(2)

ANNUAL REPORT

Joint Services Electronics Program
Contract DAAL-03-87-K-0059
January 1, 1989 - December 31, 1989

TWO-DIMENSIONAL SIGNAL PROCESSING AND STORAGE AND THEORY AND APPLICATIONS OF ELECTROMAGNETIC MEASUREMENTS

January 1, 1990

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF ELECTRICAL ENGINEERING
ATLANTA, GEORGIA 30332

ANNUAL REPORT

Joint Services Electronics Program
Contract DAAL-03-87-K-0059
January 1, 1989 - December 31, 1989

TWO-DIMENSIONAL SIGNAL PROCESSING AND STORAGE
AND
THEORY AND APPLICATIONS OF ELECTROMAGNETIC MEASUREMENTS

January 1, 1990

Georgia Institute of Technology
School of Electrical Engineering
Atlanta, Georgia 30332

Approved for public release.
Distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approval for public release. Distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Georgia Institute of Technology 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION U.S. Army Research Office
6c. ADDRESS (City, State and ZIP Code) School of Electrical Engineering Atlanta, Georgia 30332-0250	7b. ADDRESS (City, State and ZIP Code) Research Triangle Park, NC 27709
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Research Office 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAAL-03-87-K-0059
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification) Two-Dimensional Signal Processing & Storage & Theory & Application of EM Measurements	
12. PERSONAL AUTHOR(S) R. W. Schafer and D. T. Paris	
13a. TYPE OF REPORT Annual 13b. TIME COVERED FROM 1-1-89 TO 12-31-89	14. DATE OF REPORT (Yr., Mo., Day) 90-1-1 15. PAGE COUNT
16. SUPPLEMENTARY NOTATION The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.	
17. COSATI CODES FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Signal processing, holographic storage, digital filtering, iterative signal restoration, multiprocessors optical computing, electromagnetic measurements.

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

This is an annual report on research conducted under the auspices of the Joint Services Electronics Program. Specific topics covered are: digital signal processing, parallel processing architectures, two-dimensional optical storage and processing, hybrid optical/digital signal processing, electromagnetic measurements in the time domain, and automatic radiation measurements for near-field and far-field transformations.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL R. W. Schafer	22b. TELEPHONE NUMBER (Include Area Code) (404)894-2917 / 2920 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

Contents

1 Overview

This annual report covers the third year of research carried out under Contract DAAL-03-87-K00059. The research is part of the Joint Services Electronics Program and is administered by the U. S. Army Research Office. Research activities in this program are concentrated in the areas, (1) Two-Dimensional Signal Processing and (2) Theory and Application of Electromagnetic Measurements.

The research in two-dimensional signal processing is carried out in six work units. These work units are complementary, and research activities interact and reinforce one another in many ways. Research in Work Unit One, *Iterative Signal Restoration and Estimation*, and Work Unit Two, *Representation, Coding, and Analysis of Images*, is concerned with the theory, design, and implementation of multidimensional digital signal representations and digital signal processing algorithms and systems. Work Unit Number Three, *Multiprocessor Architectures for Digital Signal Processing*, focuses on hardware and software problems in the use of multiple processors for high-speed implementations of digital signal processing algorithms. The research in Work Unit Four, *Two-Dimensional Optical Information Processing*, is concerned with theoretical and experimental aspects of two-dimensional optical information processing. Work Unit Five, *Two-Dimensional Optical/Electronic Signal Processing*, is concerned with the theory, implementation, and application of hybrid optical/electronic methods for high throughput signal processing. Work Unit Eight, *Angular Spectrum Analysis for Non-Uniform Arrays*, seeks to develop new angle-of-arrival estimation techniques that are not restricted to uniform arrays of sensors.

The other two work units in the JSEP program are concerned with electromagnetic measurements. In Work Unit Six, *Electromagnetic Measurements in the Time- and Frequency-Domains*, research is concerned with the development of new methodology for making electromagnetic measurements directly in the time domain or over a wide bandwidth in the frequency domain. Work Unit Seven, *Automated Radiation Measurements for Near- and Far-Field Transformations*, is concerned with developing an understanding of the near-field and far-field coupling between antennas in the presence of scatterers.

The report begins with a summary of the most significant accomplishments (in the judgment of the lab directors) during the period January 1, 1989 to December 31, 1989. Following this are reports on the individual work units. These reports list personnel supported and discuss in general terms the research that was carried out during the reporting period. Also included in each work unit report is a complete list of publications on the research during this period. These publications are reprinted in the Annual Report Appendix, which is available in micro-fiche form as a separate document.

2 Significant Research Accomplishments

The following accomplishments are, in the judgment of the laboratory directors, of particular significance and potential and are therefore worthy of special mention.

2.1 Optical Interconnection of VLSI Systems

The development of optically interconnected VLSI and wafer scale systems has as a prerequisite the existence of a light modulator technology compatible with logic circuit technology that will enable the unconstrained placement of a large number of logic-driven light sources on a die or wafer. Ferroelectric liquid crystals (FLC's) possess low switching energies and driving voltages in the surface-stabilized cell configuration and are therefore well-matched to the output characteristics of silicon CMOS logic gates. Rapid development of faster materials is under way, with switching times of 3 microseconds at 15V/micron available now. Within the last six months a procedure that incorporates fabrication of most components of a reflective surface-stabilized FLC light modulator cell into an unmodified bulk CMOS process flow has been successfully prototyped.

2.2 Signal Estimation for Non-Uniform Arrays

The Steiglitz-McBride method of iterative exponential signal modeling has been extended to the case of missing measurements or non-uniform arrays. This situation is applicable to array processing in the case where a linear array has been thinned by removing sensors, or where some sensors in the linear array have failed. The Steiglitz-McBride method (also known as Iterative Pre-filtering) was originally formulated as a filtering problem and, hence, demands a regularly spaced grid with no missing measurements. In the case of missing measurements, an interpolation step must be added to the iterative algorithm. This step can be inserted by using the exponential model that is being built during the iteration. In the course of this work, it has also been discovered that the Steiglitz-McBride iteration is *exactly* equivalent to several existing algorithms, including the Iterative Quadratic Maximum Likelihood (IQML) algorithm.

2.3 Pulse Excited Antennas

Radars that use base band pulses, such as ground penetrating radars, require antennas that can radiate and receive temporally short, wide bandwidth pulses. These antennas have traditionally been analyzed using approximate methods, such as transmission line models or assumed current/aperture distributions. The antennas are generally electrically large (many wavelengths long) at the highest frequencies contained in the pulse. Thus, conventional frequency-domain numerical techniques are not efficient for analyzing these antennas. However, the finite-difference time-domain method (FD-TD) is a practical technique for solving three-dimensional structures that can be used to study

antennas for radiating wide bandwidth pulses. The cylindrical monopole and the conical monopole antennas are rotationally symmetric antennas and therefore require only a two-dimensional analysis. The FD-TD method was used to determine the response of these antennas to a Gaussian pulse, and the results were compared with experimental measurements to verify that the FD-TD method accurately predicts the performance of these simple antennas, and suggests that a full three-dimensional analysis of more complicated antennas, like the TEM horn, will be feasible.

Work Unit One

TITLE: Iterative Signal Restoration and Estimation

SENIOR PRINCIPAL INVESTIGATORS:

Russell M. Mersereau, Regents' Professor and Rockwell Fellow
Monson H. Hayes, Associate Professor

SCIENTIFIC PERSONNEL:

C. Auyeung, (Ph.D. received, December 1988)
S. J. Reeves, (Ph.D. Candidate)
W. S. Kim (Ph.D. Candidate)

SCIENTIFIC OBJECTIVE:

The objective of this research is to study and develop a broad class of iterative signal restoration techniques to remove linear and non-linear degradations from signals through a knowledge of the distortion operator and the class of allowable signal. This work brings together ideas from signal processing, functional analysis, optimization theory, statistical filtering, and signal modeling.

RESEARCH ACCOMPLISHMENTS:

- *Iterative Deblurring of Images*

In many applications our ability to extract information from an image is limited because of the way it was recorded. In addition to the additive noise that may be introduced by film grain or by an electronic front end, there is blur introduced by atmospheric turbulence, out-of-focus optics, or relative motion between the camera and the object. For several years doctoral students in the program have examined means for removing the effects of blur and noise from images using iterative algorithms which can incorporate knowledge of the blur, statistics of the noise, and prior knowledge about the class of allowable images. Originally it was assumed that the blurring operator was known exactly and our attention centered on means for incorporating knowledge about the class of allowable images into the restoration. During the past year a tutorial paper describing this early work was accepted for publication.

During the period of the present JSEP contract two students looked at further extensions of this work. One began to examine the case where the blurring operator is not known, but must be determined from the blurred image itself. This is the situation in most practical applications. This work is described below. The other student considered generalizations of this work to a number of related signal restoration problems. His work is described in the next section.

Let $f(x,y)$ denote the image to be restored from a linearly blurred and noisy copy, $g(x,y)$. Then f and g are related by

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y;s,t)f(s,t)dsdt + n(x,y).$$

The function $h(x,y;s,t)$ implements the blurring operation, $n(x,y)$ is the measurement noise, and the goal is to estimate f from g. Even when the blurring operator is known, this is an ill-posed problem, since small changes in g can cause large perturbations in the restoration, $\hat{f}$. Restrictions on the class of allowable signals f and knowledge of the statistics of the noise can both be used to help stabilize the estimate. An approach, that we have found to be particularly helpful, inverts this system using an iteration. In its simplest form this can be written as

1. Set $\hat{f}_0 = \lambda g$.

2. For $k = 1, 2, \ldots$

$$\begin{aligned} f_k &= \hat{f}_{k-1} + \lambda(g - H\hat{f}_{k-1}) \\ \hat{f}_k &= P_s f_k. \end{aligned}$$

H is the blurring operator defined by the integral above without the noise term and P_s is a *projection onto* the convex space of allowable signals.

The problem becomes even more difficult when the blurring operator is not known. One approach in this case is to produce a 'dual iteration' in which an estimate of the blur is used to produce an updated estimate of the signal that is then used to produce an updated estimate of the blur, etc. This can be done using a generalization of the above iteration or it can be done using an iteration which will converge more quickly, known as the conjugate gradient algorithm. We developed a dual-iteration technique for deconvolving two finite-extent signals using both of these algorithms. This approach worked quite well when the extent of the blur was small (five pixels or less). For larger blurs, the procedure converged extremely slowly.

The assumption that the original image f is of finite extent is an unrealistic one. A better assumption is that the image is of infinite extent, but that it is observed over a limited observation window; this leads to a more appropriate treatment of edge effects. Our algorithms were modified to treat this case, but to date the results have been unsatisfactory. The finite-extent property represents a powerful constraint on the deconvolution process and additional constraints on the image appear to be necessary when this property does not hold.

Satisfactory solutions to ill-posed problems are often obtained by means of regularization. Regularization introduces a bias toward a stable solution. This bias represents a tradeoff between fidelity to the original data and the stability of the solution. In the context of iterative image deblurring, this tradeoff can be expressed as the weighted sum of two expressions to be minimized—one representing the squared error on the data and

the other representing a stabilizing functional. This weighting is controlled by a scalar parameter. Once the parameter value is chosen, the deblurring can be performed using an iteration similar to the one outlined above. To date no automatic procedures have been developed for determining the optimal values of this parameter; instead trial and error methods are used. We have developed a method, based on the method of cross-validation from regression analysis, which can determine a nearly optimal choice for this regularization parameter based on measurements of the noisy blurred image when the blurring function is known. The full cross-validation procedure is computationally prohibitive, but we have developed a simplified version that seems to work satisfactorily. Cross-validation does not require any *a priori* knowledge of either the amount of noise or its distribution. A journal and a conference paper have been prepared that describe these early results.

We have also conducted some promising experiments which make use of cross-validation for problems where the functional form of the blurring operator is known but its parameters are not. In these problems cross-validation is a robust means for determining the values of these parameters. These algorithms are appropriate for removing out-of-focus blurs when the degree of defocus is unknown or for removing motion blurs when the extent of the motion is unknown.

- *Constrained Image Restoration*

The blur removal problem discussed in the previous section is a special case of the more general problem of estimating (restoring) a signal from a number of linear functional measurements. In abstract terms the problem can be cast as that of estimating a signal $f(r)$ from the measurements

$$g_k = \int_K f(r) w_k(r) dr + n_k \qquad k = 1, 2, \ldots, N.$$

The kernels $w_k(r)$ represent the effect of the blurring operator. These are assumed to be known. Some prior knowledge concerning the class of allowable signals $f(r)$ and the measurement noise n_k may also be available. We can find an optimal estimate of $f(r)$ by minimizing a functional of the form

$$\int_K h(f(r)) dr$$

where $h(x)$ is a strictly convex, continuous function of a real variable x.

The motivation for considering the signal restoration problem at this level of generalization is that a number of other important, well studied problems can be expressed the same way. These other problems include the reconstruction of a multidimensional signal from its projections and the one- and two-dimensional power spectrum estimation problems. By varying the choice of the function $h(x)$, we can produce an optimal estimate based on a least-squares, maximum entropy, or minimum cross entropy criterion. It was felt that by studying these problems under a common framework we might

gain a deeper understanding of all of them. Solution methods that worked well for one of these specific problems might be extendable to others. For example, the efficient solution methods of maximum entropy power spectrum estimation might be applicable to the deconvolution problem. This research was partially successful when measured against this goal. We were able to produce a unified mathematical basis for all of these problems and to transfer some results from one problem domain to another, but, in the most general cases, the algorithms were too numerically intensive to be useful on realistic problems with presently available computers.

The thesis by AuYeung produced a number of useful results. It was shown that the general constrained signal restoration problem could be solved using a dual method. With this approach, if the observations are consistent with each other and with the constraints on the signal and noise, the optimal restoration must be of a certain mathematical form with unknown parameters, which can be found by solving a nonlinear optimization problem. The number of unknown parameters is equal to the number of observations; this can be extremely large for must multidimensional problems. Thus, even though the dual approach considerably simplifies the solution of the restoration problem, it still cannot produce a procedure that can be implemented with a reasonable amount of computation, except in certain special cases. Auyeung's thesis was completed at the end of 1988. During the current year, however, three papers were written to describe his results.

- *The Phase Retrieval Problem* [1,2]

The phase retrieval problem is concerned with the recovery of a signal from the intensity (magnitude) of its Fourier transform and is a long-standing problem with applications in many different areas [3,4]. The phase retrieval problem may be cast in the framework of recovering a signal x from an observation $y = Dx$ where D is a nonlinear operator. Although it has been shown that, except for some trivial ambiguities, almost all two-dimensional signals are uniquely defined by their Fourier intensity [5], so far the success in reconstructing an arbitrary two-dimensional signal from its Fourier intensity has been limited [3,4,5,8]. Therefore, we have recently begun to focus on the problem of reconstructing a two-dimensional signal (image) from more than one Fourier intensity function (magnitude-only image). This work builts upon earlier JSEP research by Katsaggelos who studied the problem of reconstructing an image from more than one linearly distorted observation [7]. Here, however, we have been investigating the problem of reconstructing an image from a set of nonlinearly distorted images, i.e., two or more magnitude-only images. The specific problems that we have been looking at is described below.

It is well known that an image can be easily recovered from its Fourier intensity if the original image contains an appropriately placed latent reference point [4,5,8,9]. We have recently modified this latent reference point condition by considering the problem of image recovery given the Fourier intensity of the image and the Fourier intensity of

the image after a known reference point has been added to it. Specifically, given the Fourier intensity of the image $x(m,n)$ and the Fourier intensity of the modified image,

$$y(m,n) = x(m,n) + A\delta(m - m_0, n - n_0) \tag{1}$$

we have shown that with very modest constraints, the image may be easily and uniquely recovered by simply solving a set of linear equations [1,2]. Unlike the latent reference point condition, however, the superimposed point source is allowed to be located almost anywhere within the support of the image. The latent reference point condition of [8,9] and the off-axis holography condition are special cases of this work.

Generalizing (1) we have considered replacing the single point source $A\delta(m-m_0, n-n_0)$ with a superposition of point sources. Specifically, given the Fourier intensity of $x(m,n)$ and the Fourier intensity of

$$y(m,n) = x(m,n) + \sum_{k=1}^{N} A_k\delta(m - m_k, n - n_k) \tag{2}$$

we have generated a set of conditions that are sufficient for the unique reconstruction of the image $x(m,n)$. In addition, we have developed an algorithm that will perform this reconstruction that only requires finding the solution to a set of linear equations [1,2].

PUBLICATIONS:

Chapters in Books:

1. C. Auyeung and R. M. Mersereau. "A dual approach to signal restoration," in *Topics in Applied Physics*, vol. xx, Iterative Signal Restoration, (A. K. Katsaggelos, ed.) to be published by Springer-Verlag, Berlin, 1990.

Journal Articles Published or Accepted:

1. J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," to appear in *Proc. IEEE*, 1990 (invited).

2. S. J. Reeves, and R. M. Mersereau, "Optimal estimation of the regularization parameter and stabilizing functional for regularized image restoration," to appear in *Optical Engineering*, 1990 (invited).

3. W.S. Kim and M.H. Hayes, "Phase retrieval using two Fourier transform intensities," accepted for publication in *Journal Optical Society of America: A*. March, 1990.

Conference Proceedings:

1. C. Auyeung and R. M. Mersereau, "The dual approach to signal restoration," *Proc. 1989 International Conference on Acoustics, Speech and Signal Processing*, pp. 1326-1329.

2. C. Auyeung and R. M. Mersereau, "Efficient algorithms for least squares restoration," *SPIE Conference on Visual Communication and Image Processing IV*, pp. 1534-1540, 1989.

3. S. J. Reeves and R. M. Mersereau, "Regularization parameter estimation for iterative image restoration in a weighted Hilbert space," accepted for publication *1990 IEEE International Conference on Acoustics, Speech and Signal Processing.*

4. W.S. Kim and M.H. Hayes, "Phase retrieval using two Fourier transform intensities," *Proc. Opt. Soc. Am. Topical Conf. on Signal Recovery*, pp. 138-141, Cape Cod, June 1989.

5. W. Kim and M.H. Hayes, "Phase retrieval using two Fourier intensities" to be presented at 1990 Int. Conf. on Acoustics, Speech, and Sig. Proc.

References

[1] W.S. Kim and M.H. Hayes, "Phase retrieval using two Fourier transform intensities", *Proc. Opt. Soc. Am. Topical Conf. on Signal Recovery*, pp. 138-141, Cape Cod, June 1989.

[2] W.S. Kim and M.H. Hayes, "Phase retrieval using two Fourier transform intensities", submitted for publication in *Journal of the Optical Society of America*

[3] M.H. Hayes, "Signal reconstruction from spectral phase or spectral magnitude", *Advances in computer vision and image processing*, Chapter 3, vol. 1, Edited by T.S. Huang, JAI Press, 1984.

[4] M.H. Hayes, "The unique reconstruction of multidimensional sequences from Fourier transform magnitude or phase", *Image recovery: Theory and application*, Chapter 6, Edited by H. Stark, Academic Press, 1986.

[5] M.H. Hayes, "Inverse problems: An overview", *J. Soc. Instr. and Control Eng.* , vol. 15, no. 12, pp. 1089-1094, Dec. 1986, JAPAN (invited).

[6] M.H.Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform", *IEEE Transaction on Acoustics, Speech, and Signal Processing*, ASSP-30(No. 2):pp140–154, April 1982.

[7] A.K. Katsaggelos "Constrained iterative image restoration algorithms", Ph.D. Thesis, Georgia Institute of Technology School of Electrical Engineering, August 1985.

[8] M.H. Hayes and T.F. Quatieri, "Recursive Phase Retrieval Using Boundary Conditions", *Journal Opt. Soc. Am.*, vol. 73, no. 11, pp. 1427-1433, Nov. 1983.

[9] J.R. Fienup, "Reconstruction of objects having latent reference points", *Journal Opt. Soc. Am.*, vol. 73, no. 11, pp. 1421-1426, Nov. 1983.

Work Unit Two

TITLE: Representation, Coding, and Analysis of Images

SENIOR PRINCIPAL INVESTIGATORS:

R. W. Schafer, Regents' Professor
R. M. Mersereau, Regents' Professor and Rockwell Fellow

SCIENTIFIC PERSONNEL:

F. J. Malassenet, (Ph.D. Candidate)
L. Hertz, (Ph.D. Candidate)
C. H. Richardson, (Ph.D. Candidate)
K. Truong, (Ph.D. Candidate)
D. Suh, (Ph.D. Candidate)

SCIENTIFIC OBJECTIVE:

The first major objective of this work unit is to study and develop new approaches to the modeling of images and to the representation of the information contained in images. A second major objective is to use such models and representations to solve problems in automatic extraction of information from images and in efficient digital coding of images.

RESEARCH ACCOMPLISHMENTS:

- *Modeling of Textured Images*

The coding and detection of textures are important in a number of applications in automatic image analysis, coding, and in computer graphics. Textures can be important in the segmentation of images and in the determination of depth information from a single view. The problem is also a particularly difficult one. Researchers have used a variety of different models with some success. One common model is a Markov random field model. One difficulty with this model is that it does not account for the scale invariant property of textures. Markov models do not perform well when used for natural textures in three-dimensional scenes which are scaled by perspective transformations. Mandelbrot and Voss used a model which exploited fractals that could generate textures with stunning accuracy. Pentland used fractal Brownian motion to model textures. In the scope of our study we have used deterministic fractal textures, which can incorporate the self-affine invariant property of textures. A good model for this purpose is the Hyperbolic Iterated Function System (IFS) model developed by Barnsley. Natural textures have been generated using this method. All of these models have been used

primarily for the generation of textures. By varying the parameters of the models, an album of textures can be generated that can be compared with a desired texture. We have been interested in the inverse problem, i.e. the problem of estimating the model parameters from natural textures.

The IFS model expresses the image f using a generalized autoregressive model of the form

$$f(\mathbf{x}) = \sum_i \frac{p_i}{|w_i|} f(w_i(\mathbf{x})) \tag{1}$$

The $\{w_i\}$ are a set of affine transformations of the form

$$w_i(\mathbf{x}) = \mathbf{A}_i\mathbf{x} + \mathbf{b}_i$$

The quantities $|w_i|$ are equal to the absolute value of the determinant of the matrix $\mathbf{A}$, which must be less than one, and the p_i's are a set of weights (probabilities) which are bounded between zero and one. The parameters to be estimated are the number of maps N, the probabilities p_i, and the map parameters.

In our first attempt we worked with a synthetic one-dimensional texture that satisfied the model exactly with all of the scaling parameters being equal. In this case there is a rather straightforward method for solution. We can define a discrimination measure as the inverse Fourier transform of the ratio of the Fourier transforms of $f(x)$ and $f(ax)$, which will be the sum of impulses at $x = b_i$ whose areas are p_i. This is shown in Figure 1. This approach does not generalize to textures which have more than one scaling parameter and the performance of the algorithm is quite sensitive to the manner by which the images and the models are digitized.

In the more general case the problem can be viewed as fitting $7N - 1$ unknowns to a nonlinear model. Iterative methods for nonlinear optimization can be applied to this problem, but they are not computationally efficient and they encounter numerical difficulties for more than a very few maps. We are continuing to look for approaches where the parameters can be estimated sequentially.

During the past year, as a part of this study, an algorithm has been developed for the discrete computation of the Hutchinson metric which can measure the distance between two images that are represented as measures.

- *Knowledge-Based Image Analysis*

This project is concerned with the identification and labeling of features in noisy imagery. Because this work was partially supported from other sources, the specific application involves the identification and isolation of the inner and outer walls of the left ventricle of the beating heart from three-dimensional nuclear magnetic resonance images. Our primary interest, however, has been in the methodology used. A statistical approach, based on the Dempster-Shafer theory of evidence has been used to combine measurements from a variety of sources to confirm hypotheses for a series of labels for line segments in the scene. As additional measurements are acquired the likelihood for

a certain set of labels becomes sufficiently overwhelming that labels can be confidently assigned. These assignments then affect the likelihoods of other labels on other frames or in other parts of the images. This methodology is applicable to a number of other problem domains including automatic target recognition, robotics, image understanding, and computer vision.

Image analysis systems usually begin with a low-level analysis. The results of this analysis are then combined, refined, and finally labeled. In the specific medical imaging application that we have been considering, 96 images of cross sectional views of the heart form the raw input data. These are taken at six different vertical positions at sixteen different times (phases) within the cardiac cycle. The ventricle walls that we are looking for are quasi-circular in shape. For this reason the images are first mapped onto an $r - \theta$ coordinate system centered in the center of the ventricle. An edge detector, consisting of a Sobel operator, adjusted to find vertical edges, is then applied to the resulting images. The output of the Sobel operator is thresholded and the resulting line segments (corresponding to circular arcs in the original) are stored in a list. Associated with each are a number of measurements – radial position, arc length, gradient value, intensity values on left and right sides of the edge, etc. These arcs are non-ideal in a number of ways. A large number of spurious features are found and none of the contours are complete (i.e. there are gaps).

The goal of the higher-level processing is to attach a label to each of the arcs. Each should be labeled as belonging to the inner wall, the outer wall, or extraneous. (The extraneous group contains spurious line segments as well as those due to boundaries other than the one of interest.) Once all of the segments are labeled, all of those with the same label should be connected by removing gaps. In the intermediate stages there might be additional classes as well. For example the outer wall may be divided into the boundary with the lung, boundary with the right ventricle, boundary with the liver, etc.

In the formalism of the Dempster-Shafer theory with each line segment for each hypothesis we associate three numbers, the degree of belief, the degree of disbelief, and the degree of ignorance. Initially, maximum ignorance is assigned to each hypothesis, but as the measurements are evaluated, the other two quantities take over until each hypothesis is either accepted or rejected for each line segment. Quantities similar to conditional probabilities that have been measured from training data are used for updating the belief functions. As with human judgments, once enough evidence has been received to assure complete belief, no further evidence is evaluated. Once all of the line segments are labeled, gaps can be filled by lowering the threshold on the Sobel operator and repeating the process in selected regions. Heuristic rules can also be useful.

We have encoded a knowledge-based system for analyzing MRI images using these principles. It works well on simple frames but makes errors on some difficult ones. This is because a number of key issues remain to be addressed. Our early results have been presented in four papers [1,2,3,4]. The first of these extends the Dempster-Shafer theory to include measurements of quality on a particular data base. When evidence is

combined from several data bases it is important because of computational efficiency, that the most important data bases be combined first. These quality measures make this possible [3]. The other three papers describe our system [1,2,4].

- *Self-Organizing Feature Maps*

One of the goals of neural network research is to study the mechanisms by which the human brain processes information in order to model the functionality of the brain for perceptual tasks. Possible applications of neural network algorithms are almost uncountable but major areas are pattern (speech and image) recognition and machine vision. The recent impetus for this research area was caused by recent findings in brain physiology over the past few decades combined with the recent hardware advances (electronic and optical) that make it possible to implement neural net algorithms.

Many experimental findings have led to the conclusion that both how the neurons are interconnected and where the neurons are located are important in order to understand and analyze human perception. Neurons only interact locally with their close neighbors and the interaction is very structured. The degree of interaction decreases as the distance between cells increases.

The idea of the SOFM algorithm is to devise an adaptive system that self-organizes a given set of stimuli (input patterns) into a compact and structural representation called the feature map of the inputs. The formation of the feature map is through the interactions of each unit with its neighbors. A given set of observable input patterns X is mapped onto the output set Y via the system (weight) matrix W. The approach is to adapt W such that each region of the output becomes most sensitive to a particular set of input patterns; different regions will respond to different subsets of X. In order to find W, some sort of training algorithm is needed. The simplest approach is to find the unit that best matches a given input and then define a neighborhood around this unit to account for lateral interaction. For a given input pattern x at iteration t, the best-matched neuron is the one $y_{\max}$ which has the smallest norm with x. Thus

$$y_{\max}(t) = \min_{1 \le i \le N} ||w_i(t) - x(t)|| \ . \tag{2}$$

If neuron i is in the neighborhood of $y_{\max}$, then the weight matrix is updated as

$$w_i(t+1) = w_i(t) + g(t)[w_i(t) - x(t)] \ , \tag{3}$$

otherwise

$$w_i(t+1) = w_i(t) \ , \tag{4}$$

where $g(t)$ is a gain function that controls the rate of adaptation.

Two of the parameters that govern the training process are the gain function and the neighborhood function. Unfortunately, there are currently no analytic solutions for the best gain and neighborhood size. All of the functions that have been reported upon in the literature are based upon simulation experience.

We have been applying the use of self-organizing feature maps to the design of codebooks for vector quantization for images. Some of the tasks considered are related to properties of SOFM's themselves and others are related to the application. It must be emphasized that our primary interest is in the neural nets, but that such a study without reference to an application is of little value.

Codebooks designed using this approach can be created and searched efficiently and the coded imagery appears to be subjectively more pleasing than with codebooks generated by more traditional means. Our preliminary results are described in [4].

- *An Environment for Design and Analysis of Morphological Systems*

As a first step toward developing a design methodology for morphological systems, we have begun to develop a LISP-based design/analysis environment capable of performing symbolic and numerical manipulations of complex morphological operators. The symbolic manipulations consist of expression rearrangements and simplifications as well as signal and system analysis using knowledge of the transformations and structuring elements. This involves the use of object-oriented programming techniques for compiling and applying knowledge about morphological operators and their interconnections and for incorporating knowledge of signal properties and system implementation constraints. Symbolic manipulations are performed by applying classes of morphological rules to compound morphological expressions. The symbolic manipulations are performed in an expert system manner with a forward chaining control strategy for generating both both simplifications and equivalent forms of expressions. The rule base of the environment is already quite extensive and the system is capable of simplifying and manipulating highly complex expressions representing compound morphological operators.

The numerical processing component of this environment, which uses deferred evaluation of expressions, serves as the base upon which the symbolic processing capabilities are constructed, and provides the means for testing morphological systems on real signals. Numerical manipulations are facilitated by a signal representation based on the notion of *signal classes* and *abstract data objects*. Signals are organized into signal classes in order to allow the possibility of inheritance of computational methods and other signal properties. The completion of this software environment will make it possible to do such things as construct and test morphological systems, explore design trade-offs, minimize computational complexity, and explore alternative implementation configurations.

- *Multilevel Thresholding for Image Analysis*

Thresholding is an effective method for simplifying images while retaining shape and geometrical structure. Multiple thresholds are needed when thresholding images containing several objects of differing brightness or reflectivity. Also, to account for variations of grey levels due to non-uniform illumination, it is often necessary to allow the thresholds to vary across the image. A technique called *edge matching* has

been developed for the class of images consisting of untextured objects on a untextured background. The multiple thresholds are adjusted so that the edges of the thresholded image closely match the edges of the original grey-tone image. The result is a multi-level thresholded image that preserves the shape and geometrical structure of the objects in the image. It can then be "sliced" to extract objects that occupy the different threshold bands. Morphological systems can then be used to process these slices.

- *Thresholding Greytone Edge Maps*

Many techniques are available to produce a grey-scale edge strength image; e.g., convolution with first derivative of a gaussian, Sobel operators, and differences of morphological operators. In a large number of applications, it is necessary to threshold this edge image into a binary image representing the edge locations. Since the edges appear in the greytone edge strength image as peaks of varying magnitudes, determining the edge locations is a peak picking problem rather than a thresholding problem. A simple method using the morphological top hat transformation to accomplish this peak picking has been developed. By choosing a structuring element of the appropriate size, the resultant edges may have any desired width. This method may generalize to other one and two dimensional peak picking applications.

- *A New Morphological Edge Detector*

As a part of the above mentioned research on multilevel thresholding, it has been necessary to develop an edge detector that is computationally fast, provides edge markers of a prescribed uniform width, and is insensitive to noise in the image. Such an edge detector has been under development, with highly promising results. The method combines morphological operations similar that of Lee, et. al. with the thresholding algorithm described above. Additional morphological operations are incorporated to automatically eliminate many of the spurious edge indications, and thereby focus on the "dominant edges" in the image. Figure 2 shows an example of the performance of this edge detector for smoothing with several different structuring elements. It can be seen that the smoothing is effective in removing small unconnected edges.

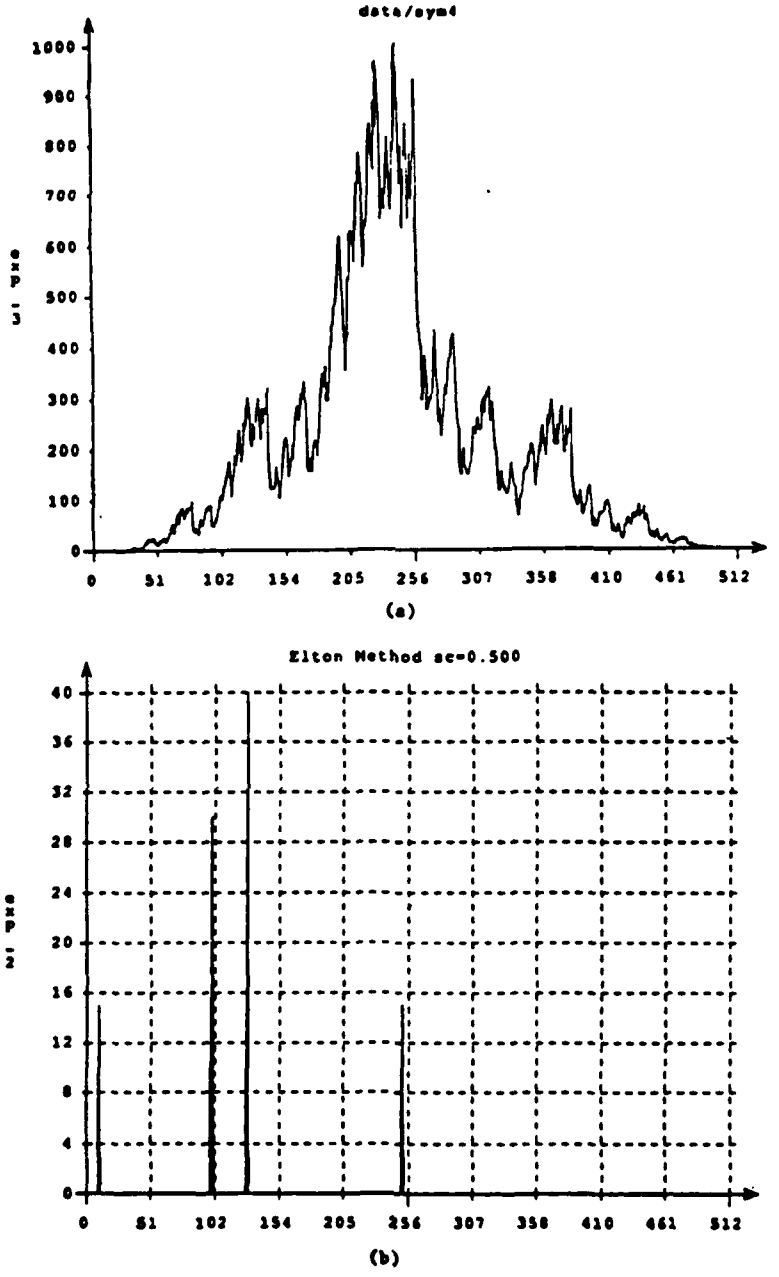
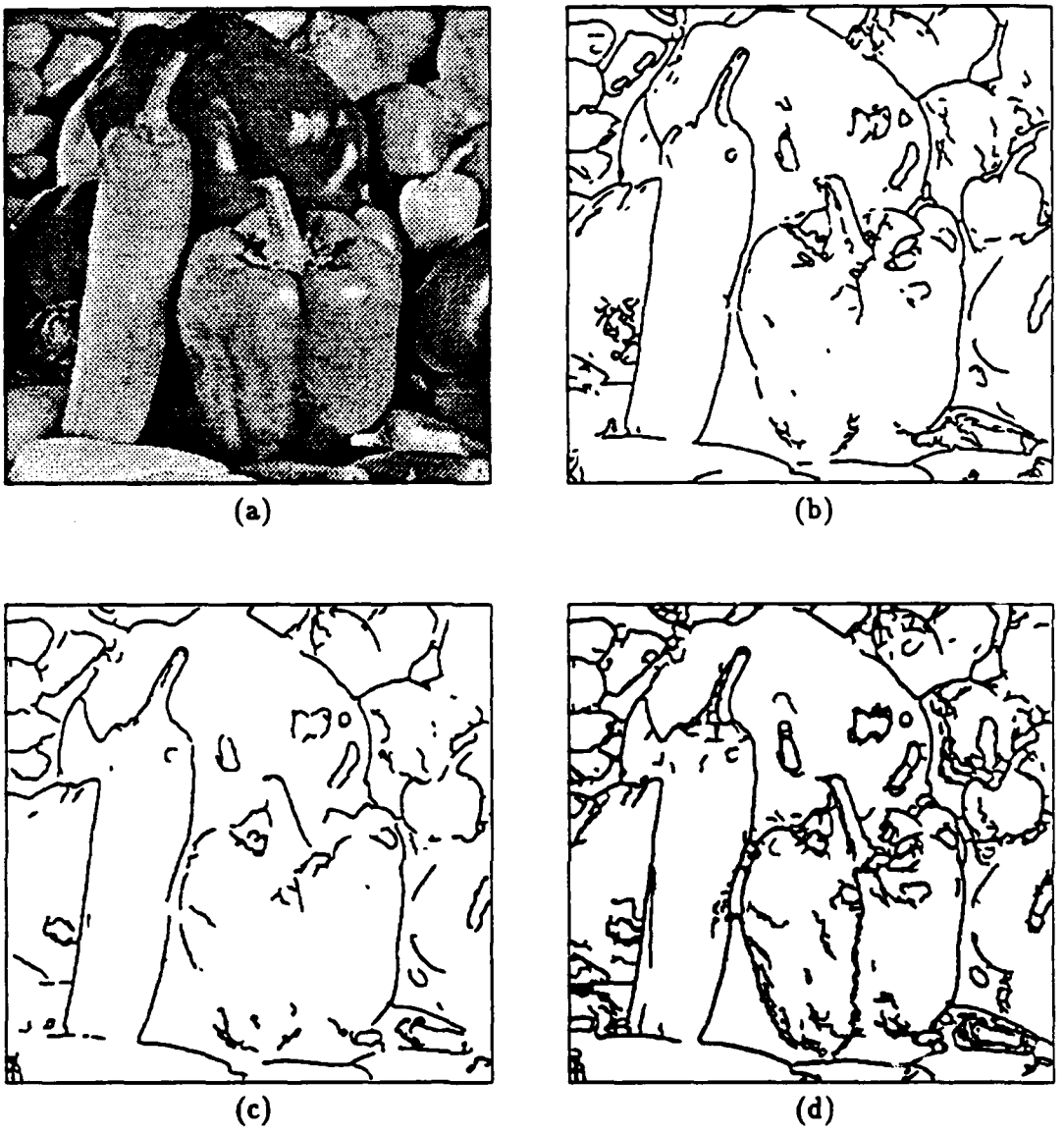


Figure 1. (a) Synthetic one-dimensional texture, $f(x)$. (b) Discrimination measure for estimating the texture parameters.



(a) Original greytone image.
(b) Extracted edges after smoothing with a circular neighborhood of size 1 (one pixel wide edges).
(c) Extracted edges after smoothing with a circular neighborhood of size 2 (one pixel wide edges).
(d) Extracted edges after smoothing with a circular neighborhood of size 2 (two pixels wide edges).

Figure 2. Illustration of a new morphological edge detector.

PUBLICATIONS:

Journal Articles:

1. P. A. Maragos, "A representation theory for morphological image and signal processing," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, pp. 586-599, June, 1989.

2. P. A. Maragos and R. W. Schafer, "Morphological systems for multidimensional signal processing," accepted for *IEEE Proceedings* special issue on multidimensional signal processing.

3. D. Y. Suh and R. M. Mersereau, "Accuracy measure in evidence theory," submitted to *J. Math. Anal. and Applications.*

4. L. Hertz and R. W. Schafer, "On thresholding edge images," submitted to *Computer Vision, Graphics, and Image Processing*, 1989.

Conference Proceedings:

1. D. Y. Suh, R. M. Mersereau, R. L. Eisner, and R. I. Pettigrew, "Knowledge-based boundary detection applied to cardiac magnetic resonance image sequences," *Proc. 1989 International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 1783-1786.

2. D. Y. Suh, R. M. Mersereau, R. L. Eisner, and R. I. Pettigrew, "A system for knowledge-based bounary detection of cardiac magnetic resonance image sequences," accepted for publication *1990 International Conference on Acoustics, Speech and Signal Processing*, April 3-6, 1990, Albuquerque, NM.

3. K. Truong and R. M. Mersereau, "Structural image codebooks based upon the Kohonen self-organizing feature map algorithm," accepted for publication *1990 IEEE International Conference on Acoustics, Speech and Signal Processing.* April 3-6, 1990, Albuquerque, NM.

4. C. H. Richardson and R. W. Schafer, "Symbolic Manipulation and Analysis of Morphological Expressions," to be presented at the *IEEE 1990 International Conference on Acoustics, Speech and Signal Processing*, April 3-6, 1990, Albuquerque, NM.

5. D. Y. Suh and R. M. Mersereau, "Accuracy measure in evidence theory," *Journal of Mathematical Analysis and Applications*, submitted (abstract, 1 page).

6. D. Y. Suh, R. M. Mersereau, R. L. Eisner, and R. I. Pettigrew, "Automatic boundary-detection on cardiac magnetic resonance image sequences for four dimensional visualization of the left ventricle," *Visualization in Biomedical Computing*, accepted, Atlanta, Georgia, May, 1990, (abstract, 1 page).

Talks and Papers without Proceedings:

1. C. H. Richardson and R. W. Schafer, "Morphological signal processing systems," talk at ESL, Sunnyvale, CA, October, 1989.

References

[1] D. Y. Suh, R. M. Mersereau, R. L. Eisner, and R. I. Pettigrew, newblock "Knowledge-based boundary detection applied to cardiac magnetic resonance image sequences," *Proc. 1989 International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 1783-1786.

[2] D. Y. Suh, R. M. Mersereau, R. L. Eisner, and R. I. Pettigrew, "A system for knowledge-based bounary detection of cardiac magnetic resonance image sequences," accepted for publication *1990 International Conference on Acoustics, Speech and Signal Processing*, April 3-6, 1990, Albuquerque, NM.

[3] D. Y. Suh and R. M. Mersereau, "Accuracy measure in evidence theory," *Journal of Mathematical Analysis and Applications*, submitted (abstract, 1 page).

[4] D. Y. Suh, R. M. Mersereau, R. L. Eisner, and R. I. Pettigrew, "Automatic boundary-detection on cardiac magnetic resonance image sequences for four dimensional visualization of the left ventricle," *Visualization in Biomedical Computing*, accepted, Atlanta, Georgia, May, 1990, (abstract, 1 page).

Work Unit Three

TITLE: Multiprocessor Architectures for Digital Signal Processing

SENIOR PRINCIPAL INVESTIGATOR:

T. P. Barnwell, III, Professor and Rockwell Fellow

SCIENTIFIC PERSONNEL:

P. R. Gelabert, (Graduate Research Assistant and Ph.D. Candidate)
C. P. Hong, (Graduate Research Assistant and Ph.D. Candidate)
B. M. Kim, (Graduate Research Assistant and Ph.D. Candidate)
H. P. Kim, (Graduate Research Assistant and Ph.D. Candidate)

SCIENTIFIC OBJECTIVE:

The primary objective of this work unit is to develop systematic techniques for the automatic generation of provably optimal multiprocessor implementations for a broad class digital signal processing (DSP) algorithms and for a broad class of multiprocessors systems. Stated in another fashion, the goal of this research is to develop DSP "compilers," where the input is an algorithm specification and the output is a complete, optimal multiprocessor implementation. An additional long-term goal is to make these optimal multiprocessor compilers available to a broad class of DSP system designers.

A basic philosophy of this research has always been to perform the theoretical developments in the context of an actual multiprocessor system. The first system used for this purpose, an eight processor LSI-11/2 system, is now obsolete. The current research is being tested using the OSCAR-32, a personal computer (PC) based multiprocessor utilizing the AT&T DSP-32 floating point DSP microprocessor [1].

RESEARCH ACCOMPLISHMENTS:

Over the past year, work on the area of multiprocessor architectures for DSP has been centered in seven areas: scheduling theory for parallel pipeline architectures; scheduling theory for time skewed multiprocessors; graph compilation for commercial DSP chip based multiprocessors; graph transformations for more efficient total implementations; optimal scheduling of graphs with data dependent branching; the development of an automated system for designing integrated circuits for multiprocessor systems; and the development of the OSCAR-32 multiprocessor. In addition, considerable progress has been made toward the development of a standard (portable) environment for multiprocessing.

- *Scheduling Theory for Parallel Pipeline Architectures*

A new research topic which was initiated last year was the extension of cyclic multiprocessor scheduling techniques such as those used in the cyclo-static [2] and generalized SSIMD [3] compilers to pipeline and parallel pipeline architectures. This new theory is based on a set of transformations which represent pipeline and parallel pipeline architectures as parallel MIMD systems. These transformations represent individual pipeline processors (such as the AT&T DSP32 processor used in the OSCAR-32) as a number of *pseudo-processors* in a parallel structure. After the transformations, the result is a fully parallel architecture with no pipelined elements. Thus, parallel processor scheduling techniques can be applied directly to the transformed architectures to achieve optimal or near-optimal schedules. These multiprocessor realizations are then transformed back into their equivalent optimal or near optimal pipeline or parallel pipeline realization.

The multiprocessor representation needed to accommodate the transformed pipeline and parallel pipeline systems is nearly identical to that needed for parallel systems. The one difference is that the extended representation must allow for a synchronous clock skew between (groups of) individual processors. The leads to the concept of *clock classes*, which are the sets of all processors which share the same clock skew. In a traditional multiprocessor system, the clock skew is zero and all the processors belong to the same clock class. The effect of the clock classes is to provide a new set of constraints which are applied during the pruned tree search used in the cyclic scheduling algorithms.

At this point in time, a compiler has been written for finding the optimum pipelined implementation for parallel pipelined processors in which the arithmetic pipeline has no feedback data paths [5]. The eventual goal is to write a compiler for general architectures in which there are a large number of identical pipelined processors.

- *Scheduling Theory for Time Skewed Multiprocessors*

In the research on parallel pipelined machines discussed above, it was discovered that pipelined multiprocessors can be abstractly transformed into equivalent synchronous

multiprocessor systems with time varying communications constraints. In the course of that research, it was also observed that a class of synchronous nonpipelied multiprocessors exists which can transformed into exactly the same abstract class of constrained synchronous multiprocessors. This is the class of all *clock skewed* multiprocessors in which each processor in the system operates with a fixed and unique time skew with respect to its neighbors.

The basic equivalence of *clock skewed* and *pipelined* multiprocessors means that any compiler that can generate optimal realizations for one can also be used to generate optimal realizations for the other. Clock skewed multiprocessors turn out to be of great interest in their own right. Such systems have two major advantages. First, they can use multiple shared busses without conflict, which vastly simplifies the difficult multiprocessor communications problem. Second, for the same basic hardware, clock skewed systems often attain better performance that the equivalent system without clock skew. A discussion of this research can be found in [8]. The OSCAR-32 is being modified to have a clock skewed mode.

- *Graph Compilation for Commercial DSP Chip Based Multiprocessors*

While single chip DSP processors are approaching a significant level of maturity, they are not well designed to met the needs of multiprocessing, particularly with respect to fine grain multiprocessing. Based on the WE DSP32 floating point DSP chip a small multiprocessor, the OSCAR-32, has been implemented. The OSCAR-32 has several features that make it difficult to practically apply the (P)SSIMD scheduling methodology. A practical graph based compiler for the system has now been development. The compiler extends well known code generation techniques and combines them with extensions to list based CPM scheduling. The main goal is to determine what are the bottlenecks in the architecture in terms of a realization and in terms of a practical compiler. A preliminary version is currently functioning and produces good code for the OSCAR-32 multiprocessor, as well as several other currently available DSP chips.

This research work can be divided into three parts. The first is the development of the practical multiprocessor compiler for DSP chips discussed above. This compiler will be used to generate code for several current DSP multiprocessors. The second part is the development of a "modification of optimal" compiler also for multiprocessors based on current DSP chips. This compiler will begin with an optimal implementation for an ideal processing element, and will systematically modify the solution so that it will fit the target processor. The final part will be to systematically explore a set of architectural choices with the goal being a practical single chip DSP for fine grain multiprocessing with practical compiler/scheduling tools.

The research is now quite mature, having incorporated a large number of locally optimal and suboptimal techniques in a single compiler environment. A number of major performance improvements have been obtained by applying optimal cyclic (cyclo-static) scheduling techniques locally in inner loops of acyclic graphs [7].

- *Code Simplification for More Efficient Total Implementations*

The task of the *Code Simplification System* is to increase the computational speed of a given program by using application specific information. This *a priori* information can be used to make local or global code simplifications to remove redundant or trivial operations. In the generation of fast Fourier transform (FFT) code, for example, information about the data to be processed is frequently known before execution time could be used to remove trivial calculations introduced by known inputs or outputs (e.g., data size, data type, and known location of zero valued inputs or don't care outputs).

In order to effectively utilize the application specific information, a proper set of simplification rules must be chosen that can be used to symbolically rewrite the software. The rules should be modular in the form of a rule-base to yield a system that could easily use several different algorithm specific rule-bases. A *rule-driver* is required to apply these rules, and a representation for the code of the given program is needed to allow these types of simplifications to take place. The representation that is chosen must retain enough information about the original code to enable the system to reconstruct a working program. Finally, the representation should be flexible, allowing for the code to be structured in a hierarchical fashion. In this way, high level rules could be implemented that handle large program sections as units.

The results of the code simplification research are reported in [4] and [6].

- *Optimal Scheduling of Graphs with Data Dependent Branching*

Much of the previous research in synchronous multiprocessor implementations for DSP algorithms has assumed that the algorithm of interest can be represented as a cyclic graph. This assumption excludes all algorithms in which data dependent branching occurs and all machines in which the durations of the arithmetic operations are data dependent. The goal of this research is to extend the existing multiprocessor scheduling theory to include algorithms with small amounts of data dependent branching. The approach in this area is to combine recent techniques developed for scheduling very wide instruction word machines (VLIW) with data dependent branching with our own techniques for scheduling fully specified flow graphs in order to develop a compiler which can find optimal or near-optimal schedules for graphs with data dependencies. The first goal is to solve the problem for graphs with only one data dependency.

At this point, the emphasis is on developing an extended formalism which extends our previous concepts of optimality to graphs with data dependencies. The data dependencies include both deterministic dependencies and stochastic dependencies. An initial compiler for deterministic data dependencies has also been demonstrated.

- *The Development of an Automated System for Designing Integrated Circuits for Multiprocessor Systems*

We have known for some time that it is possible, in principle, to do direct compilation from fully specified flow graphs to generate optimal integrated circuit realizations. However, because of the lack of good tools in both the graph compilation and integrated circuit design areas, such complete systems have been prohibitively complex. Now, however, the tools in these two areas have matured to the point that such a system is feasible.

Our system will be based on a combination of our existing cyclo-static compilers and the LAGER system from the University of California at Berkeley. The LAGER system has been installed and tested. The next step is to develop a set of macro cells on which the signal processing multiprocessors will be based. The initial application area will be time-invariant and adaptive filters.

- *Development of the OSCAR-32*

The OSCAR-32 synchronous multiprocessor system is the current target system for the multiprocessor research. It is hosted on personal computers, and is composed of processing elements designed and constructed at Georgia Tech. Several small OSCAR-32 systems have been configured (up to three boards and nine processors). In 1988, considerable progress was made toward the construction of a much larger OSCAR-32 system.

Figure 1 shows a block diagram of an OSCAR-32 constituent processor. This processor board contains three AT&T WE-DSP32 processing chips which can deliver up to 37.5 MFLOPS. This board was designed in 1986 and 1987, and was extensively tested in small multiprocessor systems in late 1987 and early 1988. In late 1988, twenty constituent processors were constructed and tested.

In 1989, a sixteen processor system was constructed. The system is hosted on a SUN 386i computer operating in the DSP Laboratory computer network. This system is currently functioning, and is being used in the research.

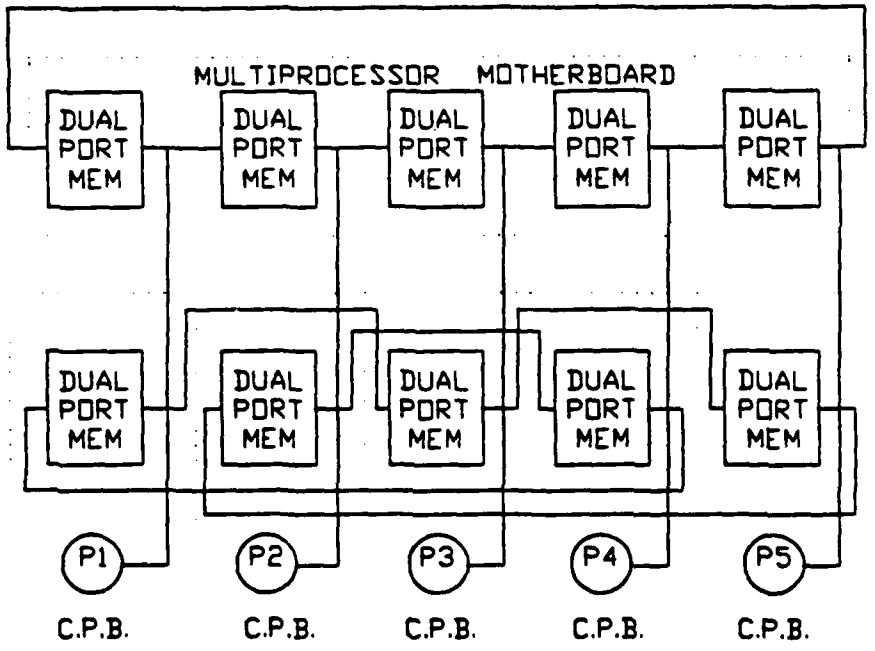


Figure 1. Block diagram of the OSCAR-32 constituent processor board.

PUBLICATIONS:

Journal Articles:

1. D. J. Pepper, T. P. Barnwell III and M. A. Clements, "A Ring Parallel Processor for Hidden Markov Model Training," accepted for publication in *IEEE Transactions on ASSP.*

Papers in Conference Proceedings:

1. L. P. Heck, D. A. Schwartz, R. M. Mersereau, and J. H. McClellan, "Symbolic Simplification of Digital Signal Processing Software," *Proc. 1989 Intern. Symp. on Circuits and Systems*, Portland, OR, May, 1989.

2. S. J. A. McGrath, C. P. Hong and T. P. Barnwell III,"A Scheduling Methodology for a Synchronous Cyclo-Static Multiprocessor," *1989 International Conference on Systolic Processors*, Kilarny, Ireland, May 1989.

3. B. M. Kim and L. P. Heck, "Automatic Design of Parallel Implementations of DSP Algorithms," *International Conference on Acoustics, Speech and Signal Processing*, Albuquerque, New Mexico, April 3-6, 1990.

4. K. Nayebi, T. P. Barnwell III, and M. J. T. Smith, "The Time Domain Analysis and Design of Exactly Reconstructing Analysis/Synthesis Filter Banks," *International Conference on Acoustics, Speech and Signal Processing*, Albuquerque, New Mexico, April 3-6, 1990.

5. K. Nayebi, T. P. Barnwell III, and M. J. T. Smith, "A General Time Domain Analysis and Design Framework for Exactly Reconstructing FIR Analysis/Synthesis Filter Banks," *International Symposium on Circuits and Systems*, New Orleans, Louisiana, May 1-3, 1990.

6. T. P. Barnwell and M. J. T. Smith, "Filter Banks for Analysis-Reconstruction of Systems: A Tutorial," (invited) *International Symposium on Circuits and Systems*, New Orleans, Louisiana, May 1-3, 1990.

7. B. M. Kim and T. P. Barnwell, "Resource Allocations and Code Generation for Pointer Based Pipelined DSP Multiprocessors," *International Symposium on Circuits and Systems*, New Orleans, Louisiana, May 1-3, 1990.

8. C. P. Hong and T. P. Barnwell, "The Implementation of Shift-Invariant Flow Graphs on Clock-Skewed Multiprocessors," *International Symposium on Circuits and Systems*, New Orleans, Louisiana, May 1-3, 1990.

References

[1] S . J. A. McGrath, T. P. Barnwell III and D. A. Schwartz, "A WE-DSP-32 Based, Low Cost Multiprocessors for Cyclo-Static Implementations, "*Proc. of the International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX, April, 1987.

[2] D . A. Schwartz, "Cyclo-Static Realizations, Loop Unrolling and CPM: Optimal Multiprocessor Scheduling," *Published by Prentice Hall in 1988 as a chapter in a book edited by Stuart Tewksbury* (based on an invited presentation at the 1987 Princeton Workshop on Algorithm, Architecture and Technology Issues in Models of Concurrent Computations, Sept. 1987, Princeton, NJ).

[3] H . R. Forren, "Multiprocessor Design Methodology for Real Time Digital Signal Processing Systems Represented by Shift Invariant Flow Graphs" Ph.D. Thesis, Georgia Institute of Technology, Atlanta, Georgia, May, 1988.

[4] L . P. Heck, D. A. Schwartz, R. M. Mersereau, and J. H. McClellan, "Symbolic Simplification of Digital Signal Processing Software," *Proc. 1989 Intern. Symp. on Circuits and Systems*, Portland, OR, May, 1989.

[5] S . J. A. McGrath, C. P. Hong and T. P. Barnwell III,"A Scheduling Methodology for a Synchronous Cyclo-Static Multiprocessor," *1989 International Conference on Systolic Processors*, Kilarny, Ireland, May 1989.

[6] B . M. Kim and L. P. Heck, "Automatic Design of Parallel Implementations of DSP Algorithms," *International Conference on Acoustics, Speech and Signal Processing*, Albuquerque, New Mexico, April 3-6, 1990.

[7] B . M. Kim and T. P. Barnwell, "Resource Allocations and Code Generation for Pointer Based Pipelined DSP Multiprocessors," *International Symposium on Circuits and Systems*, New Orleans, Louisiana, May 1-3, 1990.

[8] C . P. Hong and T. P. Barnwell, "The Implementation of Shift-Invariant Flow Graphs on Clock-Skewed Multiprocessors," *International Symposium on Circuits and Systems*, New Orleans, Louisiana, May 1-3, 1990.

Work Unit Four

TITLE: Two-Dimensional Optical Information Processing

SENIOR PRINCIPAL INVESTIGATORS:

Thomas K. Gaylord, Regents' Professor
E. I. Verriest, Associate Professor
K. F. Brennan, Associate Professor
E. N. Glytsis, Assistant Professor

SCIENTIFIC PERSONNEL:

T. J. Drabik, Graduate Research Assistant (Ph.D. Candidate)
M. C. Fazio, Graduate Research Assistant (Ph.D. Candidate)
T. A. Maldonado, A. R. O. Fellow (Ph.D. Candidate)
D. W. Wilson, Graduate Research Assistant (Ph.D. Candidate)
G. N. Henderson, Graduate Research Assistant (Ph.D. Candidate)
Karim Diff, Post Doctoral Fellow

SCIENTIFIC OBJECTIVE:

The long-term objective of this research is to develop broadly-based, theoretical and experimental knowledge of two-dimensional optical information processing including algorithms, architectures, systems, and devices. This brings together a range of concepts from basic physics to information processing in its most generalized form. Optical systems based on content-addressable memory processing, associative processing, and Givens rotations are being analyzed starting from basic physical principles and extending through experimental systems performance.

Another phase of this research is to generate the necessary knowledge of nanometer-scale semiconductor quantum wave devices that will allow the practical development and use of these structures. Utilization of these quantum devices would be both for ultra-small high-speed versions of present-day devices and for future guided electron wave integrated circuits. This activity would bring together understanding of the underlying physics, device modeling, development of software tools, device design methodology, evaluation of existing quantum interference experiments, crystal growth, device fabrication, and device testing.

RESEARCH ACCOMPLISHMENTS:

In truth-table look-up data processing all possible binary inputs are considered together with their corresponding outputs. A truth table can be realized either with a content-addressable memory (CAM) in which the input represents the data, and the

output is all of the addresses where those data occur. These addresses are provided simultaneously in parallel. Direct implementation of truth tables provides faster computation, since the entire operation is performed in a single step and not as a series of sequential steps. In the last three years, advances in number representation, multi-level coding, large-scale truth-table reduction techniques have stimulated further interest in table look-up implementations. It has been shown that in order to determine the practicality of truth-table look-up techniques for a particular operation, the sizes of the logically reduced tables must be determined. In turn, the sizes of the truth tables may be reduced by one or a combination of the following techniques: 1) using the residue number system, 2) using the modified signed-digit number system, 3) applying multi-level coding (ternary, etc.), and 4) applying logical minimization techniques.

The development of optically interconnected VLSI and wafer scale systems has a prerequisite the existence of a light modulator technology compatible with logic circuit technology that will enable the unconstrained placement of a large number of logic-driven light sources on a die or wafer. Ferroelectric liquid crystals (FLC's) possess low switching energies and driving voltages in the surface-stabilized cell configuration and are therefore well-matched to the output characteristics of silicon CMOS logic gates. Rapid development of faster materials is underway, with switching times of 3 microseconds at 15V/micron available now. Within the last six months a procedure that incorporates fabrication of most components of a reflective surface-stabilized FLC light modulator cell into an unmodified bulk CMOS process flow has been successfully prototyped.

In digital optical parallel processing, the magneto-optic spatial light modulator (MOSLM) is a relatively low-cost, small-pixel optical parallel input device that is capable of binary amplitude or binary phase modulation of each pixel as is needed in a general content-addressable memory processor as described above. However, potential problems associated with its use have included: lack of a design for a general minicomputer interface, lack of interactive data-page file generation software, lack of ability to switch single pixels in some cases, lack of remote-from-computer operation capability, spurious pixel switching, and catastrophic failure due to thermal effects. The hardware and software components of a versatile minicomputer interface that has been successful in overcoming all of these problems have now been constructed at Georgia Tech.

Development of the rigorous coupled-wave analysis for analyzing grating diffraction in holographic content-addressable memory applications described has produced an unanticipated extra benefit. Recently, a systematic procedure using the effective index method and impedance matching was developed for the design of antireflection high spatial-frequency rectangular-groove gratings on lossy materials including high conductivity metals. The design procedure, in turn, can be used as a starting point to design antireflection metallic gratings with lower spatial-frequencies using rigorous coupled-wave analysis. These lower spatial-frequency gratings have the advantage of being easier to fabricate. A particular antireflection gold grating design (having a period of 1.0 μm, a filling factor of 50%, and a groove depth of 147.5 nm for use at a freespace

wavelength of 500 *nm*, normal incidence, and polarization parallel to the grooves) was fabricated and its diffraction characteristics experimentally measured. The grating indeed showed remarkable very nearly zero specular reflection in the blue region of the spectrum. Unlike previously reported antireflection "anomalies," the effect is broadband occurring over a broad range of wavelengths and angles of incidence, and for both orthogonal polarizations. This work clearly shows that the systematic design of zero specular reflection grating surfaces is possible.

Present-day semiconductor devices are rapidly approaching a fundamental size limitation. At dimensions of less than about a quarter of a micron, "troublesome" electron quantum wave effects start to dominate device characteristics making further reduction in their size impossible. These quantum wave interference effects have been experimentally observed by investigators at numerous laboratories worldwide. These wavae effects limit the ultimate speed of conventional devices. Further advances require fundamentally new device concepts. A new technology that utilizes the "troublesome" quantum wave effects to advantage and allows scaling of devices down to the nanometer level is needed.

Digital Optical Content-Addressable Memory Processing

- *Content-Addressable Memory Processing Concept*

Compared to a location-addressable memory, less storage is generally required when a truth-table is implemented using a content-addressable memory (CAM). The unity-result truth-tables for each output bit are stored in the CAM. A unity-result or a null-result truth-table may be constructed from those combinations of inputs which cause a particular output bit to be a "one" or a "zero," respectively. The unity-result truth-table represents the canonical sum-of-products expression for the logical function corresponding to each output bit. Likewise, the null-result truth-table represents the canonical product-of-sums expression for each output bit. In a content-addressable memory, inputs are compared with the stored tables and detected matches cause the appropriate output bits to be a "one" (if unity-result truth-table is stored) or a "zero" (if null-result truth-table is stored). The stored input words are then function minterms in the sum-of-products expression (unity-result truth-table) or the function maxterms in the product-of-sums expression (null-result truth-table). In the optical holographic implementation of content-addressable memory, the number of function minterms represents the number of holograms that need to be stored in the system [1]. Using thick holographic recording media, such as photorefractive lithium niobate, the number of stored holograms can be on the order of a thousand [2].

In a location-addressable memory, the input word(s) represent the address of the data. The output is the data at that address. However, in a content-addressable memory, the input word(s) represent the data. The output is all of the addresses at which those data occur. The addresses stored represent the word patterns that make

each bit of the answer a "one." Thus, it is determined whether each bit of the answer is a "one" or not and, therefore, the answer is uniquely specified.

The implementation of a content-addressable memory for truth-table look-up processing required parallel pattern recognition. Two ways of achieving this parallel pattern recognition are: analog correlative pattern recognition and logic-based digital pattern recognition.

- *Correlative Pattern Recognition*

Optical parallel analog correlative pattern recognition is achieved by constructing an optical matched filter. This was first accomplished by Vander Lugt [3]. In this method, the complex conjugate of the Fourier transform of the reference pattern is produced by recording an off-axis Fourier transform hologram of the pattern. The complex conjugate of the transform of the reference pattern and the Fourier transform of the input pattern are multiplied in the Fourier plane in a standard 4-f optical processing configuration. The second Fourier transform lens then produces in parallel the cross-correlation between the input pattern and the reference pattern. If these patterns are the same, an auto-correlation is produced. The presence of the auto-correlation can be identified by threshold detection since its peak value is greater than the peak value of any cross-correlation. However, the difference in levels can be very small, resulting in significant false alarm rates and making this approach impractical.

- *Logic-Based Word Recognition*

Since either Boolean similarity logic functions (such as EXCLUSIVE OR or EXCLUSIVE NOR) or Boolean uniqueness logic functions (such as NOR and NAND) can be used to find matches, they can thus be used to implement digital pattern recognition. The use of the EXCLUSIVE OR and NAND functions are described in this paper. Both of these logic functions can be implemented in parallel through the physical process of complex wavefront addition and subtraction. An early demonstration of optical coherent wavefront addition and subtraction was done by Gabor *et al.* [4] and it has been widely used in many optical systems since that time. The digital pattern recognition is used, in turn, to perform truth-table look-up processing and thus optical digital parallel data processing. Unlike correlative pattern recognition, this logic-based pattern recognition has the inherent advantages of arbitrary precision, flexibility, and noise immunity that are consequences of its digital nature.

- *EXCLUSIVE OR Operation*

Gabor *et al.* [4] demonstrated "image synthesis" by multiple holographic recordings in photographic film. They presented an example of image subtraction in this pioneer paper. In a similar manner, "coherent selective erasure" was shown in lithium niobate

by Huignard *et al.* [5]. They pointed out that the complex amplitude subtraction is equivalent to the logical EXCLUSIVE OR operation. There were approximately 10,000 bits per page in their demonstration and thus 10,000 operations were being performed in parallel. The EXCLUSIVE OR result in both of these works was recorded during the process (in the film or in the crystal). Thus, the prestored data was altered by the operation and was no longer available for further processing. In the work described here [1,6], the prestored data is left unaltered and is continuously available for the processing of new data as it enters the system.

A schematic diagram of an optical system that performs two-dimensional EXCLUSIVE OR processing is shown in Figure 1. Transparent and opaque apertures in the data mask represent binary "ones" and "zeros," respectively. During the recording step, the Fourier transform of the spatial distribution of the amplitude at the input mask is produced in an electrooptic crystal, such as $LiNbO_3$, and the interference pattern between this Fourier transform and a reference wave is recorded as a thick hologram by the photorefractive effect [7]. During the processing step, both the reconstructed image of the recorded mask and the direct image of a new input mask are simultaneously produced on a photodetector array. The relative amplitude of the two images is controlled by an amplitude modulator so the light from a binary "one" in either image has the same amplitude. Also, using a phase modulator, the phase difference between the two image wavefronts is adjusted to be 180°. As a result, at locations in the image plane where the bits from the two data pages are equal, the image will be dark, while at locations where the bits are unequal, it will be bright. If the dark regions and the bright regions are considered to represent binary "zeros" and "ones," respectively, then bit-by-bit EXCLUSIVE OR processing between the two data pages is achieved at the photodetector array.

The EXCLUSIVE OR processor described above is an excellent candidate for word/signature detection. To accomplish this type of detection, all the reference words are recorded (e.g. as different rows of the data page). To see if an input word is among the recorded patterns, the word is repeated at the input page composer (e.g. in all rows). If the input word matches any of the stored patterns a row of zero elements will be present in the matrix at the detector array and can be detected electronically. This property can be used to implement numerical operations optically. For example, consider the addition of a pair of two-bit words and an input carry. This operation results in a two-bit output word and an output carry. For each output bit, all the input patterns that provide "one" at that bit location may be grouped together and recorded. An EXCLUSIVE OR process between the input pattern and these recorded patterns would determine whether each output bit is "one" or "zero."

The above EXCLUSIVE OR processor can be extended to operate on a number of input patterns simultaneously. This can be achieved by stacking the input words in a direction perpendicular to the plane of the recording beams to take advantage of the small angular selectivity in that direction [8]. Using different angles for the reference beam, it is also possible to record data pages that correspond to different operations at

the same location in the crystal. Then, these operations can be performed on the input data by simply changing the angle of the reference beam.

- *NAND Processing*

The use of holography in performing Boolean logic operations was first described by Preston [9]. He recorded Fourier holograms that upon reconstruction, produced specific phase shifts between the input light waves. The logic states "0" and "1" at the input of the system were modeled as pinhole apertures that introduced phase shifts of 0^o and 180^o in the light, respectively. During the process, the light beams that passed through these apertures experienced specific phase shifts as they were diffracted by the hologram. As a result of the complex amplitude addition of the diffracted waves, the required logic output were obtained on the detector array in the from of bright and dark spots, representing "ones" and "zeros," respectively. Using this technique, Preston performed the logic operations EXCLUSIVE NOR and EXCLUSIVE OR on two input bits and described how the operation OR could be obtained. From these, the other logic operations can all be constructed. The NAND processor work of this proposal [1,10], is an extension of Preston's method to the construction of a content-addressable memory capable of implementing any arithmetic operation.

A schematic diagram of a NAND-based optical processor is shown in Figure 2. The system is a $4f$-setup and functions as a content-addressable memory. During the recording step, for each output bit of the required function, all the input patterns that produce a "one" at that bit location are stored. At the processing step, the input pattern is compared with these stored reference patterns, and detected matches cause the appropriate output bits to be "one." Although any number system can be used with this type of processor, the use of residue number system has the advantage of requiring significantly fewer reference patterns.

Each reference pattern is recorded by using a phase-shifting line composer (a particular type of spatial light modulator) containing one more element than the number of bits in each pattern. The extra element is designated as a reference bit and is set to a phase shift of 0^o for all recordings. All the bits that are "ones" in the reference pattern are recorded with a phase of 180^o, all the bits that are "zeros" are recorded with a phase of 0^o (the same phase as the reference bit), and "don't-care" bits are not recorded (spatial light modulator does not pass light at those bit positions). The diffracted amplitudes associated with all the "one" and "zero" bits are equal. The diffracted amplitude associated with the holographic recording of the reference bit must be R times the diffracted amplitude of each "one" (or "zero") bit, where R is the total number of "ones" in the particular reference pattern being recorded. A complete reference pattern is recorded as described for a single position of the reference beam. The reference beam is then stepped to a new location, and a new reference pattern is entered into the spatial light modulator and a holographic recording made.

After all reference patterns are recorded, the phase-shifting line composer is replaced by an on-off amplitude line composer as shown in Fig. 2(b). The system is then able to function directly as a truth-table look-up processor. A binary data pattern is entered into the on-off amplitude line composer. For "one" bits in the input data pattern, the spatial light modulator is transparent, and for "zero" bits it is opaque. If (and only if) the input data pattern matches a prerecorded reference pattern, then, upon reconstruction, there will be wavefront cancellation in the direction of the particular reference beam that corresponds to that reference pattern. This is detected by an array of photodetectors with each position in the array corresponding to a different reference beam position during recording. Thus, each detector corresponds to a particular binary reference pattern, and when wavefront cancellation (a null) occurs at that detector, a match of the input data to that reference pattern has been found. Each recording may be thought of as producing an optical NAND gate. The phase shifts provided by the line composer during recording determine which data bits during processing will be presented to the NAND gate in complemented form (the "zeros" in the input data pattern) and which will be presented to the NAND gate in uncomplemented form (the "ones" in the input data pattern). During processing, the output of the NAND operation then occurs in the detector array plane at the reference beam position for that recording.

The NAND-based processor, like the EXCLUSIVE OR-based processor, may be extended to operate on many input data words simultaneously (in parallel). The set of holograms recorded as described above may be read out using multiple binary amplitude line composers displaced above and below the recording plane of incidence so that the fundamental Bragg condition is still satisfied.

- *Coding for Truth-Table Look-Up Processing*

Advances [11,12,13,14] in number representation, multi-level coding, large-scale truth-table reduction techniques together with the growing need for faster parallel computation have stimulated strong interest in table look-up implementations. It has been shown that in order to determine the practicality of truth-table look-up techniques for a particular operation, the sizes of the logically reduced tables must be determined. In turn, the sizes of the truth tables may be reduced by one or a combination of the following techniques: 1) using the residue number system, 2) using the modified signed-digit number system, 3) applying multi-level coding (ternary, etc.), and 4) applying logical minimization techniques. Furthermore, it has been shown that *depending on the operation being implemented*, each of these can produce very significant reductions.

In the last year, work completed at Georgia Tech [15] has provided the first complete comparison of number systems for a specific operation (8-bit addition) when all the steps of encoding, processing, and decoding are included. The input was two 8-bit binary numbers together with an input carry. The output is a full precision 9-bit binary sum. Three number representations were treated: binary, residue, and modi-

fied signed-digit. The numbers in all three representations were in binary-coded form throughout the processing. The critically important steps of encoding the numbers into the residue and modified signed-digit systems and then decoding the results back into direct binary were also treated using truth-table look-up methods. For the direct binary representation, a total of 2545 gates (2519 AND gates or holograms in holographic content-addressable memory implementation) are required. For the residue representation, a total of 1764 gates (1686 holograms) are required. For the modified signed-digit representation, a total of 4142 gates (4052 holograms) are required. Thus for 8-bit addition, using the residue number system allows this operation to be done with the least amount of hardware.

- *Hybrid VLSI-Optical Processor*

Considerable interest exists in the implementation of massively parallel systems tailored to problems having high inherent parallelism, that use optics to implement some or all of the tasks of interconnection, logic, and clocking [16,17,18]. We are presently exploring the design of digital systems based on two-dimensional arrays of electronic processing elements (PE's) produced at or near the wafer-scale level of integration, that are optically interconnected and synchronized by means of light modulators and detectors incorporated within the PE's and an external optical routing network. Provided a fast, reliable light modulator technology is developed that is compatible with silicon or GaAs VLSI circuit technology and signal levels, such systems will compete with all-electronic and all-optical computing systems in the regime of highly parallel and structured computation.

The shift-connected SIMD array is based on 3 micron bulk silicon CMOS technology. The PE, as shown in Figure 3, consists of a 1-bit wide ALU capable of full addition and other logic operations, a 64-bit static RAM, an optical detector, and an optical modulator. In addition, a shift register cell is included that can be read or written by the RAM. Shift register cells of all PE's are concatenated to provide the array with electrical upload and download capability. Optical detection is implemented with silicon photodiodes or phototransistors. Optical outputs are reflective FLC light modulator cells integrated on to the silicon chip. The approximate PE layout is shown in Figure 4. The 950 micron dimension permits an 8×8 array of PE's to be fabricated on the largest standard die size provided by the MOS Implementation Service (MOSIS). Clock and control are electrically broadcast to all PE's and provided through the package pins. The overall system is illustrated schematically in Figure 5. The modulator array is illuminated by a lenslet array provided by Corning Glass Works, which also collimates the beamlets exiting the modulators. The polarizing beamsplitter converts the polarization rotation of the FLC modulators to intensity modulation. Shifting of the output image is performed by programmable deflectors (galvanometers or acousto-optic devices) in the Fourier plane, and the data are then imaged onto the lenslets, which concentrate the light onto the detectors. An IBM PC XT effects control and data communication

with the PE array and controls the deflectors by means of a plug-in interface board.

- *Silicon-Compatible Spatial Light Modulator Technology*

The development of optically interconnected VLSI and wafer scale systems has as a prerequisite the existence of a light modulator technology compatible with logic circuit technology that will enable the unconstrained placement of a large number of logic-driven light sources on a die or wafer. Integrated lasers, multiple quantum well electro-absorption modulators, and organic electro-optic materials exhibit high speed but are insufficiently mature to permit system prototyping, especially in silicon. Ferroelectric liquid crystals (FLC's) possess low switching energies and driving voltages in the surface-stabilized cell configuration and are therefore well-matched to the output characteristics of silicon CMOS logic gates. Rapid development of faster materials is underway, with switching times of 3 microseconds at 15V/micron available now. A procedure has been developed [19] that incorporates fabrication of most components of a reflective surface-stabilized FLC light modulator cell into an unmodified bulk CMOS process flow. The reflective electrode and 1-1.5 micron thick spacer can be obtained by specifying features in the mask description of the die. Light sources can thus be placed in the same manner as transistors.

Prototype 8×8 arrays of FLC light modulator cells have been fabricated and evaluated. The cell structure is shown in Figure 6. A 1.3 micron layer of SiO_2 was deposited over a 120 *nm* layer of aluminum evaporated onto a silicon wafer. 100 micron $\times$ 100 micron windows on 1 mm centers were etched through the SiO_2, exposing the aluminum. A 25 *nm* aligning layer of silicon monoxide was deposited on both the silicon substrate and on the transparent cover electrode of indium tin oxide on glass by evaporation at a 60^o incident angle. The cell array was assembled and filled *in vacuo* with E. Merck ferroelectric smectic mixture ZLI-3489. Cells typically exhibited an intensity contrast ratio of 40:1.

- *Magneto-Optic Spatial Light Modulator Evaluation*

The magneto-optic spatial light modulator (MOSLM) [20,21] is a relatively new device capable of binary amplitude (on-off) or binary phase modulation. This type of operation makes it well suited for the coherent content-addressable memory processing described in the section *Content-Addressable Memory Processing Concept* earlier in this report. However, its performance in this type of coherent processing has yet to be determined. It is proposed to measure the performance of the MOSLM in this highly phase-sensitive type of processing.

The MOSLM consists of an array of square mesas (pixels) of magneto-optic iron garnet grown on a non-magnetic substrate. In the absence of an external magnetic field, each pixel is in one of three stable states: (1) magnetized with the magnetic vector parallel to light propagation, (2) magnetized with the magnetic vector anti-

parallel to light propagation, or (3) demagnetized. As linearly polarized light passes through a pixel in state "1" or "2" the angle of polarization is rotated counterclockwise or clockwise respectively. By placing an appropriately oriented polarizer at the output of the system, each pixel will appear either transparent or opaque depending on the state of the pixel. The operation of an MOSLM is illustrated in Figure 7.

The state of any selected pixel can be changed using an external bias coil and the conductors that lie between the rows and columns of the spatial light modulator matrix. First the pixel must be nucleated by pulsing current through the row and the column conductors corresponding to the desired pixel. The resulting magnetic field at the intersection of the conductors causes a domain wall to form in the desired pixel. This domain wall will spread until half of the pixel is in state "1" and the other half of the pixel is in state "2". This half and half stable state corresponds to the demagnetized state "3". To complete the switching, the pixel must be saturated by the presence of an external magnetic field supplied by the pulsing of the bias coil that encircles the device. This mode of operation lends itself to computer control for inputing specific patterns as required in logic information or alternatively to image inputing such as video information as might be required for recognition operations.

The utilization of a minicomputer to control individual pixels and input patterns into a MOSLM has been demonstrated at Georgia Tech [24]. This system provides a flexible controller that is essential in evaluating the operating characteristics of an MOSLM. The speed of the MOSLM is controlled primarily by the rate at which a pixel can be addressed. Pixel addressing times are of the order of 200 nanoseconds and frame rates to 140 per second are possible with even higher rates achievable using external cooling of the MOSLM [25]. A pixel space bandwidth product of 4×10^6 and a frame space bandwidth product of 10^6 to 10^{11} has been cited in the literature [21] for these devices. Thus high data rates and speeds are available with the MOSLM.

Coherent optical processing techniques are critically dependent on the amplitude and phase transmittance characteristics of an SLM. Laboratory characterization of an MOSLM [17] has shown these devices to be extremely stable in terms of both phase properties and amplitude transmittance. It is the laboratory configuration illustrated in Figure 8. The apparatus is essentially a Mach Zehnder interferometer configuration with the MOSLM inserted in one arm. The output from the MOSLM is magnified and imaged onto the face of the vidicon tube of an optical multichannel analyzer (OMA). The vidicon image is then scanned, digitized and stored in a 512×512 format by the OMA controller. Built in software routines are used to process the data and in particular provide intensity profile slices through the image. The phase stability of the MOSLM was evaluated by initiating various switching operations, storing the resultant image after each operation and comparing the resulting intensity profiles associated with a particular MOSLM pixel. The effects of repeated switching of the full MOSLM array, single pixel switching effects and effects of neighboring pixel switching on the phase stability would be evident by changes in the intensity profiles.

The laboratory apparatus has been designed to be capable of detecting phase shifts as small as 5 degrees. The limiting factor in the phase shift accuracy will result primarily from high frequency noise associated with the OMA. Any phase shifts due to switching effects will be crucially important since MOSLM must exhibit very good phase stability in order to be usable in coherent processing applications. Similar measurements have been made on the amplitude transmittance and its stability. Likewise, the centrally important contrast ratio will also be measured. Extremely high contrast ratios of 40,000:1 have been cited by the manufacturer [25]. Measurements of (1) the full-frame contrast ratio, (2) the within-frame contrast ratio, and (3) the within-pixel contrast ratio are being made.

- *Antireflection Surface-Relief Gratings*

Antireflection surfaces on lossy materials such as semiconductors and metals are useful for eliminating specular reflections. A periodic corrugation on the surface of a material can produce antireflection behavior. Such surface-relief diffraction gratings are known to absorb specific wavelength ranges of the incident radiation rather than to diffract them. Wood [26] reported this absorption effect in 1902. Later Wood [27] showed that this type of effect can occur when a higher-order diffracted wave is at cutoff. These types of power absorption effects are commonly called "anomalies" in the literature. They have been experimentally observed for both transverse electric (TE) polarization and transverse magnetic (TM) polarization. Furthermore, these "anomalies" have been extensively theoretically investigated. These predicted and observed "anomalies" generally occur abruptly in wavelength and/or angle of incidence but sometimes are broadband. Antireflection behavior also has been predicted theoretically [28,29] and observed experimentally [30,31] due to surface-relief gratings on dielectric surfaces.

A systematic design procedure using the effective index method and impedance matching has recently been developed by us [32] for antireflection high spatial-frequency rectangular-groove surface-relief gratings on lossy materials including high conductivity metals. These high spatial-frequency gratings have periods that are small compared to the incident wavelength and therefore all diffracted orders are cut off except the zero-order forward and backward diffracted waves. In this long wavelength limit, the results from the effective index method rapidly approach the exact results. The design procedure, in turn, can be used as a starting point to design antireflection metallic gratings with lower spatial-frequencies using rigorous coupled-wave analysis. These lower spatial-frequency gratings are therefore more easily and economically fabricated for applications in the visible and infrared regions of the spectrum. These larger period gratings, in general, have propagating higher diffracted orders present in addition to the zero-order waves. More degrees of freedom are available and additional constraints on the grating design are possible.

A particular antireflection gold grating designed by us (having a period of 1.0 μm, a filling factor of 50%, and a groove depth of 147.5 nm for use at a freespace wavelength of

500 nm, normal incidence, and polarization parallel to the grooves) was fabricated and its diffraction characteristics experimentally measured. The grating indeed showed very nearly zero specular reflection in the blue region of the spectrum. Unlike previously reported antireflection "anomalies," the effect is broadband occurring over a broad range of wavelengths and angles of incidence, and for both orthogonal polarizations. This work clearly showed that the systematic design of zero specular reflection grating surfaces is possible.

Underlying Physics

The basic premise of electron wave optics in semiconductors is that through the analogies forged between electromagnetics and quantum mechanics, new electron superlattice structures can be made which mimic corresponding devices in electromagnetic optics. Specifically, electron energy filters, impedance transformers, waveguides, etc., can be constructed using previously developed optical or microwave designs by mapping from optics to quantum mechanics. Analogous quantities between optics and quantum mechanics such as the index of refraction are used to transform optical or microwave device designs into semiconductor superlattice devices. Implicit in this approach is the assumption that an electron wave in a semiconductor behaves like an electromagnetic wave. This assumption is brought into question due to the rather dramatic difference between an electron and the corresponding electromagnetic quantum, the photon.

The successful realization of electron optical devices depends upon understanding the basic physical differences between electromagnetics and quantum mechanics which blur the otherwise simple correspondence between the two. The principle difference between optics and quantum mechanics is the nature of the basic quantum in each case. In electron optics, an electron is a spin 1/2, charged particle which behaves in a fundamentally different way from an uncharged, spin 1 photon. It is of course important to consider how the electron's spin and charge will effect the design of new superlattice structures. In addition, other effects such as scattering and time dependence, which are common to both optics and electron optics, must be considered. All in all, five basic physical effects must be addressed in any complete model of superlattice transport. Such a model must be time dependent and include the effect of the electron's charge and spin, the periodicity of the crystal, and an accurate treatment of electron scattering events. It is proposed to develop an understanding of the effect of each of these mechanisms on the transport properties and hence on the performance of semiconductor superlattice electron optic devices.

The fact that an electron is a charged, spin 1/2 particle greatly alters its behavior from that of a photon. The negative charge of the electron leads to a direct Coulomb interaction between any one electron and all others present in the device. The presence of the electron-electron interaction acts to reduce the transmittance of the superlattice [33]. In addition to the direct Coulomb interaction, the exchange interaction, which arises from the fact that an electron is a spin 1/2 particle, further alters the electron energy

eigenvalue and hence its transmission coefficient [34,35]. In this case, the exchange interaction reduces the effect of the direct Coulomb interaction. The effect of space charge due to depleted dopants near the contacts must also be included in a full analysis of a realistic device.

At first one might expect that electron wave optics in semiconductors is not possible since electrons are Fermions and as such cannot occupy the same state simultaneously. Therefore, it might seem that an electron waveguide or filter would be able to support only one electron at a time. This is, of course, not the case since a miniband of allowable energy states forms in the superlattice. Each miniband state can support an electron. The current is still restricted by the Pauli Principle but the maximum number of carriers which the waveguide or filter can contain depends upon the density of states in a fashion similar to the that of bulk material. Hence, these device structures should be capable of carrying significant current densities.

Another key assumption made in the electron optics formulation is that the electron behaves as a plane wave. As is well known from solid state theory, an electron in a crystal is most appropriately described by a Bloch wavefunction having the periodicity of the crystalline potential. The assumption that the electron can be modeled as a plane wave is then not necessarily correct since the electron's dynamics is modulated by a periodic potential. Of even greater importance is the assumption that the electron state can be treated using the effective mass approximation within the superlattice. The validity of the effective mass approximation is questionable in structures which consist of only a few monolayers of each material. Clearly, in a very thin layer it is questionable as to whether the electron "senses" the bulk material and as such can be treated as if it was in bulk material while within that layer. Alternative formulations of electron transport in thin layers which avoid the effective mass approximation either pursue the problem from a quantum mechanical transport perspective, i.e., solve the quantum Liouville Equation using Wigner functions [36,37], path integrals [38] or via the density matrix [39] or use a wavefunction derived from the tight-binding method [40,41] and proceed by solving the Schroedinger Equation. It is further necessary to formulate the problem taking into account time dependence. Time dependence arises naturally in the quantum transport models but their complexity makes them somewhat difficult to apply. Alternatively, the time-dependent Schroedinger equation can be solved numerically [42,43,44,45]. The carrier tunneling time can be determined naturally from the time-dependent formulations and the behavior of the charge as it moves through the structure can be clearly examined. In addition, the time dependent formulation has the important advantage in that it is not necessary to assume that the electron state is a plane wave. The electron can be treated as a wave packet which provides a more realistic portrayal of the carrier. Recent calculations have shown [42,43] that the use of a wave packet alters the width of the resonant state and hence the lifetime of the state. Though the time-dependent Schroedinger equation is easier to solve than any of the quantum transport models, it implicitly neglects dissipation effects such as electron-phonon scatterings. In many of the superlattice structures of interest in "electron optics"

the electron mean free path between collisions is larger than the dimensions of the device particularly if the structure is designed so as to involve only electrons within the Γ valley. Nevertheless, owing to the stochastic aspect of electron-phonon scattering processes, scatterings will occur within small superlattice structures. It is necessary then to add correctly these effects.

From the above discussion, it is clear that a more detailed model of electron transport in superlattice structures is necessary from that used in previous studies [46]. It is proposed to develop a more refined model using a numerical solution of the time-dependent Schroedinger equation. Specifically, the new model will include dissipation effects following the approach of Yalabik *et al.* [43] as well as treat the electron as a wave packet. The Schroedinger equation will be solved self-consistently with the Poisson equation in order to take into account properly the charge of the electron. Being that the vast majority of devices of interest are only two terminal structures, a one dimensional analysis will provide a reasonable model. In this way, the new model will address the effects of the electron charge, finite coherence length, time dependence and limitations of the plane wave model.

The Poisson and time-dependent Schroedinger equations will be written using the finite difference method. The boundary conditions at the contacts will be treated following the approach of Mains and Haddad [45]. In that work they assume that the reflected wave at the initial contact can be treated as having a spatially varying amplitude. If it is further assumed that this variation is linear, the initial solution of the Schroedinger equation can be readily approximated. In this way, a recursive solution can be generated.

Dissipation effects can be included by making two assumptions. The first assumption is that the occupation numbers of electrons at each time step can be updated through use of time-dependent perturbation theory [43],

$$\partial|\alpha_k|^2/\partial t \sim (2\pi/\hbar)\sum_{k'}(|V_{kk'}|^2|\alpha_{k'}|^2 - |V_{k'k}|^2|\alpha_k|^2)\,\delta[E(k) - E(k') \pm \hbar\omega] \qquad (1)$$

where α_k is the amplitude of the k component of the wave function, $V_{kk'}$ is the matrix element of the electron-phonon scattering and the delta function ensures energy conservation in the process. The above equation gives the rate at which an electron will transfer out of an initial state, k, into some final state, k'. Though the above equation does not hold rigorously over short time intervals, like those time steps used in the simulation, it nevertheless provides a correct equilibrium distribution as well as numerical stability [43]. The second approximation is that the phonons act to alter randomly the phase of the electrons in accordance with a random noise term, $\xi_k(t)$

$$\phi_k \sim \xi_k(t) \qquad (2)$$

where ϕ_k is the phase of the k component of the wave function. This term leads to an expression for the electron self-energy. The advantage of the above approximations is that they enable a means of treating dissipation effects within the context of the Schroedinger equation and yield the correct field theoretical results in equilibrium.

Through the inclusion of the above details into a more refined model, it will be possible to better design and optimize new "electron optic" devices.

Device Modeling and Software Development

A quantitative parallelism exists between electron waves in quantum mechanics and optical waves in electromagnetics [47,48]. The analogies between electron waves and optical waves can be expressed with the introduction of two electron wave refractive indices: one for phase effects, $n_e(phase)$ (which is proportional to the square root of the product of the electron kinetic energy and the elctron effective mass) and one for amplitude effects, $n_e(amplitude)$ (which is proportional to the square root of the ratio of the electron kinetic energy and the electron effective mass). Due to the analogies between electron quantum mechanical waves and electromagnetic waves, existing optical design and analysis techniques can be used for the analysis, modeling, and design of novel semiconductor quantum wave devices. In order for the quantum wave effects to prevail, these new devices must have dimensions smaller than the electron coherence length. For these nanometer scale devices special considerations have to be taken into account due to spatial quantization, carrier heating, phonon scattering, space-charge effects, and others. Some of the underlying physical problems have already been discussed in the previous section.

Device modeling and simulation plays an essential role in the design and understanding of practical quantum wave devices. Device modeling is necessary to evaluate the expected performance of these devices and to design novel devices. Models and the corresponding software are being developed for the following devices which can be the basic components of more sophisticated structures.

(a) Unbiased superlattice: computation of the total transmittance and reflectance, with the same or dissimilar input and output regions.

(b) Biased superlattice: calculation of the total transmittance and reflectance using Airy functions and a stair-step approximation. In addition, a sensitivity analysis will be performed to examine the performance of these structures in the presence of fabrication variations. Models that can take into account carrier heating will be also pursued.

(c) Unbiased and biased superlattices: models including the nonparabolicity of the conduction band.

(d) Unbiased and biased superlattices: models including the space charge effects solving iteratively the Poisson and Schroedinger equations for self consistency.

(e) Unbiased and biased superlattices: models taking into account the electron-electron interaction.

(f) Unbiased and biased superlattices: models taking into account phonon scattering.

(g) Electron wave slab waveguides: mode characterization and computation, current flow calculation, and evaluation of density of states.

(h) Modeling of one-, two-, and three-dimensional electron wave devices like electron diffraction gratings, electron beamsplitters, impedance transformers between dissimilar epitaxial layers, mirrors.

(i) Modeling of integrated guided-electron wave devices that are analogous to integrated optics devices such as channel waveguides.

First-order models have been developed for both unbiased and biased superlattices for the design of Fabry-Perot type filters [46,48,49], and for electron-wave slab waveguides [50,51]. These models are all based on the rigorous solution of the Schroedinger's equation

$$\nabla^2\psi + (2m^*/\hbar)(E - V)\psi = 0 \ , \tag{3}$$

where ψ, E, V, and m^* are the electron wavefunction, the electron total energy, the electron potential energy, and the electron effective mass, respectively ($\hbar$ being the Planck's constant h divided by 2π). A summary of these first-order models will be presented in the next subsections.

- *Unbiased and Biased Superlattices*

The geometric configuration of a biased superlattice is given in Fig. 9. For the case that the applied bias potential energy, V_{bias}, is zero the structure corresponds to an unbiased superlattice. For the design of an electron filter (for example a bandpass filter) the number of the superlattice layers, their compositions and thicknesses have to be found.

For the design of semiconductor superlattice electron filters (unbiased superlattices), and filter/emitters (biased superlattices), it is important to calculate the total transmittance and reflectance of the device. If the electron coherence length is larger than the dimensions of the device then scattering can be neglected. In addition, by neglecting electron-electron interactions, space charge effects and carrier heating, the transmittance and the reflectance of the device can be found by successively solving Eq. (3) for each layer and using the boundary conditions [ψ and $(1/m^*)\psi'$ must be continuous at the layer interfaces]. Using this first-order model the transmittance t and the reflectance r of the unbiased superlattice can be found in the form [46,48]

$$\begin{pmatrix} 1 \\ r \end{pmatrix} = \prod_{m=1}^{M} \frac{1}{t_{e,m}} \begin{pmatrix} 1 & r_{e,m} \\ r_{e,m} & 1 \end{pmatrix} \begin{pmatrix} \exp(jk_m d_m \cos\theta_m) & 0 \\ 0 & \exp(-jk_m d_m \cos\theta_m) \end{pmatrix} \times \frac{1}{t_{e,M+1}} \begin{pmatrix} 1 & r_{e,M+1} \\ r_{e,M+1} & 1 \end{pmatrix} \begin{pmatrix} t \\ 0 \end{pmatrix} , \tag{4}$$

where $r_{e,m}$ and $t_{e,m}$ are the reflection and transmission coefficients between the $m-1$ and the m layer [16], k_m is the electron wavevector in the m-th layer, d_m is the thickness of the m-th layer, and M is the total number of layers of the superlattice. Equations of the form of Eqs. (4) have been widely used over many years for the design of thin film optical coatings and filters [52]. In case of a biased superlattice the solution of Eq. (3) is more complicated due to the linearly varying bias potential energy. In this case the transmittance and the reflectance are given by [49,53]

$$\begin{pmatrix} 1 \\ r \end{pmatrix} = \frac{1}{2jK_1} \begin{pmatrix} jK_1 & -1 \\ jK_1 & 1 \end{pmatrix} [S_1(z_0)][S_1(z_1)]^{-1}[S_2(z_1)][S_2(z_2)]^{-1}[S_3(z_2)]\cdots$$
$$\cdots[S_{M-1}(z_{M-1})]^{-1}[S_M(z_{M-1})][S_M(z_M)]^{-1} \begin{pmatrix} 1 & 1 \\ -jK_2 & jK_2 \end{pmatrix} \begin{pmatrix} t \\ 0 \end{pmatrix} , \qquad (5)$$

where

$$[S_j(z)] = \begin{pmatrix} Ai[-\rho_j(z)] & Bi[-\rho_j(z)] \\ (1/M_j^*)^{2/3}\, Ai'[-\rho_j(z)] & (1/M_j^*)^{2/3}\, Bi'[-\rho_j(z)] \end{pmatrix} , \qquad (6)$$

and where $M_j^* = m_j^*/m_0$, $K_1 = [(2m_0/\hbar^2)^{1/3}(E-V_0-V_{bias})/M_0^*]^{1/2}(L/V_{bias})^{1/3}$ and $K_2 = [(2m_0/\hbar^2)^{1/3}(E-V_0)/M_{M+1}^*]^{1/2}(L/V_{bias})^{1/3}$, $\rho_j(z) = (2m_j^*V_{bias}/\hbar^2L)^{1/3}[z + (E - V_{bias} - V_j)L/V_{bias}]$, and Ai' and Bi' are the first derivatives of the Airy, and complimentary Airy functions, respectively.

For designing Fabry-Perot type electron energy bandpass filters and filter/emitters the thickness of the superlattice layers has to be an odd multiple of a quarter electron wavelength or odd multiple of a half electron wavelength [46,49]. This requirement can be expressed by the equation

$$\int_{z_{j-1}}^{z_j} k_j(z)dz = \int_{z_{j-1}}^{z_j} (1/\hbar)\{2m_j^*[E - V_j(z)]\}^{1/2}dz = (2q_j - 1)\pi/2 \ , \qquad (7)$$

where q_j is an integer (and $\pi/2$ is replaced by π when half electron wavelength is needed). In the zero bias case the electron wavelength is constant inside the layer and Eq. (7) reduces to $d_j = (2q_j-1)(\lambda_j/4)$ (where λ_j is the electron wavelength in layer j) which is a very common equation in optical thin films [52]. However, due to the ultra small dimensions of these superlattices their thickness should be an integer multiple of a monolayer thickness. In addition, for practical materials like $Ga_{1-x}Al_xAs$ there is a restriction in the usable composition range (as shown in Fig. 9, $0 < x < x_{max} = 0.45$). For these type of materials the electron potential energy and the electron effective mass are related with the composition through the equations $V_j = Ax_j$ and $m_j^* = m_0(B + Cx_j)$, respectively (where A, B, C are material constants and m_0 is the free electron mass). Requiring each thickness d_j to be an integer multiple of a monolayer thickness, and using Eq. (7) the compositions and the thicknesses of the superlattice layers can be calculated through

procedures described in Refs. [46] and [49]. The response of a bandpass filter/emitter consisting of 9 layers is shown in Fig. 10. The characteristics of the filter are given in Ref. [49].

- *Electron Wave Slab Waveguides*

A quantum well in a semiconductor material system is capable of performing as an electron waveguide [50,51] for ballistic electrons due to the total internal reflection [48] of the electron wavefunction. The general case of an asymmetric slab electron waveguide is shown in Fig. 11. Using the waveguide terminology [54], the three regions are denoted substrate (s), film (f), and cover (c). When total internal reflection occurs the electron wavefunction decays exponentially in the cover and substrate regions and at steady state all of the electron current is reflected back into the film region. The electron potential energy at the bottom of the quantum well (film) is V_f while the potential energies associated with the cover and the substrate are denoted by V_c and V_s respectively (Fig. 11a). The material system is assumed to be $Ga_{1-x}Al_xAs$ for the cover, the substrate, and the film regions. The compositions of the substrate, the film, and the cover, are x_s, x_f, and x_c, respectively.

For an infinite medium, the electron propagation constant is $\beta_i = [2m_i^*(E-V_i)]^{1/2}/\hbar$, where $i = c, f, s$. These three propagation constants are shown in Fig. 12. For a given total electron energy E, the propagation constant of a guided mode can not exceed β_f. If the propagation constant is larger than β_f this corresponds to the evanescent modes similar to those in a dielectric slab waveguide.

An electron guided wave mode can become cutoff by decreasing the electron energy to the lower-energy cutoff which occurs when the zig-zag angle becomes zero ($\theta = 0$ in Fig. 11b). The propagation constant β_ν ($\nu = 0, 1, 2, \cdots$) of the ν-th guided mode, M_ν, becomes zero at the lower-energy cutoff and the wave becomes a standing wave. In this sense, the lower-energy cutoff is like the cutoff in an electromagnetic hollow metallic waveguide with finite conductivity walls. As the electron energy is increased, an upper-energy cutoff will also occur. The upper-energy cutoff can be of three types: (1) cutoff to a substrate mode which is like the cutoff in an electromagnetic asymmetric dielectric waveguide with the substrate index higher than the cover index, (2) cutoff to a radiation mode which is like the cutoff in an electromagnetic symmetric dielectric waveguide, (3) cutoff to a cover mode which is like the cutoff in an electromagnetic asymmetric dielectric waveguide with the cover index higher than the substrate index. The type of the upper-energy cutoff that occurs depends on the material parameters. In Fig. 12, for example, the upper-energy cutoff will be to substrate modes since β_s occurs at a lower energy, in general, than does β_c. For a two-dimensional (x_w , z_w) quantum well guided electron wave (Fig. 11b), the wavefunction has a sinusoidal dependence in the z_w direction and this can be expressed as $\psi_\nu(x_w, z_w) = \psi_\nu \exp(j\beta_\nu z_w)$ where β_ν is the guided electron wave propagation constant. Using Schroedinger's equation [Eq. (3)] solution of the form $\psi_{\nu s} = A_s \exp(\gamma_s x_w)$ for the substrate, $\psi_{\nu f} = A_{f1} \exp(j\kappa_f x_w) + A_{f2} \exp(-j\kappa_f x_w)$

for the film, and $\psi_{\nu c} = A_c \exp(-\gamma_c(x_w - d)]$ for the cover are needed, where $\gamma_s^2 = \beta_\nu^2 - [(2m_s^*/\hbar^2)(E - V_s)]$, $\kappa_f^2 = [(2m_f^*/\hbar^2)(E - V_f)] - \beta_\nu^2$, and $\gamma_c^2 = \beta_\nu^2 - [(2m_c^*/\hbar^2)(E - V_c)]$. By applying the boundary conditions that ψ and $(1/m^*)d\psi/dx_w$ must be continuous across the cover-film and substrate-film boundaries, the dispersion equation is found to be

$$\kappa_f d - \tan^{-1}[(\gamma_s/m_s^*)/(\kappa_f/m_f^*)] - \tan^{-1}[(\gamma_c/m_c^*)/(\kappa_f/m_f^*)] = \nu\pi. \qquad (8)$$

For $V_f = 0$ and normal incidence ($\theta = 0$ and $\beta_\nu = 0$), this dispersion equation predicts, in the limit as V_c and V_s approach infinity, that $E = E_\nu = (\nu+1)^2\hbar^2\pi^2/2m_f^*d^2$ in agreement with the well-known one-dimensional infinite potential well results. The lower-energy cutoff electron energy is designated E_{Lco} and can be calculated from Eq. (8) for $\beta_\nu = 0$, by solving a transcendental equation [50,51]. The electron energy at which the upper-energy cutoff occurs is designated E_{Uco} and the condition of this type of cutoff (for the $Ga_{1-x}Al_xAs$ material system) can be derived from Eq. (8) for $\gamma_s = 0$ (the mode leaks into the substrate) through the solution of another transendental equation [50,51]. For a given set of material parameters, as the waveguide thickness is increased, a guided mode M_ν first starts to propagate at an energy $E = V_s$. This corresponds to the highest possible value of the cutoff energy for the lower-energy type cutoff as well as the lowest possible value of the cutoff energy for the higher-energy type cutoff. The cutoff thickness of the slab electron waveguide can be found from Eq. (8) for the M_ν mode by substituting $E = V_s$.

Semiconductor electron wave slab waveguides can perform as described in this section provided that ballistic transport can be achieved over sufficient distances and that the density of electrons is small enough to make electron-electron interactions negligible. The electron wave-guide model described is again a simple first-order model and higher-order models need to be investigated in order to take into account more physical effects like scattering, electron-electron interaction and others. Electron waveguides are potentially useful in high-speed electronic circuitry. They could also be a central component in the future electron guided-wave integrated circuits which could perform *optical like* processing functions like those of present-day integrated optical circuits.

Device Design

- *The Need for Quantum Wave Device Design*

Electron wave propagation at energies above the potential barriers can be described by the analogies that exist between electron waves in semiconductors and electromagnetic waves in dielectrics presented in the previous section. However, semiconductor quantum wave devices cannot simply be copies of electromagnetic (microwave and optical) designs. In the design of semiconductor quantum wave devices, the thicknesses of the superlattice layers are restricted to be integer multiples of the monolayer thickness. In addition, there is typically only a limited usable composition range available. Fur-

thermore, in the electron wave case the dispersion effects (refractive index as a function of energy) are very pronounced compared to typical electromagnetic situations. Systematic design procedures are clearly required.

- *Fabry-Perot Interference Filter Design*

As an example of the design considerations that are required, the case of a Fabry-Perot interference filter will be treated. This device would consist of a half-wavelength layer surrounded by quarter-wavelength stacks. In electromagnetic optics it would be called an all-dielectric Fabry-Perot filter. The overall pass kinetic energy of the filter (as measured in the material surrounding the filter) is simply $E_p - V_0$ where E_p is the total electron energy and V_0 is the potential energy in the surrounding regions. This is the pass kinetic energy that would be specified by the user and is thus the starting point in the design procedure. The pass wavelength is given by

$$(\lambda_p)_i = h/\{2m_0[-ACx_i^2 + (CE_p - AB)x_i + BE_p]\}^{1/2} \ , \quad i = 0,1,2 \ . \tag{9}$$

The thicknesses of the superlattice layers are designated d_i $(i = 1,2)$. These thicknesses must be integer multiples of the monolayer thicknesses, r_i. Furthermore, these thicknesses must also be odd multiples of a quarter wavelength as measured in these regions. These constraints may be expressed as

$$d_i = p_i r_i = (2q_i - 1)(\lambda_p)_i/4 \ , \quad i = 1,2 \tag{10}$$

where p_i is the integer number of monolayers for the i-th region and q_i is a positive integer $(q_i = 1,2,3,\ldots)$. Eliminating $(\lambda_p)_i$ between Eqs. (9) and (10) gives the following quadratic equation in the composition x_i

$$ACx_i^2 + (AB - CE_p)x_i + (h^2/32m_0)[(2q_i - 1)^2/p_i^2 r_i^2] - BE_p = 0 \ . \tag{11}$$

The solution for the composition x_i is

$$x_i = [-b \pm (b^2 - 4ac_i)^{1/2}]/2a \tag{12}$$

where $a = AC$, $b = AB - CE_p$, and $c_i = (h^2/32m_0)[(2q_i - 1)^2/p_i^2 r_i^2] - BE_p$. In order to design a superlattice interference filter, at least two solutions for x_i must be found in the range $0 \leq x_i \leq x_{max}$. The smallest value of x_i within this range will become x_1, the composition of the high index material. The value of p_i that produces x_1 becomes p_1, the number of monolayers of type 1 material used to make a quarter-wavelength layer. Similarly, the largest value of x_i within this range will become x_2, the composition of the low index material. The value of p_i that produces x_2 becomes p_2, the number of monolayers of type 2 material used to make a quarter-wavelength layer.

To illustrate the principles developed, an electron wave filter that is the counterpart of a multilayer quarter-wave stack thin film optical interference filter has been designed

as a variable bandgap and variable thickness semiconductor superlattice. The realization of this filter is also subject to the practical composition constraint and the constraint that the thicknesses be integer multiples of the monolayer thickness. It consists of electron quarter-wavelength layers of GaAs (6 monolayers thick) and $Ga_{0.55}Al_{0.45}As$ (9 monolayers thick) and a half-wavelength layer of GaAs (12 monolayers thick). The pass electron energy is 0.14eV and the passband is only 0.003eV (2.2% of pass energy). These results have been published in *Applied Physics Letters* [47].

- *Narrowband Filter/Emitter Design*

As an example of the design considerations that are required for a biased device, the case of a tunable filter/emitter as described in section B will be treated. A systematic procedure for designing a biased superlattice emitter/filter for a given output kinetic energy $[(KE)_{out} = V_{bias} + (KE)_{in}]$ is now described. The material system of the superlattice is taken to be $Ga_{1-x}Al_xAs$ with the surrounding regions being the same and having $V_0 = V_{max} = 0.3479\ eV$ corresponding to an aluminum composition of $x_0 = 0.45$. The characteristics of (unbiased) superlattice electron wave interference filters [37,38] may be used as a starting point in estimating the number of regions M and the number of monolayers i_M to be used in the emitter/filter. Following the example presented in Ref. 5, a nine-layer filter ($M = 9$) will be designed. The estimated number of monolayers needed is $i_M = 72$ (from Ref. 15). The device is to emit electrons with a kinetic energy of 0.20eV $[(KE)_{out} = 0.20eV]$. The input kinetic energy is taken to be $(KE)_{in} = 0.10$eV and thus the bias potential energy $V_{bias} = (KE)_{out} - (KE)_{in} = 0.10$eV. The parameters q_j are set equal to unity. Starting with the first layer, $i_{j-1} = 0$, the value of the monolayer index i_j is incremented $1, 2, 3, \ldots$ and Eq. (7) solved for x_j for each value of i_j. Since the $j = 1$ layer is to be a high electron refractive index layer, the resulting positive real value of x_j closest to zero is selected. The corresponding value of i_j then becomes the value of i_{j-1} for the next layer. Equation (7) is then solved again as i_j is further incremented. Since the $j = 2$ layer is to be a low refractive index layer, the resulting real value of x_j closest to, but less than, 0.45 is selected. This process is continued for all layers. For the half wavelength layer, $\pi/2$ in Eq. (7) is replaced by π. For the last ($j = M$) layer, the number of monolayers left may be too few or too many. The value of i_M must then be revised and the design process restarted at the first ($j = 1$) layer. This process is then repeated until the optimum thickness (corresponding to the value of x_M closest to zero) of the last region produces a total thickness in self-consistent agreement with the value of i_M used. The results for this nine-layer emitter/filter are shown in Table I. The total thickness is 71 monolayers ($L = 20.0692nm$).

Experimental Comparisons, Crystal Growth, Device Fabrication, and Device Testing

Electron wave quantum interference effects have been experimentally observed by numerous investigators worldwide. These effects have been noted with great interest but have not been explained *quantitatively* in terms of the electromagnetic analogies. Direct comparisons of theoretical calculations using the analogies described in the previous sections with experimental data will be crucial for establishing the validity of the quantum wave description that has been developed [14,15,16,17,18,19] and for verifying the modeling tools.

- *Existing Devices*

A wide (60 *nm*) quantum well of $In_{0.53}Ga_{0.47}As$ between two $In_{0.52}Al_{0.48}As$ barriers (7.3 *nm* thick) has been fabricated. The design includes a monoenergetic emitter of electrons which tunnel through the first barrier when the device is biased. In addition, the structure includes an undoped output region (40 *nm*) before the ohmic contact. The measured properties of the structure including the voltage-current characteristic of this device have been published [55] and have been submitted for publication [56]. For bias voltages up to 0.750 V, a series of resonant tunneling peaks are observed in the voltage-current characteristic. For voltages greater than approximately 0.750 V, the bias is sufficiently large so that the injected electrons pass over the second barrier. For bias voltages above 0.75 V traveling wave quantum interference effects are observed associated with the composite four-layer structure (barrier-well-barrier-output). It is proposed to analyze this structure using the quantum wave treatment of the previous sections in order to make an experimental-theoretical comparison.

A variable-period superlattice Coherent Hetero-Interfaces for Reflection and Penetration (CHIRP) device consisting of 60 $Ga_{1-x}Al_xAs$ layers has been fabricated by Nakagawa *et al.* [57,58]. The active region of this device is 655 monolayers thick and is again designed to have a monoenergetic emitter of electrons. The electron transmitance through this device is high but decreases with increasing bias due to low transmittance band becoming aligned in energy with the emitter energy. This gives rise to negative differential resistivity in the voltage-current characteristic. It is proposed to analyze this structure using the quantum wave treatment of the previous sections in order to make an experimental-theoretical comparison.

- *New Demonstration Quantum Wave Devices*

Presently existing molecular beam epitaxial equipment and techniques are being refined to attain the ultra-high purity requirements needed for the proposed quantum wave devices and guided electron wave integrated circuits. This equipment is housed in the Microelectronics Research Center. If further funding is obtained beyond JSEP and

internal support this equipment will be used to grow some quantum wave demonstration devices. Initially two devices are being considered for this effort: a two-terminal negative resistance diode and a three-terminal transistor-type device.

- *Two-Terminal Negative Differential Resistance Diode*

  A two-terminal semiconductor superlattice structure that can exhibit negative differential resistance is being considered. This diode would incorporate a monoenergetic electron emitter which may be one of various possible types (such as a doped region followed by a thin barrier through which electrons tunnel). The emitter is followed by an undoped superlattice (as shown in Fig. 13) that acts as an electron wave low-pass filter. Together they can produce what might be called a Quantum Electron Wave Interference Diode (QEWID). At zero bias, the electron transmittance (and thus the current through the diode) is high. As the bias is increased, the electron transmittance decreases, thus producing negative differential resistivity. The systematic design procedure for selecting the quantum well and barrier widths to be alternately high and low electron refractive indices and quarter (or half) of an electron wavelength in thickness that has previously been described will be used to design a low pass filter.

  This device would appear to be superior to the CHIRP device in that it is based on "optimum" electron optical design (quarter and half wavelength layers) rather than miniband theory for a superlattice with a large number of periods. The device described here is therefore much thinner and thus more appropriate for ballistic transport operation.

- *Three-Terminal Quantum Interference Transistor Device*

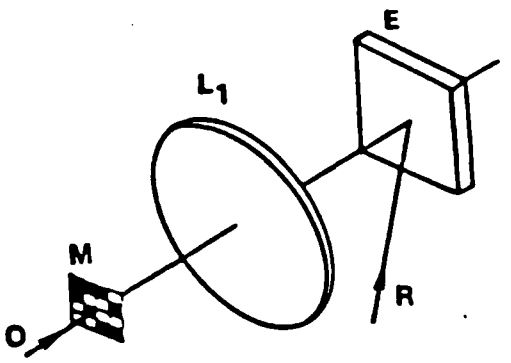
  A three-terminal semiconductor superlattice structure that can exhibit transistor action is being considered. The emitter-base "junction" is a superlattice (as shown in Fig. 14) that acts as a tunable electron wave interference filter/emitter [49]. The base-collector "junction" is a superlattice that acts as a narrowband tunable electron spectrum analyzer. Together they can produce what might be called a Quantum Electron Wave Interference Transistor (QEWIT). The systematic design procedure for selecting the quantum well and barrier widths to be alternately high and low electron refractive indices and quarter (or half) of an electron wavelength in thickness that has previously been developed [49] will be applied to the design of the emitter and the analyzer. The design of a practical quantum electron wave transistor consisting of layers of $Ga_{1-x}Al_xAs$ is presented. Emitter designs of 9 layers with the additional constraint that the electron energies be an optical phonon energy below the (111) L band minimum will be developed. Preliminary quantum electron wave interference transistor characteristics that have been calculated show the possiblity of implementing Boolean logical NOR, EXCLUSIVE OR,

AND, and EXCLUSIVE NOR operations as well as amplification.

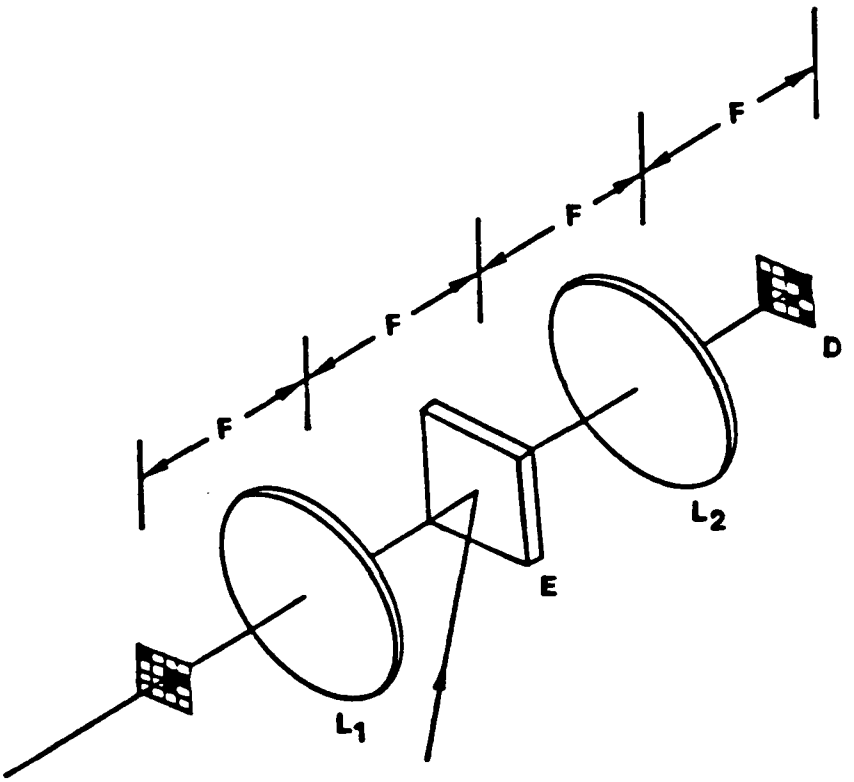
This device has some of the features of a Tunneling Hot Electron Transfer Amplifier (THETA) device [59] and the Resonant Hot Electron Transfer Amplifer (RHETA) device [60] except that the present device has a much more efficient emitter and the collector is a narrowband spectrum analyzer. That is, the collector is an "equal to" in energy spectrum analyzer in contrast to present day "equal to or greater than" type solid state electron energy analyzers [59,61,62,63].

Table I: Design Parameters of Electron Interference Filter/Emitter Consisting of Nine Layers in [H L H L HH L H L H] Configuration Surrounded by $Ga_{0.55}Al_{0.45}As$ and Designed to Emit 0.200 eV Electrons when Biased at 0.100 V.

Layer Number j	Layer Type	Starting Monolayer Number i_{j-1}	Ending Monolayer Number i_j	Number Monolayers Thick P_j	Aluminum Composition x_j	Unbiased Electron Potential Energy V_j	Normalised Effective Mass m_j^*/m_0
1	H	0	7	7	0.2222	0.1718	0.0854
2	L	7	16	9	0.4151	0.3209	0.1015
3	H	16	23	7	0.2663	0.2059	0.0891
4	L	23	32	9	0.4493	0.3473	0.1043
5	HH	32	44	12	0.0639	0.0494	0.0723
6	L	44	52	8	0.4364	0.3374	0.1032
7	H	52	58	6	0.1442	0.1115	0.0790
8	L	58	65	7	0.3748	0.2898	0.0981
9	H	65	71	6	0.1951	0.1508	0.0832



(a)



(b)

Figure 1: EXCLUSIVE OR processing configuration. (a) Recording hologram of a data pattern. (b) EXCLUSIVE OR processing between the recorded pattern and an input data pattern.

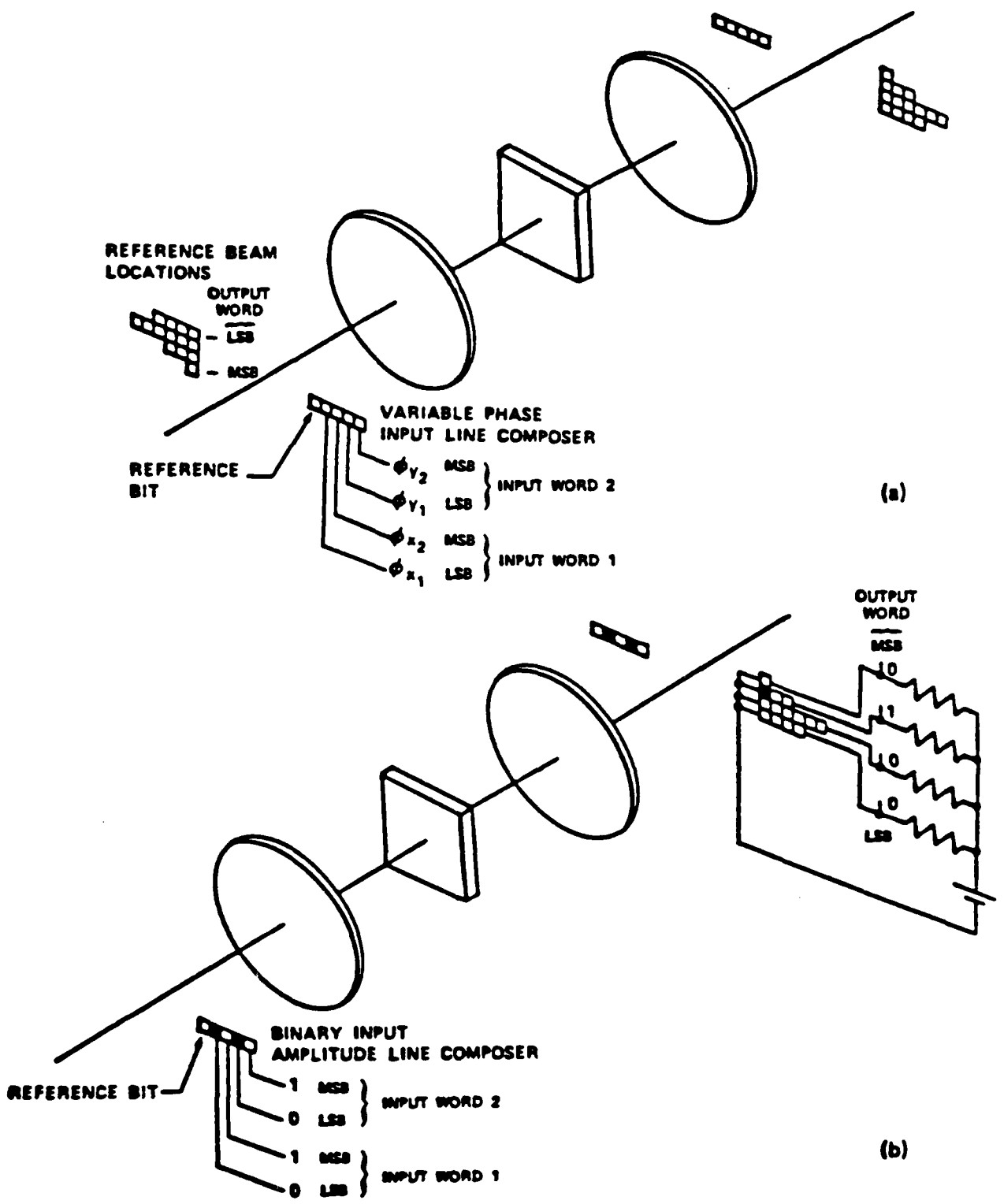


Figure 2: NAND-based numerical optical processor. (a) Recording the truth table holograms. (b) Example multiplication with the processor.

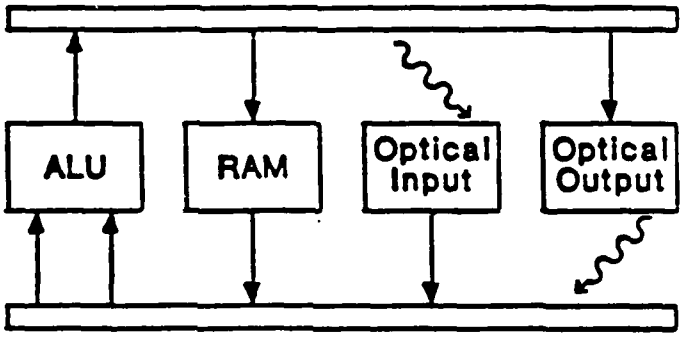


Figure 3: Processing element functional block diagram for shift-connected SIMD array implementation.

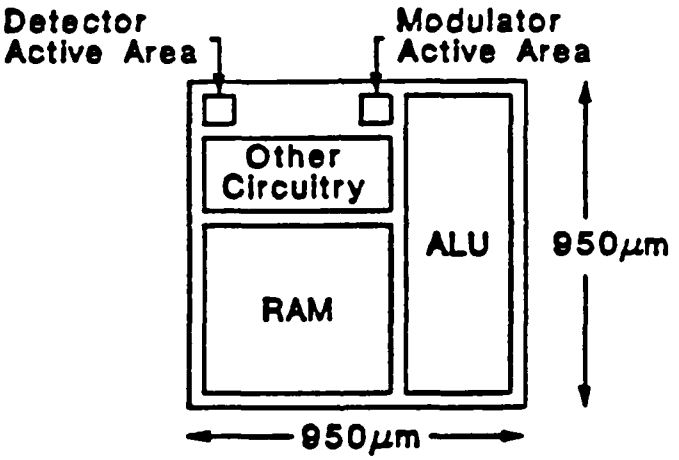


Figure 4: Processing element layout for prototype shift-connected SIMD array.

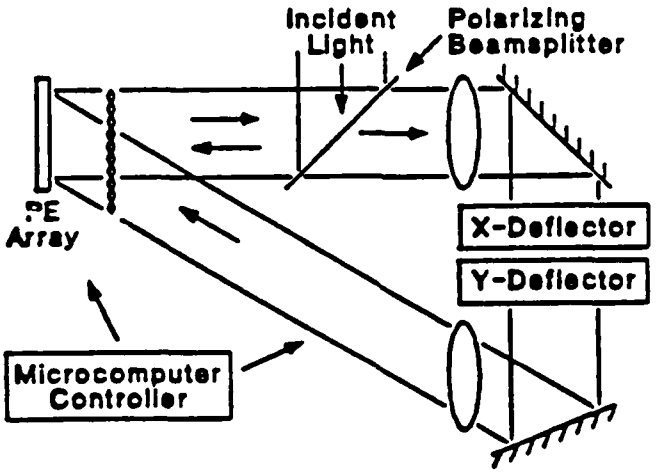


Figure 5: Prototpye shift-connected SIMD system design.

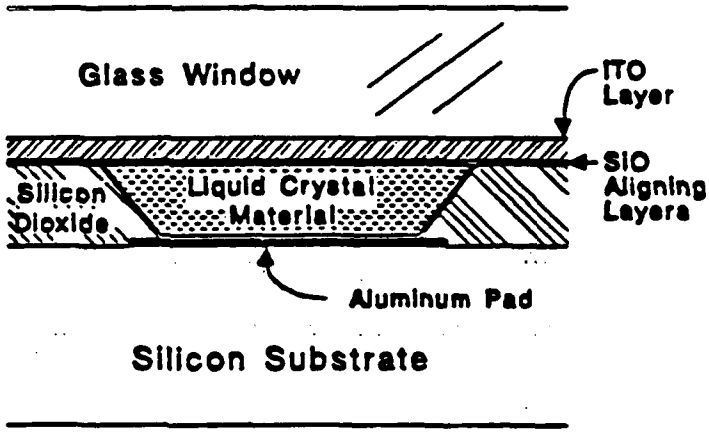


Figure 6: Ferroelectric liquid crystal/silicon reflective light modulator structure.

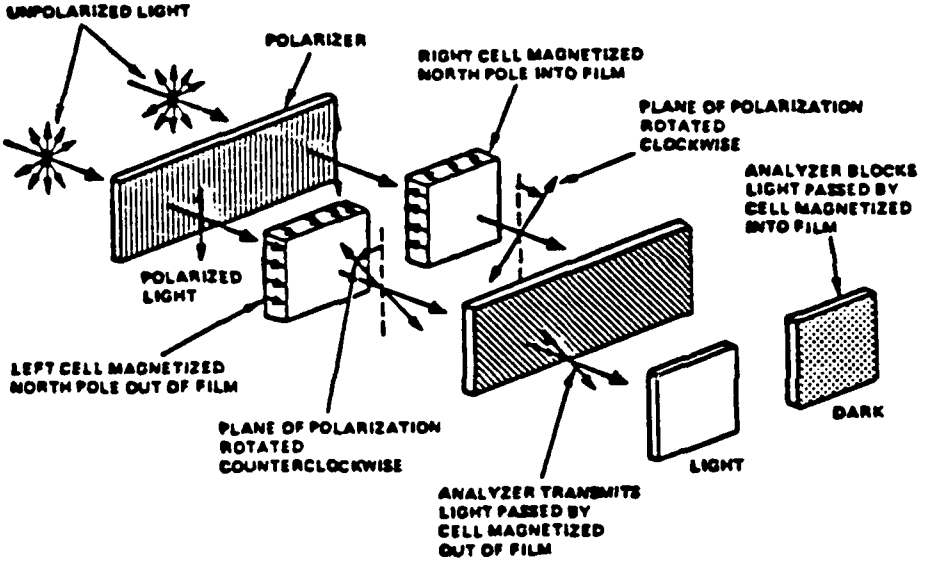


Figure 7: Operating configuration of a magneto-optic spatial light modulator. The device can be operated in a binary amplitude or a binary phase transmittance mode depending on analyzer orientation.

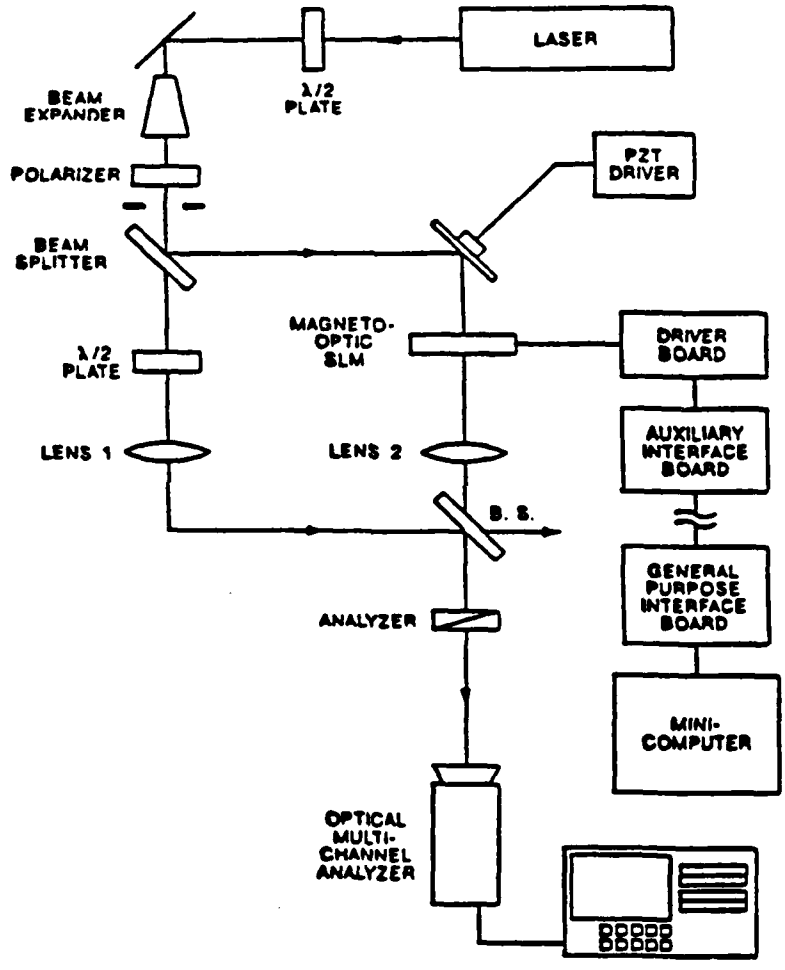


Figure 8: Laboratory configuration to be used to characterize the phase and amplitude transmittance of a magneto-optic spatial light modulator. The computer control system includes the sections composed of the driver board through the minicomputer.

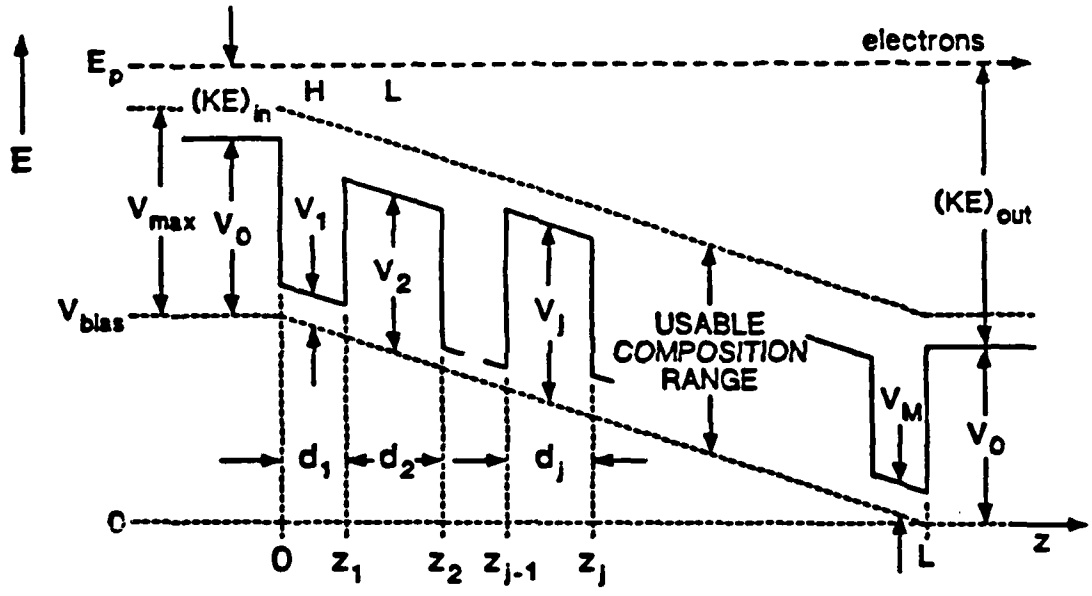


Figure 9: The conduction band energy diagram of a quantum well biased superlattice. The structure consists of M layers surrounded by bulk semiconductor material. The device becomes an unbiased superlattice when the applied bias potential energy, V_{bias}, is zero.

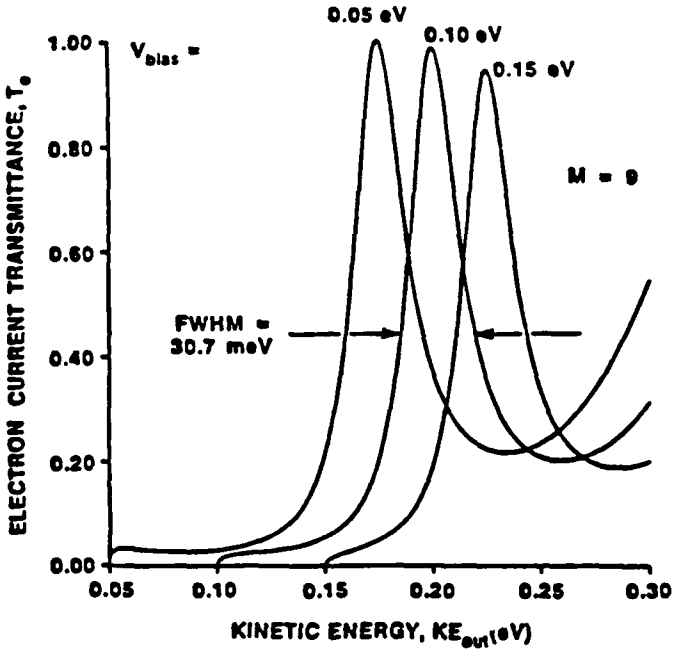


Figure 10: Electron current transmission of a 9 layer ($M = 9$) $Ga_{1-x}Al_xAs$ superlattice interference filter/emitter as a function of the electron output kinetic energy for $V_{bias} = 0.05$, 0.10, and 0.15eV. The spectral tuning with the bias is apparent.

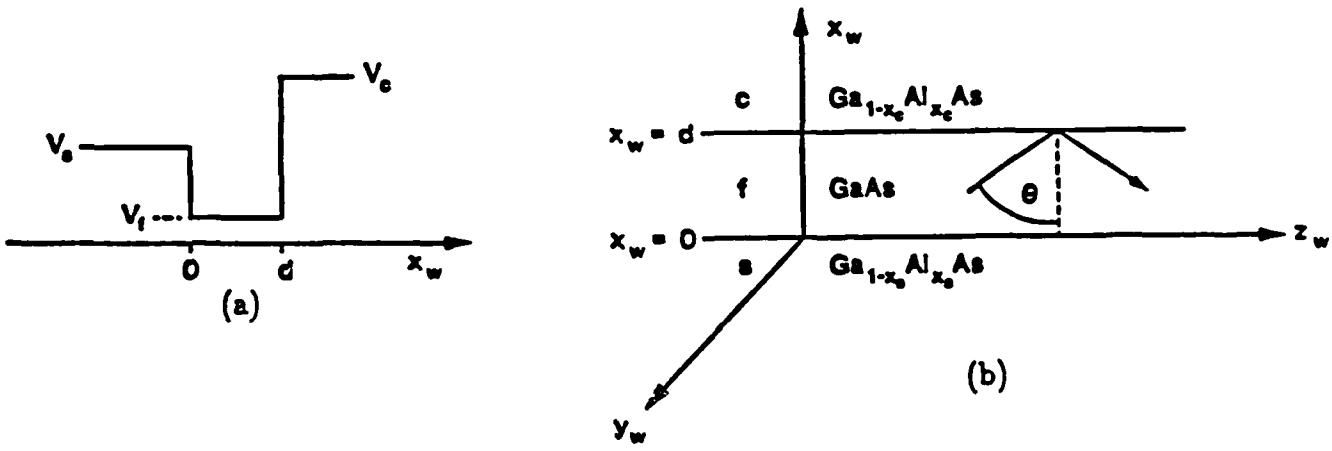


Figure 11: (a) Asymmetric quantum well showing electron potential energy in the three regions that in waveguide terminology are denoted as substrate (s), film (m), and cover (c). (b) Quantum well in three-dimensional waveguide coordinate system. A guided electron wave is composed of two plane-wave components each making an angle of incidence θ with respect to the waveguide walls ($x_w = 0,\ d$). The electron guided mode propagation direction is z_w.

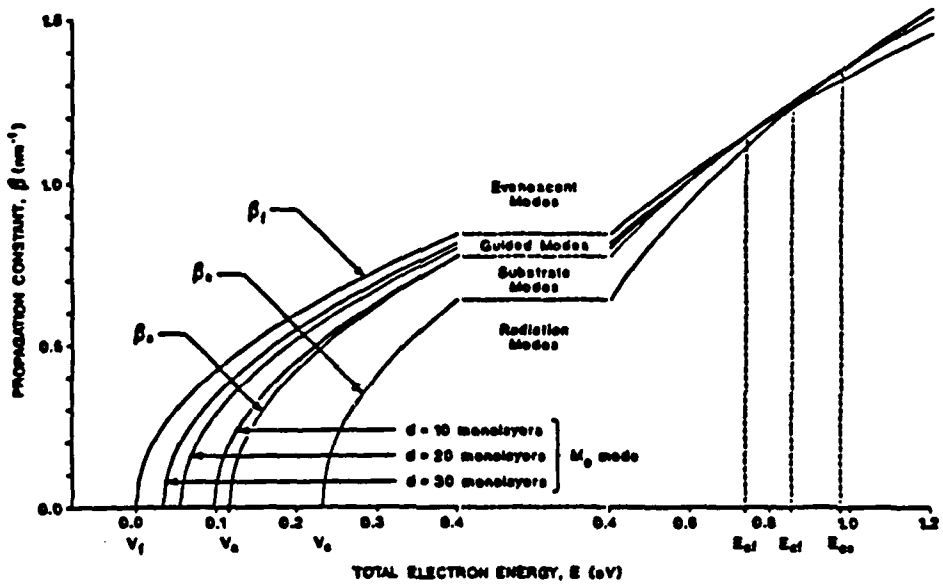


Figure 12: Electron guided mode propagation constant as a function of total electron energy showing the regions of evanescent modes, guided modes, substrate modes, and radiation modes for a quantum well waveguide composed of $Ga_{0.85}Al_{0.15}As$ (substrate), GaAs (film), and $Ga_{0.70}Al_{0.30}As$ (cover). The mode dispersion curves for the fundamental mode, M_0, are shown for film thicknesses of 10, 20, and 30 monolayers, and the monolayer thickness is 0.282665 nm. (Ref. [18])

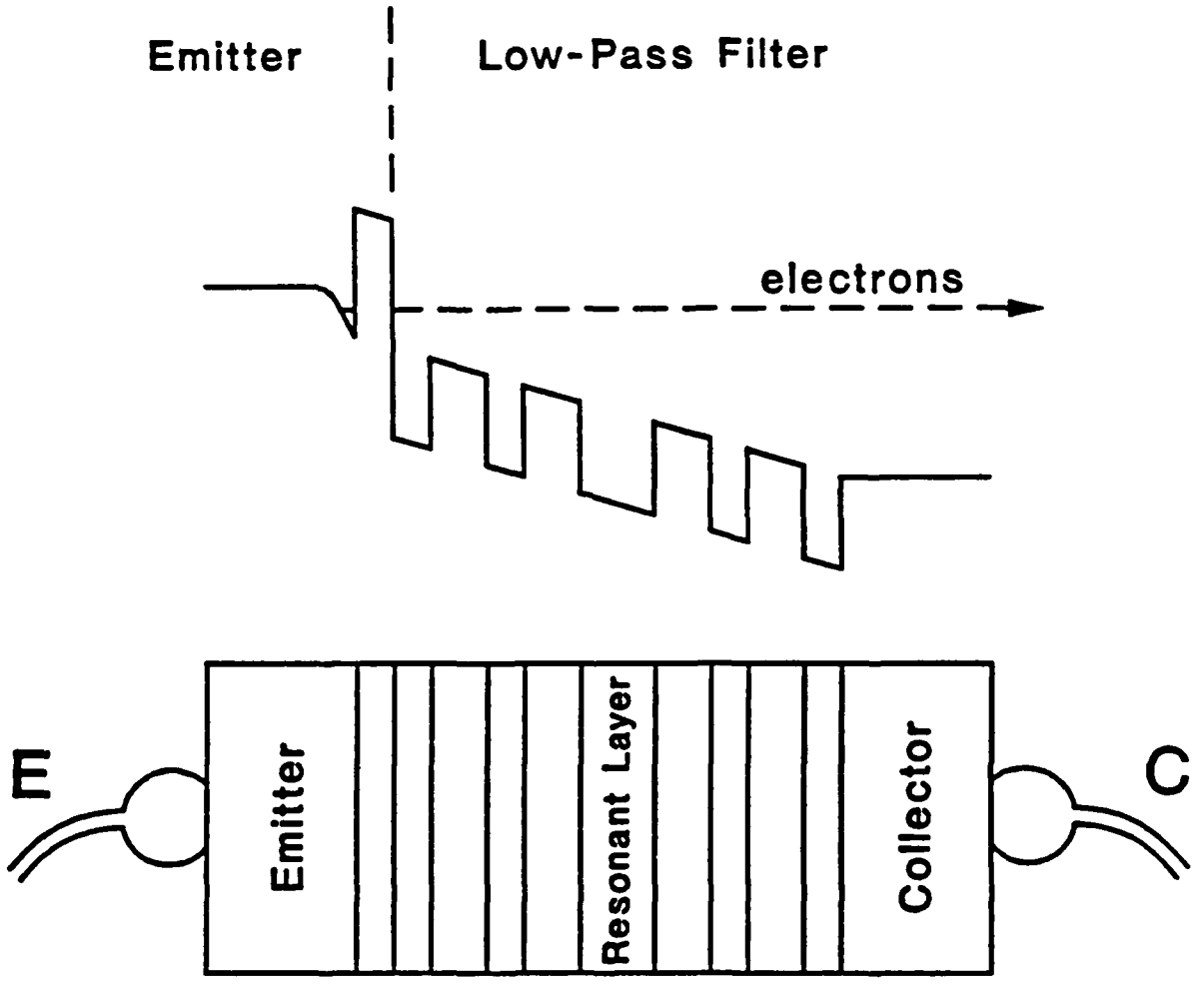


Figure 13: Two-terminal quantum electron wave interference negative resistance diode.

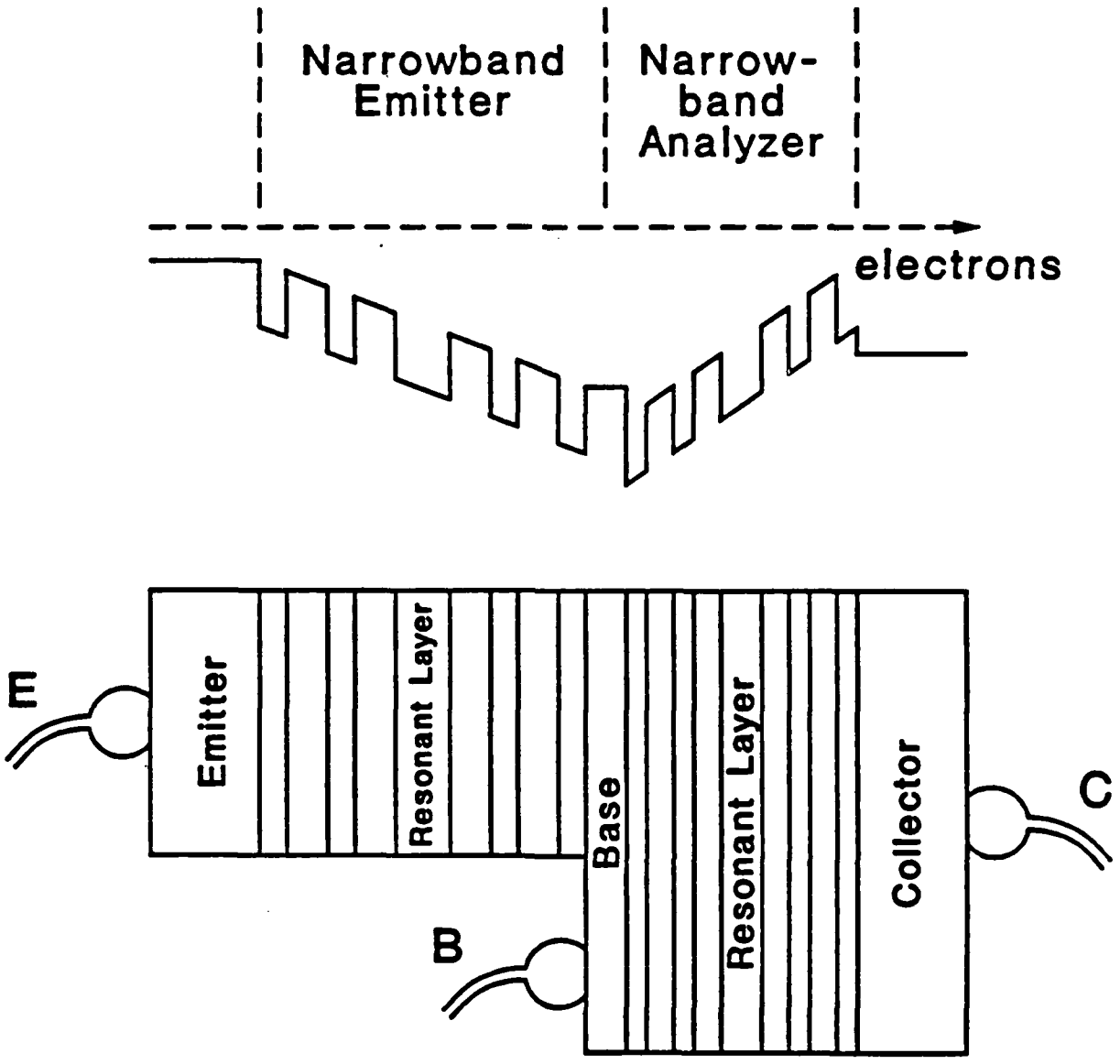


Figure 14: Three-terminal quantum interference transistor-type device.

PUBLICATIONS:

Journal Articles:

1. M. M. Mirsalehi, T. K. Gaylord, D. C. Fielder, and C. C. Guest, "Number representation effects in truth table look-up processing: 8-bit addition example," *Appl. Opt.*, vol. 28, pp. 1931-1939, May 15, 1989.

2. R. S. Weis and T. K. Gaylord, "Magneto-optic multilayered memory structure with a birefringent superstrate: A rigorous analysis," *Appl. Opt.*, vol. 28, pp. 1926-1929, May 15, 1989.

3. T. A. Maldonado and T. K. Gaylord, "Accurate method to determine the eigenstates of polarization in gyrotropic media," *Appl. Opt.*, vol. 28, pp. 2075-2086, June 1, 1989.

4. E. N. Glytsis and T. K. Gaylord, "Rigorous three-dimensional coupled-wave diffraction analysis of multiple superposed anisotropic gratings," *Appl. Opt.*, vol. 28, pp. 2401-2421, June 15, 1989.

5. M. M. Mirsalehi and T. K. Gaylord, "Analytic expressions for the sizes of logically minimized truth tables for binary addition and subtraction," *Appl. Opt.*, vol. 29, pp. xxx-xxx, 1990. (submitted)

6. N. F. Hartman and T. K. Gaylord, "Coherent optical charaacterization of magneto-optic spatial light modulator," *Appl. Opt.*, vol. 29, pp. xxxx-xxxx, 1990. (submitted)

7. M. R. Feldman, C. C. Guest, T. J. Drabik, and S. C. Esener, "A comparison between optical and electrical interconnects for fine grain processor arrays based on interconnect density capabilities," *Appl. Opt.*, vol. 29, pp. xxxx-xxxx, 1990. (accepted)

8. T. J. Drabik and M. R. Feldman, "Holographic optical time-multiplexed interconnection of hypercubes," *IEEE Trans. on Comput.*, vol. 39, pp. xxx-xxx, 1990. (submitted)

9. M. A. Handschy, T. J. Drabik, and L. Cotter, "Ferroelectric liquid crystal/silicon integrated circuit spatial light modulator," *Opt. Letts.*, vol. 15, pp. xxx-xxx, 1990. (accepted)

10. E. I. Verriest, E. N. Glytsis, and T. K. Gaylord, "Performance analysis of Givens rotation integrated optical interdigitated-electrode cross-channel Bragg diffraction devices: Intrinsic accuracy," *Appl. Opt.*, vol. 29, pp. xxxx-xxxx, 1990. (submitted)

11. T. K. Gaylord and K. F. Brennan, "Electron wave optics in semiconductors," *J. Appl. Phys.*, vol. 65, pp. 814-820, Jan. 15, 1989.

12. T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor superlattice interference filter design," *J. Appl. Phys.*, vol. 65, pp. 2535-2540, March 15, 1989.

13. E. N. Glytsis, T. K. Gaylord, and K. F. Brennan, "Semiconductor biased superlattice tunable electron interference filter/emitter," *J. Appl. Phys.*, vol. 66, pp. 1494-1497, Aug. 1, 1989.

14. T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor electron wave slab waveguides," *J. Appl. Phys. (Commun.)*, vol. 66, pp. 1483-1485, Aug. 1, 1989.

15. T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor quantum wells as electron slab waveguides," *J. Appl. Phys.*, vol. 66, pp. 1842-1848, Aug. 15, 1989.

16. E. N. Glytsis, T. K. Gaylord, and K. F. Brennan, "Theory and design of semiconductor electron-wave interference filter/emitters," *J. Appl. Phys.*, vol. 66, pp. 6158-6167, Dec. 15, 1989.

17. T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Electron-wave quarter-wavelength quantum well impedance transformers between differing energy-gap semiconductors," *J. Appl. Phys.*, vol. 67, pp. xxxx-xxxx, Mar. 1, 1990. (accepted)

Conference Proceedings:

1. T. J. Drabik and T. K. Gaylord, "Optoelectronic parallel processing arrays: System architecture and progress toward a prototype," *Optical Computing Technical Digest*, pp. TuH41-TuH44, Salt Lake City, Utah, March 1989.

2. N. F. Hartman and T. K. Gaylord, "Magneto-optic spatial light modulator characterization," (Abstract) *Optical Society of America Annual Meeting Technical Digest Series*, vol. 18, pg. THI2, Oct. 1989.

3. T. A. Maldonado and T. K. Gaylord, "Hybrid guided modes in biaxial dielectric waveguides," (Abstract) *Optical Socity of America Annual Meeting Technical Digest Series*, vol. 18, pg. FH2, Oct. 1989.

4. T. A. Maldonado and T. K. Gaylord, "Light propagation characteristics in biaxial crystals by a simple coordinate-free approach," (Abstract) *Optical Society of America Annual Meeting Technical Digest Series*, vol. 18, pg. ThG7, Oct. 1989.

5. M. M. Mirsalehi, T. K. Gaylord, D. C. Fielder, and C. C. Guest, "Comparison of number systems in truth-table look-up processing," (Abstract) *Optical Society of America Annual Meeting Technical Digest Series*, vol. 18, pg. TuDD5, Oct. 1989.

6. T. J. Drabik and T. K. Gaylord, "CMOS-compatible detector building blocks for optically interconnected VLSI and WSI," *Optical Society of America Annual Meeting Technical Digest Series*, vol. 18, pg. TuQ3, Oct. 1989.

7. E. N. Glytsis, T. K. Gaylord, and K. F. Brennan, "Semiconductor biased superlattices as electron wave interference filter/emitters," (Abstract) *Optical Society of America Annual Meeting Technical Digest Series*, vol. 18, pg. ML3, Oct. 1989.

8. T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Guided electron waves in semiconductor quantum wells," (Abstract) *Optical Society of America Annual Meeting Technical Digest Series*, vol. 18, pg. MB5, Oct. 1989.

List of Patent Applications Citing JSEP Sponsorship

1. T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor biased superlattice tunable interference filter/emitter," Georgia Tech Research Corporation patent application filed June 30, 1989, serial no. 07/374,476.

2. T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor quantum well electron and hole waveguides," Georgia Tech Research Corporation patent application filed June 30, 1989, serial no. 07/374,437.

References

[1] C. C. Guest and T. K. Gaylord, "Truth-table look-up processors using binary and residue arithmetic," *Appl. Opt.*, vol. 19, pp. 1201-1207, Apr. 1, 1980.

[2] D. L. Staebler, W. J. Burke, W. Phillips, and J. J. Amodei, "Multiple storage and erasure of fixed holograms in Fe-doped $LiNbO_3$," *Appl. Phys. Lett.*, vol. 26, pp. 182-184, Feb. 1975.

[3] A. Vander Lugt, "Signal detection by complex spatial filtering," *IEEE Trans. Inform. Theory.*, vol. IT-10, pp. 139-145, Apr. 1964.

[4] D. Gabor, G. W. Stroke, R. Restrick, A. Funkhouser, and D. Brumm, "Optical image synthesis (complex amplitude addition and subtraction) by holographic Fourier transformation," *Phys. Lett.*, vol. 18, pp. 116-118, Aug. 15, 1965.

[5] J. P. Huignard, J. P. Herriau, and F. Micheron, "Coherent selective erasure of superimposed volume holograms in $LiNbO_3$," *Appl. Phys. Lett.*, vol. 26, pp. 256-258, Mar. 1, 1975.

[6] C. C. Guest, M. M. Mirsalehi, and T. K. Gaylord, "EXCLUSIVE OR processing (binary image subtraction) using thick Fourier holograms," *Appl. Opt.*, vol. 23, pp. 3444-3454, Oct. 1, 1984.

[7] A. M. Glass, "The photorefractive effect," *Opt.Engr.*, vol. 17, pp. 470-479, Sept./Oct. 1978.

[8] H. H. Gallagher, T. K. Gaylord, M. G. Moharam, and C. C. Guest, "Reconstruction of binary-data-page holograms for an arbitrarily-oriented reference beam," *Appl. Opt.*, vol. 20, pp. 300-306, Jan. 15, 1981.

[9] K. Preston, *Coherent Optical Computers*, Chap. 8, McGraw-Hill, New York (1972).

[10] M. M. Mirsalehi, C. C. Guest, and T. K. Gaylord, "Residue number system holographic truth-table look-up processing: Detector threshold setting and probability of error due to amplitude and phase variations," *Appl. Opt.*, vol. 22, pp. 3583-3592, Nov. 15, 1983.

[11] T. K. Gaylord, M. M. Mirsalehi, and C. C. Guest, "Optical digital truth-table look-up processing," *Opt. Engr.*, vol. 24, pp. 48-58, Jan./Feb. 1985. (invited).

[12] M. M. Mirsalehi and T. K. Gaylord, "Multi-level coded residue-based content-addressable-memory optical computing, " *Appl. Opt.*, vol. 25, pp. 2277-2283, July 15, 1986.

[13] T. K. Gaylord and M. M. Mirsalehi, "Truth-table look-up processing: Number representation, multi-level coding, and logical minimization," *Opt. Engr.*, vol. 25, pp. 22-28, Jan. 1986. (invited).

[14] M. M. Mirsalehi and T. K. Gaylord, "Logical minimization of multilevel coded functions," *Appl. Opt.*, vol. 25, pp. 3078-3088, Sept. 15, 1986. (invited).

[15] M. M. Mirsalehi, T. K. Gaylord, D. C. Fielder, and C. C. Guest, "Number representation effects in truth-table look-up processing: 8-bit addition example," *Appl. Opt.*, vol. 28, pp. 1931-1939, May 15, 1989.

[16] T. J. Drabik and S. H. Lee, "Shift-connected SIMD array architectures for digital optical computing systems, with algorithms for numerical transforms and partial differential equations," *Appl. Opt.*, vol. 25, pp. 4053-4064, Nov. 15, 1986.

[17] R. K. Kostuk, J. W. Goodman, and L. Hesselink, "Design considerations for holographic optical interconnects," *Appl. Opt.*, vol. 26, pp. 3947-3953, Sept. 15, 1987.

[18] K. H. Brenner, "Programmable optical processor based on symbolic substitution," *Appl. Opt.*, vol. 27, pp. 1687-1691, May 1, 1988.

[19] T. J. Drabik and T. K. Gaylord, "Silicon VLSI-compatible ferroelectric liquid crystal light modulator process," (Abstract) *Optical Society of America Annual Meeting Technical Digest Series*, vol. 11, pg. TuAA4, Oct. 1988.

[20] W. E. Ross, D. Psaltis, and R. H. Anderson, "2D magneto-optic spatial light modulator for signal processing," *Opt. Engr.*, vol. 22, pp. 485-490, July/Aug. 1983.

[21] W. E. Ross, K. M. Snapp, and R. H. Anderson, "Fundamental characteristics of the Litton iron garnet magneto-optic spatial light modulator," *Proc. S.P.I.E.*, vol. 388, pp. 55-64, 1983.

[22] J. Horner and J. Legner, "Pattern recognition with binary phase-only filters," *Appl. Opt.*, vol. 24, pp. 609-611, Mar. 1, 1985.

[23] D. Flannery, A. Biernacki, and J. Loomis, "Real-time coherent correlator using binary magneto-optic spatial light modulators in the input and Fourier planes," *Appl. Opt.*, vol. 25, pg. 466, Feb. 15, 1986.

[24] A. Knoesen, N. F. Hartman, T. K. Gaylord, and C. C. Guest, "Minicomputer interface for magneto-optic spatial light modulator," *Rev. Sci. Instr.*, vol. 58, pp. 1843-1851, Oct. 1987.

[25] Semetex Corporation, 3450 Fujita Street, Torrance, CA 90505.

[26] R. W. Wood, "On a remarkable case of uneven distribution of light in a diffraction grating spectrum," *Phil. Mag.*, vol. 4, pp. 396-402, 1902.

[27] R. W. Wood, "Anomalous diffraction gratings," *Phys. Rev.*, vol. 48, pp. 928-936, Dec. 15, 1935.

[28] M. G. Moharam and T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings," *J. Opt. Soc. Amer.*, vol. 72, pp. 1385-1392, Oct. 1982.

[29] T. K. Gaylord, W. E. Baird, and M. G. Moharam, "Zero-reflectivity high spatial-frequency rectangular-groove dielectric surface-relief gratings," *Appl. Opt.*, vol. 25, pp. 4562-4567, Dec. 15, 1986.

[30] R. C. Enger and S. K. Case, "High frequency holographic transmission gratings in photoresist," *J. Opt. Soc. Amer.*, vol. 73, pp. 1113-1118, Sept. 1983.

[31] R. C. Enger and S. K. Case, "Optical elements with ultrahigh spatial-frequency surface corrugations," *Appl. Opt.*, vol. 22, pp. 3220-3228, Oct. 15, 1983.

[32] T. K. Gaylord, E. N. Glytsis, and M. G. Moharam, "Zero-reflectivity homogeneous layers and high spatial-frequency surface-relief gratings on lossy materials," *Appl. Opt.*, vol. 26, pp. 3123-3135, Aug. 1, 1987.

[33] M. Cahay, M. McLennan, S. Datta, and M. S. Lundstrom, "Importance of space-charge effects in resonant tunneling devices," *Appl. Phys. Lett.*, vol. 50, pp. 612-614, Mar. 9, 1987.

[34] K. M. S. V. Bandara, D. D. Coon, Byungsung O., Y. F. Lin, and M. H. Francombe, "Exchange interactions in quantum well subbands," *Appl. Phys. Lett.*, vol. 53, pp. 1931-1933, Nov. 14, 1988.

[35] K. M. S. V. Bandara and D. D. Coon, "Electron-electron interactions and resonant tunneling in heterostructures," *Appl. Phys. Lett.*, vol. 53, pp. 1865-1867, Nov. 7, 1988.

[36] W. R. Frensley, "Transient response of a tunneling device obtained from the Wigner function," *Phys. Rev. Lett.*, vol. 57, pp. 2853-2856, Dec. 1, 1986.

[37] W. R. Frensley, "Quantum transport modeling of resonant tunneling devices," *Solid-State Electron.*, vol. 31, pp. 739-742, March 1988.

[38] B. A. Mason, K. Hess, R. E. Cline, Jr., and P. G. Wolynes, "A new technique for the calculation of real-time path integrals and applications to electron transport," *Superlattices and Microstructures*, vol. 3, no. 4, pp. 421-428, 1987.

[39] G. J. Iafrate and J. B. Krieger, "Quantum transport and solid-state dynamics for Bloch electrons in an electric field," *Solid-State Electron.*, vol. 31, pp. 517-521, March 1988.

[40] P. A. Schulz and C. E. T. Goncalves da Silva, "Simple model for resonant tunneling beyond the effective mass approximation," *Phys. Rev. B.*, vol. 35, pp. 8126-8130, May 15, 1987.

[41] J. N. Schulman and C. L. Anderson, "HgTe-CdTe double barrier tunneling structures," *Appl. Phys. Lett.*, vol. 48, pp. 1684-1686, June 16, 1986.

[42] H. Guo, K. Diff, G. Neofotistos, and J. D. Gunton, "Time-dependent investigation of the resonant tunneling in a double-barrier quantum well," *Appl. Phys. Lett.*, vol. 53, pp. 131-133, July 11, 1988.

[43] M. C. Yalabik, G. Neofotistos, K. Diff, H. Guo, and J. D. Gunton, "Quantum mechanical simulation of charge transport in very small semiconductor structures," *IEEE Trans. Electron Dev.*, vol. 36, pp. 1009 - 1014, June 1989.

[44] A. P. Jauho, and M. M. Nieto, "Time-dependent tunneling of wave-packets through heterostructures in an applied field," *Superlattices and Microstructures*, vol. 2, no. 5, pp. 407-413, 1986.

[45] R. K. Mains and G. I. Haddad, "Time-dependent modeling of resonant-tunneling diodes from direct solution of the Schroedinger equation," *J. Appl. Phys.*, vol. 64, pp. 3564-3569, Oct. 1, 1988.

[46] T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor superlattice interference filter design," *J. Appl. Phys.*, vol. 65, pp. 2535-2540, March 15, 1989.

[47] T. K. Gaylord and K. F. Brennan, "Semiconductor superlattice electron wave interference filters," *Applied Physics Letters*, vol. 53, pp. 2047-2049, Nov. 21, 1988.

[48] T. K. Gaylord and K. F. Brennan, "Electron wave optics in semiconductors," *Journal of Applied Physics*, vol. 65, pp. 814-820, Jan. 15, 1989.

[49] E. N. Glytsis, T. K. Gaylord, and K. F. Brennan, "Semiconductor biased superlattice tunable electron interference filter/emitter," *J. Appl. Phys.*, vol. 66, pp.1494-1497, Aug. 1, 1989.

[50] T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor electron wave slab waveguides," *J. Appl. Phys. (Commun.)*, vol. 66, pp. 1483-1485, Aug. 1, 1989.

[51] T. K. Gaylord, E. N. Glytsis, and K. F. Brennan, "Semiconductor quantum wells as electron slab waveguides," *J. Appl. Phys.*, vol. 66, pp. 1842-1848, Aug. 15, 1989.

[52] H.A. Macleod, *Thin Film Optical Filters*, (Macmillan, New York, 1986).

[53] K. F. Brennan and C. J. Summers, "Theory of resonant tunneling in a variably spaced multiquantum well structure: An Airy function approach," *J. Appl. Phys.*, vol. 61, 614-623, Jan. 15, 1987.

[54] H. Kogelnik, *Integrated Optics*, edited by T. Tamir (Springer-Verlag, New York, 1975), pg. 13.

[55] R. C. Potter and A. A. Lakhani, "Observation of electron quantum interference effects due to virtual states in a double-barrier heterostructure at room temperature," *Appl. Phys. Lett.*, vol. 52, pp. 1349-1351, Apr. 18, 1988.

[56] S. Ben Amor, K. P. Martin, J. J. L. Rascol, R. J. Higgins, R. C. Potter, A. A. Lakhani, and H. Hier, "Magnetotransport studies of charge accumulation in an AlInAs/GaInAs tunneling structure," *Appl. Phys. Lett.*, vol. 54, 1989. (submitted)

[57] T. Nakagawa, N. J. Kawai, and K. Ohta, "Design principles for CHIRP superlattice devices," *Superlattices Microstructures*, vol. 1, pp. 187-192, 1985.

[58] T. Nakagawa, H. Imamoto, T. Sakamoto, T. Kojima, K. Ohta, and N. J. Kawai, "Observation of negative differential resistance in CHIRP superlattices," *Electron. Lett.*, vol. 21, pp. 882-884, Sept. 12, 1985.

[59] M. Heiblum, "Tunneling hot electron transfer amplifiers (theta): Amplifiers operating up to the infrared," *Solid-State Electron.*, vol. 24, pp. 343-366, 1981.

[60] C. S. Lent, "The resonant hot electron transfer amplifier: A continuum resonance device," *Superlattices Microstructures*, vol. 3, pp. 387-389, July 1987.

[61] M. Heiblum, M. I. Nathan, D. C. Thomas, and C. M. Knoedler, "Direct observation of ballistic transport in GaAs," *Phys. Rev. Lett.*, vol. 55, pp. 2200-2203, Nov. 11, 1985.

[62] M. Heiblum and L. F. Eastman, "Ballistic electrons in semiconductors," *Scientif. Amer.*, vol. 256, pp. 102-110, Feb. 1987.

[63] M. Heiblum, "Ballistic electrons and holes observed in a semiconductor," *Opt. News*, vol. 14, pp. 13-16, Oct. 1988.

Work Unit Five

TITLE: Two-Dimensional Optical/Electronic Signal Processing

SENIOR PRINCIPAL INVESTIGATORS:

W. T. Rhodes, Professor

SCIENTIFIC PERSONNEL:

S. D. Goodman, Graduate Research Assistant (Ph.D. Candidate)
J. N. Hereford, Graduate Research Assistant (Ph.D. Candidate)
R. W. Stroud, Graduate Research Assistant (Ph.D. Candidate)
J. van der Gracht, Graduate Research Assistant (Ph.D. Candidate)

SCIENTIFIC OBJECTIVE:

The long term scientific objective of this research is to advance understanding and extend state-of-the-art capabilities in real-time nonlinear processing of 2-D images and high-speed linear processing of 3-D images, particularly as applied to pattern recognition and to optical computing interconnections.

RESEARCH ACCOMPLISHMENTS:

Research this past year has been concentrated in four areas: (1) three-dimensional image processing; (2) nonlinear image processing, (3) partially coherent optical image processing, and (4) new methods for exploiting spatial light modulators. These areas are discussed in the following sections.

- *Three-Dimensional Image Processing*

Three-dimensional image processing and display have grown in importance as research topics during the past several years. The development of high-speed computer graphics hardware has played a major role in this growth, as have new developments in 3-D imaging techniques in microscopy. An important related development is the interest in 3-D interconnects for optical computing. For example, Lohmann has shown how 3-D imaging might be used to couple multi-layer integrated optics processing units [1]; Psaltis has proved the critical nature of volumetric interconnections for optical neural-network processors [2].

An important aspect of 3-D imaging is that – unlike 2-D imaging – it only rarely satisfies the conditions for space invariance. If a 3-D imaging operation is shift-invariant (to

within transverse and longitudinal magnification factors), the object-image relationship can be described by a 3-D convolution:

$$i(x_i, y_i, z_i) = \int_{-\infty}^{\infty} o(x_o, y_o, z_o) h(Mx_i - x_o, My_i - y_o, M^2 z_i - z_o) dx_o dy_o dz_o \quad , \qquad (1)$$

where $o(x_o, y_o, z_o)$ is the object distribution, $i(x_i, y_i, z_i)$ is the image distribution, $h(x, y, z)$ represents the 3-D impulse response (point spread function) and M is a magnification factor. Three-dimensional Fourier techniques can easily be applied to the processing (e.g., restoration) of such imagery. Unfortunately, most 3-D imaging cannot be described by a convolution and must instead by described by a more general space-variant superposition integral [3,4],

$$i(x_i, x_i, x_i) = \int_{-\infty}^{\infty} o(x_o, y_o, z_o) h(x_i, x_i, x_i, x_o, y_o, z_o) dx_o dy_o dz_o \quad . \qquad (2)$$

Typically, a 3-D imaging operation is strongly space-variant. Figure 1 shows how this is true even for the imaging operation performed by a single thin lens. For single-lens imaging, both transverse and longitudinal magnification change with z, as manifested by the different ways the point images go out of focus. Note also that the two image-space cones of light, which represent the 3-D responses to point sources 1 and 2, have different cone angles and different tilt angles for their axes. For space-invariant 3-D imaging, the cone angles and tilt angles of the axes must be the same. The typical absence of space invariance in 3-D imaging complicates the application of lens systems to interconnections in optical computing and neural-network systems and makes much more difficult the restoration of 3-D images obtained with conventional imaging systems. Lohmann has pointed out that afocal-telecentric lens systems are space-invariant over a limited region of object space [1], and Streible has analyzed the image-forming characteristics of such systems for partially-coherent objects [5,6]. (It should be noted that even with imaging systems of this kind, space invariance in the imaging operation can be destroyed by multiple scattering or by one part of the object occluding another.) However, so far as we can ascertain, no one has studied in detail the conditions on object size and shape imposed by lens and limiting aperture diameters that must be satisfied if such systems are to be truly shift-invariant, and researchers continue to propose interconnection schemes and 3-D image restoration schemes that will not work as believed (or may not work at all) because an underlying assumption of space invariance is not satisfied.

During this past year we began investigating 3-D imaging systems in detail and have obtained several important results. First, we have clarified the conditions that must be satisfied by an optical imaging system if true 3-D space invariance is to hold, both for coherent and for incoherent imaging [7,8]. We have shown, for example, that *only* lens systems that are both afocal and telecentric are capable of 3-D space-invariant imaging over large object fields. (A two-lens imaging system is afocal if the lenses are separated by the sum of their focal lengths. The system is also telecentric if the limiting aperture is in the common focal plane – i.e., the back focal plane of the objective lens and the front

focal plane of the eyepiece lens.) The afocal nature of the system guarantees that the transverse and longitudinal magnification factors not change with longitudinal position of the object. The telecentric nature of the system means that both the entrance and exit pupils of the system appear to be at infinity, thereby guaranteeing that the cone angle and the tilt of the cone axis of the light that comes to a focus on the Gaussian image point will not vary with the location of that point. This condition is illustrated in Figure 2. In an important practical extension of this work, we have also determined clearly the relationship between the diameter of the limiting aperture, the diameters of the lenses, and the size of the region in object space for which space invariance holds [9,10]. In the process of working out these details we have also corrected previously published misconceptions [1].

Of even greater importance in this research, we have determined that any aberration-free, non-afocal space-variant 3-D optical imaging operation can be represented by a geometrical coordinate transform, followed by a space-invariant operation, followed in turn by another geometrical coordinate transform [7,8]. Figure 3 models the basic operation, and a physical interpretation of the model is presented in Figure 4. In Figure 4(a) a generic non-afocal imaging system is illustrated. The modified system shown in Figure 4(b) would in general have different imaging characteristics. However, in each lens pair we choose the negative lens to have the same power as the positive lens, and, consequently, the lens pairs have no effect on the imaging operation. Conceptually, however, we can now break the imaging system into three parts, as shown in Figure 4(c). Because the two positive lenses make the entrance and exit pupils appear to be at infinity, the middle part is afocal and telecentric and is thus modeled mathematically by a space-invariant superposition integral. The negative lenses, as separated out, introduce an easily-modeled distortion in their respective spaces.

The equivalence just described is extremely important from a practical standpoint, for it means that high-speed FFT-based image processing and restoration algorithms can be applied to 3-D images obtained by conventional imaging systems. In the absence of such a model, methods that are much more computationally intensive have to be used.

We have also developed a new method for finding the surface and brightness functions associated with an incoherently-radiating 3-D surface object using a simple imaging system. (A surface object can be described by the distribution $o(x,y,z) = h(x,y)\delta[z - f(x,y)]$, where the equation $z = f(x,y)$ describes the surface, and the function $h(x,y)$ represents the brightness on that surface.) In this scheme the limiting aperture of the imaging system is moved, introducing a lateral shear in the 3-D point spread function in all planes except the true conjugate image plane. The proposed method consists of measuring the light intensity along lines parallel to the optical axis for a number of shifted positions of the limiting aperture. The variance of these measurements is taken for each longitudinal position; when the variance reaches a minimum the image surface has been located. We have yet to test this method experimentally, but we believe that it is robust and will work over a wider range of conditions than other methods that have

been proposed [11].

- *Nonlinear Image Processing*

Nonlinear techniques have always played an important role in image processing. Significant recent advances have come through the application of mathematical morphology [12,13,14] combined with threshold decomposition [15,16]. Increased understanding of the interrelationship between convolution (with a nonlinearity), symbolic substitution [17], and mathematical morphology has resulted in considerable research on these subjects in the context of optical processing systems [18,19,20,21,22]. The basic methods are powerful and are beginning to be used for pattern recognition and image enhancement operations.

Our interest in optical implementation of nonlinear image processing stems from the parallelism and high intrinsic speed of optical processing in analog applications. This is particularly important when morphological (shape-changing) operations are applied to continuous-tone images, where, if spatial light modulator technology progresses as expected, optical-electronic systems should have a distinct speed advantage over their all-electronic counterparts.

Most of our research in nonlinear image processing has concentrated on the generic nonlinear processing system illustrated in Figure 5. A binary (denoted by subscript "b") input image is convolved with a binary kernel and the result subjected to an ideal thresholding/hardliming operation to produce a binary output image. With this system one can perform a group of operations known as ranked-order filtering operations [23] simply by changing the level of the threshold. The ranked-order operations include the basic mathematical morphology operations of erosion and dilation, as well as the median filter operation. The convolution of the first block in the figure can be performed in parallel and at extremely high speed using optical methods; hence our particular interest in the subject.

We have greatly increased the power of the system of Figure 5 by allowing the convolution kernel to assume a continuous range of values. Gray-scale kernels are as easily implemented with the optical approach as binary kernels, and they increase significantly the flexibility and power of the system. Since gray-scale kernels are difficult to implement using all-electronic systems (which then require integer or floating-point arithmetic capabilities), they are thus particularly attractive for investigation for optical processors. Figure 6(a) shows an example of a one-dimensional discrete nonlinear filtering operation – an asymmetrical median filter – using a multi-valued convolution kernel. The kernel is given by the sequence of {1 2 1} and convolved with a sequence of 1's and 0's. The output of the convolution (second line) equals 0, 1, 2, 3, or 4. If the threshold is set at 2.5, the result is the removal of isolated positive-going pulses (third line). Negative-going pulses, however, remain. If, on the other hand, the threshold level is set at 1.5, isolated donward-going pulses are removed while upward-going pulses remain (fourth line). The filter behaves like an asymmetric version of a median

filter, the median filter removing both positive-going and negative-going isolated pulses, as shown in Figure 6(b). Although operations similar to asymmetric median filtering can be performed using sequences of erosions and dilations, impiementation using a gray-scale kernel requires only a single pass.

In an extention of this method to two-dimensional inputs using an optical processor, we have successfully implemented asymmetrical median filtering on a binary input image that was corrupted by salt-and-pepper noise. The 2-D convolution kernel had the form of a central bright spot (essentially a delta-function) surrounded by a lower-level circular plateau. A manuscript on this work has been submitted for publication [24].

Quite recently we have been looking at a new class of nonlinear optical image filtering operations that show great promise for edge detection and, we suspect, for general feature detection and pattern recognition. In these operations, referred to as rotating-kernel min-max (RKM) transformations, we first calculate the convolution of the object distribution with a rotated version of the kernel $K(x,y)$:

$$S_i(x,y) = 0(x,y) ** R_{\theta_i}\{K(x,y)\} \ , \tag{3}$$

where $**$ denotes 2-D convolution, and where R_{θ_i} denotes rotation through the angle θ_i. This operation can be performed optically in parallel. For each value of x and y the maximum and minimum values achieved by $S_i(x,y)$ over the range of values of i are found, and their ratio, $\alpha = \min / \max$ calculated. The final output image distribution $i(x,y)$ is calculated in accord with one of a variety of possible mappings of the general form

$$i(x,y) = f_n(\alpha^m(x,y)) \ . \tag{4}$$

Thus far we have considered the following six forms for $f_n(\alpha^m(x,y))$:

$$\begin{aligned}
f_1(\alpha^m) &= \alpha^m & 0 \le \alpha \le 1 \\
f_2(\alpha^m) &= 1 - \alpha^m & 0 \le \alpha \le 1 \\
f_3(\alpha^m) &= \frac{1-\alpha^m}{1+\alpha^m} & 0 \le \alpha \le 1 \\
f_4(\alpha^m) &= f_1(\alpha^m) e^{[1-f_1(\alpha^m)]} & 0 \le \alpha \le 1 \\
f_5(\alpha^m) &= f_2(\alpha^m) e^{[1-f_2(\alpha^m)]} & 0 \le \alpha \le 1 \\
f_6(\alpha^m) &= f_3(\alpha^m) e^{[1-f_3(\alpha^m)]} & 0 \le \alpha \le 1
\end{aligned}$$

Very preliminary investigations indicate that these transformations can be used for (a) contrast modification, (b) thinning of line segments, (c) suppression of slowly-varying background distributions, (d) line segment and edge extraction, (e) filling in of broken line segments.

Our sense at this time is that successful implementation of the RKM transformation will require that the convolution kernel be rotated through a large number of angles. Such rotation is not at all easily implemented digitally, but can be effected optically

with readily-available image rotators (rotating dove prism, for example).

- *Partially Coherent Optical Image Processing*

Work in this area has concentrated on understanding better the characteristics of partially coherent imaging systems and in the application of such systems to image enhancement. We feel that full two-dimensional simulations of partially-coherent imaging operations are essential to our understanding of the complicated nature of these systems. As a consequence, we have looked at alternative mathematical models for Koehler-type partially coherent imaging from the standpoint of computer simulation. Early in the project we clarified the sampling conditions that apply to the source and pupil distributions in such bilinear systems. We have shown that the minimum sampling rate depends on the reciprocal of the smaller of the width of the object and the width of the imaging system point spread function. A paper on the subject was published this past year [25].

During this past year we have written and tested most of the simulation software. In this work we have the important goal of minimizing the computation effort, which is historically tremendous for such simulations. (Full 2-D simulations are rarely carried out for distributions greater than 64×64.) The software assumes that object, pupil, and source distributions are separable in x and y, consistent with our belief that major computational savings will result from outer product expansions of these distributions using truncated singular value decomposition representations. Calculations of computational complexity indicate that dramatic savings in processing effort occur if the SVD expansions can be truncated at roughly the 10%-of-full-rank level. This work was presented at a conference this fall [26], and a manuscript is in preparation for publication.

The most recent development in this area is the successful experimental testing of a method for image enhancement using source–as opposed to pupil–modification in a partially coherent imaging system. The technique has limited processing capability but offers several potential advantages over more conventional techniques. In particular, requirements on spatial light modulators are greatly reduced, and noise performance is better compared to that associated with coherent spatial filtering. The technique blocks unwanted spatial frequencies in the object amplitude transmittance. The fixed pupil distribution is a pseudo-random binary amplitude pattern. The random nature of the pupil ensures that a large number of object amplitude spatial frequencies will participate in the image formation process. The appropriate corresponding source distribution is calculated and scanned out to block a particular object amplitude frequency. The operating principle is a variation of darkfield imaging. This work was also presented at a conference this fall [27], and a journal manuscript is in preparation.

- *New Methods for Exploiting Spatial Light Modulators*

This past year we have been looking at new ways of using spatial light modulators that overcome some of the typical deficiencies connected with such devices. One student (here on a U.S. Army fellowship) has begun a detailed study of the errors that result in image correlation operations when SLM's are used with finely-quantized phase but only coarsely-quantized magnitude transmittance values. Initial results suggest that currently-available SLM's (including the Texas Instruments deformable membrane SLM and a modified version of the Hughes liquid-crystal SLM) used in this mode can outperform the binary-phase and phase-only SLM's that have been reported as being so successful in coherent optical correlators. A manuscript has been submitted for publication [28].

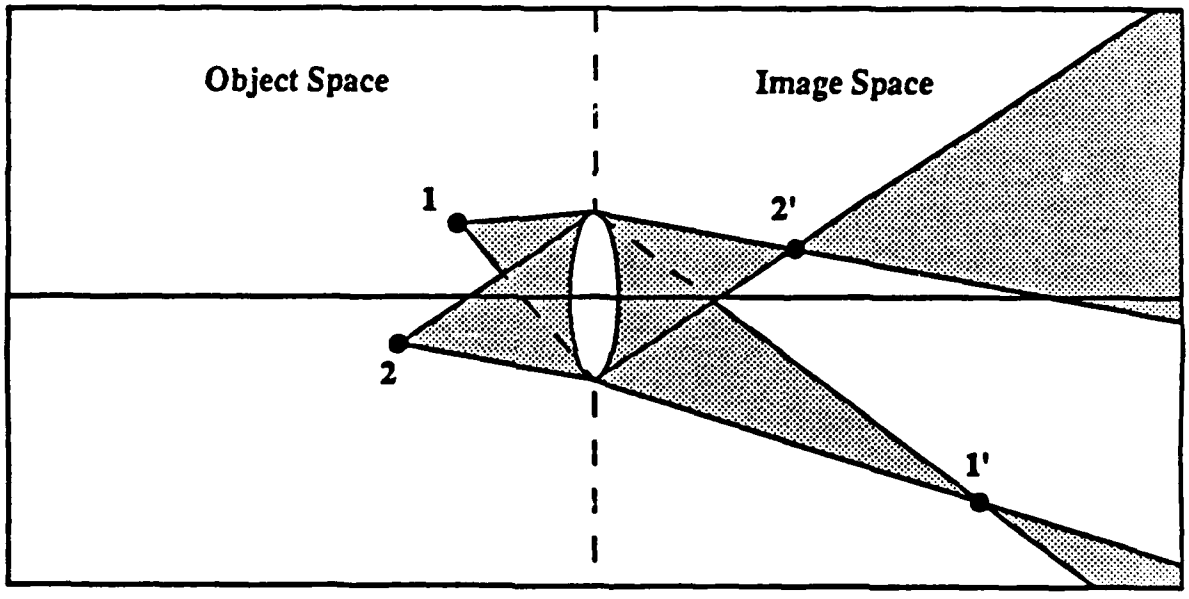


Figure 1: Example of space-variant 3-D imaging with single-lens imaging system.

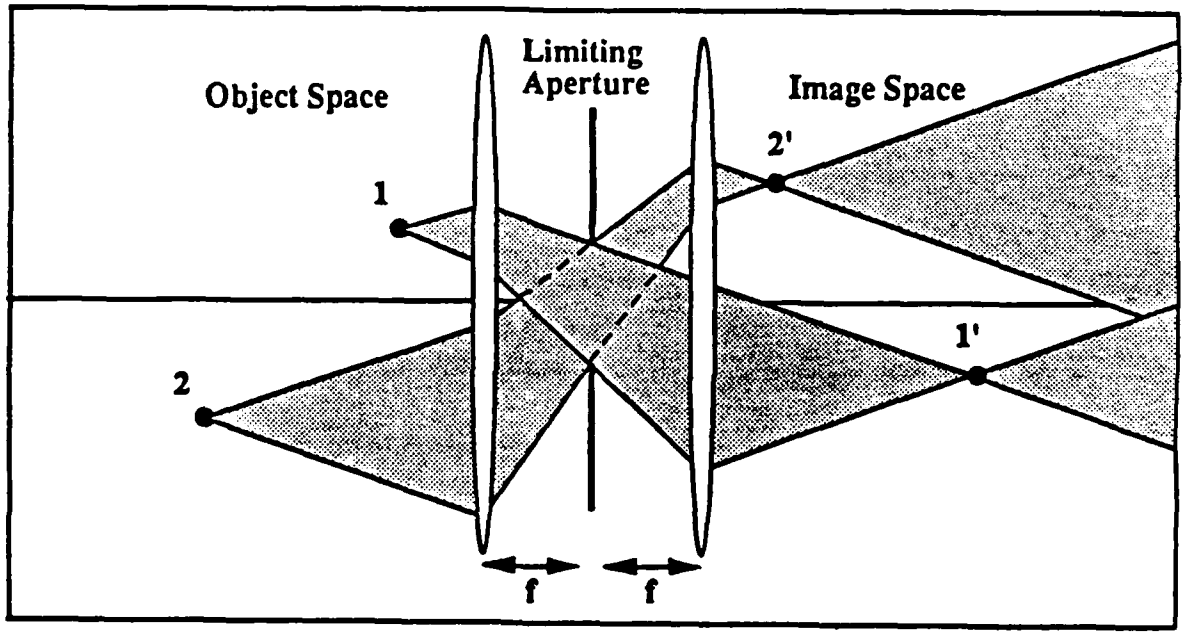


Figure 2: The afocal telecentric imaging system illustrated satisfies the conditions for space-invariant imaging.

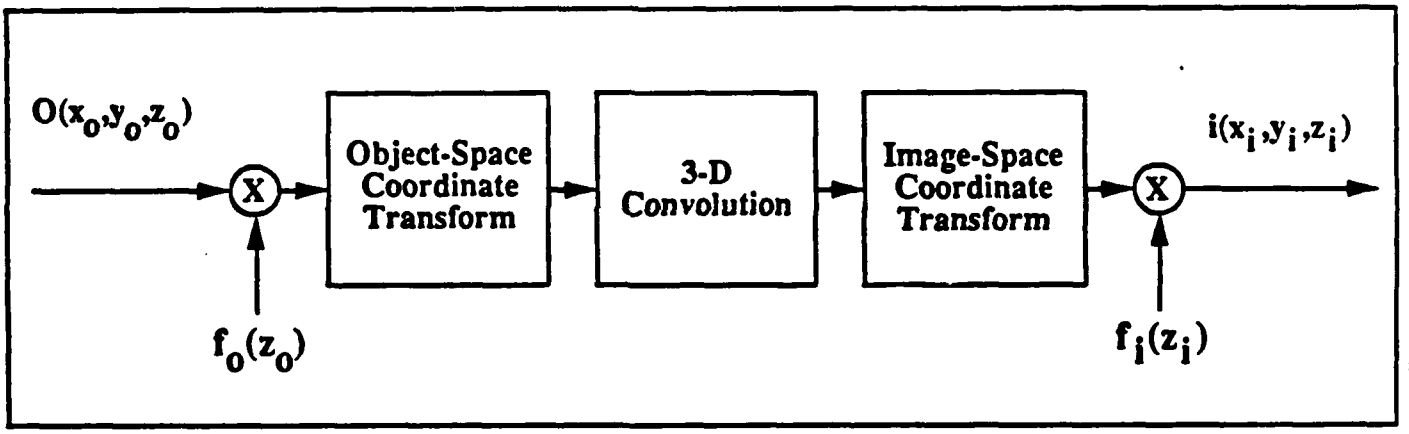


Figure 3: Model for general 3-D optical imaging system. The space-variant operation relating object and image distributions can be replaced by a space-invariant operation preceded and followed by z-dependent intensity scalings and (invertible) coordinate transformations.

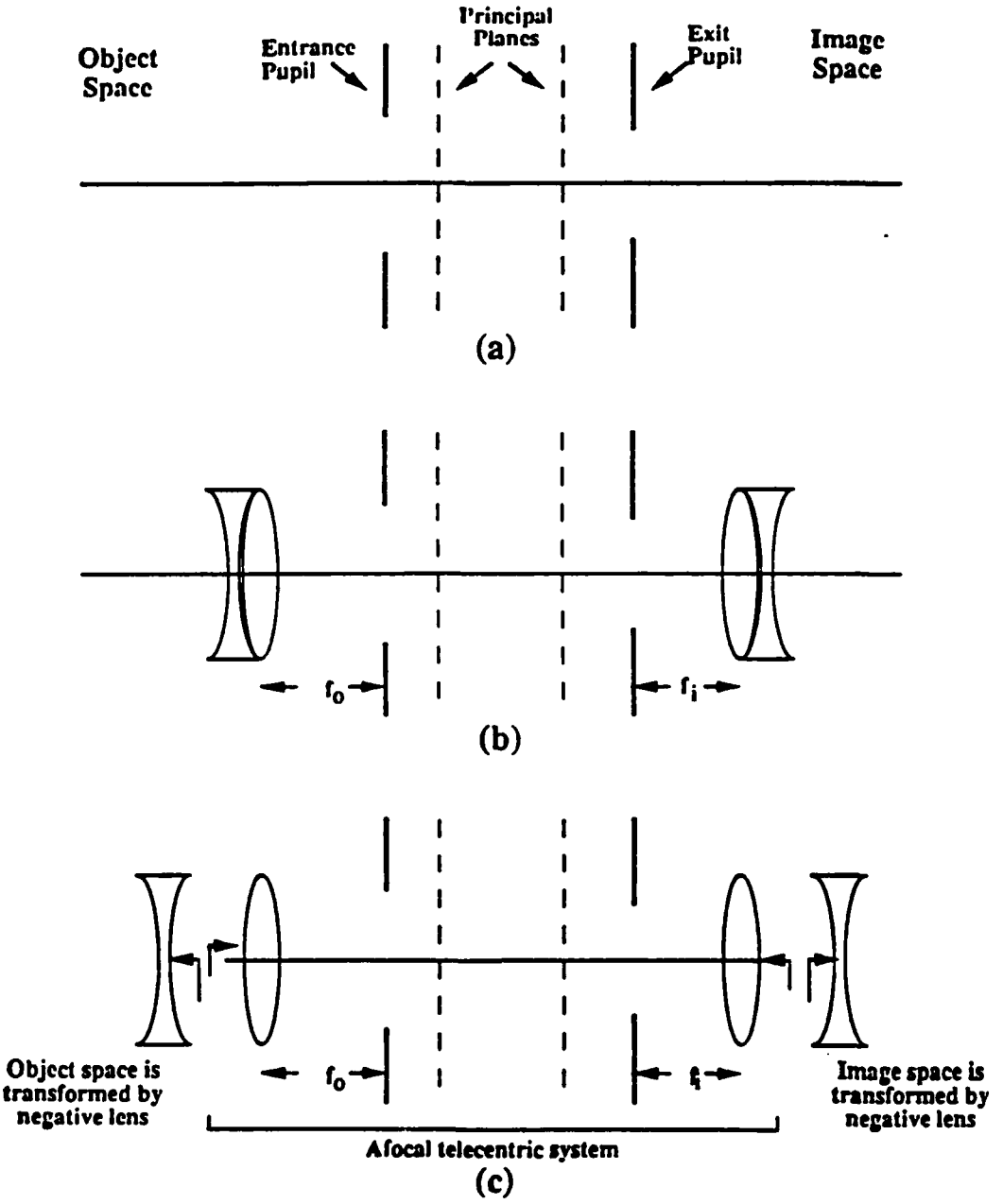


Figure 4: Physical interpretation of model of Figure 3.

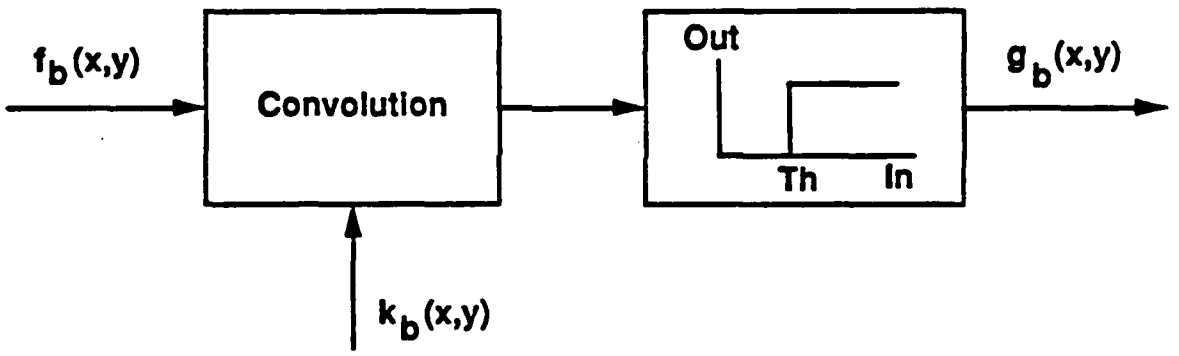


Figure 5: Block diagram of nonlinear image processing system.

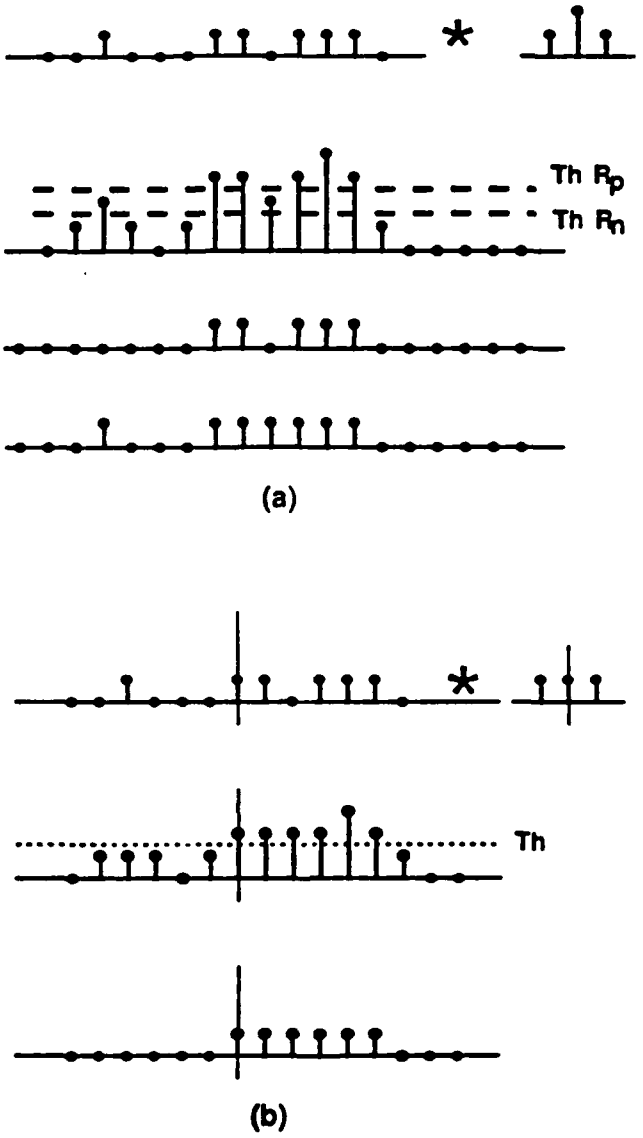


Figure 6: Example of one-dimensional asymmetric median filtering: (a) Positive-pulse removing and negative pulse removing asymmetric filtering operations controlled by choice of threshold. (b) Conventional median filtering for comparison.

PUBLICATIONS:

Journal Articles Published or Accepted:

1. J. N. Mait and W. T. Rhodes, "Pupil Function Design Algorithm for Bipolar Incoherent Spatial Filtering," *Appl. Opt.*, vol. 28, no. 8, 15 April 1989, pp. 1474-1499.

2. D. N. Sitter and W. T. Rhodes, "Generalization of the Falling Raster-Folded Spectrum Relationship," accepted for publication in *Applied Optics.*

3. J. van der Gracht and W. T. Rhodes, "Source Sampling for Incoherent Imaging and Spatial Filtering," *J. Opt. Soc. Am. A.*, vol. 6, August 1989.

Journal Articles Submitted:

1. M. A. Kaura and W. T. Rhodes, "The role of magnitude content for optical correlators demonstrated using phase-with-constrained-magnitude filters," submitted to *Optical Engineering*, December 1989.

2. J. M. Hereford and W. T. Rhodes, "Optical assymetrical median filtering using gray-scale convolution kernels," submitted to *Optics Letters.*

References

[1] A. Lohmann, "Parallel Interfacing of Integrated Optics with Free-Space Optics," *Optik*, vol. 75, 1987, pp. 1-4.

[2] X. -G Gu, and D. Psaltis, "Local and Asymmetric Interconnections Using Volume Holograms," paper THZ5, *1988 Annual Meeting of the Optical Society of America.*

[3] H. H. Hopkins, "The Frequency Response of a Defocused Optical System," *Proc. Roy. Soc. A*, vol. 231, 1955, pp. 91-103.

[4] B. Roy Frieden, "Optical Transfer of the Three-Dimensional Object," *J. Opt. Soc. Am.*, vol. 57, pp. 56-66, January, 1966.

[5] N. Streibl, "Depth Transfer by an Imaging System," *Optica Acta*, vol. 31, no. 11, Nov. 1984, pp. 1233-1241.

[6] N. Streibl, "Three-Dimensional Imaging by a Microscope," *J. Opt. Soc. Am. A*, vol. 2, Feb. 1985, pp. 121-127.

[7] W. T. Rhodes, E. H. Rothenheber, and D. Sitter, "Conditions for Space-Invariant 3-D Imaging," presented at *1989 Annual Meeting of the Optical Society of America.*

[8] W. T. Rhodes, E. H. Rothenheber, and D. Sitter, "Space-Invariant Imaging with Afocal-Telecentric Lens Systems," in preparation.

[9] D. Sitter, and W. T. Rhodes, "Model for Imaging Three-Dimensional, Incoherently Radiating Objects with Focal Imaging Systems," presented at *1989 Annual Meeting of the Optical Society of America.*

[10] D. Sitter, and W. T. Rhodes, "Space-Invariant Model for Non-Afocal 3-D Imaging Systems," in preparation.

[11] G. Häulser, and E. Kõrner, "Simple Focusing Criterion," *Appl. Opt.*, vol. 23, no. 15, pp. 2468-2469, 1 Aug. 1984.

[12] J. Serra, *Image Analysis and Mathematical Morphology*, (Academic Press, London, 1982).

[13] P. Maragos, and R. W. Schafer, "Morphological Filters - Part I: Their Set-Theoretic Analysis and Relations to Linear Shift-Invariant Filters," *IEEE Trans. on ASSP*, vol. ASSP-35, Aug. 1987, pp. 1153-1169.

[14] P. Maragos, and R. W. Schafer, "Morphological Filters - Part II: Their Relations to Median, Order-Statistic, and Stack Filters," *IEEE Trans. on ASSP*, vol. ASSP-35, Aug. 1987, pp. 1170-1187.

[15] J. P. Fitch, E. J. Coyle, and N. C. Gallagher, "Median Filtering by Threshold Decomposition," *IEEE Trans. on ASSP*, vol. ASSP-32, Dec. 1984, pp. 1183-1188.

[16] J. P. Fitch, E. J. Coyle, and N. C. Gallagher, "Threshold Decomposition of Multidimensional Ranked Order Operations," *IEEE Trans. on Circuits and Systems*, vol. CAS-32, May 1985, pp. 445-450.

[17] K. -H. Brenner, A. Huang, and N. Streibl, "Digital Optical Computing with Symbolic Substitution," *Appl. Opt.*, vol. 25, no. 18, pp. 3054-3060, 15 Sep. 1986.

[18] K. S. O'Neill, and W. T. Rhodes, "Morphological Transformations by Hybrid Optical-Electronic Methods," *Proc. SPIE*, vol. 638, 1986, pp. 41-44.

[19] Ochoa, E., J. P. Allebach and D. W. Sweeney, "Optical Median Filtering Using Threshold Decomposition," *Appl. Opt.*, vol. 26, 1987, pp. 252-260.

[20] S. D. Goodman, and W. T. Rhodes, "Symbolic Substitution Applications to Image Processing," *Appl. Opt.*, vol. 27, no. 9, 1 May 1988, pp. 1708-1714.

[21] J. M. Hereford, and W. T. Rhodes, "Nonlinear Optical Image Filtering by Time-Sequential Threshold Decomposition," *Opt. Engrng.*, vol. 27, no. 2, April 1988, pp. 274-279.

[22] E. Botha, D. Casasent, and E. Barnard, "Optical Symbolic Substitution Using Multichannel Correlators," *Appl. Opt.*, vol. 27, no. 5, pp. 817-818, 1 Mar. 1988.

[23] T. A. Nodes, and N. C. Gallagher, "Median Filters: Some Modifications and Their Properties," *IEEE Trans. on ASSP*, vol. ASSP-30, Oct. 1982, pp. 739-746.

[24] J. M. Hereford and W. T. Rhodes, "Optical Asymmetrical Median Filtering Using Gray-Scale Convolution Kernels," submitted to *Optics Letters*.

[25] J. van der Gracht and W. T. Rhodes, "Source Sampling for Incoherent Imaging and Spatial Filtering," *J. Opt. Soc. Am. A.*, vol. 6, August 1989, pp. 1165-1167.

[26] J. van der Gracht and W. T. Rhodes, "Computer Simulation of Partialiy Coherent Imaging by Outer-Product Expansions," presented at the *1989 Annual Meeting of the Optical Society of America.*

[27] J. van der Gracht, "Source Modification for Partially Coherent Image Enhancement," presented at the *1989 Annual Meeting of the Optical Society of America.*

[28] M. A. Kaura and W. T. Rhodes, "Improved Performance in an Optical Correlator Using a Phase-with-Quantized Magnitude Complex Spatial Filter," submitted tc *Applied Optics.*

Work Unit Six

TITLE: Electromagnetic Measurements in the Time and Frequency Domains

SENIOR PRINCIPAL INVESTIGATOR:

G. S. Smith, Regents' Professor

SCIENTIFIC PERSONNEL:

W.R. Scott, Jr., Assistant Professor
M. Gouker, Graduate Research Assistant (Ph.D. Candidate)
J.G. Maloney, JSEP Fwllow (Ph.D. Candidate)
G.P. Zhou, Graduate Research Assistant (Ph.D. Candidate)

SCIENTIFIC OBJECTIVE:

Electromagnetic measurements play an important role in the advancement of electromagnetic technology. Measurements are used to verify new theoretical results and, ultimately, to test all new devices and systems. New measurement techniques are needed to support the current interest in miniaturization (integrated circuits), increases in operating frequency and bandwidth (millimeter waves), and the introduction of new materials.

The broad objective of this research is to develop new methodology for making electromagnetic measurements directly in the time domain or over a wide bandwidth in the frequency domain. This research includes the development of the theoretical analyses necessary to support the measurement techniques. One aspect of the research is the systematic study of radiating structures placed near or embedded in material bodies. In a practical situation the radiator might serve as a diagnostic tool for determining the geometry, composition or electrical constitutive parameters of the body.

RESEARCH ACCOMPLISHMENTS:

- *Materials for Electromagnetic Scale Models*

Electromagnetic scale models are of value in the design and testing of electromagnetic systems that are on a scale too large or too small for routine laboratory investigation. For example, the performance of antennas over or buried in the earth can be evaluated in the laboratory using reduced size scale models operated at frequencies higher than those used for the actual antennas. A longstanding problem in experimental electromagnetics has been the synthesis of materials with adjustable electrical parameters that can be used in scale models, that is, materials whose electrical parameters (ε, σ) can be adjusted.

In this research oil-in-water (O/W) emulsions, which are suspensions of small oil droplets in water, were examined as materials for this purpose. The relative permittivity of oil is around two, while that of water is around eighty. Thus, a wide range of relative permittivity can be obtained by mixing these components. The conductivity of the emulsion can be adjusted by changing the concentration of an electrolyte, such as sodium chloride, in the aqueous phase. As one would expect from practical experience, a major problem with emulsions is stabilizing the mixture so that the water and oil do not separate. This is usually accomplished by making the oil droplets very small and adding a stabilizing agent, an emulsifier.

A series of O/W emulsions was formulated for use in scale models; this involved the selection of an appropriate emulsifier. The ingredients for the emulsions are mineral oil, a saline solution (NaCl) and a nonionic emulsifier. The electrical constitutive parameters of these emulsions are predicted by a modified Hanai-Bruggeman formula for a wide range of volume fraction of oil Φ_1 and normality of the saline solution N. The emulsions are stable over a period of one week or more without remixing, i.e., the electrical constitutive parameters do not vary significantly during this period.

Figure 1 is a graph of the low-frequency constitutive parameters (relative permittivity $\bar{\varepsilon}_{rLF}$ and conductivity $\bar{\sigma}_{LF}$) with the volume fraction of the oil Φ_1 and the normality of the saline solution N as parameters. The solid lines show the predictions from the theoretical formulas. Results for 16 measured emulsions are shown as open dots; they are to be compared with the corresponding theoretical predictions, solid dots. From these data it is clear that this particular series of emulsions can be used to model materials with a wide range of electrical parameters $10 \lesssim \bar{\varepsilon}_{rLF} \lesssim 80$, 5×10^{-3} S/m $\lesssim \bar{\sigma}_{LF} \lesssim 5$ S/m.

As a practical example an emulsion of the type described above was used to represent red-clay earth in a one-third size scale model ($k_\ell = 0.333\ldots$). The frequencies in the full-sized system were 50 MHz $\leq f_f \leq$ 500 MHz, while those in the model were 150 MHz $\leq f_m \leq$ 1.5 GHz. The measured electrical parameters, ε_r and σ, for the clay are shown in Figure 2 (open dots). Note that the upper and right-hand scales on the graphs apply to the full-sized system. The electrical parameters for the model are also shown in this figure. Two sets of curves are presented, the theoretical predictions (solid line) and the measured results (dots). Note that the lower and left-hand scales on the graphs apply to the emulsion (scale model).

The emulsion is seen to closely match the properly scaled electrical constitutive parameters of the clay, including the dispersion in the conductivity, over the ten-to-one frequency range for which the clay was measured. As these results show, the series of emulsions developed should be particularly useful for representing geophysical materials in models for buried antennas, subsurface radars, etc.

Emulsion science is an active area of research with vast application in the food, pharmaceutical, cosmetics, etc., industries. The literature on this subject is voluminous. Only the simplest O/W emulsions were examined in this research; there are certainly other emulsions that have interesting electrical properties and possible electromagnetic

applications.

This research was also supported by the Air Force Rome Air Development Center.

- *A Scale Model for Studying Ground Penetrating Radars*

A scale model was developed for experimentally studying ground penetrating radars. The model is one-third full size and is used with transient signals that have significant frequency content within the range 150 MHz to 1.5 GHz. A unique feature is that the earth in the model is represented by an emulsion, which is a mixture of mineral oil, saline solution, and a stabilizing agent (described above). This emulsion is a scale model for red-clay earth; it matches the electrical parameters of the clay, including the dispersion in the conductivity, over a ten-to-one frequency range.

Typical results measured with the model include the pattern for the electric field transmitted by the radar into the earth and the radar signatures for pipes of various composition buried in the earth.

The model permits easy experimental comparison of different configurations for a radar and the eventual optimization of the radar for a particular application.

- *Pulse Excited Antennas*

Radars that use base band pulses, such as ground penetrating radars, require antennas that can radiate and receive temporally short, wide bandwidth pulses. These antennas have traditionally been analyzed using approximate methods, such as transmission line models or assumed current/aperture distributions. The antennas are generally electrically large (many wavelengths long) at the highest frequencies contained in the pulse. Thus, conventional frequency-domain numerical techniques, such as the method of moments, coupled with a fast Fourier transform are not efficient for analyzing these antennas.

The finite-difference time-domain method (FD-TD) is a numerical procedure used to solve Maxwell's equations directly in the time domain. While this technique has been used for several years, only recently has it become practical for solving three-dimensional structures due to the increases in the speed and storage of digital computers.

This technique (FD-TD) is being used to study antennas, like the TEM horn, for radiating wide bandwith pulses. The objectives are to obtain a better understanding of how pulses are radiated from such structures and to apply this knowledge to improve the performance of these antennas (lower internal reflections, reduce dispersion in the radiated pulse, etc). All theoretical predictions will be verified with accurate experimental measurements.

As a beginning a few simple antennas have been investigated – the cylindrical monopole and the conical monopole. These are rotationally symmetric antennas and therefore require only a two-dimensional analysis. Theoretical models formulated for these antennas are very good representations of the experimental models used for lab-

oratory measurements. The FD-TD method was used to determine the response of these antennas to a Gaussian pulse, and the results were compared with experimental measurements to determine the accuracy of the FD-TD method. Figure 3 shows the response of a conical monopole – the reflected voltage within the coaxial feed line for a 1V incident Gaussian pulse. The cone has half angle $\alpha = 30°$ and slant height $h/c = 4.37\tau_p$, where τ_p characterizes the width of the pulse. The solid line is the FD-TD results; the solid dots are the measured data.

These results are encouraging; they show that the FD-TD method can accurately predict the performance of these simple antennas. Of course, a full three-dimensional analysis of more complicated antennas, like the TEM horn, will offer new challenges.

- *Millimeter Wave Substrate Mounted Antennas*

Investigations of millimeter wave substrate mounted antennas have only recently begun to be undertaken. The lack of precise experimental data and loosely knit theories indicate more research is needed in this area.

The applications discussed in the literature for these substrate mounted antennas are primarily millimeter wave radars and imaging systems. We are studying these antennas for a different application – free space energy transfer at millimeter wavelengths. For this application the antenna is combined with a rectifying element (diode) and referred to as a "rectenna."

The rectenna concept was originally developed for use in the Solar Power Satellite System. This system proposed to beam microwave energy generated from solar powered satellites in a geosynchronous orbit to the Earth. The earth-based receiving portion of the system is an array of rectennas which converts the incoming microwave power to dc power. Current efficiencies for power conversion by a single rectenna element at 2.45 GHz are approaching 80%.

Applications for higher frequency (millimeter wave) rectennas have initiated research on integrated or monolithic rectenna elements. In this case many rectenna elements are fabricated on a single substrate using integrated circuit techniques. This is in contrast to the lower frequency rectennas where the antenna is essentially suspended in free space, and a discrete diode is used in the circuit. Many questions about the substrate mounted antennas and higher frequency rectifying circuits must be resolved before the integrated circuit rectenna becomes practical. The rectenna has been suggested for use in point to point energy transfer in space at millimeter wave frequencies. Higher frequency operation permits a small beamwidth making it feasible for satellites or remote aircraft to receive power from a distant power station.

The initial phase of this research is concentrated on the antenna portion of the rectenna system, with experiments being performed at a frequency of 230 GHz (λ_o = 1.30 mm). Simple substrate mounted, resonant dipole elements are being studied first. Note that these antennas differ from microstrip antennas in that the substrate is not grounded. The millimeter wave signal is being detected at the terminals of the dipole

with a bismuth bolometer, and the low-frequency modulation removed via an integrated low-pass filter and parallel strip transmission line. The field patterns and gains of the antennas are measured on a specially designed millimeter wave antenna range.

Preliminary measurements show that the measured patterns for the substrate mounted dipole are in good agreement with simple theoretical results (geometrical optics). Before more sophisticated antennas could be evaluated, refinements were required in the experimental apparatus. These include more precise fabrication for the bolometer detectors and more accurate positioning and signal measurement for the antenna range. A new evaporator has now been built for depositing the bismuth bolometers, and the antenna range has been automated with more precise positioners and a more accurate power meter.

The effect of substrate thickness on antenna patterns and gain is now being investigated. Future research will concentrate on dipole antennas combined with monolithic focusing elements for increased gain/efficiency.

This research is a coordinated effort between the School of Electrical Engineering and the Georgia Tech Research Institute (GTRI). NASA has provided low level support to GTRI for research on millimeter wave rectennas. The effort proposed here complements this research by including personnel from the School of Electrical Engineering. Faculty and graduate research assistants (doctoral students) from the School of Electrical Engineering are performing theoretical and experimental investigations of the substrate mounted antennas. The microfabrication facilities and millimeter wave test equipment at GTRI are used in the evaluation of the antennas.

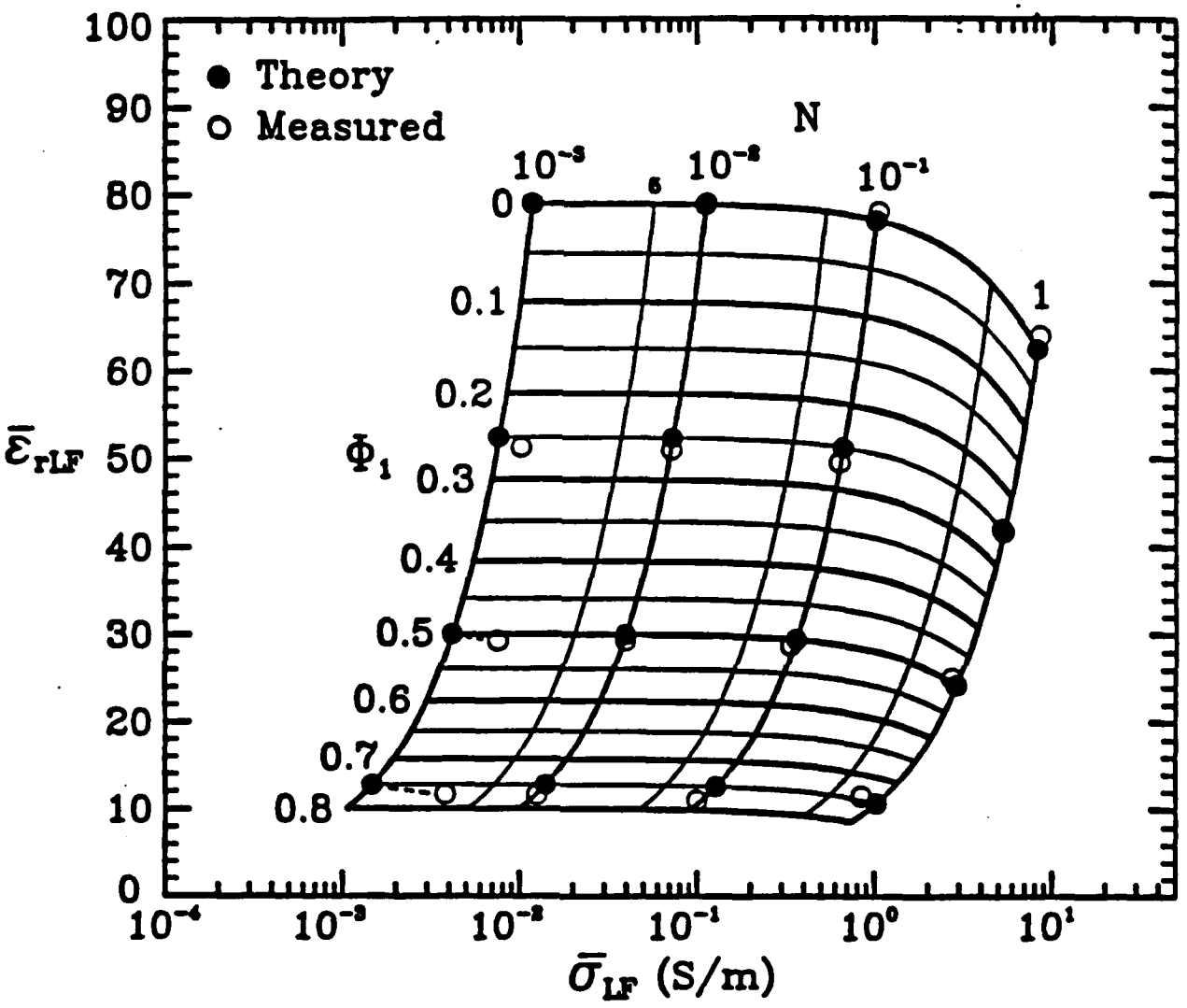


Fig. 1: Comparison of Theoretical and Measured Low-Frequency Electrical Constitutive Parameters of Mineral Oil, Saline Solution O/W Emulsions, $T = 23 \pm 1°C$.

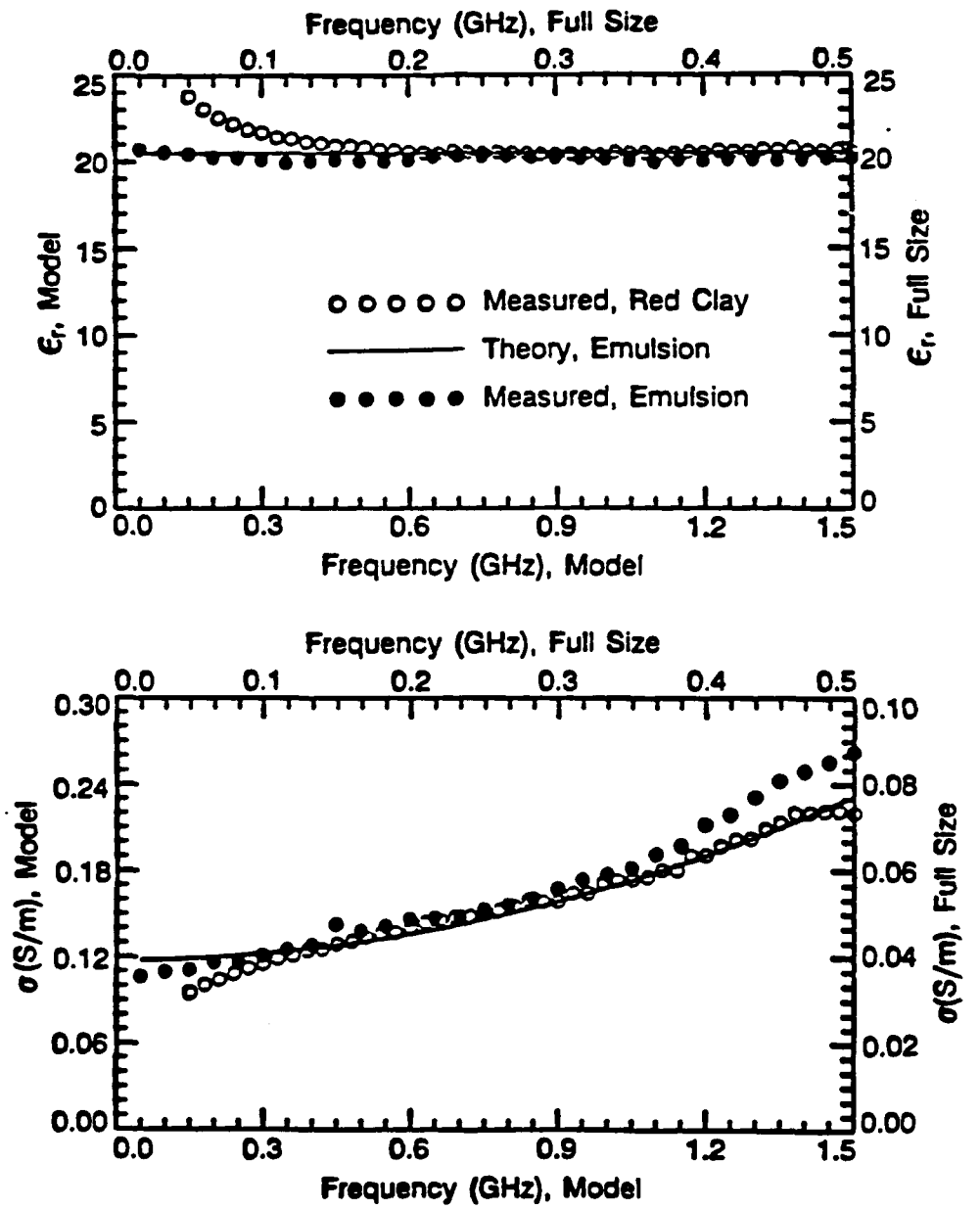


Fig. 2: Measured Constitutive Parameters for Red-Clay Earth (Full-Sized system), and Theoretical and Measured Constitutive Parameters for Emulsion (Scale Model), $T \approx 23°C$.

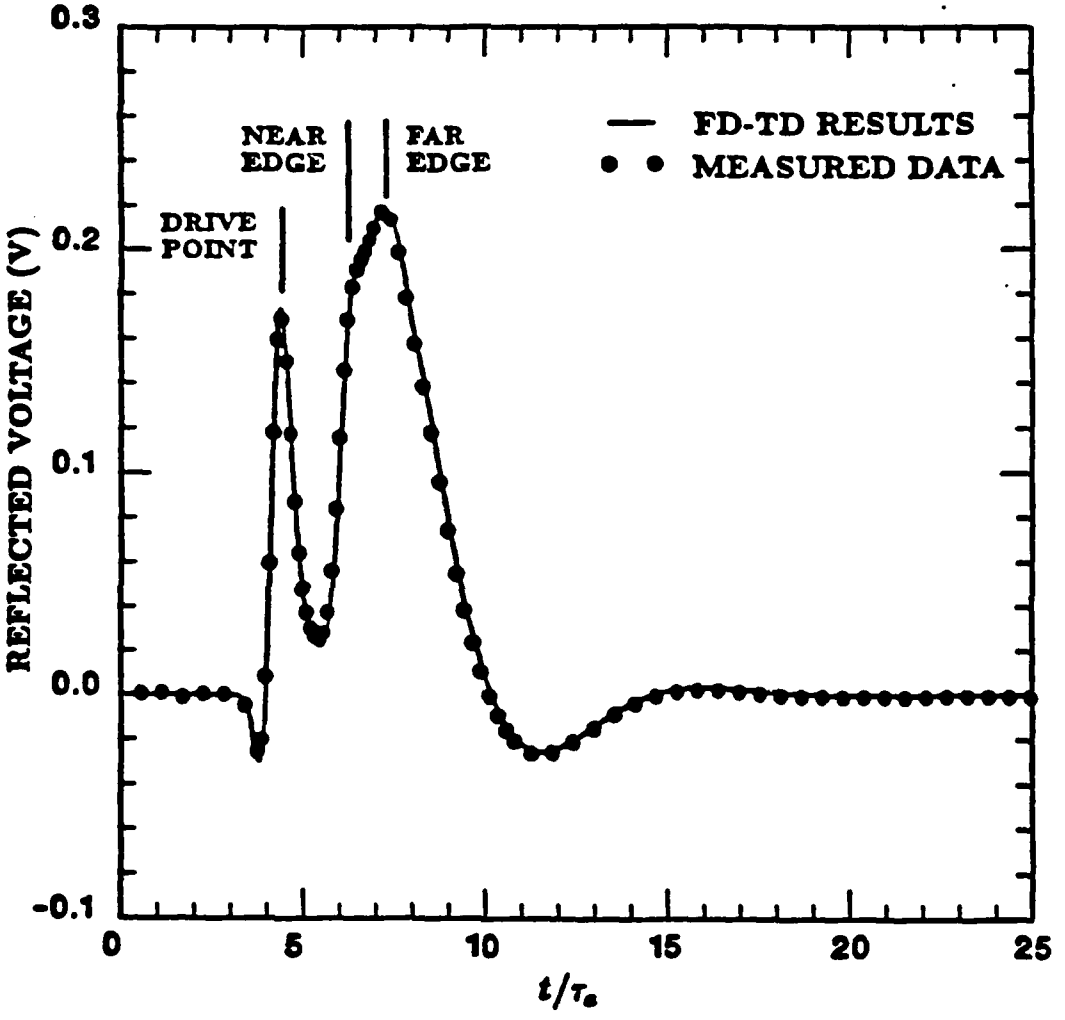


Fig. 3: A Comparison of the Theoretical (FD-TD) and the Measured Reflected Voltages in the Coaxial Line for a Conical Monopole Antenna Excited by a 1V Gaussian Pulse: $b/a = 2.30, h/a = 2.31 \times 10^{1}, h'/h = 8.63 \times 10^{-2}, \alpha = 30°$, and $\tau_p/\tau_a = 2.29 \times 10^{-1}$.

PUBLICATIONS:

Journal Articles Published or Accepted:

1. G.S. Smith and W.R. Scott, Jr., "The Use of Emulsions to Represent Dielectric Materials in Electromagnetic Scale Models," to be published in *IEEE Trans. on Antennas and Propagation*, March 1990.

2. G.S. Smith and W.R. Scott, Jr., "A Scale Model for Studying Ground Penetrating Radars," *IEEE Trans. Geoscience and Remote Sensing*, vol. 27, pp. 358-363, July 1989.

3. W.R. Scott, Jr., "A General Program for Plotting Three-Dimensional Antenna Patterns," to be published in *IEEE Trans. Antennas and Propagation Society Magazine*, pp. 6-11, December 1989.

4. J. G. Maloney, G. S. Smith, and W. R. Scott, Jr., "Accurate Computation of the Radiation from Simple Antennas Using the Finite-Difference Time-Domain Method," *IEEE Trans. on Antennas and Propagation*, August 1990.

Papers in Conference Proceedings:

1. J.G. Maloney, G.S. Smith, and W.R. Scott, Jr., "Accurate Computation of the Radiation from Simple Antennas Using the Finite-Difference Time-Domain Method," 1989 Spring URSI, IEEE AP-S Symposium, San Jose, CA, pp. 42-45, June 1989.

2. G. Zhou and G. S. Smith, "An Accurate Theoretical Model for the Thin-Wire Circular Loop Antenna," to be presented, 1990 Spring URSI, IEEE AP-S Symposium, Dallas, TX, 1990.

3. J. G. Maloney and G. S. Smith, "Implementation of Surface Impedance Concepts in the Finite Difference Time-Domain (FD-TD) Technique," to be presented, 1990 Spring URSI, IEEE AP-S Symposium, Dallas, TX, 1990.

Work Unit Seven

TITLE: Automated Radiation Measurements for Near- and Far-Field Transformations

SENIOR PRINCIPAL INVESTIGATOR:

E. B. Joy, Professor

SCIENTIFIC PERSONNEL:

R. E. Wilson, (Ph.D. Candidate)
Mike G. Guler, Graduate Research Assistant (Ph.D. Candidate)
Donald Black, Graduate Research Assistant (Ph.D. Candidate)
Angela R. Dominy, Graduate Research Assistant (Ph.D. Candidate)
Joe Epple, Graduate Research Assistant (Ph.D. Candidate)
Ashley L. Slappy (M.S. Candidate)
John R. Dubberley (Undergraduate)
Scott C. Waid (Undergraduate)

SCIENTIFIC OBJECTIVE:

The long term objective of this research is to understand the near field and far field coupling between antennas in the presence of scatterers. Special emphasis is placed on determination of limits of accuracy in the measurement of the fields radiated or scattered by an antenna-under-test by a second antenna and to develop techniques and computer algorithms for compensation of such measurements due to known geometrical or electromagnetic anomalies.

Three application areas are pursued: a) antenna measurements, where the effects of scatterers are suppressed or compensated, b) scattering measurements, where the effects of scatterers are enhanced, and c) radome measurement, where the effects of the scatterer (the radome) are of equal importance to the antenna measurement.

STATE OF THE ART:

The state-of-the-art in antenna pattern measurement, radar cross-section measurement and in radome measurement is rapidly advancing. The rapid advance has been forced in an attempt to keep up with the rapidly advancing state-of-the-art in antennas such as: low sidelobe level phased array antennas, conformal arrays, and adaptive arrays; the rapidly advancing state-of-the-art in radar cross section reduction: stealth; and the rapidly advancing state-of-the-art in radomes and radome materials, especially for high speed aircraft and missiles. Measurements are increasingly being used not only to verify performance of these state-of-the-art antennas, reduced radar cross-sections and radomes, but are also used to learn about the fields on, in and near the antenna,

scatterer or radome. This backward transform of measurements from the measurement surface to the antenna, scatterer or radome developed under JSEP sponsorship, is a major new thrust in electromagnetic measurements and has already lead to significant improvements in phased array antennas and radomes. This area of research is in its infancy and will become a powerful tool for the study of advanced antennas, radar cross-section signature identification and reduction, and the scattering and coupling of electromagnetic fields into and from complex structures. We have already made significant contributions toward the development of the theory and techinque for the electromagnetic characterization of complex shaped radome materials using backward transformed, spherical surface, near-field measurements.

RESEARCH ACCOMPLISHMENTS:

- *Near-Field Measurement of Radome Anomalies*

Radome performance assessment using spherical near field measurements has been an ongoing research project at Georgia Tech for the past five years. Radome flaws such as wall thickness variations or inhomogeneities can cause serious pattern degradation. A method has been developed to locate such anomalies using spherical near field range measurements. A measurement sphere is chosen such that it completely encloses the radome. Since the near field probe's response to the evanescent fields is unknown, the measurement sphere should be at least several wavelengths beyond the radome surface so these fields can decay to an insignificant level. A broad beamwidth source is used to completely illuminate the radome from within. Tangential electric fields are sampled on the measurement sphere with and without the radome in place. The data is then backward transformed onto the surface dimensions of the radome. The backward transform is performed the same as a forward transform except more care must be taken in the number of spherical modes used. Measurement noise can result in small, but non-zero, modal coefficients for high order modes that should not exist within the minimum radius enclosing the radome. In a forward transform, these modes decay and are negligible. In a backward transform they can increase exponentially and cause large errors. A calculation must be performed to determine the expected number of non-zero modes for a given backward transform radius, then the higher order modal coefficients are set to zero to prevent noise generated errors. After backward transforming the two data sets, they are subtracted to identify differences. The phase difference over the radome surface is a measure of the phase shift introduced by the radome. Variations in this phase shift represent the undesirable variations in the radome wall. Figure 1 shows a phase subtraction between a flawed radome and no radome. The overall phase shift of about 100 degrees is the phase shift caused by the radome. The two distinct bumps are 1 inch diameter dielectric disks placed on the radome surface to simulate flaws. The increase in phase shift near the $\theta = 0$ line is due to the increased thickness of the ogival shaped radome's tip.

Such an analysis could be used as a non-destructive, non-intrusive method for acceptance testing of radomes. The phase difference plot could also be used as a map for computer controlled surfacing of the radome's interior to correct electrical variations.

Two results stemming from this research were published in 1989. The first was the demonstrated results of a new hybrid plane-wave/spherical mode backward transform technique to determine antenna aperture field distributions from spherical surface near-field measurements. These near-field distributions are required for radome analysis and design and have been impossible to obtain with accuracy until now.

The second publication shows the results of a major improvement to radome analysis algorithms. The paper presents the theory and results of the inclusion of inner radome reflections on the predicted boresight error of a radome. The results show that the radome induced boresight error depends heavily on this neglected electromagnetic scattering mechanism.

PUBLICATIONS:

Books or Chapters in Books:

1. E. B. Joy, "Near Field RCS Measurement Ranges," Chapter 12, *Techniques of Radar Reflectivity Measurement*, N. C. Currie (Ed.), Artech House, 1989.

Conference Proceedings:

1. M. G. Guler, E. B. Joy, and D. N. Black, "Planar Surface Near-Field Data Determined from Spherical Surface Near-Field Measurement," *Proceedings of the 19898 Antenna Measurement Techniques Association Meeting*, Monterey, CA, October 9-13, 1989, pp. 14-9 through 14-12.

2. J. W. Epple, Jr., E. B. Joy, M. G. Guler and R. E. Wilson, "Effect of Radome Wall Reflections on Boresight Error," *Proceedings of the 3rd DoD EM Windows Symposium*, Huntsville, AL, November 14-16, 1989. (not yet published).

Invited Seminars:

1. E. B. Joy, "Antenna Measurement Techniques," Atlanta Chapter Meeting of International Test and Evaluation Association, Atlanta, Georgia, November 10, 1989.

2. E. B. Joy, "Dipole Arrays," Bendix Corporation, Columbia, Maryland, February 15, 1989.

Work Unit Eight

TITLE: Angular Spectrum Analysis for Non-Uniform Arrays

SENIOR PRINCIPAL INVESTIGATOR:

James H. McClellan, Professor

SCIENTIFIC PERSONNEL:

D. Lee, Graduate Research Assistant (Ph.D. Candidate)
K. Blanton, (Ph.D. Candidate)
L. Fertig, (Ph.D. Candidate)

SCIENTIFIC OBJECTIVE:

The objective of this research is to develop new methods of multi-dimensional spectrum estimation that are applicable to space-time data recorded from sensor arrays (e.g., antennas, sonar, etc.). Several major areas are being emphasized: wideband processing, adaptive arrays, non-uniform array geometries, 2-D arrays, and large distributed arrays. New techniques for wideband processing are being sought in the area of coherent processing. The potential application for non-uniform, or random, arrays includes linear arrays with failing sensors, or with sensor position uncertainties. In the case of very large arrays, the problems of interest are direction finding and/or source localization algorithms using localized processing with (possibly) non-uniform grouping of sensors. In all cases, we are attempting to model realistic effects that are encountered in actual arrays in the field. Models of various array imperfections are needed to design more robust processing schemes and to derive array calibration techniques.

RESEARCH ACCOMPLISHMENTS:

- *Signal Estimation for Non-Uniform Arrays*

The Steiglitz-McBride method [5] of iterative exponential signal modeling has been extended to the case of missing measurements or non-uniform arrays. This situation is applicable to array processing in the case where a linear array has been thinned by removing sensors, or where some sensors in the linear array have failed. The Steiglitz-McBride method (also known as Iterative Pre-filtering) was originally formulated as a filtering problem and, hence, demands a regularly spaced grid with no missing measurements. In the case of missing measurements, an interpolation step must be added to the iterative algorithm. This step can be inserted by using the exponential model that is being built during the iteration. This strategy is similar to an extension of the E-M algorithm recently published [7] for the non-uniform array case. In the course of this work, we have also discovered that the Steiglitz-McBride iteration is *exactly* equivalent

to several existing algorithms, including the Iterative Quadratic Maximum Likelihood (IQML) algorithm [1]. IQML was recently re-discovered, but has also been published under other names by various authors over the past 15 years [2,4]. The Steiglitz-McBride method (1965) pre-dates all of these.

- *Dual Form Adaptive Filtering*

A new form of the classic LMS (Wiener) adaptive filter [3] has been derived by solving the iterative minimization problem in its dual form. This dual algorithm has an advantage when the number of parameters in the array weight vector is much larger than the number of constraints, because the dimensionality of the update vector in the dual-form adaptive filter is equal to that of the constraint space. The resulting algorithm has convergence properties that are similar to a Recursive Least-Squares (RLS) adaptive filter. An extensive analysis of the convergence issue as well as misadjustment has been carried out. These will be presented, along with a derivation of the dual form, in a paper to be presented at the 1990 *International Conference on Acoustics, Speech and Signal Processing.*

- *Minimum Redundancy Array Design*

This thesis project is directed at a study of the theory of non-uniform array design when the criterion of interest is to minimize the number of sensors, while maximizing the array aperture. The constraint is to measure all the correlation lags possible from 1 to the aperture length. Only a few theoretical results are available for the general case, because the problem seems to need exhaustive enumeration. Therefore, most practical results have been restricted to the problem of generating useful, but suboptimal, arrays. We have succeeded in generating, by exhaustive search, some large arrays (e.g., length 58 with 13 sensors), but this has been done primarily in an attempt to disprove some conjectures made by others. As a part of this exercise we have developed some simplifications needed to pare down the search space when generating larger arrays.

- *Time-Space Domain Processing*

We are investigating methods of 2-D filter design to develop angle-of-arrival filters for wideband processing, and to develop new structures for wideband adaptive arrays. Presently, many methods for wideband processing are based on extensions of narrowband processing, (e.g., focusing). The implementation is performed in the frequency domain by processing one frequency at a time. One exception is the coherent signal-subspace method [6] in which a focusing transformation is used to map various frequency components to a common reference frequency. If this method is recast as a 2-D (time-space) filtering operation, a new structure results. Thus there is a different implementation strategy in which the focusing transformation is expressed as a block-

ing matrix followed by a steering matrix. The resulting design strategy is based on 1-D FIR filter design for realizing the fractional delays needed in the blocking matrix. This result will be presented at the *International Symposium on Circuits and Systems*. Since the wideband operation needed is a constant AoA filter, these structures can also be be related to the fan filter.

- *Related Work*

We are also investigating the use of symbolic algebra methods for DSP via the program Mathematica. Some results on a multi-dimensional z-transform package will be presented at the 1990 *International Conference on Acoustics, Speech and Signal Processing*. We are planning to extend this computer tool to handle the analysis of fan filters and other common 2-D operators.

PUBLICATIONS:

Journal Articles Submitted

1. D. Lee and J. H. McClellan, "Exact equivalence of Steiglitz-McBride and IQML," submitted to *IEEE Trans. Acoust., Speech, Signal Processing.*

Conference Proceedings:

1. L. P. Heck, D. A. Schwartz, R. M. Mersereau and J. H. McClellan, "Symbolic Simplification of Digital Signal Processing Software," *Proc. 1989 International Symposium on Circuits and Systems*, Portland, OR, May, 1989.

2. L. B. Fertig and J. H. McClellan, "A dual form adaptive filter," *Proc. 1990 International Conference on Acoustics, Speech and Signal Processing*, Albuquerque, NM.

3. K. Baudendistel and J. H. McClellan, "Code generation for the AT&T DSP32," *Proc. 1990 International Conference on Acoustics, Speech and Signal Processing*, Albuquerque, NM.

4. B. L. Evans, J. H. McClellan, and W. B. McClure, "Symbolic Z-transforms using DSP knowledge bases," *Proc. 1990 International Conference on Acoustics, Speech and Signal Processing*, Albuquerque, NM.

5. D. Lee and J. H. McClellan, "Decomposition of the focusing transformation in coherent signal subspace DOA estimation," accepted *ISCAS-90.*

6. J. H. McClellan and D. Lee, "Exact equivalence of iterative prefiltering and IQML," accepted *ISCAS-90.*

References

[1] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of super-imposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1081-1089, October 1986.

[2] A. G. Evans and R. Fischl, "Optimal least squares time-domain synthesis of recursive digital filters," *IEEE Trans. Audio Electroacoustics*, vol. AU-21, pp. 61-65, February 1984.

[3] O. L. Frost, "An algorithm for linearly constrained adaptive beamforming," *IEEE Trans. Antennas and Propagation*, AP-30, pp. 27-34, Jan. 1982.

[4] L. L. Scharf, R. Kumaresan and A. K. Shaw, "An algorithm for pole-zero modeling and spectral analysis," *IEEE Trans. Acoust., Speech and Signal Processing*, vol. ASSP-34, pp. 637-640, June 1986.

[5] K. Steiglitz and L. E. McBride, "A technique for the identification of linear systems," *IEEE Trans. Automatic Control*, pp. 461-464, October 1965.

[6] H. Wang and M. Kaveh, "Coherent signal subspace processing for the detection and estimation of angles or arrival of multiple wideband sources," *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-33, no. 4, pp. 823-831, Aug. 1985.

[7] A. Weiss, B. C. Levy and A. S. Willsky, "Non uniform array processing via the polynomial approach," *IEEE Trans. Aerospace and Electronic Systems*, Jan. 1989.