6].

If the splitter produces two beams with a power variation of $2\Delta P$ and the second array generator is perfect, then for the devices to operate as NOR gates

$$\frac{1+\frac{\Delta P}{P}}{1-\frac{\Delta P}{P}} < \frac{\alpha}{S} < C_{SW}\frac{1-\frac{\Delta P}{P}}{1+\frac{\Delta P}{P}} \tag{3}$$

This is the same result as if one array generation element was used and it produced an array of beams with a spread of $2\Delta P$ in power.

If the splitter is perfect and the second array generator produces beams with a spread of $2\delta P$ the for proper device operation

$$\frac{(1+\frac{\delta P}{P})C_{SW}+(1-\frac{\delta P}{P})}{(1+\frac{\delta P}{P})+(1-\frac{\delta P}{P})C_{SW}} < \frac{\alpha}{S} < C_{SW} \tag{4}$$

It can be seen that the constraints 4 and 3 on $\frac{\alpha}{S}$ for a given $\frac{\delta P}{P}$ are less rigorous than those imposed by 3 on $\frac{\alpha}{S}$ for the same $\frac{\Delta P}{P}$. To examine this further we plotted these constraints for different values of C_{SW} Figure 4.

Conclusions Looking at the figures we can see that the splitter + array generator method is much more tolerant of non-uniform array generation particularly at lower contrast, since we can make an accuarate splitter. We can also see that as the range in power supplies (the error) approaches its maximum allowable value the allowable range of $\frac{\alpha}{S}$ becomes much smaller. As one would expect, at higher contrasts the system tolerances become much better.

To set a system up with devices of a given contrast we draw two graphs showing the constraints on $\frac{\delta P}{P}$ and $\frac{\Delta P}{P}$ such as in Figure 4. We then look at our array generator errors and find $\frac{\Delta P}{P}$ and $\frac{\delta P}{P}$. This will give us a range of $\frac{\alpha}{S}$ which is accepatable. We can then pick α so we are in the middle of the range. Then we can look at the allowable range in switching ratio S between the upper and lower constraints and see if our devices can match this. We may optimise for α being as close to 1 as possible. This may be more energy efficient/faster. We may of course find that the system simply does not work and we have to improve the array generation and/or make better devices.

Any variations in the coupling into different devices on the array can be thought of as a variation in α and can also be examined using the graphs. The introduction of the attenuator α allows another degree of freedom in

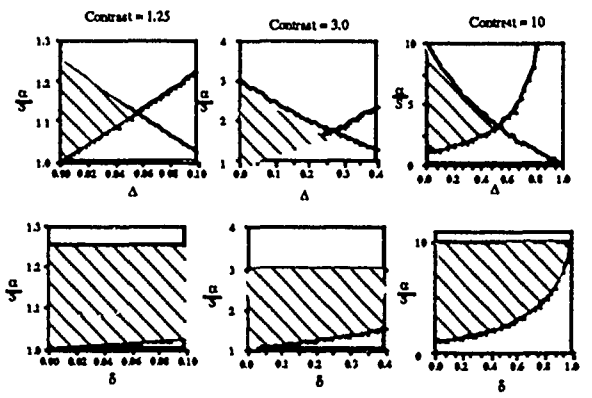


Figure 4: Allowable range of $\frac{\alpha}{S}$ for a given $\frac{\Delta P}{P}$ and $\frac{\delta P}{P}$.

the case of the s-SEED device and may be essential for some other differential devices. It allows us to take full advantage of the differential nature of the device.

We believe that is a very useful way to examine the interelations between the device and system parameters involved in making a digital optical system using any form of differential thresholding logic device and will prove invaluable in the study of issues such as variations in device characteristics with speed and temperature. As a footnote we found that setting up the constraint equations on a spread sheet program (in our case Microsoft Excel) proved a very useful way of looking at all the dependencies.

It is evident that further work is necessary on the dynamics of the devices. The switching speed of devices as a function of switching ratio is important since it may mean we are required to *overswitch* the devices. This could be included in our treatment of system tolerances. From the point of view of device characteristics we can say that higher contrast would make building systems easier (but not at the expense of all the other desirable charateristics such as low energy, high speed etc.). At the conference we will discuss where the systems we are building fall within these constraints.

References

[1] A. L. Lentine, D. A. B. Miller, J. E. Henry and J. E. Cunningham, "Photonic Ring Counter and Differential Logic Gate Using The Symmetric Self Electro-optic Effect Device",*Conference on Lasers and Electro-optics 1988*, paper TUE4

[2] M. E. Prise, M. M. Downs, F. B. McCormick, S. J. Walker and N. Streibl, " Design of an Optical Digital Computer", in *Optical Bistability IV* edited by W. J. Firth et al., published by Les Editions de Physique 1988

[3] M. J. Murdocca , N. Streibl, A. Huang and J. Jahns, "Algorithmic Design Technique for a Free- Space Regularly Interconnected Digital Optical Computer", *Applied Optics*, 27, (9), 1651

[4] S. D. Smith, A. C. Walker, B. S. Wherrett, F. A. P. Tooley, N. C. Craft, G. Mathew, M. R. Tahizadeh, I. Redmond, R. G. Campbell, "Restoring Optical Logic- The Demonstration of Extensible All-optical Digital Systems",*Optical Engineering*, 26(1), 41, 1987

[5] N. Streibl, "Generation of an Array of light spots out of a laser beam", unpublished work

[6] J. Jahns, M.E. Prise, M.M. Downs, S.J. Walker, N. Streibl, "Dammann gratings as array generators", Optical Society of America Annual Meeting, Rochester 1987, paper WJ3.

Tolerance Analysis and Design of Optical Processors

J.F. Snowdon and B.S. Wherrett

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

In recent years the literature in optical computing has grown vastly and diversified widely. Many architectures have been proposed and many different devices fabricated, but remarkably few optical processors have been constructed. One reason for this is the lack of a close link between architectural considerations and device fabrication considerations. A tolerance design strategy can provide this link and enable comparative evaluation of different systems. In this paper a strategy is presented for the tolerance design of prototype optical computing circuits containing programmable logic arrays, full adders and threshold elements [1,2,3].

Tolerance analysis and device optimisation have been considered [4,5,6], but only in terms of simple single gate networks or specific device parameters. This paper presents a tolerance design strategy which takes into account the interaction of each element within the entire processor. The methodology can be generalised to include any type of optical device and is sufficiently extensible to enable complete simulation of any particular optical processor. This paper focusses on the use of such a strategy to design circuits from nonlinear and/or bistable interference filter elements to perform processing operations of the cellular logic image processing (CLIP) type [7].

Table 1 shows the set of parameters that a circuit design must include. Each class of parameters is given a region of acceptance (e.g. Figure 1) by

the operational demands of the processor, and a region of tolerance by the fabrication limits of the devices. A technique known as design centering [8] is used to place the tolerance region inside the region of acceptance thus establishing optimal nominal values of the parameters for the construction of working circuits. A dynamic model of device behaviour is used to give the dependence of switching times on other parameters. As we are interested in prototype construction, a 'worst case' algorithm is employed, ensuring the design is completely centred and no portion of the tolerance region protrudes from the region of acceptance (corresponding to 100% yield in VLSI design terminology) [9,10].

Figure 2 shows a block diagram of the design strategy, demonstrating how the interactions between each design level contribute constraints and sensitivity information to the adjoining levels. This chain of design levels enables greater generality and provides for the inclusion of experimental and simulation results at each level.

In this paper we shall explain this methodology and apply it to optical processors such as those described in [1,2,7]. The simulations generated by the design strategy provide precise information at both system and component levels necessary to the construction of such processors.

References

[1] Wherrett, B.S., Appl. Opt., 24, 2876 (1985).

[2] Wherrett, B.S., Opt. Commun., 56, 87 (1985).

[3] Lewis, P.M., Coates, C.L., Threshold Logic, J. Wiley & Sons Inc. (1967).

[4] Prise, M.E., Streibl, N., Downs, M.M., Opt. & Quantum Electron., 20, 49 (1988).

[5] Wheatley, P., Ph.D. Thesis, University College London (1987).

[6] Wherrett, B.S., Hutchings, D.C., Russell, D., J. Opt. Soc. Am. B, 3, 351 (1986).

[7] Wherrett, B.S., Snowdon, J.F., SPIE, Vol. 963, in press.

[8] Wehrham, E., Spence, R., IEEE Proc. Int. Sym. Cts. Sys., 1424 (1984).

[9] Spence, R., Soin, S.S., Tolerance Design of Electronic Circuits, Addison-Wesley (1988).

[10] Lightner, M.R., IEEE Proc. 75, No. 6, 786 (1987).

Table 1. Categorisation of Parameters

Processor Requirements	Adjustable Beam Parameters	Optical characteristics
f_i fan in	P_H holding power	P_{SU} switch up power
f_t threshold	P_B bias power	P_{SD} switch down power
f_o fan out	L inter gate losses	P_F minimum signal power
T clock time		P_N maximum signal power
N number of gates		P_{HO} critical holding power

Physical Parameters	Operational Conditions
R_f front reflectivity	t switch time
R_B back reflectivity	ΔP_H
αD cavity absorption	ΔP_{SU} tolerance
ϕ_o cavity detuning	ΔR_F
M nonlinear coefficient	ΔR_B, etc.
a off-axis absorption	
τ characteristic time	

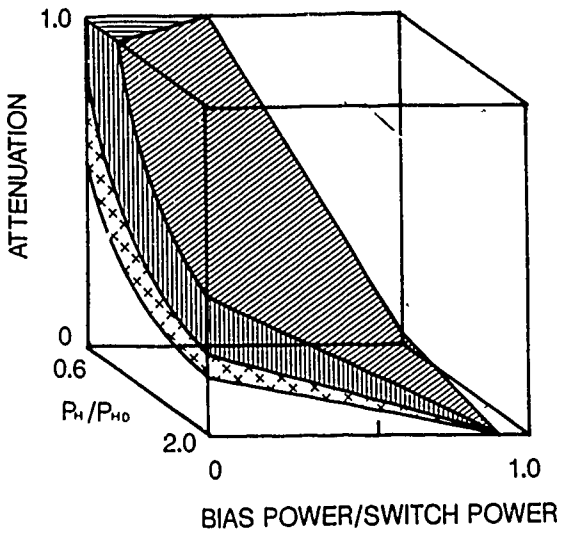


Figure 1. The region of acceptance of an idealised off-axis device with regions of OR-gate (shaded) and AND-gate (cross-hatched) operation. This region is created with respect to the 'adjustable beam' class of parameters [7].

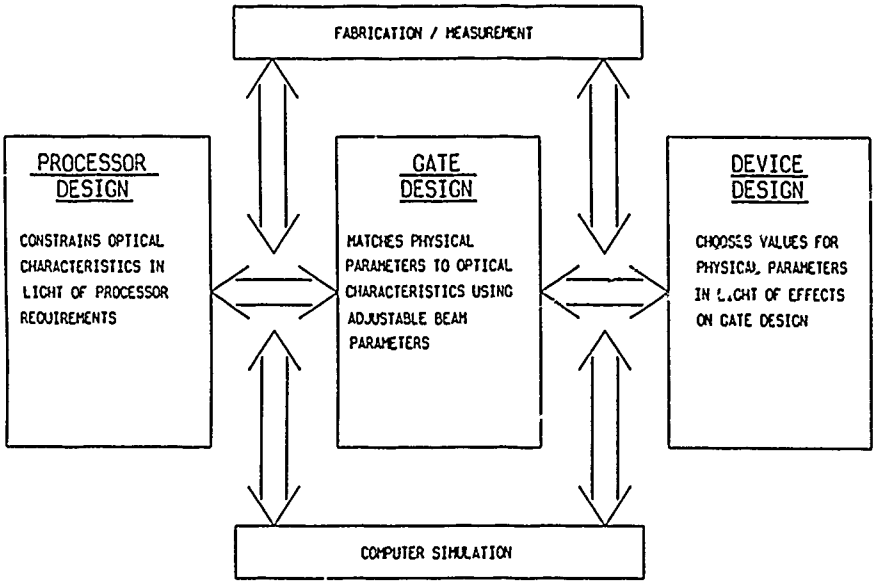


Figure 2. Block diagram of the information exchange process. The bi-directional arrows indicate the passing of design constraints and sensitivity information between each level.

OPTIMIZATION OF BINARY PHASE ONLY FILTER WITH SIMULATED ANNEALING ALGORITHM

Myung Soo Kim , Michael R. Feldman, and Clark C. Guest
Department of Electrical and Computer Engineering,
University of California San Diego , La Jolla , CA 92093

1.Introduction

Binary phase only filters (BPOF) show excellent performance for matched filtering. However, when the BPOF is synthesized to distinguish similar patterns, existing encoding methods do not produce good results because the binary state of each pixel on the filter is not optimally encoded. In this paper we examine an optimum encoding of the BPOF for matched filtering to distinguish similar patterns, using a simulated annealing algorithm [1].

2. Encoding of the BPOF

In the BPOF, two encoding methods have been used [2],[3]:

$$\hat{F}(u,v) = 1 \text{ if } Re[F(u,v)] > 0 ; \hat{F}(u,v) = -1 \text{ otherwise;} \qquad (1)$$

or $\hat{F}(u,v) = 1$ if $Im[F(u,v)] > 0$; $\hat{F}(u,v) = -1$ otherwise

where F(u,v) is a Fourier Transform of a pattern f(x,y) in the space domain and $\hat{F}(u,v)$ represents the BPOF. This means that the phase of the pixel in the BPOF is binary, 0 or π.

Let us consider that each pixel of the BPOF has a binary phase , 0 or ϕ not equal to π . To reduce computational requirements optimizing the BPOF later, the BPOF is represented as follow:

$$\hat{F}(u,v) = \tilde{F}(u,v)\ (1-\exp(j\phi)) + \exp(j\phi) \qquad (2)$$

where $\tilde{F}(u,v)$ is 0 or 1 . So, the phase of each pixel in the BPOF is

$$\hat{F}(u,v) = \begin{cases} 1 & \text{for } \tilde{F}(u,v) = 1 \\ \exp(j\phi) & \text{for } \tilde{F}(u,v) = 0. \end{cases}$$

If the BPOF consists of K x L pixels, $\tilde{F}(u,v)$ becomes

$$\tilde{F}(u,v) = \sum_{k=-K/2}^{K/2} \sum_{l=-L/2}^{L/2} F_{kl}\ \mathrm{rect}\left(\frac{u - k\Delta u}{\Delta u}, \frac{v - l\Delta v}{\Delta v}\right) \qquad (3)$$

where F_{kl} is 1 or 0. When the Fourier Transform is taken for (3),

$$\tilde{f}(x,y) = \Delta u\ \Delta v\ \mathrm{sinc}(x\Delta u, y\Delta v) \sum_k \sum_l F_{kl}\ \exp(2\pi j(kx\Delta u + ly\Delta v)) \qquad (4)$$

Then (4) is sampled at intervals $\Delta u\ \Delta x = 1/K$ and $\Delta v\ \Delta y = 1/L$,

$$\tilde{f}(m\Delta x, n\Delta y) = C\ \tilde{f}_{mn} \qquad (5)$$

where m is -K/2 to K/2 , n is -L/2 to L/2

$C = \Delta u\ \Delta v\ \mathrm{sinc}\ (m/K\ ,\ n/L)$,

$$\tilde{f}_{mn} = \sum_k \sum_l F_{kl}\ \exp(2\pi j(km/K + ln/L))$$

Then, if the Fourier Transform is also taken for $\hat{F}(u,v)$ in (2) , as for (4) and (5),

$$\hat{f}(m\Delta x, n\Delta y) = (1 - \exp(j\phi))\ C\ \tilde{f}_{mn} + \exp(j\phi) \qquad (6)$$

In this paper, it is assumed that the coefficient C in $\hat{f}$ (m Δx, n Δy) caused by the finite size of the pixel can be neglected with preprocessing of the pattern.[4] So, what we are concerned with is

$$\hat{f}_{00} = (1 - \exp(j\phi))\,\tilde{f}_{00} + \exp(j\phi)$$
$$\hat{f}_{mn} = (1 - \exp(j\phi))\,\tilde{f}_{mn} \qquad \text{for } m \neq 0, n \neq 0 \quad (7)$$

When the BPOF is synthesized for matched filtering and an input is a(x,y), the correlation of the BPOF with the input on the frequency domain is

$$\hat{G}(u,v) = A(u,v)\,F(u,v)\,(1 - \exp(j\phi)) + A(u,v)\exp(j\phi) \quad (8)$$

where A(u,v) is the Fourier Transform of a(x,y). Then the Fourier Tansform is taken for (8) and the result is sampled as in (3) to (5)

$$\hat{g}(m\,\Delta x, n\,\Delta y) = (1 - \exp(j\phi))\,C\,\tilde{g}_{mn} + \exp(j\phi)\,C\,\tilde{a}_{mn} \quad (9)$$

where

$$\tilde{g}_{mn} = \sum_k \sum_l A_{kl}\,F_{kl}\exp(2\pi j(mk/K + nl/L))$$

and

$$\tilde{a}_{mn} = \sum_k \sum_l A_{kl}\exp(2\pi j(mk/K + nl/L))$$

If the coeffient C in (9) is neglected as before,

$$\hat{g}_{mn} = (1 - \exp(j\phi))\,\tilde{g}_{mn} + \exp(j\phi)\,\tilde{a}_{mn} \quad (10)$$

So, $\hat{g}_{mn}$ is the discrete correlation of the B. OF with the input and it will be used to optimize the BPOF for matched filtering.

3. Optimization of the BPOF for matched filtering

Matched filtering has been used for pattern recognition. But conventional matched filtering doesn't produce a good result for similar patterns, for example the characters P and R in Fig.1, because P is imbedded in R. Thus the ratio of autocorrelation of P and P (AC[P,P]) to crosscorrelation of P and R (CC[P,R]) is almost 1. Existing encoding methods for BPOFs don't guarantee good performance for recognition of similar patterns even though the BPOF has an edge-enhancement property. The binary phase of the BPOF should be optimized to give a high ratio of AC[P,P] to CC[P,R]. We have used the simulated annealing (SA) algorithm for optimization of the binary phase of each pixel in the BPOF for matched filtering.

In SA, a system to be optimized is described as a physical system which has system variables u_i and an energy $E(u_i)$. When a variable is perturbed with Δu_i, the energy difference, $\Delta E = E(u_i+\Delta u_i) - E(u_i)$, is calculated. If $\Delta E < 0$, the perturbation , Δu_i , is accepted. Otherwise, it is conditionally accepted, based on the acceptance probability, $P(\Delta E) = 1 / [1 + \exp(\Delta E/T)]$ where T is a temperature parameter . Then the process is repeated for randomly chosen u_i. Decreasing T slowly as the process continues, the variables approach their optimum values which give the global energy minimum, or ground state of the system. If T decreases too fast, the system may get trapped in a local energy minima.

To apply SA to the optimization of the BPOF, the binary phase ϕ and the F_{kl} in (10) are defined to be the system variables[5]. The rate of the temperature decrease that we used[6] is

$$T = (D_T)^r\,T_{initial} \quad \text{and} \quad D_T = (T_{final}/T_{initial})^{1/q}$$

where r is the number of the iteration, q is the total number of iterations and $D_T > 0.9$. The energy function of the system is determined to achieve our goal : a high ratio of autocorrelation to crosscorrelation. The energy function was chosen as follows:

$$E = (H_A - AC[P,P])^2 + (H_A - AC[R,R])^2 + (H_C - CC[P,R])^2 + (H_C - CC[R,P])^2 \quad (11)$$

where H_A and H_C are target values for autocorrelations and crosscorrelations , and $H_A >> H_C$. As the SA process progresses, the energy approaches the global energy minimum of the system. This means that the autocorrelations and crosscorrelastions approach the target values.Two pixels of $\hat{g}_{mn}$ in (8) are chosen for AC[P,P] and AC[R,R] respectively for use in (9). So, CC[P,R] is the intensity at the pixel chosen for AC[P,P] and CC[R,P] at the pixel chosen for AC[R,R]. When F_{kl} in (8) is changed while decreasing the temperature T ,

$$\tilde{g}_{mn}^{new} = \tilde{g}_{mn}^{old} - A_{kl} \exp(2\pi j(mk/K+nl/L)) \quad \text{if } F_{kl} = 0$$
$$\tilde{g}_{mn}^{new} = \tilde{g}_{mn}^{old} + A_{kl} \exp(2\pi j(mk/K+nl/L)) \quad \text{if } F_{kl} = 1 \qquad (12)$$

Thus the new energy is calculated easily , using (10) , (11) and (12) for each randomly chosen binary phase of the pixels.

4. Results and Discussions

When the BPOF is generated with the input pattern P and R which consists of 64 x 64 pixels as in Fig.1, using (1), the reconstructed pattern is given Fig.2. The edges of P and R are enhanced as expected. When the unannealed BPOF is used as a matched filter, the ratio of AC[P,P] to CC[P,R] is almost 1 as in Table 1 so that they cannot be distinguished. After using SA on the filter as explained in the previous section, the optimized BPOF is reconstructed in Fig.3. The common part of the two characters is suppressed, while the different part is enhanced. But, in P, the common part is less suppressed than in R, while the different part is enhanced equally as the R. The results in Table 1 show that AC[P,P]/CC[P,R] is increased by more than 3 times and AC[R,R]/CC[R,P] by more than twice after SA. This means that the phase of the differing part of the patterns is encoded to be out of phase with the common part. Also, those parts are encoded to make AC[P,P] almost equal to AC[R,R] even though the energy of P is smaller than that of R.

An inherent problem of the BPOF is that it generates the Hermitian patterns as well as the original patterns. Even though the Hermitian patterns are considered noise , the overall performance is good enough to recognize two characters as the Table 1 shows.

5. Conclusion

The simulated annealing algorithm can be used to encode optimally the BPOF for distinguishing two similar patterns through matched filtering. After SA, the two patterns are optimally encoded in the binary phase of the BPOF. So, patterns that cannot be distinguished with conventional BPOF encoding methods are clearly recognized with SA. Further, the computational requirements for optimizing the BPOF are not excessive.

REFERENCES

1. S. Kirpatrick, C. D. Gelatt, and M. P. Vecchi, Science 220, 671 (1983)
2. J. L. Horner and H. O. Bartelt, Appl. Opt. 24, 2889 (1985)
3. D. Psaltis, E. G. Paek, and S. S. Venkatesh, Opt. Eng. 23, 698 (1984)
4. M. S. Seldowitz, J. P. Allebach, and D. W. Sweeney , Appl. Opt. 26, 2788 (1987)
5. M. R. Feldman and C. C. Guest, OSA annual meeting (1988)
6. C. S. Steele, M. S. Thesis, Cal. Tech. (1985)

TABLE 1

	Matched Filter	BPOF	Optimized BPOF	Overall Performance
$\frac{AC[R,R]}{AC[P,P]}$	1.2	1.4	1.0	1.0
$\frac{AC[P,P]}{CC[P,R]}$	1.0	1.1	3.5	2.2
$\frac{AC[R,R]}{CC[R,P]}$	1.2	1.7	3.5	2.2

Fig.1

Fig.2

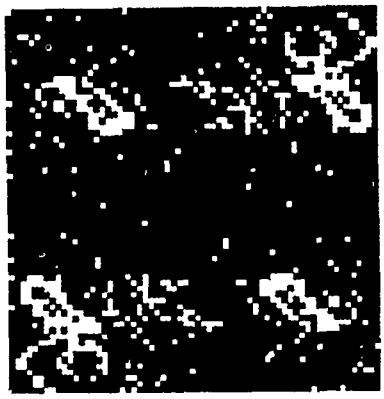


Fig.3

Fig.1 Characters to be recognized

Fig.2 Reconstruction of characters in Fig.1 with BPOF

Fig.3 Reconstructed pattern with the optimized BPOF

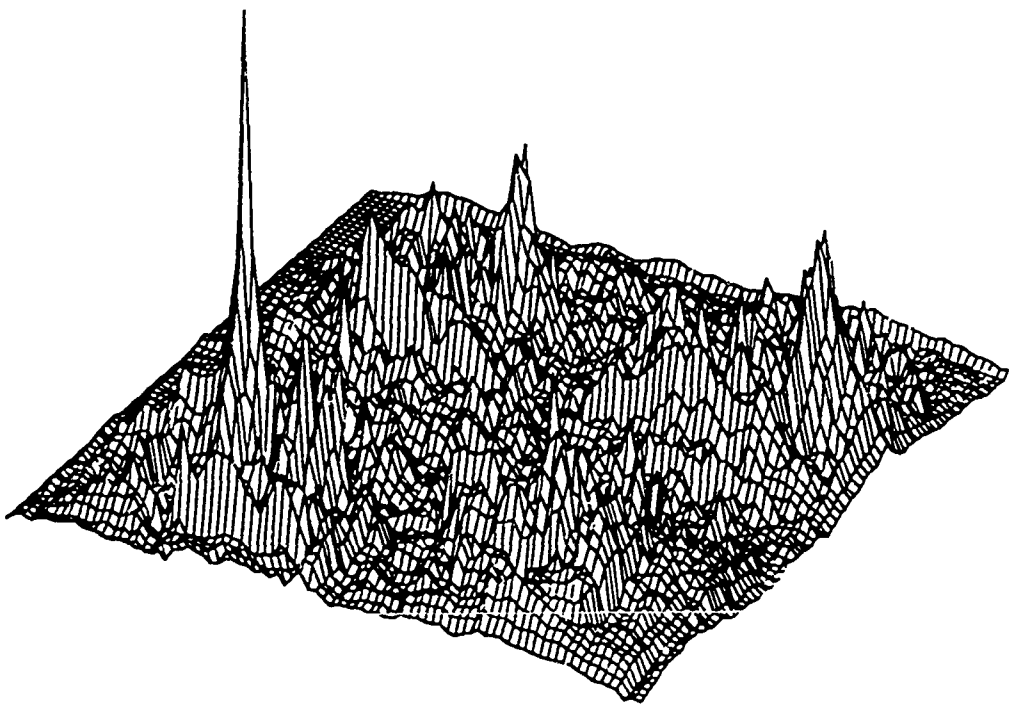


Fig.4 Correlation of the optimized BPOF with character P as input.

Computer Holographic Elements Using PostScript Laser Printers

Lawrence Domash and Philip Levin
Foster Miller, Inc.
350 Second Ave.
Waltham, MA 02154-1196
617-890-3200

Introduction

Computer generated holograms and diffractive optical elements are desired for a wide range of applications including optical interconnects, pattern recognition filters, binary phase filters, laser beam combiners, and displays. However, methods to design and produce them have been cumbersome. Typically, software been written anew for each application in Fortran on large computers and used to address e-beam lithography devices (1). This process results in submicron resolution but is slow, inflexible and costly. Only recently have efforts begun to bring to diffractive optics the highly developed computer aided design tools available for other engineering disciplines (2). Research and practical applications of diffractive/holographic elements of all kinds would be accelerated if there existed a simple, quick, low cost and easy to use method for creating optical patterns even of moderate resolution (1-3 microns).

Computer-Aided Publishing Technology

Other technologies driven by much larger markets than optics also make use of complex fine scale patterns. A recent revolution in electronic publishing based on personal computers has led to powerful new software and hardware tools for the design and output of graphic patterns. High definition industrial laser typesetters are now widely used which output arbitrary graphic patterns on film with up to 25,000 X 25,000 10 micron pixels.

Equally important, a software language, PostScript, has been widely accepted as a format to encode an arbitrary pattern and interface it to any of several dozen laser printers equipped with a PostScript interpreter/driver module (3). Numerous software applications are now available for many computer families which automatically produce PostScript output format files for word processing, scientific and mathematical graphics, graphic arts, image processing and engineering design. As an industry standard, PostScript provides a universal, device independent, resolution independent file format for arbitrary graphic patterns, including typefaces and bitmap images. Files encoded in PostScript are described in terms of analytically defined geometrical forms and can be sent to any PostScript-compatible output device for output; only the resolution of the representation will be device dependent.

We investigated the suitability of PostScript software, Macintosh computers and high definition laser typesetters for encoding and producing diffractive optics. We evaluated both the fundamental capabilities of the software environment and the optical quality of the hardcopy output.

PostScript

PostScript is not only a standard file format but also a programming language containing a powerful set of graphics instructions which we used directly to efficiently encode optically useful patterns (4). Graphic primitives include Bezier cubic curves, and procedures to perform such operations as filling a closed curve and scaling, rotating or combining complex objects. Graphic objects and their manipulations can be included in repetitive loops and other logical structures, just as numerical or text objects are in other languages. PostScript also permits combining the two sources of graphic objects of interest for optics: mathematically defined patterns and 8-bit greyscale bitmap images. PostScript possesses additional features which hold potential for optical software, including calculating the interaction of patterns according to logical rules, distorting collections of objects as a unit, and defining and manipulating "typefaces," defined as any collection of stored subpatterns.

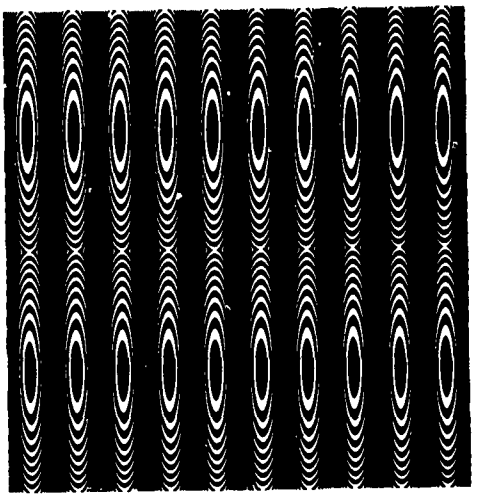
Laser Typesetters Output

PostScript is only interesting as a language for diffractive optics if output devices exist to convert files to hardcopy with sufficiently high resolution. In our research PostScript files were proofed using the relatively low cost Apple LaserWriter IINT at 300 dpi (85 micron pixels) and the same files were then sent to an Allied-Mergenthaler Linotronic 300 with a PostScript interpreter/ driver unit for film output. This device uses a HeNe laser to write 2540 dots per inch (dpi) for nominal 10 micron pixels on rolls of Agfa 2BV9K film 25 cm wide, so that binary patterns of up to 25,000 X 25,000 pixels may be produced. Access to the Linotronic 300 is widely available at electronic typesetting service bureaus in many cities on a per page basis. Photomicrographs of simple grating patterns showed that pixel quality was excellent, although some spurious periodic modulation errors were observed due to imperfections in the faceted laser scanning mirror.

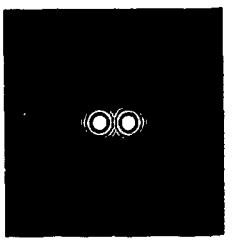
Tests of PostScript and Linotronic Output

A zone plate to focus a HeNe laser beam, a simple diffractive object, was easily designed and produced by writing a PostScript program containing only a few lines of code. A 2 cm diameter zone plate contained useful fringes almost to the 10 micron resolution limit. The binary amplitude film output was laser tested with good results.

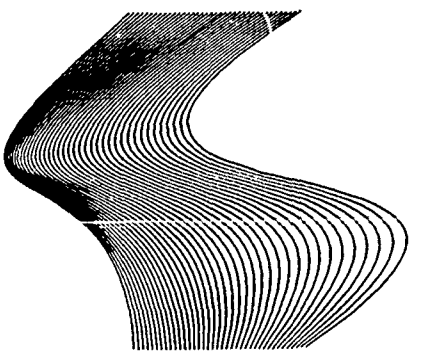
PostScript also has good facility for creating and managing fonts--defined as a library of standard subpatterns--and scaling and tiling them. An "optics font" of various interference patterns can be created in PostScript and used with almost the same ease as letters of the alphabet in a word processor. The figure shows a tiled collection of zone plates which have been distorted in one dimension with a single added instruction.



PostScript also permits two or more patterns to be combined by imposing a modulo 2 binary logic on overlapped regions through use of the "clipping" operation which yields the intersection of two regions, however complex. This means calculated holographic fringes can be combined by a direct digital graphic model of constructive and destructive optical interference, as opposed to being merely stacked incoherently as Moire patterns. However, this procedure gives rise to a large number of separate fragmented objects, which can exceed RAM capacity.

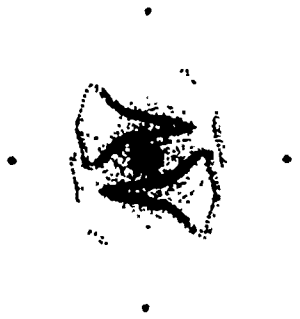


The structure of most holographic optical elements is in the form of fringes, which may be loosely defined as curved groups of parallel lines of varying width. PostScript includes instructions for grouping sets of Bezier curves of varying widths and defining distortion paths for the entire group, as illustrated. With appropriate quantitative definitions for the various parameters, this capability could be the basis for a highly efficient software approach to problems such as lens testing holograms.



As a test of the overall suitability for holographic elements in pattern recognition applications, we digitized a photo of a stealth airplane model and then computed a 2D Fourier transform of the binary 128 X 128 input image. The real part of the grey scale Fourier transform was binarized about the

zero level and converted to a PostScript file for transmission to the Linotronic. This procedure yielded an amplitude-binary filter on film which could serve as a mask for a phase-only binary filter. Samples were produced with 128 X 128 10 micron pixels (about 1 mm square) and tested by direct laser diffraction. The reconstruction is shown at right. The best results were obtained with 20 micron pixels and tiling of nine identical FFT patterns to reduce optical noise. The edge enhancement, conjugate image and central DC spot are all well known characteristics of binary amplitude filters.



Discussion

PostScript as a graphics description environment is surprisingly well suited to the production of holographic patterns. Laser typesetters such as the Linotronic have sufficient resolution to produce useful pattern recognition filters, and higher resolution holographic needs could be met through photoreduction and conversion of amplitude masks to binary phase-only filters. With appropriate software development, such tools could be extremely valuable for quickly designing and producing holographic filters and optical elements to test new ideas. The most important value of the PostScript/laser printer approach may be to offer the optics researcher access to a much larger, more sophisticated, and easier to use computer graphics technology than is likely to be developed for optics alone.

This research was supported by the Naval Surface Warfare Center under the Small Business Innovation Research program. The support and assistance of Nick Caviris are gratefully acknowledged.

References

1. S. Arnold, Opt. Eng., vol. 24, 803, 1985.
2. H. Farhoosh and S. Lee, Paper ME3-3, Topical Mtg. on Optical Computing, Incline Village, Nevada, March 1987.
3. PostScript is a trademark of Adobe Systems, Inc.
4. PostScript Language Reference Manual, Addison Wesley, 1985.

NOTES

WEDNESDAY, MARCH 1, 1989

SALON F

8:00 AM–9:00 AM

WA1–WA3

DIGITAL OPTICAL COMPUTING: 1

John A. Neff, DuPont Co., *Presider*

Digital Optical Computing with Fibers and Directional Couplers

Harry F. Jordan
Optoelectronic Computing Systems Center
University of Colorado
Boulder, CO 80309-0525

The goal of the Digital Optical Computing program of the Center for Optoelectronic Computing Systems is to design and demonstrate a prototype of a stored program optical computer using the knowledge base developed in connection with electronic digital computers. The target architecture is to be all-optical. This means that components with only optical inputs and outputs (except perhaps for power) are the basis of the architecture. The devices making up these components may have significant electronic parts which mediate the optical switching. From the architectural point of view, however, these components are treated as "black boxes" and could be replaced with any optical switching devices yielding the same functionality. The emphasis is thus on optical architecture, which we take to mean the architecture of a machine in which all information is carried between logic elements by optical signals, and in which the role of electronics interior to switching elements is minimized to the extent which is realizable.

The interesting architecture problems in the optical domain are a result of the time-space tradeoffs which become possible, and are essential, in a machine in which all signals are traveling at a fixed, finite speed. The distinction between combinational and sequential logic circuits is blurred, and ideas currently thought of as pipelining and systolic design reach their limiting case. Since computers are physical models of mathematical systems, however, one cannot study architectures in the absence of a physical realization. Our current understanding of computer architecture is dominated by explicit and tacit assumptions about the implementation capabilities of digital electronics. It is therefore necessary to have a real optical technology for implementing optical architectures. The problem at the present time is that, while optics can communicate large amounts of information at high speeds, optical switching devices are rudimentary, expensive and lack good characteristics for incorporation in large systems.

The route forward adopted by this program is to rely on architectural techniques similar to those used in electronic computers in the days when vacuum tubes represented the optimal switching elements. In the abstract, the requirements are to minimize the amount of logic and to use passive, rather than active, elements for data storage. These requirements led to the bit serial electronic designs of the late 1940s and early 1950s. This program bases its architectures on delay line loops for dynamic data storage, optical fiber for connections and delay lines, and controlled directional couplers as logic elements [1]. This technology base has a reasonable level of maturity as a result of development by the communications industry.

The use of dynamic storage is fundamental to speed of light operation. If data is to be represented optically in storage, as well as in logic operations, then storage is necessarily dynamic. The feedback logic circuits used to form electronic flip-flops are, of course, dynamic circuits at the speed of light. Bit serial architectures put emphasis on random, as opposed to regular, interconnection patterns. This makes optical fiber a good interconnection choice. Since the traversal of space and time are equivalent in optical

architectures, it is natural to use the same device for interconnection and storage. Several kilometers of optical fiber can be wound on a spool and handled conveniently in a laboratory environment, so moderate numbers of bits can be stored.

The choice of directional couplers as switching elements has the advantage that signals remain optical between the switched inputs and outputs. The only thing necessary to turn an electric field actuated directional coupler into an all-optical component is to add a photodetector and electronic amplifier to the control electrodes. A potentially promising technology is *GaAs*, in which photodetector, drive electronics, and waveguides might be built on the same substrate. However, the only device currently available in sufficient quantity to be used in a complete computer system is the $LiNbO_3$ directional coupler, whose degree of maturity is a result of its use in the communications industry. From the digital design point of view, the directional coupler is a controlled exchange element. It is not only logically complete, given the availability of constant inputs, but is natural for multiplexer based design and for switched interconnection of subunits.

Given the direction established by the above discussion, several research projects are underway within the Digital Optical Computing program. The first, and most complete is the instruction set and architecture for a 16 bit per word, microcomputer-like machine using 48 $LiNbO_3$ switches [2]. The architecture assumes ideal component characteristics, and it is understood that the number of switches will increase as actual device characteristics are taken into account. Since the machine is bit serial, the specification of 16 bits per word is really quite flexible. The 16 bit word allows one instruction per word with a single address format and a memory size of 1024 words. The architecture has been emulated and found to work correctly with ideal components. The detailed emulation is capable of taking non-ideal characteristics, such as delay, loss, and crosstalk into account. As parameters for the available components are determined, they are incorporated into the emulation and the architecture is refined.

Several binary counters are needed in the architecture, including one for the bit position in a word, one for the serial memory word position, and one for the instruction counter. Coupled with the fact that the counter is a simple feedback state machine, this makes the construction and operation of a counter an excellent first step in investigating the systems characteristics of the components to be used in the bit serial optical computer. The optical counter project was thus initiated. The long lead time to delivery of $LiNbO_3$ switches led to the decision to build a so-called mock counter which uses optical transmitters and receivers with electronic logic to emulate the operation of a $LiNbO_3$ switch [3]. Construction of a scale of 16 counter with such mock switches and fiber delay line storage led to the establishment of a step by step assembly and testing procedure for the optical counter which will enable us to assemble the optical switches without the need to break critical feedback loops to make measurements for debugging.

The technological problems which need to be overcome in implementing the bit serial optical computer have led to three distinguishable projects: $LiNbO_3$ drive electronics, delay line storage loop characteristics, and logic signal synchronization (delay distribution). The drive electronics for the directional coupler control electrodes, including the photodetector represents a significant effort. It differs from the high speed amplifier design for using such devices in the communications environment because the end to end latency through all stages of amplification is critical to the operation of a feedback circuit. In communications applications, only the bandwidth of the amplifier is critical, and

long latency can be tolerated since the systems are feed forward only, and thus amenable to pipelining. For 100 Mbit per second operation, the time from light incident on the photodetector to switching of the directional coupler must be only a few nanoseconds.

A more nearly system level project is the study of the characteristics of fiber and switches for use in delay line storage loops. The effect of temperature on physical length and index of refraction of optical fiber is an important parameter for synchronous operation. Asynchronous operation, while an alternative, has the major disadvantage of requiring expensive logic elements for recovering timing from stored data. The decision to regenerate correct data amplitude and timing from the system clock on each pass through the delay line makes signal degradation and crosstalk relatively less important in this subsystem, provided that the operating wavelength is properly matched to the fiber characteristics. A report on the study of delay line storage parameters is in preparation. A limit of about 10^4 bits per fiber loop represents the capabilities of a synchronous loop without very precise temperature control.

The problem of logic signal synchronization is being addressed both specifically, by using the bit serial computer emulation program, and in general, by development of an algorithm to produce an optimal, distributed delay design given a lumped delay architecture and a set of device delay characteristics. The emulation program will be sufficient to deal with the simple architecture we have developed. The general algorithm will be needed for more complex designs and is the first of several architectural research projects going on within the program. The initial design of architectures using signal propagation delays for all storage leads to a lumped delay system, in which delays are only introduced for the purpose of information storage. In a real system, all components and interconnections have an irreducible minimum delay associated with them. The synchronization of logic signals at their points of interaction requires that additional delays be added to some paths. An optimal system will do this by adding the minimum possible delay to the system. The network and component delay specifications can be represented as a weighted graph, and either linear programming or a shortest path algorithm used to determine minimum additional delays. A report on the general algorithm is in preparation.

Another architecture research project involves time slot interchange using delay lines and optical exchange elements. This work is a time domain equivalent of the multiport switching network research which has been so important in parallel computing. A permutation of blocks of information in different time slots of a serial data stream corresponds to switching among time multiplexed inputs and outputs. One application is to access words in a serial memory loop in a different order than the one in which they are stored. Since the perfect shuffle permutation forms the basis of several spatial switching networks, we started by studying networks for perfect shuffle of time slots using a minimum number of optical exchange elements and minimal end to end fiber delay. The report which is in preparation on this work will show that a network with delays corresponding to powers of 3 time slots is optimal for the architectures considered. This is true even for perfect shuffles of a power of 2 slots in the large number limit.

A future application of the time slot interchange study will be to an architectural project which is not yet under way. This will be the time multiplexing of several bit serial computers on the same physical hardware. Time multiplexing will give the ability to use more system bandwidth than is possible if the bit rate is limited by the end to end

latency of the shortest feedback loop which can be built. Multiplexing and demultiplexing, being feed forward operations, have a potentially higher bandwidth than the feedback operation required for stored program computing. Several machines, time multiplexed on the same hardware will demand high bandwidth components, but will relax the demand on end to end latency. Time slot interchange corresponds to interconnections among time multiplexed machines in this case.

An important application niche for the bit serial optical processing technology is in a multi-GHz optical packet transport network. Packet switched communication requires some minimal processing power at each switching node. Since information is transmitted bit serially and in optical form, there is a potential advantage to using the bit serial optical processor technology for simple routing in such a network. The use of *hot potato* routing protocols, in which no messages are stored in queues at nodes, avoids the problem of the lack of high speed optical random access memories. Planning for a project to research such a packet transport network is in progress between this program and the NSF Center for Telecommunications Research at Columbia University.

The over all research plan for this program can be summarized as follows. First, a working optoelectronic computer with an all-optical architecture will be demonstrated. The initial target is a 16 bit per word machine with a minimal instruction set operating at a rate of 100 Mbits per second. This prototype will form the basis for a faster version in the Gbit per second range as we gain experience with the technological problems involved. In parallel, we will be pursuing the switching node and network architecture for the self routing packet communication network as a short term payoff area for the bit serial optical computing technology. On an ongoing basis, but with a major study planned to coincide with the demonstration of the first prototype, we will assess the possibility of basing the switching technology on one of the promising optical devices which are currently under development. On the basis of the new implementation technology and the knowledge base gained from the initial optical architecture, we will reassess optical architectures for incorporation into a second generation optical computer, which will make even better use of the potential advantages of optically represented digital data.

REFERENCES

[1] H. F. Jordan, "A Bit Serial Optical Computer," *Second Topical Meeting on Optical Computing*, Optical Society of America, Lake Tahoe (March 16-17, 1987).

[2] V. P. Heuring, H. F. Jordan and J. Pratt, "A Bit Serial Architecture for Optical Computing," *OCS Tech. Report 88-01*, Optoelectronic Computing Systems Center, Univ of Colo., Boulder, CO 80309-0525 (March 1988).

[3] A. B. Yadlowsky, "The Mock Counter: A Hybrid Optical-Electronic Counter Using Fiber Delay Line Memory," *OCS Tech. Report 88-04*, Optoelectronic Computing Systems Center, Univ of Colo., Boulder, CO 80309-0525 (September 1988).

Dynamic Optical Processing for Parallel Digital Addition and Subtraction

Takashi KUROKAWA, Seiji FUKUSHIMA
NTT Opto-electronics Laboratories
3-1, Morinosato Wakamiya, Atsugi-shi, Kanagawa 243-01, Japan

Hideo SUZUKI
NTT Communications and Information Processing Laboratories
9-11, Midori-cho 3-chome, Musashino, Tokyo 180, Japan

Introduction

An optical arithmetic processor for digital addition will be one of the most important basic elements in optical computers. Several optical implementations of half or full adder circuits have been proposed.[1-4] However, dynamic processing has not been developed for parallel digital addition. The main problem is real-time parallel operation of the half adder and the dynamic construction by cascadable configuration.

In this paper, dynamic parallel arithmetic processing is demonstrated for digital addition and subtraction. An all-optical half adder circuit using polarization logic combines ripple carry structure with micro-channnel spatial light modulators (MSLMs: Hamamatsu photonics K. K.). Real-time operation is achieved by synchronous control of logic gates, latches and input and output ports with a microprocessor.

Optical Implementation

Sum and carry logic operations can be achieved by polarization logic, in which two logic states are represented by two orthogonal polarization states. In the MSLM operation, the polarization of the reflected readout light rotates by 90 degrees in the write-light incident region.[5] This means that several logic operations are obtained by selecting the polarization direction of the readout light from the cascaded MSLMs.

The optical circuitry of a half adder is shown in Fig. 1. The polarization of the readout light rotates by 90 degrees through a series of two MSLMs when either of two write-lights are bright, so that the XOR operation, giving the sum signal, is performed by selecting the polarization direction of readout light with a polarization beam splitter. The NAND operation, giving the negation of the carry signal, is obtained from the superposition of XOR and NOT output beams. In the half adder described above, input is binary polarized data arrays on optical patterns. Therefore, a lot of half additions can be performed in parallel.

The arithmetic processor for digital addition is designed by the ripple carry method. A schematic representation of the processor is shown in Fig. 2. The sum and carry signals from the half adder constructed by two MSLMs, M1 and M2, are fed back to the half adder through MSLMs, Ms and Mc latch memories. All MSLMs are synchronously controlled by a microprocessor. The sum signals read from the Ms are fed back to the same-bit pixel on M1 in parallel, while the carry signals read from the Mc are fed back to the upper-bit pixel on M2, shifted in parallel with a mirror. A number of k-bit word pairs can be added through k-time feedback loops in parallel.

Two liquid crystal cells, L1 and L2, and a CCD array sensor are installed as the input and output interfaces between the optical logic unit and the electronic memories. Input ports convert the electronic data stored in the LSI memories to two optical data patterns. The output port converts the parallel optical output into serial electronic data to store it in the

memories. The MSLMs, the liquid crystal cells, the array sensor and the electronic memories are controlled synchronously by a microprocessor, as shown in Fig.3. The output pattern is observed by a monitor TV .

Results and Discussions

The parallel addition of 2 word pairs consisting of 3 bit binary digits was carried out programmably. The sum and carry patterns observed in the loop are shown in Fig. 4, for addition example, (011,011)+(010,001)=(101,100). The addition of the lower words required three feedback cycles, corresponding to the longest computing time. In contrast, the addition of the upper word was finished by the second cycle. It took about 1 second for writing and erasing patterns on the MSLMs.

This processor also performed the parallel subtraction, adding a conversion process to two's complements of binary input data. The subtraction A-B was executed through a revision of the operation program for liquid crystal cells and MSLMs as follows: The vector B and an all-true pattern are input to LC1 and LC2, and then the negation of B was obtained from the sum output. By the addition of the vector of '0···01' input to LC2 and $\overline{B}$ fed back to M1, the two's complement of B was obtained. The parallel subtraction, A-B, was finally performed through the addition of the other vector A, input to LC2, and the complements of B fed back to M1. The parallel subtraction of 2 word pairs consisting of 4 bit binary digits was carried out experimentally. The processing time was about twice that of the addition, because subtraction needed one more addition and an XOR operation to make the complements.

Pixel size, which decides the parallelism, was limited by beam diffraction because the light propagation distance between spatial light modulators was quite large. Spatial optical integration and imaging optics are necessary to attain a much higher parallelism. Total addition time was limited only by the logic operation time, not by the time of serial-parallel and parallel-serial conversion between electronic memories and the optical logic unit because the switching time of the spatial light modulator was very slow. However, memory access time would limit the total operation speed if much higher parallel addition could be executed with fast logic devices.

Conclusion

Parallel arithmetic processing has been demonstrated for digital addition and subtraction. The processor had a ripple carry adder array structure. Parallel addition and subtraction of 2 word pairs were carried out in real time. This kind of processor has a high potential for parallel digital operation, if highly improved spatial light modulators are introduced.

The authors would like to thank Y. Hayashi for his helpful discussions and Dr. T. Ikegami for his encouragement.

References

1. M. T. Fatehi, K. C. Wasmundt, and S. A. Collins, Jr. , Appl. Opt. **20**, 2250 (1981)
2. B. S. Wherrett, Opt. Comm. **56**, 87 (1985)
3. J. Tanida and Y. Ichioka, J. Opt. Soc. Am. **73**, 800 (1983)
4. A. W. Lohmann and J. Weigelt, Appl. Opt. **25**, 3047 (1986)
5. T. Hara, . Shinoda, T.Kato, M. Sugiyama, and Y. Suzuki, Appl. Opt. **25**, 2306 (1986)

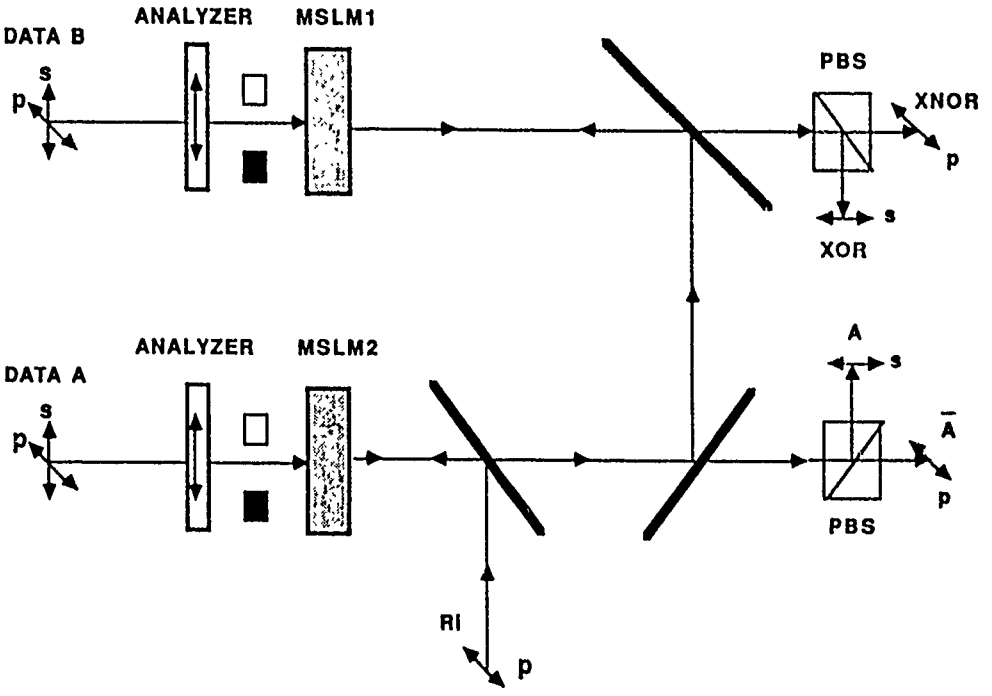


Fig. 1. Optical implementation of a half adder by polarization logic.

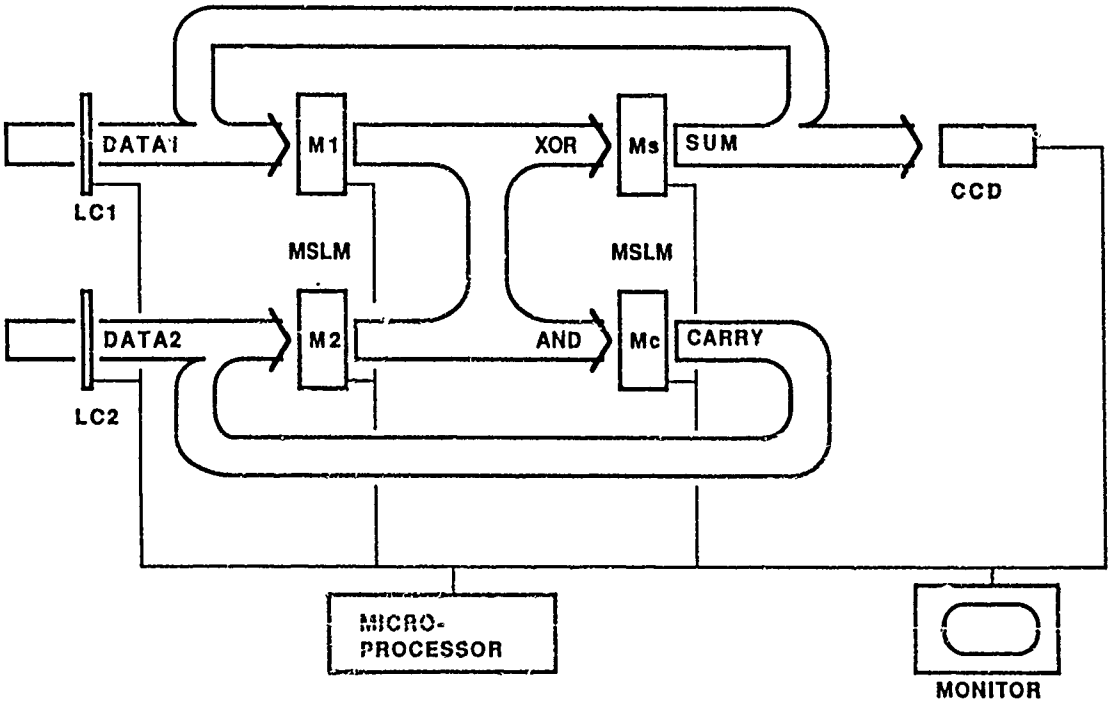


Fig. 2. A schematic representation of an optical processor for digital addition and subtraction.

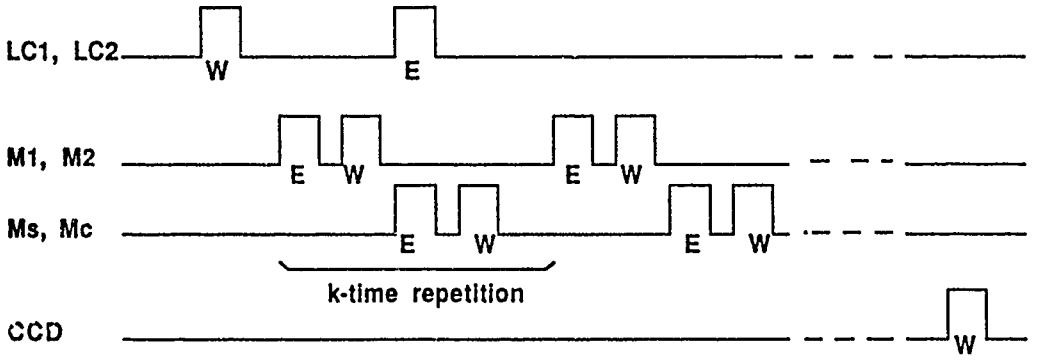


Fig. 3. Synchronous control signals in an optical processor. E and W are control pulses for erasing and writing.

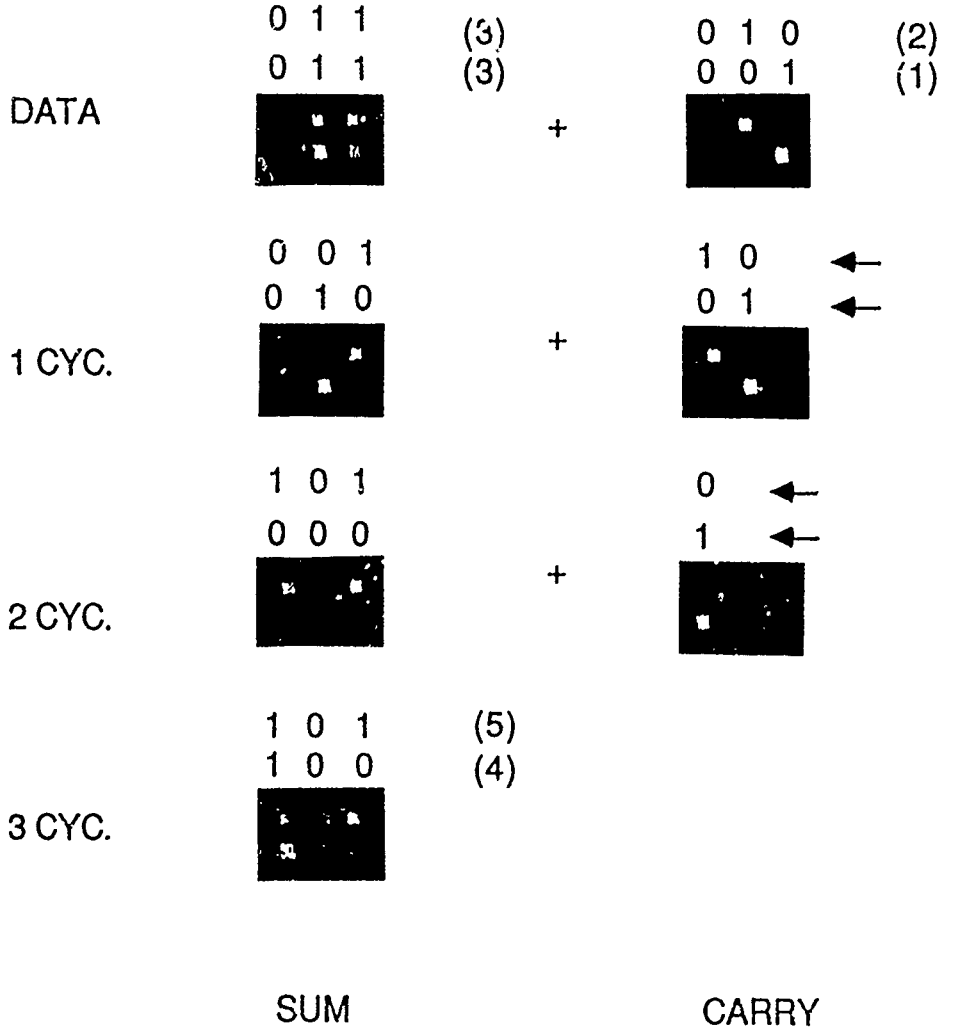


Fig. 4. Experimental results of parallel addition of 2 word pairs consiting of 3 bit binary digits.

Flexible-Structured Computation Base on Optical Array Logic

Jun Tanida, Masaki Fukui, and Yoshiki Ichioka

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

I. Introduction

Optical computing techniques have excellent features for large capacity information processing, such as massively parallelism, high speed processing, crosstalk free interconnection capability, and so on. Among those features, reconfigurability of optical processing systems should be stressed. To utilize this excellent feature, we have considered flexible-structured computation with optical array logic (OAL).[1,2] The programmability of OAL is fully utilized for designing such a flexible-structured computing system.

II. Optical Array Logic[1,2]

OAL is a technique to achieve any parallel neighborhood operation for two 2-D binary data according to the procedures shown in Fig.1. Two 2-D binary images are encoded into a coded image composed of four kinds of code patterns. The coded image is separately correlated with several pointwise functions called operation kernels. The individual correlated images are spatially sampled at 1-pixel intervals. Inverted OR operation for all the sampled images provides the result of a parallel neighborhood operation. Since all the procedures can be executed with optical system in parallel, OAL is one of promising techniques of optical computing.

The processing manner in OAL is identical to the sum of product processing, so that any logical operation for 2-D data can be achieved by a combination of operation kernels. This means that functions of OAL can be programmed with operation kernels and that OAL has great capability for parallel digital optical computing.

In addition, OAL can implement space-variant operations with a specific programming technique. The fundamental of the technique is that one of the 2-D inputs of OAL is used for data to be processed and the other is used for the selector of the operation to be executed. We call these inputs *data* and *attribute planes* according to their usage. The data and attribute planes have pixel patterns indicating the data and its operational selector, respectively. With the combination of attribute patterns and operation kernels, space-variant operations can be executed.

III. Virtual Machine Implementation

In modern architecture of parallel computers, the structure of the system reflects the structure of data to be processed because of computational efficiency. So that, many networks for parallel processing systems are proposed to process various structures of data. This means specialization of computer architecture and difficulties of making general-purpose computers. However, reconfigurability of optical system enables us to construct such a general-purpose computer with virtual hardwired logic and flexible data transfer capability. Using these ideas, we have designed a virtual computing system with OAL, which can treat various types of structured data efficiently.

Figure 2 shows an implementation of a Turing machine.[3] A Turing machine is considered as an elemental computing machine expressing every computers, so that its implementation is important to verify the capability of our scheme. Technically, a data tape, internal state, and head cells are required for such a machine. We express these cells by bit patterns and set on attribute and data planes as well as tape alignment cell indicating the position of data tape. The function of the machine is assigned with operation kernels as shown in Fig.3.

The number of terms in an operation kernel corresponds to the required number of optical correlation and indicates the step number for the program execution. Thus, the programming

efficiency is estimated by the number of terms in an operation kernels. Using this configuration, we can operate any number of Turing machines in parallel on processing planes of OAL.

VI. Systolic Computation

As an example of virtual machine on OAL, a systolic computing array[4] is demonstrated. Systolic computation has been developed for VLSI systems to avoid limitation of pin number of VLSI chips. Many types of structured data are efficiently computed with this scheme, i.e., matrix-vector multiplication, matrix-matrix multiplication, and so on. However, this scheme requires a network depending on the data structure, so that if a reformable network is realized, flexibility for data structure can be obtained.

Figure 4 is an implementation of an inner-product step processor in systolic computing array. For n-bit number computation, 8 by $2n+2$ cells are prepared as registers. Modified signed digit[5] is used for number representation because of processing efficiency. To describe the function of a inner-product step processor for n-bit, an operation kernel with $40n$ terms is required. For data transmission, 4-term operation kernel is used. Thus, $40n+4$ steps are required to drive one step operation in the inner-product step processor.

The simulating result of matrix-vector multiplication with a systolic computing array on OAL is shown in Fig.5. The processing plane is divided into three parts: store areas for vector and matrix data and that for inner-product step processor. At the final step, computing result is obtained in the area of vector data. Note that data transfer in this system is executed by shift operation in OAL and it can be controlled with operation kernels. In addition, the location of each inner-product step processor is determined by the pixel pattern in the attribute plane. Consequently, this systolic computing system has reconfigurability and flexibility for operand structure.

V. Hardware Implementation

To execute OAL effectively, we have proposed an ideal computing system named OPALS (optical parallel array logic system).[6] The OPALS has many variations for its physical implementation. Among of them, the system using birefringent phenomenon is promising because of hardware simplicity and great capability. The birefringent version of OPALS can execute correlation with a large size of operation kernel as well as image encoding. The programs presented in the previous sections can be executed optically on the birefringent version of OPALS.

VI. Summary

We have considered flexible-structured computation with optical array logic. As examples of such a system, we programmed two types of virtual machines using optical array logic and demonstrated some simulation results. The programs can be executed on the OPALS optically. The requirements for effective processing are large size of processing plane and correlation capability for large size of operation kernels.

References

1. J. Tanida and Y. Ichioka, "Programming of optical array logic: 1. Image data processing," Appl. Opt. **27**, 2926 (1988).
2. J. Tanida, M. Fukui, and Y. Ichioka, "Programming of optical array logic: 2. Numerical data processing based on pattern logic," Appl. Opt. **27**, 2931 (1988).
3. M. Minsky, *Computation: Finite and Infinite Machines* (Prentice-Hall, Englewood Cliffs, NJ, 1967).
4. H. T. Kung, "Why systolic architectures?" Computer **15**, (1), 37 (1982).
5. A. Avizienis, "Signed-digit number representations for fast parallel arithmetic," IRE Trans. Electron. Comput. **EC-10**, 389 (1961).
6. J. Tanida and Y. Ichioka, "OPALS: optical parallel array logic processor," Appl. Opt. **25**, 1565 (1986).

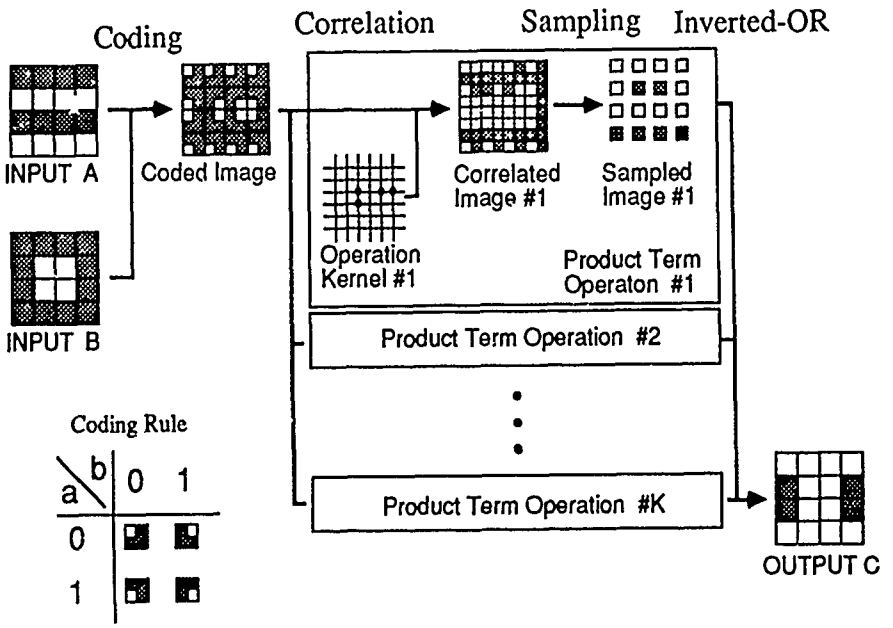


Fig.1. Schematic diagram of optical array logic

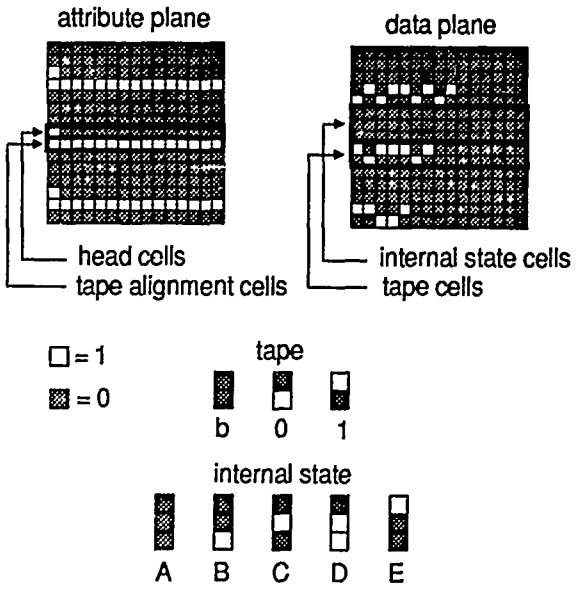


Fig.2. OAL implementation of Turing Machine

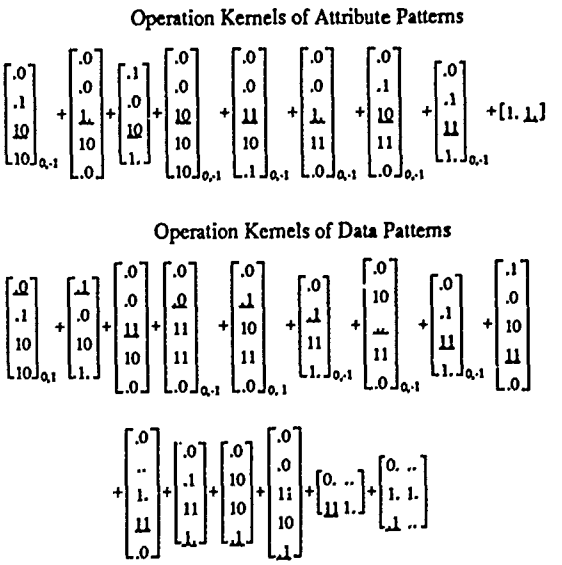


Fig.3. Operation kernels to drive Turing Machine

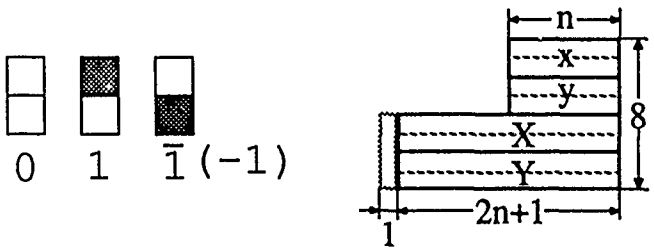
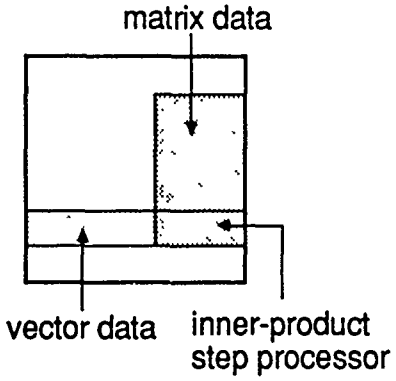


Fig.4. OAL implementation of inner-product step processor

$$\begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \\ 0 & 3 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \\ -7 \end{bmatrix}$$



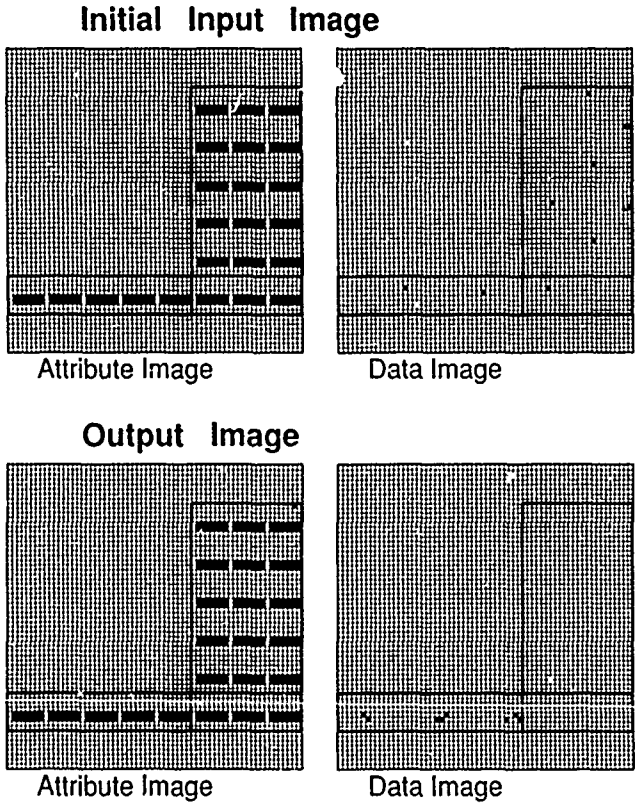


Fig.5. Simulating result of systolic computation

NOTES

WEDNESDAY, MARCH 1, 1989

SALON F

9:00 AM–10:00 AM

WB1–WB4

DIGITAL OPTICAL COMPUTING: 2

Ravindra A. Athale, BDM Corporation, *Presider*

Reconfigurable Programmable Optical Digital Computer

P.S. Guilfoyle, F. F. Zeise
OptiComp Corporation
PO Box 10779
310 Dorla Court, Suite 210
Zephyr Cove, Lake Tahoe, NV 89448

Abstract

Previous optical computing schemes offered analog or quasi-digital accuracies with a single fixed primitive. This paper describes how programmable, arbitrary bit length all digital Central ProcessingUnit (CPU) computations are now possible. In addition, the current state-of-the-art in optical computer subsystem devices such as acousto-optic modulators, detector and source arrays, posture this architecture as a revolutionary technology in and of itself, as it may be applied to an implementation plethora.

Technical Summary

Our research has produced a new class of optical computing architecture — a general purpose digital optical computer of *arbitrary* bit length. Shannon's theorem on general purpose digital computation states that all digital logic functions can be represented by two sets of equations. The first set takes the input data vector represented by bits x_1 through x_n and combines the bits in such a way to produce k output combinatorial functionals f_1 through f_k. Note that f_1 through f_k represent the logical/Boolean "multiplication" or "AND"ing of any combination of x_1 through x_n. These inputs, x_1 through x_n, are represented in "dual rail" format, i.e. both x_i and its complement (shown with a bar over them) are available. We shall refer to this first step as the combinational "AND"ing of the arbitrary input data vectors.

$$f_1 = x_1 x_2 \ldots \bar{x}_i \ldots x_{n-1} x_n$$

$$f_2 = \bar{x}_1 x_2 \ldots x_i \ldots x_{n-1} x_n$$

$$f_3 = \bar{x}_1 \bar{x}_2 \ldots x_i \ldots x_{n-1} x_n$$

$$\vdots$$

$$f_k = \bar{x}_1 \bar{x}_2 \ldots \bar{x}_i \ldots \bar{x}_{n-1} \bar{x}_n$$

The second step in Shannon's generalized formulation is to take these arbitrary combinational functionals and produce arbitrary combinational summations as shown in the second set of equations below. Inputs to the second step are the outputs from the first step above, i.e., the combinational "AND" products f_1 through f_k. These are then "OR"ed or Boolean summed as shown in arbitrary dual rail form. The equivalent function of $\bar{f}_n$ can be realized at worst as a sum of only f_m (high true) functionals.

$$y_1 = f_1 + f_2 + \ldots + f_i + \ldots + f_{n-1} + f_n$$

$$y_2 = \bar{f}_1 + f_2 + \ldots + f_i + \ldots + f_{n-1} + f_n$$

$$y_3 = \bar{f}_1 + \bar{f}_2 + \ldots + f_i + \ldots + f_{n-1} + f_n$$

$$\vdots$$

$$y_k = \bar{f}_1 + \bar{f}_2 + \ldots + \bar{f}_i + \ldots + \bar{f}_{n-1} + \bar{f}_n$$

To facilitate the selection of the appropriate terms in both sets of equations, control selection logic must be used on the dual rail input data before either of Shannon's equations can be realized. Figure 1 shows how this can be performed on a simple optical computer.

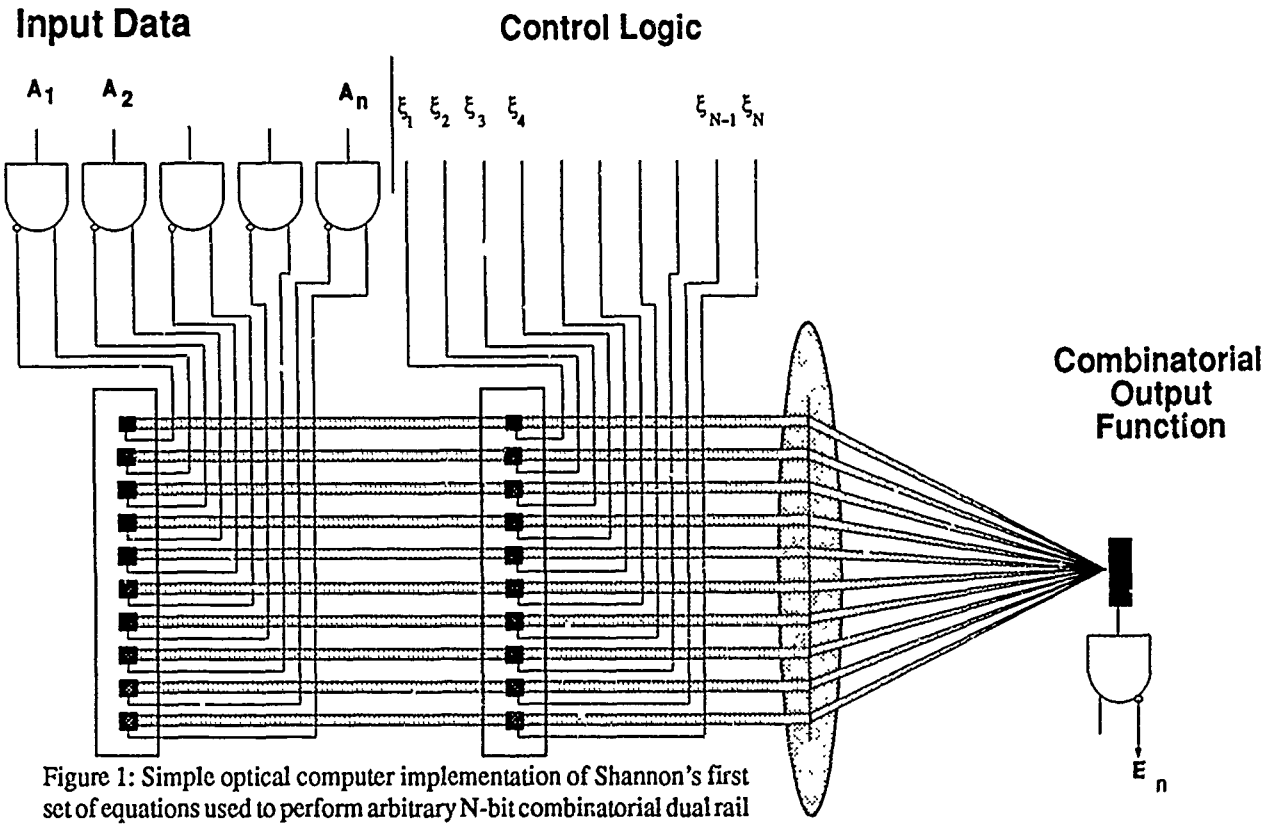


Figure 1: Simple optical computer implementation of Shannon's first set of equations used to perform arbitrary N-bit combinatorial dual rail multiplication or "AND"ing.

In figure 1 input data is fed from the data bus in dual rail format to a set of electro-optic transducers. Given n input data bits, 2n transducers are required. At the same time control logic is sent to a second set of input transducers. The optical system shown images the first set of transducers onto the second set. The resultant products, n two input "AND" gates, are then "OR"ed on the detector. The benefit of this is that the detector need only detect the presence or absence of light. Fan-in on the detector can be quite high as the off state is the required information state, i.e. a dark system. Only multiplicative modulation efficiencies of the devices determine the leakage or fan-in limitation as compared to previous summing or multi-level threshold logic schemes.

The output of the detector thus can be written:

$$E_N = \overline{A_1\xi_1 + \overline{A_1}\xi_2 + A_2\xi_3 + \overline{A_2}\xi_4 + \ldots + A_N\xi_{2N-1} + \overline{A_N}\xi_{2N}}$$

This represents 2N "AND" gates "OR"ed together. It is critical to recognize at this stage the impact of DeMorgan's laws. The particular law that should be applied at this point is the "AND" law. Simply stated this Boolean logic law is written:

$$XY = \overline{\overline{X} + \overline{Y}}$$

In other words, the inverted Boolean sum of conjugated input bits is equivalent to their Boolean product. This allows us to rewrite the output E_N from above as an N bit Boolean "AND" product!

$$E_N = \overline{A_1\xi_1} * \overline{\overline{A_1}\xi_2} * \overline{A_2\xi_3} * \overline{\overline{A_2}\xi_4} * \ldots * \overline{A_N\xi_{2N-1}} * \overline{\overline{A_N}\xi_{2N}}$$

Consequently, by producing the required control bits (microcode), ξ_1 through ξ_{2N}, it is possible to arbitrarily *program* this machine to produce any sequence of combinatorial multiplications of arbitrary bit length. Without DeMorgan's law, a sequential stack of spatial light modulators of stack height N would be required and therefore impractical. The outputs E_N now represent Shannon's combinatorial output functionals f_1 through f_k given a sequence of k control vectors of length 2N.

These combinatorial output functionals can be "OR"ed to produce Shannon's second set of equations by (1) passing the functionals back through the optical system, (2) supplying the correct microcode for the second set of equations, and (3) ignoring DeMorgan's law, i.e. do not take the inverted output. This now represents what is commonly referred to as an *instruction*. It is thus possible by downloading the correct microcode stored in a memory subsystem, to program the machine to perform instructions. Different microcoded sequences will act on the data in different fashions thereby providing the user access to a *microcode instruction set*. If this instruction set comprises a complete set of operations, a compiler code generator can be written for any desired higher level languages. A *fully general purpose optical computer can thus be realized*. However, the optical architecture as shown in figure 1, does not represent a competitive interconnect configuration which will allow optics to perform within its optimal characteristics. Parallel implementations of microcode are possible.

Most all code that exists today is Von Neuman in nature, i.e. *single instruction sequential*. What would be desired is a fast Von Neuman machine without I/O bottle necks. The architecture described here provides a solution. Parallelism is identified by the compiler and exploited at the microcode level. That is, each instruction can be written as parallel combinatorial functionals. Data re-use is achieved by operating on the data several times within one instruction thereby avoiding the I/O bottleneck.

Consider the optical matrix/vector computing architecture shown in figure 2 titled Fixed Program Flash N bit ALU. Instead of having a parallel array as shown in figure 1, this architecture utilizes the three dimensional capability of optical computing. The input source data vector is input in dual rail format to the input source array. This vertical input vector parallel illuminates a control operator plane which consists on α, N bit control sequences. In parallel all combinatorial functionals, f_1 through f_α (α could equal k if desired) are available simultaneously at the output detector array. Consequently the system is computing microcoded combinatorial functionals in parallel.

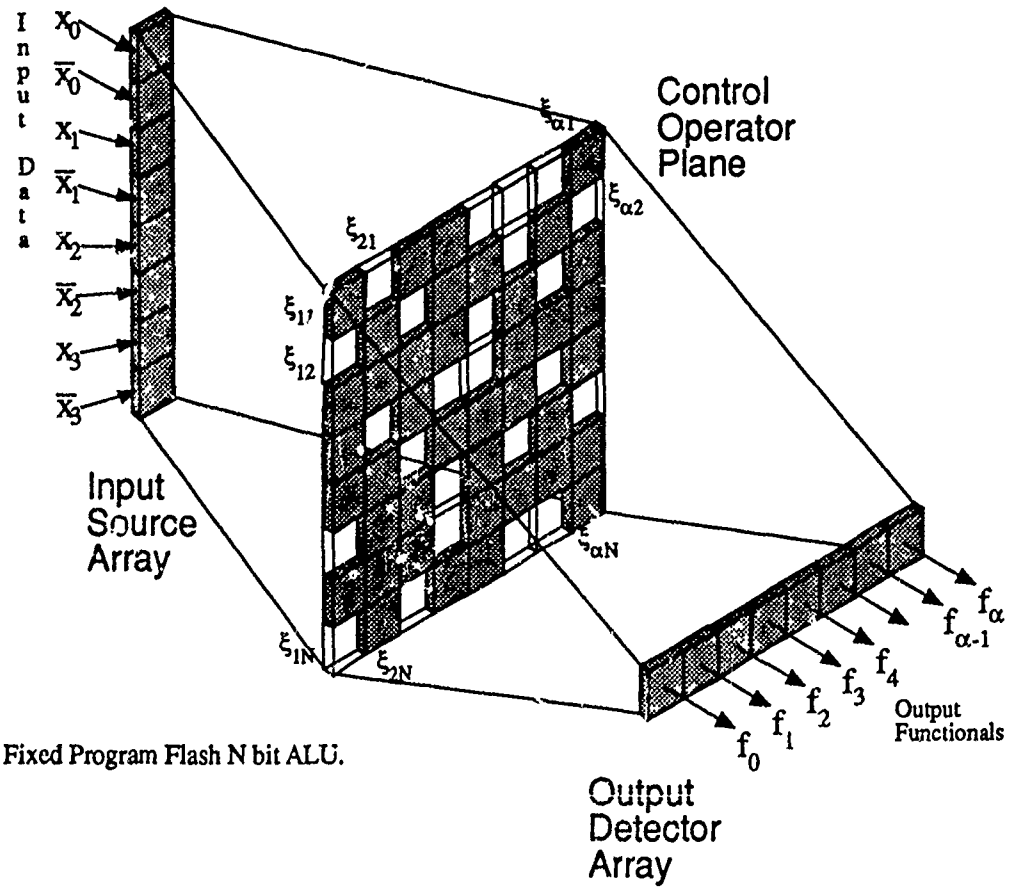


Figure 2: Fixed Program Flash N bit ALU.

This architecture can be represented as a Boolean logic matrix/vector multiplication which produces all of the combinatorial output functionals f_1 through f_k. The only difference between this matrix vector formulation and one use commonly in mathematics is that the inner product summation terms are actually threshold detections, Boolean

summations, or "OR"ings. The only precision that is needed is binary, i.e. 1 or 0. The maximum inner product answer is 1. However the effect is to have multiple parallel input "AND" gates.

$$\begin{bmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1N} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2N} \\ \vdots & \vdots & & \vdots \\ \xi_{\alpha-1,1} & \xi_{\alpha-1,2} & \cdots & \xi_{\alpha-1,N} \\ \xi_{\alpha 1} & \xi_{\alpha 2} & \cdots & \xi_{\alpha N} \end{bmatrix} \begin{bmatrix} x_0 \\ \overline{x_0} \\ \vdots \\ x_{\frac{N}{2}} \\ \overline{x_{\frac{N}{2}}} \end{bmatrix} = \begin{bmatrix} \overline{f_0} \\ \overline{f_1} \\ \vdots \\ \overline{f_{\alpha-1}} \\ \overline{f_\alpha} \end{bmatrix} =$$

$$\begin{bmatrix} x_0\xi_{11} + \overline{x_0}\xi_{12} + x_1\xi_{13} + \overline{x_1}\xi_{14} + \cdots + x_{\frac{N}{2}}\xi_{1,N-1} + \overline{x_{\frac{N}{2}}}\xi_{1,N} \\ x_0\xi_{21} + \overline{x_0}\xi_{22} + x_1\xi_{23} + \overline{x_1}\xi_{24} + \cdots + x_{\frac{N}{2}}\xi_{2,N-1} + \overline{x_{\frac{N}{2}}}\xi_{2,N} \\ \vdots \\ x_0\xi_{\alpha-1,1} + \overline{x_0}\xi_{\alpha-1,2} + x_1\xi_{\alpha-1,3} + \overline{x_1}\xi_{\alpha-1,4} + \cdots + x_{\frac{N}{2}}\xi_{\alpha-1,N-1} + \overline{x_{\frac{N}{2}}}\xi_{\alpha-1,N} \\ x_0\xi_{\alpha 1} + \overline{x_0}\xi_{\alpha 2} + x_1\xi_{\alpha 3} + \overline{x_1}\xi_{\alpha 4} + \cdots + x_{\frac{N}{2}}\xi_{\alpha,N-1} + \overline{x_{\frac{N}{2}}}\xi_{\alpha N} \end{bmatrix}$$

Control matrix — Data vector — Output functional vector

This matrix/vector formulation represents a complete instruction. Notice that all output functionals f_1 through f_α are produced. Again α could equal k if desired. Each vertical vector on the optical architecture of figure 2 is in fact producing one of the α equations shown above. Note again that the summations shown are actually "OR" functions and the detector is merely thresholding. Again applying DeMorgan's, law as shown below, after inversion the output combinatorial functionals are actually realized.

$$\begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{\alpha-1} \\ f_\alpha \end{bmatrix} = \begin{bmatrix} \left(\prod_{i=1}^{\frac{N}{2}} \overline{x_{i-1}\xi_{1,2i-1}}\right)\left(\prod_{i=1}^{\frac{N}{2}} \overline{\overline{x_{i-1}}\xi_{1,2i}}\right) \\ \vdots \\ \left(\prod_{i=1}^{\frac{N}{2}} \overline{x_{i-1}\xi_{\alpha,2i-1}}\right)\left(\prod_{i=1}^{\frac{N}{2}} \overline{\overline{x_{i-1}}\xi_{\alpha,2i}}\right) \end{bmatrix}$$

The control logic matrix here represents a *complete instruction* on the input data vector x_1 through x_n. The output is the first set of answers required by Shannon's theorem. They can be fed back to the input, the control operator changed (or downloaded as the case may be) and the second set of Shannon's equations are produced at the output, thus representing, in two computation cycles, a complete instruction.

For a complete text of this paper the reader sould review reference [1]. which is the fourth of a series of papers describing combinatorial logic based optical computing methods. For furthur background information the reader is encouraged to review in addition references 2-4 cited below. Reference 5 describes the author's original transition architecture from analog to digital optical computing.

REFERENCES

[1.] P.S. Guilfoyle, "Programmable Optical Digital Computing," Proc. of the 21st Annual Asilomar Conference on Signals, Systems and Computers, Nov. 3, 1987.

[2.] P.S. Guilfoyle, W.J. Wiley, "Globally Folding Combinatorial Logic Cells in Digital Optical Systolic Computing Arrays," Proceedings of the 1987 2nd Topical Meeting on Optical Computing, Lake Tahoe, NV., March 1987.

[3.] P.S. Guilfoyle, W.J. Wiley, "Digital Optical Linear 3 x 3 Bit Combinatorial Systolic Multiplication Array," PROCEEDINGS OF THE SPIE, Real Time Signal Processing IX, Vol. 698-30, August, 1986.

[4.] P.S. Guilfoyle, W.J. Wiley,"Combinatorial Logic Based Optical Computing," PROCEEDINGS OF THE SPIE, Vol. 639-17, April,1986.

[5.] P.S. Guilfoyle, "Systolic Acousto-Optic Binary Convolver," OPTICAL ENGINEERING, vol. 23, Number 1, pg. 20-25, Jan./ Feb., 1984.

Programmable Logic Gate Array and its Applications to Reconfigurable Network Based on Modified Sign Digit

Yoshiki Suzaki and Toyohiko Yatagai

University of Tsukuba, Institute of Applied Physics
Tsukuba, Ibaraki 305, Japan

We have developed a programmable parallel logic unit with a dynamic interconnection ability based on the truth table architecture. This enable us to design very flexible digital optical computing systems. An element cell has two input ports and three output ports, as shown in Fig. 1. Element cells are connected optically each other, and the connection network is changeable by selecting three output ports. The element cell is able to select which output ports are active or not. Because of this feature of the element cell the interconnection network is dynamically configured. It is convenient to use three-state logic or ternary logic to assign one or two of three ports to the output output port. Ternary logic is also employed to perform a modified sign digit (MSD) operation, which enables us to make a full parallel algorithm

of numerical calculation.

Because two three-state inputs A and B is considered, each pixel of the inputs A and B are represented with the position of a bright luminous subpixel as shown in Fig. 2. Pixels of the input A, Aij are coded in the vertical direction, the pixel values of -1, 0 and 1 are are assigned the bright positions of bottom, center and top, respectively, while pixels of the input B, Bij are coded in the horizontal direction.

We consider a cellular logic architecute consisting of a 3-D array of simple logic operation units, of which interconnection are changed by the results of the operations in element units. A structure of the element cell is shown in Fig. 3. The cell consists of three parts; the operation part, the condition part and the interconnention part. Two encoded input patterns incident to the cell are superimposed and their copies are transferred to the operation part and the condition part. In the operation part, the superposition of the superimposed input pattern and an operation mask is subjected to thresholding. The operation mask corresponding to a look-up table is changed or programmed by using the outer program light. The thresholded output pattern is normalized to become to the input format of the interconnection part. The similar operation is performed in the condition part, of

which result determines which output ports should be active or not. The formatted outputs from the operation part and the condition part is superimposed and thresholded. The output pattern is a three-state pattern, in which each one of three parts is bright or not. The output pattern is projected to the trapezoidal prism to separate the interconnention direction of the output. For example, if the central part is bright, the central output port is active and so the output goes straight. The position of the bright parts of the output pattern directly corresponds to the output port.

By using the programmable and ternary functions of the proposed gate array, we have implemented a MSD adder for two 4-digit numbers and a MSD multiplier. Figure 4 shows an prototype of an hybrid element cell consisting of LEDs and phototransistors. The gate delay of the experimental cell is about 100ms. We have simulated the MSD adder by the prototype system and a binary full adder. In the case of binary circuit, the interconnection patterns and the functions in element cells are determined by using a PLA CAD tool.

Analyses and experimental results suggest the proposed programmable logic gate array can yield significant advantages in terms of dynamic and reconfigurable interconnention, system efficiency and systematic design approach.

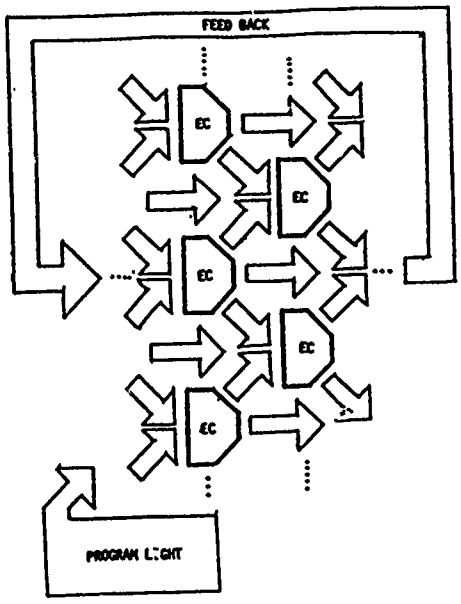


Fig. 1 Reconfigurable net using programmable gate array.

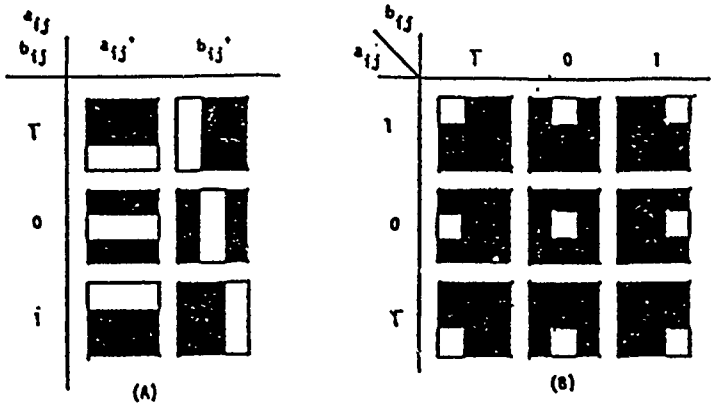


Fig. 2 Spatial coding of three-state logic.

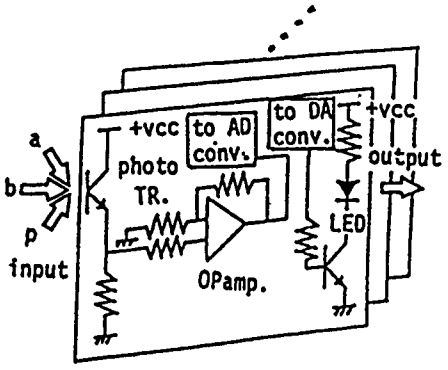


Fig. 4 Prototype of programmbale gate.

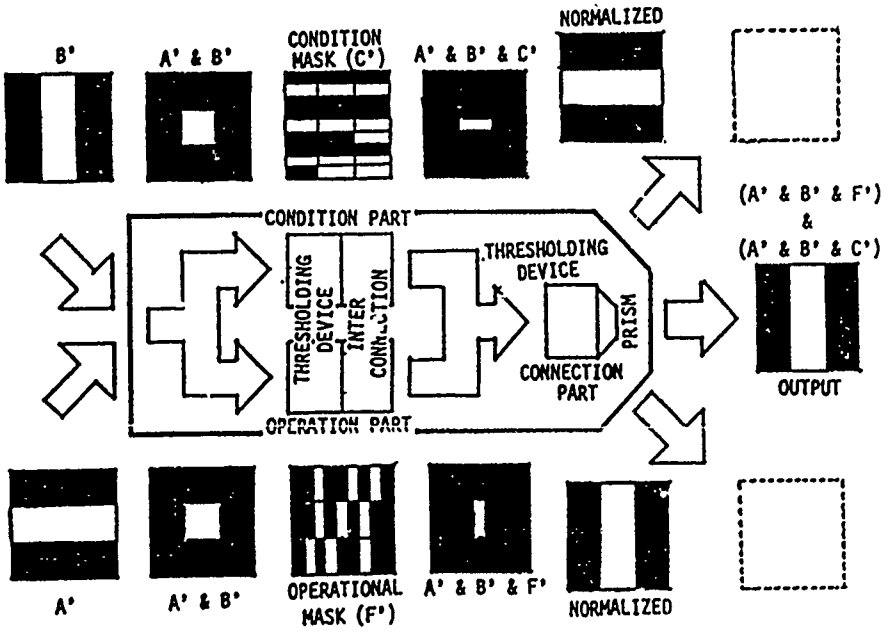


Fig. 3 Element cell of programmable gate.

An Optical Programmable Binary Symmetric Logic Module

Yao Li, Berlin Ha, and George Eichmann

Department of Electrical Engineering,
The City College of the City University of New York,
New York, New York 10031.

Binary symmetric logic function (BSLF)[1], because of its invariance under the permutation of its input variables, is an important class of Boolean logic function. The diversified BSLF applications include the synthesis of binary full adder and subtractor, the binary text comparator, the median filter, the parity checker, various threshold elements, etc.. The classical BSLF realization uses an array of regularly interconnected slow-speed electric contact network. Optical switches, because of their pico- or femtosecond switching capability, are excellent candidates for an optical BSLF (OBSLF) implementation. In this paper, various OBSLF architectures together with their applications to optical digital and symbolic computing, data communication, image processing, as well as neural networks, are described.

A switching function of n variables $f(x_1, x_2, \cdots, x_n)$ is called symmetric if and only if it is invariant under any permutation of its variables[1]. For exampie,

$$f(x, y, z) = \bar{x}\bar{y}z + x\bar{y}\bar{z} + \bar{x}y\bar{z} \quad (1)$$

where $\bar{}$ denotes the logic complement, is symmetric with respect to x, y, and z, while the function

$$g(x, y, z) = \bar{x}y\bar{z} + xyz + x\bar{y}\bar{z} \quad (2)$$

is not symmetric with respect to x, y, and z, but it is completely symmetric in x, y and $\bar{z}$.

It has been shown that a necessary and sufficient condition[1] for a function $f(x_1, x_2, \cdots, x_n)$ to be symmetric is that it can be reexpressed in the form of $S_{a_1 a_2 \cdots a_k}(x_1, x_2, \cdots, x_n)$, where $a_i = i$ with $i \in (0, 1, \ldots, n)$ are called a-numbers, such that when and only when a_i of the n variables are equal to 1, the function assumes the value 1. Using this definition. Eqs.(1) and (2) can be reexpressed as $S_1(x, y, z)$ and $S_2(x, y, \bar{z})$, respectiveiy. Since the number of its logic product terms corresponding to each a-number is

$$k = \frac{n!}{(n-a)!\,a!}, \quad (3)$$

to synthesize a n-variable BSLF, each a-number output channel must generate a logic OR function of its corresponding k product terms.

In Fig.1, an optical device that generates a BSLF (for $n = 4$) is shown. It consists of a triangular array of 50/50 splitting ratio beamsplitters and optical on/off switches. Because for a BSLF the optically switched signal must reach one of the possible output ports, all of the input optical energy is ultimately used. In order to generate, for each BSLF, the k required logic product terms, for the logic variables (their complements) on/off switches are inserted into the horizontal (vertical) sections. It can be

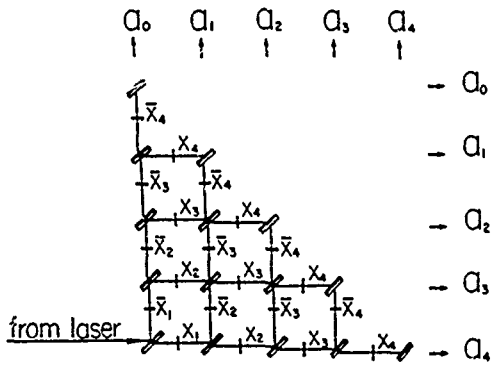


Fig.1 A free-space optical BSLF implementation. The beamsplitters form a triangular array with optical on-off switches forming its branches. The switches in the horizontal (vertical) branches are activated by the logic (complement) variables. Two output directions can be used. The a_i indicates the channel that generates, in terms of i input variables, the symmetric logic outputs.

own that to process n logic variables, beamsplitters and q optical on/off respectively, where

$$\sum_{1}^{+1} i \text{ , and } q = \sum_{i=1}^{n} 2i. \quad (4)$$

need to be employed. In order to reduce the number of optical elements, and to achieve a more compact geometry, in Fig.2, a guided-wave-optics-based approach to implement a BSLF (n =4), is shown. For the n logic variables, this device uses an array of p' waveguide directional couplers and q' waveguide Y junctions, respectively, where

$$p' = \sum_{i=1}^{n} i \text{ , and } q' = \sum_{i=1}^{n-1} i \text{ .} \quad (5)$$

Both integrated optical couplers and Y junctions are available[2]. When the directional switch is activated, the input signal is guided into one of the two output channel. An additional advantage of this this approach is that to control the array only the logic variables, and not its complements, are needed. With both free-space and guided-wave imlementation approach, an optical programmable binary symmetric logic module (OPBSLM) can be realized. To accomplish this task, an additional optical spatial light modulator (SLM) that can be programmed to select the a —number output channels is utilized.

There are potential diversified applications of an OPBSLM. For example, consider the design of a binary optical full-adder. Since a large number of arithmetic computations consist of binary additions and multiplications, where the multiplication is performed through the formation of partial products followed by either a series or a tree-structured parallel additions, a fast optical adder is a necessary building block for an optical computer. It can be shown that the bit-wise full adder's sum S_i and carry output C_i are $S_{1,3}(A_i, B_i, C_{i-1})$ and $S_{2,3}(A_i, B_i, C_{i-1})$. Using an array of three-input OPBSLMs together with two spatial light modulators (SLMs) and two cylindrical lenses that select and sum the a —number outputs (see Fig.3), an array of

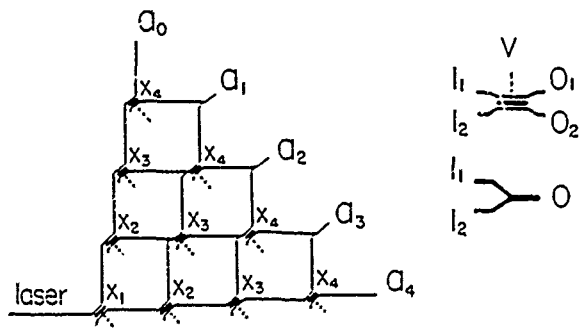


Fig.2 A guided-wave optical BSLF implementation. Upon the activation of the control signal, the coupler routes the input signal to either one of the output channels.

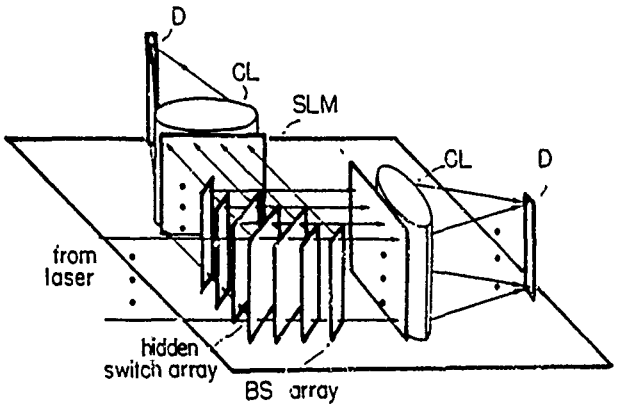


Fig.3 A schematic for a three-variable 8-bit experimental OPBSLM setup. *BS* , beam splitter; *CL* , cylindrical lens; *D* , detector; SLM, spatial light modulator. In addition, into every branches, a binary switching array is inserted. After passing through a SLM programmed to select the required a −numbers, the array generates two outputs. Finally, using the CLs, the selected a -number outputs are summed.

sum and carry bits can be optically generated. Here, the SLMs are programmed to select only a_1 and a_3 channels for the sum and a_2 and a_3 channels for the carry, and to block all the other outputs. It can be shown that for a guided-wave OPBSLM-based 1-bit adder, only six active couplers are needed, while using a conventional Boolean logic XOR and AND gates, seven active elements must be employed. Thus, for some applications, the OPBSLM-based approach is more energy efficient.

As a second OPBSLM example, a programmable optical binary weighted threshold summer can be implemented. A threshold summer nas n -input ports and one output port. For a particular threshold level, say j , when more than j (regardless of their order) of the n inputs are present, the summer will generate an ouput one. Otherwise, its output is a zero. For an OPBSLM implementation, after generating all a -number channels in parallel, a SLM mask that passes $a_{j+1}, \cdots a_n$ while blocking $a_0, \cdots a_j$ channels are used. For a final summing result, the selected (thresholded) channels then pass through a lens. A direct application of the optical binary weighted threshold element is in optical symbolic substitution (OSS)[3] and in optical neural network (ONN)[4]. For an OSS operation, to produce an intermediate result, after copying, spatial shifting, and combining the input pattern, an array of optical binary weighted threshold elements is used. While for an ONN, in addition to employing massively parallel interconnect channels, a large number of programmable threshold elements are needed. Because the OPBSLM-based threshold gate does not directly count the summed input power, one advantage is its low error accumulation rate. Additionally, it is fully programmable and its threshold performance does not change with different threshold levels. Other OPBSLM applications can be found in optical parity check for optical communication, optical text comparison for data processing, optical median filtering for image processing, etc..

To demonstrate the OPBSLM's operational principle, a three-variable beamsplitter type OPBSLM was experimentally constructed. As the source, a 4 w Ar ion laser was employed. For our experiment, the laser beam was spatially filtered, expanded and then masked to provide an approximately 2 mw power for each of the eight input channels. On a special breadboard, seven antireflectively coated 50/50 splitting ratio beamsplitters were mounted. To provide an exact beam position match at the output, a 3D beamsplitter adjustment was performed. Also, between every two beamsplitters, an identical spacing of 1 cm was used. In our proof-of-principle experiments, for the switching array, an assembly of binary masks were used. In the first experiment, an 8-bit (vertically coded) binary full-adder array was synthesized. The multidigit binary full-addition is an iterative process where initially the carry is set to zero. After the initial iteration, a non-zero carry string is generated. In our example, assuming, in the

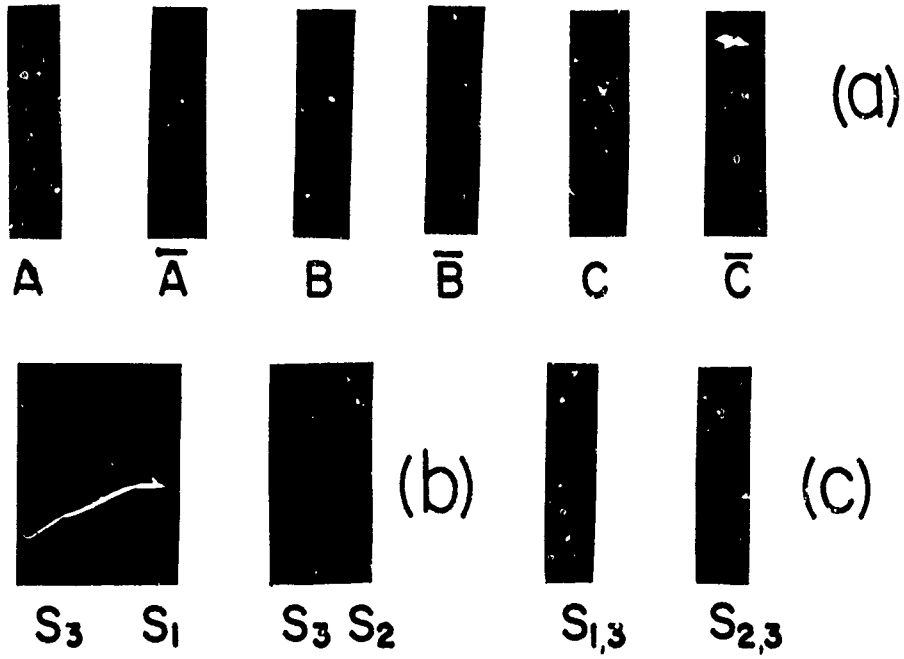


Fig.4 Proof-of-principle experimental result for a 8-bit binary full adder array. (a), Mask outputs representing the logic inputs and their complements for the three variables (vertically coded) A, B, and C_{in} data array. (b) The BSLF results S_1, S_3 and S_2, S_3 for the sum and carry outputs generated past the two selection plane. (c) the final summation results of (b) obtained at the cylindrical lens focal plane.

middle of the multi-iteration full addition operation, the addition of $A = 11001011$ and $B = 01101101$ and the carry-in $C_{in} = 01001101$ needs to be performed. The insertion of an array of binary switching masks represents the three numbers and their logic complements (see Fig.4(a)), Using two additional masks that select the $S_{1,3}(A, B, C_{in})$ and $S_{2,3}(A, B, C_{in})$ for the two outputs, in Fig.4(b) selected output patterns are displayed. Finally, at the focal planes of two cylindrical lenses, the 8-bit optical full addition sum (S) and carry (C_{out}) outputs, where $S = 11101011$ and $C_{out} = 01001101$, are obtained (see Fig.4(c)).

For a practical implementation of a OPBSLM, a number of performance factors need to considered. To satisfy real-time computation requirements, both switching speed and efficiency of the device must be considered. In addition, for the proposed free-space approach, diffraction through a gapped array of switches and beamsplitters may result in signal spatial channel broadening leading to energy loss and cross-talk. Also, for a guided-wave based approach, the signal loss due to waveguide branching and coupling efficiencies may limit the ultimate signal propagation length affecting the size of the array. Practical considerations of these implementation factors will be detailed in this paper.

This work was supported in part by a grant from the U.S. Air Force Office of Scientific Research.

References

[1] Z. Kohavi, *Switching and Finite Automata Theory* McGraw-Hill, ch.6, New York, (1978).

[2] Y. Suematsu and S. Arai, "Integrated Optics Approach for Advanced Semiconductor Lasers," Proc. IEEE, 75, 1472 (1987).

[3] K. H. Brenner, A. Huang and N. Streibl, "Digital Optical Computing with Symbolic Substitution," Appl. Opt. 25, 3054 (1986).

[4] N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical Implementation of Hopfield Model," Appl. Opt. 24, 1469 (1985).

Optical realization of arithmetic operations in ternary number system

E.M. Dianov, A.A. Kuznetsov, S.M. Nefjodov, G.G. Voevodkin

Academy of sciencies of the USSR

General Physics Institute

Vavilov street, 38 117942 Moscow, USSR

The main advantage of the ternary number system over the other ones is its the promise of higher efficiency.

For the optical realization of the main arithmetical operations (addition and multiplication) we have used a liquid-crystal light valve. It is clear, that the main difficulty in the addition is the transformation of the sum equaling 3 or 4 (in the decimal number system) into the ternary form (10 and 11) and the transfer from the lower order digit into the higher order digit. The addition was optically realized in the following way (Fig.1). Two matrices A and B with a set of input data, presented in a ternary number system (0, 1, 2), were imaged on a photosensitive layer of modulators T_1 and T_2, where the intensity summarizing of these matrices occured in the corresponding digits. A small part of the summary light flow was carried to T_1.

At the read-out from the modulator T_1 the summary intensity breaks into two groups of values: 0, 1, 2 and 3, 4. It can be achieved by means of bichromatic read-out, using the feedback [1, 2], which provides threshold discrimination of intensity.

At the bichromatic read-out from the modulator T_1 by the light of the wavelengthes λ_1 (red beam) and λ_2 (green beam), it is possible to achieve such a regime, when at the intensities on the photolayer of 0, 1 and 2 (i.e. very low due to the weak energy transfer to the modulator T_1) only the radiation λ_1 (with practically similar intensity for all the three values) will be present in the read-out flow after the analyzer [3]. It can be explained by different birefringence, observed in liquid crystals for the two wavelengthes, providing orthogonal relation of polarizations for two radiations. For λ_2 a feedback circuit is realized (λ_1 is suppressed in it by the filter Φ), which changes the modulator over into the flip-flop regime. At the illumination of a photolayer by the external signal (the sum A + B), exceeding some threshold (in our case at $I > 3$), the avalanche-type increase of the signal is observed at the output of the modulator for the light with the wavelength λ_2 to some constant level. In this case the light intensity with λ_1 after the analyzer A approaches zero in the points of the increase of light with λ_2 due to the polarization orthogonality of two radiations.

Thus, after the analyzer one light intensity will correspond to the input signals 0, 1, 2 for λ_1 and 0 for λ_2, and to signals 3 and 4 - the similar intensity for λ_2 and 0 for λ_1.

Some part of the light with the wavelength λ_2 is transferred by a semi-transparent mirror for the further summarizing (which corresponds to the unit transfer for the input signals 3 and 4). Two-colour matrix is imaged as a whole on the modulator T_2 following the condition of automorphism. In this case only the input signals 0, 1 and 2 are read by the light with the wavelength λ_1, which provides identical output signals 0, 1, 2. At λ_2 only the input signals 3 and 4 are read; the minimum of the modulator T_2 characteristic for λ_2 is tuned in signal 3. Then the read-out of 3 and 4 will give at the output 0 and 1, which corresponds to the reminants in the digit under consideration in the ternary number system.

Optical summarizing of the result of read-out from T_2 at λ_1, λ_2 with the result of read-out from T_1 at λ_2 (with the transfer) gives the result in the ternary number system. If considering the results of the transfer, it is possible to obtain once more the intensities 3 and 4; then in order to provide further transfer it is necessary to realize a sequential switch-on of a few modulators.

Note here, without explanation, that multiplication realization is possible on the base of described procedures.

The second considered way of addition is based on the method of symbol substitution [4] and binary presentation of the ternary symbols (Fig. 2). The input matrix is imaged on the photosensitive layers of two LCLV. At the bichromatic read-out after the first modulator the input matrix is transformed into a two-colour one, and the read-out of the shifted by one step image of the input matrix on the photosensitive layer of the second modulator is realized. After the Wollaston prism (which is the output analyzer) we shall obtain two matrices of different colours. Note, that for the matrix of one polarization red registers correspond to the realization of the logical operation AND, and green ones - of NOR. For the orthogonal polarization the operation $\bar{A} \cdot B$ corresponds to the red registers and $A \cdot \bar{B}$ - to green ones. It means the recognition of the following combinations:

$$\begin{matrix}1\\1\end{matrix} \rightarrow 1 \qquad \begin{matrix}1\\0\end{matrix} \rightarrow 1 \qquad \begin{matrix}0\\1\end{matrix} \rightarrow 1 \qquad \begin{matrix}0\\0\end{matrix} \rightarrow 1 \quad .$$

Here 1 corresponds to the presence of light (red or green).

We should point out the simultaneous recognition of all four possible combinations. Thus, this polarization-colour coding makes wider the area of application of the polarization coding

method, proposed in [5], and allows to reduce the number of channels used for the recognition operation. It is clear, that after the Wollaston prism the red-green matrix with the light of one polarization will correspond to 0 and 2, and the other matrix, corresponding to 1 - to the orthogonal polarization. Note, that one 1 will be green and the other 1 - red, i.e. the simultaneous recognition of all the values of the ternary system will occur. Their energetic equivalence should be also taken into consideration.

Thus, three matrices can be formed simultaneously from 0, 1 and 2. For the addition operation the recognition of 9 possible combinations is needed as well as the realization of 9 substitution laws (9 channels). If we use the polarization-colour coding, the number of necessary channels can be reduced to 4.

1. R.P. Akins, R.A. Athale, S.N. Lee. Feedback in analog and digital optical image processing - a review. Opt. Eng., v. 19, No. 3, p. 347 (1980).

2. G.G. Voevodkin, E.M. Dianov, A.A. Kuznetsov, S.M. Nefjodov, A.V. Parfjonov. Feedback in Devices with Spatial Light Modulator. Kvantovaya elektronika, v. 15, No. 4, p. 805-810 (1988).

3. G.G. Voevodkin, E.M. Dianov, A.A. Kuznetsov, S.M. Nefjodov. Optical realization of logic functions with LCLV. J. Tech. Physiki, v. 58, No. 10, p. 2082-2086 (1988).

4. A. Huang. Parallel Algorithms for Optical Digital Computers. Tech. Digest of IEEE Tenth Intern. Conf. on Optical Computing (1983), p. 13-17.

5. K.H. Brenner. New Implementation of Symbolic Substitution Logic. Appl. Opt., v. 25, No. 18 (1986).

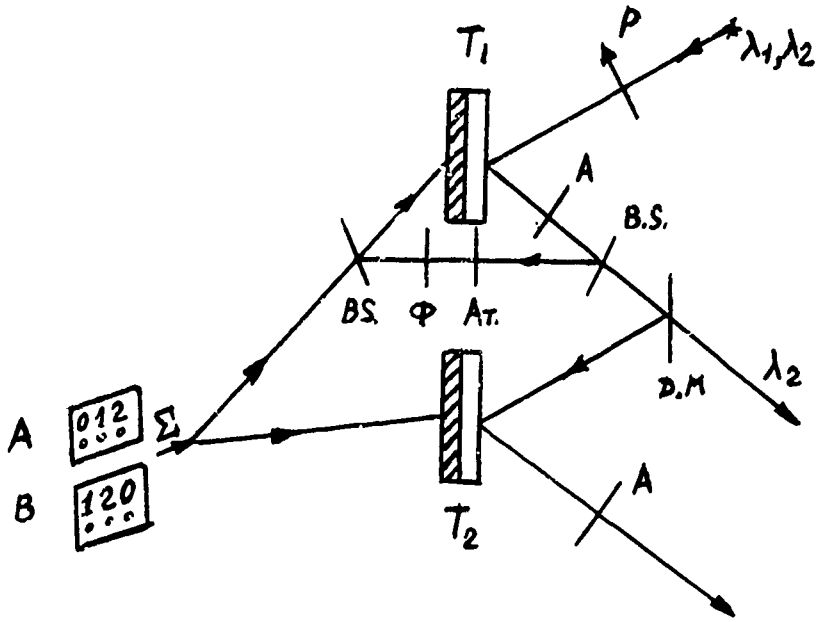

T1
P
λ1,λ2
A
B.S.
BS
Φ
Ат.
D.M
λ2
A
012
Σ1
B
120
T2
A

Fig. 1

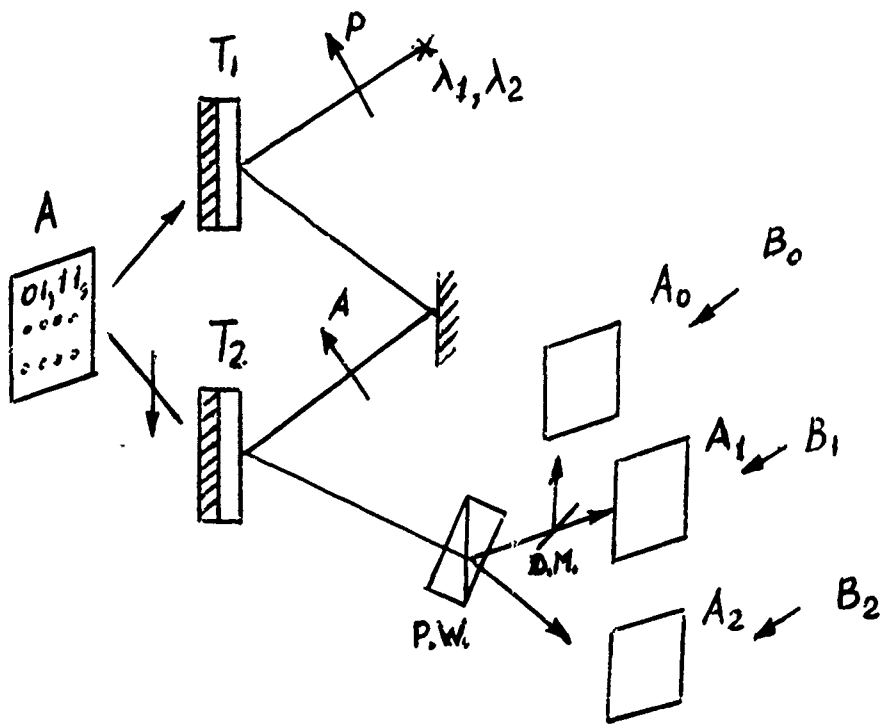

T1
P
λ1,λ2
A
01,11
T2
A
B0
A0
A1
B1
D.M.
P.W.
A2
B2

Fig. 2

NOTES

WEDNESDAY, MARCH 1, 1989

SALON F

10:30 AM–11:15 AM

WC1–WC2

DIGITAL OPTICAL COMPUTING: 3

Alexander A. Sawchuk, University of Southern California, *Presider*

Business and Technological Issues for the Commercialization of Optical Computing

By: Henry Kressel
E.M. Warburg, Pincus and Co.
466 Lexington Avenue
New York, NY 10016

The major technological elements encompassed by optical computing will be discussed in terms of their applications. Comparisons with the successful commercial introduction of other optical technologies will be made in order to highlight the elements which contributed to their success.

Optical computing will be analyzed in term of four major technological areas: analog and digital signal processing, interconnections and large scale memory. Each area offers potentially unique advantages relative to other technological approaches in certain applications. However, experience in the introduction of other technologies has shown that many factors need to be considered before a technology becomes imbedded in widely used products. Major considerations include:

1. The degree to which the new technology leads to products which perform unique and valuable functions.

2. The quality, reliability and economy of competing approaches and their power requirements, size and ease of use.

3. The perceived risk by users of switching over to a new technology.

4. The size of the investment needed by manufacturers to introduce a new technology and the perceived market size and return on investment. A major factor here is the expected life of a new product which is impacted by competitive technological evolution.

Selected applications of optical technology in computing have already demonstrated their economic viability. Other applications will mature as major hurdles are overcome and the need for new approaches become evident because existing ones fail to meet the needs of users. These will be discussed and analyzed in comparison with other optical technologies which have matured in the past decade.

All-optical Full-adder Based on Zinc Sulphide Optical Bistable Device

Wang Ruibo Zha Zizhong Zhang Lei Li Chunfei
Department of Physics, Harbin Institute of Technology
Harbin, People's Republic of China

Summary

1. Introduction

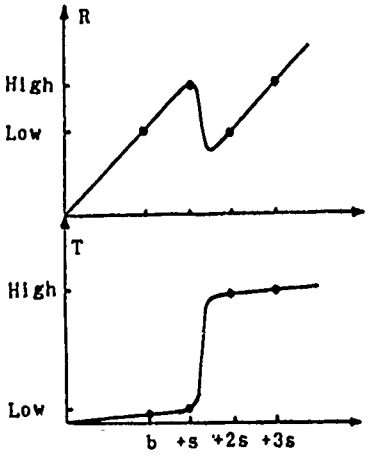
Optical computing systems constructed with relatively slow logic devices and massively parallel configurations could have very high processing rates. For this reason, great attention has been paid on circuits based on lower power, moderate speed, nonlinear interference filter logic devices. Single-gate full-adder was proposed and realized by B.S.Wherrett et al[1,2]. In their experiment, input signals were input with an incident angle. Here we report the experimental demonstration of a single-gate full-adder with on-axis input and put forward a design of multi-bit full-adder.

2. Single-gate full adder

One advantage of nonlinear etalons is the simultaneous presence of two responses -- the transmission and the reflection. These responses are almost complementary, which is unique to optical logic. On condition that the transfer characteristic and input signal levels are appropriate, the transmission and the reflection can be used in a three input mode to represent CARRY and SUM of a full-addition.

Fig.1. shows the idealized characteristics for an

interference filter in a single-gate full-adder. While the input power level is at b, b+s, b+2s or b+3s, the transmission is low, low, high, high, and the reflection is low, high, low, high. It is clear that the reflection and transmission responses satisfy the requirement of a full-addition. Here b and s represent bias power and signal power.



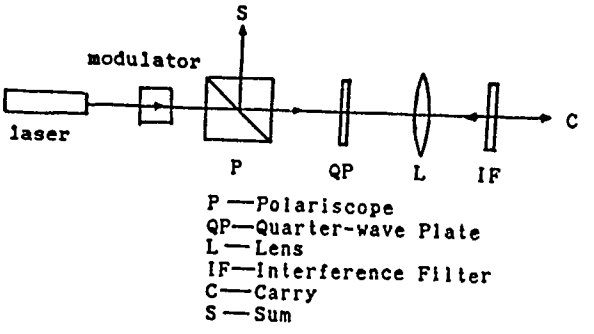


Fig.1 Transfer characteristic for a nonlinear etalon

Fig.2 Experimental set-up

In our experiment, a polariscope and a quarter-wave plate are employed to take out the reflection beam. The experimental set-up and result obtained with a 513.0nm ZnS interference filter are shown in Fig2. and Fig.3.

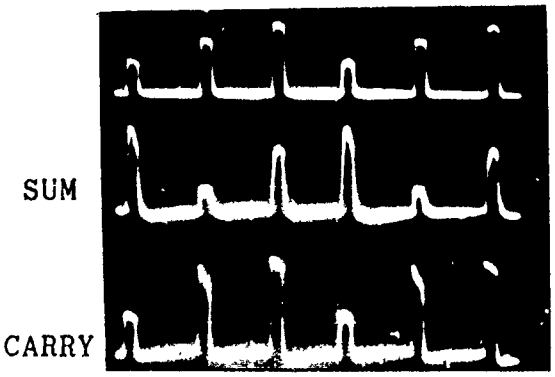


Fig3. Waveforms of input, SUM and CARRY

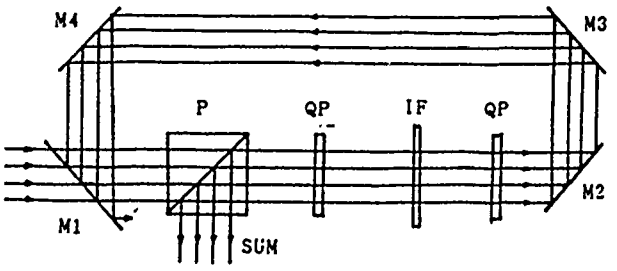


Fig4. Schematic diagram of a multi-bit full-adder

3. Multi-bit number addition

As shown in Fig.4, multi-bit full-adder can be constructed by

combining the gates in series with the CARRY beam. The CARRY signal from one gate is incident to the adjacent one by positioning mirror M4 properly. To ensure that the intensity of the CARRY signal equals the intensity of the input signal, the reflectivity of mirror M1 is dependent on the essential losses in the loop. Moreover, to avoid the wrong operation caused by the discrimination of the CARRY signals, it is necessary for the filter to have a high switch contrast and slight slopes on the branches of the transfer characteristic curve. Otherwise, another filter will be needed to standardise the CARRY signals.

4. Conclusions

Nowadays filters with high switch contrast (20:1) and 50% on-resonance transmission are available [3]. Although that transmission is not satisfactory, and other difficulties still exist, the rapid advances in fabrication of filter and other optical elements indicate that the optical computing circuit should be reliable in the near future.

References

1. B.S.Wherrett, Opt. Comm.,56,87,1985.
2. F.A.P.Tooley, N.C.Craft, S.D.Smith and B.S.Wherrett, Opt.Comm. 63,365,1987.
3. A.C.Walker, M.R.Taghizadeh et al, Opt.Eng. 27, 38, 1988.

NOTES

WEDNESDAY, MARCH 1, 1989

SALON F

11:15 AM–12:30 PM

WD1–WD5

MATRIX ALGEBRAIC PROCESSING

William T. Rhodes, Georgia Institute of Technology,
Presider

An Electro-Optical Architecture for Solving General Sparse Linear Systems[1]

M. Mary Eshaghian, V. K. Prasanna Kumar, and David W. Tang

SAL-344, Department of EE-Systems
University of Southern California
Los Angeles, CA 90089-0781, USA

1 Introduction

Many engineering problems involve the solution of a set of linear equations at some stage in the analysis. Many of these problems can be characterized by sparsity, that is the associated coefficient matrix contains a large proportion of zero elements. Examples include electric power system analysis, structural analysis, image processing, etc [1].

There are two general classes of methods for solving sparse linear systems, direct and iterative. Direct methods usually involve matrix factorizations and lead to additional nonzero elements being created during the computation. These fill-in elements cause extra required storage and an increase in computation time. On the other hand, Iterative methods preserve the sparsity of the matrix during computation and reduce the problem into some simple iterations of matrix-vector multiplications. They are preferred for solving large sparse systems because they can take advantage of zeros in the matrix and tend to be self-correcting and hence tend to minimize roundoff error.

When designing a special purpose sparse system solver, some important issues must be carefully considered. The architecture must be efficient for sparse matrix-vector multiplications and must realize any abitrary iterative matrix structure. Moreover, the structure of the iterative matrix is fixed throughout the computation. Thus, the expensive crossbar networks may be quite wasteful for this kind of computations. Even though electro-optical arrays have been designed for dense matrix computations and sparse banded matrix computations, no architectures are known parallel solution to sparse linear systems in which the nonzeroes are arbitrarily distributed.

Codenotti and Romani [2] presented a modular VLSI structure which is capable of reconfiguring for different problem sizes. The main components in their design are an array of m PEs and a mesh of switch nodes where the switch nodes are configured according to the particular structure of the coefficient matrix. The time required for one iteration is $O(m)$ for solving a set of m equations.

In this paper we present a new and efficient implementation of the iterative solution of general sparse linear systems by utilizing a regular array of VLSI implementable PEs communicating using optical beams in free space. Our design is very flexible; any iterative matrix structure can be realized by the use of holograms. Although the reconfiguration time for the hologram can be in the order of seconds in the current technology, it only needs to be done once in the preprocessing phase in which the structure of the coefficient matrix is used to define the holographic connections. The interconnection pattern remains the same throughout the computation. An optimal $O(\log m)$ time can be achieved by this design and the number of processors depends only on the number of non-zero elements in the matrix. This method is attractive when many computations are to be performed in which the structure of the coefficient matrix is fixed. It is well suited for implementation of many iterative methods such as Gauss-Jordan, Gauss-Siedel and the Conjugate method [1].

[1]This research was supported in part by the National Science Foundation under grant IRI-8710836 and by a Hughes Fellowship.

2 Proposed Architecture

We first present an abstract model of computation which closely captures currently implementable optical network of processors. This model enables us to analyze the optimality of the physical implementations in solving many problems [3].

Definition 1 *An optical model of computation represents a network of N processors each associated with a deflecting unit capable of establishing direct optical connection to any other processor. The basic assumptions used in our model are:*

1. *The processing layer consists of processing elements and I/O components. Each processor requires one unit of area. A processor can compute a simple arithmetic operation in one unit of time.*
2. *The deflecting layer is made up from a collection of deflector units, one for each processor. Each deflector unit takes one unit of area and is capable of redirecting an incident beam in one unit of time. The deflecting layer can be configured to realize any arbitray permutations.*
3. *The intercommunication is done through free space optical beams. An optical beam carries a constant amount of information in one unit of time, independent of the distance to be covered.*

We assume that the N processors are placed on the grid points in a $N^{1/2} \times N^{1/2}$ processing layer and the intercommunication beams are sent in the free space between the processing layer and the deflecting layer which is located directly above it.

This model has the full capability of a crossbar network where any processor can communicate with any other processor in one unit of time. However, it is still not feasible to design fully reconfigurable, fast, optical interconnects. Among the existing 2D spatial light modulators holograms offer a versitile way of interconnecting devices. They are not only simple to use and capable of being reconfigured to any arbitrary patterns (in order of seconds), they can also realize large amount of interconnections in a relatively small area. For the above reasons the hologram is the best choice for the implementation of this optical model for the sparse linear system applications. With holograms a processor can broadcast data to a constant number of other processors. For the sake of simplicity, we allow processors to broadcast to only two other processors. The hologram must be configured only once in the preprocessing phase and remains the same throughout the computation.

Laser diodes and light detectors are used to originate and detect light beams. Each processing unit requires two laser diodes, one detector to communicate with other processing units. The laser diodes are used to route the data to different destinations in two different steps of the algorithm which will be presented in the next section. The light detector acts as a synchronizer where upon receiving the data it signals the processor to start the computation and transmit data to other processors. Here, we restrict the processors to receive data from one other processor at a time. If optoelectronic components can be successfully integrated monolithically with high speed electronics, our proposed architecture can be directly implemented as presented in the abstract model. The main advantage is that the optical access it provides to data sources is interior to the chip rather than requiring that data be routed to the edge of the chip before being converted into optical form. However, this monolithic approach is still in a premature stage. A few research groups have demonstrated some successes of integrating optoelectronic devices with complex gallium arsenide circuits. So far only small scale devices have been built [4].

A more realizable scheme is shown in figure 1 where GaAs chips with optical sources are connected in a hybrid fashion to a Si chip with the wire bond technique. The Si chip will have

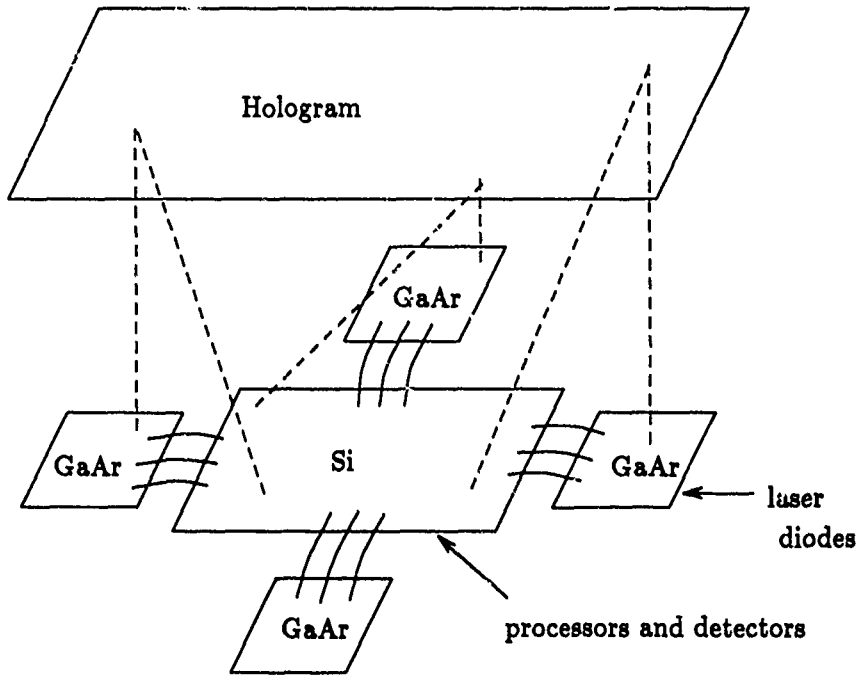


Figure 1: Physical Design of the Electro-Optical Architecture

detectors to receive the optical signals generated by the sources. Here data have to be routed to the laser diodes through wires.

3 The Algorithm

Let us consider the iterative method

$$x^{k+1} = Mx^k + g$$

where M is sparse and nonsingular. Let n_i be the number of nonzero element in the ith row of M, and let $j_1, j_2, ..., j_{n_i}$ be the columns corresponding to these elements. Thus, the above equation can be rewritten as

$$x_i^{k+1} = \sum_{s=1}^{n_i} m_{ij_s} x_{j_s}^k + g_i.$$

Suppose there are n processors and each of the processors stores exactly one nonzero element in matrix M.

Theorem 1 *The proposed electro-optical architecture can solve each iteration of the iterative solution to any general sparse linear systems with m variables and n nonzero elements in $O(\log m)$ time.*

Proof: Consider the following steps for solving the sparse matrix vector multiplication:

1. Send x_i to all processors with elements in ith column.

2. Perform multiplication in each processor.

3. Sum up all the products from the processors with elements in the same row.

The broadcast of x_i in step 1 is accomplished by sending the data to the first two processors with elements in ith column and then from those two processors to the next four processors with elements in the same column. It can be trivially shown that this step takes at most $O(\log m)$ time. The summation step can be similarly done by summing the products in groups of two processors with elements in the same row and then to the rest of the products. Therefore, the matrix-vector multiplication can be performed in $O(\log m)$. Since the iterative matrix is nonsingular, the x_i elements and g_i elements are stored in any one processor with a non-zero element in the same row. This processor is always the last one to be routed in the summation step. Thus we don't need extra processors to store x and g elements and extra connections to route the result back. After each iteration, a norm $|x_i^{k+1} - x_i^k|$, for all i, must be checked for convergence. If the maximal acceptable error is reached or the maximum allowable number iterations has been exceeded, the iteration process halts. The interconnection patterns can be reconfigured for different inputs and different sizes as long as the number of nonzero elements doesn't exceed the number of processors.

4 Conclusion

In this paper we presented an electro-optical implementation of the iterative solution of general sparse linear systems. This design utilized the newly developed optical model to communicate among the processing elements where each one can talk to any other processor in one time unit. An optimal time of $O(\log m)$ is achieved for each iteration of the method. As a comparison Codenotti and Romani's design [2] takes $O(m)$ for one iteration. Our design is well suited for implementation of many iterative methods such as Gauss-Jordan, Gauss-Siedel and the Conjugate method. Futhermore, it can be reconfigured to any arbitrary pattern in the iterative matrix. Many other sparse matrix-vector multiplication based applications can be implemented using our architecture as a building block.

References

[1] D.J. Evans, Ed. Sparsity and its applications. Cambridge University Press, Cambridge/London, 1985.

[2] B. Codenotti, F. Romani. A compact and modular VLSI design for the solution of general sparse linear systems. Integration, the VLSI journal 5 (1986) 77-86.

[3] M.M. Eshaghian, and V.K. Prasanna-Kumar. Massively Parallel Architecture with Optical Interconnects. Optical Computing 88, A Topical Meeting of the International Commission of Optics, Toulon, France, 1988.

[4] L. Hutcheson and P. Haugen. Optical Interconnects replace hardwire. IEEE Spectrum, March 1987, 30-35.

Digital Optoelectronic Processor Array Architectures for Vector-Matrix Multiplication

Michael R. Feldman and Clark C. Guest
Dept. of Electrical & Computer Engineering, Mail Code, R-007
University of California, San Diego, La Jolla, Ca. 92093

I. Introduction

High speed computation of the product of a vector and a matrix is desirable for problems in neural networks, signal processing, artificial intelligence and many other applications. If the number of elements in the vector is N, and the number in the matrix is N^2, then a serial computer requires a time which grows at least as N^2 to complete the computation. Many processing elements (PE's) can be combined to form a processor array in order to decrease the computation time. Electrically connected processor arrays are effective for locally connected parallel processing networks[1]. A linear mesh can be used to perform the multiplication in time proportional to N [2]. However to compute the product in shorter time, more complicated schemes with non local connection lengths are required. For such networks, the interconnections themselves are often responsible for a large percentage of the computation time, power dissipation and silicon area. Optically interconnected processor arrays can be employed to reduce this communication bottleneck. Previous work has shown that optical interconnections have advantages in terms of power dissipation and communication bandwidth for sufficiently long interconnection paths[3], and in terms of area for sufficiently complex connection networks[4].

However, optical communication links impose certain requirements on processor array•interconnection networks. While electrical VLSI interconnects achieve highest performance for locally connected networks, optically interconnected processor arrays are most effective for networks which (1) have a high degree of spatial invariance[4,5] and (2) require a small number of transmitters and receivers per PE.

In this paper, we present networks that can be used to produce vector-matrix multiplication in time T with $O(1) < T < O(N^{1/2})$. (O(f(N)), defined formally in Ref. 6, indicates an upper bound on the asymptotic dependence of the growth rate on N.) Networks particularly well-suited for optical interconnects have been found by considering topologies that have space-invariant properties and require minimal numbers of transmitters and detectors.

II. VLSIO Processor Array Model Description

A VLSIO (Very Large Scale Integrated Optoelectronic) processor array consists of individual electronic PE's that are interconnected by optical beams. Each PE would be identical to an electrically connected PE except instead of containing bonding pads, and pad and link drivers, the PE would require one or more optoelectronic signal transmitters(a laser or light modulator) and one or more photodetectors[3,7]. Holograms and other passive optical elements may be used to interconnect the transmitters and detectors in the desired pattern. We consider an arrangement in which all the PE's reside in the same plane. They may all be integrated on a single chip or wafer, or reside on an ensemble of VLSI circuit chips, all bonded to a common substrate. We first describe fundamental lower bounds on the area occupied by the processor array circuitry, and on the time required to complete the computation. We then discuss interconnection networks that have area and time growth rates close to these minimum values. The following processor array model provided the basis for our calculation of these area and time growth rates.

We assume the electronic processing circuitry is divided into N identical PE's. Each PE contains M_D detectors and M_T transmitters, local memory, electronic logic elements and local intra-PE wired connections. The area and latency growth rates of each PE were based on the VLSI grid model of Ullman [6]. The grid model consists of wires laid out on the lines of a rectangular grid and circuit elements occurring at the grid points. We treat the transmitters and detectors as ordinary circuit elements. Wire delay is assumed negligible, while all circuit elements impose a single unit of time delay on the passage of signals. Communication within a PE occurs along grid line wires. Communication between PE's is exclusively limited to optical beams. The optical beam delay is assumed negligible (although delays associated with the transmitter and receivers are accounted for). The area cost associated with the optical communication links is given by the area growth rate of the optical components which can be determined from the VLSIO model described in Ref. 6.

Initially the j^{th} PE receives input vector element a_j. After the computation is completed it contains the output b_j, where

$$b_j = \sum_{i=1}^{N} a_i W_{ij} \qquad (1)$$

Each PE is identical, and thus each PE contains N matrix elements (W_{ij}'s), assuming the matrix elements are stored within the processor array.

III. Lower Bounds on Area and Time

A. AT Product Lower Bound

We define the area devoted to computation of an optically interconnected processor array, A, as the area of all the PE's plus the area of each optical element. Thus, the manufacturing cost is approximately proportional to A. We denote the time to solve the vector-matrix multiply problem by T.

For either optically or electrically interconnected processor arrays, since N^2 multiplications are required for a vector-matrix multiply,

$$AT = \Omega(N^2) \qquad (2)$$

where $\Omega(f(N))$, defined formally in Ref. 6, indicates a lower bound on the asymptotic dependence of the growth rate on N.

B. Memory Based Area Lower Bounds

Since N^2 matrix elements are required for the computation, $A = \Omega(N^2)$ if these elements are stored in the PE array. However, since memory storage is comparatively efficient and, for projected values of N, is not expected to be responsible for the majority of the cost, we have neglected the area occupied by the matrix elements in calculations of A. Note that the cost penalty associated with the matrix elements is the same for all architectures.

C. Communication Time Lower Bounds

The transmitters and detectors associated with each PE are responsible for a significant portion of the manufacturing cost of a VLSIO processor array. Limitations on the number of transmitters and detectors per PE impose lower bounds on the computation time, T. In this section we determine lower bounds on T as a function of the number of detectors per PE, M_D, the number of transmitters, M_T, and the total number of PE's, N.

Based on the model described in Section II, we can prove the following theorem by induction.

Theorem 1: If during the total computation time, T, each PE receives q inputs ($q < N$), each PE must transmit at least N/q signals (either inputs, matrix elements or partial sums) to other PE's.

Since each detector and transmitter can handle at most one signal per unit time a direct consequence of theorem 1 is

$$T \geq q/M_D + N/(M_T q) \qquad (3)$$

Eq. 3 is minimized for $q = q_{opt}$,

$$q_{opt} = (NM_D/M_T)^{1/2} \qquad (4)$$

which yields,

$$T \geq 2 \cdot N / (M_D\ M_T)^{1/2}, \qquad (5)$$

Eq. 5 gives a lower bound for vector-matrix multiply consistent with our model described above.

IV. Nested Crossbar

A. Topology Description

In this section we will describe a set of architectures that come to within a factor of log N of both the AT product and the communication time lower bounds discussed above, for time growth rates between O(1) and $O(N^{1/2})$. We have termed the interconnection networks for these processor arrays nested crossbar networks. A nested crossbar is characterized by its dimension, m, and base, b, where

$$b = N^{1/m} \qquad (6)$$

The nested crossbar topology is illustrated in Fig. 1 for several values of b, m and N. The connection topology of a 2-d nested crossbar is loosely given by the following algorithm: divide the N nodes into $N^{1/2}$ groups of $N^{1/2}$ nodes per group. Connect each group in a full crossbar pattern (fully connected) forming $N^{1/2}$ sub crossbars. These will be referred to as dimension 1 connections. Next connect the corresponding elements in different groups as a full crossbar, forming $N^{1/2}$ additional sub crossbars, denoted dimension 2 connections. For higher dimensional nested crossbars, begin with N/b groups of b nodes per group, and have m levels of sub crossbars, each level containing N/b sub crossbars with b elements in each sub crossbar.

The connection pattern can be described more formally by assigning to each node a label that is an m digit integer in base b. (Each digit ranges from 0 to b - 1.) If each node is assigned a distinct label, connecting any two nodes whose labels differ in exactly 1 digit forms the nested crossbar connection pattern. Connections between nodes whose labels differ in the i^{th} bit position are termed dimension i connections.

Note that a 1-dimensional nested crossbar is a fully connected network (full crossbar) and all base 2 nested crossbars are binary hypercubes. The number of directly connected neighbors of each node in an m-dimensional

nested crossbar is b.

B. Example: 2-Dimensional Nested Crossbar, Dedicated Links

For a 2-dimensional nested crossbar (note: this is not a hypercube, see Fig. 1) each processing node is assigned a 2 digit label, kl, where $0 < k < N^{1/2} - 1$, and $0 < l < N^{1/2} - 1$. Node xy is connected to all nodes with label xl by dimension 1 connections, and to nodes with labels ky by dimension 2 connections.

A 2-dimensional nested crossbar can compute a vector-matrix multiplication in O(log N) time, if each PE contains $N^{1/2}$ detectors and $N^{1/2} + 1$ transmitters. The dimension 1 connections are formed by a single dimension 1 transmitter with a fanout of $N^{1/2}$ associated with each PE. The dimension 2 connections are implemented with $N^{1/2}$ dimension 2 transmitters at each PE. Each detector has a fan-in of two: - one signal from a dimension 1 transmitter a one from a dimension 2 transmitter. (The algorithm is designed so that dimension 1 transmissions never occur the same communication cycle as dimension 2 transmissions.)

The algorithms for vector-matrix multiplication with the 2-dimensional nested crossbars can be readily described by using the base $N^{1/2}$ node labels as the indices for the inputs, outputs and matrix elements. Thus in base $N^{1/2}$ notation Eq. (1) becomes,

$$b_{xy} = \sum_{i=0}^{\sqrt{N}-1} \sum_{j=0}^{\sqrt{N}-1} a_j W_{ijxy} \qquad (7)$$

Initially each PE broadcasts its input to its dimension 1 connected neighbors by transmission with its dimension 1 optical signal transmitter. After 1 unit of time delay a PE labeled xy (denoted PE_{xy}) receives the vector elements, a_{xi}, $i = 0, \ldots, N^{1/2} - 1$, on its $N^{1/2}$ detectors. PE_{xy} then computes $N^{1/2}$ partial sums (in $\log N^{1/2}$ time), denoted P_{xjy}, for $j = 0, \ldots, N^{1/2} - 1$, where,

$$P_{xjy} = \sum_{i=0}^{\sqrt{N}-1} a_{xi} W_{xijy} \qquad (8)$$

During the second communication cycle each PE with label xy emits the above $N^{1/2}$ partial sums - one on each of its $N^{1/2}$ dimension 2 optical signal transmitters. PE_{xy} then receives $N^{1/2}$ partial sums, P_{kxy}, for $k = 0, \ldots N^{1/2} - 1$. By summing these terms together, each node computes its final result,

$$\sum_{k=0}^{\sqrt{N}-1} P_{kxy} = \sum_{i=0}^{\sqrt{N}-1} \sum_{j=0}^{\sqrt{N}-1} a_j W_{ijxy} = b_{xy} \qquad (9)$$

Thus the complete computation is performed with only 2 communication steps, and $2\log N^{1/2}$ internal time steps.

C. Summary of Performance of Nested Crossbar Architecture

In Table I we characterize the performance of several nested crossbar architectures in terms of area, number of communication time steps, t, number of optoelectronic transmitters, and communication-related power dissipation. The communication related power dissipation is defined as the maximum instantaneous optical power required by the transmitters to drive the detectors and is proportional to the largest number of detectors required to receive signals during any one unit time step [4]. Note that although both a 2-dimensional nested crossbar and a full crossbar can complete the computation in log N time, a 2-dimensional nested crossbar requires substantially less optical power.

For optical interconnects the asymptotic area growth is determined by the larger of two areas - the area of a CGH designed to perform the interconnects and the area of the electronic PE's. The area of the PE's allows for all computation necessary to complete the algorithm in time, T = O(t log N), where t is the number of communication time steps. Algorithms for 2-d, 3-d and 4-d nested crossbar and hypercube interconnection networks were developed and are included in Table I. Note that for all networks except the hypercube, time lower bounds based on the number of optical components(given by Eq. (5)) are achieved to within a factor of log N.

The CGH area growth rates were determined from the model described in Ref. 4. Shared link nested crossbars achieve minimum area growth rates through use of a space-invariant architecture for which A = O(NF). The values listed for dedicated links are based on the use of double pass basis set CGH[4], which has a growth rate of $O(NF^{3/2})$ for the nested crossbar architecture. Note that for each architecture, the area of the CGH grows at a smaller rate than the area of the corresponding circuitry. This is a consequence of the space-invariant properties of the connection topology.

Area lower bounds for electrically connected PE's, determined from the model described in Ref. 6, are also listed in Table I. Note that lower bounds on the area growth rates of electrically connected PE's are significantly larger than upper bounds on optically connected PE's. This is because the VLSI area growth rate is determined by the area of the wires interconnecting the processors, while the VLSIO bound is limited only by the area of the PE's themselves. Also for this reason many of the optically interconnected networks come to within a factor of log N of the AT lower bound (given by Eq. 1) while VLSI implementations do not.

In short, the nested crossbar architectures meet lower bounds on time and on AT products for given numbers of optical components. Table I indicates the tradeoffs suggested by Eqns. (2) and (5). As T decreases, A increases and/or the number of optical components increase and/or the power dissipation increases.

References

1. C.Seitz, IEEE Transactions on Computers, c-33, pp.1247-1265, Dec. 1984.
2. C. Mead and L. Conway, Introduction to VLSI Systems, Menlo Park, Calif., Addison-Wesley, pp. 271-276, 1980.
3. M. R. Feldman, S. C. Esener, C. C. Guest and Sing H. Lee, Appl. Opt., 27, pp. 1742-1751, May 1,1988.
4. M. R. Feldman, C. C. Guest, T. J. Drabik and S. C. Esener, "A Comparison Between Electrical and Free-Space Optical Interconnects for Processor Arrays based on Interconnect Density Capabilities," submitted to Appl. Opt.
5. B. K. Jenkins, et. al., Appl. Opt., 23, pp. 3465-3474, 1984.
6. J. D. Ullman, Computational Aspects of VLSI, Rockville, MD, Computer Science Press, Chap. 2, 1984.
7. J. W. Goodman, F. I. Leonberger, S. Y. Kung and R. A. Athale, Proc. IEEE, 72, pp. 850-866, 1984.

Connection Technology	Architecture	H-cube	Nested Crossbar 4d	3d	3d	2d	2d	full X-bar
	shared or dedicated links	shared	dedicated	shared	dedicated	shared	dedicated	shared
Optical	# time steps	$\sqrt{N}\log N$	$N^{1/4}$	$N^{1/3}$	$N^{1/6}$	$N^{1/4}$	1	1
	Area	$N^{3/2}$	$N^{7/4}\log N$	$N^{5/3}$	$N^{11/6}\log N$	$N^{7/4}$	$N^{2}\log N$	N^{2}
	# transmitters / node	1	$N^{1/4}$	1	$N^{1/3}$	1	$\sqrt{N}$	1
	power / N	1	$N^{1/4}$	$N^{1/3}$	$N^{1/3}$	$\sqrt{N}$	$\sqrt{N}$	N
Electrical	Area	N^{2}	$N^{5/2}$	$N^{7/3}$	$N^{8/3}$	$N^{5/2}$	N^{3}	N^{3}

Table 1

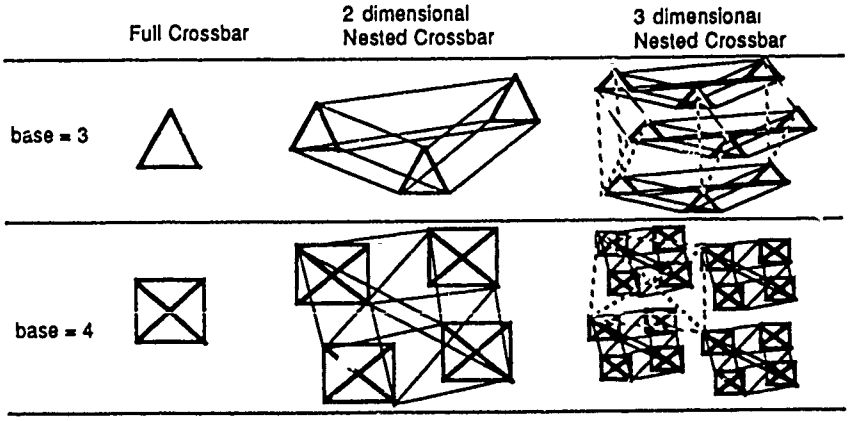


Fig. 1: Nested Crossbar topology. Not all connections shown for base 4, 2-d and 3-d and base 3, 3-d nested crossbars.

" Hybrid opto-electronic co-processor implementation inside a computer workstation "

Guy Lebreton (*), Rémy Frantz (**)

(*) GESSY, Université du Var, 639 Bd. des Armaris, 83100 TOULON, FRANCE
(**) CETIA (Compagnie Européenne des Techniques de l'Ingénierie Assistée), B.P.244, TOULON cedex, FRANCE

Introduction

The experimental realization of an optical processor is described. The goal is to implement in a computer workstation a linear algebra processor based on analog relaxation, to help the host computer solving three types of problems :
- solution to systems of linear equations (e.g. partial differential equations),
- matrix inversion,
- computation of eigenvalues and eigenvectors.

The theoretical approach is a synthesis of the previous ideas from various authors: space and frequency multiplexing in a multichannel acousto-optic Bragg cell with parallel throughput from an array of laser diodes; analog loop with continuous relaxation process rather than discrete iteration; and matrix preconditionning to operate with a moderate dynamic range.

The originality lies in both the optical implementation and the computer modeling of the processor. The experimental processor design should be achieved at the time of the meeting. The first draft on its realization, presented here, is derived from off-the shelf components end simulation results which enable to preview the attainable performances.

After a short description of the hybrid system, the results of preliminary simulations are summarized.

1. Description of the hybrid processor

The algorithm is the fully-parallel relaxation scheme proposed by W.K.Cheng and H.J.Caulfield [1], described by the differential equation

$$\frac{dx}{dt} = \frac{1}{\tau}\,(y - A.x), \quad \text{with} \quad x (t \to \infty) = A^{-1}. y$$

where the asymptotic solution exists only if the matrix A is real and symmetric, but is not restricted to the unit circle as in the discrete iteration case.
In the hybrid implementation of **Figure 1**, only the matrix-vector product A.x is performed optically in parallel. The three other operators are performed with electronic components.

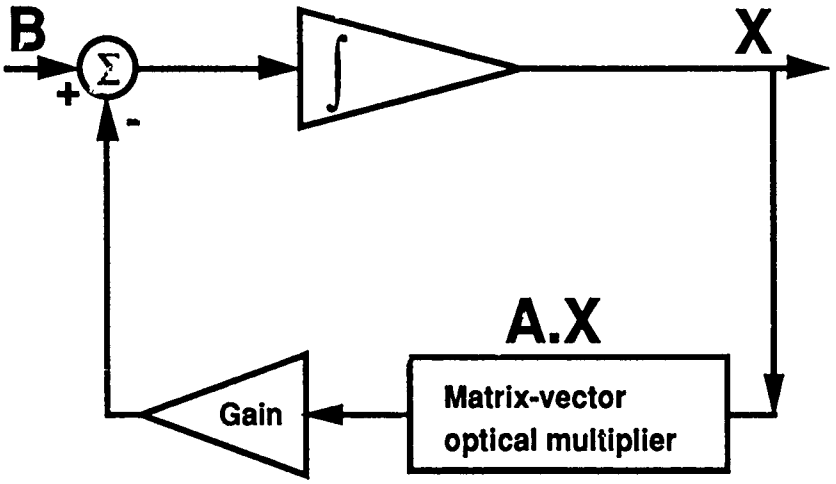


Figure 1 : hybrid analog relaxation

The vector-matrix optical multiplier is derived from the space and frequency multiplexing in a multichannel acousto-optic Bragg cell, as demonstrated already by D.Casasent and J.Jackson **[2]**. Two columns of laser diodes provide for the bipolar vector input. The matrix rows fill in parallel the acousto-optic modulator, with each pixel encoded as amplitude modulation on a separate frequency, directed on a related photo-detector by the Fourier Transform lens.
For the first experimental realization, only a 8x8 matrix will be implemented, but testing will be performed to evaluate the possibility to increase this number in further devices. The main difficulty is the minimum rate of partial coherence required to separate the channels in the optical Fourier Transform plane, already studied by B.Javidi **[3]** : experiments are in progress to utilize multimode optical fibers with connector-mounted laser diodes. This solution would enable the easy replacement of the laser sources in a well-engineered compact system. The overall system is being designed for a very compact implementation, to be included in a standard computer workstation.

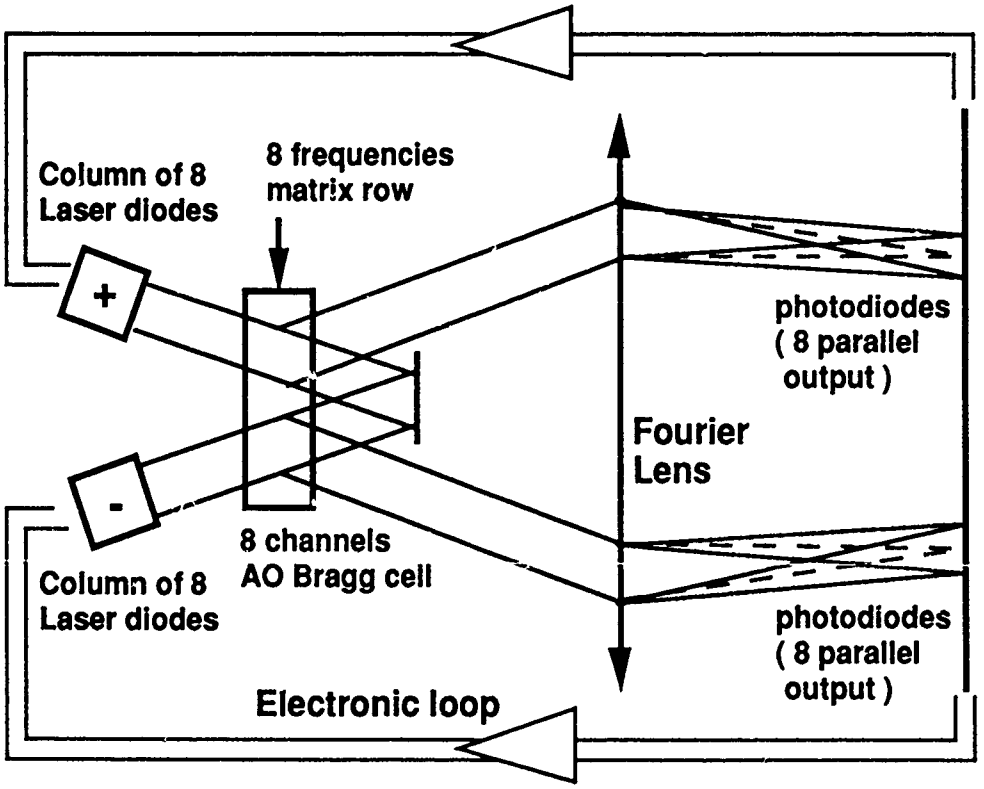


Figure 2 : optical vector-matrix multiplier

2. Computer modeling and simulated performances

The processor has been modeled accordingly to the studies of A.Ghosh, D.Casasent and P.Neuman **[4]**, but the impulse response and the bandwidth of the electronic components has been included, yielding somme additional constraints.
The laser noise appears to be of moderate influence, the overall accuracy in solving linear equations going from 1.5 to 2.5 when the laser noise increases from 0.5 % to 5 %, if the condition number of the matrix (ratio of the highest to the lowest eigenvalue) keeps under 50.
The performances simulated on Figures 3 to 5 are obtained with the time constant of the overall analog system (about 1.5 µs, slightly dependant from the condition number), multiplied by the number of operations required by the Jacobi algorithm to reach the precision of the optical analog computing. The speed of the analog processor is found to depend strongly on the condition number **(Figure 3)**, which is not the case for the digital computer.

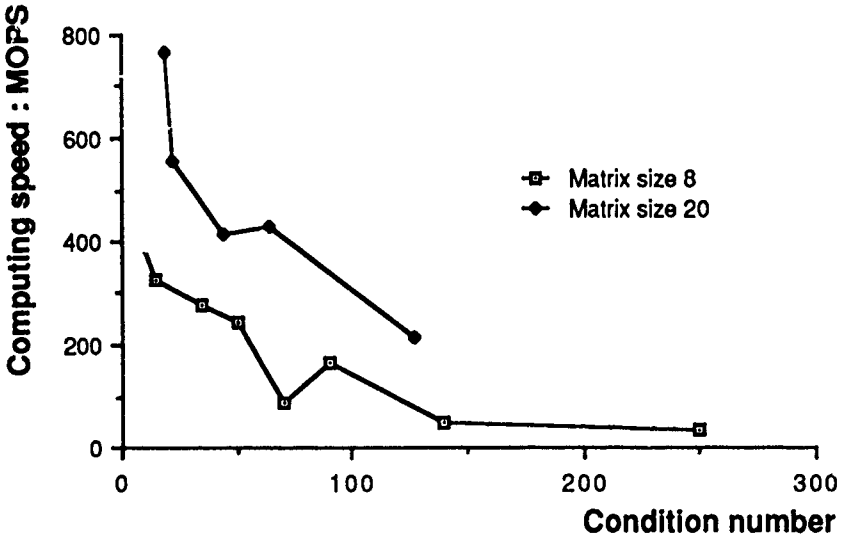


Figure 3 : optical processor speed versus the matrix condition number

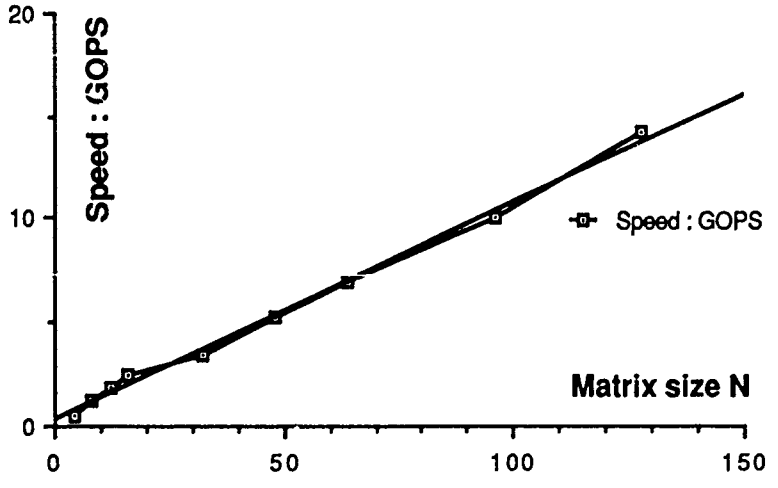


Figure 4 : optical processor speed for well-conditionned matrix

With a condition number between 2 and 3, the speed increases linearly with the size N of the parallel matrix **(Figure 4).** The overall acceleration factor for the host computer is shown on **Figure 5** for the extreme cases of a VAX 8650 (0.7 MFLOPS) and a CRAY X-MP4 (800 MFLOPS). It appears that the 8x8 optical processor is already efficient for small computers. To enhance this efficiency, pre-conditioning algorithms should be of interest, as shown by A.Ghosh and P Paparao **[5]**. To extend the application to large sizes of matrix, we are currently investigating attractive partitionning algorithms.

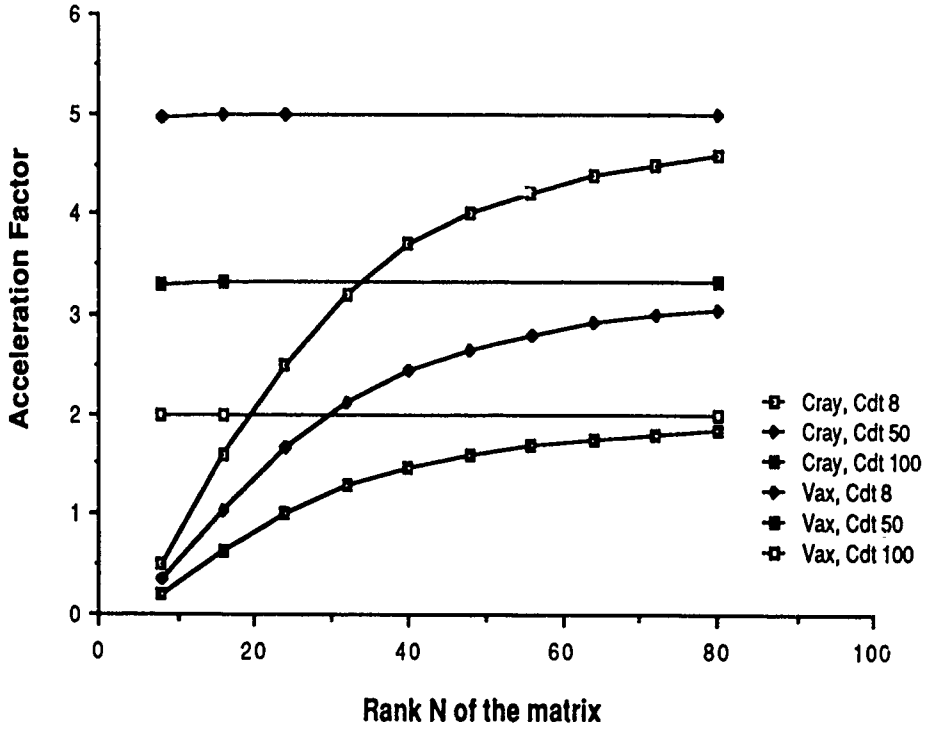


Figure 5 : acceleration factor with the optical processor in a host computer

The experimental realization is now in progress and some first results should be available at the time of the meeting. The implementation inside the work-station is expected to take one more year.

[1] W.K.Cheng, H.J.Caulfield, "Fully-parallel relaxation algebraic operations for optical computers", Opt.Com. 43/4, 15 Oct.1982, p. 251-254.

[2] D.Casasent, J.Jackson, "Space and frequency-multiplexed optical linear algebra processors: fabrication and initial tests", Appl.Opt. 25/14, 15 July 1986, p.2258-2263.

[3] B.Javidi, "Real-time joint transform correlation by partially coherent readout illumination", Appl.Opt. 26/18, 15 Sept.1987, p.3762-3771.

[4] A.Ghosh, D.Casasent and C.P.Neuman, "Performance of direct and iterative algorithms on an optical systolic processor", Appl.Opt. 24/22, 15 Nov. 1985, p. 3883-3892.

[5] A.Ghosh, p.Paparao, "Matrix preconditioning: a robust operation for optical linear algebra processors", Appl.OPt. 26/14, 15 July 1987, p.2734-2737.

Theoretical Description of the Bimodal Optical Computer

A V Scholtz and E van Rooyen

Carl and Emily Fuchs Institute for Micro Electronics
University of Pretoria, Pretoria, Rep. of South Africa, 0002

The use of optical processors for fast numerical calculations has always been hampered by the inherent inaccuracies of analog optics. Numerous techniques have been proposed to overcome this deficiency. One of these, the Bimodal Optical Computer (BOC), proposed by Caulfield et al[1], has generated considerable interest lately. The BOC (shown in fig.1) is a hybrid processor dedicated to solving a set of linear algebraic equations (LAE's). It consists of an iterative optical processor that uses continuous analog feedback, coupled to a digital processor to ensure high accuracy through a process of iterative refinement.

The analog feedback loop can be described by the following differential equation:

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{A}\mathbf{x} = \mathbf{b} \qquad (1)$$

where $\mathbf{A}$ is the known matrix, $\mathbf{b}$ is the output vector, $\mathbf{x}$ is the unknown solution vector, and t is measured in units of τ, the response time of the integrators. The solution to this set of non-homogeneous differential equations is

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + \int_0^t [e^{\mathbf{A}(t-\tau)}]\mathbf{b}\, d\tau. \qquad (2)$$

To illustrate the role of the eigenvalues of the matrix in the actual convergence of the solution vector as a function of time, Equation (2) was rewritten by Cheng and Caulfield[2] as

$$\mathbf{x}(t) = \sum_{i=1}^{N} [\, (\mathbf{v}_i \cdot \mathbf{x}_0)e^{-\lambda_i t} + \frac{\mathbf{v}_i \cdot \mathbf{b}}{\lambda_i}(1 - e^{-\lambda_i t}) \,]\, \mathbf{v}_i \qquad (3)$$

where $\mathbf{v}_i$ is the right-hand eigenvector (usually just called the eigenvector or RH eigenvector) associated with eigenvalue λ_i.

To derive this equation, Cheng and Caulfield assumed that the matrix had distinct eigenvalues, or alternatively that it can be reduced to diagonal form. But we have found that a second assumption must also be made, one that says that the matrix $\mathbf{Q}$, formed by the set of eigenvectors, must be orthogonal. This means that $\mathbf{Q}^T = \mathbf{Q}^{-1}$, which is an identity that is used to derive Equation (3). This is a very strict condition. If we do not assume that the matrix $\mathbf{Q}$ is orthogonal, Equation (3) can be rewritten as

$$\mathbf{x}(t) = \sum_{i=1}^{N} [\, (\mathbf{u}_i \cdot \mathbf{x}_0)e^{-\lambda_i t} + \frac{\mathbf{u}_i \cdot \mathbf{b}}{\lambda_i}(1 - e^{-\lambda_i t}) \,]\, \mathbf{v}_i \qquad (4)$$

where $\mathbf{u}_i$ is the left-hand (LH) eigenvector of $\mathbf{A}$ defined by $\mathbf{A}^T\mathbf{u}_i = \lambda_i\mathbf{u}_i$. It is possible to show that $\mathbf{R}^T = \mathbf{Q}^{-1}$, where $\mathbf{R}$ is the matrix formed by the set of $\mathbf{u}_i$. The only difference between Equation (3) and Equation (4) is the appearance of the LH eigenvectors. In both cases, the eigenvalues of the matrix $\mathbf{A}$ dictate the rate with which $\mathbf{x}(t)$ approaches the steady state solution.

To ensure that $\mathbf{A}$ can be reduced to diagonal form, it is required that $\mathbf{A}$ is Hermetian or real symmetric. The eigenvalues are therefore real, but not necessarily positive. To force the matrix to be positive definite, rather evaluate $\mathbf{A}^T\mathbf{A}$ (for real $\mathbf{A}$) or $\mathbf{A}^H\mathbf{A}$ (for complex $\mathbf{A}$). The result vector now becomes $\mathbf{A}^T\mathbf{y}$ or $\mathbf{A}^H\mathbf{y}$ instead of $\mathbf{y}$. This carries a penalty in the form of extra time needed to premultiply $\mathbf{A}$ with $\mathbf{A}^T$ digitally.

Simulations show that settling times of less than 1μs can be expected, independent of the size of the matrix. The convergence of the feedback loop for a 4×4 matrix is shown in fig. 2. Convergence is only assured if all eigenvalues are positive or have positive real parts, which implies that $\mathbf{A}$ must be a positive definite matrix. If that is true, the steady state solution is

$$\mathbf{x}(t\to\infty) = \sum_{i=1}^{N} \left[\frac{\mathbf{u}_i \cdot \mathbf{b}}{\lambda_i}\right] \mathbf{v}_i = \mathbf{A}^{-1}\mathbf{b}, \qquad (5)$$

which is the correct solution to the original set of linear equations.

This answer is usually not accurate enough, due to inaccuracies in the optical system. The main source of error is nonuniformity in the spatial light modulator, while nonlinearities and noise in the source and detector circuits can add significantly to the overall error. A more accurate answer is obtained with the help of a digital processor, which performs a process called iterative refinement. As can be seen in fig. 1, the answer obtained with the optical processor is read by the digital circuit via an A/D converter. A new set of LAE's is set up to calculate the difference between the correct solution and the one supplied by the optical processor. Scaling is used to keep the values within the dynamic range of the optical processor. Adding the calculated residue to the inaccurate solution, improves the accuracy of the answer. For an optical processor that is 98% accurate (6 bit resolution), one can assume that after this addition of the residue, the error will be reduced approximately 50 times. The process is repeated until the desired accuracy is obtained. Let us consider a 25×25 matrix, and an optical processor with 6 bit resolution. Simulations show that one can get to 16 bit accuracy within 5 or 6 iterations. It is possible to obtain 32 bit accuracy or better if more iterations are performed, given that a more accurate digital processor is used.

The theoretical description will not be complete without an evaluation of the case where the eigenvalues are not distinct. The Jordan canonical representation of a matrix with distinct eigenvalues is a diagonal matrix with the eigenvalues on the diagonal. When the eigenvalues are not distinct, the Jordan canonical representation is a matrix with submatrices on the diagonal, or

$$\hat{\mathbf{A}} = \begin{bmatrix} [C]_{p_{11}} & [C]_{p_{12}} & & \mathbf{0} \\ & \ddots & & \\ & & [C]_{p_{1k_1}} & \\ \mathbf{0} & & & \ddots \\ & & & [C]_{p_{mk_m}} \end{bmatrix} \qquad (6)$$

where each $[C]_{p_{ij}}$ represents the j^{th} Jordan submatrix associated with λ_i, p_{ij} is the order (dimension) of the submatrix, k_i is the number of submatrices associated with λ_i, and m is the number of distinct

eigenvalues. One can only find as many linearly independent eigenvectors for this matrix as there are Jordan submatrices. In general, the number of independent eigenvectors are less than N. It is however possible to find a full set of N generalised RH eigenvectors that are linearly independent[3], if one defines a generalised RH eigenvector of grade k associated with λ_i as a vector **v** that satisfies

$$(\mathbf{A} - \lambda_i \mathbf{I})^k \mathbf{v} = \mathbf{0} \quad \text{while} \quad (\mathbf{A} - \lambda_i \mathbf{I})^{k-1} \mathbf{v} \neq \mathbf{0}. \tag{7}$$

One can also define a generalised LH eigenvector of grade k as a vector **u** that satisfies

$$(\mathbf{A}^T - \lambda_i \mathbf{I})^k \mathbf{u} = \mathbf{0} \quad \text{while} \quad (\mathbf{A}^T - \lambda_i \mathbf{I})^{k-1} \mathbf{u} \neq \mathbf{0}. \tag{8}$$

It is possible to show that one can find a set of RH eigenvectors **Q** (not necessarily unique) and a set of LH eigenvectors **R** (also not necessarily unique), such that $\mathbf{Q}^{-1} = \mathbf{R}^T$. Using these equations, one can manipulate Equation (2) into the following form:

$$\mathbf{x}(t) = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \sum_{q=1}^{p_{ij}} \left\{ \mathbf{u}_{q+\gamma_j} \bullet \mathbf{x}_0 \, e^{-\lambda_i t} \left[\sum_{r=1}^{q} \mathbf{v}_{r+\gamma_j} \left(\frac{t^{q-r}}{[q-r]!} \right) \right] \right. \tag{9}$$

$$\left. + \mathbf{u}_{q+\gamma_j} \bullet \mathbf{b} \left[\sum_{r=1}^{q} \mathbf{v}_{r+\gamma_j} \left\{ [1/\lambda_i]^{q-r+1} + e^{-\lambda_i t} \sum_{s=0}^{q-r} (-1)^s \frac{t^{q-r-s}}{(q-r-s)!(-\lambda_i)^{r+1}} \right\} \right] \right\}$$

$$\text{with} \quad \gamma_j = \sum_{a=1}^{j-1} p_{ia}$$

If we assume that the eigenvalues of the matrix **A** are distinct, or alternatively that the Jordan canonical representation of **A** is a diagonal matrix, this expression simplifies to Equation (4).

The convergence of the solution vector in Equation (9) is still dependent on the eigenvalues of the matrix. Some of the terms are multiplied by t^k, but since these terms are firstly scaled by (k)!, and secondly multiplied by $\exp(-\lambda t)$, their values reduce to zero as $t \to \infty$.

If we now compare the results that we have obtained for the general case where the matrix has multiple eigenvalues, with that of Cheng and Caulfield, we can see that the conclusions made by them are still valid. The eigenvalues of the matrix dominate the convergence of the analog feedback loop, with the requirement that the eigenvalues must have positive real parts.

This article reports on the general theoretical description of the bimodal optical computer, showing the role that the eigenvalues of the matrix plays in the convergence of the optical processor. We have shown that for the general case the convergence of the optical processor still depend on the eigenvalues of the matrix. As time approaches infinity, an answer similar to Equation (5) is obtained. This is a confirmation that the special case discussed by Cheng and Caulfield[2], did not lead to incorrect assumptions about the convergence requirements of the optical processor. Coupling this optical processor to a digital processor to perform iterative refinement, leads to a fast processor that solves linear algebraic equations with 16-32 bit accuracy.

References

[1] Caulfield HJ, et al, "Bimodal optical computers", Appl Opt, Vol. 25, no. 18, p. 3128, 1986

[2] Cheng WK, Caulfield HJ, "Fully-parallel relaxation algebraic operations for optical computers", Opt. Comm, Vol. 43, no. 4, p. 251, 1982

[3] Chen CT, "Linear system theory and design", Holt, Rinehart and Winston, New York, 1984

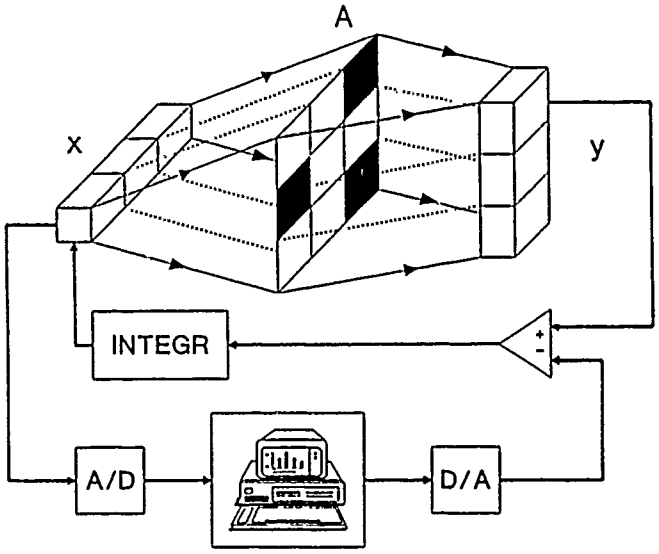


Figure 1. Bimodal optical computer

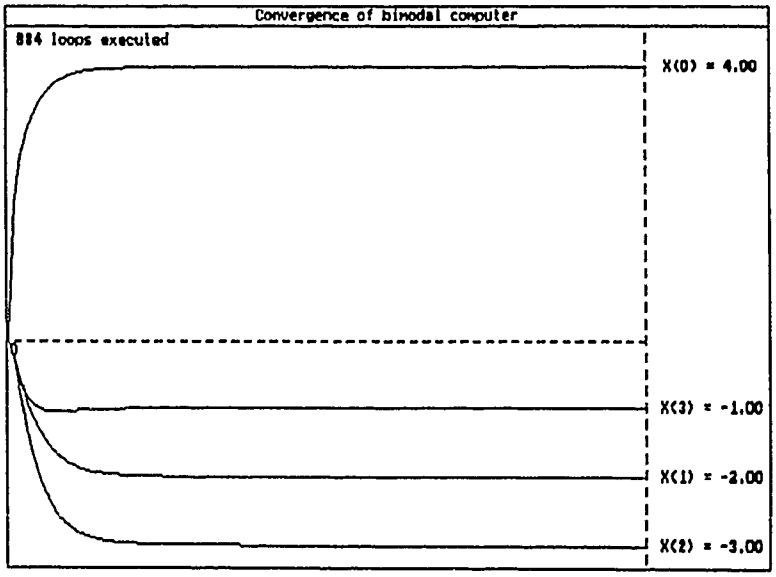


Figure 2. Convergence of the feedback loop

Sequential Logic Operation Using Optical Parallel Processor Based upon Polarization Encoding

By
Masashi Hashimoto, Ken-ichi Kitayama, and Naohisa Mukohzaka[+]
NTT Transmission Systems Laboratories,
1-2356 Take, Yokosuka-shi, Kanagawa 238-03, Japan
[+]Hamamatsu Photonics K.K.
1126-1 Ichino-chou, Hamamatsu-shi, Shizuoka, Japan

1. Introduction

Optical parallel processing based upon various polarization encodings[1] is a promissing concept for the practical implementation. The optical processor performing programmable and cascade operations on real-time basis, has been constructed.[2]

In this paper, sequential logic operation is performed using newly developed all-optical processor based upon the polarization encoding. Optical latch memory and spatial decoder in optical feedback path are the key elements. A successful operation is largely due to the precise interconnection between encoder for input and latch via the feedback path.

2. Architecture

In Fig.1(a), the basic architecture of sequential operation based upon finite state machine is shown. For sequential logic operation, the interconnection must include parallel feedback loop to obtain memories from latches. This circuitry allows to execute any combinatorial logic operation. In Fig.1(b), the optical implementation including a logic array block consisting of two polarization encoders for binary logic and operation kernel, a spatial decoder, and optical latches is shown. The logic array can execute all sixteen Boolean logics by providing instructions to the operation kernel. The spatial decoder completes the interconnection between the logic array block and latches.

3. Processing Algorithm

In Table1, the processing algorithm is shown. Assume that the logic operation is expressed by the functional form of sum of product, A•B+C•D+E•F+--. Switching of the operation kernel between product and sum is needed. Note that erasing steps for the encoder and latch must be taken before addressing new data.

4. Realignment of Optical Path using Spatial Decoder

As shown in Fig.2, the polarization encoding is carried out with SLMs(spatial light modulators) as polarization modulators and BP as a spatial separator. In the encoding process four possible combinatorial binary logics between the two inputs are allocated individually on spatially separated optical paths. Note that the encoding process is made pixel by pixel basis. To complete an optical connection for sequential logic operation, therefore, realigning the optical path of the output light into a unique position after passing through the operation kernel is requisite before proceeding to the next operation.

An optical element, so called, spatial decoder is developed for the specific spatial realignment. Without this, the SWBP of the processor deteriorates as increasing the encoding step because a pixel occupies the area for 2^N of optical paths where N is the number of encodig step. In Fig.2(a) the scheme of spatial decoder is shown. At this time, the decoder is designed to apply to the linearly polarized light after two steps of the encoding.

The path-wide stripe of half wave plate is indicated by shades area. Assume that the input light to the spatial decoder is horizontally polarized. The first and secont birefringent plates shift the vertically polarized light upward and right, respectively, to the next neighboring path of horizontally polarized light. The polarizer allows to pass only vertically polarized component out of incoming circulary polarized light. Thus, light beam on any of four possible paths shifts eventually into a unique optical path. As seen from Fig.3(b), this is confirmed experimentally.

5. Experimental Results

The experimental setup is shown in Fig.4. Two optically addressable MSLMs(micro-channel spatial light modulators)[3] 1 and 2 are used for polarization modulation of inputs. Another two MSLMs 3 and 4 are used as latched. MSLM can store data for days. LC(liquid-crystal)-SLM is used as the operation kernel for programmable operation on real-time basis. It filters spatially the light pixel by pixel according to operation instruction.

Precision of optical interconnection via the feedback path was tested. Pattern A in Fig.5(a) is addressed on MSLM1, while FALSE(0) logic is addressed on MSLM2. A OR 0 is addressed on the latch, MSLM3. It is transferred to MSLM1 via optical feedback path. To inspect the optical interconnectivity, logic operation of input A XOR read-out (A OR 0) is performed. As seen from Fig.5(c), the result is almost perfectly FALSE logic over the whole area. This shows that the optical interconnection via the feedback path is achieved precisely. On the contrary, the mismatch of the interconnection partly results in TRUE logic for the same XOR logic operation as shown in Fig.5(d). These experimental results are so encouraging to carry out the sequential logic operation with a high accuracy.

Logic operation for three patterns A, B, and C is shown in Fig.6. A OR B shown in Fig.6(c) is addressed on the latch MSLM3 after passing through the spatial decoder. Then read-out light from MSLM3 in Fig.6(d) is addressed again on MSLM1. On the other hand, input pattern C in Fig.6(e) is addressed on MSLM2. Finally, logic operation for three patterns, read-out (A OR B) EQV input C is executed. The result after the operation kernel is shown in Fig.6(f). The contrast of the same pattern read out from latch 3, shown in Fig.6(h), guarantees that optical gain provided by MSLM balances the loss per cycle.

It is found that as the step proceeds the the pixel defect increases. This is mainly due to the accumulated effect of phase-modulation uniformity of MSLM caused by partly insufficient flatness of the crystal in MSLM.

6. Conclusion

We have described the architecture, algorithm, and configuration of optical processor performing sequential logic operations. Experimental results of sequential logic operation performed by all-optical processor have been shown. We intend to demonstrate various sequential logic operations with patterns of higher contrast and less pixel defect at the conference.

References

(1)For example, A.W.Lohmann et al. Appl. Opt. 26, p.131(1987).
(2)K.Kitayama et al. OPTICAL COMPUTING 88, P_I15(August 1988, Toulon, France).
(3)T.Hara et al. Proc. SPIE 613, p.153(1986).

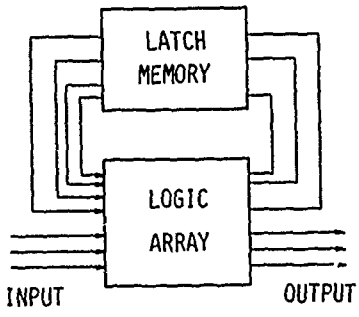


Fig. 1(a) Architecture of sequential logic operation

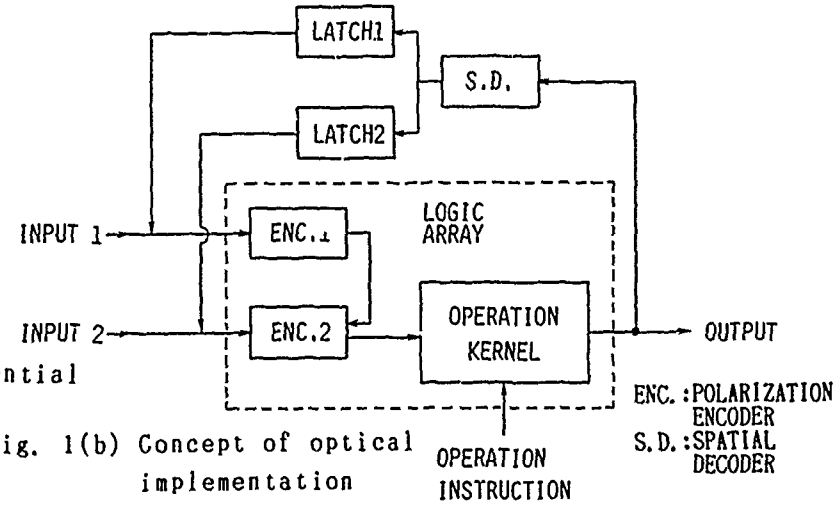


Fig. 1(b) Concept of optical implementation

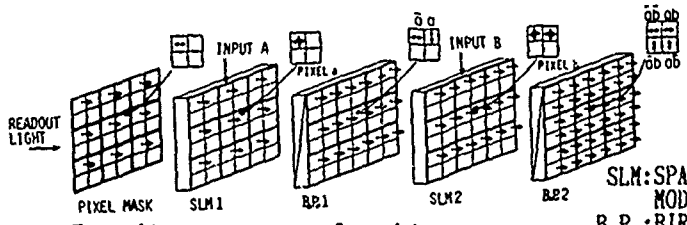


Fig. 2 Encoding concept for binary logical values using polarization state of light

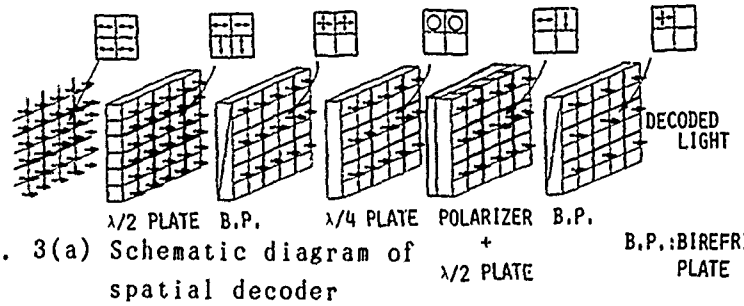


Fig. 3(a) Schematic diagram of spatial decoder

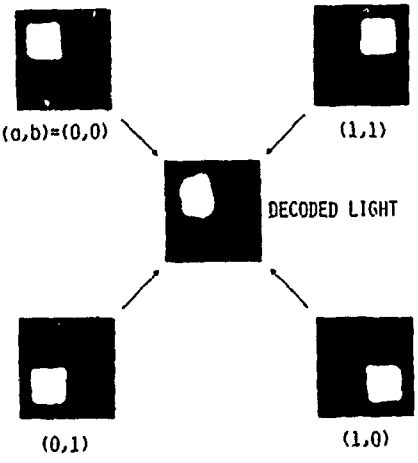


Fig. 3(b) Optical path realigned by the spatial decoder

Table 1 Algorithm for sequential logic operation of the processor

STEP	ENC. 1	ENC. 2	INST.	LATCH1	LATCH2
1	A	B			
2			AND	A · B	
3	ERASE	ERASE		A · B	
4	C	D	AND	A · B	C · D
5	ERASE	ERASE		A · B	C · D
6	A · B	C · D		A · B	C · D
7				ERASE	ERASE
8			OR	A · B+ C · D	
9	ERASE	ERASE		A · B+ C · D	
10	E	F		A · B+ C · D	
·					
·					

ENC.:ENCODER
INST.:OPERATION INSTRUCTION

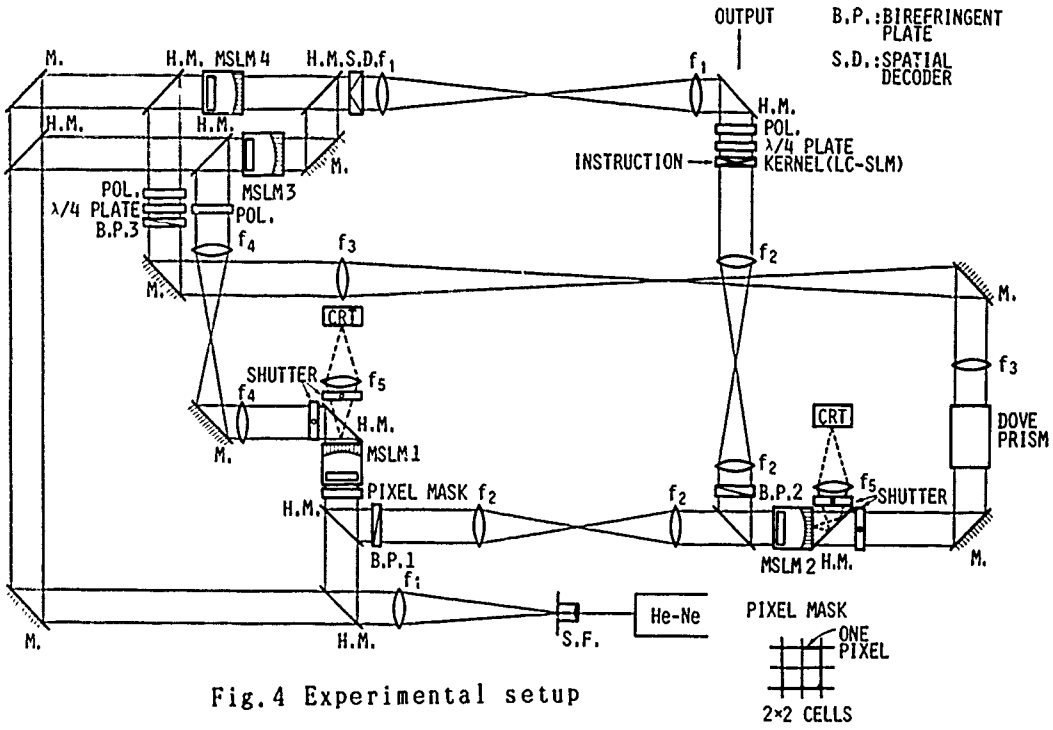


Fig.4 Experimental setup

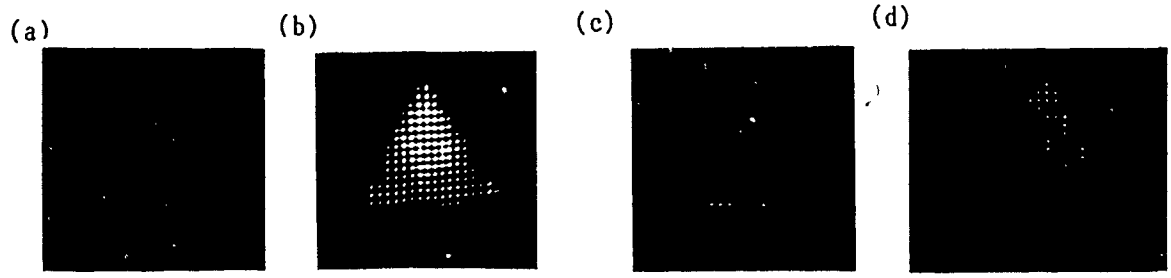


Fig.5 Optical interconnetion via feedback path
(a)INPUT A (b) A read out from latch
(c)A XOR A (d) A XOR A with mismatch

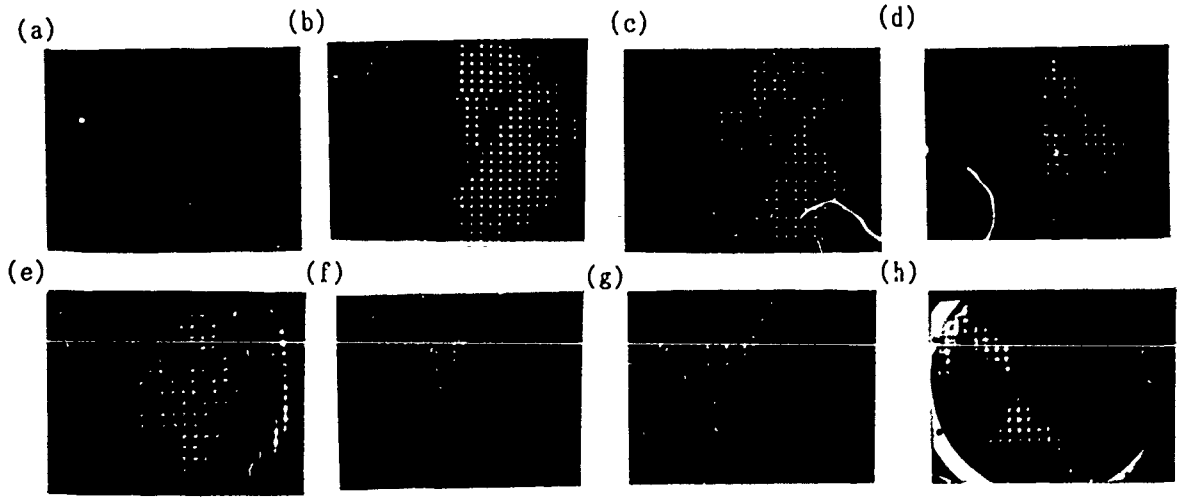


Fig.6 Experimental results of sequential logic operation
(a)INPUT A (b)INPUT B (c)A OR B (d)A OR B decoded by S.D.
(e)A OR B read out from latch 3 (f)INPUT C (g)(A OR B) EQV C
(h)(A OR B) EQV C decoded and read out from latch 3

NOTES

WEDNESDAY, MARCH 1, 1989

SALON F

2:00 PM–5:30 PM

WAA1–WAA4

JOINT PHOTONIC SWITCHING AND OPTICAL COMPUTING PLENARY SESSION

Joseph W. Goodman, Stanford University, *Co-presider*

John E. Midwinter, University College London, United Kingdom, *Co-presider*

OEIC Technology for Photonic Switching

S. Yamakoshi
Fujitsu Laboratories, Ltd.
10-1, Morinosato-Wakamiya, Atsugi 243-01, Japan

OEIC technology is promising to construct the new optical systems such as photonic switching, routing and other opitcal processing operations. The state-of-the-arts and future prospects of OEICs for photonic switching are discussed.

Quantum Well Devices for Optical Computing and Switching

D. A. B. Miller,

AT&T Bell Laboratories,

Holmdel, NJ 07733

The future prospects for both optical computing and photonic switching are clearly very dependent on advances in devices. This is especially true in the case of large scale applications requiring arrays of devices. Not only must the individual devices have good physical performance, they must also (i) operate at very low energies so that the array can be powered optically if required and have sufficiently low overall dissipation, (ii) be fabricatable in uniform arrays, and (iii) have sufficiently sophisticated functionality to allow efficient design of complex systems. Any one of these requirements suggests integration; taken together, we can see that a technology that offers straightforward integration of large numbers of flexible devices is essential for such array applications. Given that there are very few physical mechanisms that can offer sufficiently low operating energies for optical devices regardless of integration, we can see that this is a hard problem.

The potential return for a suitable technology is, however, large. Not only are there many architectural advantages in parallel optics, there are also now clear and relatively fundamental physical arguments why optics is actually better than electronics for communicating inside processors; the impedance transformation performed by optical devices can actually reduce the energy required for communication inside the processor.[1] To exploit this impedance transformation also requires integration since the capacitance of any connections between the opto-electronic devices and the electronic devices should be smaller than the device capacitances themselves. Indeed, it is arguable that much of the desire to avoid optical-electronic-optical conversions is because of lack of integration.

Quantum well devices have emerged over the past few years as strong candidates for optical switching and logic devices, especially for high-performance two-dimensional arrays compatible with free-space optics. One physical reason for this is the quantum-confined Stark effect electroabsorption mechanism,[2] which offers a low energy means for getting optical information out of a system, and is sufficiently strong that it can be used for modulation of beams propagating perpendicular to the chip surface, as required for two-dimensional arrays. A technological reason for the attractiveness of these devices is that the layered semiconductor growth techniques used to fabricate the thin semiconductor layers required for quantum wells and the lithographic techniques used for quantum well devices are also well suited to integration, both of many quantum well devices on one chip and of quantum well devices with other electronic and optical components.

The concept of combining photodetectors and quantum well modulators to give an optically controlled device with optical outputs is the principle of the self-electrooptic-effect device (SEED).[3] Such devices only offer low energy performance when they are integrated so that there are no parasitic capacitances associated with the interconnections between the different parts of the device. Integration of simple bistable devices was demonstrated first.[4] Although these large ($200\times200\ (\mu m)^2$) devices did not have particularly low switching energies (~ 1 - 2 nJ), they had the property that they could be scaled to smaller size ($60\times60(\mu m)^2$) with an approximately proportional improvement in switching energy and in increase in the size of the arrays (6×6).[5]

Systems experiments with simple bistable devices have problems because of the critical biasing requirements of simple bistable devices. The next step was therefore the symmetric SEED (S-SEED), which is a device that employs two quantum well diodes in series and is bistable in the ratio of two beam powers.[6] This S-SEED is in effect a three-terminal optical device, greatly simplifying system design. This device is now being scaled to larger arrays of smaller devices, with 16×8 arrays of devices with $13.5\times14\ (\mu m)^2$ mesas.[7] These devices are now being used for more complex optical circuit experiments.[8] There has also been recent work to extend the functionality of S-SEEDs further by including yet more diodes in series to make a multistate SEED (M-SEED).[9] Such a device can have N or 2^N stable states for N light beams on N series diodes depending on the biasing conditions.

Other opportunities with the SEED concept include the integration of more electronic components. Integration of bipolar transistors has been proposed,[3,10] and integration with field-effect transistors (F-SEED) has been demonstrated[11]. Importantly, the F-SEED integration is compatible with standard GaAs field effect transistor processing, and so we may contemplate integration of arbitrary amounts of electronics to expand the functionality of the optical module if we wish. It is also becoming increasingly likely that the quantum well modulators can be integrated with silicon circuits.[12]

In the future, we can expect continued miniaturization of SEEDs; indeed we cannot expect the necessary performance out of these devices for real applications unless and until they are fabricated with dimensions comparable to small electronic devices. We can also expect increasing flexibility in the functionality of the devices so that they are more suitable for particular systems applications. Finally, we can anticipate that a natural consequence of these developments will be to offer us the choice as to where we make the interface between optics and electronics so that we may have the best of both worlds.

[1] D. A. B. Miller, Optics Lett. (to be published, Jan. 1989)

[2] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus, Phys. Rev. **B32**, 1043 (1985); D. A. B. Miller, J. S. Weiner, and D. S. Chemla, IEEE J. Quantum Electron. **QE-22**, 1816 (1986).

[3] D. A. B. Miller, D. S. Chemla, T. C. Damen, T. H. Wood, C. A. Burrus, A. C. Gossard, and W. Wiegmann, IEEE J. Quantum Electron. **QE-21**, 1462 (1985);

D. A. B. Miller, U. S. Patents 4,546,244 and 4,716,449.

[4] D. A. B. Miller, J. E. Henry, A. C. Gossard, and J. H. English, Appl. Phys. Lett. **49**, 821 (1986).

[5] G. Livescu, D. A. B. Miller, J. E. Henry, A. C. Gossard, and J. H. English, Optic, Lett. **13**, 297 (1988).

[6] A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, Appl. Phys. Lett. **52**, 1419 (1988).

[7] L. M. F. Chirovsky, L. A. D'Asaro, C. W. Tu, A. L. Lentine, G. D. Boyd, and D. A. B. Miller, submitted to Photonic Switching Conference.

[8] F. B. McCormick, A. L. Lentine, L. M. F. Chirovsky, and L. A. D'Asaro, submitted to Photonic Switching Conference.

[9] A. L. Lentine, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, Paper FBB1, OSA Annual Meeting, Santa Clara, November 1988.

[10] P. Wheatley, P. J. Bradley, M. Whitehead, G. Parry, J. E. Midwinter, P. Mistry, M. A. Pate, and J. S. Roberts, Electron. Lett. **23**, 93 (1987).

[11] D. A. B. Miller, M. D. Feuer, T. Y. Chang, S. C. Shunk, J. E. Henry, D. J. Burrows, and D. S. Chemla, Paper TUE1, CLEO, Anaheim, April 1988.

[12] W. Dobbelaere, D. Huang, M. S. Unlu, and H. Morkoc, Appl. Phys. Lett. **53**, 94 (1988); K. W. Goossen, G. D. Boyd, J. E. Cunningham, W. Y. Jan, D. A. B. Miller, D. S. Chemla, and R. M. Lum, submitted to Quantum Wells for Optics and Optoelectronics Conference.

Switching in an Optical Interconnect Environment

Joseph W. Goodman
Department of Electrical Engineering
Stanford University

Optical interconnects are gaining importance in a wide range of applications, ranging from interconnection of supercomputers and workstations to interconnection of multiple chips on a single board. With the development of any interconnect technology, eventually the need for switching arises. Thus the switching of optical interconnects is a topic of much current interest.

Some applications of switching in the interconnect environment include, for example: 1) connection of a multitude of workstations to several shared resources, such as high-speed disk drives, laser printers, high-speed scanners, etc.; 2) connection of a multitude of backplanes in a tightly coupled multiprocessing machine; and 3) connection of a multitude of boards on an optical backplane in a single computer. The data rates and switch reconfiguration times required in these applications can differ significantly.

The lengths of the interconnects in these applications are typically quite short, ranging from perhaps hundreds of meters at one extreme (machine to workstation interconnection) to a few centimeters at the other (chip to chip on a board). The losses associated with the interconnect medium are therefore typically quite small (only a few dB), and both modal and material dispersion effects are often negligible. The consequences of these facts are several: 1) switching architectures with significant loss may still be of interest; 2) the choice of an operating wavelength (0.8 μm, 1.3 μm, etc.) is dictated by reliability rather than material dispersion; and 3) the use of multimode solutions is quite acceptable.

While switch loss appears not to be a critical parameter, nonetheless the issue does warrant further thought. When one considers the most fundamental motivations for the use of optics (as opposed to electronics) in interconnect problems, it appears that low drive power per length-bandwidth product can be one important advantage. When the internal loss of a switch is too great, optical interconnects may lose some of their

attractiveness when compared with electronic solutions, due to the increased electrical power required to drive the optical links.

With these facts in mind we examine several alternative approaches to optical switch construction for these applications. Most direct is the use of an electronic switch interfaced to optical receivers and transmitters. An intermediate electro-optic approach is a switch based on an array of forward- and back-biased detectors, interfaced to an array of optical transmitters [1,2]. All optical approaches include the optical matrix-vector [3,4], switches based on beam deflection, either through the use of stripe domain gratings in magneto-optic materials [5] or through acousto-optic deflection [6], and switches based on wavelength selective switching (e.g. [7]). Finally, LiNb switchable couplers, under intense development for long-distance telecommunications, are a candidate in this application as well (e.g. [8]). Each approach has its own unique advantages and disadvantages.

References

1. R.I. MacDonald, "Optoelectronic switch matrices: recent developments", *Optical Engineering*, Vol. 24, pp. 220-224 (1985).
2. G.L. Tangonen, V. Jones, J. Pikulski, D. Jackson, J. Persechini, G. Thornebooth, "8*8 optoelectronic crossbar switch", *Electronics Letters*, Vol. 24, No. 5, pp. 275-277 (1988).
3. A. Himeno, M. Kobayashi, "4x4 optical gate matrix switch", *J. Lighwave Technology*, Vol. LT-3, pp. 230-235 (1985).
4. A.R. Dias, R.F. Kalman, J.W. Goodman, A.A. Sawchuk, "Fiber-optic crossbar switch with broadcast capability", *Optical Engineering*, Vol. 27, No. 11, pp. 955-960 (1988).
5. E.J. Torok, J.A. Krawczak, G.L. Nelson, B.S. Fritz, W.A. Harvey, F.G. Hewitt, "Photonic switching with stripe domains", in *Photonic Switching*, T.K. Gustafson and P.W. Smith, Eds., Springer-Verlag, pp. 46-49 (1988).
6. P.C. Huang, W.E. Stephens, T.C. Banwell, L.A. Reith, "4x4 Acoustooptic photonic space switch with multicasting capability", *Paper Summaries*, OSA Annual Meeting, Paper TuK6, October 1988.
7. S. Suzuki, K. Nagashima, "Optical broadband communication network architecture utilizing wavelength-division switching technology", in *Photonic Switching*, T.K. Gustafson and P.W. Smith, Eds., Springer-Verlag, pp. 134-137 (1988).
8. H.S. Hinton, "Applications of photonic switching devices", *Proceedings of the SPIE*, Vol. 835, pp. 11-16 (1988).

The Relationship Between Photonic Switching and Optical Computing

H. S. Hinton

AT&T Bell Laboratories

Naperville, Illinois 60566P

The purpose of this talk is to outline the relationship between the hardware requirements of photonic switching and optical computing systems. The majority of the talk will address the hardware requirements of digital optical switching and computing systems with the exception of a brief discussion on analog switching and computing. It will include a review and comparison of the devices, interconnects, and systems that have been proposed for both types of systems.

KEY TO AUTHORS, PAPERS AND PRESIDERS