AD A046214

ENGINEERING

AD No.
DDC FILE COPY

USCIPI Report 770

12

b 5

UNIVERSITY OF SOUTHERN CALIFORNIA

SEMIANNUAL TECHNICAL REPORT

Harry C. Andrews
Project Director

Covering Research Activity During the Period
1 April 1977 through 30 September 1977

30 September 1977

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

Sponsored by

Advanced Research Projects Agency
Contract No. F-33615-76-C-1203
ARPA Order No. 3119

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

NOV 9 1977

IPI

IMAGE PROCESSING INSTITUTE

The views and conclusions in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

USCIPI Report 770

SEMIANNUAL TECHNICAL REPORT

Covering Research Activity During the Period
1 April 1977 through 30 September 1977

Harry C. Andrews
Project Director
(213) 741-5514

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007



30 September 1977

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Wright Patterson Air Force Base under Contract F-33615-76-C-1203, ARPA Order No. 3119. Additional support was provided by NSF Grant ENG-76-15318, AFOSR Contract AFOSR-77-3285 and WPAFB Contract F-33615-77-C-1016.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

2a. REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

(6) IMAGE UNDERSTANDING RESEARCH.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

(9) Semiannual Technical, 1 Apr[illegible]-30 Sep[illegible]77

5. AUTHOR(S) (First name, middle initial, last name)

(10) Harry C. Andrews (Project Director)

rept.

6. REPORT DATE

(11) 30 Sep[illegible]77

7a. TOTAL NO. OF PAGES

215

7b. NO. OF REFS

84

8a. CONTRACT OR GRANT NO.

(15) F-33615-76-C-1203

b. PROJECT NO.

ARPA Order [illegible]-3119

c.

d.

(12) 215p.

9a. ORIGINATOR'S REPORT NUMBER(S)

(14) USCIPI [illegible]-770

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Approved for release: distribution unlimited

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

13. ABSTRACT

This technical report summarizes the image understanding, smart sensor, and image processing research activities performed by the Image Processing Institute at the University of Southern California during the period of 1 April 1977 through 30 September 1977 under Contract Number F-33615-76-C-1203 with the Advanced Research Projects Agency Information Processing Techniques Office.

The research program has as its primary purpose, the development of techniques and systems for understanding images. Five tasks are reported: Image Understanding Projects, Image Processing Projects, Smart Sensor Projects, Recent Ph.D. Dissertations, and Recent Institute Personnel Publications. The image understanding tasks reported on include comparison of region growing versus boundary delineated segmentation, analytic results on the clustering segmentor, development of feature extractors for edge detection, circle detection, line detection, and texture detection and higher level image understanding for an interactive user system as well as a system for sharing information between existing image understanding programs. The image processing projects include degrees of freedom analyses for radar images, blind a posteriori phase restoration, psychovisual modelling for image rate distortion analysis, and optical processing procedures for on-axis holographic filtering and optical pseudocoloring of texture information. The smart sensor project describes testing of the Sobel circuit and fabrication of the spatially variant circuit. Recent Ph.D. disserta-

DD FORM 1 NOV 68 1473

391 141

UNCLASSIFIED

Security Classification

LB

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
tions are discussed in the following section, the report concluding with listings of recent publications.						
**********						
Key Words: Digital Image Processing, Image Restoration, Degrees of Freedom, Scene Analysis, Image Understanding, Edge Detection, Image Segmentation, CCD Arrays, CCD Processors.						

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. SPECIAL
A

ABSTRACT

This technical report summarizes the image understanding, smart sensor, and image processing research activities performed by the Image Processing Institute at the University of Southern California during the period of 1 April 1977 through 30 September 1977 under Contract Number F-33615-76-C-1203 with the Advanced Research Projects Agency Information Processing Techniques Office.

The research program has as its primary purpose, the development of techniques and systems for understanding images. Five tasks are reported: Image Understanding Projects, Image Processing Projects, Smart Sensor Projects, Recent Ph.D. Dissertations, and Recent Institute Personnel Publications. The image understanding tasks reported on include comparison of region growing versus boundary delineated segmentation, analytic results on the clustering segmentor, development of feature extractors for edge detection, circle detection, line detection, and texture detection and higher level image understanding for an interactive user system as well as a system for sharing information between existing image understanding programs. The image processing projects include degrees of freedom analyses for radar images, blind a posteriori phase restoration, psychovisual modelling for image rate distortion analysis, and optical processing procedures for on-axis holographic filtering and optical pseudocoloring of texture information. The smart sensor project describes testing of the Sobel circuit and fabrication of the spatially variant circuit. Recent Ph.D. dissertations are discussed in the following section, the report concluding with listings of recent publications.

PROJECT PARTICIPANTS

Project Director	Affiliation
Harry C. Andrews	Computer Science & Electrical Engineering

Research Staff	Affiliation
Nasser E. Nahi	Electrical Engineering
Ramakant Nevatia	Computer Science
William K. Pratt	Electrical Engineering
Keith E. Price	Image Processing Institute
Erica M. Rounds	Computer Science
Alexander A. Sawchuk	Electrical Engineering
Timothy C. Strand	Image Processing Institute

Support Staff

Rosemarie Bell	Clay Olmstead
Marilyn Chan	David Peck
Gary Edwards	Beverly Sanders
Eva Huang	Hector Saucedo
Scott Johnston	Ray Schmidt
Eileen Jurak	Gary Seaton
Ed Kasanjian	Joyce Seguy
Toyone Mayeda	Dennis Smith
Satyabrata Misra	Pat Stoliker
David Nagai	James Tertocha
Shankar Narayan	Amy Yiu

Students

Ikram Abdou
Ahmad Armand
Bir Bhanu
Chung-Ching Chen
Peter Chuan
Guy Coleman
David Garber
Charles Hall
Chung-Kai Hsueh
Bijan Lashgari
Kenneth Laws
Simon Lopez-Mora
Lee Martin
Dennis McCaughey
John Morton
Jin Soh

TABLE OF CONTENTS

Page

1. Research Overview..... 1

2. Image Understanding Projects..... 4
   2.1 A Comparison of Some Segmentation Techniques
   - Ramakant Nevatia and Keith Price..... 5
   2.2 Locating Structures in Aerial Images
   - Ramakant Nevatia and Keith Price..... 11
   2.3 An Interactive User System
   - Keith Price..... 17
   2.4 An Initial System for Sharing Information Between Image Understanding Programs
   - Keith Price..... 19
   2.5 Singular Value Decomposition Image Feature Extraction
   - William K. Pratt..... 20
   2.6 Some Comments on the Hueckel Operator
   - Ikram E. Abdou and William K. Pratt..... 30
   2.7 Development of Edge Detectors for the Extraction of Linear Segments
   - Peter Chuan..... 37
   2.8 Circle Detection in Noisy Images
   - Kenneth Laws..... 48
   2.9 Analytic Results of the Coleman Segmentor
   - Harry C. Andrews..... 59

3. Image Processing Projects..... 72
   3.1 Synthetic Aperture Radar and Imaging System of the Stripping Mode
   - Chung-Ching Chen..... 73
   3.2 DOF of Tomographic Projections
   - Chung-Ching Chen.....103

Page

3.3 Psychovisual Transform Coding of Images
- Charles F. Hall.....122
3.4 Statistical Analysis of a Model of the Human Visual System
- Charles F. Hall and Lloyd R. Welch.....129
3.5 A Technique for A Posteriori Restoration
- John Morton.....141
3.6 On-Axis Optical Filtering System
- Alexander A. Sawchuk and Chung-Kai Hsueh.....151
3.7 Optical Pseudocoloring of Spatial Frequency Information
- Timothy C. Strand.....161

4. Smart Sensor Projects.....172
4.1 Charge-Coupled Technology for Image Understanding
- Graham R. Nudd.....173

5. Recent Ph.D. Dissertations.....202
5.1 A Joint Detection-Estimation Approach to Boundary Estimation
- Simon Lopez-Mora.....203
5.2 Image Segmentation by Clustering
- Guy B. Coleman.....204

6. Recent Institute Personnel Publications.....206

1. Research Overview

This document represents the fourth semiannual report funded under the current ARPA Image Understanding contract. Research over the past six months (and, in fact, past two years) has been devoted to three major areas. The first, and majority of effort, has been in the area of image understanding. The second has been the area of smart sensor design, and the third has been the area of research in image processing.

Image Understanding

This aspect of our research has concentrated upon both bottom up and heterarchical methodologies for image understanding. Segmentations based on region growing and boundary delineation have been compared to test the strength and weaknesses of each. Analytic results of a specific clustering segmentor are developed. Specific feature extractors for edge detection, circle detection, line detection, and texture detection are each investigated in considerable detail. The emphasis of approach is somewhat evenly distributed between the use of mathematical tools and the use of computer science tools. At the higher levels of the system a more sophisticated viewpoint is developed for the heterarchical methods applied to locating structures in aerial images. In addition these high level techniques are being applied to the development of an interactive user system for usage by Institute personnel and visitors. Finally, the development of an initial system for sharing information between existing image understanding programs is underway. It is in these latter directions that we expect the heterarchical approach to provide fruitful results.

Smart Sensors

This section represents the smart sensor phase of research funded during the past six months. Two circuits have been fabricated for CCD analog near focal plane processing to implement a variety of front end image processing functions. The first circuit implements the Sobel operator on an image. This represents a nonlinear spatially invariant processor. The second circuit is designed to implement both nonlinear spatially invariant as well as variant processes. Both circuits have been fabricated. The Sobel circuit has been tested with success, (see accompanying section) and the second circuit is soon to experience testing.

Image Processing

This section of the report describes the efforts expended and results obtained over the past six months on the various image processing projects carried out at the Institute. Some of these projects have been funded by other sources as indicated in the appropriate title acknowledgements. Results of work in defining the degrees of freedom inherent in radar imaging systems are presented for the stripping mode of SAR. It is expected that future radar imaging model degrees of freedom analysis will lead to efficient radar image understanding, a task which for humans, is far more difficult than visual image understanding. This section also summarizes the psychovisual modelling work being applied to image coding and image rate distortion functions. Results in blind a posteriori image restoration are next presented. Finally two new optical processing procedures are described in which on-axis holographic optical filtering is developed, and optical pseudocoloring of texture information is described.

Finally a report of recent Institute Ph.D. dissertations is included as well as the listing of recent Institute personnel publications in the open literature.

2. Image Understanding Projects

This aspect of our research has concentrated upon both bottom up and heterarchical methodologies for image understanding. Segmentations based on region growing and boundary delineation have been compared to test the strength and weaknesses of each. Analytic results of a specific clustering segmentor are developed. Specific feature extractors for edge detection, circle detection, line detection, and texture detection are each investigated in considerable detail. The emphasis of approach is somewhat evenly distributed between the use of mathematical tools and the use of computer science tools. At the higher levels of the system a more sophisticated viewpoint is developed for the heterarchical methods applied to locating structures in aerial images. In addition these high level techniques are being applied to the development of an interactive user system for usage by Institute personnel and visitors. Finally, the development of an initial system for sharing information between existing image understanding programs is underway. It is in these latter directions that we expect the heterarchical approach to provide fruitful results.

2.1 A Comparison of Some Segmentation Techniques

Ramakant Nevatia and Keith Price

Segmentation is, of course, a key component in the Image Understanding process. The numerous segmentation techniques may be viewed as being either edge based or region based. The edge based techniques start by detection of local discontinuities in some attribute, such as brightness of an image and attempt to construct object boundaries from them. The region based techniques attempt to find areas in the image over which one or more attributes are constant.

It may be that in some sense the two techniques are trying to compute similar functions and that they should be capable of achieving similar performance. However, at the present state of development of these methods, one or the other technique may be more successful on certain kinds of images. This is a subject of active discussion among researchers in the field, but we are unaware of any comparative studies.

In the section, results of processing two selected, black and white pictures using the two classes of techniques are presented that lead to some expected conclusions about their suitability for different tasks. The edge based technique is that developed at the University of Southern California [1-2], and the region based technique is that of Ohlander [3], modified by K. Price [4], and developed at Carnegie-Mellon University.

A brief review of the two segmentation techniques used is provided here.

a) An Edge Based Method - In this method, a local edge operator is applied to an image first. The resulting edges are then linked in straight line segments and only segments of a minimum length or above are preserved (for details see [1]). It is hypothesized that such segments usually correspond to the desired boundaries.

The linking method is independent of the edge operator used. However, the final performance is obviously determined by the output of the local edge operator. We have used a Hueckel edge detector [5] in previous experiments. This edge detector is believed to have superior performance to many simpler edge detectors, but it is not always effective in the presence of texture. A simple edge detector, which consists of convolving an image with elongated edge masks in various directions and choosing the maximum was developed and found to perform well (for details, see [2]). This edge detector has been used in results presented later.

b) A Region Based Method - The Ohlander segmenter operates by computing histograms of various image attributes and segmenting the image into regions with a certain range of values of an attribute. The attribute with the best separation (a bimodel distribution in the histogram) is chosen for segmentation. Originally, the method was developed for color images. We have used only black and white images here and only the intensity attribute was used.

This technique is recursively applied to the segmented regions until regions become too small or cannot be further segmented according to established criteria of histogram separations. Regions smaller than a selected size are ignored. Therefore, long thin regions which are broken into

several smaller regions may be lost.

Experimental Results

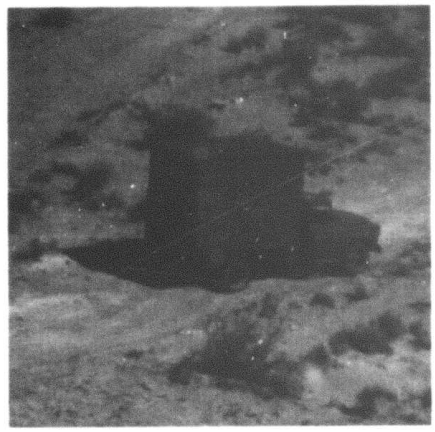
The two test images are shown in figures 1a and 2a. Figures 1b and 2b show the edges detected in the two images. Figures 1c and 2c show the regions detected by the Ohlander-Price segmenter. (Results of linking edges in line segments are not shown).

Following are some observations on the relative merits of the two approaches.

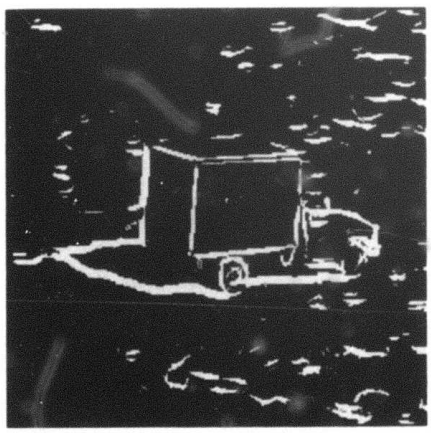
a) The performance on the simpler picture of the truck of figure 1 is comparable. The edge segmentation can be more sensitive, as in separating the front and top surface of the truck, but the boundary is fragmented into several segments. Region methods always give closed regions, by definition, which may be easier to handle for some types of objects or processing. Note that both methods fail to separate the bush and truck top surface.

b) In the more complex aerial picture, the edge technique seems to extract linear features, such as roads, with ease, whereas the region method does the same for parts of the image that are homogeneous, for example, the lakes in figure 2. Note that the major vertical road is not extracted as a region in figure 2b, but is only indicated by the boundaries of the regions on two sides of it.

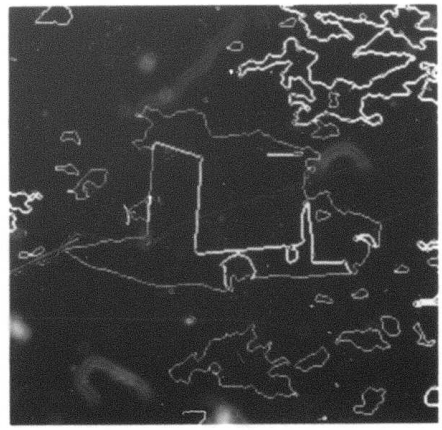
c) The more complex parts of the aerial pictures are not adequately analyzed by either technique, for example, the lower part of the river of the suburban areas in figure 2. The main difficulty seems to be due to the presence of texture and fine detail.



(a) Digitized image



(b) Detected edges

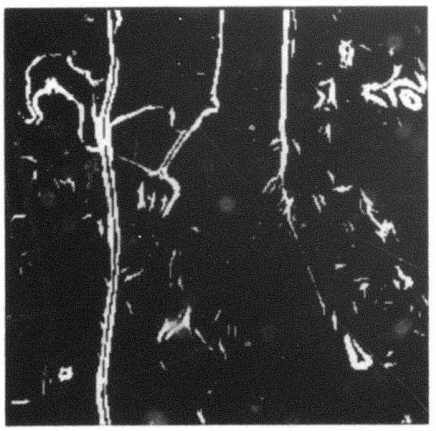


(c) Segmented regions

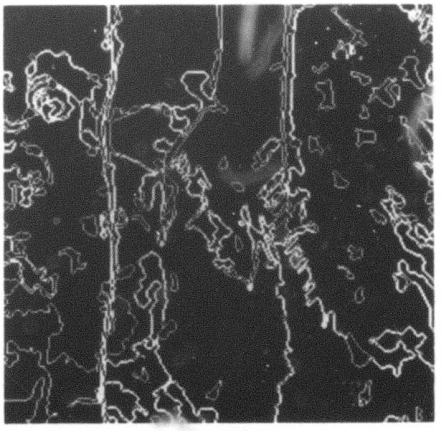
Figure 1. A Truck



(a) Digitized image



(b) Detected edges



(c) Segmented regions

Figure 2. An Aerial Picture

Conclusions

Interestingly, the two methods perform similarly on large areas of the tested images. However, specific structures are handled better with one or the other. The clear implication is that a complete Image Understanding system should utilize both depending on its goals. A straightforward method is to use a specific technique to locate particular types of objects.

The two segmentation techniques may also be able to reinforce each other at the image level, for example, using regions to bridge gaps in boundary segments or to use boundary segments to sub-divide regions. We have not examined such interaction in depth.

References

1. R. Nevatia, "Locating Object Boundaries in Textured Environments," IEEE Transactions on Computers, Vol. C-25, No. 11, November 1976, pp. 1170-1175.

2. R. Nevatia and P. Chuan, "Detection of Edges in Elongated Neighborhoods," University of Southern California, Image Processing Institute, USCIPI Report No. 740, March 1977, pp. 34-39.

3. R. Ohlander, "Analysis of Natural Scene," Ph.D. Thesis, Carnegie-Mellon University, Pittsburgh, Pennsylvania, August 1975.

4. K. Price, "Change Detection and Analysis in Multi-Spectral Images," Ph.D. Thesis, Carnegie-Mellon University, Pittsburgh, Pennsylvania, December 1976.

5. M. Hueckel, "An Operator Which Locates Edges in Digitized Pictures," Journal of the ACM, Vol. 18, No. 1, January 1971, pp. 113-125.

2.2 Locating Structures in Aerial Images

Ramakant Nevatia and Keith Price

Introduction

Analysis of aerial images is, in general, a complex task. The reasons for such complexities are many and varied. A prime cause is the presence of texture which causes difficulties for the low level processes such as edge detection and segmentation. Another source of difficulty is that the desired objects and structures may be small compared to the size of a complete image. A detailed analysis of a complete high resolution aerial image is generally prohibitive because of the computational costs.

For many applications, however, a complete and general analysis is unnecessary. Specific structures of interest may have special properties, known a priori, that allow for their easy extraction. The problem of searching for small structures is helped by locating them by their spatial relationships to larger, more easily located structures.

In previous work, we compared two segmentation techniques, the edge based and the region based methods, and concluded that one or the other may be suited for extraction of particular types of structures [1]. This describes our initial attempts to use both techniques, taking advantage of their respective strong points.

Problem Description and Representation

The problem approached is that of finding user specified structures in aerial images. The user specifies the properties useful for the location of the desired structure and also of other related structures. (An interactive, question-answer dialog system is being developed to facilitate interaction with a user, see [2].) This amount of a priori knowledge is likely to be available in many applications of guidance and photo-interpretation.

The a priori information is stored as properties of objects and their relationships to each other, and may be viewed as constituting a graph structure with the objects as nodes and relationships as arcs. The properties and relationships will, in general, need to be unrestricted. Currently, an object is described either by a collection of line segments or by its region properties. The segments are described by their length and width. The regions are described by properties such as brightness (color) and simple shape measures (area, perimeter, ratio of area to perimeter squared, elongation, etc.).

The relationships used are those of relative locations of the different objects and the symbolic relationships of left, right, above and below. Other relationships such as symmetry and similarity are obviously useful, but have not been implemented.

Our representation and use of knowledge is similar to that described by Tenenbaum [3]. The principal difference is in Tenenbaum's use of single pixel attributes to uniquely distinguish objects (in a given context). We use object attributes to aid in the segmentation of the image and then

use the attributes of larger, segmented parts for recognition.

Feature Extraction and Segmentation

Feature extraction and segmentation is guided by the properties of the desired objects to be extracted. Thus, an edge detection-line finding process is applied to extract desired linear segments (such as roads) and a region segmentor for extracting areas uniform in some property (for example lakes and other bodies of water).

Consider the aerial image shown in figure 1 (the displayed image contains 352 x 352 pixels, an image of twice the resolution is also used in the analysis). Here, an objective may be to locate the dock structure and perhaps some ships in it. As this structure consists of relatively small parts and is complex, it may be easier to extract related structures such as the river, the major highway and the lakes first, and use these to concentrate the search for docks to a smaller area of the image. (We assume such information is supplied by the user. No attempt has been made to automate the strategy generation process, as in [4].)

Edge detection processes are appropriate for the extraction of the desired roads. Figure 2 shows the results of applying a Hueckel edge detector [5] on the image of figure 1 and linking the resulting edge segments in elongated segments [6]. The road is known to be narrow enough that the edges corresponding to it are of the "line" type (as contrasted with a step edge). Restricting the linked edges to be only of line type results in fewer segments (shown in figure 2).

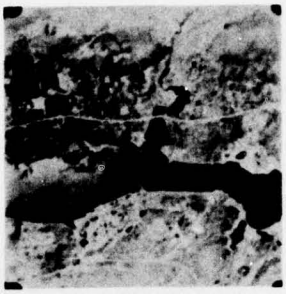


Figure 1. An aerial picture.

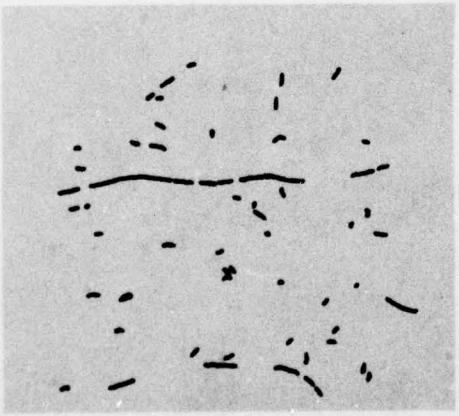


Figure 2. Lines detected in figure 1.

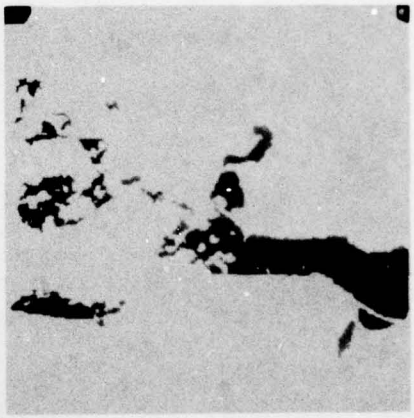


Figure 3. Dark regions detected in figure 1.

The lakes and parts of the river are conveniently extracted by using the Ohlander-Price Segmentor [7]. It is known that the desired objects are relatively dark and uniform in intensity, and the dark peak in the intensity histogram should be used for segmentation. The completed segmentation is shown in figure 3.

Matching of Segments

The next step is to match the derived line segments and regions with a model of the image. This phase of our work is in progress and experimental results are expected to be available soon. Assuming that the derived segments are distinctive enough to be easily distinguished, approximate location of the dock structures can be predicted. Now, sensitive line detectors should help locate the piers of the dock. (We have found the Hueckel edge detector to be deficient in locating small edges, perhaps because of the large neighborhood size used. Development of more sensitive edge and line detectors is being carried out concurrently, see [8].)

Conclusions

Some results of processing a complex, aerial image using both the line and the region based techniques have been shown. It appears that the use of simple techniques, specifically suited to particular objects in an image, may allow useful processing of rather complex images. This work is in initial stages of development and the array of segmentation attributes is limited. While it is hoped that the described techniques have general applicability, our experience with real images is, as yet, limited.

References

1. R. Nevatia and K. Price, "A Comparison of Some Segmentation Techniques," Proceedings: Image Understanding Workshop, Minneapolis, Minnesota, April 1977, pp. 55-57.

2. K. Price, "An Interactive User System," University of Southern California, Image Processing Institute, USCIPI Report 770, October 1977.

3. J.M. Tenenbaum, "Locating Objects by Their Distinguishing Features in Multisensory Images," Computer Graphics and Image Processing, Vol. 2, No. 3, December 1973.

4. T.D. Garvey and J.M. Tenenbaum, "On the Automatic Generation of Programs for Locating Objects in Office Scenes," Proceedings of the Second International Joint Conference on Pattern Recognition, Copenhagen, Denmark, August 1974, pp. 162-168.

5. M. Hueckel, "A Local Operator Which Recognizes Edges and Lines," Journal of the ACM, October 1973, pp. 634-647.

6. R. Nevatia, "Locating Object Boundaries in Textured Environments," IEEE Transactions on Computers, Vol. 25, No. 11, November 1976, pp. 829-832.

7. R. Ohlander, "Analysis of Natural Scenes," Ph.D. Thesis, Department of Computer Science, Carnegie-Mellon University, April 1975.

8. P. Chuan, "Development of Edge Detectors for the Extraction of Linear Segments," University of Southern California, Image Processing Institute, USCIPI Report 770,

October 1977.

2.3 An Interactive User System

Keith Price

One major problem with the several subsystems at the Image Processing Institute is that there is no way to easily run these programs. All these systems require program specific inputs which are meaningful only to the programmer. We have implemented an initial system which is intended to partially eliminate this problem. This system asks questions which are necessary to obtain the information needed to operate the program so that it acquires the proper information. The system is then able to initiate the execution of the subsystem and give the input that has been specified by the user's answers.

This program is not developed to the extent of systems such as MYCIN for medical analysis [1] or PROSPECTOR for geological analysis [2]. This program is intended to provide an inexperienced user with a guide to the information that is required by various programs, not deduce an interpretation.

For a program, or the set of programs, a user must provide a set of questions that may be asked. Each of the questions has a set of valid responses with actions for each response. These questions, responses, and actions are given in a file, not compiled, thus modification of the individual questions or changing the entire set of questions is simplified. The actions may include new questions which are necessary because of the answer, e.g. if the question is about the type of segments which are desired, and the user

indicates "lines" then the questions dealing with region based segmentations are not necessary and those dealing with line extraction must be asked. Actions also include remembering the program inputs that are specified by the user's answer. Some special operations such as running the specific program are compiled rather than included in the question-response file.

The future goals for this, or a similar system, are to learn how a user would solve particular problems, how certain types of objects might be specified and to allow interaction with the several distinct subsystems which are used in the image understanding task. The user should be allowed to modify or add to the legal questions, responses, or actions so that the user can request the segmentation of a particular object after describing it once rather than describing the method in detail each time. Currently a small set of questions are available for segmentation and matching operations.

References

1. E.H. Shortliffe, Computer-Based Medical Consultations: MYCIN, Elsevier, New York, 1976.

2. R.O. Duda, P.E. Hart, N.J. Nilsson, G.L. Sutherland, "Semantic Network Representations in Rule-Based Inference Systems," SRI AI Center Technical Note 136, January 1977.

2.4 An Initial System for Sharing Information Between Image Understanding Programs

Keith Price

This is a report on the initial efforts to allow easier interaction of a user between several different programs working on the same problem. The system allows the user to interact with several programs operating in parallel on one or more problems and to share information between programs.

The TENEX system provides many of the features necessary for this system. A program can create an inferior fork, start another program running in this fork, and also continue running itself. The program in the inferior fork does not require any modification to be run in this mode since the superior program controls the source and destination of all input and output which would usually go through the user terminal. The current system allows the user to send terminal input to a specific fork or connect to a particular fork so that all terminal input go directly to that fork.

Sharing of a portion of the address space to programs and inferior forks is also provided by the TENEX monitor. A program can map certain pages, a page is the basic TENEX unit for storage allocation, of its core area or a file into other pages in the program address space. Additionally, pages in the address space of a program in a superior fork can be mapped into the address space of an inferior fork or from one inferior fork to another. The mapping of pages in the address space of one fork to the address space of another fork means that references to addresses within a mapped page, by the program in either fork, are references to the same actual location.

If several programs are going to share data or other results in this manner, then it is usually necessary to modify these programs, since the shared data must be placed in a particular location in the address space. Such data can thus be hard to access symbolically, for example as a simple variable, but SAIL provides a partial solution to this problem, that is the access of this global data symbolically. The RECORD construct in SAIL has provisions for the user specification of the allocation method for individual records, such as determining the address of these records. Thus some record structures may be allocated in a global area and referenced symbolically by several current processes. The program in a superior fork can also access any address in its inferior forks. This feature is more useful as a de-bugging tool than for sharing data.

The question/answer program described elsewhere in this report may be included as a sub-program in this system. When that program is required to run another system, it uses the multiple process facilities built into this system.

References

1. K. Price, "An Interactive User Interface System," University of Southern California, Image Processing Institute, USCIPI Report 770, October 1977.

2.5 Singular Value Decomposition Image Feature Extraction

William K. Pratt

The singular value decomposition (SVD) is a numerical technique of matrix transformation by which an arbitrary matrix can be expressed in outer product form. SVD

expansions have been applied in the solutions of ill-conditioned sets of linear equations [1]. There has also been recent application of the SVD concept to image restoration and coding [2]. Another application, explored here, is the use of the SVD to extract descriptive features from an image region for purposes of classification or analysis.

Singular Value Matrix Decomposition

Consider an N x N matrix $\underline{F}$ of rank R. The SVD transform of $\underline{F}$ is defined to be

$$\underline{\Lambda}^{\frac{1}{2}} = \underline{U}^T \underline{F} \underline{V} \tag{1}$$

where $\underline{U}$ and $\underline{V}$ are unitary matrices and $\underline{\Lambda}^{\frac{1}{2}}$ is a diagonal matrix whose diagonal elements $\lambda^{\frac{1}{2}}(1) \geq \lambda^{\frac{1}{2}}(2) \geq \ldots \geq \lambda^{\frac{1}{2}}(N)$ are called the singular values of $\underline{F}$. Since $\underline{U}$ and $\underline{V}$ are unitary matrices, the inverse transform yields

$$\underline{F} = \underline{U} \underline{\Lambda}^{\frac{1}{2}} \underline{V}^T \tag{2}$$

The unitary matrix $\underline{U}$ contains columns $\underline{u}_m$ composed of eigenvectors of the symmetric matrix product $\underline{F}\underline{F}^T$. Similarly, the columns $\underline{v}_n$ of $\underline{V}$ are eigenvectors of $\underline{F}^T\underline{F}$. The defining relations are

$$\underline{U}^T [\, \underline{F}\, \underline{F}^T \,] \underline{U} = \underline{\Lambda} \tag{3a}$$

$$\underline{V}^T [\, \underline{F}^T \underline{F} \,] \underline{V} = \underline{\Lambda} \tag{3b}$$

The matrix decomposition of eq.(2) can also be expressed in the series form

$$\underline{F} = \sum_{j=1}^{R} \lambda^{\frac{1}{2}}(j)\, \underline{u}_j\, \underline{v}_j^T \tag{4}$$

where the sum is over the rank R of $\underline{F}$.

SVD Image Feature Extraction Concept

The singular values of a matrix can be considered as descriptors or features of the matrix elements and their inter-relationships. If the matrix is composed of randomly chosen real numbers, the singular values will tend toward equality. On the other hand, a highly structured matrix will exhibit a few dominating singular values. Figure 1 sketches the qualitative structural relationship of the singular values. This observation forms the basis for utilization of the SVD as a means of forming image features. An SVD expansion is formed over a block of N x N pixels, and some measure of the skewness of the singular value distribution is formed to characterize the spatial "coherence" of the pixel values. This vague concept will be solidified shortly.

Deterministic Properties

Certain deterministic image patterns possess SVD expansions that can be easily computed. For example, an array of unit value pixels can be generated by the outer product expansion

$$\begin{bmatrix} 1 \\ 1 \\ . \\ . \\ . \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & . & . & . & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ 1 & 1 & . & . & . & 1 \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 1 & 1 & . & . & . & 1 \end{bmatrix} \qquad (5)$$

Such a matrix is of unit rank, and therefore, possesses only one non-zero singular value. A matrix containing alternating vertical stripes of zeroes can be formed by the single outer product expansion

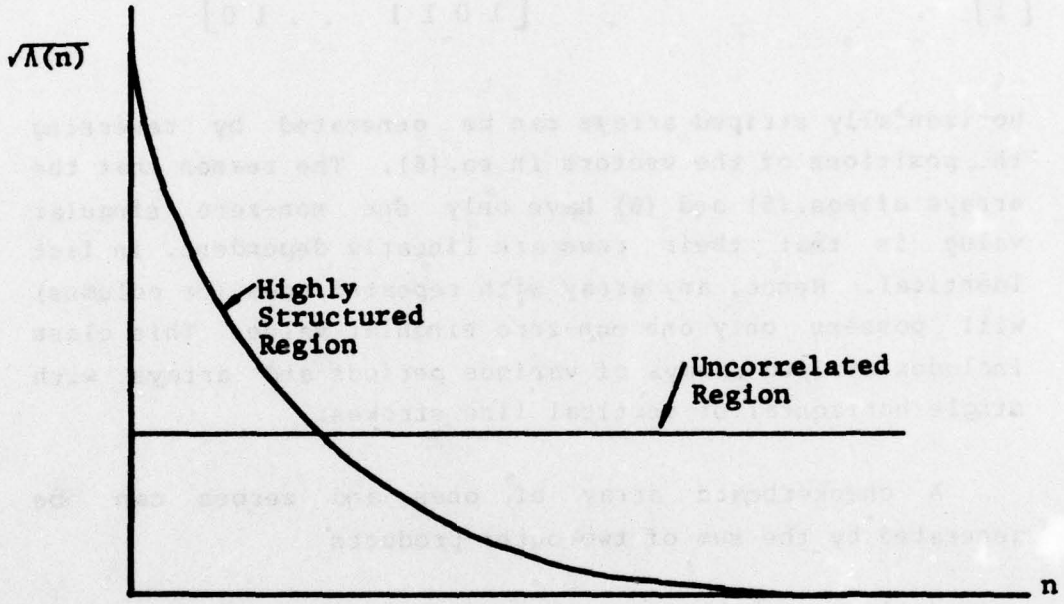


Figure 1. Example of Singular Value Behavior of Image Regions.

$$\begin{bmatrix}1\\1\\.\\.\\.\\1\end{bmatrix}\begin{bmatrix}1&0&1&0&.&.&.&1&0\end{bmatrix} = \begin{bmatrix}1&0&1&0&.&.&.&1&0\\1&0&1&0&.&.&.&1&0\\.&.&.&.&&&&.&.\\.&.&.&.&&&&.&.\\.&.&.&.&&&&.&.\\1&0&1&1&.&.&.&1&0\end{bmatrix} \qquad (6)$$

Horizontally striped arrays can be generated by reversing the positions of the vectors in eq.(6). The reason that the arrays of eqs.(5) and (6) have only one non-zero singular value is that their rows are linearly dependent, in fact identical. Hence, any array with repeated rows (or columns) will possess only one non-zero singular value. This class includes striped arrays of various periods and arrays with single horizontal or vertical line strokes.

A checkerboard array of ones and zeroes can be generated by the sum of two outer products

$$\begin{bmatrix}1\\0\\1\\0\\.\\.\\.\\1\\0\end{bmatrix}\begin{bmatrix}1&0&1&0&.&.&.&1&0\end{bmatrix} + \begin{bmatrix}0\\1\\0\\1\\.\\.\\.\\0\\1\end{bmatrix}\begin{bmatrix}0&1&0&1&.&.&.&0&1\end{bmatrix}$$

$$= \begin{bmatrix} 1\,0\,1\,0\,\ldots\,1\,0 \\ 0\,1\,0\,1\,\ldots\,0\,1 \\ \ldots\quad\quad\quad

non-zero singular values, while horizontally or vertically striped arrays only have one non-zero singular value. As an even more extreme example consider the N x N diagonal matrix with a single diagonal line. It has N identical singular values. But, an array with a single pixel vertical line possesses one non-zero singular value.

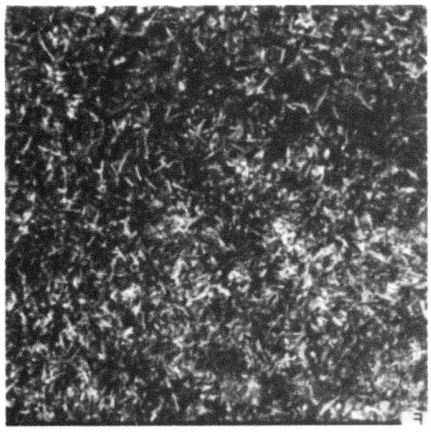
The conclusion of these observations is that the singular value distribution is not well suited for the characterization of deterministic line, spot, or shape structure. But rather, the SVD does provide an indication of the structural dependence between rows and columns of a pixel array. This property, as will be shown, is extremely important for the characterization of texture-like regions of an image.

Statistical Properties

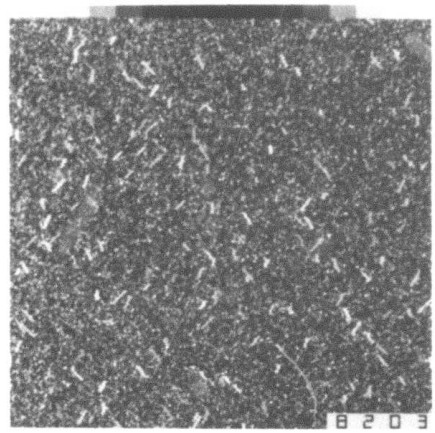
If the image block $\underline{F}$ is considered to be a sample of a two-dimensional random process then the singular values $\underline{\Lambda}^{\frac{1}{2}}$ as defined by eq.(1) will be random variables since the SVD is a linear transformation. Attempts are underway to determine the covariance matrix of the singular values in terms of the covariance function of the image block.

Experimental Results

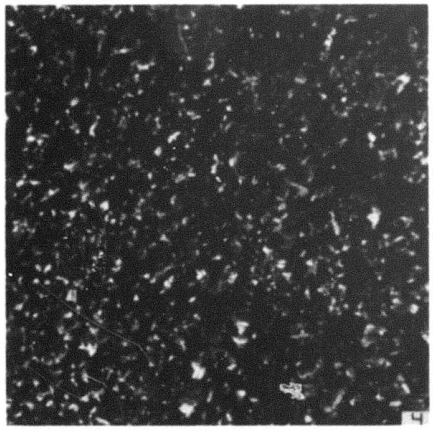
Figure 2 contains two images of natural texture, grass and ivy, along with two artificially generated fields obtained by convolutional processing of two-dimensional fields of randomly generated pixels. The subjective match between the natural and artificial fields appears to be in reasonable agreement. Figures 3 and 4 contain plots of the singular values extracted from 16 x 16 pixel blocks in the center of each image. The distributions of the natural



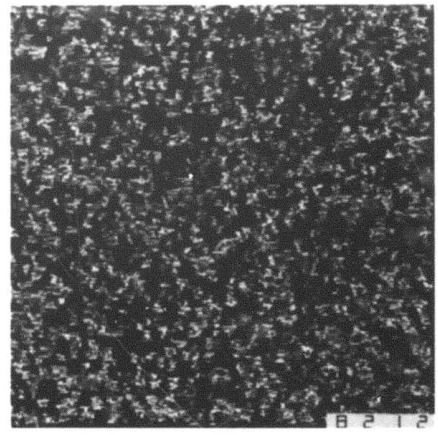
(a) natural grass



(b) artificial grass



(c) natural ivy



(d) artificial ivy

Figure 2. Examples of Natural and Artificial Texture.

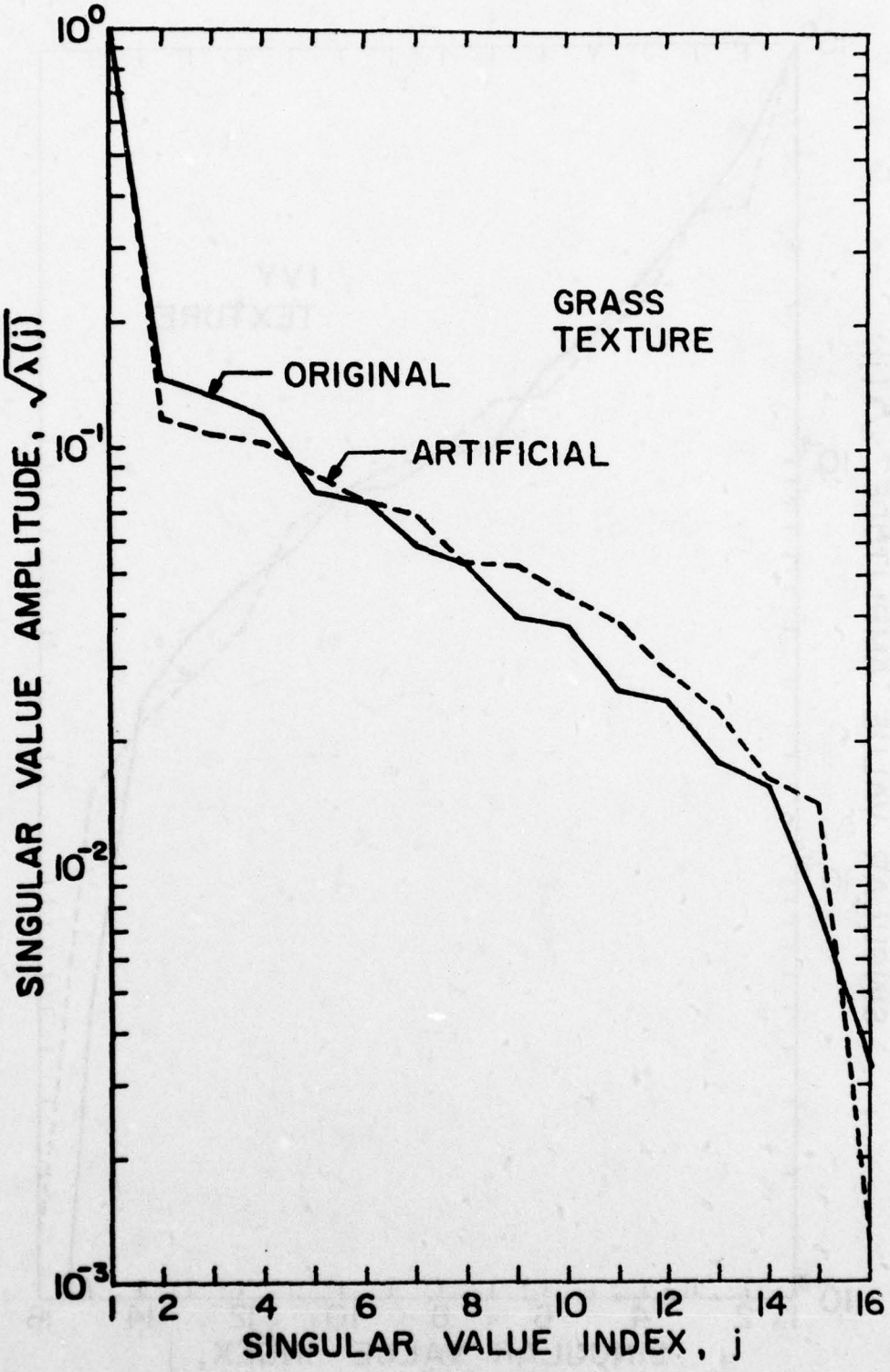


Figure 3. Singular Values of Natural and Artificial Grass Texture.

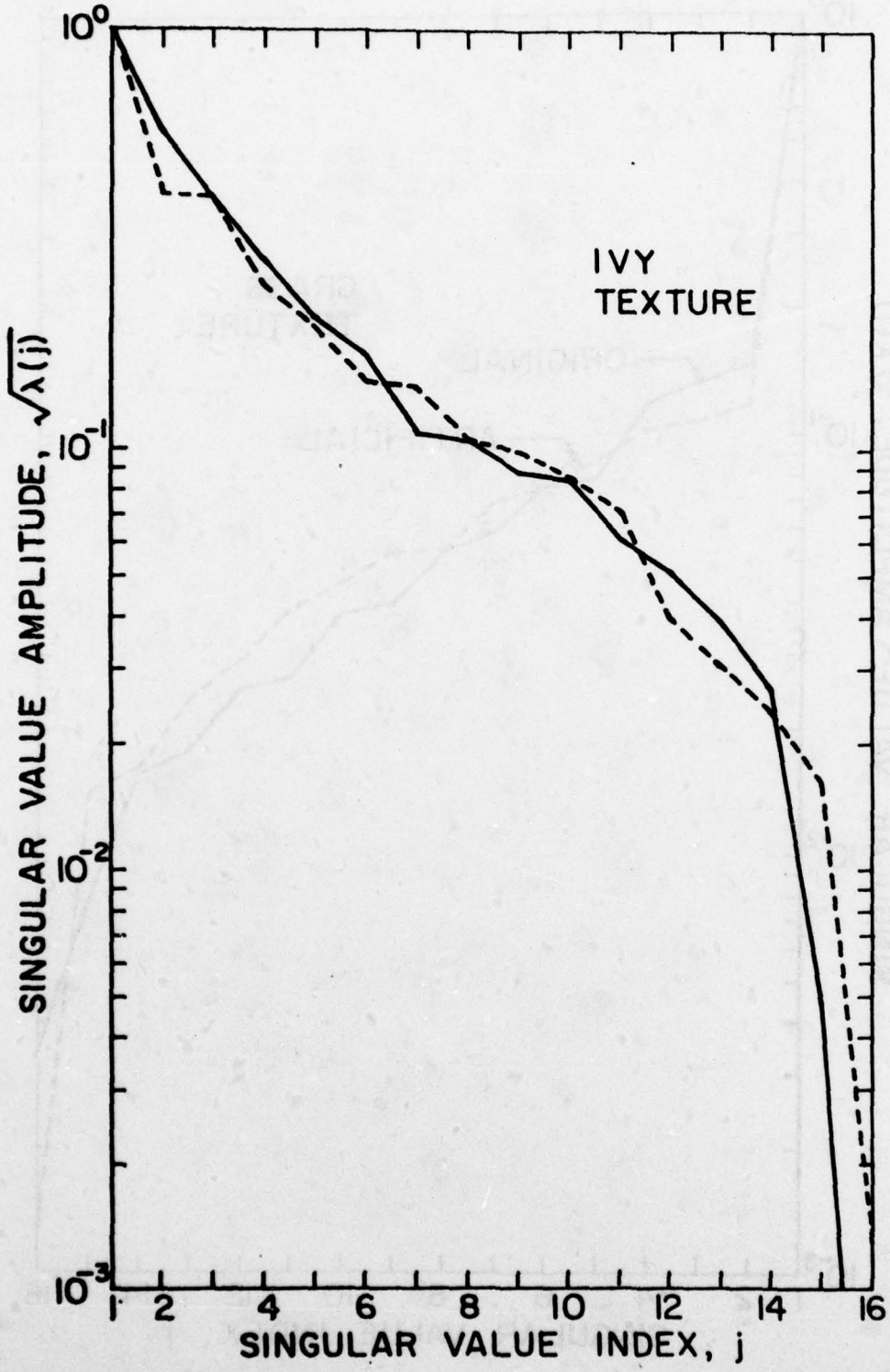


Figure 4. Singular Values of Natural and Artificial Ivy Texture.

between the actual and the ideal template in the decision procedure. Although this assumption is correct for the exact Hueckel's operator, two basic approximations have their effect on actual results. The first is the discretization procedure in which a continuous picture is replaced by discrete intensities at sampling points. The second is the use of a finite number of Fourier coefficients in the optimization procedure. In this paper the effect of these two approximations is discussed. In addition, some advantages of the Hueckel operator are explained.

Effect of Continuous to Discrete Mapping

This effect is pronounced in two points. The first is that a circular disk cannot be represented in the discrete domain. The second point is that the orthogonal bases used in the optimization procedure are derived in the continuous domain, and orthogonality is not exactly valid in the discrete domain. Although Hueckel has used averaging methods to reduce the previous effects, it is not clear that his procedure can be better than another operator designed completely in the discrete domain.

The Effect of Using Finite Number of Coefficients

Hueckel has used only nine coefficients in the optimization procedure. This choice was based on the fact that high frequency components are generally the result of noise. To determine the effect of neglecting high frequency components, simple edges and lines are reconstructed using the first nine components only. Results are shown in figure 1. It appears that the resulting images differ from the original, especially in the case of thin lines. This affects the optimality of the procedure. In addition it is not clear, in the case of finite number of components, that

grass and ivy are seen to be significantly different, but the distributions of the natural and artificial pairs are quite close.

Summary

The results presented are quite preliminary, but very encouraging. The SVD singular value distribution does indeed seem to be a viable means of characterizing spatial structure within an image block. Further work is underway to analyze the statistical properties of the singular values and to develop a quantitative measure of the singular value distribution.

References

1. G.H. Golub and C. Reinsch, "Singular Value Decomposition and Least Squares Solutions," Numerical Mathematics, Vol. 14, 1970, pp. 403-420.

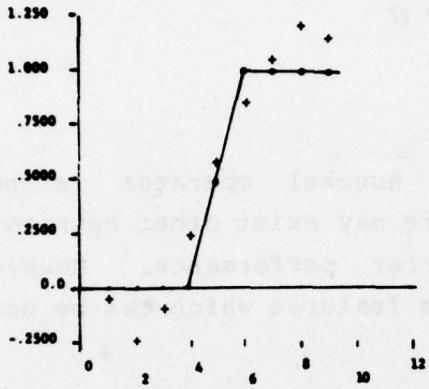
2. H.C. Andrews and C.L. Patterson, "Outer Product Expansions and Their Uses in Digital Image Processing," American Mathematical Monthly, Vol. 1, No. 82, January 1975, pp. 1-13.

2.6 Some Comments on the Hueckel Operator

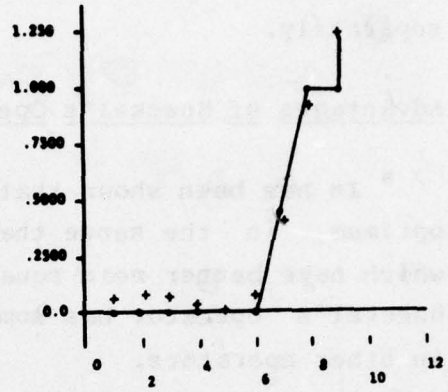
Ikram E. Abdou

Introduction

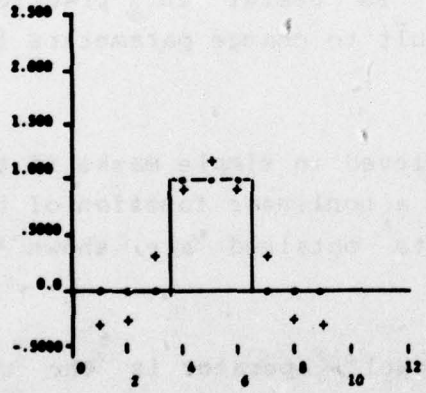
The Hueckel operator is one of the classical methods for edge detection [1,2]. It has been considered as an optimum method [3] because it uses the mean square error



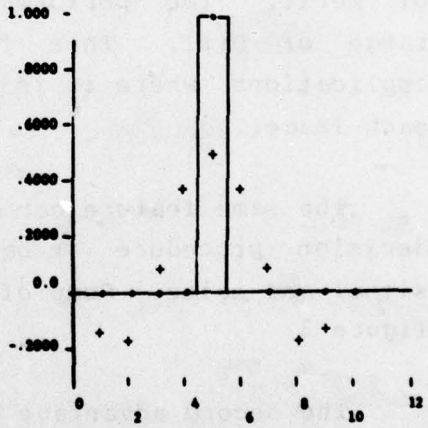
(a) Edge at center



(b) Edge at 7



(c) Line width = 3



(d) Line width = 1

Figure 1. • Original
+ Reconstruction

equal weighing of the error in different components result in optimum response. Other methods are to use different weights, or even to base the decision on each component separately.

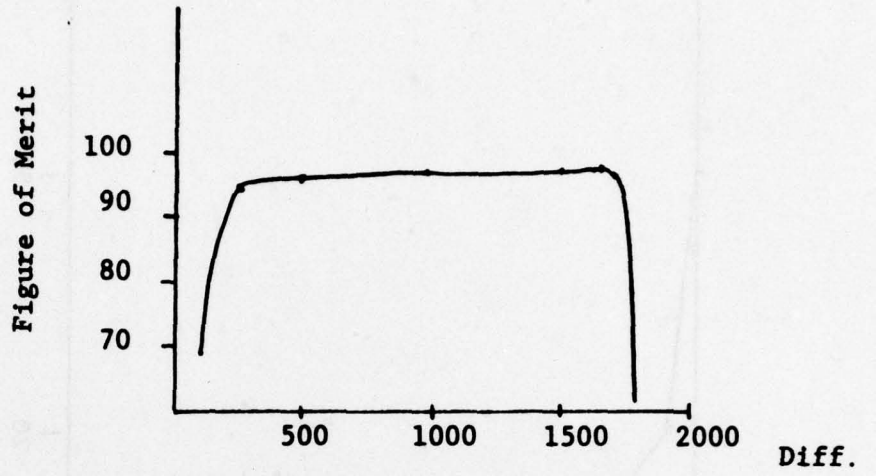
Advantages of Hueckel's Operator

It has been shown that the Hueckel operator is not optimum, in the sense that there may exist other operators which have better mean square error performance. However Hueckel's operator has some nice features which can be used in other operators.

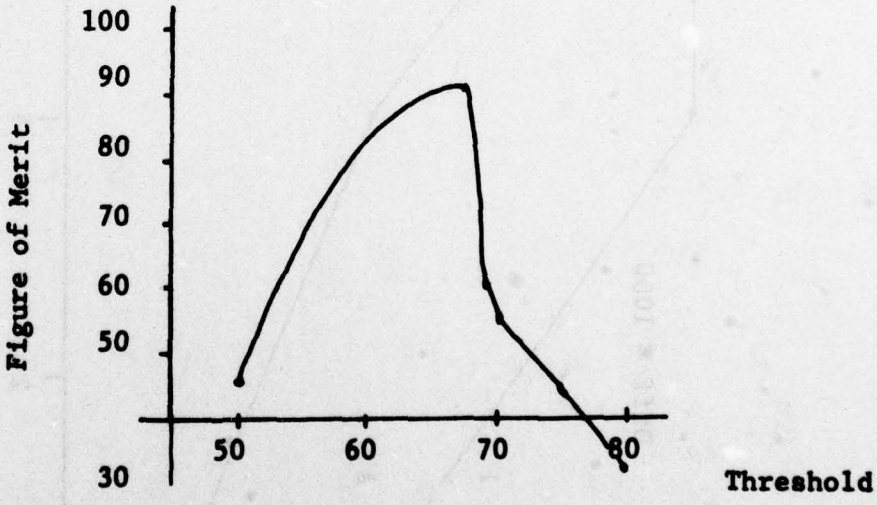
The first is that Hueckel's performance is less sensitive to parameter variation. Figure 2 shows a typical relation between the Hueckel parameter Diff and the figure of merit. The performance is almost constant over a wide range of Diff. This feature is useful in practical applications where it is difficult to change parameters for each image.

The same feature can be achieved in simple masks if the decision procedure is based on a nonlinear function of the signal and noise. Some of results obtained are shown in figure 3.

The second advantage of Hueckel's operator is the use of a large mask. This reduces the effect of noise in case of low SNR. However, large masks reduce the resolution of the operator, and the optimum size should be a compromise between SNR and resolution.



a. Hueckel (Conf = .85)



b. Simple (3x3) mask

Figure 2. Variation of figure of merit with edge detector parameters

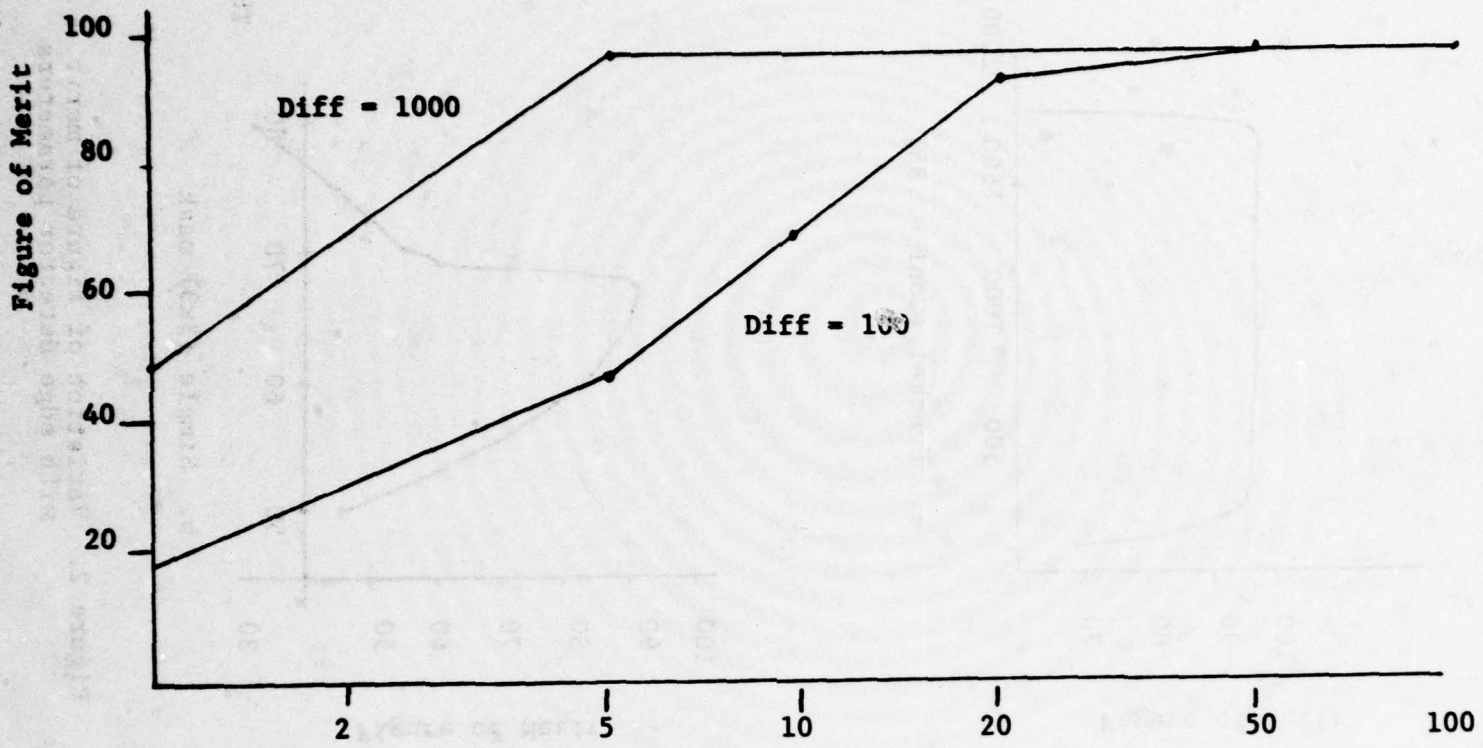


Figure 3. Figure of merit as a function of SNR for the Hueckel operator

Conclusion

In this paper some of the features of Hueckel's operator are discussed. It appears that the operator is not a standard method to which other operators should be compared. However it has its own advantageous properties which can be extended successfully to other edge detectors.

Acknowledgement

Dr. R. Nevatia has helped with many ideas about the use of Hueckel's operator.

References

1. M.H. Hueckel, "An Operator Which Locates Edges in *Digitized Pictures*," J. Association for Computing Machinery, Vol. 18, No. 1, January 1971, pp. 113-125.

2. M.H. Hueckel, "A Local Visual Operator Which Recognizes Edges and Lines," J. Association for Computing Machinery, Vol. 20, No. 4, October 1973, pp. 634-647.

3. L. Mero and Z. Vassy, "A Simplified and Fast Version of the Hueckel Operator for Finding Optimal Edges in Pictures," Proceedings 4th Internationl Joint Conference on Artificial Intelligence, 1975, pp. 650-655.

4. W.K. Pratt, "Figure of Merit for Edge Location," University of Southern California, Image Processing Institute, USCIPI Report 660, March 1976, pp. 85-93.

2.7 Development of Edge Detectors for the Extraction of Linear Segments

Peter Chuan

Numerous edge detectors have been suggested in the past. However, it is not clear as to which technique is most useful under different given conditions. Evaluation of edge detectors on real pictures is difficult and subjective in nature because a model for the image does not exist. Such evaluation cannot be separated from the goals of the processing. Here we describe an approach to the detection of edges that correspond to long, linear boundaries.

One often used edge detector is due to Hueckel [1] and has been extensively used in some of our previous work. For the purpose of detecting long, linear edges, its performance has been poor in the presence of noise. Figures 1a and 1b show a noisy image and the corresponding Hueckel edges, respectively. Here, the Hueckel operator is shown to be indiscriminately sensitive to all kinds of edges.

This calls for a particular edge detector tailored to be selectively sensitive to long, linear edges. Some crudely constructed edge detectors were used in the past [2] and have shown encouraging results. Six directional edge masks angularly distributed in approximately 30 degrees were convolved with the picture of interest. An approximated 60 degree edge mask is shown in figure 2. The maximum among the outputs of the six directional edge masks is taken and thresholded to show strong edges only. One such result obtained from figure 1a is shown in figure 3. Because the masks are binary (excluding 0) masks, gross approximations occur both in angular orientation and in spatial linearity of the edge line. Moreover, such edge masks are difficult

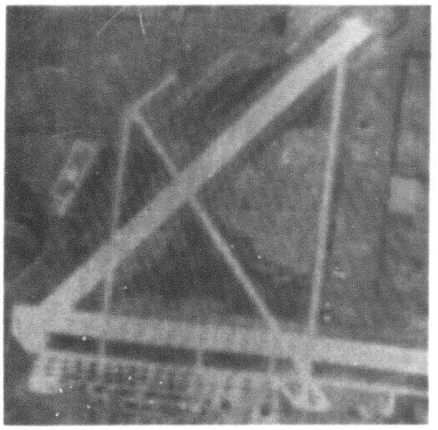


Figure 1a. Original image



Figure 1b. Hueckel edges

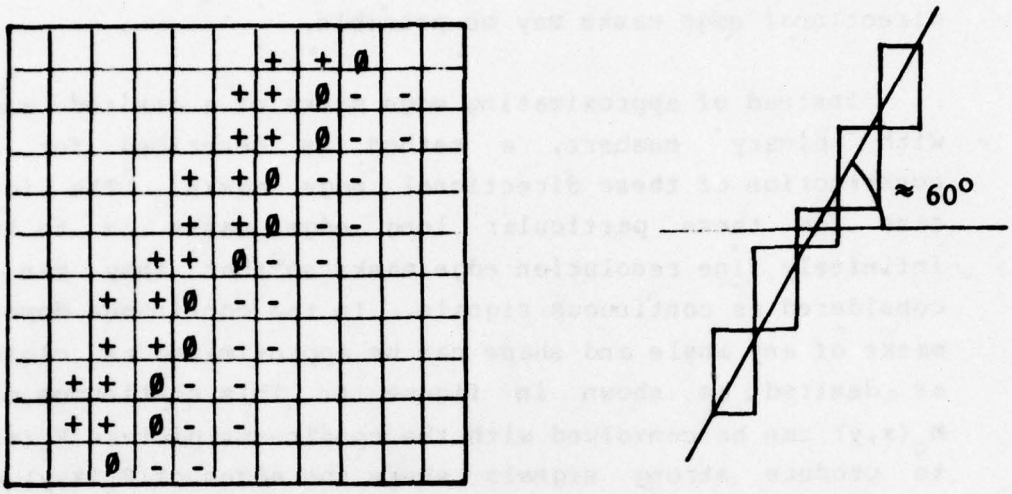


Figure 2. A grossly approximated binary (excluding 0) mask with an approximately straight edge of 60 degrees.



Figure 3. Output from long, linear edge detector of the kind shown in figure 2.

to construct and for a finite size mask, only a few directional edge masks may be possible.

Instead of approximating edge masks of a desired angle with binary numbers, a method is described for the construction of these directional edge masks. The ideal case for these particular long edge masks is to have infinitely fine resolution edge masks so that they can be considered as continuous signals. In the continuous domain, masks of any angle and shape can be approximated as closely as desired as shown in figure 4. This continuous mask $M_c(x,y)$ can be convolved with the continuous picture $F_c(x,y)$ to produce strong signals where the edges of $F_c(x,y)$ are long and continuous. This convolved output can be sampled and quantized to produce our ideal discretized edge signal $E_{cc}(m,n)$. For clarity, this process is illustrated in figure 5a.

Since we only have the sampled and quantized version $F_d(m,n)$ of the continuous picture $F_c(x,y)$ to work with, the ideal situation can only be approximated. The continuous edge mask $M_c(x,y)$ is therefore sampled and quantized and the discrete output $M_d(m,n)$ is convolved with $F_d(m,n)$ instead, to produce $E_{dd}(m,n)$. Again for clarity, this process is illustrated in figure 5b.

The price paid for using this approximate process versus the use of binary edge masks (figure 2) is in the greater number of quantization levels needed for the edge masks. However, the advantage gained is that this technique will allow the generation of directional edge masks of any desired size and orientation, with the possible application being the confirmation of the presence or absence of an edge with a certain direction or shape. Moreover, this model also opens up the possibility of an analytical evaluation of

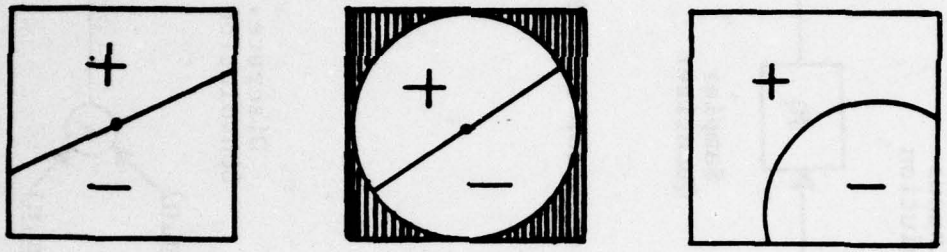


Figure 4. Examples of continuous masks that can be constructed for any angle and shape. The shaded region is weighted zero.

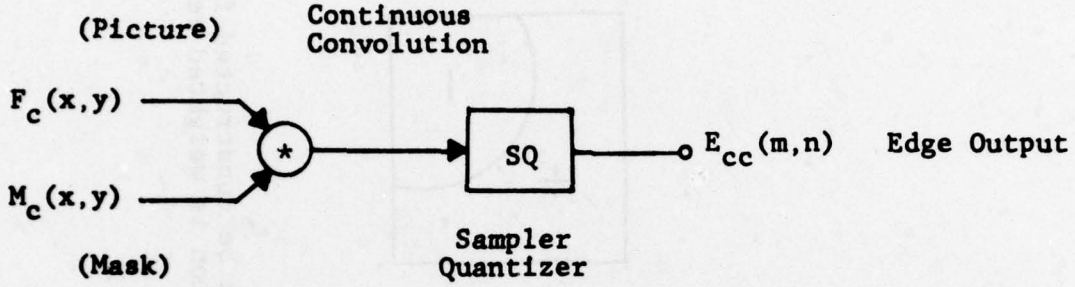


Figure 5a. Ideal case for edge detection.

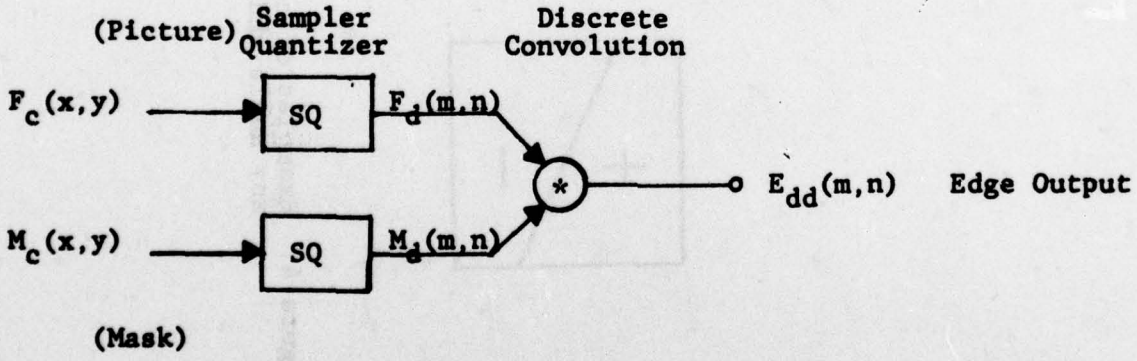


Figure 5b. Model for actual edge detection.

edge masks. Figure 6a shows a 60 degree edge mask generated through this model. This directional edge mask has weights obtained by integrating the continuous edge masks (shown in figure 6b) over an area covered by each square pixel. The continuous edge mask is two pixels wide on each side. Integration is carried out over each square pixel because this is the same process that will be carried out if the continuous mask were physically digitized by an Optronics scanner. Figure 7 shows the edge of an airport picture obtained by applying masks generated by this model.

One consideration in edge detection problems is on deciding the appropriate size of edge mask to use. Given a picture with small signal to noise ratio, larger masks should be allowed for detecting edges. Depending on the Fourier bandwidth (e.g. the closeness of two edges) of the information in the picture, the maximum mask size can be determined that will average out the noise but retain most of the signals. For example, comparing figures 8a and 8b, the mask with the size shown in broken lines will work fine in picking out edges in figure 8a but will obviously give weaker edges in between the two circles in figure 8b. Masks of two different sizes have been applied on a test pattern with signal to noise ratio (Signal Energy/Variance of Noise) equal to 2.7 and on the airport picture. Results are shown in figure 9.

Square masks have been used conventionally for directional edge masks. This is illustrated in figure 10. Convolutional outputs from each of these individual masks were then compared to extract the maximal value. It should be realized that mask 1 and mask 2 not only have different edge directions but as a result of this, they also have different shapes. Comparing mask 1 with mask 2 would be equivalent to comparing matched filter outputs using two

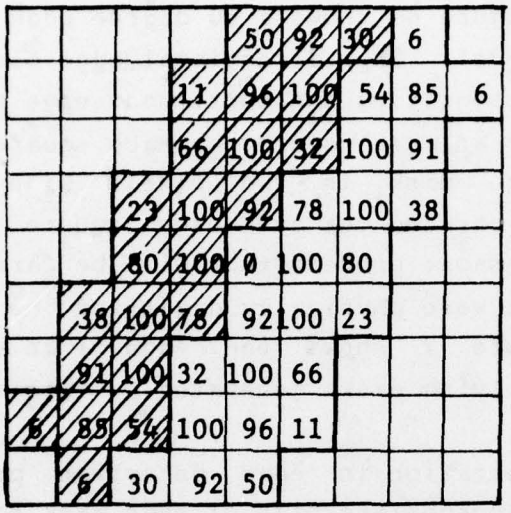


Figure 6a. Approximated 60^{o} edge mask (9x9) generated through the model. The shaded region contains negative weights.

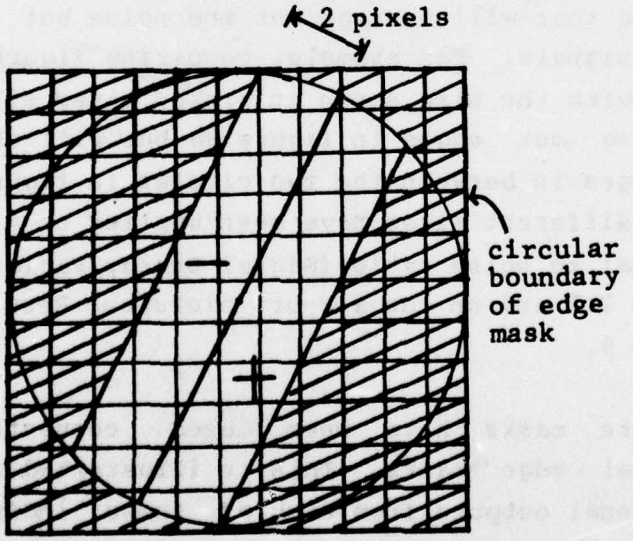


Figure 6b. Continuous edge mask. Shaded region contains 0 weight, positive region is weighted 1, negative region is weighted -1.

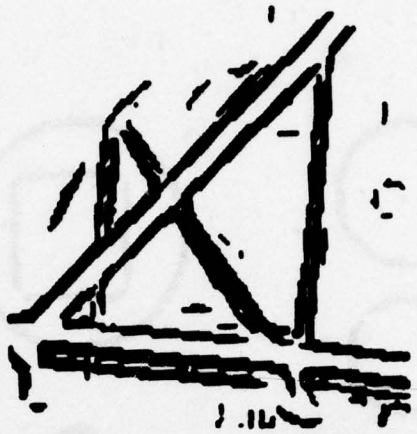
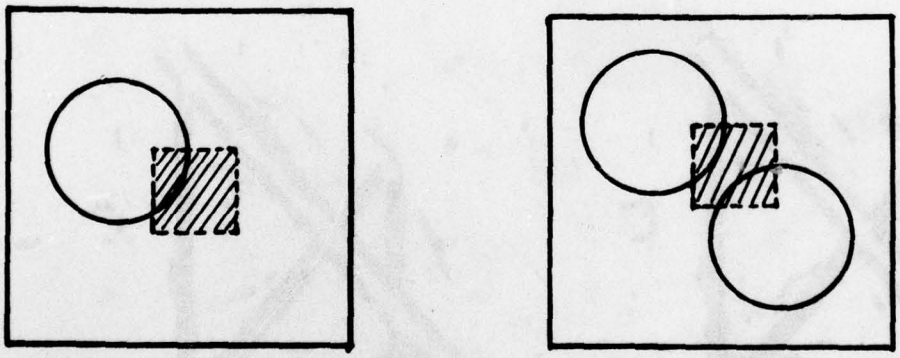
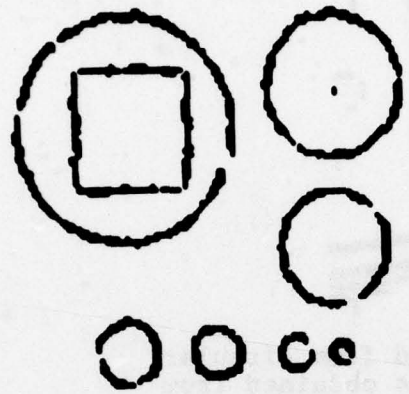


Figure 7. Edges obtained from circular weighted masks obtained from our proposed model.

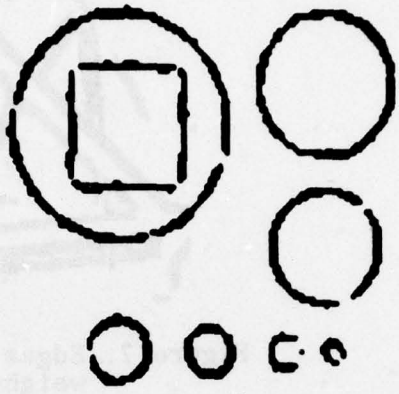


(a) (b)

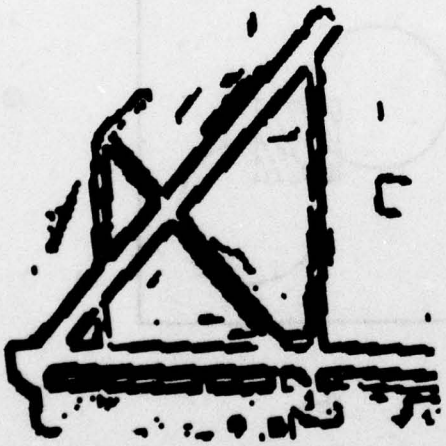
Figure 8. The square boxes show the same edge detecting mask. This mask is applied on both figures (a) and (b). At the location indicated by the position of the mask, the edge output from figure (b) will be lower than the edge output from figure (a).



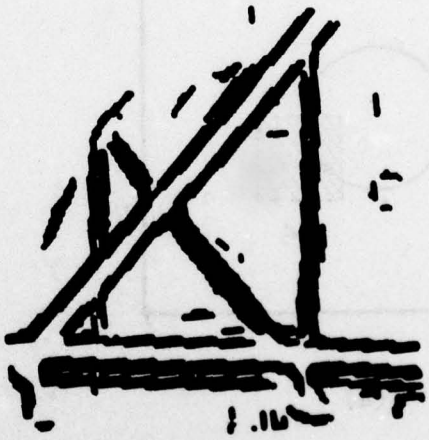
(a) Edges from 5x5 circular weighted masks on test pattern (SNR=0.36)



(b) Edges from 9x9 circular weighted masks on test pattern



(c) Edges from 5x5 circular weighted masks on airport picture



(d) Edges from 9x9 circular weighted masks on airport picture

Figure 9. Effect of using masks with sizes 5x5 and 9x9.

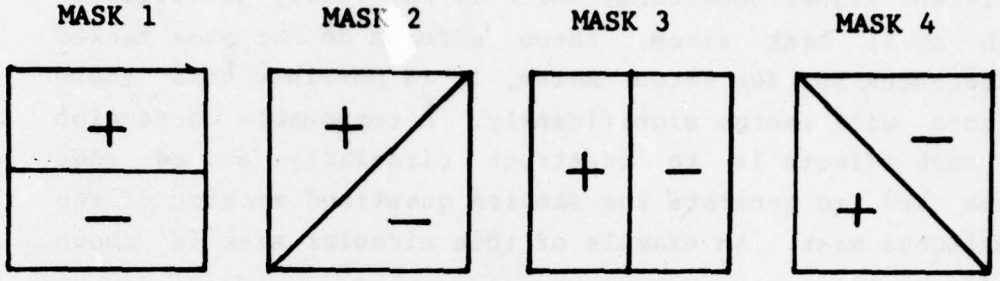


Figure 10. Square masks with zero entries only along the edge line.

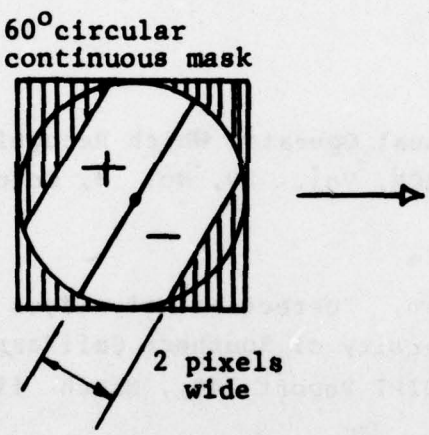


Figure 11a. Shaded region is weighted 0

60° circular discrete mask

Ø	64	100	25	17
23	100	92	78	83
80	100	Ø	100	80
83	78	92	100	23
17	25	100	64	Ø

Figure 11b. Shaded region is weighted negative

100	100	100	25	17
100	100	92	78	100
100	100	Ø	100	100
100	78	92	100	100
17	25	100	100	100

Figure 11c. Square mask

different signal detectors, which is not really justifiable. With small mask sizes, these effects do not show marked differences but for larger masks, it is possible that these effects will emerge significantly. A reasonable correction for such effects is to construct circularly shaped edge masks and to generate the sampled quantized version of the continuous mask. An example of this circular mask is shown in figure 11.

7 x 7 square and circularly bounded masks (listed in figure 13) have been used to show the effect of the different mask shapes. This is illustrated in figure 12.

References

1. M.H. Hueckel, "A Local Visual Operator Which Recognizes Edges and Lines," Journal ACM, Vol. 20, No. 4, October 1973, pp. 634-647.

2. R. Nevatia and P. Chuan, "Detection of Edges in Elongated Neighborhoods," University of Southern California, Image Processing Institute, USCIPI Report 740, March 1977, p. 34.

2.8 Circle Detection in Noisy Images

Kenneth I. Laws

This paper presents two methods of identifying circles within a noisy or textured image. A variant of the Hough transform method is outlined, then a new method for directed edge elements is developed and demonstrated.

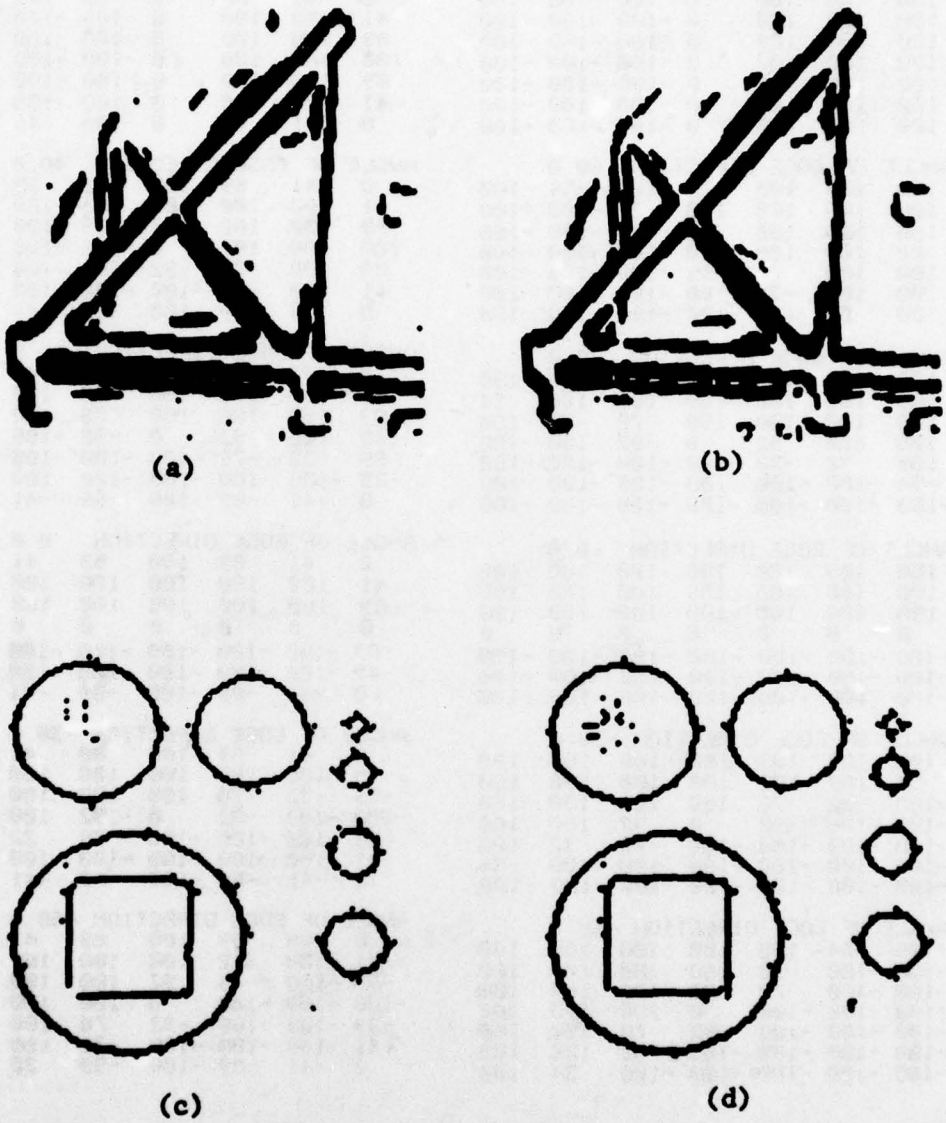


Figure 12a,12c Thresholded output from applying square masks.
12b,12d Thresholded output from applying circularly bounded masks.

The edges in (c) and (d) have been obtained by applying a thinning procedure.

ANGLE OF EDGE DIRECTION 90.0

100	100	100	0	-100	-100	-100
100	100	100	0	-100	-100	-100
100	100	100	0	-100	-100	-100
100	100	100	0	-100	-100	-100
100	100	100	0	-100	-100	-100
100	100	100	0	-100	-100	-100
100	100	100	0	-100	-100	-100

ANGLE OF EDGE DIRECTION 60.0

100	100	100	100	100	-54	-100
100	100	100	100	32	-100	-100
100	100	100	92	-78	-100	-100
100	100	100	0	-100	-100	-100
100	100	78	-92	-100	-100	-100
100	100	-32	-100	-100	-100	-100
100	54	-100	-100	-100	-100	-100

ANGLE OF EDGE DIRECTION 30.0

100	100	100	100	100	100	100
100	100	100	100	100	100	54
100	100	100	100	78	-32	-100
100	100	92	0	-92	-100	-100
100	32	-78	-100	-100	-100	-100
-54	-100	-100	-100	-100	-100	-100
-100	-100	-100	-100	-100	-100	-100

ANGLE OF EDGE DIRECTION 0.0

100	100	100	100	100	100	100
100	100	100	100	100	100	100
100	100	100	100	100	100	100
0	0	0	0	0	0	0
-100	-100	-100	-100	-100	-100	-100
-100	-100	-100	-100	-100	-100	-100
-100	-100	-100	-100	-100	-100	-100

ANGLE OF EDGE DIPECTION -30.0

100	100	100	100	100	100	100
54	100	100	100	100	100	100
-100	-32	78	100	100	100	100
-100	-100	-92	0	92	100	100
-100	-100	-100	-100	-78	32	100
-100	-100	-100	-100	-100	-100	-54
-100	-100	-100	-100	-100	-100	-100

ANGLE OF EDGE DIRECTION -60.0

-100	-54	100	100	100	100	100
-100	-100	32	100	100	100	100
-100	-100	-78	92	100	100	100
-100	-100	-100	0	100	100	100
-100	-100	-100	-92	78	100	100
-100	-100	-100	-100	-32	100	100
-100	-100	-100	-100	-100	54	100

(a) 7x7 Square edge masks

ANGLE OF EDGE DIRECTION 90.0

0	41	89	0	-89	-41	0
41	100	100	0	-100	-100	-41
89	100	100	0	-100	-100	-89
100	100	100	0	-100	-100	-100
89	100	100	0	-100	-100	-89
41	100	100	0	-100	-100	-41
0	41	89	0	-89	-41	0

ANGLE OF EDGE DIRECTION 60.0

0	41	89	100	89	-25	0
41	100	100	100	32	-100	-41
89	100	100	92	-78	-100	-89
100	100	100	0	-100	-100	-100
89	100	78	-92	-100	-100	-89
41	100	-32	-100	-100	-100	-41
0	25	-89	-100	-89	-41	0

ANGLE OF EDGE DIRECTION 30.0

0	41	89	100	89	41	0
41	100	100	100	100	100	25
89	100	100	100	78	-32	-89
100	100	92	0	-92	-100	-100
89	32	-78	-100	-100	-100	-89
-25	-100	-100	-100	-100	-100	-41
0	-41	-89	-100	-89	-41	0

ANGLE OF EDGE DIRECTION 0.0

0	41	89	100	89	41	0
41	100	100	100	100	100	41
89	100	100	100	100	100	89
0	0	0	0	0	0	0
-89	-100	-100	-100	-100	-100	-89
-41	-100	-100	-100	-100	-100	-41
0	-41	-89	-100	-89	-41	0

ANGLE OF EDGE DIRECTION -30.0

0	41	89	100	89	41	0
25	100	100	100	100	100	41
-89	-32	78	100	100	100	89
-100	-100	-92	0	92	100	100
-89	-100	-100	-100	-78	32	89
-41	-100	-100	-100	-100	-100	-25
0	-41	-89	-100	-89	-41	0

ANGLE OF EDGE DIRECTION -60.0

0	-25	89	100	89	41	0
-41	-100	32	100	100	100	41
-89	-100	-78	92	100	100	89
-100	-100	-100	0	100	100	100
-89	-100	-100	-92	78	100	89
-41	-100	-100	-100	-32	100	41
0	-41	-89	-100	-89	25	0

(b) 7x7 Circular edge masks

Figure 13. Edge masks used in figure 12.

A circle is characterized by its uniform boundary. The form of a circle may be perceived by humans even if the interior region is not uniform. It seems appropriate to search for circles with edge detection techniques rather than region growing methods.

A circle, or even a circular segment, is an extended entity. No local property of an edge element is sufficient to determine whether it is part of a circle. Curvature might be a sufficient local property, but it is not available from the standard edge detectors. It is the relationship among edge points which constitutes a circle. The edge elements must be equidistant from some center and the direction of each, if known, must be perpendicular to a line through that center.

True circles are also characterized by local continuity. If the images of circular objects were continuous there would be little difficulty in tracing the boundaries. Several processes interfere, however. The object itself may be slightly irregular, and the image may be corrupted by noise. Conversion to digital image form introduces quantization error. Finally, edge detectors can be applied to the image at only a finite number of points, and may be confused by texture in the image.

Fitting Circles Through Directed Edge Elements

Most edge detectors identify the direction of maximum gradient at each edge point. Adding 90 degrees to this angle gives the edge direction, d. The edge element (x,y,d) is thus a short directed line segment at (x,y), with the image known to be dark on one side and bright on the other. If the image contains a bright disk, the edge elements will circle it in a counter-clockwise (or positive) direction.

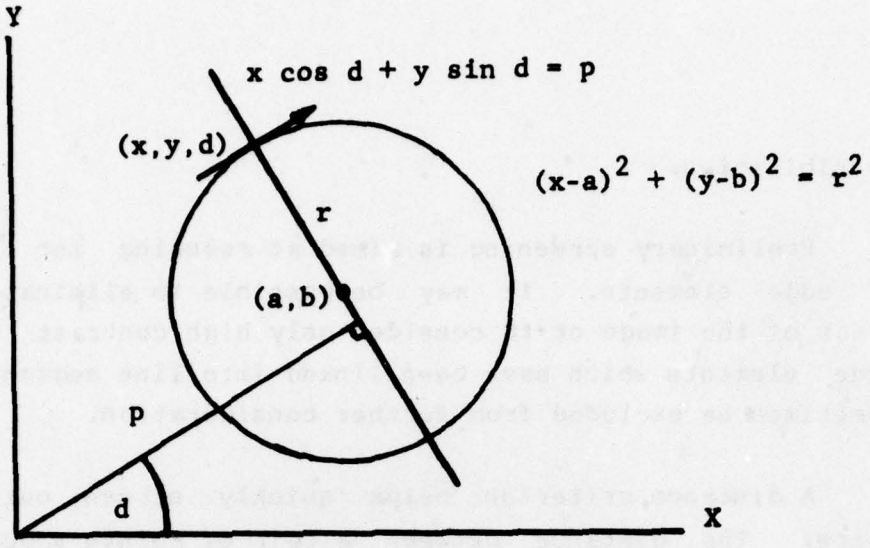
Circles through a given (x,y,d) have centers (a,b) on the line $a \cos(d) + b \sin(d) = p$, where $p = x \cos(d) + y \sin(d)$. Figure 1a shows this locus in X-Y space. This parameterization may be made unique either by restricting p to be positive or d to be between 0 and 180 degrees.

The Hough transform method [1,2] can be used to locate image circles. Each edge element generates a line in A-B-R space. The locus consists of points $(x + r \sin(d), y - r \cos(d), r)$, where the sign of r corresponds to the circle direction. If r is restricted to positive values, the locus also includes the line $(x - r \sin(d), y + r \cos(d), r)$. If edge directions are not known the locus of an edge point is a cone in A-B-R space.

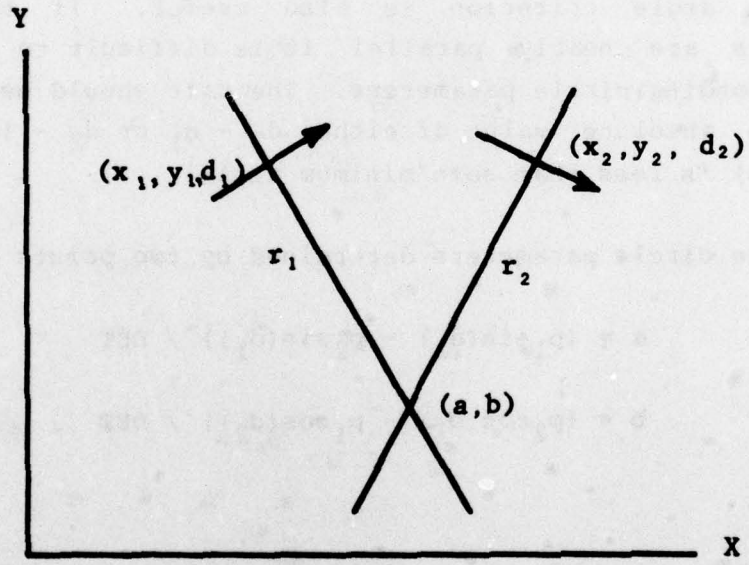
The A-B-R space is quantized to form an accumulator array. For each edge element the counters along the transform locus are incremented. Cells with high counts then correspond to image circles.

The above method requires the search of a three-dimensional transform space for clusters. The space must be quantized finely enough to provide one cell for each possible circle in the image. A more direct method utilizing pairs of edge elements will now be developed. It avoids quantization errors by searching for clusters in a continuous space.

A circle is determined by two of its directed edge elements, as shown in figure 1b. A space with N edge elements contains $N(N-1)/2$ possible circles -- one for each pair of points. A typical picture may have several thousand edge points, or millions of hypothesized circles. Simple screening can reduce this to a manageable number of



a) Locus of Circle Centers



b) Determination of Circle Parameters

Figure 1. Parametric Representations in X-Y space

possibilities.

Preliminary screening is aimed at reducing the number of edge elements. It may be possible to eliminate some areas of the image or to consider only high contrast edges. Edge elements which have been linked into line segments may sometimes be excluded from further consideration.

A distance criterion helps quickly screen out edge pairs. The distance between a pair of points should not exceed Dmax, the largest diameter being sought. This parameter is critical since further processing time is proportional to the square of Dmax. Larger circles can be found by line detection techniques.

An angle criterion is also useful. If two edge elements are nearly parallel it is difficult to find the corresponding circle parameters. The pair should be skipped if the absolute value of either $d_2 - d_1$ or $d_2 - (d_1 + 180$ degrees) is less than some minimum angle.

The circle parameters determined by two points are

$$a = (p_1 \sin(d_2) - p_2 \sin(d_1)) \; / \; DET$$

$$b = (p_2 \cos(d_1) - p_1 \cos(d_2)) \; / \; DET$$

where

$$p_i = x_i \cos(d_i) + y_i \sin(d_i),$$

$$DET = \cos(d_1)\sin(d_2) - \cos(d_2)\sin(d_1).$$

The radius may be computed separately for each edge point:

$$r_i^2 = (a - x_i)^2 + (b - y_i)^2 .$$

If the two values differ greatly, the two edge elements cannot be from the same circle. This fact may be used to screen the hypothesized circle centers and considerably reduce the number which must be processed further.

The direction of rotation of each edge element is another important parameter. The angle

$$d_i - \arctan(y_i - b) \; / \; (x_i - a)$$

will equal +90 degrees for positive circles and -90 degrees for negative circles. Two edge points do not belong to the same circle if they have opposite rotations about the common center. For line drawings these directions are not defined. Even if defined, it may be desirable to ignore them since disks do not always appear against uniformly darker or lighter backgrounds.

The identification of image circles has now been reduced to a standard clustering problem. The parameter points corresponding to an image circle should form globular clusters, possibly elongated in the R dimension. There is a relationship between the radius of an image circle and the number of edge pairs it produces, but this serves only as an upper bound on cluster membership if circular segments are also being sought.

Many algorithms exist for finding globular clusters. The chosen method must be insensitive to a large number of outliers or noise points. The method should also be fast and use little memory. It should identify the number of

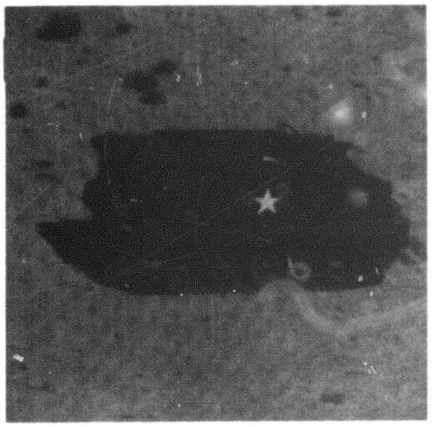
clusters a posteriori, but need not do so hierarchically. The clusters found should be independent of the order in which points are given, although this is not critical.

Experimental Results

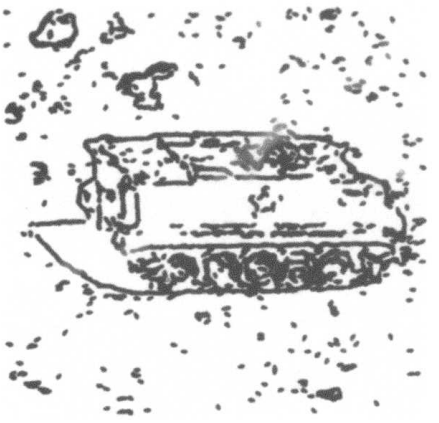
Figure 2a shows an armored personnel carrier against a desert background. Figure 2b shows the edge elements identified by a Hueckel operator [3]. This data base was taken as the starting point for the circle location problem. The image scale is considered to be 256 pixels in each direction.

The edge map contains 1,916 edge elements. There are thus 1,834,570 pairs of edge points to be screened. Figure 2c shows the X-Y positions of the 523 parameter points found by screening for 49 seconds on a PDP-10 KI. Edge pairs more than 15 pixels apart or deviating less than 60 degrees from each other were not considered. Radii computed for the two elements had to be within 10 percent of each other. The direction of rotation had to be the same for each element, and was used to form the sign of the radius value.

Clustering was then done in the A-B-R space using a Euclidean distance measure and a variant of Wishart's convergent K-Means algorithm for a variable number of clusters [4]. A circle with a radius of r pixels should produce a cluster of approximately r^2 parameter points, depending upon the screening criteria. Clusters were only kept if they contained more than three points and more than $0.25 * r^2$. Five clustering iterations were used, taking 15 seconds of computing time. Figure 2d is a plot of the ten circles corresponding to cluster centroids. The reference lines in the figure were found independently by the method of Nevatia [5].



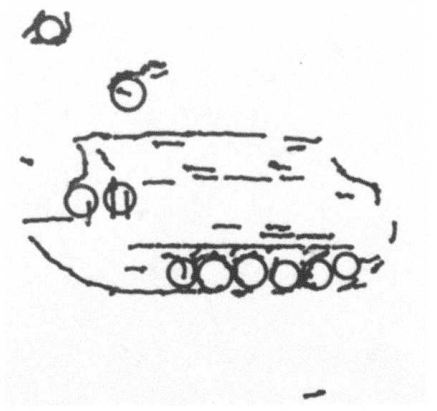
(a) Original Image



(b) Edge Data



(c) Hypothesized Circle Centers



(d) Accepted Circles

Figure 2. Experimental Results

The practical limit of 15 pixels for Dmax is adequate, since larger circles may be found by line detection and linking techniques. Better screening may be possible using edge strength or even curvature information. Experiments indicate that removal of straight segments from this data base greatly reduces the circle finding time, but larger circles are missed. Removal of edges belonging to circles might similarly aid line finding programs.

Without reference back to the edge data the circles can only be classified as complete and incomplete. Locating the corresponding edge elements is a simple best-fit search problem. Each set of points along a circle must be sorted or linked, then checked for continuity. Some circles, such as the vehicle wheels, will be dense and complete. Circles corresponding to rounded corners can be identified as short circular segments. Others, such as the bushes in the upper left corner of figure 2, may be identifiable as noise circles because of their lack of continuity.

References

1. R.O. Duda and P.E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," Communications of the ACM, Vol. 15, January 1972, pp. 11-15.

2. D.H. Ballard and J. Sklansky, "A Ladder-Structured Decision Tree for Recognizing Tumors in Chest Radiographs," IEEE Transactions on Computers, Vol. C-25, No. 5, May 1976, pp. 503-513.

3. M.H. Hueckel, "An Operator Which Locates Edges in Digitized Pictures," Journal of the ACM, Vol. 18, 1971, pp. 113-125.

Best possible scan

4. M.R. Anderberg, Cluster Analysis for Applications, Academic Press, New York, 1973.

5. R. Nevatia, "Locating Object Boundaries in Textured Environments," IEEE Transactions on Computers, Vol. C-25, No. 11, November 1976, pp. 1170-1175.

2.9 Analytic Results of the Coleman Segmentor

Harry C. Andrews

Automatic bottom up human unassisted image segmentation has been developed by Coleman [1] for the Image Understanding program. The system utilizes pattern recognition techniques in N dimensional vector space to perform decorrelation, clustering, feature rejection and ultimate segmentation. The only underlying assumption for the process is that homogeneous clusters in N space are representative of homogeneous regions of an image in perceptual space. The system is designed to operate with any set of computable features and will automatically select the best subset of those features to develop tightly clustered homogeneous regions in N space which then serve to define the segmentation of the original image. In the interest of smart sensor implementation, the system has been designed for frame-to-frame segmentation for real time television-like sensors.

Segmentor Configuration

Figure 1 presents a block diagram representative of the system design of the segmentor. The first component of the system is the "feature computation" phase. This process computes the features that the designer feels will be

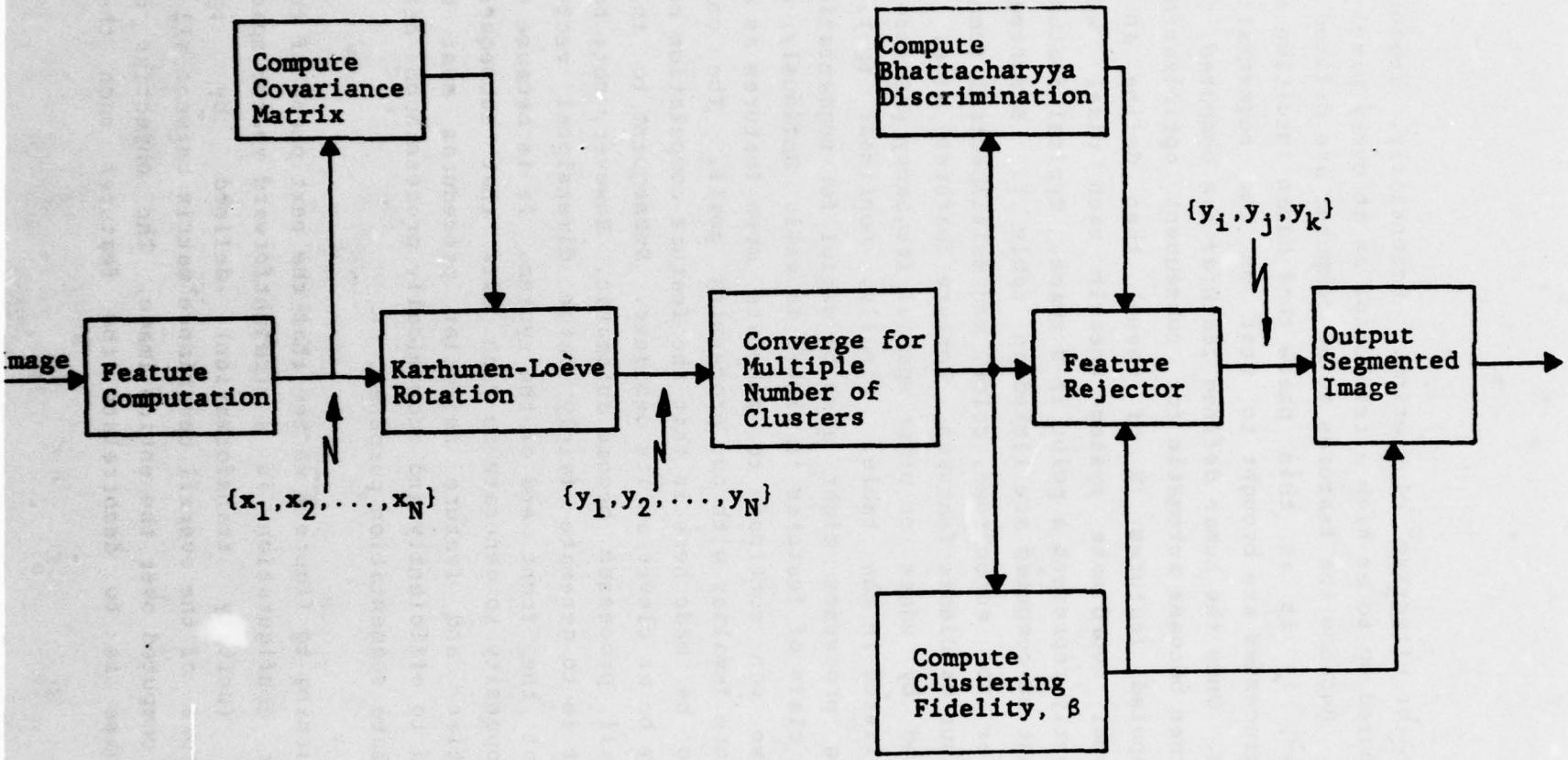


Figure 1. Segmentor Configuration

Best possible scan

relevant for effective clustering. Essentially, features are computed up to as high a resolution as at every pixel if desired. Because the features to be computed are defined by the user, it is at this phase that human intuitive and design processes are brought to bear on the segmentation problems. Once the human defined features are computed, the system then becomes automatic for subsequent optimization. The computed features (N of these) then define an N dimensional coordinate system wherein each pixel will subsequently represent a point in N space. Typical features that might be computed are listed in table 1. Brightness amplitudes for monochrome, color, and multispectral scenes are obvious candidate features. Texture features might be delineated by edges or other spatial frequency processors and are listed in the table. Finally, nonlinear spatial filtering processes might also be useful for segmentation and this class of features is listed as well. Obviously, as humans we can continue to generate more features as we become more familiar with our processing goals. The only point to be made here is that the feature computation box will only be as clever as its designer. Subsequent to this phase, all processes become automatic. However, note how simple it is to generate fairly large dimensional vector spaces at the front end of the system. It is because of man's propensity to generate so much data that subsequent optimization and feature rejection procedures must be developed to efficiently and economically process such data for ultimate segmentation purposes.

Returning to figure 1 we see that the next phase of the segmentor configuration is a straightforward vector space rotation (unitary transformation) defined by the eigenvectors of the overall covariance matrix between all N features computed over the entire image. The objective of this phase is to decorrelate the features such that

FEATURE TABLE 1

INDEX	FEATURE DESCRIPTION	FEATURE CLASS
x_1	monochrome brightness	monochromatic amplitude
x_2	red color brightness	color amplitude
x_3	green color brightness	
x_4	blue color brightness	
x_5	band 1 brightness	multispectral amplitude
x_6		
⋮		
x_{10}	band 6 brightness	
x_{11}	Sobel magnitude on x_1	texture feature
x_{12}	Sobel magnitude on x_2	
⋮		
x_{20}	Sobel magnitude on x_{10}	
x_{21}	Sobel phase on x_1	texture orientation
⋮		
x_{30}	Sobel phase on x_{10}	
x_{31}	mode filter on x_1	nonlinearly filtered feature
⋮		
x_{40}	mode filter on x_{10}	
x_{41}	dispersion filter on x_1	
⋮		
x_{50}	dispersion filter on x_{10}	

clustering is implemented in N dimensional decorrelated space. In this way good features can be selected individually and bad features rejected individually without concern as to correlation properties with other features. This will allow efficient compaction of good clustering features into a few parameters thereby providing a large dimensionality reduction. However it is important to realize that feature reduction does not occur immediately following the rotation process but only subsequent to clustering analysis.

This brings us to the next step in the system which is a k-means clustering algorithm in N dimensional rotated space. This algorithm converges to a set of k-mean points describing the best assignment of pixel features to k-clusters such that the sum of within cluster distances is the smallest. The disadvantage of the algorithm is that it requires knowledge of the number of clusters, k, in advance. Clearly this is unknown and consequently the k-means clustering routine must be implemented for all reasonable values of k (i.e. k=1,...,16). Subsequent blocks in the figure are designed to determine the best number of cluster and the best features to provide the tightest cluster distributions.

Once the k-means cluster algorithm has converged to the minimum spread of points in N space, a fidelity measure, β, is computed to establish the tightness of the points within the clusters and the degree of spread or separateness of the clusters one from another. This fidelity measure is given by

$$\beta(k) = tr[S_w(k)] \quad tr[S_b(k)]$$

where $[S_w(k)]$ is the within cluster scatter matrix and $[S_b(k)]$ is the between cluster scatter matrix [2]. It can

be shown that β is everywhere nonnegative, has at least one maximum, and achieves that maximum where the ratio of the within cluster scatter equals the between cluster scatter. Therefore it is hypothesized that the optimal number of clusters (k) occurs at β equal to its maximum. Therefore these values of β and k are used to control the output segmentor and the feature rejector.

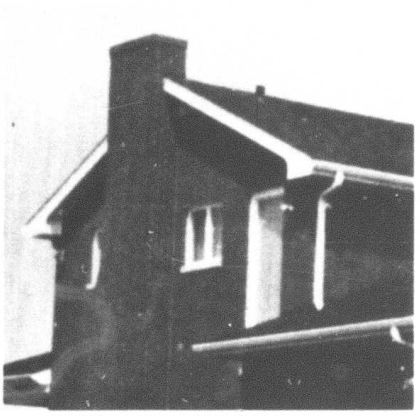
The feature rejector provides the function of removing those features which do not contribute to tight homogeneous clusters. Consequently, this process borrows from supervised pattern recognition theory in which feature selection/rejection is often implemented through the use of the Bhattacharyya distance function [3]. This function provides a measure of the usefulness of a particular dimension or feature by investigating that feature's ability to separate the data points into the proper clusters determined by the k-means convergence algorithm. This measure is provided by mean and variance parameters determined by each dimension for all the clusters. Those features or dimensions which do not provide well-defined clusters (due to separate means and tight variances) are rejected, thereby leaving good features for more tightly homogeneous clusters.

Experimental Results

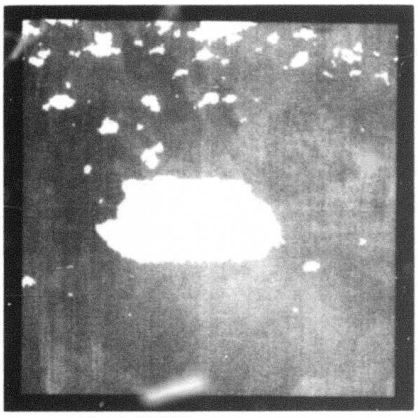
A variety of images have been segmented using the above clustering algorithm with varying degrees of perceptual success. Figures 2 and 3 present these results in pictorial form. Figure 2a and 3d were original monochrome images while figure 2d was a color image and figure 3a was a ten band multispectral image. Various clustering results are presented for each image for viewer inspection. The last sequence in figure 3 represents clustering on frame-to-frame



a) Original



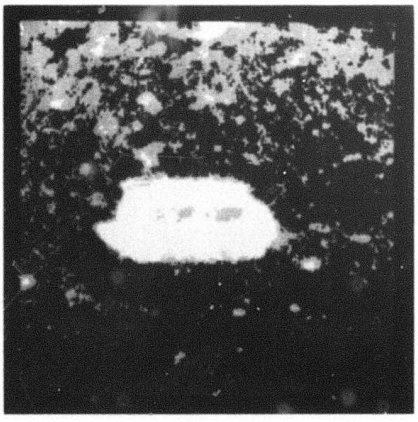
d) Original



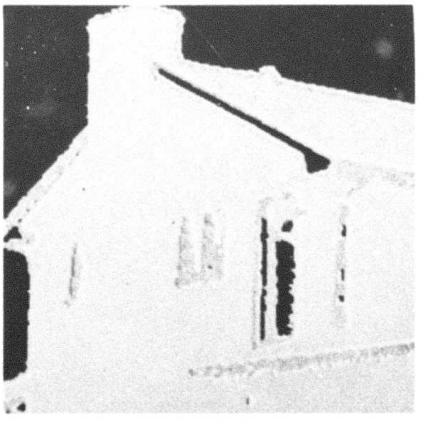
b) 2 Clusters



e) 2 Clusters

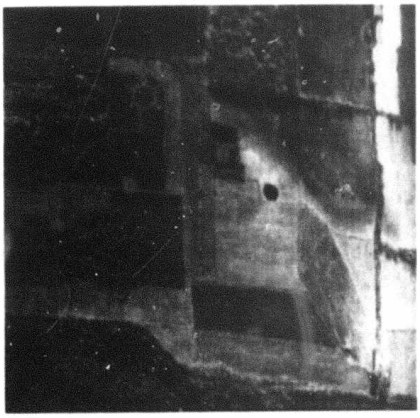


c) 3 Clusters

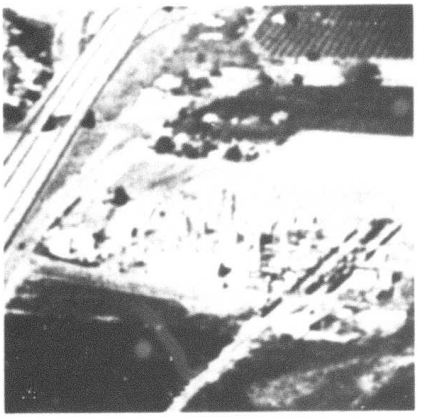


f) 3 Clusters

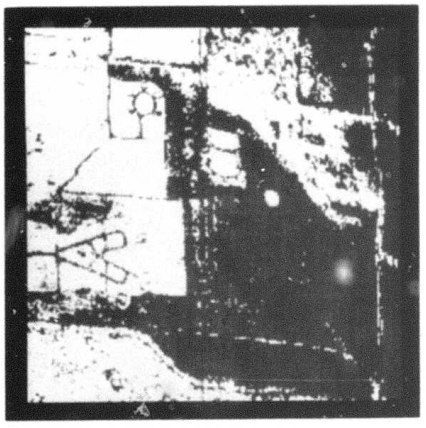
Figure 2. Pictorial Clustering Results



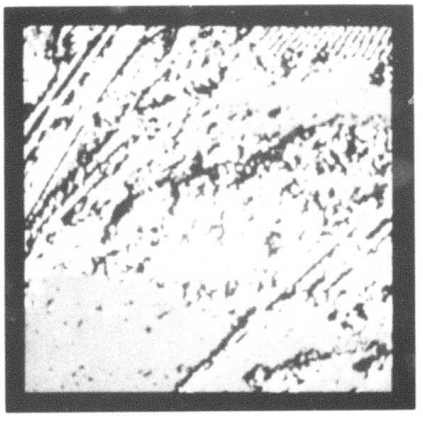
a) Original



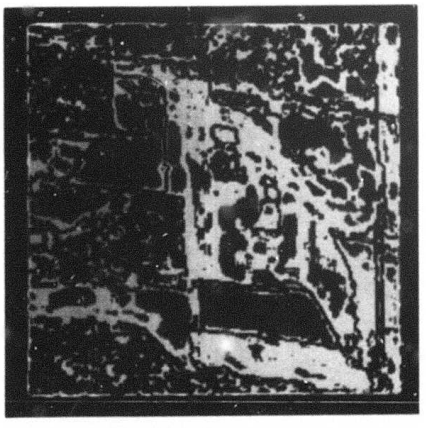
d) Original



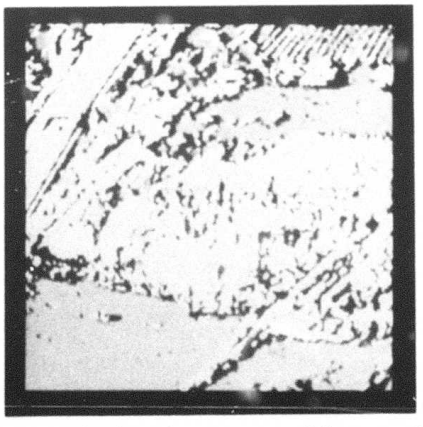
b) 2 Clusters



e) 4 Clusters (frame 1)



c) 3 Clusters



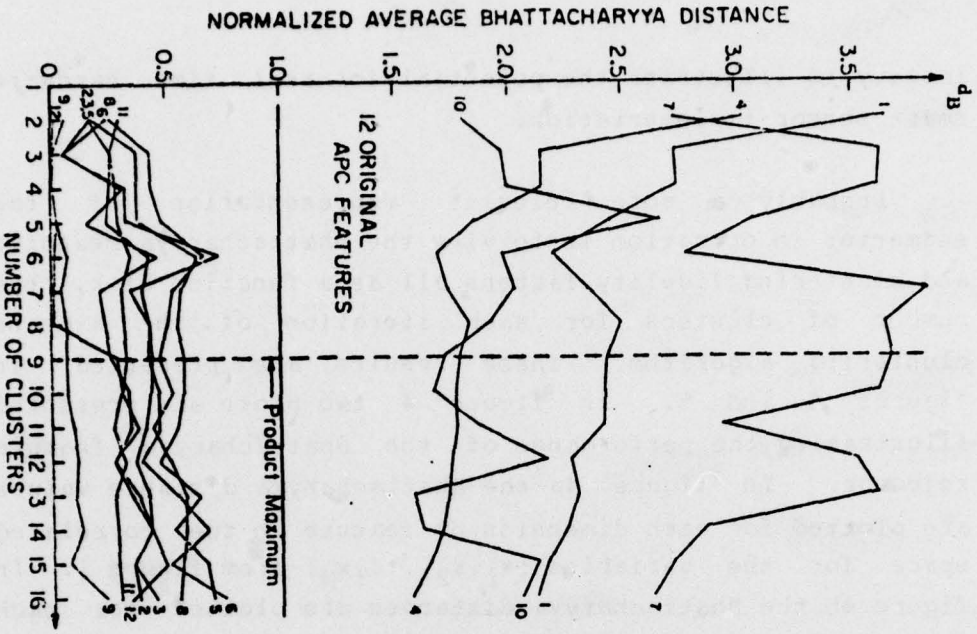
f) 4 Clusters (frame 5)

Figure 3. Pictorial Clustering Results

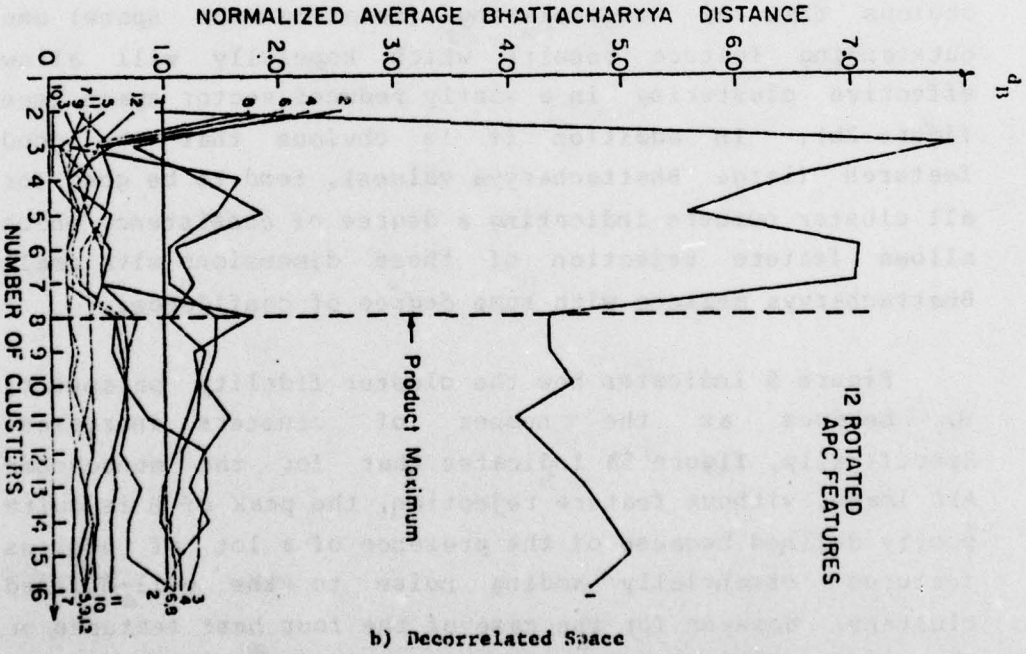
imagery to illustrate the potential for real time hardware smart sensor implementation.

Probably a more relevant representation of the segmentor in operation is to view the Bhattacharyya measures and clustering fidelity factors all as a function of k, the number of clusters for each iteration of the k-means clustering algorithm. These results are presented in figures 4 and 5. In figure 4 two plots are presented illustrating the performance of the Bhattacharyya feature rejector. In figure 4a the Bhattacharyya distance values are plotted for each dimension or feature in the correlated space for the variables $\{x_1, x_2, \ldots, x_N\}$ from figure 1. In figure 4b the Bhattacharyya distances are plotted for each rotated dimension or feature in the decorrelated space for the variable $\{y_1, y_2, \ldots, y_N\}$ of figure 1. It is immediately obvious that by decorrelating (rotating the space) one outstanding feature results which hopefully will allow effective clustering in a vastly reduced vector space (see figure 2b). In addition it is obvious that the good features (large Bhattacharyya values), tend to be good for all cluster numbers indicating a degree of consistency which allows feature rejection of those dimensions with small Bhattacharyya measure with some degree of confidence.

Figure 5 indicates how the cluster fidelity parameter, β, behaves as the number of clusters increases. Specifically, figure 5a indicates that for the monochrome APC image, without feature rejection, the peak of β is quite poorly defined because of the presence of a lot of useless features essentially adding noise to the well-defined clusters. However for the case of the four best features or the single best feature, a much more marked peak results at a lower cluster number. A similar effect occurs for the colored house of figure 5b. However from the curves of all

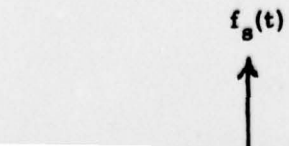


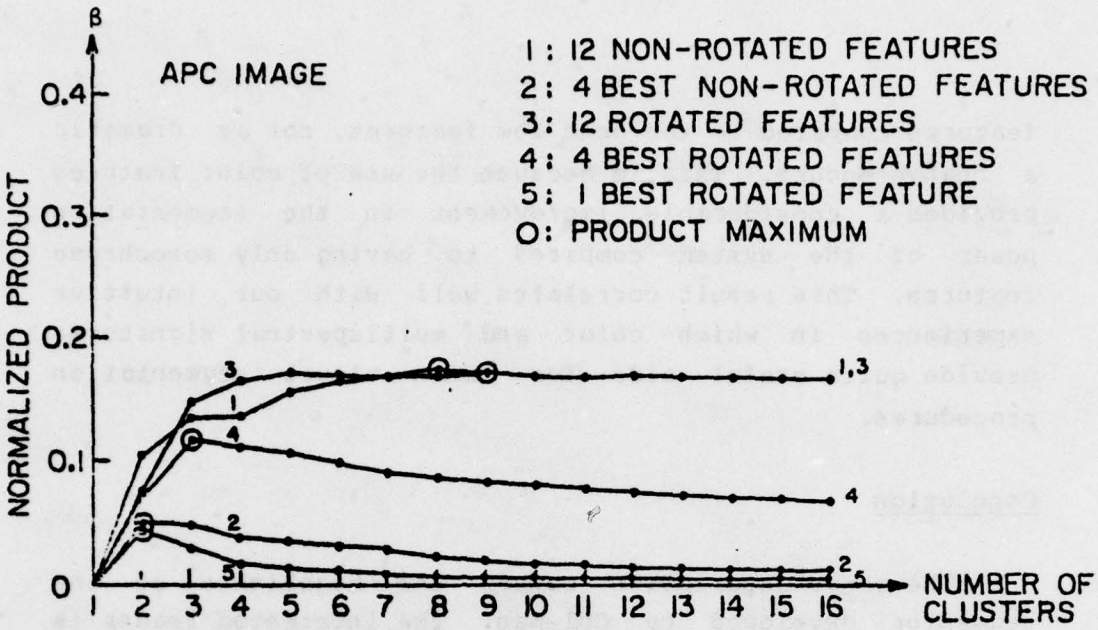
a) Correlated Space



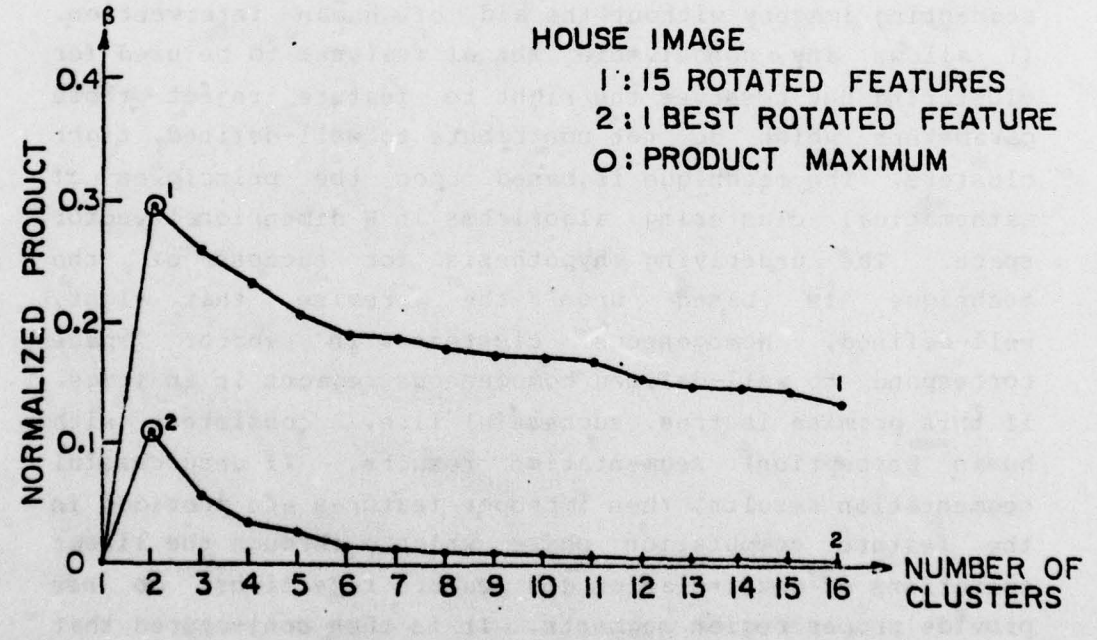
b) Decorrelated Space

Figure 4. Bhattacharyya Feature Rejector for the APC Image.





a) APC



b) Colored House

Figure 5. Clustering Fidelity Measure.

features compared to the best few features, not as dramatic a change occurs. This is because the use of color features provides a considerable improvement in the segmentation power of the system compared to having only monochrome features. This result correlates well with our intuitive experiences in which color and multispectral signatures provide quite useful aids for human visual segmentation procedures.

Conclusion

The above description covers the highlights of the segmentor developed by Coleman. The interested reader is referred to reference [1] for details of the system. The algorithm represents a bottom up attempt at automatically segmenting imagery without the aid of human intervention. It allows any conceivable set of features to be used for clustering but reserves the right to feature reject those parameters which do not contribute to well-defined, tight clusters. The technique is based upon the principles of mathematical clustering algorithms in N dimensional vector space. The underlying hypothesis for success of the technique is based upon the premise that tight, well-defined, homogeneous clusters in vector space correspond to well-defined homogeneous regions in an image. If this premise is true, successful (i.e. consistent with human perception) segmentation results. If unsuccessful segmentation results, then improper features are provided in the feature computation phase which, through the linear operations of decorrelation and feature rejections, do not provide proper region segments. It is then conjectured that nonlinear transformations (or other features) are necessary. Finally, the segmentor has been designed with smart sensor real time implementation in mind. The hardware construction of such a system is under contemplation.

References

1. G. Coleman, "Image Segmentation by Clustering," University of Southern California, Image Processing Institute, USCIPI Report 750, August 1977.

2. For the rather involved details of this fidelity measure and its derivation the interested reader is referred to reference [1] above.

3. H.C. Andrews, Introduction to Mathematical Techniques in Pattern Recognition, Wiley-Interscience, New York, New York, 1972.

3. Image Processing Projects

This section of the report describes the efforts expended and results obtained over the past six months on the various image processing projects carried out at the Institute. Some of these projects have been funded by other sources as indicated in the appropriate title acknowledgements. Results of work in defining the degrees of freedom inherent in radar imaging systems are presented for the stripping mode of SAR. It is expected that future radar imaging model degrees of freedom analysis will lead to efficient radar image understanding, a task which for humans, is far more difficult than visual image understanding. This section also summarizes the psychovisual modelling work being applied to image coding and image rate distortion functions. Results in blind a posteriori image restoration are next presented. Finally two new optical processing procedures are described in which on-axis holographic optical filtering is developed, and optical pseudocoloring of texture information is described.

3.1 Synthetic Aperture Radar and Imaging System of the Stripping Mode

Chung-Ching Chen

Introduction

The concept of radar is relatively simple although its implementation often is not. It is an active device which operates by radiating electromagnetic waves and estimating the characteristics of the target by the echoes returned from the objects. Since its appearance in World War II, radar has played a very important role in both military and civilian applications, such as the target detection, navigation of the ships and aircraft, etc [1]. The purposes of the radar can be dichotomized as target detection and parameter estimation. Detection of a target is the determination of its presence in the unavoidable noisy situation, and parameter estimation is the measuring of characteristics of the targets, e.g., their ranges, relative velocities, angular directions, sizes, etc, by the extraction of available information from the received echoes [1].

The power of an airborne or spaceborne ground mapping radar is limited by its resolving abilities in both azimuth and range directions, whereby azimuth is "along the flight track" and range is that perpendicular to it on the ground. Range resolution can be achieved by the radiation of a short pulse and the accurate measurements of the time of its return. Alternatively, the pulse-compression techniques can be applied to obtain similar resolution with greatly increased signal power [2]. Here a suitable modulation, usually a linear F.M. or chirp signal, because of its high efficiency in terms of time-bandwidth product, easy analysis

and implementation, is imposed upon a CW of moderate time duration. In this technique the large bandwidth which is necessary for accurate range resolution is created while a large signal power is also achieved because of the "stretching" of the signal duration compared to short pulse modulation, which has a large bandwidth but short time duration. Upon receiving the echoes, the data processor "compresses" the large time-bandwidth product pulse to reach virtually the same range resolution. It can be shown that the signal is essentially the point spread function (PSF) of the radar imaging system in the range dimension. Although the signal is long in time and thus the PSF is wide in range space, the range resolution is not degraded because of the high-bandwidth property of the signal. Hence the simultaneous achievements of high resolution and large signal power are possible. It is pointed out, however, that because the pulse is in coded form, a decoding scheme which in this case is pulse compression which is coherent in principle, has to be adopted, as contrasted to noncoherent processing for traditional range resolution schemes. The coherent processing provides the possibility of another point of view on the imaging system - the optical or hologram processing.

From antenna theory it is well known that the half power beam width β in radians of a physical antenna of length L is (see figure 1)

$$\beta = \lambda/L \qquad (1)$$

where λ is the wavelength of the radiation. The groundwidth projected by the half power beam is then $\Delta_z = \beta \cdot R = \frac{R\lambda}{L}$, where R is the range.

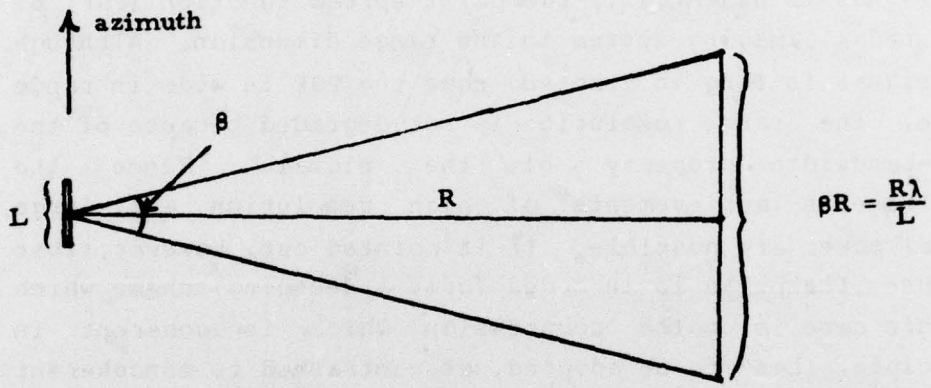


Fig. 1. Antenna Illumination Geometry

Traditional ground mapping radar achieves azimuth resolution by using antenna with beam narrow in azimuth. Physically speaking, the narrower the beam width, the narrower the point spread function and thus the better resolution obtainable. This is the same situation as discussed in the case of uncoded short pulses used to obtain range information. It is again noted that here the illumination imaging is noncoherent among the patches illuminated at different antenna locations. Thus the wider beam width means more blurring, hence fine azimuth resolution demands a very long physical antenna (L large in $\frac{R\lambda}{L}$), most of the time not practically available.

A comparison between the imaging schemes along range and azimuth directions suggests that high azimuth resolution is possible by the use of some forms of coding. However, since the ground is two dimensional (range and azimuth) while the signal is only one dimensional (a function of time only) a direct simultaneous compression upon the signal itself is not conceivable, if not impossible. One solution is to azimuth-modulate the returned echo, instead of modulating the signal before radiation. This is made possible by the constantly changing relative location between the ground points and the radar. The technique thus developed is called synthetic aperture radar (SAR).

The azimuth resolution abilities can be evaluated from another point of view. The aircraft during its flight radiates pulses at different locations and receives their echoes shortly after. Because of the azimuth modulation incurred by the relative motion of the craft and ground, coherent processing upon the received data to obtain maximum possible resolution is required as explained earlier. This is analogous to the case of an antenna array where the received signals at each array element are coherently

processed and summed [3].

From figure 2 it is easily seen that the width of the azimuth beam at range R gives the maximum value for the length of synthetic aperture that can be used at that range. Hence

$$L_{eff}=R\beta=\frac{R\lambda}{L} \qquad (2)$$

is the effective length of the synthetic antenna aperture where L is the azimuth size of the physical antenna. However, since the operation of SAR utilizes the two-way beam pattern in the sense that phase shift is introduced on both the paths and from the target, this round-trip phase shift effectively reduces the wavelength by a factor of 2. Thus

$$\beta_{eff} = \frac{\lambda}{2L_{eff}} \qquad (3)$$

The azimuth resolution δ_a is the effective beam width projected on the ground at range R

$$\delta_a = \beta_{eff}R = \frac{\lambda R}{2L_{eff}} = \frac{\lambda R}{2\frac{R\lambda}{L}} = \frac{L}{2} \qquad (4)$$

which is independent of λ and R, and is proportional to the azimuth size of the physical antenna. Thus to achieve higher azimuth resolution, a shorter antenna is to be used, in contrast to the single antenna case in eq.(1).

Basically there are three modes of ground mappings of the SAR [4,5]: spot mapping, strip mapping, and Doppler beam-sharpening. Strip mapping is the most general one and has received the most attention. Some of its principles

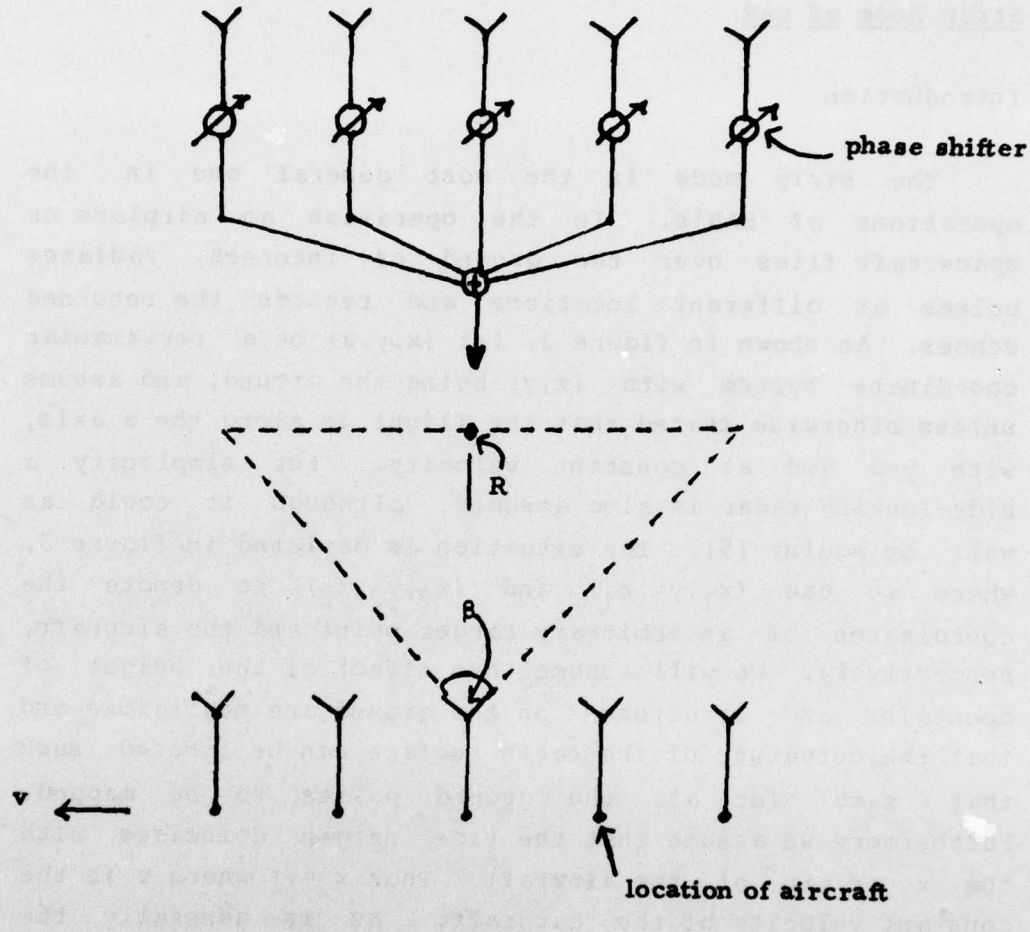


Fig. 2. Analogy between antenna array and synthetic antenna

have been given in the preceding discussion. In the following sections we will formulate the strip mapping mode from a system point of view.

Strip Mode of SAR

Introduction

The strip mode is the most general one in the operations of SAR's. In the operation an airplane or spacecraft flies over the ground of interest, radiates pulses at different locations and records the returned echoes. As shown in figure 3, let (x,y,z) be a rectangular coordinate system with (x,y) being the ground, and assume unless otherwise stated that the flight is along the x axis, with y=0 and at constant velocity. For simplicity a side-looking radar is also assumed, although it could as well be squint [5]. The situation is depicted in figure 3, where we use (x_1,y_1,z_1) and (x_2,y_2,z_2) to denote the coordinates of an arbitrary target point and the aircraft, respectively. We will assume that effect of the height of mountains and structures on the ground are negligible and that the curvature of the earth surface can be ignored such that $z_1=0$ for all the ground points to be mapped. Furthermore we assume that the time origin coincides with the x origin of the aircraft. Thus $x_2=vt$ where v is the constant velocity of the aircraft. As is generally the case, the antenna is assumed to be shared by the transmitter and the receiver. This necessitates the pulsed nature of the signal waveform and inevitably creates blind ranges [1]. Echoes from targets in blind ranges reach the radar while it is transmitting and not receiving and thus are lost. The pulse repetition frequency (PRF) also sets an upper bound to the maximum range without range ambiguities. For simplicity of analysis we assume that the signal pulse train consists

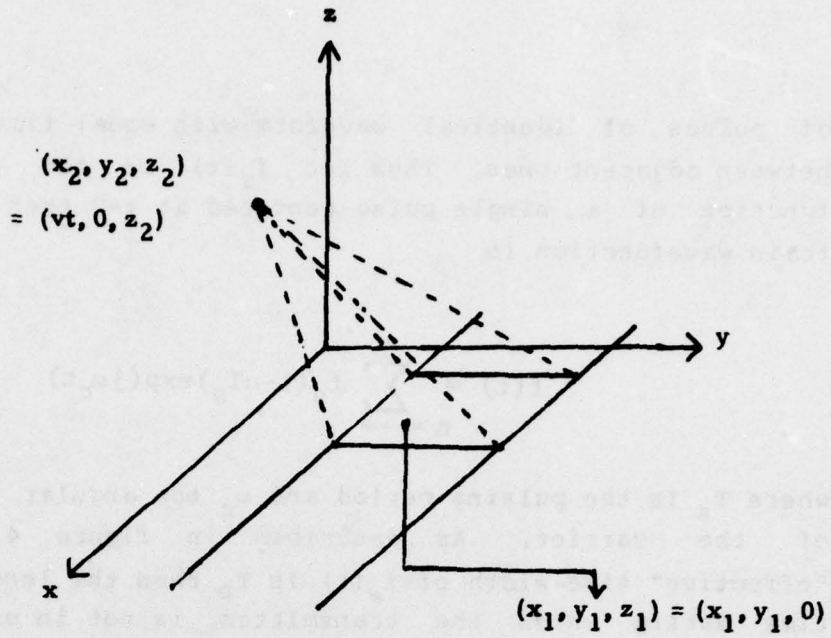


Fig. 3, Flight-path geometry

of pulses of identical waveform with equal time interval between adjacent ones. Thus let $f_s(t)$ be the modulation function of a single pulse centered at t=0 then the pulse train wavefunction is

$$f(t) = \sum_{n=-\infty}^{\infty} f_s(t-nT_s)\exp(j\omega_c t) \qquad (5)$$

where T_s is the pulsing period and ω_c the angular frequency of the carrier. As described in figure 4, if the "effective" time width of $f_s(t)$ is T_p then the length of the time during which the transmitter is not in use between consecutive pulses is $T_s - T_p$ which decides the maximum range deviation without range ambiguities.

In addition to those factors, the depression angle ψ, which is the angle between the horizontal plane and the radiated beam, and the antenna pattern also affect the performance of the system; the geometry is depicted in figures 5 and 6. All those parameters and their interdependence come into the picture of the SAR imaging system to make it extremely complicated to evaluate its capabilities and estimate its performance accurately. However, if we model the point spread function (PSF) of the system in some desirable way by appropriate geometrical considerations and approximations, we will be able to simplify the description of the system, making the evaluation relatively easier and the reconstruction more feasible. Of course, by so doing we also distort the system by an inexact modelling, thus an incomplete or nonoptimal (in some sense) reconstruction is to be expected. It is obvious that the more approximations made, the more degradation will result in the image reconstruction. In this section we will give a hierarchy of the system models

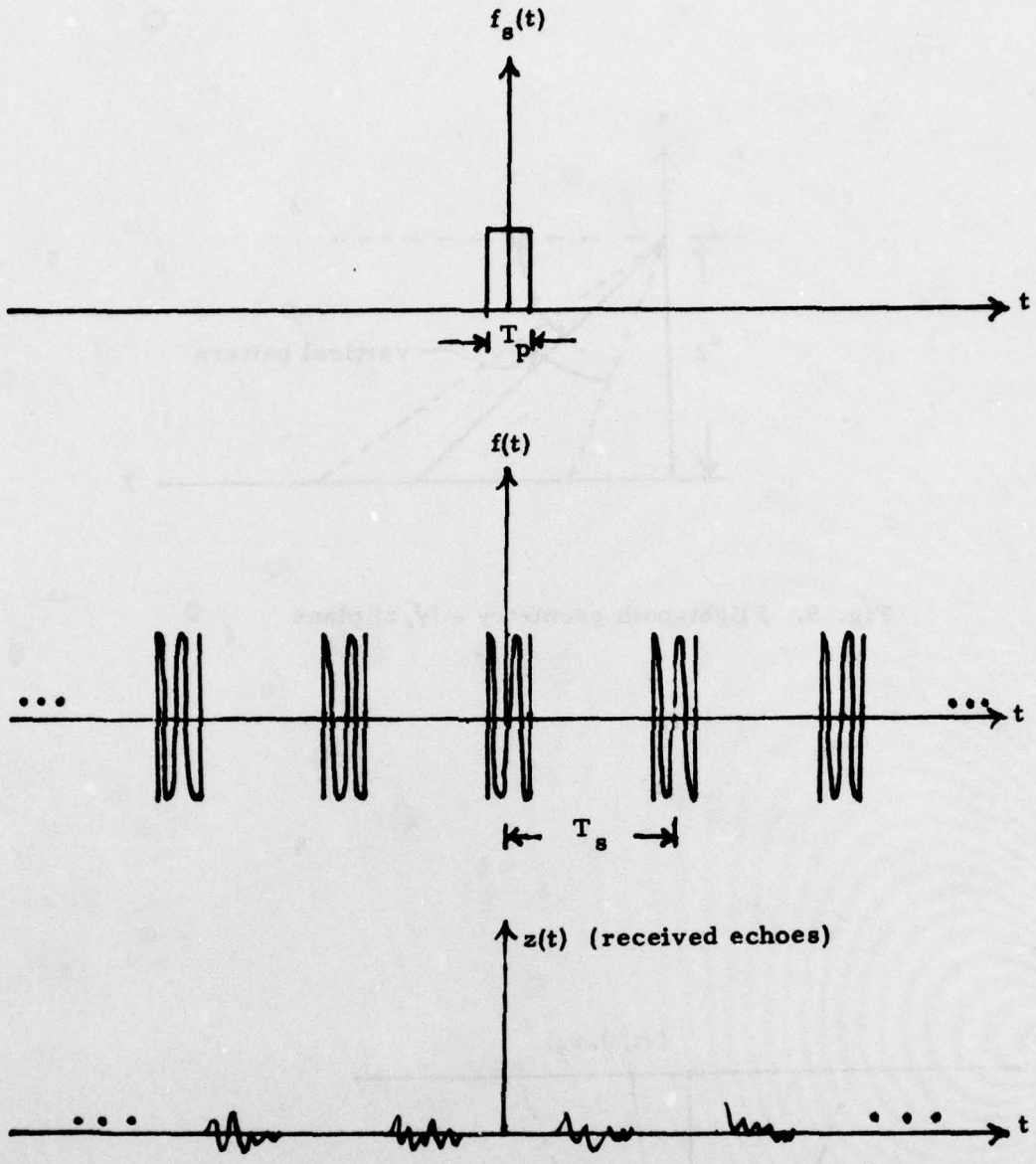


Fig. 4. Waveforms of Signals

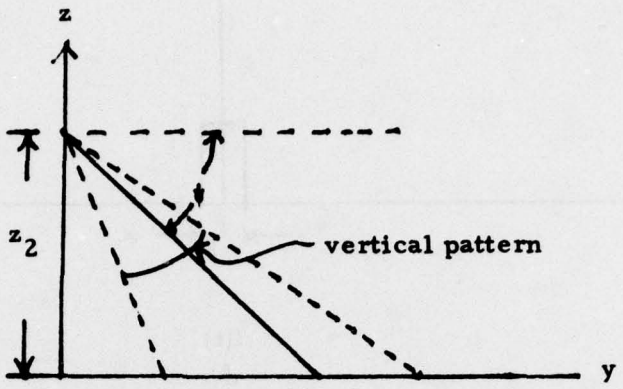


Fig. 5. Flight-path geometry - (y, z) plane

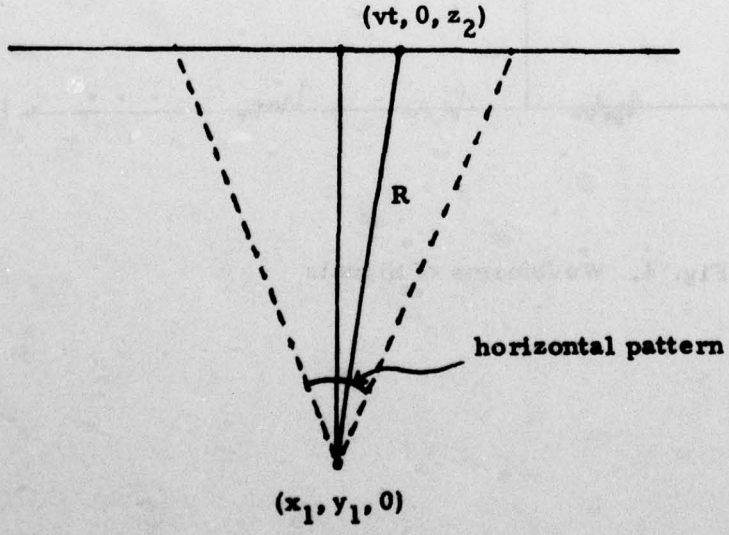


Fig. 6. Flight-path geometry - slant plane

in progressive inaccuracies. We will also tabulate the approximations made and their justifications. We will find that in its simplest form the system is separable space-invariant.

Method of system evaluation

In general, gathering more data provides more information to solve for the unknowns at the expense of increased requirements in storage and complexity in computation. On the other hand, intuition suggests that after some "threshold amount" of data is obtained, the additional observations do not always provide equal amounts of new information. This is due to the inherent "blurring" of the imaging systems and unavoidable observation noise, etc. Thus the concept of degrees of freedom (DOF) has arisen to measure the number of truly independent samples of data one gathers under a particular imaging system [6,7].

In our system evaluation we will adopt the concept of eigenvalues of a correlation matrix or the Gramian matrix [8,9]. For example, in the continuous-discrete case, we equate the number of degrees of freedom with the number of the eigenvalues of the Gramian matrix whose magnitudes are larger than some threshold determined by the noise level of the system. This is equivalent to the singular value analysis of the system. Except for possible permutation the singular value spectrum remains invariant regardless of whether the object and/or the image have gone through orthogonal transformation before and after the imaging system, respectively. To show this, consider the discrete-discrete case for the sake of ease in proof: Let H be the matrix of the linear system and P,Q be orthogonal matrices compatible with the dimensions of the output and input vector sizes, respectively. Then $P^tP=PP^t=I$ and

$Q^tQ=QQ^t=I$ where the superscript t denotes a transposition. The set of eigenvalues of $PHQ(PHQ)^t=PHQQ^tH^tP^t=PHH^tP^t$ is the same set of eigenvalues of $P^tPHH^t=HH^t$ except for additional zeroes due to possible size difference of P and Q [10]. Note that the sets of eigenvalues of PHQ and H differ in general.

Derivation of point spread function

Referring to figure 3, $z_1=0$, $x_2=vt$, $y_2=0$ and z_2 is flight height. Define ground range

$$R_g \triangleq \left[(x_1-x_2)^2+(y_1-y_2)^2\right]^{\frac{1}{2}}$$

$$= \left[(x_1-vt)^2+y_1^2\right]^{\frac{1}{2}} \qquad (6)$$

and slant range

$$R = \left[(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2\right]^{\frac{1}{2}}$$

$$= \left[(x_1-vt)^2+y_1^2+z_2^2\right]^{\frac{1}{2}} \qquad (7)$$

$$= \left[R_g^2+z_2^2\right]^{\frac{1}{2}}$$

The propagation delay associated with a point source at (x_1,y_1,z_1) with range R defined above is $\frac{2R}{C}$ where the factor 2 is because of round trip to and from the target. Let $\rho(x_1,y_1)$ be the reflectivity function of the terrain and $A(x_1,y_1,x_2,z_2)$ be the illuminating pattern of the antenna beam on the terrain. If the antenna pattern remains the same during the flight, it is easily seen that $A(x_1,y_1,x_2,z_2)=A(x_1-x_2,y_1,z_2)$. The received echoes as a function of t are the product of illuminating pattern A,

terrain reflectivity ρ and the delayed signal function f, summed over the ground coordinates (x_1, y_1):

$$z(t) = \int_{-\infty}^{\infty}\int A(x_1-x_2, y_1, z_2)\rho(x_1, y_1) f(t-\frac{2R}{c})\,dx_1 dy_1 \tag{8}$$

Substituting eq.(5) into eq.(8),

$$z(t) = \int_{-\infty}^{\infty}\int A(x_1-x_2, y_1, z_2)\rho(x_1, y_1) \sum_{n=-\infty}^{\infty} f_s(t-\frac{2R}{c}-nT_s)$$
$$\exp\left\{j\omega_c(t-\frac{2R}{c})\right\} dx_1 dy_1 \tag{9}$$

If we interpret eq.(9) as a system with $\rho(x_1, y_1)$ as input and $z(t)$ as output of the system, it is obvious that the system is linear with point spread function

$$h(t; x_1, y_1) = \sum_{n=-\infty}^{\infty} A(x_1-x_2, y_1, z_2) f_s\left(t-nT_s-\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right)$$
$$\exp\left\{j\omega_c\left(t-\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right)\right\} \tag{10}$$

$$= \exp\left\{j\omega_c t\right\} \sum_{n=-\infty}^{\infty} A(x_1-nvT_s, y_1, z_2) f_s\left(t-nT_s-\right.$$
$$\left.\frac{2\left[(x_1-nvT_s)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right)\exp\left\{-j\omega_c\frac{2\left[(x_1-nvT_s)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\}$$

where nvT_s is the x coordinate of the aircraft at which n-th pulse is being radiated. It is assumed that during the transmission and receiving of a single pulse the aircraft is approximately stationary so that x_2 is substituted by nvT_s

in eq.(10). This is valid if (sufficient conditions)

a) $A(x_1-nvT_s, y_1, z_2) \approx A(x_1-nvT_s \pm \Delta x_2, y_1, z_2)$

b) $(|x_1-nvT_s| + \Delta x_2)^2-(x_1-nvT_s)^2 << (y_1+\Delta y_1)^2-y_1^2$ and

c) $2\left[(|x_1-nvT_s|+\Delta x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}-2\left[(x_1-nvT_s)^2+y_1^2+z_2^2\right]^{\frac{1}{2}} << \lambda_c$

where Δx_2 is the maximum distance the aircraft travelled during transmission and receiving of a single pulse. Δy_1 is the range resolution desired. See figure 7. Note that $\Delta x_2 < vT_s$ as assumed earlier.

(a) is easily satisfied by noting that $L_{eff}=\frac{R\lambda}{L}$ is of the order of hundred or thousand meters, and is therefore greatly larger than Δx_2, which is, at most, of the same order of azimuth resolution desired.

(b)

$$(|x_1-nvT_s|+\Delta x_2)^2 - (x_1-nvT_s)^2$$
$$= 2\Delta x_2|x_1-nvT_s|+\Delta x_2^2$$
$$\leq 2\Delta x_2(L_{eff}+\Delta x_2) \approx 2\Delta x_2 Leff$$

while

$$(y_1+\Delta y_1)^2-y_1^2$$
$$= 2\Delta y_1.y + \Delta y_1^2$$
$$\approx 2\Delta y_1.y$$

Since Δx_2 is of the same order as Δy_1 (potentially best azimuth resolution vs. range resolution), (b) will be valid if

$$L_{eff} << y_1 \qquad (11)$$

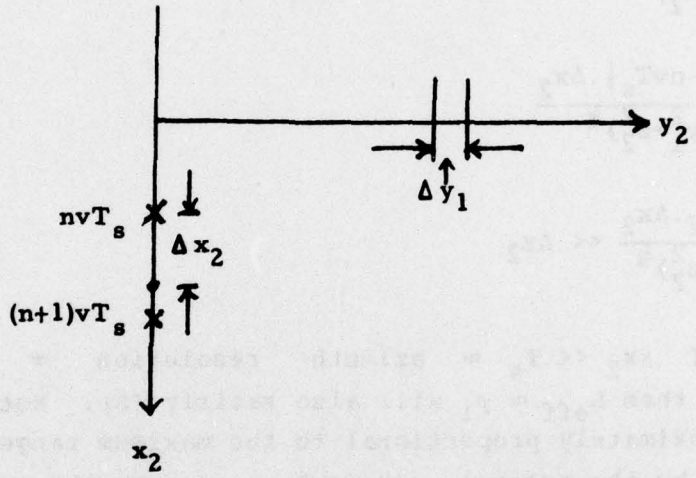


Fig. 7.

(c) $2\left\{\left[(|x_1-nvT_s|+\Delta x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}} - \left[(x_1-nvT_s)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}\right\}$

$$\simeq 2\left\{(y_1^2+z_2^2)^{\frac{1}{2}} + \frac{1}{2}\frac{(|x_1-nvT_s|+\Delta x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}} - (y_1^2+z_2^2)^{\frac{1}{2}} - \frac{1}{2}\frac{(x_1-nvT_s)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}\right\}$$

$$= \frac{2}{(y_1^2+z_2^2)^{\frac{1}{2}}}\left(|x_1-nvT_s|.\Delta x_2+\Delta x_2^2\right)$$

$$\approx \frac{2|x_1-nvT_s|.\Delta x_2}{(y_1^2+z_2^2)^{\frac{1}{2}}}$$

$$\leq \frac{2L_{eff}.\Delta x_2}{(y_1^2+z_2^2)^{\frac{1}{2}}} << \Delta x_2$$

However, if $\Delta x_2 << T_s \approx$ azimuth resolution $\approx$ range resolution, then $L_{eff} \approx y_1$ will also satisfy (b). Note that Δx_2 is approximately proportional to the maximum range of y_1 illuminated by the antenna where $L_{eff} << y$ by the validity of (b). Thus if we let Δx_2 be of the same order as λ_c, (c) will be satisfied. In fact, the achievable azimuth resolution is of the same order as λ_c [13].

Because the sinusoidal phase term $\exp\{j\omega_c t\}$ in eq.(10) does not carry any information on $\rho(x_1,y_1)$, it can be shifted to any lower frequency ω_o desired. In optical processing upon SAR data, the "offset" frequency $\omega_o \neq 0$ is to separate reconstructed twin images from each other and from other useless images [11,12].

Thus,

$$h(t;x_1,y_1) = \exp\{j\omega_0 t\}\sum A(x_1-nvT_s,y_1,z_2)$$

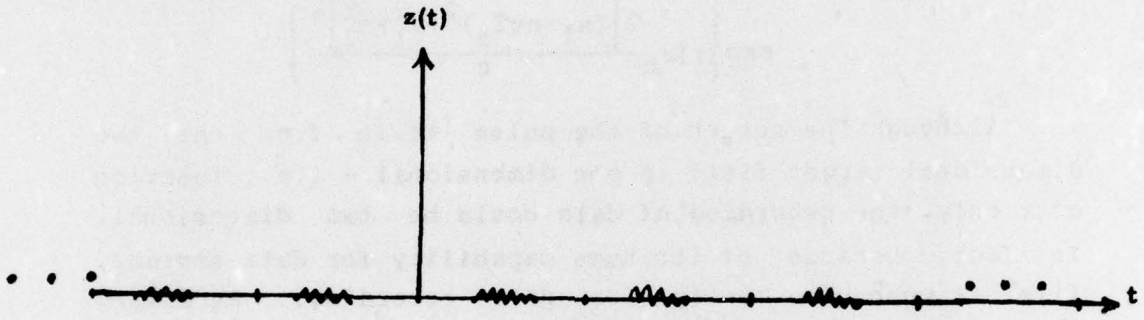
$$f_s\left(t-nT_s-\frac{2\left[(x_1-nvT_s)^2+y_1^2+z_1^2\right]^{\frac{1}{2}}}{c}\right) \qquad (12)$$

$$\exp\left\{-j\omega_c\frac{2\left[(x_1-nvT_s)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\}$$

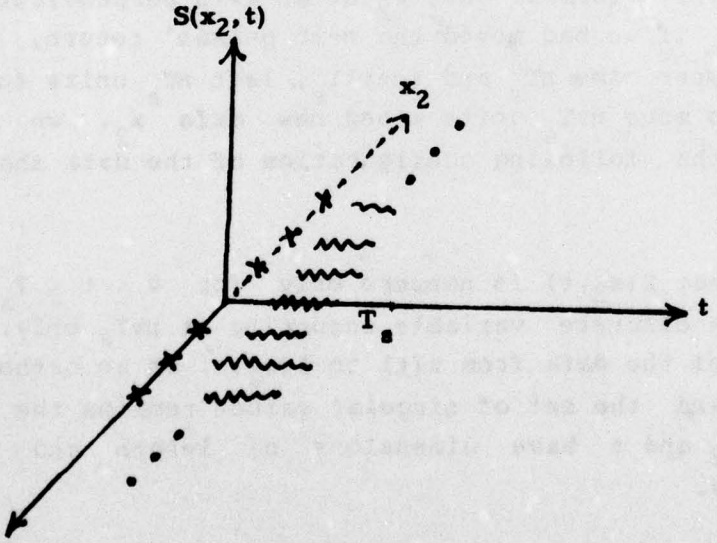
Although the return of the pulse train from the two dimensional target field is one dimensional - i.e., function of t only, the recording of data could be two dimensional. In fact, because of its huge capability for data storage, film has been used mostly for data recording. Recalling that the signal returns from different pulses do not overlap, we could arrange them in a more compact two-dimensional format: Let x_2 be an axis perpendicular to the t axis. If we had moved the n-th pulses' return, which occurs between time nT_s and $(n+1)T_s$, left nT_s units towards t=0 and also move nvT_s units along new axis x_2, we would have had the following configuration of the data shown in figure 8.

Note that $S(x_2,t)$ is nonzero only for $0 \leq t \leq T_s$ and that x_2 is a discrete variable occurring at nvT_s only. The reordering of the data from z(t) to $S(x_2,t)$ is an orthogonal transform and the set of singular values remains the same. Note that x_2 and t have dimensions of length and time, respectively.

The PSF expressed in (x_2,t) variables now becomes



(a) received signal



(b) reordering of (a)

Fig. 8. Reordering of received signal

$$h(x_2,t;x_1,y_1) = \sum_{n=-\infty}^{\infty} \delta(x_2-nvT_s)\exp\left[j\omega_0(t+nT_s)\right]A(x_1-x_2,y_1,z_2)$$
$$f_x\left(t-\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right)\exp\left\{-j\omega_c\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\} \quad (13)$$

where $0 \le t \le T_s$

Multiplying $\exp(-j\omega_o nT_s)=\exp(-j\omega_o\frac{x_2}{v})$ and $\exp(-j\omega_o t)$ respectively to $h(x_2,t;x_1,y_1)$ yields range offset and azimuth offset cases, respectively [14]. Thus

$h(x_2,t;x_1,y_1)$ =

$$\exp(j\omega_0 t)\sum_{n=-\infty}^{\infty}\delta(x_2-nvT_s)A(x_1-x_2,y_1,z_2)$$
$$f_s\left(t-\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right)\exp\left\{-j\omega_c\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\}$$

range offset case (14)

$h(x_2,t;x_1,y_1)$ =

$$\exp(j\frac{\omega_0}{v}x_2)\sum_{n=-\infty}^{\infty}\delta(x_2-nvT_s)A(x_1-x_2,y_1,z_2)f_s\left(t-\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right)$$

$$\exp\left\{-j\omega_c\frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\}, \text{ azimuth offset case} \quad (15)$$

In the following analysis we shall assume that the range offset case is used.

Simplifications of PSF

In this section we continue simplifying the PSF (eq.(14)) of the SAR imaging system. In the meantime a

hierarchy of models of PSF's with decreasing complexity will be derived along with their associated assumptions and approximations.

We start by noting the strong relations between variables x_1 and x_2, y_1 and t, respectively: the argument of f_s in eq.(14), $t-\left[2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}\right]/c$, were it not for the factor $(x_1-x_2)^2$, would have yielded a propagation delay which connects y_1 with t only to provide the range information, and is independent of azimuth modulation. This is the only way by which $h(x_2,t;x_1,y_1)$ is not separate in azimuth and range in eq.(14). If this fact can be ignored, e.g., if the propagation delay induced by the variation in (x_1-x_2) is much smaller than the range resolution interested, then the PSF can be considered separate in azimuth (x_1-x_2) and range $(y-t)$:

$$h(x_2,t;x_1,y_1) = \exp(j\omega_o t) \sum_{n=-\infty}^{\infty} \delta(x_2-nvT_s)A(x_1-x_2,y_1,z_2)$$

$$f_s\left(t-\frac{2(y_1^2+z_2^2)}{c}\right) \exp\left\{-j\omega_c \frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\} \tag{16}$$

Physically speaking, this means that the range resolution cells do not move to overlap as the flight goes on. The situation is depicted in figure 9, where the range γ of a target point is plotted as a function of the position x of the aircraft. Various ways have been proposed to alleviate the problem of range-azimuth coupling [15,16,17]. If ρ_r is the range resolution pursued and β is the effective beam width then this range curvature effect will be ignored if $\left|\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}-(y_1^2+z_2^2)^{\frac{1}{2}}\right| \leq \rho_r$ for all valid x_1-x_2 and

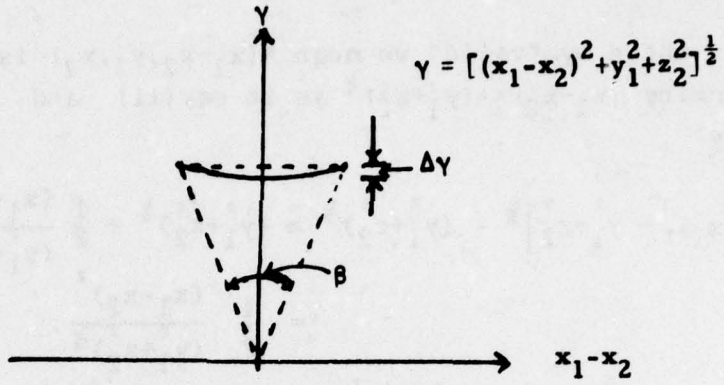


Fig. 9. Range variation of a point target. Δγ is less than range resolution.

y_1, where by "valid" we mean $A(x_1-x_2,y_1,z_2)$ is significant. Assuming $(x_1-x_2)<<(y_1^2+z_2^2)^{\frac{1}{2}}$ as in eq.(11) and β small, we have

$$\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}} - (y_1^2+z_2^2)^{\frac{1}{2}} \approx (y_1^2+z_2^2)^{\frac{1}{2}} + \frac{1}{2}\frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}} - (y_1^2+z_2^2)^{\frac{1}{2}}$$

$$= \frac{1}{2}\frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}$$

$$\leq \frac{1}{8}\frac{L_{eff}^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}$$

$$\approx \frac{\beta}{8} L_{eff}$$

We require that

$$\rho_r > \frac{\beta}{8} L_{eff} \qquad (17)$$

i.e., PSF can be approximated by eq.(16), which is separate in azimuth and range, if eq.(17) holds.

To see that the system with kernel eq.(16) is separate in azimuth and range, we rewrite eq.(16) as

$$h(x_2,t;x_1,y_1) = \exp(j\omega_o t) f_s\left(t-\frac{2(y_1^2+z_2^2)}{c}\right)$$

$$\sum_{n=-\infty}^{\infty} \delta(x_2-nvT_s) A(x_1-x_2,y_1,z_2) \exp\left\{-j\omega_c \frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\}$$

$$= h_t(t;y_1) h_z(x_2;x_1,y_1) \qquad (18)$$

where $h_t(t;y_1) \triangleq \exp(j\omega_o t) f_s(t-\frac{2(y_1^2+z_2^2)^{\frac{1}{2}}}{c})$ is the impulse response of t due to a unit point source at y_1, which is independent of azimuth dimension; and

$$h_z(x_2;x_1,y_1) \triangleq \sum_{n=-\infty}^{\infty} \delta(x_2-nvT_s)\, A(x_1-x_2,y_1,z_2) \exp\left\{-j\omega_c \frac{2\left[(x_1-x_2)^2+y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\}$$

is the impulse response of x_2 due to a unit point source at (x_1,y_1). Note that $h_z(x_2;x_1,y_1)=h_z(x_1-x_2;y_1)$ is space invariant in x_1 and x_2, but varies its form as y_1 changes. If we use eq.(18) as the kernel of SAR, then the input-output relation will be

$$z(x_2,t) = \iint_{-\infty}^{\infty} h_t(t;y_1) h_z(x_1-x_2;y_1) \rho(x_1,y_1)\, dx_1\, dy_1$$

$$= \int_{-\infty}^{\infty} h_t(t;y_1) \left[\int_{-\infty}^{\infty} h_z(x_1-x_2;y_1)\ \rho(x_1,y_1)\ dx_1 \right] dy_1$$

Thus the imaging system of SAR is to transform $\rho(x_1,y_1)$ into $z(x_2,t)$ in a sequential order: azimuth transformation followed by range transformation. However, because of the dependence of h_z upon y_1, h_t and h_z are not separable and thus in general their order cannot be interchanged in modelling the system. Accordingly, the reconstruction of the ground reflectivity function $\rho(x_1,y_1)$ from its image $z(x_2,t)$ has to follow the reversed order.

From eq.(11), $|x_1-x_2| << y_1^2+z_2^2$ hence

$$\exp\left\{-j\omega_c \frac{2\left[(x_1-x_2)^2 + y_1^2+z_2^2\right]^{\frac{1}{2}}}{c}\right\} \approx \exp\left\{-j\omega_c \frac{2(y_1^2+z_2^2)^{\frac{1}{2}} + \frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}}{c}\right\}$$

$$= \exp\left\{-j\omega_c \frac{2(y_1^2+z_2^2)^{\frac{1}{2}}}{c}\right\} \exp\left\{-j\omega_c \frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}\right\}$$

and thus $h_z(x_2;x_1,y_1)$ can be approximated by

$$h_z(x_2;x_1,y_1) \approx \exp\left\{-j\omega_c \frac{2(y_1^2+z_2^2)^{\frac{1}{2}}}{c}\right\} \sum_{n=-\infty}^{\infty} \delta(x_2-nvT_s) A(x_1-x_2,y_1,z_2)$$

$$\exp\left\{-j\frac{\omega_c}{c} \frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}\right\}$$

Because the phase term $\exp\left\{-j\omega_c \frac{2(y_1^2+z_2^2)^{\frac{1}{2}}}{c}\right\}$ is independent of azimuth its effect can be taken out and absorbed in the reflectivity function $\rho(x,y)$. It is a nonlinear operation on $\rho(x,y)$, however, its effect on the singular values of SAR will be ignored. In terms of block diagrams, we have figure 10 where

$$\begin{cases} h_t(t;y_1) = \exp(j\omega_o t) f_s\left(t - \frac{2(y_1^2+z_2^2)^{\frac{1}{2}}}{c}\right) \\ h_z'(x_1-x_2,y_1) \triangleq \sum_{n=-\infty}^{\infty} \delta(x_2-nvT_s) A(x_1-x_2,y_1,z_2) \\ \qquad \exp\left\{-j\frac{\omega_c}{c} \frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}\right\} \end{cases} \qquad (19)$$

Equation (19) is the form assumed for most of the processing

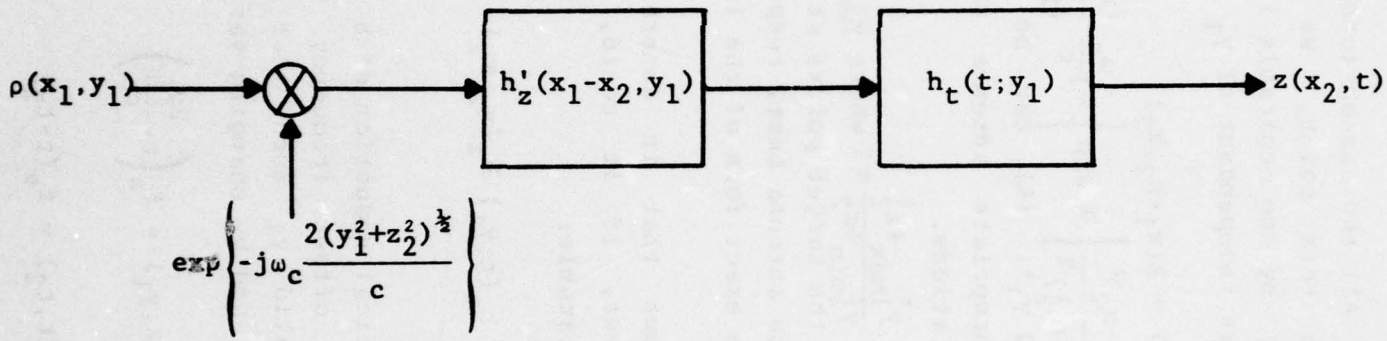


Fig. 10. Block Diagram of Equation (17).

of the SAR data. All the assumptions required are easily justified to reach this point. We shall proceed to approximate eq.(19) by one separable in azimuth and range, i.e. such that h_z is independent of y_1 in eq.(19). This will be true if

(A) $A(x_1-x_2, y_1, z_2) = A(x_1-x_2, z_2)$

(B) $$\exp\left\{-j\frac{\omega_c}{c}\frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}\right\} \cong \exp\left\{-j\frac{\omega_c}{c}\frac{(x_1-x_2)^2}{(y_1^2+z_2^2)^{\frac{1}{2}}}\right\}$$

for all valid y_1 and y_1'. (A) can be made approximately true by an appropriate antenna pattern design with geometrical considerations.

(B) will be true if $\frac{y_{1max}^2+z_2^2}{y_{1min}^2+z_2^2} \approx 1$ where y_{1min} and y_{1max} mean the y_1 coordinates of the target points at maximum and minimum ranges covered by the antenna beam, respectively. We shall not elaborate on the exact form of the inequality.

It is pointed out that in general (B) cannot hold practically. However, if it could, then the system of eq.(19) would be separable:

$$h(x_2,t;x_1,y_1) = h_t(t;y_1)\, h_z(x_2-x_1) \tag{20}$$

Further theoretical reduction of h is still possible: if we ignore the offset frequency term and assume that $y_1 >> z_2$ for all valid y_1 such that $2(y_1^2+z_2^2)^{\frac{1}{2}}/c$ can be approximated by $\frac{2y_1}{c}$ and by changing variable $t_2 \triangleq \frac{2y_1}{c}$

then $$h_t(t;y_1) = f_s\left(t-\frac{2y_1}{c}\right)$$

or $$h_t(t;t_2) = f_s(t-t_2)$$

$$h(x_2,t;x_1,y_1) = h_1(x_2,t;x_1,t_1)$$

and
$$= h_t(t-t_2)h_z(x_2-x_1) \qquad (21)$$

is a separable space invariant PSF (SSIPSF).

A summary of the properties of the PSF's under various assumptions are listed in Table 1.

References

1. M.I. Skolnik, Radar Handbook, McGraw-Hill, 1970.

2. A.W. Rihaczek, Principles of High-Resolution Radar, McGraw-Hill, 1969.

3. S.L. Johnston, "Matched-Filter Response to a Linear FM Signal Transmitted Through a Phased Array," IEEE Transactions Aeronautical and Electrical Systems, Vol. AE-12, No.1, January 1976, pp. 73-77.

4. J.C. Kirk, Jr., "Motion Compensation for SAR," IEEE Transactions Aeronautical and Electrical Systems, Vol 11, No. 3, May 1975.

5. J.C. Kirk, Jr., "A Discussion of Digital Processing in SAR," IEEE Transactions Aeronautical and Electrical Systems, Vol 11, No. 3, May 1975.

6. F. Gori and G. Guattari, "Shannon Number and Degrees of Freedom of an Image," Optics Communication, Vol. 7, No. 2, 1973, pp. 163-165.

7. S. Twomey, "Information Content in Remote Sensing," Applied Optics, Vol. 13, No. 4, 1976, pp. 942-945.

8. S. Hou and H.C. Andrews, "Fundamental Limits and

TABLE 1

EQUATION NUMBER	FORM OF PSF	ASSUMPTIONS
(13) (14) (15)	NSVPSF[2]	1. separate pulse return do not overlap 2. aircraft stationary in one T_s time or $L_{eff} << y_1$
(16) (18)	NSVPSF separate in azimuth and range (order counts) invariant in azimuth	1. $\rho_r > \frac{\beta}{8} L_{eff}$
(20)	SSVPSF[3] invariant in azimuth	1. $A(x_1-x_2, y_1, z_2) \approx A(x_1-x_2, z_2)$ 2. $\frac{y_{1\,max}^2+z_2^2}{y_{1\,min}^2+z_2^2} \approx 1$
(21)	SSIPSF[4]	1. range offset frequency $\omega_o=0$ 2. $\frac{\left(y_1^2+z_2^2\right)^{\frac{1}{2}}}{y_1} \approx 1$

1. All assumptions in upper blocks remain true in lower ones
2. Nonseparable space variant PSF
3. Separable space variant PSF
4. Separable space invariant PSF

Degrees of Freedom of Imaging Systems," Proceedings of SPIE, Vol. 74, 1976.

9. D.G. McCaughey and H.C. Andrews, "Degrees of Freedom for Projection Imaging," IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-25, No. 1, February 1977.

10. M. Marcus and H. Ming, A Survey of Matrix Theory and Matrix Inequalities, Allyn and Bacon, 1964, pp. 24.

11. J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, 1968.

12. E.N. Leith, "Quasi-Holographic Techniques in the Microwave Region," Proceedings of the IEEE, Vol. 59, No. 9, September 1971.

13. R.O. Harger, Synthetic Aperture Radar Systems, Academic Press, 1970, pp. 37-42.

14. R.O. Harger, Synthetic Aperture Radar Systems, Academic Press, 1970, pp. 64-65.

15. W.M. Brown and R.J. Fredricks, "Range-Doppler Imaging with Motion Through Resolution Cells," IEEE Transactions on Aeronautical and Electrical Systems, Vol. AES-5, No. 1, January 1969, pp. 98-102.

16. E.N. Leith, "Range-Azimuth-Coupling Aberrations in Pulse-Scanned Imaging Systems," JOSA, Vol. 63, No. 2, February 1973, pp. 119-126.

17. E.N. Leith, "Complex Spatial Filters for Image Deconvolution," Proceedings of IEEE, Vol. 65, No. 1,

January 1977, pp. 18-28.

3.2 DOF of Tomographic Projections

Chung-Ching Chen

Introduction

Recently the reconstruction of an image from its projections has gained more and more interest in many fields such as electron microscopy [1], radio astronomy [2], medical tomography [3], etc. If the projection systems are ideal, e.g. if the projection beam is infinitely narrow, it is possible to reconstruct the image perfectly by direct [2] or indirect [1] manipulations of the projections. However, in reality, the projections are blurred and usually only a finite number of projections are available. Although in general we could obtain more information by more projections, it seems that because of the inherent blurring of the system, there is an "optimal" amount of projections beyond which the cost of increased data storage and manipulation is not worth the rapidly decreasing information return. The concept of degrees of freedom (DOF) has thus arisen to quantify the number of truly independent data gathered under a certain geometry [4,5]. By examining the eigenvalue spectrum of the correlation matrix (or the Gramian) associated with the linear imaging system, one is able to evaluate the system performance in the sense of DOF. Roughly speaking, in the eigenspace of the system only those components (or singular values) whose magnitudes are larger than the noise level provide information about the input with high enough confidence. Discarding those singular values of relatively small magnitude also guarantees an optimum bandwdith reduction or data compression in the

minimum mean square sense, among other advantages of eigenspace manipulations. As is well known, the complexity of the diagonalization process of an arbitrary N x N matrix increases very rapidly with N (of the order N^3) [6], in addition to the fact that only approximate values can be obtained because of the iterative nature of the algorithms. This limits the practicability of eigenspace analysis even for moderately large N unless the system (and hence its matrix) is well-structured and algorithms to fully utilize its structure to speed up the computations exist.

In this paper we discuss the principles of operation of a tomographic projection system, formulate its Gramian matrix and seek fast ways to find its eigenvalues for certain sampling functions.

Formulations

Referring to figure 1 let the rectangular coordinate system (ξ,η) be fixed on the object whose size has been normalized to be contained within a unit circle and (x,y) be another rectangular coordinate system with the y-axis parallel to the lines of projection for a given source orientation angle θ defined as the angle between axes x and ξ as shown in figure 1. The relations between the coordinates (x,y) and (ξ,η) are

$$\begin{cases} x = \xi\cos\theta+\eta\sin\theta \\ y = -\xi\sin\theta+\eta\cos\theta \end{cases} \quad \text{or} \quad \begin{cases} \xi = x\cos\theta-y\sin\theta \\ \eta = x\sin\theta+y\cos\theta \end{cases} \qquad (1)$$

For each θ M_r samples $\gamma_1,\gamma_2,\ldots\gamma_{M_r}$ are taken corresponding to M_r non-ideal line projections perpendicular to the x-axis with $0 \le \gamma_k \le 1$. The number of samples in θ is M_θ, with θ_i, $i=1,2,\ldots,M_\theta$ ranging from 0 to 2π. Thus the total number of samples is $M_r M_\theta$ and the system is a continuous-discrete one [7] because the input variables (ξ,η) are continuous and the

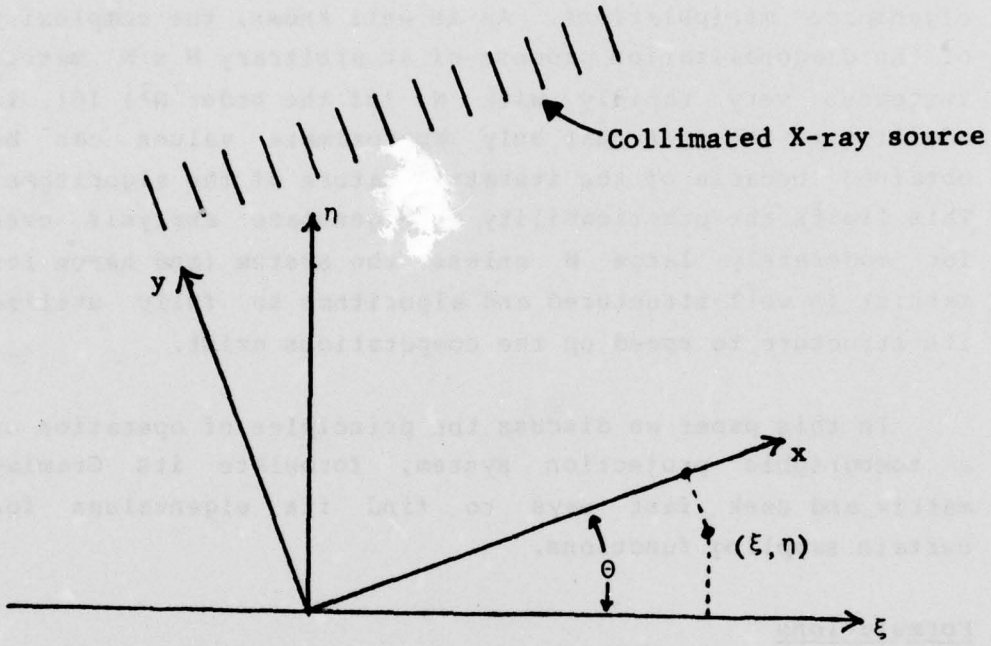


Figure 1. Coordinate systems

output variables (γ,θ) are discrete. Let $f(\xi,\eta)$ be the object density function, then the projection $z(\gamma_k,\theta_i)$ is related to $f(\xi,\eta)$ as a two variable Fredholm integral

$$z(\gamma_k,\theta_i) = \iint_R f(\xi,\eta)\ h(\theta_i,\gamma_k;\xi,\eta)\ d\xi d\eta \qquad (2)$$

where $h(\theta_i,\gamma_k;\xi,\eta)$ is the point spread function (PSF) designating the contribution of a unit magnitude point spread source at (ξ,η) to the output variables (θ_i,γ_k) and R denotes the unit circle. Note $h(\theta_i,\gamma_k;\xi,\eta)$ can be rewritten as $h(\theta_i,\gamma_k;x,y)$ through the transformation of eq.(1). We assume that the projections are aimed at γ_k, $k=1,\ldots,M_r$, and assume the same waveform $\omega(x)$ in the x direction. Then $h(\theta_i,\gamma_k;x,y) = \omega(x-\gamma_k)$ is a shifted version of $\omega(x)$. Thus eq.(2) becomes

$$z(\gamma_k,\theta_i) = \iint_R f(\xi,\eta)\ \omega(\xi\cos\theta_i+\eta\sin\theta_i-\gamma_k)\ d\xi d\eta \qquad (3)$$

where we have used the relation of eq.(1) with the addition of subscript i. Note that x_i and y_i are continuous functions of ξ and η, although the functions depend on discrete variables θ_i because of the discrete sampling over the whole circle.

In terms of matrix notation, eq.(3) assumes the form

$$[Z] = \int_R\int \left[H(\xi,\eta)\right] f(\xi,\eta)\,d\xi d\eta \qquad (4)$$

where $[Z]$ and $[H(\xi,\eta)]$ are matrices of size $M_r \times M_\theta$. For our purpose it is more useful to put eq.(4) in a vector

$$\underline{z} = \int_R\int \underline{h}(\xi,\eta)\,f(\xi,\eta)\,d\xi d\eta \qquad (5)$$

where $\underline{z}$ and $\underline{h}$ are $M_r M_\theta \times 1$ matrices. Degrees of freedom of

the imaging system can be derived from the eigenvalues of the following correlation matrix

$$[\Gamma] = \iint_R \underline{h}(\xi,\eta)\underline{h}^t(\xi,\eta)\,d\xi d\eta \tag{6}$$

If the image vector $\underline{z}$ is formed in a lexicographic order then the matrix will be

$$[\Gamma] = \begin{bmatrix} [\Gamma]^{(1,1)}, & [\Gamma]^{(1,2)} \cdots\cdots [\Gamma]^{(1,M_\theta)} \\ \vdots & \\ [\Gamma]^{(M_\theta,1)}, [\Gamma]^{(M_\theta,2)} \cdots\cdots [\Gamma]^{(M_\theta,M_\theta)} \end{bmatrix} \tag{7}$$

where $[\Gamma]^{i,m}$ is the correlation matrix between i-th an m-th samples in θ. The (k,1)-th entry of $[\Gamma]^{(i,m)}$ is

$$\gamma_{(k,l)}^{(i,m)} = \iint_R \omega(x_i-\gamma_k)\,\omega^*(x_m-\gamma_l)\,d\xi d\eta \tag{8}$$

where x_i and x_m are defined in eq.(1) with $\theta = \theta_i$, $\theta = \theta_m$, respectively and $\omega(x_i-\gamma_k)$ is the pulse function along the x_i axis shifted by γ_k units. A similar geometric interpretation for $\omega(x_m-\gamma_l)$ exists. The situation is depicted in figure 2 where for pictorial clarity, finite width of $\omega(t)$ is assumed. The integration of eq.(6) is over the overlapped region of $\omega(x_i-\gamma_k)$ and $\omega^*(x_m-\gamma_l)$ which lies within the unit circle. Unfortunately there is no closed form for the expression of eq.(8) for arbitrary function $\omega(x)$. A major difficulty is that the integrand is limited to the unit circle instead of the whole plane. Appropriate assumptions on the shape of $\omega(x)$, as are usually practical,

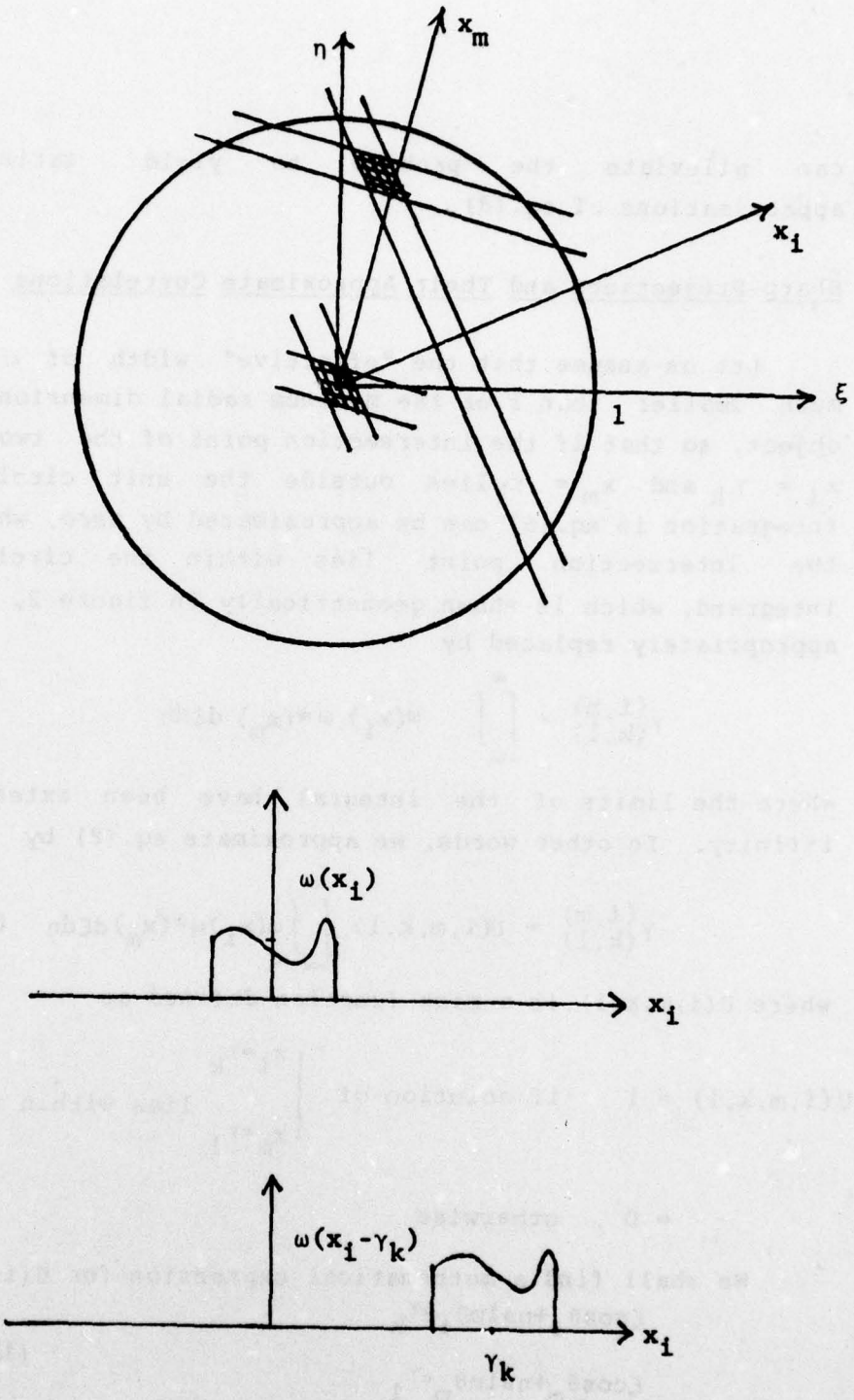


Figure 2. Geometric interpretation of $\gamma_{(k,l)}^{(i,m)}$

can alleviate the problem to yield satisfactory approximations of eq.(8).

Sharp Projections and Their Approximate Correlations

Let us assume that the "effective" width of $\omega(x)$ is much smaller than 1 or the maximum radial dimension of the object, so that if the intersection point of the two lines $x_i = \gamma_k$ and $x_m = \gamma_l$ lies outside the unit circle, the integration in eq.(6) can be approximated by zero, while if the intersection point lies within the circle, the integrand, which is shown geometrically in figure 2, can be appropriately replaced by

$$\gamma_{(k,l)}^{(i,m)} = \int_{-\infty}^{\infty}\int \omega(x_i)\,\omega^*(x_m)\,d\xi d\eta \qquad (9)$$

where the limits of the integral have been extended to infinity. In other words, we approximate eq.(8) by

$$\gamma_{(k,l)}^{(i,m)} = U(i,m,k,l) \int_{-\infty}^{\infty}\int \omega(x_i)\omega^*(x_m)\,d\xi d\eta \qquad (10)$$

where U(i,m,k,l) is a mask function defined as

$$U(i,m,k,l) = 1 \quad \text{if solution of} \quad \begin{cases} x_i=\gamma_k \\ x_m=\gamma_l \end{cases} \text{lies within unit circle}$$

$$= 0\,, \quad \text{otherwise}$$

We shall find a mathematical expression for U(i,m,k,l):

$$\begin{matrix} \xi\cos\theta_i+\eta\sin\theta_i=\gamma_k \\ \xi\cos\theta_m+\eta\sin\theta_m=\gamma_l \end{matrix} \qquad (11)$$

The solution (ξ_o,η_o) of eq.(11) is

$$\xi_0 = \frac{\gamma_k \sin\theta_m - \gamma_1 \sin\theta_i}{\sin(\theta_m - \theta_i)}$$

$$\eta_0 = \frac{\gamma_1 \cos\theta_i - \gamma_k \cos\theta_m}{\sin(\theta_m - \theta_i)}$$

and (ξ_o, η_o) lies in a unit circle if and only if

$$\xi_0^2 + \eta_0^2 \leq 1 \quad \text{iff}$$

$$\left[\gamma_k^2 + \gamma_1^2 - 2\gamma_k\gamma_1 \cos(\theta_m - \theta_i)\right]^{\frac{1}{2}} \leq |\sin(\theta_m - \theta_i)| \tag{12}$$

Figure 3 shows a geometric interpretation of the above inequality, which is an even function of $\theta_m - \theta_i$. Thus

$$U(i,m,k,1) = 1, \quad \text{if } \left[\gamma_k^2 + \gamma_1^2 - 2\gamma_k\gamma_1 \cos(\theta_m - \theta_i)\right]^{\frac{1}{2}} \leq |\sin(\theta_m - \theta_i)|$$

$$= 0, \quad \text{otherwise} \tag{13}$$

Diagonalization of the Gramian for an Infinite Domain w(x) Function

Let us assume that $\omega(t)$ has a normalized Gaussian shape with variance σ^2, i.e.

$$\omega(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-x^2}{2\sigma^2}\right\} \tag{14}$$

Let u be the bisector of the angle formed by x_i and x_m and v be the axis 90^o counterclockwise from u, as in figure 4. Since the angle between u and x_i is $\frac{\theta_m - \theta_i}{2}$, we have the following transformation relations:

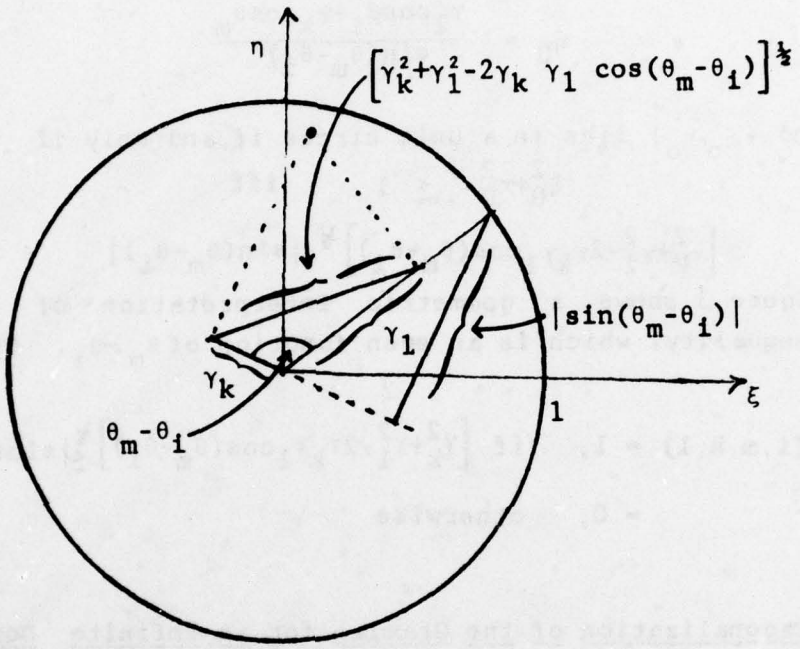


Figure 3. Interpretation of U(i,m,k,1)

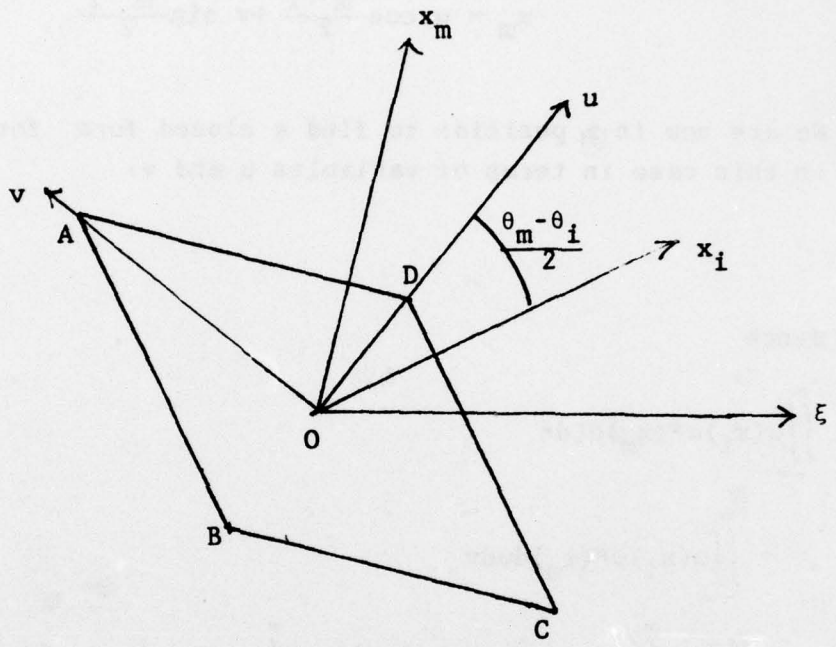


Figure 4. Definitions of u and v

$$x_i = u\cos\frac{\theta_m-\theta_i}{2} - v\sin\frac{\theta_m-\theta_i}{2}$$

$$x_m = u\cos\frac{\theta_m-\theta_i}{2} + v\sin\frac{\theta_m-\theta_i}{2} \qquad (15)$$

We are now in a position to find a closed form for eq.(9), in this case in terms of variables u and v:

Hence

$$\iint_{-\infty}^{\infty}\omega(x_i)\omega^*(x_m)\,d\xi\,d\eta$$

$$= \iint_{-\infty}^{\infty}\omega(x_i)\omega^*(x_m)\,du\,dv$$

$$= \frac{1}{2\pi\sigma^2}\int_{\infty}^{\infty}\exp\{-\frac{u^2}{\sigma^2}\cos^2\frac{\theta_m-\theta_i}{2}\}du\int_{\infty}^{\infty}\exp\{-\frac{v^2}{\sigma^2}\sin^2\frac{\theta_m-\theta_i}{2}\}\,dv$$

$$= \csc|\theta_m-\theta_i| \qquad (16)$$

Because eq.(16) blows up at $\theta_m = \theta_i$, a more accurate approximation for this case is required. Assuming negligible overlap between adjacent pulses with the same orientation as before, we consider the case $\gamma_k = \gamma_l$, $x_i = x_m$ only as in figure 5. In this case we modify the limits of the integral of variable v to get

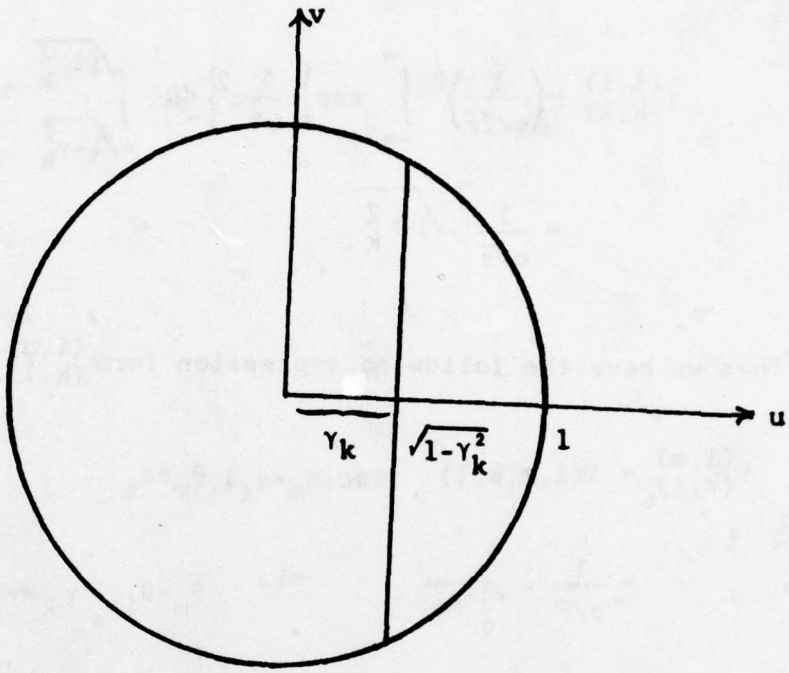


Figure 5. The case $\theta_i = \theta_m$, $t_k = t_l$

$$\gamma_{(k,k)}^{(i,i)} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^2 \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{\sigma^2}u^2\right\} du \int_{-\sqrt{1-\gamma_k^2}}^{\sqrt{1-\gamma_k^2}} 1 \, dv$$

$$= \frac{1}{\sigma\sqrt{\pi}} \cdot \sqrt{1-\gamma_k^2} \qquad (17)$$

Thus we have the following expression for $\gamma_{(k,l)}^{(i,m)}$

$$\gamma_{(k,l)}^{(i,m)} = U(i,m,k,l) \quad \csc|\theta_m-\theta_i|, \theta_m \neq \theta_i$$

$$= \frac{1}{\sigma\sqrt{\pi}} \cdot \sqrt{1-\gamma_k^2} \qquad \theta_m=\theta_i, \ \gamma_k=\gamma_l$$

$$= 0 \qquad \theta_m=\theta_i, \ \gamma_k \neq \gamma_l \qquad (18)$$

where U(i,m,k,l) is given by eq.(13). Note that eq.(18) implies that the adjacent θ's are not so close as to make $\frac{1}{2}U(i,m,k,l)\csc(\theta_m-\theta_i)$ larger than either $\frac{1}{\sigma\sqrt{\pi}}\sqrt{1-\gamma_k^2}$ or $\frac{1}{\sigma\sqrt{\pi}}\sqrt{1-\gamma_l^2}$. If this should happen in eq.(18), an appropriate thresholding function should be incorporated in the expression. This is the result of the Schwartz inequality

$$\iint_R \omega(x_i-\gamma_k)\omega(x_m-\gamma_l)\,d\xi\,d\eta$$

$$\leq \left[\iint_R \omega^2(x_i-\gamma_k)\,d\xi\,d\eta \iint_R \omega^2(x_m-\gamma_l)\,d\xi\,d\eta\right]^{\frac{1}{2}}$$

$$\leq \max\left[\left(\iint_R \omega^2(x_i-\gamma_k)\,d\xi\,d\eta\right)^2 \iint_R \omega^2(x_m-\gamma_l)\,d\xi\,d\eta^2\right]$$

where $\omega(x)$ is a nonnegative function. Note that $\gamma_{(k,l)}^{(i,m)}$ depends on θ_m and θ_i only through their difference, a fact very useful in diagonalizing $[\Gamma]$ to find its eigenvalues. In addition $\gamma_{(k,l)}^{(i,m)}$ is an even function of $\theta_m-\theta_i$. These phenomena make the Gramian matrix look like

$$[\Gamma] = \begin{bmatrix} [\Gamma]^{(1,1)}, & [\Gamma]^{(1,2)}, & [\Gamma]^{(1,3)}, \ldots, [\Gamma]^{(1,3)}, & [\Gamma]^{(1,2)} \\ [\Gamma]^{(1,2)}, & [\Gamma]^{(1,1)}, & [\Gamma]^{(1,2)}, \ldots, [\Gamma]^{(1,4)}, & [\Gamma]^{(1,3)} \\ \vdots & & & \\ [\Gamma]^{(1,2)}, & [\Gamma]^{(1,3)}, & [\Gamma]^{(1,4)}, \ldots, [\Gamma]^{(1,2)}, & [\Gamma]^{(1,1)} \end{bmatrix} \quad (19)$$

a circulant matrix whose elements are symmetric matrices but are not necessarily circulant themselves. It is pointed out that the thresholding function of eq.(18) as mentioned earlier does not affect the circulant property of eq.(19).

Fast algorithms to diagonalize eq.(19) exist [8]. Because of its circulant property with elements $[\Gamma]^{(i,m)}$, $[\Gamma]$ can first be block diagonalized, a process similar to the one diagonalizing a circulant matrix except that here, entries of $[\Gamma]$ are matrices of sizes $M_r \times M_r$ instead of scalars. After block diagonalization, the block diagonal matrix can be diagonalized by individual block diagonalization processes, greatly reducing the computation required.

Summary

The imaging system of tomographic projections has been reviewed. The structure among the data gathered is utilized to yield a computational reduction in diagonalizing the Gramian matrix of the system for projections with Gaussian-shaped blur. Sharp projections are assumed to come up with a closed form expression of the elements of the Gramian matrix, although other approaches based on different assumptions also exist [8]. In the Appendix, a rectangular pulse form is assumed and the corresponding Gramian derived, although many other waveforms are also possible to get a closed form for $[\Gamma]$. It is observed that the $\gamma^{(i,i)}_{(k,k)}$, of the "autocorrelation" of the pulses increases as the width of pulses decreases*, while $\gamma^{(i,m)}_{(k,l)}$ $i \neq m$ or the "cross-correlation" is independent of the pulse width due to cancellation of the area under pulses with widths of pulses.

In the ideal case where the pulse is infinitely narrow,

*this phenomena is due to the fact that the value of a diagonal entry is equal to the peak value of the response of a matched filter which is a measure of the total energy of the signal. the total energy of a normalized Gaussian-shaped signal increases as its "variance" decreases as can be easily verified.

[Γ] resembles a diagonal matrix.

Appendix: Gramian Analysis for Rectangular-Shaped Projections

We assume $\omega(x) = b\ \mathrm{rect}(bx)$ as shown in figure 6 where

$$\mathrm{rect}(x) = 1 \qquad -\frac{1}{2} \leq x \leq \frac{1}{2}$$
$$= 0 \qquad \text{otherwise}$$

and $\omega(x)$ has been normalized such that $\int_{-\infty}^{\infty} \omega(t)\, dt = 1$. Similar to the case with Gaussian-shaped projections, we assume $\frac{1}{b} << 1$ so that the approximation schemes carry over here.

Referring to figure 4

$$\omega(x_i)\omega(x_m) = b^2 \qquad \text{if } (x_i, x_m) \in \text{parallelogram ABCD}$$
$$= 0 \qquad \text{otherwise}$$

Note that (x_i, x_m) are not rectangular coordinates

$$\iint_{-\infty}^{\infty} \omega(x_i)\omega^*(x_m)\, d\xi d\eta$$
$$= b^2 \iint_{ABCD} dvdu \qquad \langle 1 \rangle$$
$$= b^2 \times \mathrm{Area(ABCD)}$$

Since $OD = \frac{1}{2b} \sec\frac{\theta_m - \theta_i}{2}$ and $AO = \frac{1}{2b} \csc\frac{\theta_m - \theta_i}{2}$,

$$\iint_{-\infty}^{\infty} \omega(x_i)\omega^*(x_m)\, d\xi d\eta = 2b^2 \times AO \times OD \qquad \langle 2 \rangle$$
$$= \csc|\theta_m - \theta_i|$$

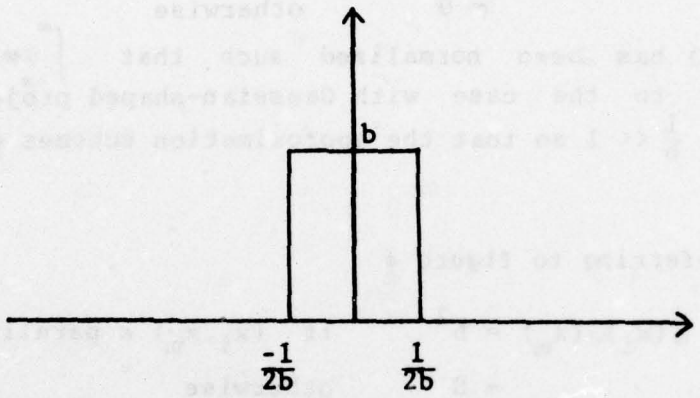


Figure 6. Rectangular function

Again, eq.<2> blows up at $\theta_m = \theta_i$. So for i=m, k=l, we again modify the limits of the integral in eq.<1>.

$$\gamma_{(k,k)}^{(i,i)} = \int_{-\infty}^{\infty} b^2 \mathrm{rect}(bu)\,du \int_{-\sqrt{1-\gamma_k^2}}^{\sqrt{1-\gamma_k^2}} 1\,dv$$

$$= 2b\sqrt{1-\gamma_k^2}$$

Thus

$$\gamma_{(k,l)}^{(i,m)} = U(i,m,k,l)\csc|\theta_m-\theta_i| \quad , \quad \theta_m \neq \theta_i$$

$$= 2b\sqrt{1-\gamma_k^2} \quad , \quad \theta_m=\theta_i,\ \gamma_k=\gamma_l$$

$$= 0 \quad , \quad \theta_m=\theta_i,\ \gamma_k \neq \gamma_l \qquad \text{<3>}$$

Again $\gamma_{(k,l)}^{(i,m)}$ depends on i,m only through the difference $\theta_i-\theta_m$. The diagonalization of [Γ] is thus similar to that discussed earlier. Thresholding upon the values of off-diagonal entries may be necessary if the projections are very densely sampled compared to the width of the projection functions.

Comparison of eq.(18) and eq.<3> shows that if $\sigma = \frac{1}{2b\sqrt{\pi}}$ and is sufficiently small then Gaussian projections and rectangular projections have the same Gramian matrix and thus from the DOF point of view, they are of the same performance.

References

1. R.A. Crowther, et.al., "The Reconstruction of a Three-Dimensional Structure from Projections and Its Application to Electron Microscopy," Proceedings Royal Society of London, Vol. A317, 1970, pp. 319-340.

2. R.N. Bracewell and A.C. Riddle, "Inversion of Fan Beam Scans in Radio Astronomy," Astrophysics Journal, Vol. 150, 1967, pp. 427-434.

3. G.N. Hounsfeld, "Computerized Transverse Axial Scanning (Tomography): Part I - Description of System," British Journal of Radiology, Vol. 46, 1973, pp. 1016-1022.

4. S. Twomey, "Information Content in Remote Sensing," Applied Optics, Vol. 13, 1974.

5. S. Hou and H.C. Andrews, "Fundamental Limits and Degrees of Freedom Imaging Systems," Proceedings of SPIE, Vol. 74, 1976.

6. E.K. Blum, Numerical Analysis and Computation: Theory and Practice, Addison-Wesley, 1972.

7. H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice-Hall, 1977.

8. D.G. McCaughey and H.C. Andrews, "Degrees of Freedom for Projection Imaging," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-25, No. 1, February 1977, pp. 63-73.

3.3 Psychovisual Transform Coding of Images (Supported by WPAFB under Contract F-33615-77-C-1016)

Charles F. Hall

Several recent papers have discussed the merits of transform image coding [1-6]. This particular research is concerned with an extension of this technique. The images to be coded are first preprocessed with an algorithm which is based on a model of the human visual system (HVS) [7] and [8]. The system to be used is shown in figure 1.

The first two blocks of figure 1 represent the HVS. The bandpass filter portion is implemented with the equation

$$A(f_r) = 2.6[0.0192+0.114f_r]\exp[-(0.114f_r)^{1.1}] \qquad (1)$$

(see [8] for an illustration of this isotropic filter function). The next three blocks represent the coding portion of the system. The preprocessed image is cosine transformed, quantized, and then inverse transformed. The cosine transform may be taken in any block size up to the full image size, which for this work was 256 x 256. The quantizer was a Max quantizer which assumed a Gaussian pdf for all transform coefficients. For the D.C. term a mean was estimated and subtracted from the D.C. value prior to quantization and then added back subsequently. After the transform coefficients are quantized they are inverse transformed and then passed through the final two stages which represent the inverse HVS.

There are several advantages to coding by transform blocks smaller than the image size. If one transforms 8 x 8 blocks for example, only eight lines are required before processing begins. In addition only two 8 x 8 covariance

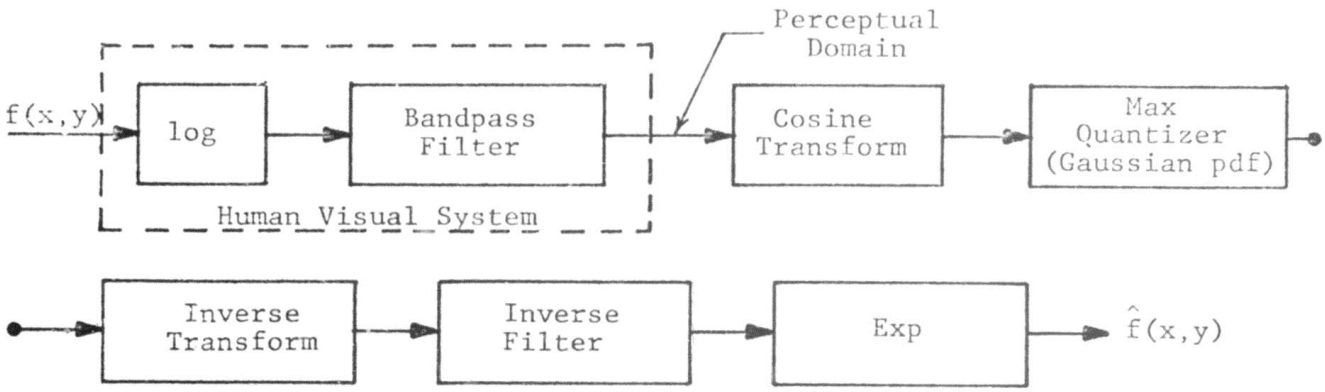


Figure 1. Psychovisual Cosine Coder

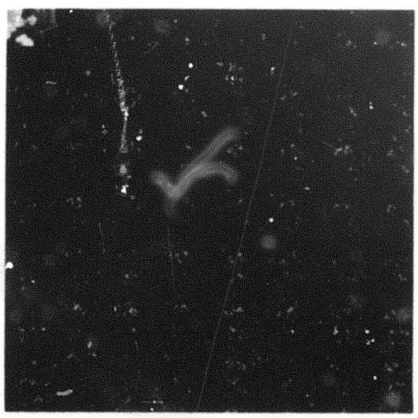


Fig 2. Mandala Cosine Transform
(8 x 8 blocks)



Fig 3. Cosine Coded 1 bit/pixel
(8 x 8 blocksize)
NMSE = .395%

matrices need to be diagonalized and used to compute an 8 x 8 bit map. To visualize how this bit map is used the entire cosine transform domain may be reordered for display purposes into a two-dimensional Mandala transform domain [9]. Such an ordering of an 8 x 8 block cosine transform of the GIRL picture is shown in figure 2. The input image was 256 x 256, thus there are 1024 terms for D.C. and every component. For the cosine transform there are 64 unique frequencies including "0", i.e. D.C. The 1024 terms make up a 32 x 32 array which, when scaled for visible display, form a subimage in figure 2. Every term in each subimage is coded with the same number of bits and since there are only 8 x 8 subimages our bit map is only 8 x 8. Note how the increasing harmonics (left to right and top to bottom) represent more and more "edge" information and the highest harmonic (lower right subimage) is almost random noise. These particular subimages are set to zero in the coding process. (The reader should note that <u>each</u> subimage has been scaled for display. There were more than 6 orders of magnitude differences between points in upper left and lower right before scaling).

If we actually quantize the block cosine transform of the GIRL for an average rate of 1 bit/pixel we would get the result shown in figure 3 (after inverse cosine transforming). The normalized mean square error (NMSE) in this case is 0.395%. If one examines the picture carefully, particularly at edges, small 8 x 8 blocks can be detected. These blocks become visible when errors in quantizing the D.C. term for that cell become large enough to shift the average grey level a detectible amount. If we code the image with a block size equal to the image size, obviously, this problem disappears. Figure 4 contains such an image. In this case the bit rate was 0.7 and the NMSE was 0.33%. Thus, we have reduced the bit rate by 30% while decreasing



Fig 4. Cosine Coded .7 bit/pixel
(256 x 256 blocksize)
NMSE = .33%

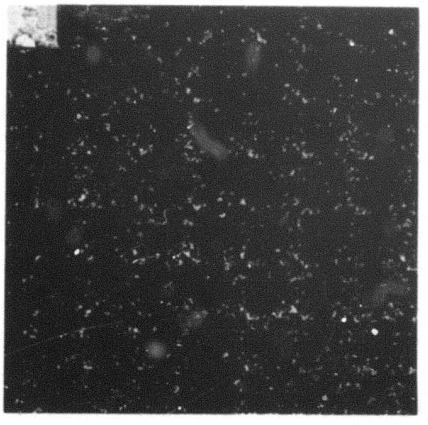


Fig 5. Psychovisual Cosine Coded
1 bit/pixel
(8 x 8 blocksize)
NMSE = .434%



Fig 6. Psychovisual Cosine Coded
.7 bit/pixel
(256 x 256 blocksize)
NMSE = .35%

the NMSE about 20%.

In the previous paragraph we considered an image which had not been preprocessed by the human visual system model. If we now preprocess the GIRL with the HVS and 8 x 8 block code as shown in figure 1 we obtain the result shown in figure 5. This image was coded at 1 bit/pixel and the NMSE was 0.434%. Although the NMSE is larger than the 8 x 8 block result in figure 3, a close inspection reveals that the 8 x 8 subblocks are not as visible in figure 5. Again, a 256 x 256 block size was used and the result is shown in figure 6. In this case the rate was 0.7 bit/pixel and the NMSE was 0.35%. The reduction in NMSE is again approximately 20% with a 30% decrease in bit rate.

If one refers to figure 1 and considers the implementation of the HVS linear filter is in the Fourier domain the logical question is, "why not code in the filtered Fourier domain?" This coding scheme eliminates the cosine transforming completely and also eliminates a forward and an inverse Fourier transform. The system is therefore reduced to that shown in figure 7.

The GIRL picture was coded in this manner. The result is shown in figure 8. This image has a rate of 1 bit/pixel and a NMSE of 0.26%. The quality of this image compared to the other results within this paper indicates that coding the HVS filtered Fourier domain (as proposed in [7]) results in excellent quality images. The technique is being extended to color imagery as proposed in [7] and initial results are encouraging.

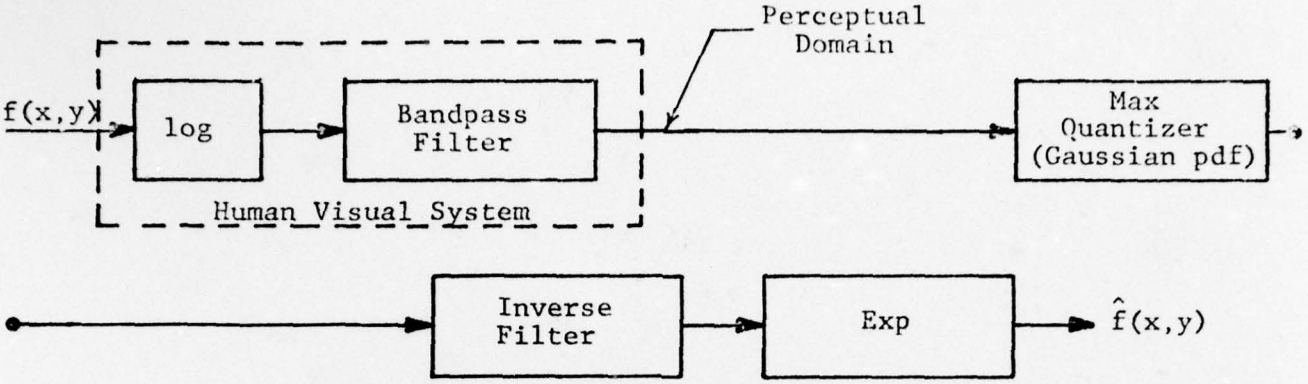


Figure 7. Psychovisual Fourier Coder



Fig 8. Psychovisual Fourier Coded 1 bit/pixel
(256 x 256 blocksize)
NMSE = .26%

References

1. W.K. Pratt, J. Kane, and H.C. Andrews, "Hadamard Transform Image Coding," Proceedings IEEE, Vol. 57, January 1969, pp. 58-68.

2. A. Habibi and P.A. Wintz, "Image Coding by Linear Transformation and Block Quantization," IEEE Transactions on Communication Technology, Vol. COM-19, No. 1, February 1971, pp. 50-62.

3. W. K. Pratt, "Spatial Transform Coding of Color Images," IEEE Transactions on Communication Technology, COM-19, No. 6, December 1971, pp. 980-992.

4. P.A. Wintz, "Transform Picture Coding," Proceedings IEEE, Vol. 60, No. 7, July 1972, pp 809-820.

5. W.K. Pratt, W. Chen, and L.R. Welch, "Slant Transform Image Coding," IEEE Transactions on Communications, Vol. COM-22, No. 8, August 1974, pp. 1075-1093.

6. W. Chen and C.H. Smith, "Adaptive Coding of Color Images Using Cosine Transform," Proceedings International Communications Conference, Philadelphia, June 1976, pp. 47-7 to 47-13.

7. C.F. Hall and E.L. Hall, "A Nonlinear Model for the Spatial Characteristics of the Human Visual System," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-7, No. 3, March 1977, pp. 161-170.

8. C.F. Hall, "Image Filtering Based on Psychophysical Characteristics of the Human Visual System," University of Southern California, Image Processing Institute, USCIPI

Report 740, March 1977, p. 79-88.

9. J.T. Kajiya, "Group Representations and the Modeling of Visual Perception," in Image Science Mathematics, C.O. Wilde and E. Barrett, eds., North Hollywood Western Periodicals Company, 1977, pp. 67-70.

3.4 Statistical Analysis of a Model of the Human Visual System (Supported by WPAFB under Contract F-33615-77-C-1016)

Charles F. Hall and Lloyd R. Welch

In a recent paper Hall and Hall have discussed several models for the HVS [1]. In this section we will analyze one of these models (shown in figure 1), both statistically and deterministically. The main objective will be to determine the assumptions which may be made about the characteristics of images after they are passed through the model. In particular, those characteristics consistent with the assumptions must be made to apply the known solutions to the parametric set of equations which are the heart of rate-distortion theory [2,3].

We will begin by first considering the assumption of Gaussian pdf. The image will be assumed to be Gaussian after passing through the nonlinearity. This assumption presents no problem since the filter of figure 1 is linear and hence the output of the filter will be Gaussian if the input is. When the output of the nonlinearity is Gaussian, what is the pdf of the input? This question is answered quite simply by applying a fundamental theorem discussed in section 5-2 of Papoulis [4].

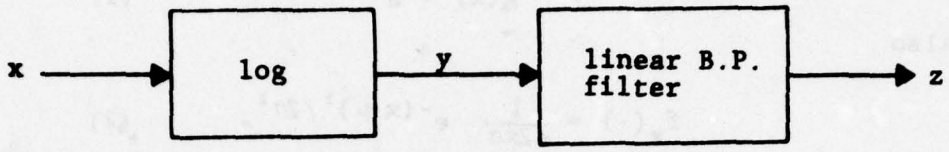


Figure 1. Nonlinear model of HVS.



Figure 2.

For the analysis consider figure 2. Let x be the Gaussian distributed image. From figure 2,

$$y = g(x) = e^{x} \tag{1}$$

also

$$f_x(\cdot) = \frac{1}{\sqrt{2\pi}\sigma} \; e^{-(x-\mu)^2/2\sigma^2} \tag{2}$$

Solving for x in terms of y

$$x_1 = \ln y_1 \tag{3}$$

also

$$g'(x) = e^{x} \tag{4}$$

Thus

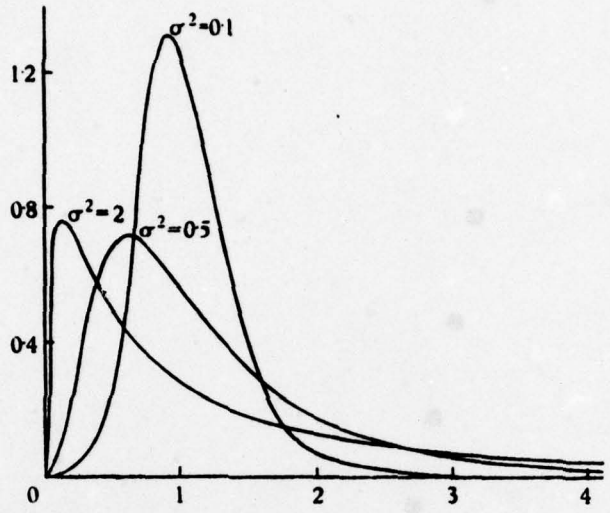
$$f_y(\cdot) = \frac{f_x(x_1)}{|g'(x_1)|}$$

becomes

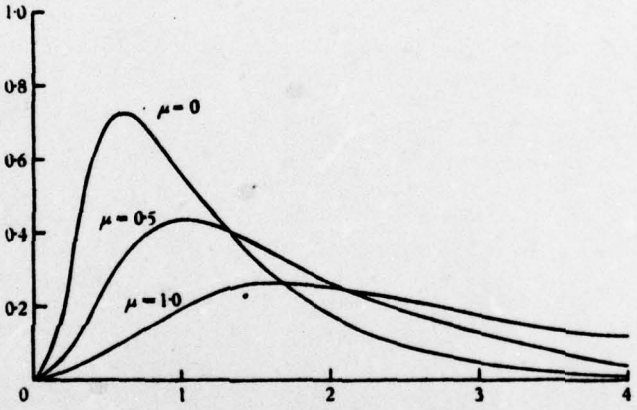
$$\frac{1}{\sqrt{2\pi}\sigma y_1} \; e^{-(\ln y_1 - \mu)^2/2\sigma^2} \qquad y_1 \geq 0 \tag{5}$$

This pdf is known as the lognormal distribution and has several interesting characteristics [5]. Plots of this function for several values of μ and σ^2 are shown in figure 3. The similarity to image histogram data is immediately apparent. The histogram data for the Kodak GIRL image was plotted on log probability paper and is shown in figure 4. Note that the data points are essentially three straight lines over the 1% to 99% range. This indicates the data is strongly lognormal. Thus, we see the HVS models help satisfy the common assumption (which is unrealistic for an unprocessed image) that the data is Gaussian.

Next, let us consider the entropy of the two processes. We will use the common definition for differentiated entropy



(a)



(b)

Figure 3. Frequency curves of the lognormal distribution for a) three values of σ^2 and b) three values of μ.

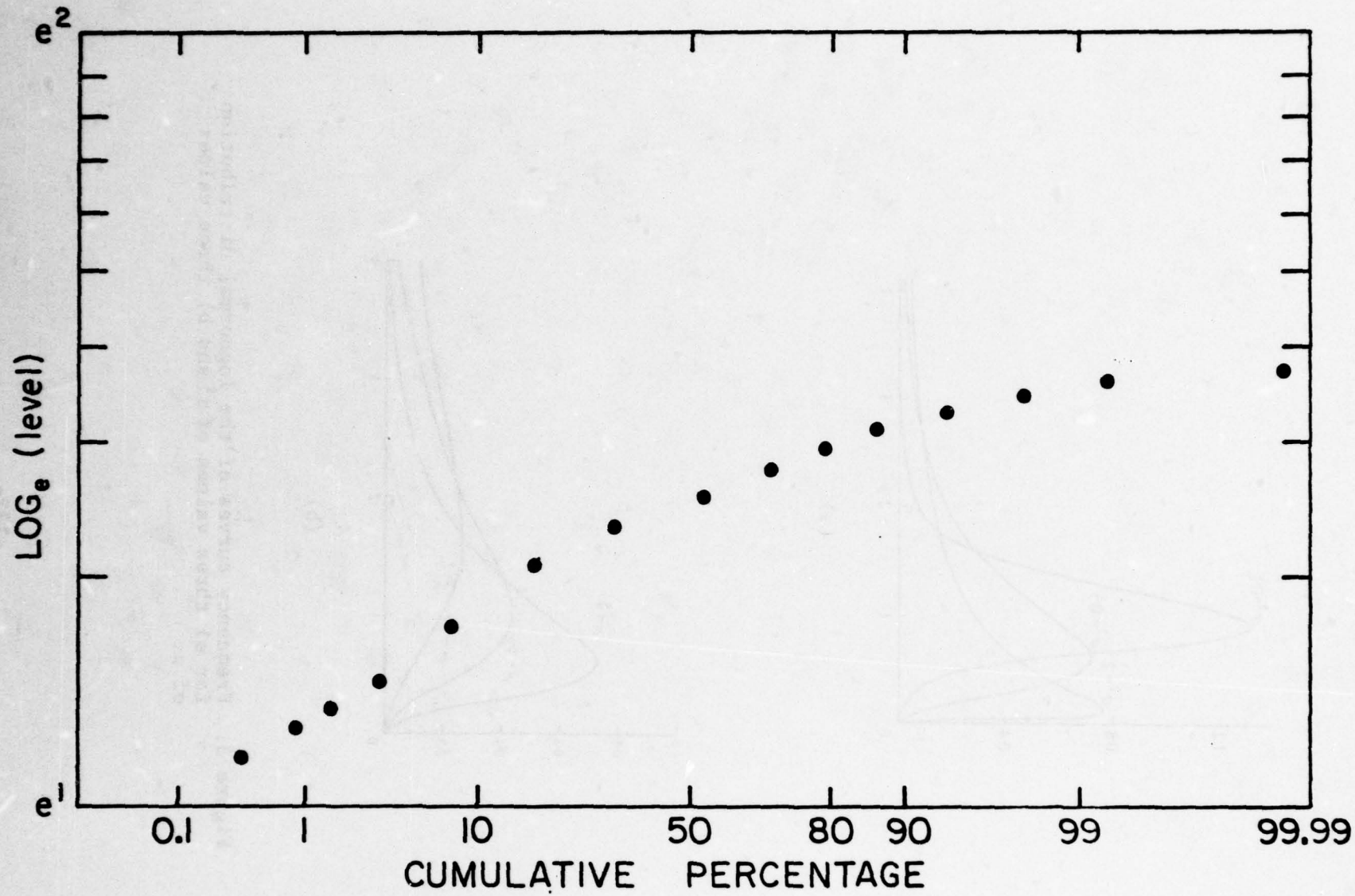


Figure 4. First-order histogram of Kodak GIRL plotted on log-probability coordinate system.

$$H(x) = -\int_{-\infty}^{\infty} p(x) \ln p(x)\, dx \qquad (6)$$

where $p(\cdot)$ denotes the pdf of x. Shannon has shown for the Gaussian case we get [6]

$$H(x) = \sqrt{2\pi e}\sigma \qquad (7)$$

Consider the lognormal distribution,

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}x}\ e^{-(\ln x-\mu)^2/2\sigma^2} \qquad (8)$$

then

$$\ln p(x) = -\ln[\sigma\sqrt{2\pi}x] - \frac{(\ln x-\mu)^2}{2\sigma^2} \qquad (9)$$

and

$$H(x) = \int_0^{\infty} p(x)\left[\ln(\sigma\sqrt{2\pi}x) + \frac{(\ln x-\mu)^2}{2\sigma^2}\right] dx \qquad (10)$$

where the lower limit of integration has been changed to 0 since x ranges from 0 to ∞ for the lognormal pdf. Thus,

$$H(x) = \int_0^{\infty} p(x)\ln(\sigma\sqrt{2\pi})dx + \int_0^{\infty} p(x)\ln x\, dx + \int_0^{\infty} p(x)\,\frac{(\ln x-\mu)^2}{2\sigma^2}\, dx \qquad (11)$$

but

$$\int_0^{\infty} p(x)dx = 1 \qquad (12)$$

therefore

$$H(x) = \ln(\sigma\sqrt{2\pi}) + \int_0^{\infty} p(x)\ln x\, dx + \int_0^{\infty} p(x)\,\frac{(\ln x-\mu)^2}{2\sigma^2}\, dx \qquad (13)$$

Now let $y = \ln x$ which implies $x = e^y$ and that $dx = e^y dy$. Also, when $x = 0$, $y = -\infty$ and when $x = \infty$, $y = \infty$. Substituting into eq.(8) gives

$$p(\cdot) = \frac{1}{\sqrt{2\pi}\sigma \exp(y)} e^{-(y-\mu)^2/2\sigma^2} \tag{14}$$

Completing the substitution in H gives

$$H(\cdot) = \ln(\sigma\sqrt{2\pi}) + \int_{-\infty}^{\infty} \frac{y}{\sqrt{2\pi}\sigma} e^{-(y-\mu)^2/2\sigma^2} dy + \frac{1}{2\sigma^2} \int_{-\infty}^{\infty} \frac{(y-\mu)^2}{\sqrt{2\pi}\sigma} e^{-(y-\mu)^2/2\sigma^2} dy \tag{15}$$

The first integral is just the mean of a Gaussian pdf and the second integral is the variance, therefore,

$$H(\cdot) = \ln(\sigma\sqrt{2\pi}) + \mu + \frac{1}{2\sigma^2}\sigma^2 = \ln(\sigma\sqrt{2\pi e}) + \mu \tag{16}$$

Thus, for a nonzero μ we have an entropy change going through the system which is equal to μ, the mean of the resultant Gaussian pdf.

Next we will consider the autocorrelation and power spectrum for such a system. Now

$$R_x(\tau) = E\{x(t)x(t+\tau)\}$$

$$= \int_{-\infty}^{\infty}\int\cdots\int^{\infty} \frac{e^{y_1} e^{y_2} e^{-\frac{1}{2}(\vec{y}-\mu)^T C^{-1}(\vec{y}-\mu)}}{(2\pi)^{N/2}|C|^{\frac{1}{2}}} d\vec{y} \tag{17}$$

where N is the dimension of the system and equals 2 for our case. Also,

$$\vec{y}^T = (y_1\ y_2) \tag{18}$$

and

$$C = \begin{bmatrix} \sigma^2 & \sigma^2\rho_y(\tau) \\ \sigma^2\rho_y(\tau) & \sigma^2 \end{bmatrix} \tag{19}$$

In this case $\rho(\tau)$ is the normalized autocovariance defined as

$$\rho_y(\tau) = E\left\{[y(t)-\mu]\,[y(t+\tau)-\mu]\right\} \tag{20}$$

Now let $\lambda_1 = \lambda_2 = -j$ and rewrite $R_x(\tau)$,

$$R_x(\tau) = E\left\{e^{j\lambda_1 y_1 + j\lambda_2 y_2}\right\}$$

$$= \int_{-\infty}^{\infty}\int \frac{e^{j\lambda_1 y_1 + j\lambda_2 y_2} e^{-\frac{1}{2}(\vec{y}-\mu)^T C^{-1}(\vec{y}-\mu)}}{\sqrt{2\pi}\,|C|^{\frac{1}{2}}}\, dy_1\, dy_2 \tag{21}$$

The above equation is in the form of a characteristic equation. The characteristic function for a two-dimensional Gaussian of nonzero mean is (see [4], p. 255)

$$\Phi(\lambda) = e^{-\frac{1}{2}\vec{\lambda}^T C\vec{\lambda}}\, e^{j\vec{\lambda}^T\mu} \tag{22}$$

where $\lambda^{-\tau} = (-j\ -j)$, therefore,

$$R_x(\tau) = \Phi\begin{pmatrix} -j \\ -j \end{pmatrix} = e^{\sigma^2[1+\rho_y(\tau)]+\mu(t)+\mu(t+\tau)} \tag{23}$$

For the input process x, we will assume the general form for the autocorrelation in terms of the covariance of the process,

$$R_x(\tau) = \mu_x^2 + C_x(\tau) = \mu_x^2 + \sigma_x^2\,\rho_x(\tau) \tag{24}$$

Then, from eq.(23) we have

$$R_x(\tau) = \mu_x^2 + \sigma_x^2 \rho_x(\tau) = e^{2\mu_y + \sigma_y^2[1+\rho_y(\tau)]} \quad (25)$$

Now

$$E\{x\} = \mu_x = E\{e^y\} = e^{\mu_y + \frac{1}{2}\sigma_y^2} \quad (26)$$

therefore

$$\mu_x^2 = e^{2\mu_y + \sigma_y^2} \quad (27)$$

and substituting in eq.(25) yields

$$\mu_x^2 + \sigma_x^2 \rho_x(\tau) = \mu_x^2 e^{\sigma_y^2 \rho_y(\tau)} \quad (28)$$

or

$$1 + \frac{\sigma_x^2}{\mu_x^2} \rho_x(\tau) = e^{\sigma_y^2 \rho_y(\tau)} \quad (29)$$

Expanding the right side of this equation gives

$$1 + \frac{\sigma_x^2}{\mu_x^2} \rho_x(\tau) = 1 + \sigma_y^2 \rho_y(\tau) + \sum_{k=2}^{\infty} \sigma_y^{2k} \frac{\rho_y{}^k(\tau)}{k!} \quad (30)$$

The sum of eq.(30) represents the error if we use only the first two terms of the expansion. The normalized covariance of any process will have an upper bound of 1. Also, for typical picture data $\sigma_y^2 \approx 0.5$. Thus, the worst case expansion is on $e^{0.5}$ and the error in taking only the first two terms is less than 10%. This is a very conservative error bound, particularly since it assumes the data are completely correlated. We will neglect the error term in eq.(30), giving

$$1 + \frac{\sigma_x^2}{\mu_x^2} \rho_x(\tau) = 1 + \sigma_y^2 \rho_y(\tau) \quad (31)$$

From eq.(29), if we take the natural logarithm of both sides,

$$\ln\left\{1 + \frac{\sigma_x^2}{\mu_x^2} \rho_x(\tau)\right\} = \sigma_y^2 \rho_y(\tau) \quad (32)$$

Typical image data for the x process gives a $\frac{\sigma_x^2}{\mu_x^2}$ ratio of 0.16. Again for a worst case $\rho_x(\tau) = 1$ we get $\ln(1+0.16) = 0.14842$ which is within 8% of 0.16. Thus, within experimental error

$$\ln\left(1 + \frac{\sigma_x^2}{\mu_x^2}\,\rho_x(\tau)\right) \approx \frac{\sigma_x^2}{\mu_x^2}\,\rho_x(\tau) = \sigma_y^2\,\rho_y(\tau) \tag{33}$$

Now $\rho_x(\tau) = 1$ for $\tau = 0$ for any valid covariance function, therefore

$$\frac{\sigma_x^2}{\mu_x^2} \approx \sigma_y^2 \tag{34}$$

Substituting eq.(34) into eq.(31) gives

$$1 + \sigma_y^2\,\rho_x(\tau) = 1 + \sigma_y^2\,\rho_y(\tau) \tag{35}$$

which implies that $\rho_x(\tau) = \rho_y(\tau)$. Thus, the output autocorrelation becomes

$$R_y(\tau) = \mu_y^2 + \sigma_y^2\,\rho_x(\tau) \tag{36}$$

By definition the power spectrum of y process is

$$S_y(\omega) = \int_{-\infty}^{\infty} R_y(\tau) e^{-j\omega\tau} d\tau \tag{37}$$

Therefore,

$$\begin{aligned} S_y(\omega) &= \int_{-\infty}^{\infty} \left[\mu_y^2 + \sigma_y^2\,\rho_x(\tau)\right] e^{-j\omega\tau} d\tau \\ &= 2\pi\mu_y^2\,\delta(\omega) + \sigma_y^2 \int_{-\infty}^{\infty} \rho_x(\tau)\, e^{-j\omega\tau} d\tau \end{aligned} \tag{38}$$

This relationship is of great importance in rate-distortion applications. Given an input autocorrelation, we can compute the output power spectrum which can be used in the equations (see [2], p. 117)

$$D_\theta = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\min\ \theta,\ \Phi(\omega)]\, d\omega \tag{39}$$

and

$$R(D_\theta) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \left[\max\ \theta,\ \log\frac{\Phi(\omega)}{\theta}\right] d\omega \quad (40)$$

Thus, the rate-distortion curve of the source x, after passing through the logarithmic nonlinearity, can be easily obtained.

A common assumption for imagery data is that it is Markov, in particular,

$$\rho_x(\tau) = e^{-\alpha|\tau|} \quad (41)$$

Substituting this form into the power spectrum equation gives

$$\begin{aligned} S_y(\omega) &= \sigma_y^2 \int_{-\infty}^{\infty} e^{-\alpha|\tau|} e^{-j\omega\tau} d\tau + 2\pi\mu_y^2\ \delta(\omega) \\ &= \frac{2\alpha\sigma_y^2}{\alpha^2+\omega^2} + 2\pi\mu_y^2\ \delta(\omega) \end{aligned} \quad (42)$$

Referring to figure 1 and letting $H(\omega)$ be the linear filter function we see that

$$S_z(\omega) = S_y(\omega)\ |H(\omega)|^2 = \left[\frac{2\alpha\sigma_y^2}{\alpha^2+\omega^2} + 2\pi\mu_y^2\ \delta(\omega)\right] |H(\omega)|^2 \quad (43)$$

We have shown that if an image source is lognormal and Markov, then after passing through the HVS model of figure 1

it will be Gaussian Markov with a power spectrum defined by eq.(43). Furthermore, the entropy of the original source will be changed by μ, the mean of the resultant Gaussian pdf. These results are being used to find theoretical bounds on the coding improvements which can be realized when images are preprocessed with models of the HVS.

References

1. C.F. Hall and E.L. Hall, "A Nonlinear Model for the Spatial Characteristics of the Human Visual System," Transactions on Systems, Man, and Cybernetics, Vol. SMC-7, No. 3, March 1977, pp. 161-170.

2. T. Berger, Rate Distortion Theory, Prentice-Hall, Englewood Cliffs, 1971.

3. L.D. Davisson, "Rate-Distortion Theory and Application," Proceedings of the IEEE, Vol. 60, No. 7, July 1972, pp. 800-808.

4. A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.

5. J. Aitchison and J.A.C. Brown, The Lognormal Distribution, University Press, Cambridge, 1957.

6. C.E. Shannon, A Mathematical Theory of Communication, University of Illinois Press, Urbana, 1949.

3.5 A Technique of A Posteriori Restoration

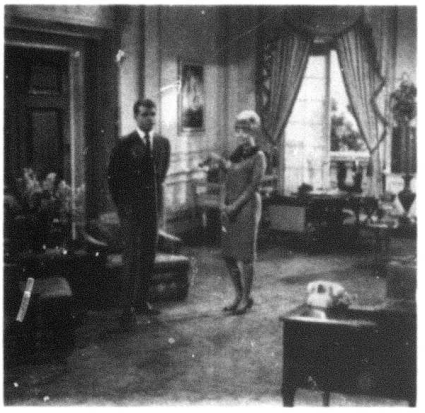
John Morton

This effort is concerned with restoring a degraded image while assuming a minimum of a priori information. The previous reports [1-2] have outlined the approach to be investigated and in addition, contain summaries of the progress achieved up to the respective report dates. This report will present some preliminary results.

Figure 1 contains four statistically similar images, statistically similar in the sense that the power spectra corresponding to the four images are similar. Note that the subject content of one image is different from the other images. The difference in subject matter was intentional to emphasize the point that different image subject matter may corrrespond to similar power spectra.

A contour plot of the log of the power spectrum calculated from figure 1a is presented in Figure 2. Figure 3 contains the log of the average power spectrum calculated from figures 1b, 1c, and 1d. Note the close similarity between figures 2 and 3. The power spectrum depicted in figure 3 will be used to estimate the magnitude of the optical transfer function (OTF) via the method of Cannon [3].

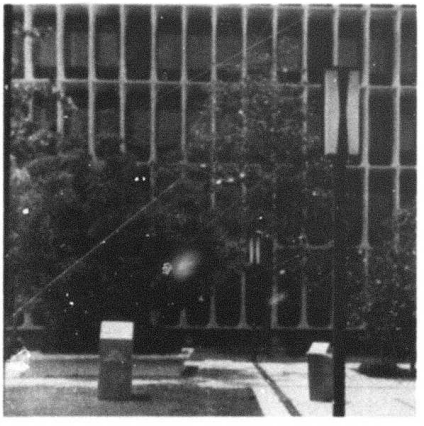
Figure 4 contains a perspective plot of the point spread function (PSF) to be used to degrade the image presented in figure 1a. Figures 5, 6 and 7 contain the image of figure 1a degraded by the PSF of figure 4 together with restorations assuming different combinations of a priori information. Table 1 defines the assumptions associated with the different restorations.



a)



b)



c)



d)

Figure 1. Four statistically similar images.

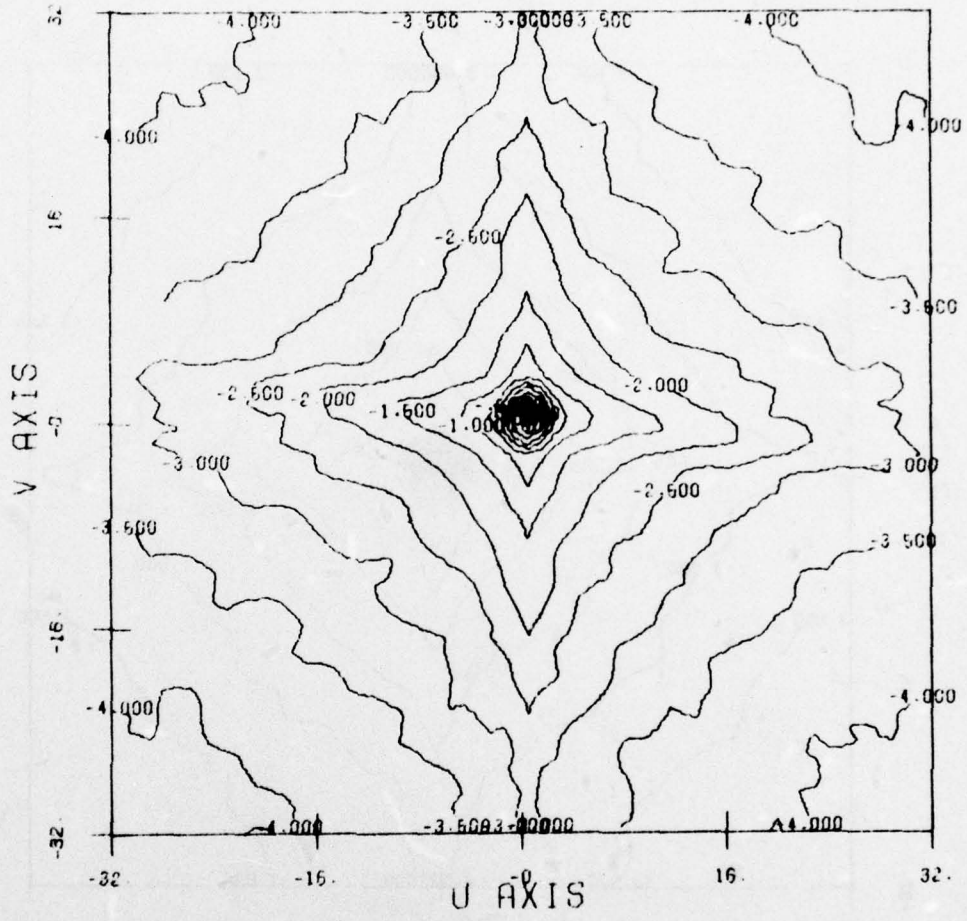


Figure 2. Contour plot of the $\log_{10}$ of the power spectrum calculated from Figure 1a.

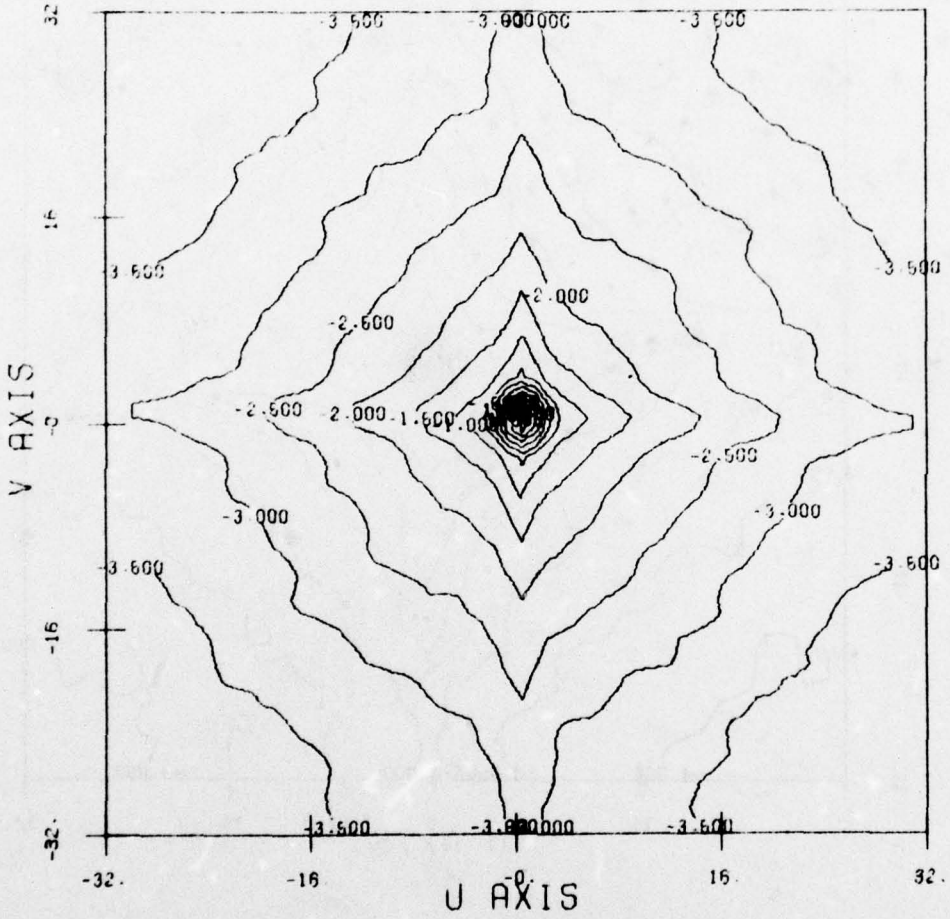


Figure 3. *Contour plot of the* $\log_{10}$ *of the average power spectrum calculated from Figures 1b, 1c, and 1d.*

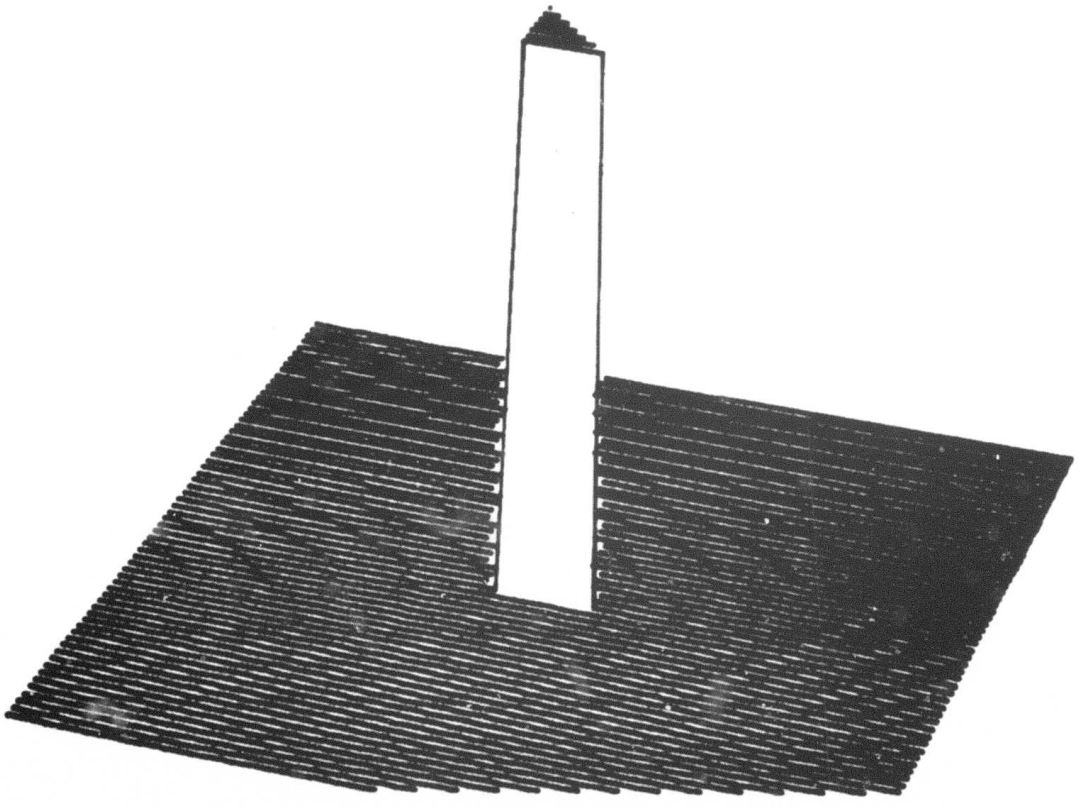
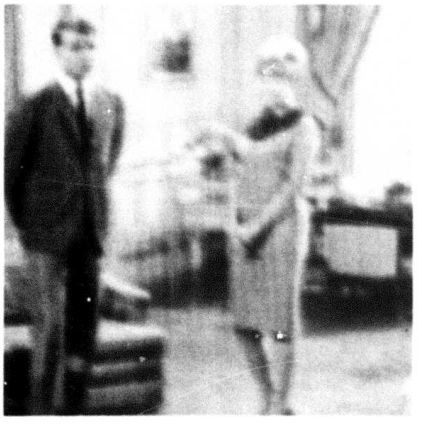
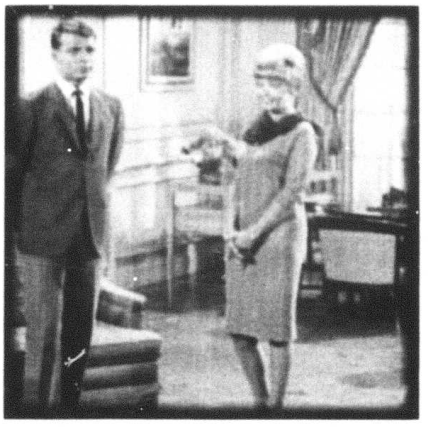


Figure 4. Perspective plot of the PSF to be used to degrade the image presented in Figure 1a.



a)



b)

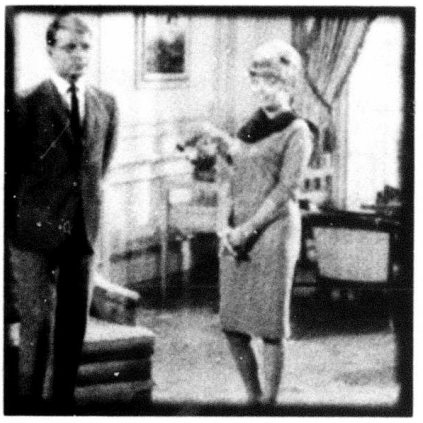


c)

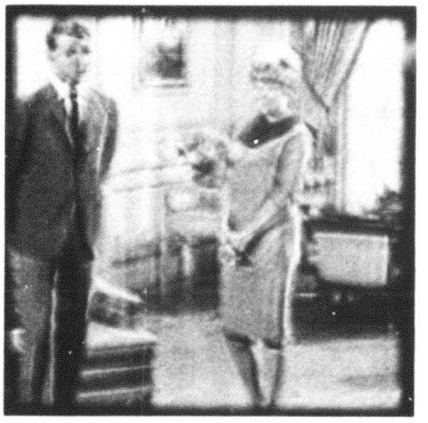


d)

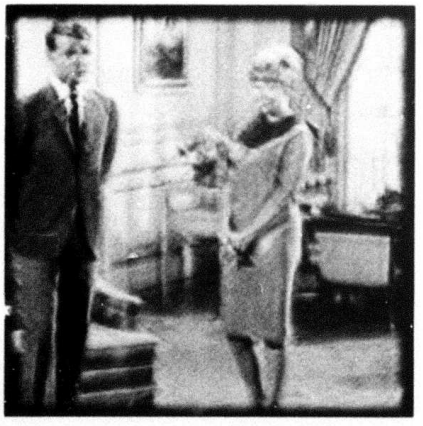
Figure 5. a) Figure 1a degraded by point spread function of Figure 4. b), c), and d) restorations see Table 1.



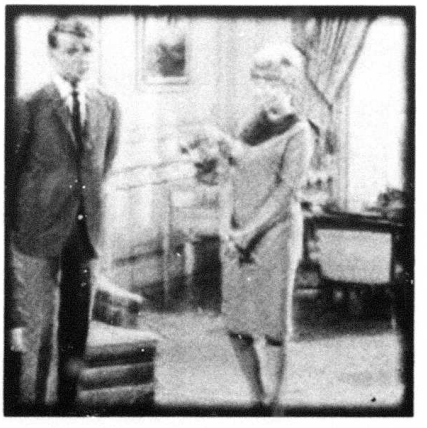
a)



b)

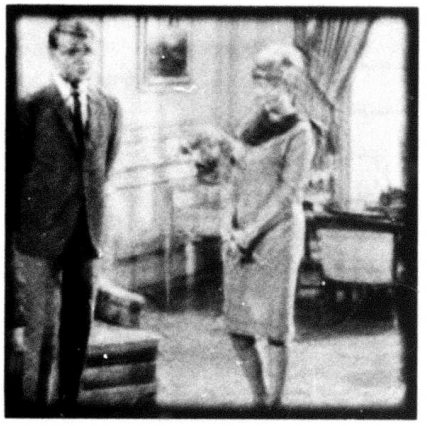


c)

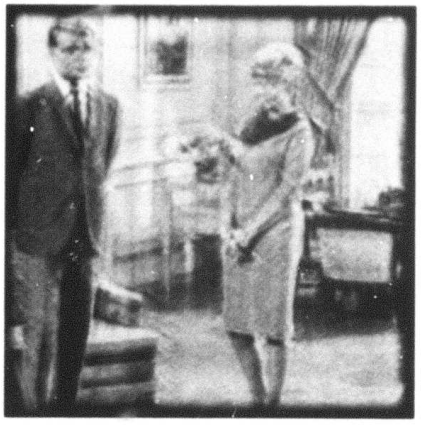


d)

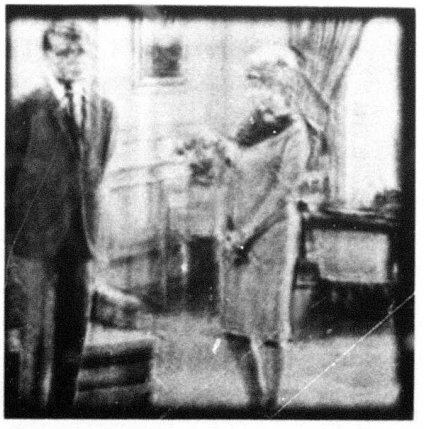
Figure 6. a), b), c), and d) restorations see Table 1.



a)



b)



c)

Figure 7. a), b), and c) restorations see Table 1.

FIGURE	PSF MAGNITUDE	PSF PHASE	POWER SPECTRUM OF UNDEGRADED IMAGE
5b	given	given	given
5c	given	estimated as 0	given
5d	given	estimated	given
6a	estimated	given	estimated
6b	estimated	estimated	estimated
6c	estimated	estimated as 0	estimated
6d	estimated	estimate multiplied by -1	estimated
7a	estimated	estimate multiplied by -2	estimated
7b	estimated	estimate multiplied by -3	estimated
7c	estimated	estimate multiplied by -4	estimated

Table 1. Key to Restorations of Figures 5, 6, and 7.

If one has knowledge of the OTF and also knowledge of the power spectrum of the undegraded image, one can obtain the restoration of figure 5b. Figure 5c assumes knowledge of the magnitude of the OTF, knowledge of the power spectrum of the undegraded image, and an estimate of the phase of the OTF of zero everywhere. Using the algorithm previously reported [1-2] to estimate the phase of the OTF and assuming knowledge of the magnitude of the OTF and knowledge of the power spectrum of the undegraded image, one obtains figure 5d. Upon comparison of figures 5c and 5d it is apparent that the phase estimate per se does not afford any improvement over a phase estimate of zero everywhere.

Upon examination of figure 6a where the magnitude of the OTF is estimated using the power spectrum depicted in figure 3, and where in addition the power spectrum of Figure 3 is used as an estimate of the power spectrum of the undegraded image, and lastly assuming the phase of the OTF is known, it is evident that estimation of the magnitude of the OTF is very good.

Figure 6b assumes no a priori knowledge. Figure 6c assumes no apriori knowledge of the magnitude of the OTF, nor any knowledge of the power spectrum of the undegraded image and assumes an estimate of zero everywhere as an estimate of the phase of the OTF.

The phase estimate can be shown experimentally to converge to a phase which is of incorrect sign and to a value less than, in some cases much less than, the correct value. Thus, it is not unrealistic to change the sign of the estimated phase and boost the estimates by factors. Figures 6d, 7a, 7b, and 7c correspond to the phase estimates with a sign change and multiplications of 1, 2, 3, and 4 respectively. As judged by the increased sharpness in the

collar of the man and less distortion in the face of the man, perhaps some improvement has been obtained.

References

1. Semiannual Technical Report, Harry C. Andrews - project director, September 30, 1976, USCIPI Report 720.

2. Semiannual Technical Report, Harry C. Andrews - project director, March 31, 1977, USCIPI Report 740.

3. T.M. Cannon, "Digital Imaging Deblurring by Nonlinear Homomorphic Filtering," Department of Computer Science, University of Utah, ARPA Technical Report UTEC-CSC-74-091, August 1974.

3.6 On-Axis Optical Filtering System (Supported by NSF Grant ENG-76-15318 and AFOSR under Contract AFOSR-77-3285)

Alexander A. Sawchuk and Chung-Kai Hsueh

Computer generated holograms have many advantages over optical generated holograms. If the hologram is used in a filtering system, the filter H(f) can be specified without a physical specimen of the impulse response. Computer generated holograms also eliminate the need of very stable and low noise photographic recording set up. With computer plotting, the nonlinearity of the film and other effects of the system can be precompensated and more flexibility can be achieved.

However most of the computer generated holograms are off-axis. The desired output appears on the first diffraction order which has a maximum diffraction efficiency of 41% when square wave phase grating is used. Since most of the energy is concentrated at the center, to get useful results, we have to increase the carrier frequency and

restrict the size of the impulse response and the object to avoid overlapping between two adjacent diffraction orders.

On-axis holograms, on the other hand, do not have the problems mentioned above. The desired output is on the axis and contains most of the energy. As will be explained later if the high order diffractions can be eliminated properly the whole output plane is available and the size of the object need not be restricted. The sampling rate of the filter then depends on the extent of the impulse response rather than on the object extent which is generally much larger. Therefore the sampling rate is much lower compared to other holographic type filters. However for proper suppression of the higher diffraction orders, the sampling rate still has to be considerably high as will become clear in the following analysis.

To simplify the notation, we use one-dimensional variables in the following formulation. The complex filter $H_s(u)$ is designed to have coherent impulse response $\underline{h}(x)$ and incoherent response $|h(x)|^2$ in the image plane. Denoting the Fourier transform of $h(x)$ by $H(u)$, the digital hologram $H_s(u)$ approximates $H(u)$ by a series of weighted pulses

$$H_s(u) = \left\{\left[\sum_n H(n\delta u)\,\delta(u-n\delta u)\right] * p(u)\right\} \cdot M(u) \qquad (1)$$

shown in figure 1. The pulse spacing is δu, each pulse has shape $p(u)$ with weights given by sampled values $H(n\delta u)$. In general the pulse has a rectangular shape with width δu, i.e.,

$$p(u) = \text{rect}\,\frac{u}{\delta u} \qquad (2)$$

The function $M(u)$ is a mask representing the physical size limit of the hologram. If we assume that $H(u)$ is

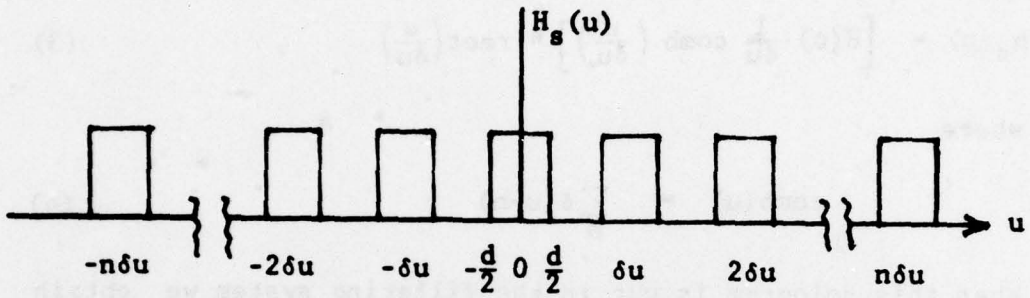


Figure 1. Complex Pupil Mask.

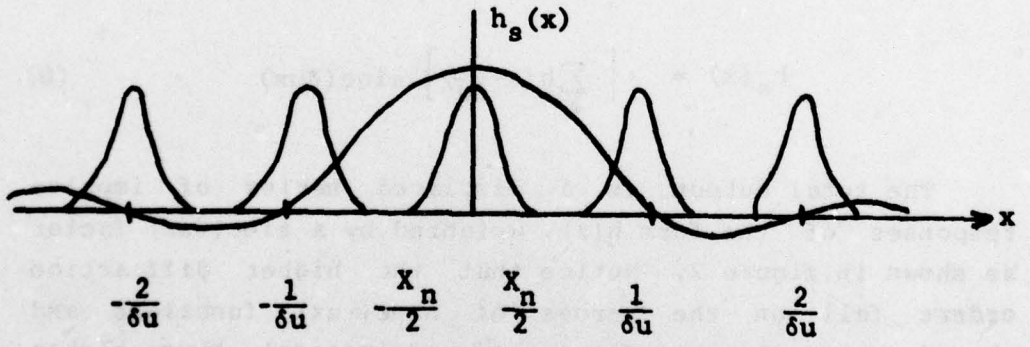


Figure 2. Output of Fourier Transform Lens.

bandlimited and the size of the hologram is large enough to cover the whole spectrum, then we may drop the mask function. Thus (1) becomes

$$H_s(u) = \left[H(u)\cdot\frac{1}{\delta u}\ \mathrm{comb}\left(\frac{u}{\delta u}\right)\right] * \mathrm{rect}\left(\frac{u}{\delta u}\right) \tag{3}$$

where

$$\mathrm{comb}(u) = \sum_n \delta(u-n) \tag{4}$$

When this hologram is put in the filtering system we obtain the impulse response $h_s(x)$ which is the inverse Fourier transform of $H_s(u)$

$$h_s(x) = \mathcal{F}^{-1}\left\{H_s(u)\right\}$$

$$= \left[h(x) * \mathrm{comb}(\delta u x)\right] \cdot (\delta u)\,\mathrm{sinc}(\delta u x) \tag{5}$$

or

$$h_s(x) = \left[\sum_k h(x-\frac{k}{\delta u})\right] \mathrm{sinc}(\delta u x) \tag{6}$$

The total output is a displaced series of impulse responses of the form h(x), weighted by a sinc(δux) factor as shown in figure 2. Notice that the higher diffraction orders fall on the zeroes of sinc(δux) functions and therefore are strongly attenuated. In general, these higher order responses can be neglected and the whole image plane is available. Therefore the size of the object is no longer restricted and the sampling interval of the filter is determined by the extent of the impulse response x_n rather than by the extent of the object x_o. We need

$$\delta u < \frac{1}{x_n} \tag{7}$$

However, to properly eliminate the higher order responses, we require the higher order responses to be concentrated around the zeroes of the sinc(δux) function and δu is generally smaller than the requirement imposed by (7).

If the higher order responses cannot be neglected, the sampling rate must be much higher. To avoid overlapping between adjacent orders we require

$$\frac{1}{\delta u} > x_n + x_o \tag{8}$$

Generally x_n is much smaller than x_o so this condition becomes

$$\delta u > \frac{1}{x_o} \tag{9}$$

The sampling interval is determined by the extent of the object rather than the impulse response.

To further reduce the intensity of higher diffraction orders, we may use a smoother microdensitometer plot. For example if a simple linear interpolator is used then the impulse response becomes

$$h_s(x) = \left[\sum_k h(x - \frac{k}{\delta u}) \right] \text{sinc}^2(\delta u x) \tag{10}$$

Additional sidelobe reduction is achieved by the $\text{sinc}^2(\delta ux)$ factor. Basically this problem is similar to the apodization problem in which a proper window is used in order to reduce the sidelobes. However these different profiles may be difficult to implement by optical plotting.

Besides the reduction of the sidelobes the sinc function also modifies the impulse response slightly due to the modulation near the peak of $\mathrm{sinc}(\delta ux)$, but this effect is small. If desired, the function h(x) could be premultiplied by a window function correction term. The transform of

$$h'(x) = \frac{h(x)}{\mathrm{sinc}(\delta ux)} \qquad |x| < x_n/2 \qquad (11)$$

could be plotted so that h(x) would appear correctly. This modification also increases the intensity in the higher diffraction orders. However this effect is small due to the fact that high diffraction orders are strongly attenuated by the zeroes of the sinc function as illustrated in figure 2.

To achieve on-axis complex operations, several new techniques have been proposed. The most straightforward method is to plot the amplitude and phase of the transfer function separately. The phase plot is bleached and then superimposed on the amplitude plot. The registration problem can be simplified by plotting grating patterns outside the aperture. The Moire fringes produced by these patterns can be used for alignment. The sandwich of the two holograms is then permanently fixed together.

Other on-axis holograms including kinoform [1] and ROACH [2] have been used. Kinoforms are made by discarding the amplitude of the transform of a diffused object. Due to the lack of amplitude information, degradation occurs in the reconstruction. Besides using a random phase diffuser or deterministic diffuser [3] to even the spectrum, several iterative methods [2,4] can be used to achieve a flat spectrum. These methods are convergent in general, however, we do not expect to get perfect reconstruction in a finite

number of steps. Therefore instead of discarding the slow amplitude variation of the spectrum, we may implement it as a sandwich hologram. Since the slowly varying amplitude has a low dynamic range, the limited dynamic range of the film is no longer a problem. The ROACH stores amplitude and phase information on different layers of the color film. Theoretically it is a perfect filter without any registration problems. However it suffers from the low space-bandwidth product.

A two-kinoform system has been proposed [5] which sums up the responses due to two kinoforms. Therefore the light efficiency is improved while the correct amplitude information is preserved. Suppose the amplitude of the transfer function is normalized to unity, then each vector $Ae^{j\theta}$ in the unit circle can be decomposed into two vectors with lengths of 1/2 so that

$$Ae^{j\theta} = \frac{1}{2} e^{j(\theta+\psi)} + \frac{1}{2} e^{j(\theta-\psi)} \qquad (12)$$

$$\psi = \cos^{-1} A \qquad 0 \leq \psi \leq \frac{\pi}{2} \qquad (13)$$

This decompositon and the optical set up have been shown in figure 1 and figure 2 of reference 5 except the circle should be a unit circle.

Similar ideas have been discussed by Severcan [6] and Shmarev [7]. Severcan has discussed a two-channel phase only filtering system. In the two-channel phase only filtering system he used a grating to modulate the input signal into two higher frequency channels. Two filters corresponding to the two kinoforms used are placed in these channels. To extract the desired output, one requires another grating and low pass filter in another imaging

system. This system has diffraction efficiency of 66% and suffers from the requirement of very accurate alignment of the second grating. In the on-axis phase only filtering system, he plotted the phase only function side by side in a cell. Upon reconstruction the noise due to the parity term can be neglected for the zero-th diffraction order, however, this noise term dominates in the first diffraction order. To get useful results we have to either limit the object size or increase the sampling rate for the impulse response.

In a recent paper [7] Shmarev discussed a similar kinoform system using a balanced grating. However he decomposed the vector into two unit vectors, thereby reducing the light efficiency by 50%. The use of the grating further reduces the intensity of the output.

In general, there are many ways to decompose a vector in the unit circle into two constant vectors. A vector $Ae^{j\theta}$ in the unit circle can be written as

$$Ae^{j\theta} = \frac{n}{2}\left[e^{j(\theta+\psi)} + e^{j(\theta-\psi)}\right] \qquad 0 \leq A \leq 1 \tag{14}$$

where

$$\psi = \cos^{-1}\left[A/n\right] \qquad n \geq 1 \tag{15}$$

In the two-kinoform system n=1 and the results are eqs.(12) and (13). Shmarev uses n=2 and obtains

$$Ae^{j\theta} = e^{j(\theta+\psi)} + e^{j(\theta-\psi)} \tag{16}$$

where

$$\psi = \cos^{-1}A/2 \qquad \frac{\pi}{3} \leq \psi \leq \frac{\pi}{2} \tag{17}$$

In the set up shown in figure 2 of reference 5, if both channels have the same intensities, then every decomposition for a particular n would give the correct vector except a constant factor. For example n=2 results in a vector with half of the magnitude of the vector resulted from n=1. Light efficiency is therefore reduced for n>1. Furthermore, for larger n, ψ varies over a smaller range as illustrated by eq.(17). For computer generated holograms this means more quantization levels in order to obtain the same amount of accuracy. Therefore it is better to use the decomposition with n=1.

When n=1; the angle ψ is related to amplitude A by

$$A = \cos \psi \qquad 0 \leq \psi \leq \frac{\pi}{2} \tag{18}$$

Its derivative is given by

$$A' = \sin \psi \qquad 0 \leq \psi \leq \frac{\pi}{2} \tag{19}$$

For a computer generated hologram, ψ is quantized to a certain level $\hat{\psi}$ and the amplitude becomes $\hat{A}=\cos\hat{\psi}$ rather than A. The derivative of $\hat{A}$ is an increasing function with the minimum at $\psi=0$ and the maximum at $\psi=\pi/2$. Equivalently, the same amount of quantization error in ψ results in small error in A when A=1 ($\psi=0$) and larger error in A when A=0 ($\psi=\pi/2$). Therefore we can conclude that the phase coding techniques can be applied here to even the spectrum and therefore to reduce the quantization error.

In the above analysis the intensity of the impulse response is of main concern. We can thus take advantage of

allowing the phase of the impulse response to be arbitrary in order to obtain the desired transform distribution. One particular application is the continuous display of the discrete data [5]. The interpolators used for these purposes are separable in general. As a matter of fact, the separated functions have the same form. Therefore phase coding can be done in one dimension and the final result is the product of two one-dimensional functions. If Hirsch's method [1] is used for phase coding of an N x N picture with m iterations, the number of multiplications needed is

$$N_1 = 2m \cdot 2N^2 \log N^2 = 2N(2m \cdot 2N \log N) \qquad (20)$$

When the function is separable, the number of multiplications becomes

$$N_2 = 2m \cdot 2N \log N + N^2 \qquad (21)$$

If m is large, the second term of N^2 can be neglected. The computing time is then reduced by a factor of 2N. In fact the N^2 operations are simply additions for kinoforms.

Although the phase coding methods have significantly increased the light efficiency and reduced the quantization error, the phase associated with the impulse function tends to be irregular. This might result in 'built-in speckle' noise even if incoherent processing is used. Future work will include the study and implementation of other phase coding methods which would result in less noise.

References

1. L.B. Lesem, P.M. Hirsch and J.A. Jordan, Jr., "The Kinoform: A New Wavefront Reconstruction Device," IBM J. Res. Develop., Vol. 13, No. 2, March 1969, pp. 150-55.

2. J.R. Fienup, "Improved Synthesis and Computational Methods for Computer-Generated Holograms," Ph.D. Dissertation, EE Department, Stanford University, May 1975.

3. W.J. Dallas, "Deterministic Diffusers for Holography," Applied Optics, Vol. 12, No. 6, June 1973, p. 1179-1187.

4. N.C. Gallagher and B. Liu, "Method for Computing Kinoforms that Reduces Image Reconstruction Error," pplied Optics, Vol. 12, No. 10, October 1973, pp. 2328-2335.

5. A.A. Sawchuk and C.K. Hsueh, "Optical Filters for Image Reconstruction," University of Southern California, Image Processing Institute, USCIPI Report 740, March 1977, pp. 89-99.

6. M. Severcan, "Computer Generation of Coherent Optical Filters with High Light Efficiency and Large Dynamic Range," Ph.D. Dissertation, EE Department, Stanford University, December 1973.

7. E.K. Shmarev, "Kinoform in Filtering and Image-Synthesis Systems," Opt. Spectrosc., Vol. 41, No. 5, November 1976, pp. 535-536.

3.7 Optical Pseudocoloring of Spatial Frequency Information

(Supported by NSF Grant ENG-76-15318 and AFOSR under Contract AFOSR-77-3285)

Julian Bescos and Timothy C. Strand

Introduction

The use of pseudocolor to improve the information transfer from a display to the human observer is well established. Although its potential applications are numerous, pseudocolor has to date been used almost exclusively to encode image intensity variations. Another application that has been considered is to encode variations in spatial frequency content of an image by pseudocoloring in the Fourier transform domain [1]. Although this showed great promise for various applications such as texture analysis it was never actively pursued. The reason for this is the difficulty and expense of implementing Fourier domain pseudocoloring in a digital system. This is because it involves one forward two-dimensional Fourier transform, followed by the generation of three color filter planes each of which must finally be inverse Fourier transformed to generate the final color image. All of this is very time consuming on a digital computer. However this entire process can be very easily implemented optically. The basic idea is to have a spatial filtering setup which uses a multi-color filter in the Fourier plane to encode different spatial frequencies with different colors. Such a system can be developed using coherent, incoherent, or partially coherent illumination. Preliminary analysis indicates that this is one application where a partially coherent system can be used to particular advantage, combining the best features of coherent and incoherent systems without many of the disadvantages of the respective systems. In particular, a properly designed system using partially coherent illumination will result in a system that combines good chromatic differentiation of various spatial frequency components with excellent signal-to-noise ratio characteristics.

In the following section we present a theoretical analysis of a Fourier plane pseudocolor system with

partially coherent illumination and study the effects of going to the coherent or incoherent limits.

Theory

Although theoretical analyses of partially coherent filtering systems are straightforward in concept, the resultant calculations are usually intractable for any general treatment. However if enough restrictions are placed on the system, a tractable problem can be found. In this section we wish to outline the analysis of a system which is admittedly very restricted, but which gives results which are easily understandable and which offer at least a qualitative explanation of the experimental systems described here.

The system we wish to analyze is the one-dimensional system shown in figure 1. Two-dimensional analysis does not alter the results in essence but makes the evaluation quite cumbersome and less transparent. The source will be considered to be spatially incoherent with a temporal spectral distribution which can be approximated as three monofrequencies $\nu_1=c/\lambda_1$, $\nu_2=c/\lambda_2$ and $\nu_3=c/\lambda_3$ which we will refer to as color 1, color 2 and color 3 respectively. The source is assumed to have a uniform intensity over the region $-\mu_s \leq \mu \leq \mu_s$.

The object we will consider has the following amplitude transmittance (independent of illumination wavelength)

$$u_0(x) = A_0 + A_1 \cos(2\pi\mu_0 x) \qquad (1a)$$

where A_0, A_1, and μ_o are real constants.

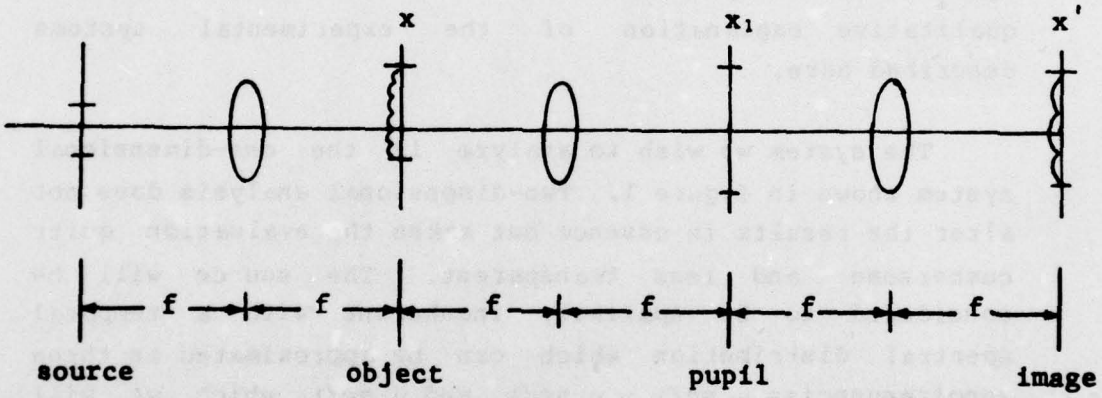


Figure 1. Optical system. An extended white light source is used to illuminate the object. A color filter is placed in the pupil plane.

Thus the intensity of the object is

$$I_0(x) = A_0^2 + \frac{A_1^2}{2} + 2A_0A_1 \cos(2\pi\mu_0 x) + \frac{A_1^2}{2} \cos(4\pi\mu_0 x) \quad (1b)$$

In the pupil plane we have basically three independent filters, one which transmits only in the spectral region of ν_1 (color 1), one which transmits only at ν_2 (color 2) and one which transmits only at ν_3 (color 3). These filters have the following binary transmittances for colors 1, 2 and 3 respectively (see figure 2):

$$f_1(\mu) = \text{rect}\left(\frac{\mu}{2\mu_c}\right) \quad (2a)$$

$$f_2(\mu) = \text{rect}\left(\frac{\mu}{4\mu_c}\right)\left(1-\text{rect}\left(\frac{\mu}{2\mu_c}\right)\right) \quad (2b)$$

$$f_3(\mu) = \text{rect}\left(\frac{\mu}{6\mu_c}\right)\left(1-\text{rect}\left(\frac{\mu}{4\mu_c}\right)\right) \quad (2c)$$

where

$$\text{rect}\left(\frac{\mu}{2\mu_c}\right) = \begin{cases} 1 & \text{if } -\mu_c \leq \mu \leq \mu_c \\ 0 & \text{otherwise} \end{cases}$$

The color 1 filter f_1 is a low pass filter with cutoff frequency μ_c where as the color 2 and color 3 filters are high-pass filters with the cutoff frequencies $2\mu_c$ and $3\mu_c$. Here μ represents a spatial frequency variable. The corresponding variable in the physical system is x_1 where

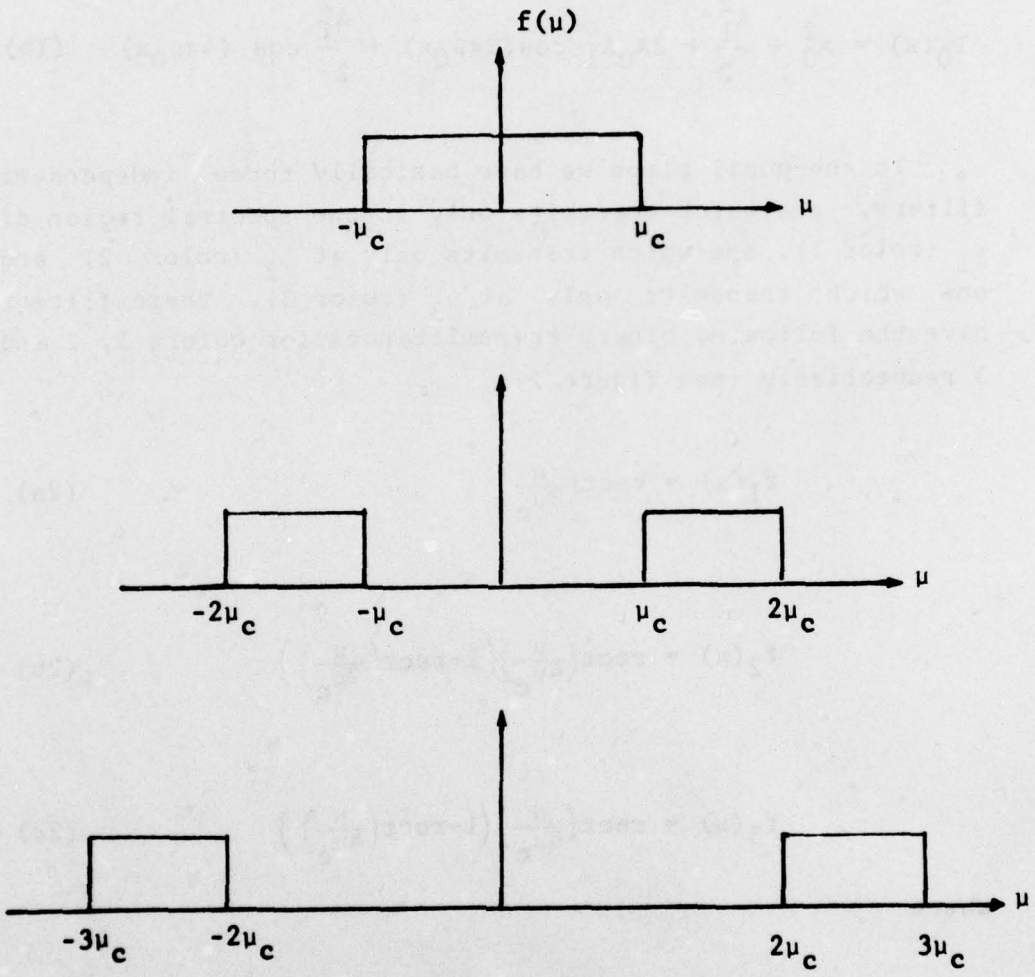


Figure 2. Three-color filter in the pupil plane. The coordinate μ represents spatial frequency in the image plane.

$x_1 = k\mu/\lambda j$, i.e. μ must be scaled by the appropriate illumination wavelength λ_j.

For a spatially incoherent and quasimonochromatic source, we can write the Fourier transform of the image intensity in color j as follows [2]:

$$I_j(\mu) = \int_{-\infty}^{\infty}\int f_j(\mu')f_j^*(\mu'-\mu)v(\mu'')v^*(\mu''-\mu)S(\mu'-\mu'')d\mu'd\mu'' \quad (3)$$

where * denotes complex conjugate, $S = k\mu_s \,\mathrm{rect}\left(\frac{\mu}{2\mu_s}\right)$ is the source intensity and v is the Fourier transform of the object

$$v(\mu) = c\int_{-\infty}^{\infty} u(x)\exp(-2\pi i\mu x)dx$$

From eq.(1a) we see that

$$v(\mu) = c\int_{-\infty}^{\infty}(A_0+A_1\cos 2\pi\mu_0 x)\exp(-2\pi i\mu x)dx$$

$$= c\left[A_0\delta(\mu)+\frac{A_1}{2}\left(\delta(\mu-\mu_0)+\delta(\mu+\mu_0)\right)\right] \quad (4)$$

Substituting eq.(4) into eq.(3) gives us an expression for the Fourier transform of the intensity of the color j. If one assumes the source intensity S and the filter function f_j are symmetric, the expression simplifies to the following (where constant factors have been dropped):

$$I_j(\mu) = \alpha_j(\mu_0)\delta(\mu)+\frac{\beta_j(\mu_0)}{2}\left[\delta(\mu-\mu_0)+\delta(\mu+\mu_0)\right]$$

$$+\frac{\gamma_j(\mu_0)}{2}\left[\delta(\mu-2\mu_0)+\delta(\mu+2\mu_0)\right] \quad (5)$$

The coefficients are defined as follows

$$\alpha_j(\mu_0) = |A_0|^2 \int_{-\infty}^{\infty} |f_j(\mu')|^2 S(\mu')\,d\mu' + \left(\frac{|A_1|^2}{2}\right) \int_{-\infty}^{\infty} |f_j(\mu')|^2 S(\mu_0-\mu')\,d\mu'$$

$$= |A_0|^2 F_{0j}(0) + \left(\frac{|A_1|^2}{2}\right) F_{0j}(\mu_0) \qquad (6a)$$

where $F_{0j}(\mu_o)$ is the convolution of the source with the filter intensity transmission $|f_j|^2$

$$\beta_j(\mu_0) = 2\mathrm{Re}\left[A_0 A_1^* \int f_j(\mu') S(\mu') f_j^*(\mu-\mu')\,d\mu'\right]$$

$$= 2\mathrm{Re}\left[A_0 A_1^* F_{1j}\right] \qquad (6b)$$

where F_{1j} is the filter function multiplied by the source function convolved with the (complex conjugated) filter function.

$$\gamma_j(\mu_0) = \left(\frac{|A_1|^2}{2}\right) \int_{-\infty}^{\infty} f_j(\mu') f_j^*(2\mu_0-\mu') S(\mu_0-\mu')\,d\mu'$$

$$= \left(\frac{|A_1|^2}{2}\right) F_{2j}(\mu_0) \qquad (6c)$$

From this we can write an expression for the intensity of the three color images as

$$I_j(x') = \int_{-\infty}^{\infty} \tilde{I}_j(\mu) \exp(-2\pi i \mu x')\,d\mu$$

$$= \alpha_j(\mu_0) + \beta_j(\mu_0)\cos(2\pi\mu_0 x') + \gamma_j(\mu_0)\cos(4\pi\mu_0 x') \qquad (7)$$

It should be remembered that the expressions for α_j, β_j, and γ_j are valid only for the particular input object we

have considered. Keeping this in mind we will refer to α_j, β_j, and γ_j as pseudo-transfer functions for each of the three components of eq.(7). In figure 3 we plot these pseudo-transfer functions for the filter described by eq.(2) and for three different source sizes. Figure 3a represents the limiting case as the source size approaches zero, i.e. the coherent illumination limit. In this case we see that there is always a certain bias of color 1 in the image, but otherwise the color of the image is going to be purely color 1, 2, or 3 depending upon the spatial frequency μ_o of the object. Thus we would expect a test target made up of sinusoidal gratings to have a highly saturated color image. The problem with such a system is that it is subject to all the noise problems of a coherent systems such as speckle and diffraction from dust, etc. Figure 3b shows the pseudo-transfer functions for a partially coherent system where the image of the source in the filter plane is half the size of the central filter (color 1). Although there is some overlap of the pseudo-transfer functions for different colors, in general, this system is seen to also yield fairly saturated colors. It is important to note that unlike the coherent system, this partially coherent system yields slightly different colors and varying degrees of modulation for two gratings at different but approximately equal frequencies. Finally, figure 3c shows the pseudo-transfer functions for the limiting case as μ_s becomes very large, i.e. spatially incoherent illumination. In this case β_j is the standard modulation transfer function (MTF) for the incoherent imaging system. α_j and γ_j follow directly from β_j. The important thing to note is that there is a constant bias term with equal contributions from all three colors (independent of μ_o). This implies poor saturation in the color image of a sinusoidal grating. Thus in many respects, the partially coherent system combines the good qualities of both the coherent and incoherent systems, specifically good

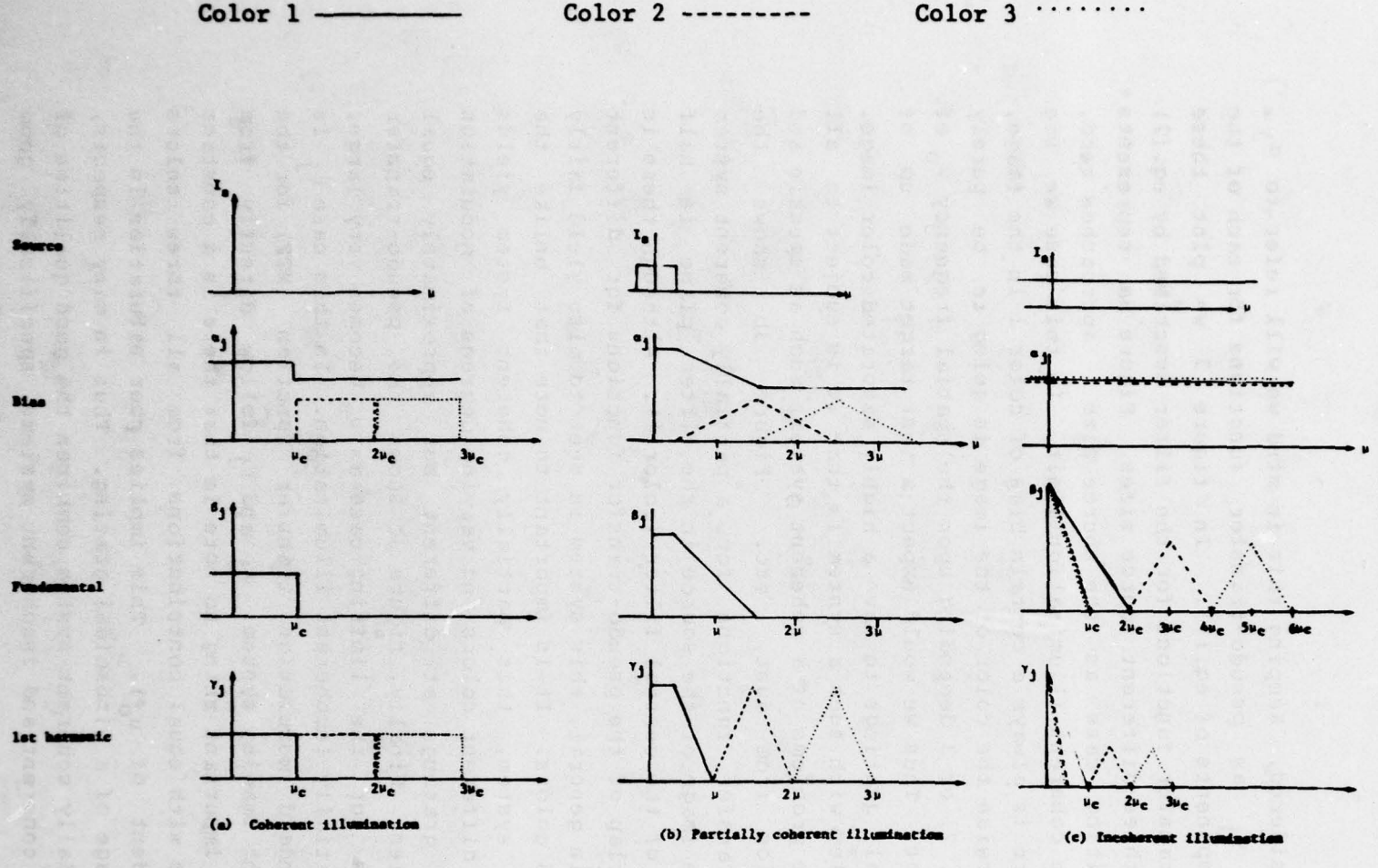


Figure 3. Pseudo-transfer functions for a cosinusoidal target $u(x) = A_0 + A_1 \cos(2\pi\mu x)$ imaged through the filter shown in figure 2. The top drawing indicates the spatial extent of the source. Successive drawings indicate the relative intensities of the bias component, the fundamental and the first harmonic respectively.

saturation with low noise and good color differentiation of adjacent spatial frequencies. It should be noted that, as can be seen from the plots of β_j, the fundamental frequency term tends to be eliminated in the coherent and partially coherent systems. This leads to false imaging effects for these systems which can be disturbing in certain applications.

The color saturation in the image is generally high as long as the image of the source is confined to one color in the filter plane. Thus a source intensity distribution described by eqs.(2a), (2b), or (2c) would give good saturation with those filters.

Conclusions

We have discussed an optical system for color encoding of the spatial frequency content of an image. A theoretical analysis is given for the general partially coherent illumination case. Preliminary experimental work has verified the theoretical results. This work will be reported on later.

It is felt that this system could have several significant applications in image analysis. Its application to texture analysis, for instance, seems particularly promising.

References

1. H.C. Andrews, A.B. Tescher and R.P. Kruger, IEEE Spectrum, 9 7, 20 (1972).

2. K. Dutta and J.W. Goodman, Journal Optical Society of America, 67, 796 (1977).

4. Smart Sensor Projects

This section represents the smart sensor phase of research funded during the past six months. Two circuits have been fabricated for CCD analog near focal plane processing to implement a variety of front end image processing functions. The first circuit implements the Sobel operator on an image. This represents a nonlinear spatially invariant processor. The second circuit is designed to implement both nonlinear spatially invariant as well as variant processes. Both circuits have been fabricated. The Sobel circuit has been tested with success, (see accompanying section) and the second circuit is soon to experience testing.

4.1 Charge-Coupled Technology for Image Understanding

Graham R. Nudd

Principal participants in this contract during this reporting period were Jerry Erickson, Paul Nygaard, Dale Sipma, Gary Thurmond, and Graham Nudd.

Three principal tasks have been undertaken since the last report, including:

the fabrication and testing of Test Circuit I;

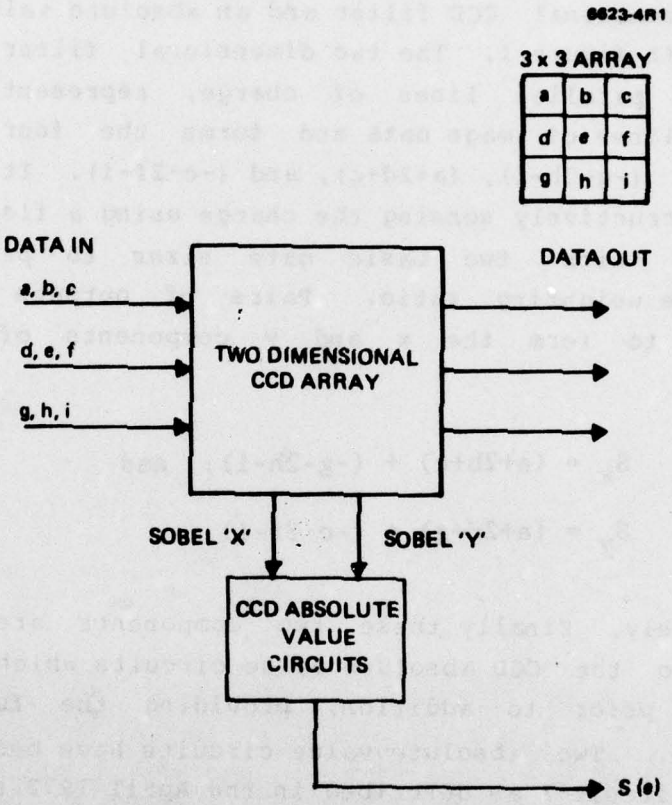
the detailed design, layout and processing of Test Circuit II; and

the design and construction of our test facilities.

We have made significant progress in each of these areas. The first test circuit is now operating on test images generated on the USC PDP-10 computer, and the photographs of the processed data are included here, together with the detailed test results. The second test circuit has now been designed and initial circuits are currently available. Details of this are also included. Finally, as the testing has progressed, we have made considerable modifications to our test facilities to provide more flexibility and also added a good deal of software to interface between the microprocessor, the PDP-10 and the microprocessor and the CCD circuits.

Test Circuit I

Details of the first test circuit, a three by three Sobel edge detection circuit, have been given in the semi-annual report dated April 1977. The basic algorithm operates on the nine picture elements shown in figure 1 and provides the output



$$S(e) = |(a + 2b + c) - (g + 2h + i)| + |(a + 2d + g) - (c + 2f + i)|$$

Figure 1. Schematic of the CCD sobel circuit.

$$S(e) = |(a+2b+c) - (g+2h+i)| + |(a+2d+g) - (c+2f+i)| \quad (1)$$

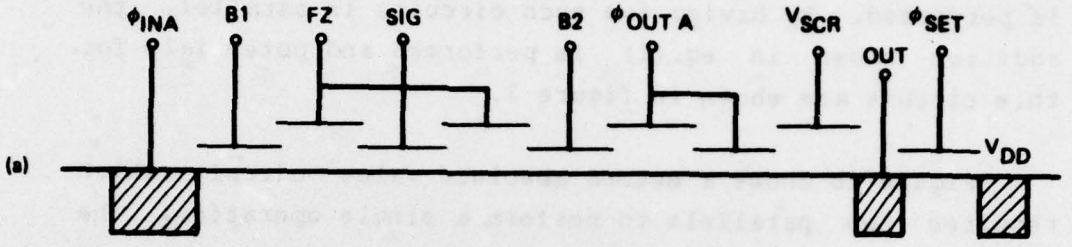
It performs this by the combination of two circuit elements, a two-dimensional CCD filter and an absolute value circuit as shown in figure 1. The two dimensional filter operates on three parallel lines of charge, representing three adjacent lines of image data and forms the four outputs: (a+2b+c), (-g-2h-i), (a+2d+g), and (-c-2f-i). It does this by nondestructively sensing the charge using a floating gate structure with two basic gate sizes to provide the two-to-one weighting ratio. Pairs of outputs are then combined to form the x and y components of the Sobel operator

$$S_x = (a+2b+c) + (-g-2h-i); \text{ and}$$

$$S_y = (a+2d+g) + (-c-2f-i) \quad (2)$$

respectively. Finally these two components are used as inputs to the CCD absolute value circuits which forms (S_x and S_y) prior to addition, providing the full Sobel operation. Two absolute value circuits have been included on Test Circuit I as described in the April 1977 semi-annual report. A schematic of the first circuit is shown in figure 2. It consists of a modified Tompsett input with two input gates B1 and SIG. A reference signal, equivalent to the zero level of the absolute value, is applied to B1 and the input signal is applied to SIG. When the input is more negative than the reference, B1, a potential well of capacity $(B1 - SIG)\cdot\varepsilon A/t_{ox}$, coulombs is formed where A is the gate area, t_{ox} is the oxide thickness and ε is the dielectric constant. Hence, when the diffusion is pulsed, this amount of charge collects under gates B2 and FZ. If,

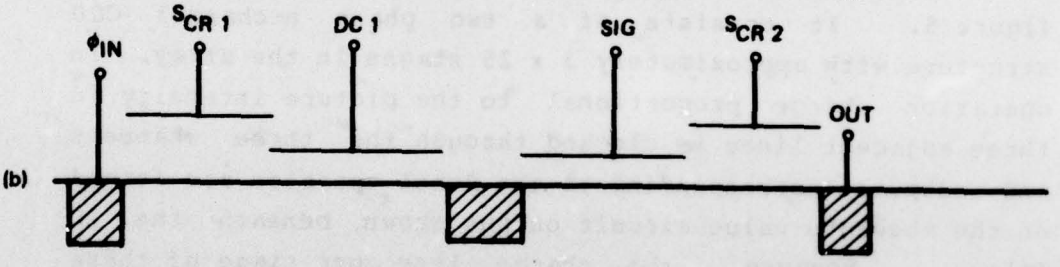
5931-10



SINGLE CHANNEL DEVICE

5931-11

V_{REF} = SIG WHEN ϕ_{IN} LOW



DUAL CHANNEL

Figure 2. CCD absolute value circuits

however, SIG is positive with respect to the reference, a charge $(SIG - B1)\cdot\varepsilon A/t_{ox}$ coulombs collects under electrodes FZ and SIG. Hence if the area of all the electrodes are equal, a charge equivalent to $|(SIG - B1)\cdot\varepsilon A/t_{ox}|$ is generated in either case and an absolute magnitude operation is performed. By having two such circuits in parallel, the addition shown in eq.(1) is performed and potentials for this circuit are shown in figure 3.

Figure 2b shows a second absolute value circuit which requires two parallels to perform a single operation. The potential profiles through the two channels are shown in figure 4. The technique relies on any voltage change in V_{sig} changing the potential profile under the SIG electrode and spilling charge = $\Delta V_{sig}\varepsilon A/t_{ox}$ through only one channel. For example, as shown in the figure, a positive change in V_{sig} will create a charge flow in the top channel only.

A photograph of the processed circuit is shown in figure 5. It consists of a two phase n-channel CCD structure with approximately 3 x 25 stages in the array. In operation charge proportional to the picture intensity in three adjacent lines is clocked through the three channels and outputs corresponding to the Sobel operator are formed on the absolute value circuit output shown beneath the 2D filter. Because the charge loss per stage of these circuits can be made very low, less than one part in 10^3, other processing functions have been included by adding circuits in series. In this case, a high-pass temporal filter, a two-dimensional Laplacian operator and an aperture corrector have been included.

During this period of the contract, the initial testing and performance analysis of the circuit has been performed. The full masking process consists of eight levels and two

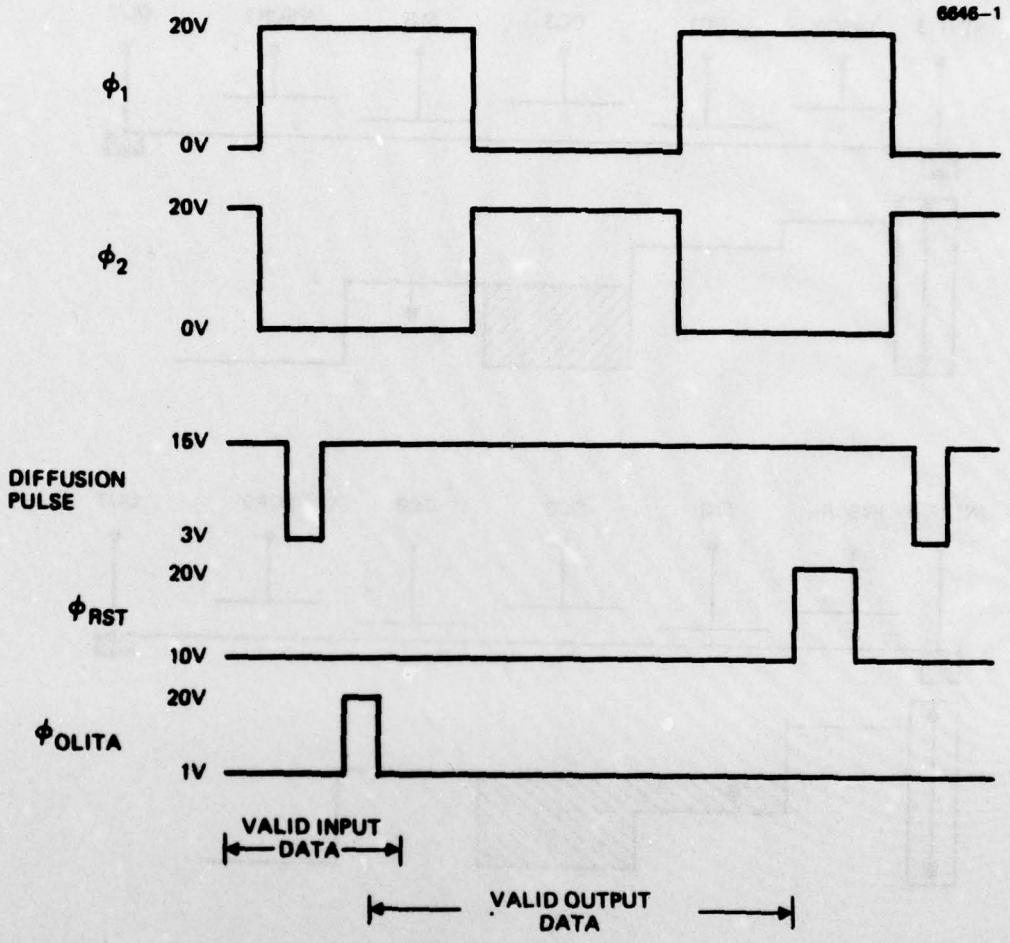


Figure 3. Timing diagram for absolute value circuit

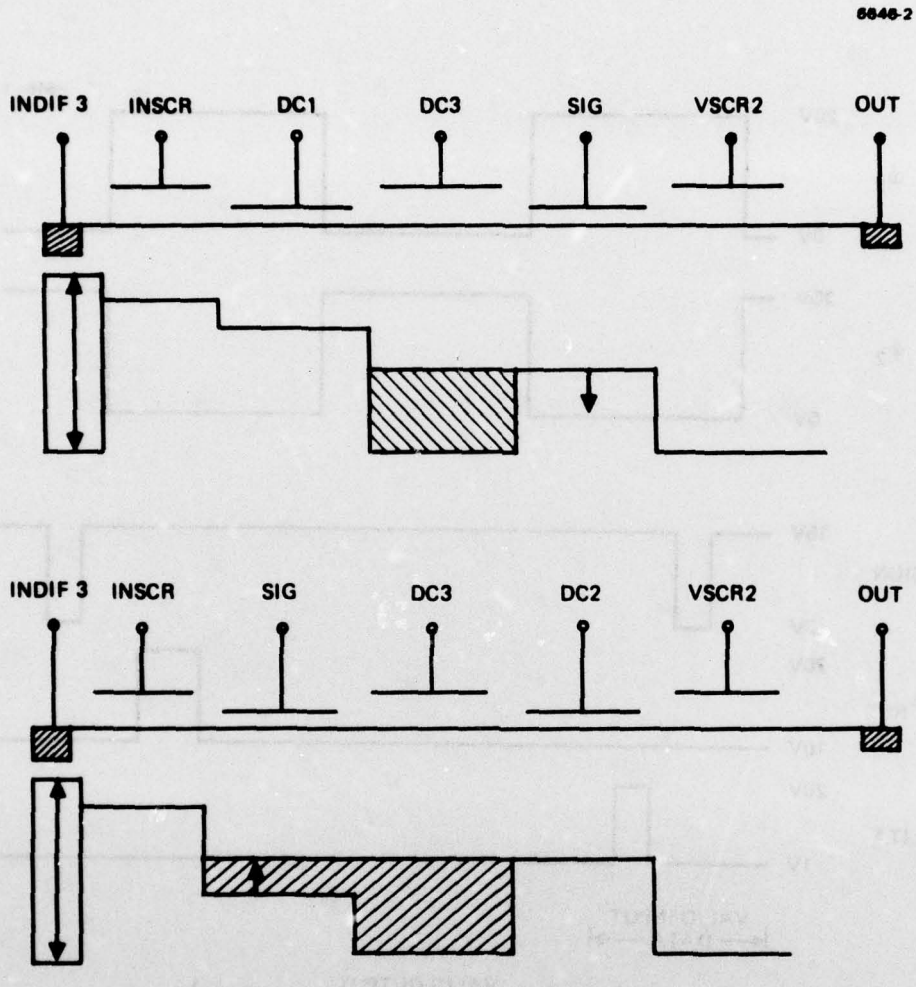


Figure 4. Operation of dual channel absolute value circuit.

6237-1

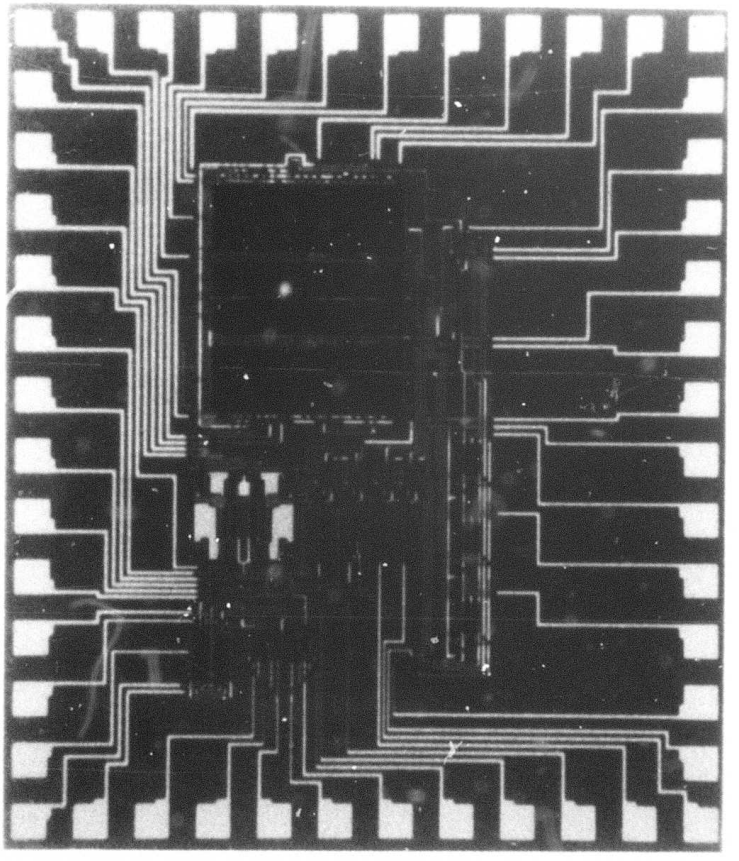


Figure 5. Photograph of Test Circuit I.

omissions have been encountered on the original masks which have now been corrected to provide working parts. The first mask error was an open circuit on one of the clock lines, ϕ_2, which prevented charge from being properly clocked through the full array. Initially attempts were made to wire bond directly to the clock line, but this proved to be very difficult and the results were unreliable, and a new mask was purchased. Fortunately this error occurred on the upper level and, since we had other partially processed wafers, we were able to provide corrected circuits in about six weeks. A later problem occurred on absolute value circuit 1 which prevented the x component of the Sobel from operating. This was an unconnected diffusion which eliminated the charge source to the Tompsett input. Here we were able to use a wire bond which provided satisfactory operation, and the full Sobel circuit is now operating. These circuit problems, which are to be expected in the initial runs of circuits of this complexity, have caused two to three months delay in our original plans. We have now however established the operation of the initial concepts, and a brief outline of the test procedures are included.

The circuit performance has been evaluated with the computer controlled test facilities described in a later section. In concept, image data, either from the USC-IPI data base or from specially prepared test patterns, is accessed via a standard 30 characters/second data link and stored in the random access memory (RAM) of the microcomputer controlling the test set-up. Special purpose CCD drivers and clocks have been built to drive the CCD circuits, request data from the RAM and return processed data back to the computer. Details concerning the timing and data flow are given in that section.

The initial test for the circuit was to ensure that charge was being clocked through the 2D filter from the Tompsett input. Accordingly, the voltage in each input gate for the three parallel arrays was varied as the outputs on Sobel 'x' (S_x) and Sobel 'y' (S_y) were monitored.

In this way, the effective operation of the top and bottom channels can be monitored using a dc voltage. It should be noted that the gate connection on the center channel requires a time varying waveform to test its operation. The output from the three floating gates on the top channel is shown in figure 6a. Note that here three gates with weightings 1/2, 1 and 1/2 are connected so the effective output has a weighting of 2. Also since the weightings are all positive, the output has the same polarity as the input signal. A schematic of the waveform is provided in figure 6b for reference. The linear range of this operation was from +5.5V to 7.3V. The charge storage for a single gate is given by

$$Q_s = C_{ox} V_s \qquad (3)$$

where $C_{ox} = \varepsilon A/t_{ox}$ is the oxide capacitance. In our case, 830 μm^2 and hence C_{ox} = 0.27 pF, resulting in a charge storage of 0.54×10^{-12} coulombs for a 2 volt gate swing. This is equivalent to 3×10^6 electrons. The noise associated with just filling this well is $Q_n = \sqrt{KT\, C_{ox}}$, which is equivalent to 690 electrons. In practice, other sources of noise exist, associated with the trapping states and thermal generation, etc., which might increase this number to approximately two or three thousand. Even so, a dynamic range at the input in excess of eight bits should be attainable.

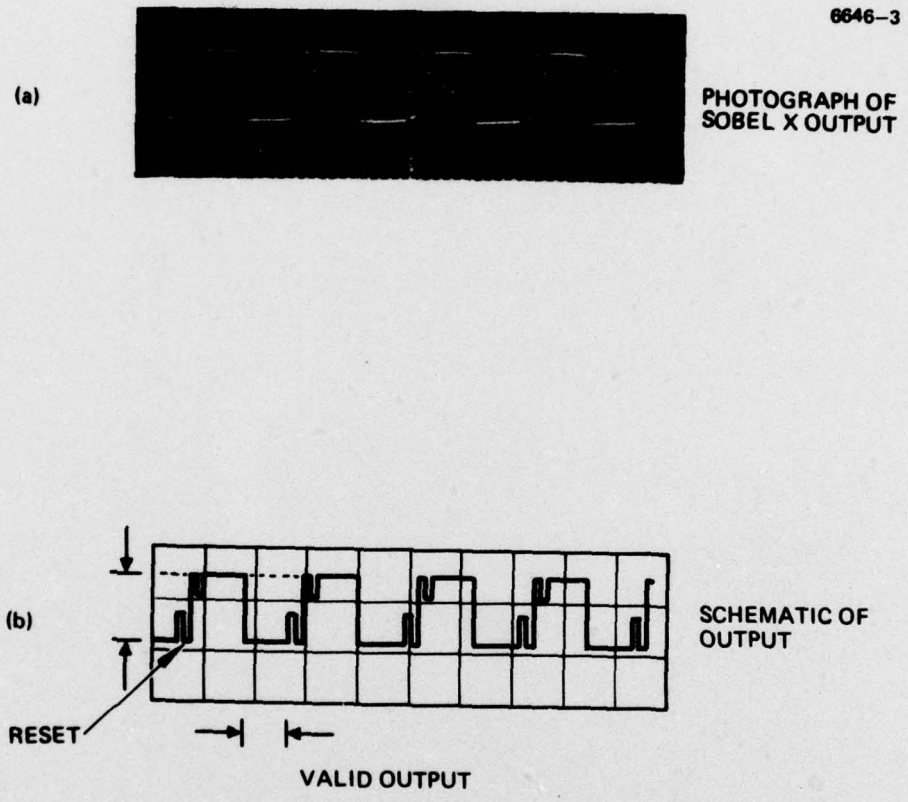


Figure 6. Output from floating gate array on top channel

Similarly, the operation of the lower channel is shown in figure 7a. Note the weighting on this channel corresponds to -1/2, -1, -1/2 or -2 and hence the polarity change. The linear range for this circuit is from 5.2V to 7.1V. Figures 6 and 7 indicate the effective clocking of charge through the channels. The desired weightings can only be analyzed using a time varying signal. Because of the speed limitation on the INTEL 8080A microprocessor, a data rate of approximately 5 kHz was used in these tests. This avoids the use of direct memory access, which will be incorporated in the next phase of the program.

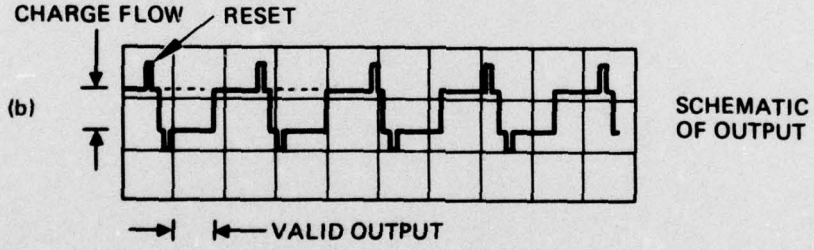
The appropriate weightings for the Sobel x and y components are illustrated in figure 8, together with the impulse response for each channel when measured at both the x and y outputs. The time dependent outputs for a single impulse are shown in figure 9, from which it can be seen that there are three functions of interest; Sobel "x", channel 1 and 3, and Sobel "y". The output waveforms for channels 1 and 3 at Sobel "x" and channel 2 at Sobel "y", in response to a single impulse of one picture element duration are shown in figure 10. It can be seen that the form of the weightings is current although the accuracy is not sufficient to provide operation equivalent to eight bits. We believe that the loss of accuracy is due to incomplete charge transfer through the full array and we are currently investigating techniques to improve this. In particular, one of the subsequent circuits in the array (the temporal high pass filter) has been designed to operate with a clocking waveform which is slightly shifted in phase from the prime clocks, and adjustment of these clocks should lead to better transfer. At present, we are operating all clocks from two phases.

6646-9

(a)



PHOTOGRAPH OF
SOBEL X OUTPUT



SCHEMATIC
OF OUTPUT

Figure 7. Output from floating gate array on bottom channel

6646–4

SOBEL "X"

1/8	1/4	1/8
-1/8	-1/4	-1/8

SOBEL "Y"

1/8		-1/8
1/4		-1/4
1/8		-1/8

SOBEL 'X'
IMPULSE RESPONSE

TOP CHANNEL:	1/8	1/4	1/8
MIDDLE CHANNEL:	0	0	0
BOTTOM CHANNEL:	-1/8	-1/4	-1/8

SOBEL 'Y'
IMPULSE RESPONSE

TOP CHANNEL:	1/8	0	-1/8
MIDDLE CHANNEL:	1/4	0	-1/4
BOTTOM CHANNEL:	1/8	0	-1/8

Figure 8. Weightings required for the Sobel algorithm

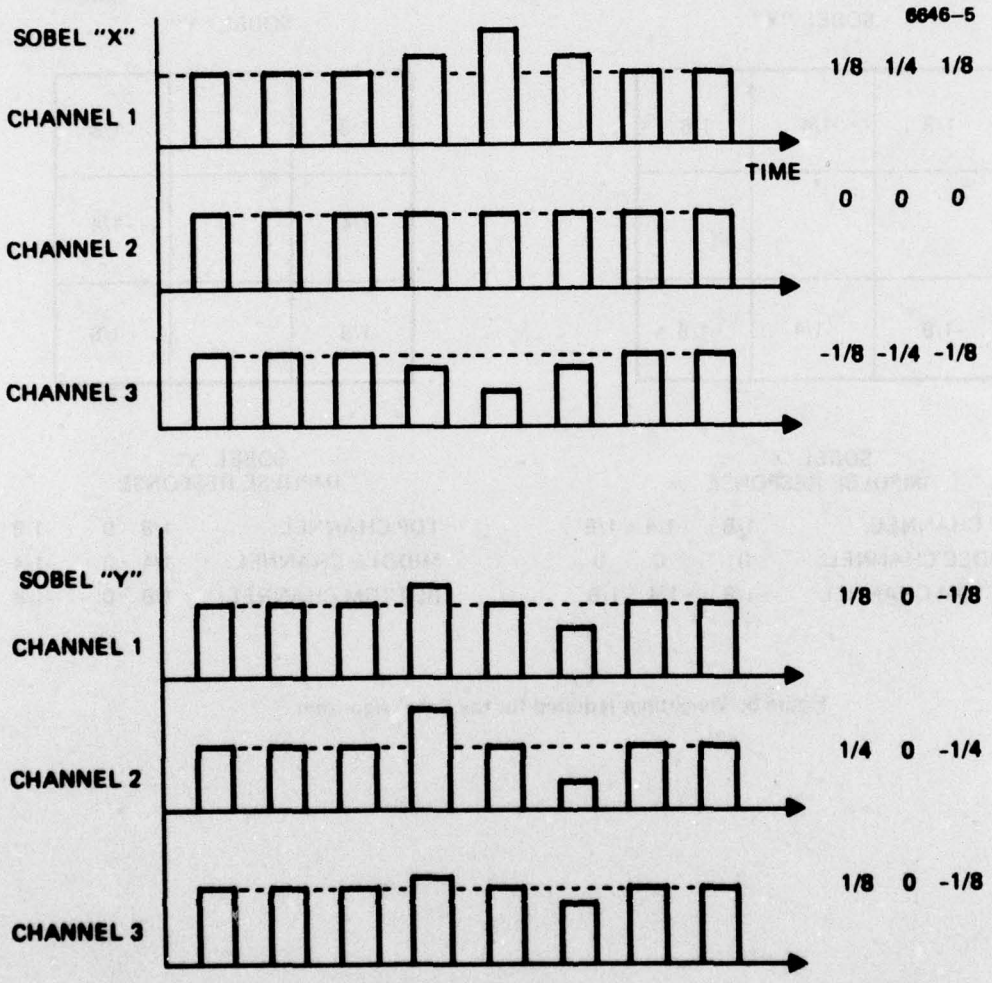


Figure 9. Impulse response of the three channels measured at Sobel "X" and Sobel "Y" outputs

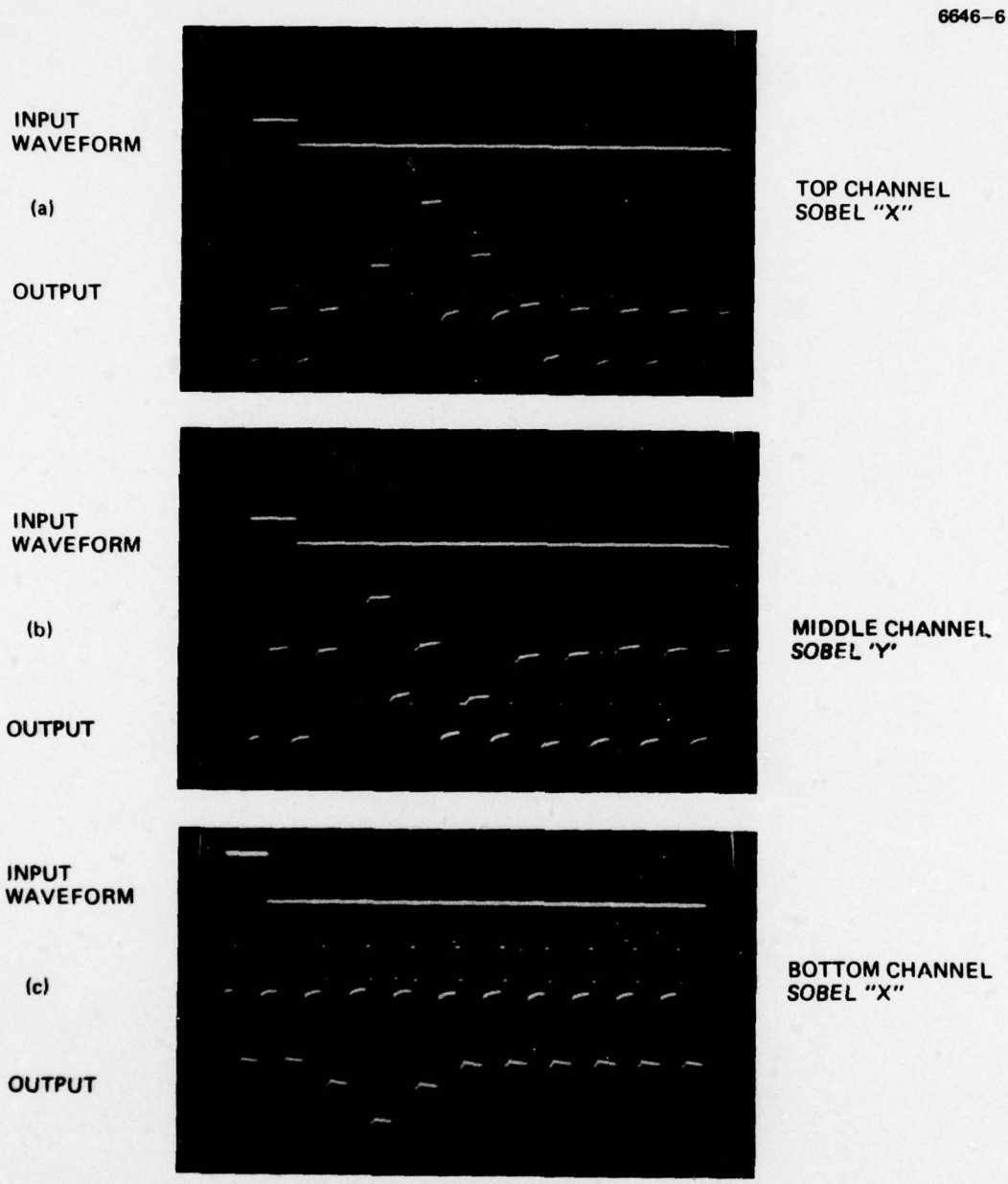


Figure 10. Measured impulse response of three channels

To provide the full Sobel output, the absolute magnitude of each component must be taken and the output summed, using the CCD absolute value circuit described above. An example of the operation of this circuit is given in figure 11. Here only the Sobel "y" component is used for illustration and the input to the array consists of an impulse of one pixel duration. This results in a positive going signal, representing the transition at the leveling edge of the pulse, zero output when the discontinuity is at the center of the array (i.e., at location "e") and a negative signal corresponding to the transition on the trailing edge, as shown in the upper trace. The positive and negative pulses in effect locate the edges of the discontinuity. The Sobel operator, however, requires that both edges have the same polarity. This corresponds to the output shown in the lower trace, where both signals are shown as negative going outputs. The polarity is merely convention and the negative outputs are a direct result of an n-channel device operating with electrons as the carriers. The form of this lower waveform corresponds precisely to the full Sobel algorithm for a single discontinuity. In terms of an optical image, the input would correspond to a vertical grid, each element having a width equal to one picture element, and spaced at the pulse repetition interval. The output is equivalent to a grid pattern of double lines representing the edges of the original pattern.

The two components of the Sobel operator have been tested using patterns with both vertical and horizontal symmetry. Specific examples of this are given in figures 12 and 13. In figure 12a, vertical grid of lines each of which is one pixel wide is shown as the unprocessed image. The processed image (figures 12b) consists of pairs of vertical lines spaced by a single picture element width and represent

6623-2 R1

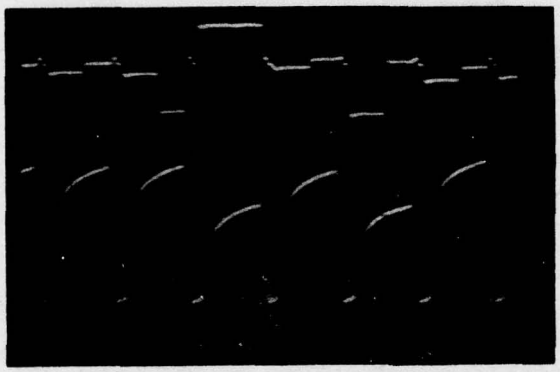
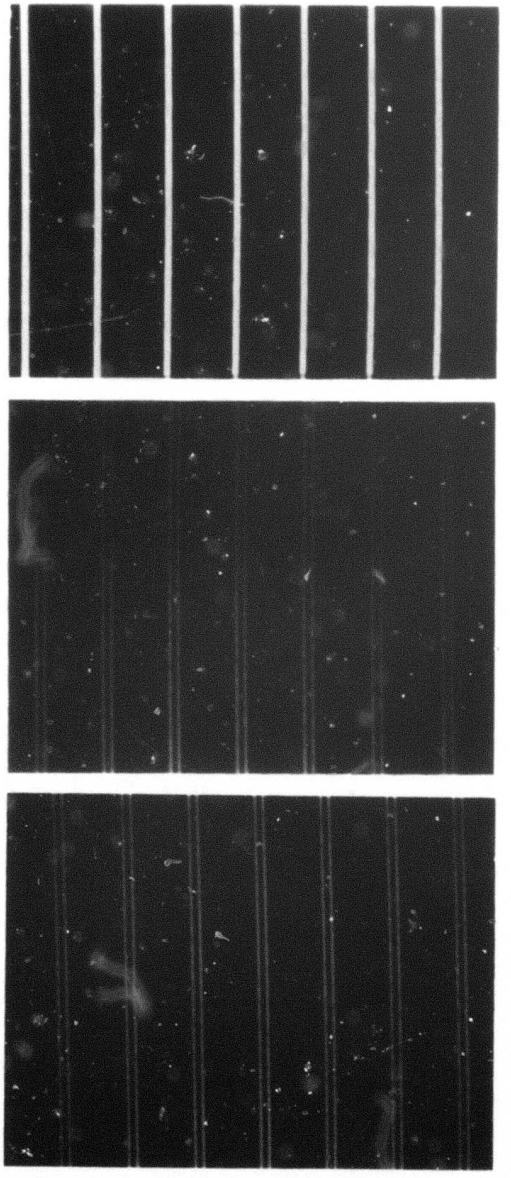


Figure 11. Operation of CCD absolute value circuit.



(a) ORIGINAL IMAGE

(b) IMAGE PROCESSED WITH CCD SOBEL

(c) COMPUTER PROCESSED IMAGE

Figure 12. Example of CCD Sobel operation on test pattern with vertical symmetry

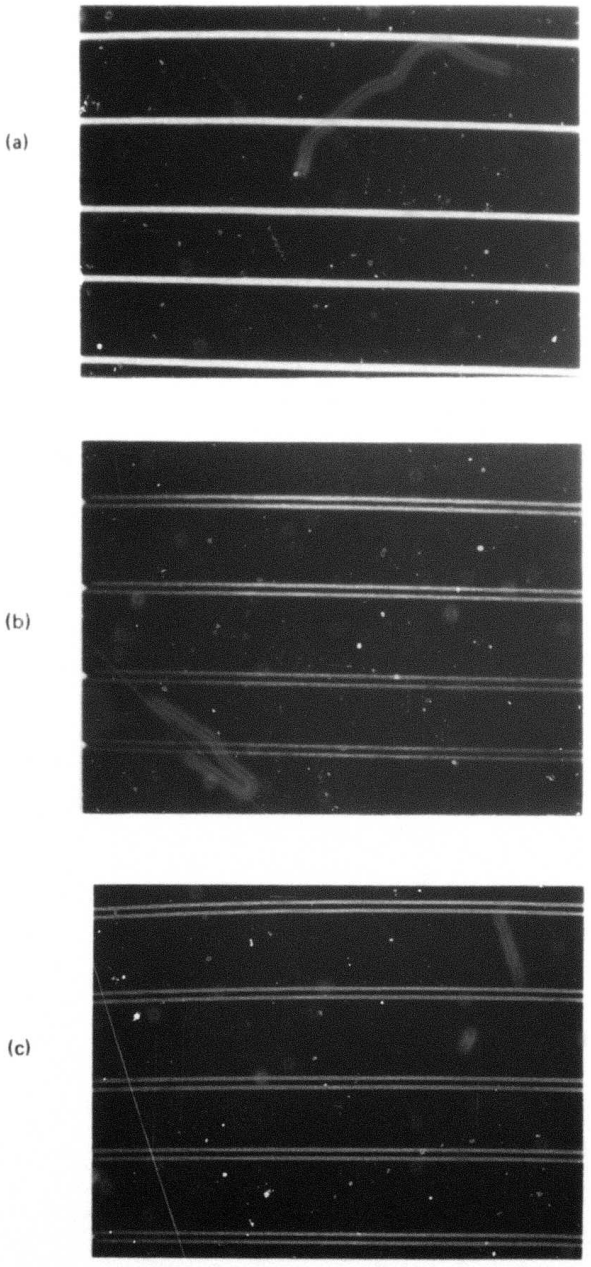


Figure 13. Example of CCD Sobel operation on test pattern with horizontal symmetry

the output from the CCD Sobel circuit operated at approximately 5 kHz. For comparison, the computer generated Sobel is shown in figure 12c. Figure 13 shows the corresponding results for a horizontal grid, thereby testing Sobel "x".

These simple patterns serve to verify the CCD Sobel operation and provide data for the circuit analysis. Our present work in this area is concentrated in three specific areas:

testing the circuit on a more extensive base;

increasing the dynamic range and speed of the present devices, and providing a more rapid data transfer between the microcomputer and the CCD circuits.

Test Circuit II

The aim of the second test circuit is to investigate the possibility of performing adaptive algorithms; in this case, based largely on the local mean. Again a kernel of three by three picture elements is used and the five algorithms included are:

Sobel Edge Detection

$$S(e) = 1/8\{|(a+2b+c) - (g+2h+i)| + |(a+2d+g) - (c+2f+i)|\}$$

Low Pass Filtering

$$F_M = 1/9(a + b + c + d + e + f + g + h + i)$$

Unsharp Masking

$$F_{USM} = (1-\alpha)S(i,j) + \alpha F_m$$

Adaptive Binariser

$$F_b = \begin{cases} 1 & F_M \leq e \\ 0 & F_M > e \end{cases}$$

Adaptive Stretching

$$F_a = \begin{cases} 2 \text{ Min } |e, r/2| & F_M \leq r/2 \\ 2 \text{ Max } |(e-r/2), 0| & F_M > r/2 \end{cases}$$

Because of the complexity of building all of these functions in a single circuit, a modular approach has been employed as shown in figure 14. Here each circuit is built as a separate unit on a single chip and the full operation can be obtained by using either wire bonds at the crystal surface or coaxial interconnects. The design rules for this circuit are given in Table 1 and, where appropriate, the performance of the circuit elements has been computer simulated using the "SPICE" simulation programs. The detailed design and layout of this circuit is now completed and the first lost has been processed. A photograph of the processed parts is shown in figure 15. Schematics of each circuit concept were included in the previous semi-annual report. We anticipate that preliminary testing and performance evaluation will be completed in the next two months. The circuits themselves will be bonded in forty-eight pin dual-in-line packages identical to that used for Test Circuit I and it should be possible to use the drivers and computer interfaces already developed. However to achieve the adaptive processing with all circuit elements, it will be necessary to initially provide external coaxial interconnects between several break-out boxes. Then, as testing proceeds, internal bonds at the chip surface will be employed to enable all functions to be obtained from a single circuit. We anticipate that further development of the peripheral circuitry such as clocks, rest and diffusion pulses necessary to operate this circuit will be relatively minor.

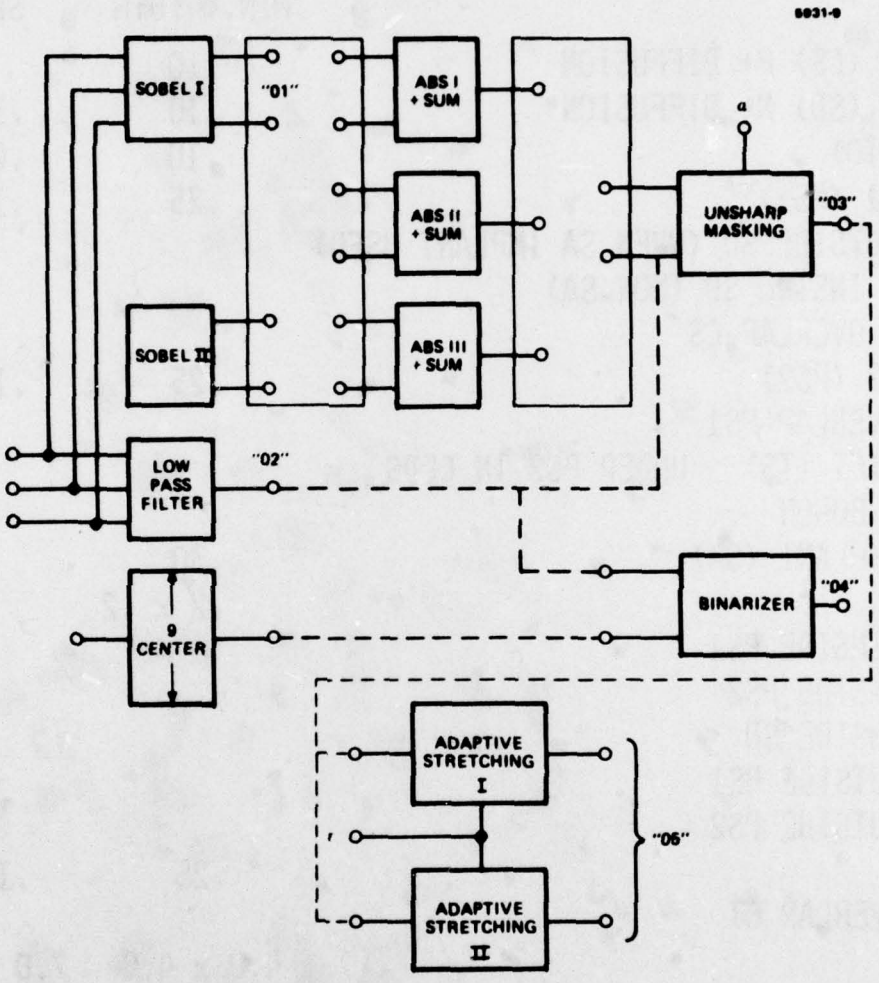


Figure 14. Block schematic of CCD processor.

STARTING MATERIAL 10 OHM-CM P-TYPE (100)

N-CHANNEL 1000 Å GATE OXIDE WITH SELF-ALIGN IMPLANT

LAYOUT RULES (MILS)

	MIN. WIDTH	SPACE
CHANNEL STOP (CS) P+ DIFFUSION	.10	
SOURCE/DRAIN (SD) N+ DIFFUSION	.10	.3 TO CS
GATE OXIDE (TO)	.10	.025 INSIDE SD
POLYSILICON 1 (PS1)	.25	.125 TO PS1
PS1 .1 OUTSIDE SD (WHEN SA IMPLANT USED)		
PS1 .075 INSIDE SD (NON SA)		
PS1 .125 OVERLAP CS		
POLYSILICON 2 (PS2)	.25	.125
PS2 .1 OVERLAP PS1		
THRESHOLD SHIFT (TS) UNDER PS2 IN CCDS		
IMPLANT, BORON		
SELF-ALIGN IMPLANT (SA)	.30	
CONTACT (CT)	.2 x .2	
CT .125 INSIDE PS1		
CT .15 INSIDE PS2		
CT .10 INSIDE SD		
CT .10 OUTSIDE PS1		
CT .15 OUTSIDE PS2		
METAL (AL)	.25	.15
AL .15 OVERLAP CT		
PAD	4.0 x 4.0	7.0 CENTER TO CENTE

DEVICES: L = .30, MIN W = .2

CCD: PS1 L = .35 GAP = .175
PS2 L = .375 W = 2.4

1.05 MILS/BIT

Table 1. Design Rules for Test Circuit II

6623-6

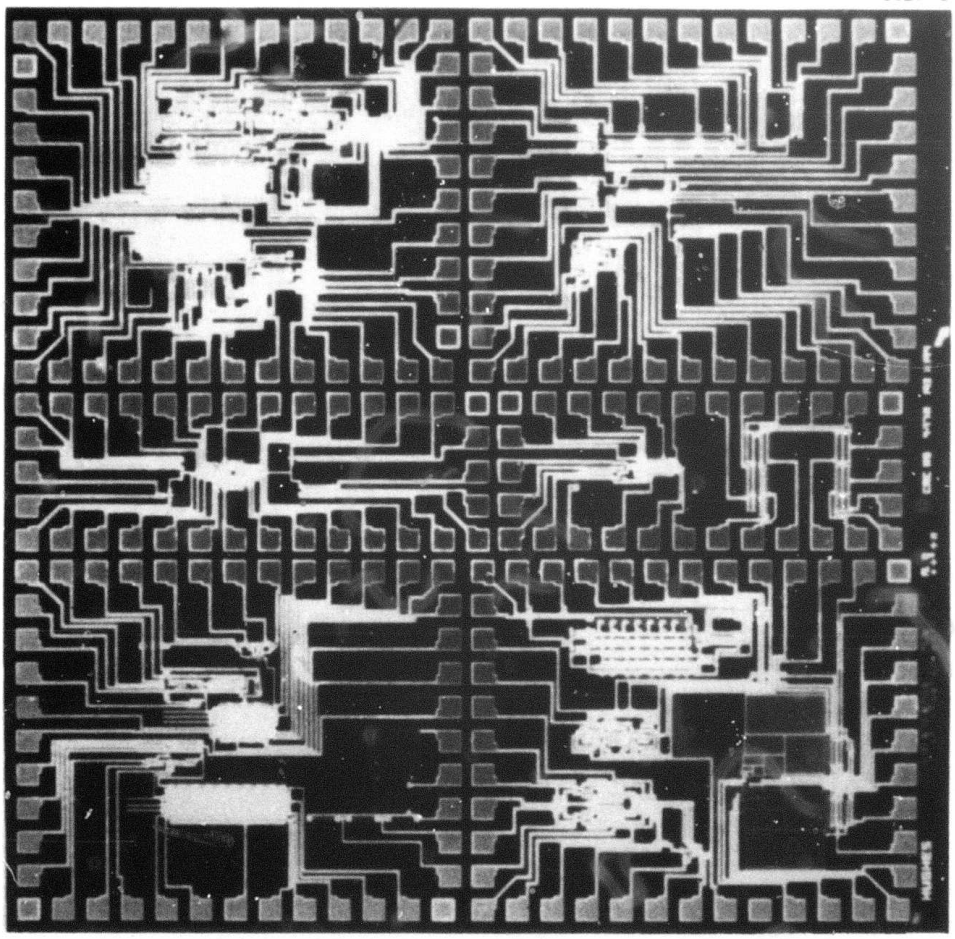


Figure 15. Photograph of test circuit II (size approximately 190 x 190 $mils^2$)

Development of Test Facilities

In addition to the development of Test Circuits I and II, we have made considerable progress with our test facilities, and most of the necessary equipment has been designed and built. The CCD test facility has been set up to accomplish two main objectives:

1. Provide the necessary signals, dc levels, and adjustments for functionally testing the operation of the CRC-111 and CRC-115 chips.
2. Demonstrate the operation of these chips on real images.

The following tasks have been done to accomplish these objectives:

- A box was built to hold the 48 pin chips and connect their leads to 48 bnc outlets.
- Two boxes were built to provide 12 independent, adjustable dc levels.
- A "CCD driver" box was built to provide the necessary periodic waveforms required for operating the chips. These waveforms include two clocks, two resets, two diffusion signals and clocking signals. These signals have adjustments for gain, offset and timing for optimizing the operation of the chips.
- The microprocessor (IMSAI 8080) was interfaced with the "CCD driver" box so that it can output three analog channels and read in one analog channel synchronously with the "CCD driver" box. Thus data in computer memory can be formatted into three parallel output channels, fed to the CCD chips, and the resulting operation (Sobel, median, etc.) can be read and stored back into computer memory.
- A display with 16 grey levels and a 128 x 128 pixel resolution was built for the microprocessor. Each

128 x128 x 4 bit picture occupies 8K of 8080 memory. Pictures and programs are stored permanently on cassette tape.

- A means of transmitting a picture from the USC time sharing system to the microprocessor was developed. To accomplish this, a version of the Scene Analysis System used by Hughes has been modified to contain a command which will encode a picture into ASCII and output it on the terminal. The microprocessor then intercepts the picture as it is transmitted.
- Several programs for the microprocessor test facilities have been written. These include the following:

A. CALIB: This program is used for calibrating the input and output analog channels of the microprocessor.

B. BLOCK: This program does the Sobel operation on a 3 x 3 block of data using the CRC-111 chip.

C. PACK2: This program intercepts a 128 x 128 x 4 bit picture from USC.

D. SOBEL: This program does a Sobel operation on a 128 x 128 bit picture using the CRC-111 chip. The results are displayed on the 128 x 128 x 4 bit display.

E. CSOBL: This program calculates exactly the Sobel operator on a 128 x 128 x 4 bit picture. The results are displayed on the 128 x 128 x 4 bit display and can be compared with the results of program SOBEL.

A block diagram of the test facilities are shown in figure 16. Also a number of test images have been generated to provide data of the circuits performance.

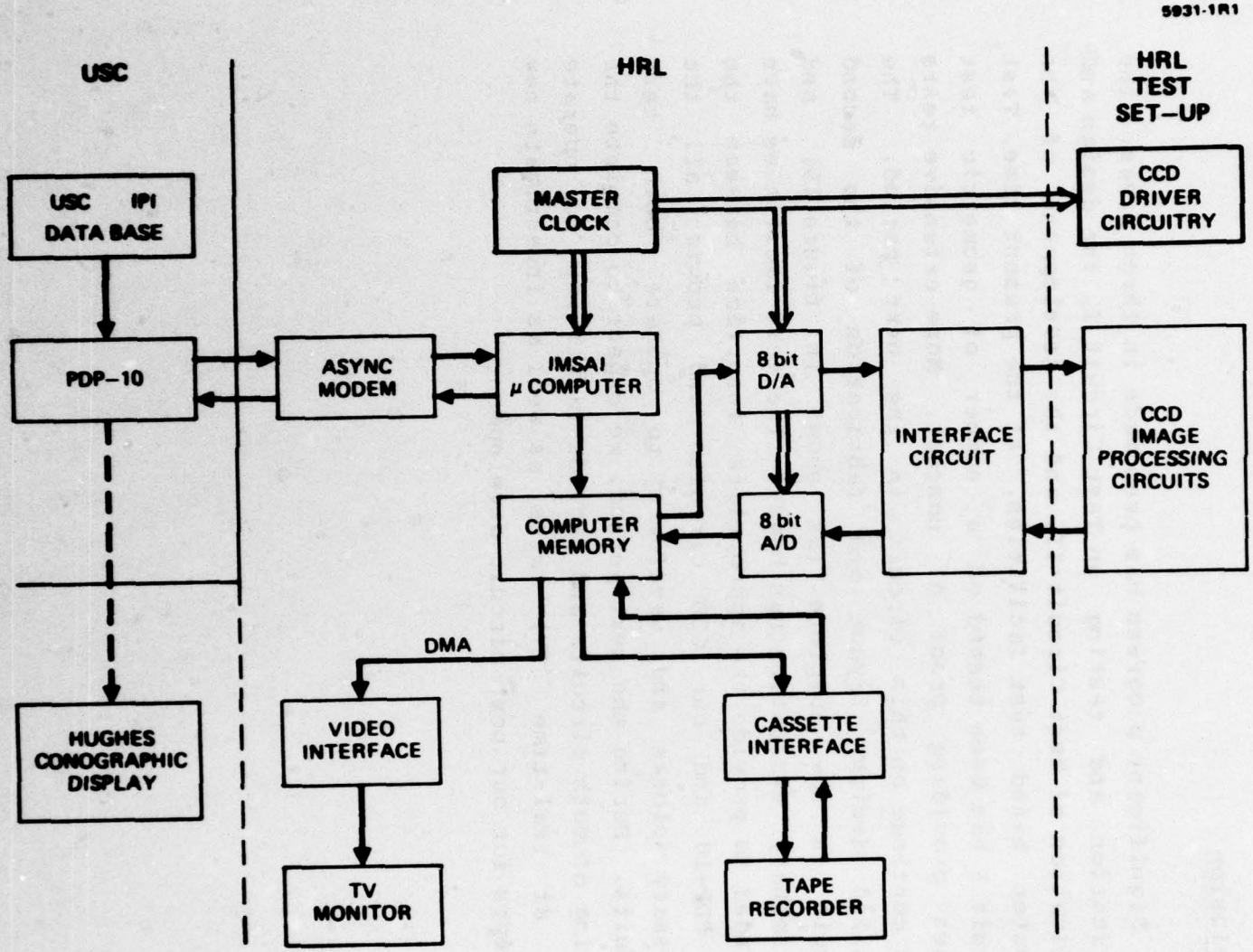


Figure 16. Schematic of test set–up.

Conclusion

Significant progress has been made in three areas: the fabrication and testing on Test Circuit I, the design and fabrication of Test Circuit II, and the development of the computer based test facilities. At the present time, Test Circuit I has been tested on a number of geometric test images providing proof of concept. More extensive tests will continue on this circuit in the next period. The detailed design, layout and fabrication of the second circuit is now complete (as shown in figure 15) and preliminary evaluation has begun. The test facilities have expanded to provide the appropriate interface between the USC PDP-10 and our CCD circuits and provide all the necessary clocks and waveforms to operate both test circuits. During the next period, we expect to complete the testing of both circuits and provide facilities to operate them at real-time (TV) rates as well as investigate new concepts for our next circuit development.

5. Recent Ph.D. Dissertations

This section includes those dissertations completed since the last reporting period. The ones listed here reflect results in two areas of our image understanding project. The first report describes a joint detection-estimation approach to boundary estimation. The methods of nonlinear estimation theory are applied to the detection of objects in extremely poor signal-to-noise ratio images with surprisingly good results. The second report describes the subject of image segmentation by clustering. This work utilizes the methodologies of pattern recognition, clustering, principal components analysis, feature selection and image pre-processing to automatically segment images into homogeneous regions.

Results of the above two research topics are described in detail in USCIPI Report 760 and USCIPI Report 750, respectively.

5.1 A Joint Detection-Estimation Approach to Boundary Estimation

Simon Lopez-Mora

The estimation of object boundaries based on noisy observations is considered in the context of joint detection and estimation.

The images are expressed as replacement processes and the boundaries modelled in terms of geometrical parameters associated with the object. The images studied have two textures, object and background, characterized by their first and second order statistics. A boundary processor consisting of optima estimator and detector is derived, for an appropriately chosen cost function. Differences between the cost function and resultant processor with other costs and estimator-detector pairs used previously in other applications is indicated. The optimal solution involves a nonlinear estimator and a detector with a variable threshold dependent on the estimator output.

Further, because of information restrictions imposed on the estimator that alleviate its computational requirements, a recursive, easily implementable algorithm, updating only the first two moments is derived, and subsequently used to evaluate the estimate as well as to perform detection.

Experimental results are illustrated. Of particular significance is the applicability of said processor under very low signal to noise ratio conditions.

5.2 Image Segmentation by Clustering

Guy B. Coleman

The segmentation of imagery into homogeneous regions using digital techniques has been a goal of researchers for the past several years. Pattern recognition approaches using mathematical models have achieved results which are only partially satisfactory. The large dimension of the pattern space and the quantity of data involved in the digital representation of images are in part responsible for the limited applicability of these approaches. Other shortcomings are related to the demands for data with which to train the classifier.

Approaches based on linguistic models have also been tried, again with results which are partially satisfactory. The most serious shortcomings are related to the performance of these approaches in the presence of noise, a phenomenon with which man has learned to function effectively.

This dissertation describes a procedure for segmenting imagery using digital techniques and is based on the mathematical model. The classifier does not require training prototypes, that is, it operates in an "unsupervised" mode. The procedure is general in that the features most useful for the particular image to be segmented are selected by the algorithm. The algorithm operates without any human interaction.

The features used are based on brightness and texture in regions centered on every picture element in the image. To perform an elementary pre-classification of local regions, a filter based on the mode of the local area histogram is proposed and used in segmenting images.

The basic procedure is a K-means clustering algorithm which converges to a local minimum in the average squared inter-cluster distance for a specified number of clusters. The algorithm iterates on the number of clusters, evaluating the clustering based on a parameter of clustering quality. The parameter proposed is a product of between and within cluster scatter measures, which achieves a maximum value that is postulated to represent an intrinsic number of clusters in the data.

It has been impossible in the past to compare different segmentations of the same image. A comparison measure based on the joint histogram of the two segmentations is proposed and examples of its use are presented.

It is within the state of the art to adapt the segmentation procedure described herein to operate in hardware at television rates. A functional diagram of such a system is presented, and estimates of the required capacities are given.

6. Recent Institute Personnel Publications

1. H.C. Andrews, "An Educational Digital Image Processing Facility," Computer, Vol. 10, No. 8, August 1977, pp. 48-53.

2. H.C. Andrews, "An Educational Digital Image Processing Facility," Photogrammetric Engineering and Remote Sensing, September 1977.

3. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Analysis and Synthesis," Applied Optics, Vol. 16, 1977, p.1009.

4. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Nonmonotonic Halftone Cells and Phase Halftones," Applied Optics, Vol. 16, 1977, p.1936.

5. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Effects of Input Medium and Precompensation," Applied Optics, Vol. 16, 1977.

6. H.S. Hou and H.C. Andrews, "Least Squares Image Restoration Using Spline Basis Functions," IEEE Transactions on Computers, September 1977.

7. F. Naderi and A.A. Sawchuk, "Estimation of Images Degraded by Film-Grain Noise," submitted to Applied Optics.

8. F. Naderi and A.A. Sawchuk, "Detection of Low-Contrast Images in Film-Grain Noise," submitted to Applied Optics.

9. R. Nevatia, "Characterization and Requirements of Computer Vision Systems," Advance Papers of the Workshop on Computer Vision Systems, University of Massachusetts,

Amherst, Massachusetts, June 1977.

10. R. Nevatia, "A Color Edge Detector and Its Use in Scene Segmentation," accepted for publication in IEEE Transactions on Systems, Man, and Cybernetics.

11. R. Nevatia and K. Price, "A Comparison of Some Segmentation Techniques," Proceedings of the Image Understanding Workshop, Minneapolis, April 1977, pp. 55-57.

12. R. Nevatia, "Segmentation Processes for Image Understanding," presented at the USA-Japan Vision Workshop, MIT, Cambridge, Massachusetts, August 1977.

13. M.J. Peyrovian and A.A. Sawchuk, "Restoration of Noisy Blurred Images by a Smoothing Spline Filter," accepted for publication in Applied Optics.

14. M.J. Peyrovian and A.A. Sawchuk, "Image Restoration by Spline Functions," accepted for publication in Applied Optics.

15. W.K. Pratt and F. Davarian, "Fast Computational Techniques for Pseudoinverse and Weiner Image Restoration," IEEE Transactions on Computers, Vol. C-26, No. 6, June 1977, pp. 571-580.

16. K. Price, "Change Detection and Analysis in Multi-Spectral Images," Proceedings IJCAI-77, Cambridge, Massachusetts, August 1977.

17. T.C. Strand, A.F. Turner, R.P. Kruger, H. Meyers, and W. Ahmed, "Automatic Detection of Alterations in Pulmonary Venous Pressure," Proceeding of the San Diego Biomedical Symposium, February 2-4, 1977.

18. T.C. Strand and A.F. Turner, "Hybrid Optical-Electronic Processing Applied to Chest Radiographs," Proceedings SPSE Tokyo Symposium '77 on Photo and Electronic Imaging, September 26-30, 1977.