1.0  4.5  2.8  2.5
     5.0
     5.6
1.1  6.3  3.2  2.2
          3.6
          4.0  2.0
               1.8
1.25  1.4  1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

| REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|

| 1. REPORT NUMBER AFOSR-TR- 86-1000 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
|---|---|---|

| 4. TITLE *(and Subtitle)* "Optical Data Processing" | 5. TYPE OF REPORT & PERIOD COVERED Annual Report 9/30/8⅜ - 9/30/85 |
|---|---|
|  | 6. PERFORMING ORG. REPORT NUMBER |

| 7. AUTHOR(s) David Casasent | 8. CONTRACT OR GRANT NUMBER(s) AFOSR-84-0293 |
|---|---|

| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Department of Electrical and Computer Engineering Pittsburgh, PA 15213 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2305/B1 |
|---|---|

| 11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR/NE, Dr. C. Lee Giles Building 410 Bolling AFB, Washington, D.C. 20332 | 12. REPORT DATE 1 October 1985 |
|---|---|
|  | 13. NUMBER OF PAGES 136 |

| 14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* Same as 11 | 15. SECURITY CLASS. *(of this report)* Unclassified |
|---|---|
|  | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A |

16. DISTRIBUTION STATEMENT *(of this Report)*

Unlimited

**DISTRIBUTION STATEMENT A**

Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

Unlimited

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Acousto-Optic Processors, Chords, Computer Generated Holograms, Correlation, Feature Extraction, Fourier Coefficient, Moments, Optical Linear Algebra, Optical Pattern Recognition, Synthetic Discriminant Functions.

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Research on optical data processing for missile guidance and robotics is described. Our major emphasis is pattern recognition using feature extraction (Fourier coefficients, moments and chord features) and correlation (using distortion-invariant synthetic discriminant function matched spatial filters). All research in pattern recognition concerns multi-class distortion-invariant pattern recognition. Recent research includes: different feature extraction post-processors, new algorithms to extract distortion parameters from chord features and a hierarchical moment feature processor for distortion parameter

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

estimation. Extensive database tests of moments and synthetic discriminant functions have been performed. Component research has addressed AO cells with performance measures and detector effects described. Matrix-vector research includes: error source analysis, a new quadratic matrix algorithm, and initial laboratory system results with attention to the electronic support system and the laboratory system fabrication.

AFOSR·TR· 86-1000

ANNUAL REPORT

# OPTICAL DATA PROCESSING - 1

SUBMITTED BY:

**David Casasent**, Principal Investigator
Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, PA 15213
Telephone: (412) 578-2464

SUBMITTED TO:

Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C. 20332

ATTN: Dr. C. Lee Giles, AFOSR/NE

Date: 1 October 1985

Period Covered
30 September 1983 - 30 September 1985

| Accesion For | | |
|---|---|---|
| NTIS CRA&I | | ☒ |
| DTIC TAB | | ☐ |
| Unannounced | | ☐ |
| Justification | | |
| By | | |
| Distribution/ | | |
| Availability Codes | | |
| Dist | Avail a· d / or Special | |
| A-1 | | |

COPY INSPECTED

86 11 25 337

# Table of Contents

# ABSTRACT

Research on optical data processing for missile guidance and robotics is described. Our major emphasis is pattern recognition using feature extraction (Fourier coefficients, moments and chord features) and correlation (using distortion-invariant synthetic discriminant function matched spatial filters). All of our research in pattern recognition concerns multi-class distortion-invariant processors. It includes new algorithms to extract distortion parameters from chord features and a hierarchical moment feature processor for distortion parameter estimation. Extensive database tests of moments and synthetic discriminant functions have been performed. Component research has addressed AO cells with performance measures and detector effects described. Matrix-vector research includes: error source analysis, a new quadratic matrix algorithm, and initial laboratory system results with attention to the electronic support system and the laboratory system fabrication.

# 1. INTRODUCTION

During this first year (September 1984 - September 1985) of our new research contract in optical data processing for missile guidance, we have addressed the major key issues and aspects required and associated with this technology. This research includes:

- real-time devices and components,
- new system architectures,
- new algorithms,
- new high-speed general-purpose optical data processing techniques and systems,
- tests on new and extensive image databases,
- plus new pattern recognition techniques, architectures, algorithms and concepts.

As in past years, we have been quite faithful in reporting our *AFOSR* sponsored research in various journals and conference publications. 24 publications (an average of 2 per month) have resulted from this *AFOSR* research (Chapter 18). Copies of the more relevant papers we have published over the past year are included as various chapters of this report. These are included to provide complete documentation of the different aspects of our work.

In Chapter 2, we provide a summary and overview of our research progress achieved during the past year. This work addressed 6 vital areas of optical data processing research:

1. real-time spatial light modulators (Section 2.2 and Chapter 3),

2. optical pattern recognition (Section 2.3 and Chapter 4),

3. computer generated holograms (Section 2.4 and Chapter 5),

4. optical feature extraction (Section 2.5 and Chapters 6-10),

5. optical correlation (Section 2.6 and Chapters 11-14), and

6. optical linear algebra processors (Section 2.7 and Chapters 15-17).

Topic (1) concerns the vital issue of real-time spatial light modulators. Topics (2)-(5) address pattern recognition for ATR using new optical pattern recognition (OPR) techniques. In this work, we have been faithful to address vital problems such as multi-class distortion-invariant pattern recognition of military targets, the acquisition and importance of the use of a large database and the effects of noise on the algorithms used. Topic (6) concerns the most attractive item in optical processing at present and a potentially quite general-purpose optical processor concept.

Details of the more salient results of our research are provided in Chapters 3-17. In Chapter 18, we enumerate our *AFOSR* sponsored publications, the presentations given on this research at conferences and seminars during the past year, and the Master's and PhD students that this grant has supported.

Our level of *AFOSR* research support on this grant has not increased for several years and our optical artificial intelligence separate research *AFOSR* proposal was not funded. This will significantly impact our research program. Other funds are being sought to allow support of this research we feel is necessary. The aforementioned remarks, plus the unavailability of funding from Eglin AFB for our Kalman filtering research are expected to result in a reduction in the quantity of research we are able to produce for *AFOSR*. We anticipate that we will still remain considerably above the output level of other researchers however.

During the past year, the principal investigator (PI) presented invited talks on our *AFOSR* sponsored research at various conferences including the *Critical Review of Technology SPIE Conference* on *Digital Image Processing* and the *Critical Review of Technology SPIE Conference* on *Computer Generated Holograms* (*SPIE*, Los Angeles, California, January 1985) and the *DoD* conference on *Parallel Algorithms and Architectures for ATR* (Leesburg, Virginia, conference proceedings published February 1985), plus other *OSA* and *SPIE* optical computing and robotic conferences during the year. The PI has chaired conference sessions and seminars and served on the organizing committees for the following conferences and topics:

- *SPIE* (robotics),
- *Optical Society of America* (optical computing),
- *Optical Society of America* (machine vision),
- *SPIE* (digital image processing),
- *SPIE* (computer generated holograms).

The PI was also guest editor of a special issue of *Optical Engineering* on robotics and computer vision. He was invited to submit papers to the journal *Optical Engineering* special issues on pattern recognition (November 1984), optical computing (January 1985) and computer generated holograms (October 1985).

# 2. OVERVIEW AND SUMMARY

## 2.1 INTRODUCTION

Our six major research areas and our recent progress in each are highlighted in Sections 2.2 - 2.7. Details of each aspect of our fifteen work topics follows in Chapters 3 - 17.

## 2.2 SPATIAL LIGHT MODULATORS (ACOUSTO-OPTIC CELLS, CHAPTER 3)

Recently, our spatial light modulator research has emphasized acousto-optic cells. In [1], we considered the salient acousto-optic architectures (spectrum analyzers and correlators). The various acousto-optic cell and acousto-optic architecture component errors have been enumerated, grouped into different classes and combined into several new models. New performance measures for acousto-optic correlators and spectrum analyzers were defined and detailed (spectrum estimation, delay estimation, and detection). Each is an appropriate performance measure for a different application. General error-free formulae for each of these performance measures were derived and the performance obtained with each was described and quantified as a function of the various system parameters. Our new work [2] in this area (Chapter 3) addressed component error source effects on performance (specifically detector effects). We plan to apply AO processors to optical image processing in our future research.

## 2.3 OPTICAL PATTERN RECOGNITION REVIEWS (CHAPTER 4)

Our *AFOSR* optical pattern recognition research is at the forefront. Our paper [3] in Chapter 4 on coherent optical pattern recognition was included in the recent *Critical Review of Technology* series on Digital Image Processing. A more recent review [4] was one of only two optical pattern recognition papers at a recent *DoD* conference on parallel architectures and algorithms for ATR. A journal OPR paper was invited and published in the *Optical Engineering* issue on optical computing [5] in January 1985. Chapter 4 [3] is a complete review of optical techniques for feature extraction and correlation and includes new algorithms, architectures and hybrid optical/digital processing concepts. Sections 2.4-2.6 and Chapters 5-14 detail specific aspects of our recent OPR research.

## 2.4 CGHs FOR OPR (CHAPTER 5)

Our 1984-1985 research has increased the use of computer generated holograms (CGHs) for optical pattern recognition (OPR). We were selected to present a review of this area [6] in a recent *Critical Review of Technology* conference on CGHs. A detailed revised version [7] of this paper was invited for submission to a upcoming journal special issue. This review will be included in our 1985-1986 annual report. In Chapter 5, we include new recent work on the use of a CGH as a wedge ring detector for diffraction pattern sampling [8].

## 2.5 OPTICAL PATTERN RECOGNITION FEATURE EXTRACTION (CHAPTERS 6 - 10)

Three new optical feature extraction techniques have been detailed in our recent research:

1. the use of multiple feature extractors and dimensionality reduction techniques (we consider the specific case of a wedge ring detector-sampled optically produced Fourier transform feature space) (Chapter 6 and Ref.[9]);

2. a new method to measure distortions from a chord distribution feature space (Chapter 7 and Ref.[10]); and

3. a hierarchical two-level hybrid optical/digital moment feature processor (Chapter 8 and Refs.[11] and [12]).

Our optical Fourier transform space and multiple feature space work (Chapter 6) includes four different dimensionality reduction and feature extraction techniques. A new classifier concept, quantitative data on the importance of amplitude versus phase Fourier coefficients (for pattern recognition, rather than image reconstruction) and the performance of each in the presence of noise. These represent quite novel results which have thus far not been published for any other feature extractor (optical or digital). Experimental results for two letters and two vehicles with 25 images of each at different scale and in-plane rotational differences were obtained. In Chapter 7, new techniques to obtain distortion parameters from chord features are detailed [10].

In Chapter 8, our new hybrid optical/digital moment processor, our new hierarchical moment-based class estimator technique, and a new two-level classifier using moments are detailed and the results obtained on a set of ship images are presented [11]. Robotic part data on the same system are contained in Ref.[12]. The performance of this system on non-controlled imagery and a new segmentation

processing technique were recently published [13] and are included in Chapter 9 for completeness. The accuracy with which the distortion parameter estimates can be obtained is summarized [14] in Chapter 10.

## 2.6 OPTICAL PATTERN RECOGNITION CORRELATORS (CHAPTERS 11-14)

Our distortion-invariant multi-class multi-object correlator research emphasizes synthetic discriminant functions (SDFs). Our tests and algorithms for projection SDFs on ship images with data on noise performance with new guidelines for the suggestion of projection values were included in a recent journal special issue on pattern recognition [15] and are provided in Chapter 11. New related SDFs that optimize various performance measures [16] are detailed in Chapter 12. New correlation SDFs have been described and initial results with them have been obtained for a tank and APC image database [17]. These results are summarized in Chapter 13. We were directed to perform tests on aircraft images by *AFOSR*. These results [18] are included in Chapter 14.

## 2.7 OPTICAL LINEAR ALGEBRA PROCESSORS (CHAPTERS 15 - 17)

This optical data processing application area has received very much recent attention.

A first vital aspect of optical linear algebra research that we initiated was the error source modeling and simulation of OLAP (optical linear algebra processor) architectures and algorithms [19]. Chapter 15 details this work. A second novel facet of our OLAP research has concerned specific applications. The application chosen for major attention was Kalman filtering and the specific application of it was missile guidance, control and state estimation. Support for future research in this area is questionable at present. A third facet of our research is new parallel algorithms. A new parallel algorithm for the solution of quadratic nonlinear matrix equations using a finite number of steps has been devised [20] and is detailed in Chapter 16.

The fourth and final aspect of our OLAP research has been attention to fabrication of an OLAP. We recently [21] discussed our laboratory processor and its electronic support and initial results. This is detailed in Chapter 17. A lengthy version of this work is in preparation for a journal special issue. This

is one of the few laboratory results published on OLAP processors. We are thus quite novel in the work we have generated on each of these OLAP areas.

# 3. DETECTOR EFFECTS IN REAL-TIME AO PROCESSORS

Ref. 2

# Detector effects on time-integrating correlator performance

Anastasios Goutzoulis, David Casasent, and B. V. K. Vijaya Kumar

Detector array effects are considered for a time-integrating acoustooptic correlator used for signal detection. Effects such as detector area integration, detector element, spatial response, and the location of the correlation peak within a detector element are included. General SNR, $P_D$, and $P_{FA}$ expressions are derived as a function of various system and detector parameters. Quantitative data are provided for a Gaussian-Markov signal, and initial experimental confirmation is included.

## I. Introduction

Acoustooptic (AO) devices have been suggested for use in many new signal processing architectures and applications.[1] This interest is motivated by the commercial availability, good reliability, and performance of new AO cells.[1,2] One of the most attractive AO signal processors is the time-integrating (TI) correlator.[3,4] This architecture is attractive because of the large processing gain it provides and the large signal deviations it can accommodate. However, only limited statistical analyses,[5] error source consideration,[6] and quantitative performance data have been published on this system. Published work has considered the effects of signal time bandwidth product (TBWP), input signal noise, detector noise,[4] and finite detector area effects on time delay estimation applications.[7]

In this paper we consider detector effects in a signal detection application of a TI correlator. We consider correlators using AO cells operated in the linear intensity mode (since these architectures yield analytical results). Detector effects for AO cells operated in the amplitude mode can be analyzed following the procedures and models advanced herein. In Sec. II, we review the linear intensity TI correlator and derive an expression for its output including the finite area $D$ of each detector element (and the associated spatial integration and sampling), the spatial weighting function

When this work was done all authors were with Carnegie-Mellon University, Department of Electrical & Computer Engineering, Pittsburgh, Pennsylvania 15213; A. Goutzoulis is now with Westinghouse Research & Development Laboratories, 1310 Beulah Road, Pittsburgh, Pennsylvania 15235.

$w_n(\tau)$ for an individual detector, and conventional signal and system parameters. Our performance measures used are probability of detection $P_D$ and probability of false alarm $P_{FA}$. Our prior statistical analysis[6] related these to measurable correlator SNR values and showed that these factors completely characterize the system's performance. We do not consider detector noise and detector element cross talk, since earlier detector noise analyses[4] can easily include such effects. The statistics (mean and variance) of the correlator's output are then evaluated in Sec. III for the case of Gaussian-distributed signal and noise. In Sec. IV, performance expressions are derived for the case of Gaussian-Markov signals. The effect of the finite detector area (Sec. V), the location of the correlation peak within a detector element (Sec. VI), and spatial weighting across each detector (Sec. VII) are then analyzed, and quantitative analytical results are provided. Brief experimental results are included (Sec. VIII), and then our summary and conclusions are advanced (Sec. IX) on the design of a TI correlator for detection applications. Signal, systems, and output detector parameters are considered throughout. Emphasis is given to our general analyses, the quantitative effect of different parameters, and the analyses of various initial quantitative results.

## II. TI AO Correlator for Signal Detection

A simplified schematic of a linear intensity TI AO correlator[3] is shown in Fig. 1. We denote the reference signal by $s(t)$ and the received signal by $s(t - \tau_0) + n(t)$, where $\tau_0$ is the delay and $n(t)$ is additive noise. For linear intensity modulation[3] of the AO cells, the signals are added to two biases $B_1$ and $B_2$. The signal driving the point modulator at $P_1$ is

$$s_2(t) = B_2 + s(t - \tau_0) + n(t). \tag{1}$$

The light intensity leaving $P_1$ is then proportional to $s_2(t)$. This light beam is expanded by lens $L_1$ and
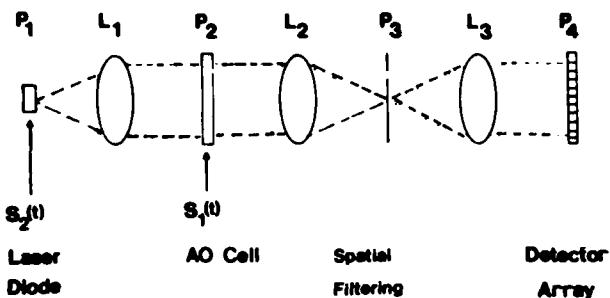
Fig. 1. Schematic of a time-integrating acoustooptic correlator.

uniformly illuminates the AO cell at $P_2$. The signal $s_1(t) = B_1 + s(t)$ modulates a rf carrier and drives the AO cell at $P_2$. Lenses $L_2, L_3$ and the spatial filter at $P_3$ separate the diffracted and undiffracted orders, block the undiffracted order, and image the +1 diffracted order onto a linear detector array at plane $P_4$. The detector array at $P_4$ provides the time integration over $T_I$ of the resulting light intensity $s_1(t)s_2(t)$. Including the finite area $D$ of the detector elements, we write the $P_4$ output from the $n$th detector as

$$I(n) = \frac{1}{T_I} \int_{-T_I/2}^{T_I/2} \int_{(n-1/2)D}^{(n+1/2)D} w_n(\tau)[B_1 + s(t - \tau)]$$
$$\times [B_2 + s(t - \tau_0) + n(t)]d\tau dt, \quad (2)$$

where $\tau = x/v_s$, $x$ denotes the direction of the sound propagation, $v_s$ is the speed of sound in the AO crystal, $w_n(\tau)$ is the spatial response weighting function for the $n$th detector element, and $n = -N/2, \ldots, 0, \ldots, N/2$ is the index for the $N + 1$ detectors. We note that $D = D_s/v_s$ has units of time. ($D_s$ is the detector area in distance units.) The normalization factor $1/T_I$ is included to simplify our results and does not affect the system's detection performance.

Equation (2) contains all parameters necessary to study the effects of all detector parameters [i.e., $D, w_n(\tau)$ as well as the value of $\tau_0$ with respect to $D$] and various signal and system parameters (such as $T_I$, signal bandwidth, and TBWP) on the system's performance. Other AO system component errors can be treated individually in the input or frequency plane as shown earlier.[6] Dead spaces between detector elements can be included by allowing $w_n(\tau)$ to become zero at the edges of each detector element. In our analyses, we assume 1:1 imaging from $P_2$ to $P_4$ in Fig. 1. Operation of the AO cell in the linear amplitude mode is also possible. In this case, the correlation output is present on a spatial carrier, and after postdetection processing the correlation obtained is still given by Eq. (2) with a different signal-to-bias ratio. Thus, our results can be extended to apply to both amplitude and intensity mode AO cell operation.

As performance measures, we use the parameters[6] $SNR_1$, $SNR_2$, $P_D$, and $P_{FA}$. $SNR_1$ is the typical SNR measure[8] used in communications (the square of the ratio of the average correlation value at the peak to the

standard deviation in the peak value). $SNR_2$ is the same as $SNR_1$ except the standard deviation is computed far from the peak. (It is thus similar to the peak-to-sidelobe ratio.[9]) The probability of detection $P_D$ and probability of false alarm $P_{FA}$ are related to these two SNR measures by[6]

$$P_D = \frac{1}{\sqrt{2\pi E^2[C(0)]/SNR_1}} \int_\theta^\infty \exp\left(\frac{-SNR_1|x - E[C(0)]|^2}{2E^2[C(0)]}\right)dx, \quad (3)$$

$$P_{FA} = \frac{1}{\sqrt{2\pi E^2[C(0)]/SNR_2}} \int_\theta^\infty \exp\left(\frac{-SNR_2|x - E[C(\tau)]|^2}{2E^2[C(0)]}\right)dx, \quad (4)$$

where $E[C(0)]$ and $E[C(\tau)]$ are the means of the signal and noise, respectively, and $\theta$ is the detection threshold. By increasing $\theta$, $P_{FA}$ will be reduced, but $P_D$ will also decrease. Note that $E[C(\tau)]$ and $E[C(0)]$ can be estimated by evaluating the correlation $C(\tau)$ far from the peak $\tau \gg 0$ and at the peak $\tau = 0$, respectively. We choose to express $P_D$ and $P_{FA}$ in terms of $SNR_1$ and $SNR_2$ because of the considerable ease with which these two SNR terms can be measured experimentally on an optical correlator. In our statistical analysis in Sec. III, we derive expressions for $SNR_1$ and $SNR_2$ and from these obtain $P_D$ and $P_{FA}$ expressions.

## III. Statistical Analysis

To simplify our statistical analysis, we assume uniform weighting across each detector, $w_n(\tau) = 1$, and equal biases $B_1 = B_2 = B$ and that constant bias terms are subtracted from the $P_4$ output. Equation (2) for the $n$th detector output now contains the following five terms:

$$I(n) = \frac{B}{T_I} \int_{-T_I/2}^{T_I/2} \int_{(n-1/2)D}^{(n+1/2)D} s(t - \tau)d\tau dt$$
$$+ \frac{B}{T_I} \int_{-T_I/2}^{T_I/2} \int_{(n-1/2)D}^{(n+1/2)D} s(t - \tau_0)d\tau dt$$
$$+ \frac{B}{T_I} \int_{-T_I/2}^{T_I/2} \int_{(n-1/2)D}^{(n+1/2)D} n(t)d\tau dt$$
$$+ \frac{1}{T_I} \int_{-T_I/2}^{T_I/2} \int_{(n-1/2)D}^{(n+1/2)D} s(t - \tau)n(t)d\tau dt$$
$$+ \frac{1}{T_I} \int_{-T_I/2}^{T_I/2} \int_{(n-1/2)D}^{(n+1/2)D} s(t - \tau)s(t - \tau_0)d\tau dt. \quad (5)$$

For the case of zero-mean independent signal $s(t)$ and noise $n(t)$, the square of the expected value involves only the last terms in Eq. (5), i.e.,

$$E^2[I(n)] = \left[\int_{(n-1/2)D}^{(n+1/2)D} R_s(\tau - \tau_0)d\tau\right]^2, \quad (6)$$

where $R_s$ is the signal autocorrelation function. The variance of $I(n)$ is found from Eq. (5) to be

$$\text{var}[I(n)] = E[I(n)]^2 - E^2[I(n)]$$

$$= \frac{B^2}{T_I^2} \iint_{-T_I/2}^{T_I/2} \iint_{(n-1/2)D}^{(n+1/2)D} R_s(t - u - \tau + \tau')d\tau d\tau'dtdu$$

$$+ \frac{B^2D^2}{T_I^2} \iint_{-T_I/2}^{T_I/2} R_s(t-u)dtdu + \frac{B^2D^2}{T_I^2} \iint_{-T_I/2}^{T_I/2} R_n(t-u)dtdu$$

$$+ \frac{1}{T_I^2} \iint_{-T_I/2}^{T_I/2} \iint_{(n-1/2)D}^{(n+1/2)D} R_s(t - u - \tau + \tau')R_n(t-u)d\tau d\tau'dtdu$$

$$+ \frac{1}{T_I^2} \iint_{-T_I/2}^{T_I/2} \iint_{(n-1/2)D}^{(n+1/2)D} R_s(t-u)R_s(t - u - \tau + \tau')d\tau d\tau'dtdu$$

$$+ \frac{1}{T_I^2} \iint_{-T_I/2}^{T_I/2} \iint_{(n-1/2)D}^{(n+1/2)D} R_s(t - u - \tau + \tau_0)R_s(t - u + \tau' - \tau_0)d\tau d\tau'dtdu$$

$$+ \frac{2B^2D}{T_I^2} \iint_{-T_I/2}^{T_I/2} \int_{(n-1/2)D}^{(n+1/2)D} R_s(t - u - \tau + \tau_0)d\tau dtdu, \tag{7}$$

where Gaussian-distributed signals were assumed (third-order moments are zero, and the fourth moment theorem[8] can be used) and where $R_n(\tau)$ is the noise autocorrelation function. Assuming that the signal and noise have similarly shaped autocorrelation functions, Eq. (7) simplifies to

$$\text{var}[I(n)] = \frac{B^2}{T_I^2} \int_{-T_I}^{T_I} \iint_{(n-1/2)D}^{(n+1/2)D} (T_I - |z|)R_s(z - \tau + \tau')d\tau d\tau'dz + \frac{B^2D^2}{T_I^2}\left(1 + \frac{1}{\text{SNR}_I}\right) \int_{-T_I}^{T_I} (T_I - |z|)R_s(z)dz$$

$$+ \frac{1}{T_I^2}\left(1 + \frac{1}{\text{SNR}_I}\right) \int_{-T_I}^{T_I} \iint_{(n-1/2)D}^{(n+1/2)D} (T_I - |z|)R_s(z)R_s(z - \tau + \tau')d\tau d\tau'dz$$

$$+ \frac{1}{T_I^2} \int_{-T_I}^{T_I} \iint_{(n-1/2)D}^{(n+1/2)D} (T_I - |z|)R_s(z - \tau + \tau_0)R_s(z + \tau' - \tau_0)d\tau d\tau'dz$$

$$+ \frac{2B^2D}{T_I^2} \int_{-T_I}^{T_I} \int_{(n-1/2)D}^{(n+1/2)D} (T_I - |z|)R_s(z - \tau + \tau_0)d\tau dz, \tag{8}$$

where the input $\text{SNR}_I$ is the ratio of the peak signal power to the peak noise power. If the assumption of similar correlation functions for the signal and noise is removed, the $\text{SNR}_I$ expression can be appropriately modified.[10] Assuming $T_I \gg 1/\beta$, where $\beta = \text{BW}_S$ is the signal bandwidth, we can omit the $|z|$ factors in Eq. (8) and obtain

$$\text{var}[I(n)] = \frac{B^2}{T_I} \int_{-T_I}^{T_I} \int_{-D}^{D} (D - |q|)R_s(z + q)dqdz$$

$$+ \frac{B^2D^2}{T_I}\left(1 + \frac{1}{\text{SNR}_I}\right) \int_{-T_I}^{T_I} R_s(z)dz$$

$$+ \frac{1}{T_I}\left(1 + \frac{1}{\text{SNR}_I}\right) \int_{-T_I}^{T_I} \int_{-D}^{D} (D - |q|)R_s(z)$$

$$\times R_s(z + q)dqdz$$

$$+ \frac{1}{T_I} \int_{-T_I}^{T_I} \iint_{(n-1/2)D}^{(n+1/2)D} R_s(z - \tau + \tau_0)$$

$$\times R_s(z + \tau' - \tau_0)d\tau d\tau'dz$$

$$+ \frac{2B^2D}{T_I} \int_{-T_I}^{T_I} \int_{(n-1/2)D}^{(n+1/2)D} R_s(z - \tau + \tau_0)d\tau dz. \tag{9}$$

With no loss of generality, we assume that the correlation peak occurs at the $n = 0$ detector element. Then

$$E^2[I(0)] = \left[ \int_{-D/2}^{D/2} R_s(\tau - \tau_0)d\tau \right]^2, \tag{10}$$

where now $-(D/2) \leq \tau_0 \leq (D/2)$. The variance at the peak var$[I(0)]$ is thus given by Eq. (9) with $n = 0$ and the variance far from the peak var$[I(n)]$ by neglecting the

fourth term in Eq. (9). $\text{SNR}_1$ and $\text{SNR}_2$ can now be obtained from the ratio of Eq. (10) to var$[I(0)]$ and Eq. (10) to Eq. (9), respectively. A numerical evaluation shows that the fourth term in Eq. (9) has a negligible 3% contribution to the total variance far from the peak. This is logical because $R_s$ is sharply peaked and because the two factors in term four diverge as $\tau$ changes. For generality, we retain all terms in Eq. (9).

## IV. Gaussian-Markov Case Study

We now use the results of our statistical SNR analysis in Sec. III to derive $P_D$ and $P_{FA}$ expressions. We consider the case of signals with a Gaussian-Markov autocorrelation function[11]:

$$R_s(z) = R_0 \exp(-\beta|z|), \tag{11}$$

where $\beta$ is the signal's 3-dB bandwidth and $R_0$ is the signal power. This signal model was chosen because it allows an analytical evaluation of both $\text{SNR}_1$ and $\text{SNR}_2$ without the need for numerical evaluation. We have also numerically evaluated our results for a Gaussian-shaped autocorrelation signal model and obtained results similar to those obtained herein, where we include only the analytical results for the Gaussian-Markov model.

Using the model in Eq. (11), the average peak power in Eq. (10) can be shown by a simple but tedious analysis to be

$$E^2[I(0)] = \frac{R_0^2}{2} \{2 - \exp[-\beta(D/2 + \tau_0)] - \exp[\beta(-D/2 + \tau_0)]\}^2.$$

$$(12)$$

The var$[I(0)]$ and var$[I(n)]$ expressions now become

$$\text{var}[I(0)] = \frac{R_0 B^2 D^2}{T_I \beta}\left(8 + \frac{2}{\text{SNR}_I}\right) + \frac{4R_0^2 D}{T_I \beta^2}\left(2 + \frac{1}{\text{SNR}_I}\right)$$

$$+ \frac{R_0^2}{T_I \beta^2} A_1 + \frac{R_0^2}{T_I \beta^2}\left(1 + \frac{1}{\text{SNR}_I}\right) A_2 \qquad (13)$$

$$\text{var}[I(n)] = \frac{R_0 B^2 D^2}{T_I \beta}\left(8 + \frac{2}{\text{SNR}_I}\right) + \frac{R_0^2}{T_I \beta^2}\left(1 + \frac{1}{\text{SNR}_I}\right) A_3, \quad (14)$$

where

$$A_1 = -8\tau_0 - \frac{6}{\beta}\exp(-2\beta\tau_0) - 4\tau_0 \exp(-2\beta\tau_0)$$

$$+ \exp(-\beta D)[\exp(2\beta\tau_0) + \exp(-2\beta\tau_0)]\left(D + \frac{3}{\beta}\right)$$

$$- \exp(-\beta D)[\exp(2\beta\tau_0) - \exp(-2\beta\tau_0)]2\tau_0, \qquad (15a)$$

$$A_2 = \frac{6}{\beta} + \frac{6}{\beta}\exp(-\beta D) + 2D \exp(-\beta D), \qquad (15b)$$

$$A_3 = 4D - \frac{6}{\beta} + \frac{6}{\beta}\exp(-\beta D) + 2D \exp(-\beta D). \qquad (15c)$$

From Eqs. (12)–(15), we find

$$\text{SNR}_1 = \frac{\{2 - \exp[-\beta(D/2 + \tau_0)] - \exp[\beta(-D/2 + \tau_0)]\}^2}{\dfrac{D^2\beta}{T_I(\text{SBR})^2}\left(8 + \dfrac{2}{\text{SNR}_I}\right) + \dfrac{4D}{T_I}\left(2 + \dfrac{1}{\text{SNR}_I}\right) + \dfrac{A_1}{T_I} + \dfrac{A_2}{T_I}\left(1 + \dfrac{1}{\text{SNR}_I}\right)},$$

$$(16)$$

$$\text{SNR}_2 = \frac{\{2 - \exp[-\beta(D/2 + \tau_0)] - \exp[\beta(-D/2 + \tau_0)]\}^2}{\dfrac{D^2\beta}{T_I(\text{SBR})^2}\left(8 + \dfrac{2}{\text{SNR}_I}\right) + \dfrac{A_3}{T_I}\left(1 + \dfrac{1}{\text{SNR}_I}\right)},$$

$$(17)$$

where $\text{SBR} = \sqrt{R_0}/B$ is the signal-to-bias ratio for the input data to the AO cell. The error-free $\text{SNR}_1$ and $\text{SNR}_2$ expressions are found (by applying l'Hopital's rule with $D = 0$ and $\tau_0 = 0$) to be

$$\text{SNR}_1 = \frac{T_I \beta}{\left(2 + \dfrac{1}{\text{SNR}_I}\right) + \left(8 + \dfrac{2}{\text{SNR}_I}\right)\dfrac{1}{(\text{SBR})^2}}, \qquad (18)$$

$$\text{SNR}_2 = \frac{T_I \beta}{\left(1 + \dfrac{1}{\text{SNR}_I}\right) + \left(8 + \dfrac{2}{\text{SNR}_I}\right)\dfrac{1}{(\text{SBR})^2}}. \qquad (19)$$

These error-free expressions are useful for measuring the loss incurred when $D \neq 0$ and $\tau_0 \neq 0$.

From Eqs. (12), (16), and (17), we can now quantify the $P_D$ and $P_{FA}$ performance to be expected as a function of the different signal and system parameters and the different detector effects. $P_D$ is obtained by substituting Eqs. (12) and (16) into Eq. (3), and $P_{FA}$ is found by substituting Eqs. (12) and (17) into Eq. (4). In calculating $P_{FA}$ we assume $E\{C(\tau)\} = 0$. This follows from our zero-mean signal and noise assumption and the fact that $R_s(\tau)$ will be sharply peaked. We also note that the $P_{FA}$ we calculate corresponds to $P_{FA}$ for one detector element. The total $P_{FA}$ for the entire output ($P_{FAT}$) of $N + 1$ detectors can be obtained from our $P_{FA}$ by
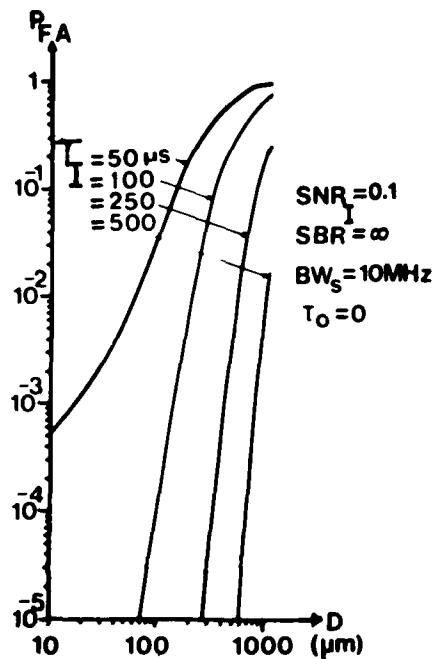


Fig. 2.   Effect of detector size $D$ and integration time $T_I$ on $P_{FA}$ (for $P_D = 0.999$).

$$P_{FAT} = 1 - (1 - P_{FA})^{N+1}. \qquad (20)$$

The three detector effects we consider are the finite detector size $D$ (Sec. V), the location of the correlation peak within a detector element (Sec. VI), and the spatial response across a detector element (Sec. VII). Each of these detector effects is treated separately, since for each case the integration time $T_I$, the input $\text{SNR}_I$, the signal bandwidth $\beta = \text{BW}_s$, the signal-to-bias ratio (SBR), and other such parameters affect the results. Our purpose in these next three sections is to quantify the effect of these various parameters on the detection performance (measured through $P_D$ and $P_{FA}$) of a linear intensity TI AO correlator and to provide guidelines for TI AO correlator design.

## V.  Area Integration Effects

In this section, the effect of the finite detector element size is quantified. Graphic presentations are used to provide quantitative performance data. The trends observed are then noted and discussed. We include only $P_{FA}$ data rather than $P_D$ data to reduce the length of our text.

In Fig. 2 we show the variation of $P_{FA}$ with $D$ for different $T_I$ values. Both $P_{FA}$ and $P_D$ improve as $T_I$ increases (as expected since longer integration time reduces noise and enhances signal). $P_{FA}$ and $P_D$ also improve as $D$ decreases. This is less immediately obvious but can be explained by realizing that increasing $D$ increases the noise more than the signal (per detector element). This occurs since the noise is relatively
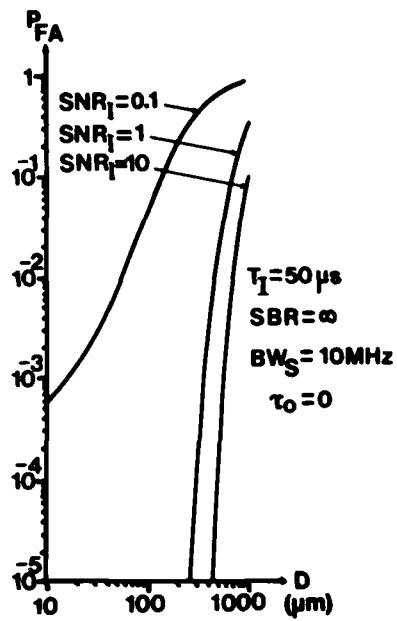
Fig. 3. Effects of input $SNR_I$ (amplitude) and detector size $D$ on $P_{FA}$ (for $P_D = 0.999$).

uniform over the correlation plane, whereas the signal correlation is of narrow and finite width. Thus, for any $D \geq 0$ (not just for $D$ greater than the width of the correlation peak), a larger $D$ degrades $P_{FA}$ and $P_D$ performance. This effect is more pronounced when $D$ is larger than the width of the correlation peak (123 $\mu$m for $BW_s = 10$ MHz and the shear $TeO_2$ AO cell assumed). The data in Fig. 2 verify this and quantify this effect.

As noted at the outset that many system and signal parameters exist and affect performance. Next we consider the effect of $D$ and input $SNR_I$ on $P_{FA}$. As expected, we find (Fig. 3) that $P_{FA}$ improves as $SNR_I$ increases (for a fixed $D$). For the case considered ($T_I = 50$ $\mu$sec, $BW_s = 10$ MHz, or TBWP = 500), we find that a smaller $D$ is needed (and oversampling of the correlation is required) when $SNR_I$ is low (below 1.0 for the case chosen). For example, if $P_{FA} = 0.001$ is desired (with $P_D = 0.999$), the detector size must satisfy $D \leq 18$ $\mu$m if $SNR_I = 0.1$. ($D = 18$ $\mu$m is much less than the 123-$\mu$m width of the correlation peak.) We note (from Figs. 2 and 3) that $T_I$ and $SNR_I$ have a much more significant effect on $P_{FA}$ than does $D$. For example, for $D = 70$ $\mu$m, doubling $T_I$ from 50 to 100 $\mu$sec (Fig. 2) results in a quite significant $P_{FA}$ improvement (from $10^{-2}$ to $10^{-5}$). Conversely, reducing $D$ by a factor of 2 to 35 $\mu$m improves $P_{FA}$ from $10^{-2}$ to only $2 \times 10^{-3}$. Thus, as a general system design guideline, if the desired $P_{FA}$ for a given $SNR_I$ cannot be achieved with a reasonable $D$, a slight increase in $T_I$ can often overcome finite detector element effects (assuming that the signal duration is sufficient). For large $SNR_I$, the size $D$ is of concern. However, low $SNR_I$ is the scenario of most concern.

As our next signal parameter, we consider the effect of $D$ and the signal bandwidth $BW_s$ on $P_{FA}$. Our results are shown on Fig. 4. Recall that the width $D_c$ of the correlation peak decreases as $BW_s$ increases, specifically $D_c = (2/BW_s)v_s$. For $BW_s = 1$ MHz, $D_c = 1230$ $\mu$m, and all $D$ values shown are much less than $D_c$, and hence the variation of $P_{FA}$ with $D$ is neglibible. As $BW_s$ increase, $P_{FA}$ improves (due to the increased TBWP). For $BW_s > 10$ MHz achieving $P_{FA} < 10^{-3}$ is easy for a wide range of detector sizes $D$. For $BW_s = 10$ MHz, $D_c = 123$ $\mu$m, and we see that any detector size $D \leq 100$ $\mu$m (or $D$ less than approximately $D_c$) yields good $P_{FA} < 10^{-3}$ performance. However, as $D$ is increased further, the degradation in $P_{FA}$ is more severe for larger $BW_s$ (since the width of the correlation peak becomes increasingly less than the width $D$ of a detector and thus more correlation noise enters the detector). For $BW_s = 40$ MHz, $D_c = 31$ $\mu$m, and for any $D \leq 120$ $\mu$m we find $P_{FA} < 10^{-3}$. Thus, as $BW_s$ increases, the maximum allowable $D$ for a given $P_{FA}$ increases. This occurs because the improvement in $P_{FA}$ (with increasing $BW_s$) is larger than the degradation in $P_{FA}$ (with increasing $D$). For lower $SNR_I$ cases, smaller $D$ values than those shown are expected to be required (as we found in Fig. 3). For $BW_s = 5$ MHz, $D_c = 246$ $\mu$m, and we find that $D < 35$ $\mu$m (one-seventh of the width of the correlation peak) is required to obtain $P_{FA} < 10^{-3}$. Thus, as $BW_s$ decreases, we require finer sampling of the correlation peak to maintain a given $P_{FA}$.

The number of detector samples required within the correlation peak and the $P_{FA}$ obtainable thus interact significantly as $BW_s$ varies. The quantitative data in Fig. 4 show this clearly (for the $SNR_I$ and $T_I$ values
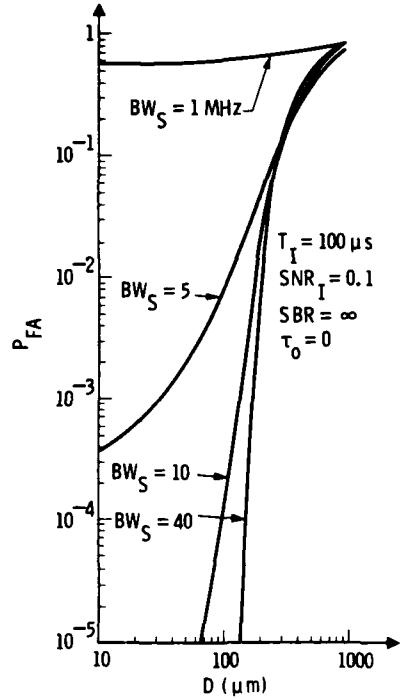


Fig. 4. Effects of bandwidth $BW_s$ and detector size $D$ on $P_{FA}$ (for $T_I = 100$ $\mu$sec, $SNR_I = 0.1$, SBR = $\infty$).
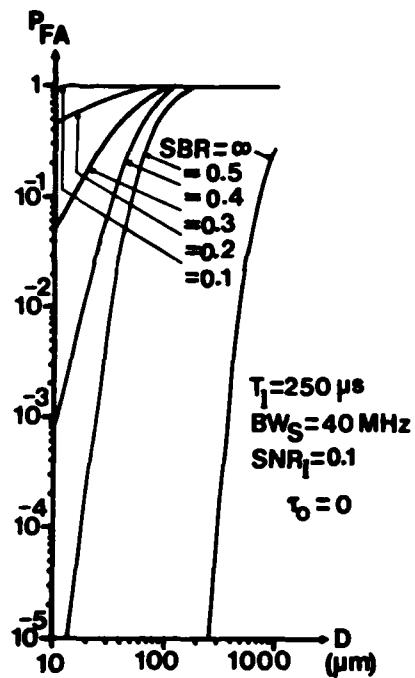
Fig. 5. Effects of signal-to-bias ratio and detector size $D$ on $P_{FA}$ for $P_D = 0.999$.

selected). This vividly demonstrates the importance of obtaining such plots for the parameters of the signal of concern. Without this, the detector sampling required for a given $P_{FA}$ would be quite difficult to assess. In general, for signals with large TBWP > 1000 and moderately low $SNR_I \geq 0.1$, the detector size can be chosen to be less than or equal to the width of the correlation peak, and excellent $P_{FA} < 10^{-3}$ will result. For signals with moderate TBWP = 500, increased over-sampling of the correlation plane is required.

Last, we consider how $D$ and the final and most dominant system parameter (the SBR of the input data to the AO cell) affect our $P_{FA}$ performance measure. Recall that SBR = $\infty$ for operation of the AO cell in the amplitude modulation mode and that the best value for the intensity modulation mode is SBR = 0.5. In Fig. 5, we show how $P_{FA}$ varies with $D$ and SBR. We immediately note that the amplitude modulation mode (SBR = $\infty$) yields much better performance for any $D$ value and allows much larger $D$ values. This must be qualified by noting that the output of a TI correlator appears on a spatial carrier[4] when the AO cells are operated in the linear amplitude mode. Thus the detector size in this case must be sufficiently small to detect the spatial carrier. (This effect is not included in our present data.) However, in a detection (compared to a delay estimation) application of a correlator, we often know where the correlation will occur (once we are in synchronization), thus considerably reducing the number of detectors required.

From Fig. 5 we see that $P_{FA}$ degrades as $D$ increases (as explained before). The decrease in $P_{FA}$ performance as SBR decreases is due to the increase in the

signal-dependent noise present in the output of a TI correlator. This bias cannot be simply subtracted form the system's output. The slope of the SBR = 0.5 curve in Fig. 5 is comparable with that of the high $BW_s$, $SNR_I$ and $T_I$ curves in our prior (SBR = $\infty$) figures. However, the associated $D$ values are an order of magnitude smaller. Thus quite small detectors and quite fine correlation plane sampling are required for intensity mode TI AO correlators operation. For example, for the signal considered, the width of the correlation peak is 31 $\mu$m, whereas the maximum detector size for $P_{FA}$ = 0.001 is 24 $\mu$m or approximately the width of the correlation peak (for SBR = 0.5). A change in $D$ by only 5–19 $\mu$m (with SBR = 0.5) changes $P_{FA}$ from $10^{-3}$ to $10^{-4}$. For the smallest realistic 10-$\mu$m detector size shown, the $P_{FA}$ values obtained are quite large (for SBR $\leq 0.4$). Thus a finite detector size significantly affects $P_{FA}$ performance for intensity mode AO operation. The low $SNR_I$, the large $T_I$ and large $BW_s$ scenario used in Fig. 5 is typical of most spread spectrum signal cases.

## VI. Effects of Correlation Peak Location

In this section, we consider our second detector effect (the location $\tau_0$ of the correlation peak within one detector element of finite area $D$). We first consider $P_{FA}$ as a function of delay $\tau_0$ (where $-D/2 \leq \tau_0 \leq D/2$) within one detector element for several signal bandwidths $BW_s$ and several detector sizes $D$. A delay $\tau_0$ = 0.0 corresponds to a correlation peak located in the center of a detector element, whereas a delay of $\pm 0.5$ corresponds to a peak located at the edge of a detector (between two detectors). In Fig. 6, we summarize our quantitative $P_{FA}$ performance as a function of $D$ and
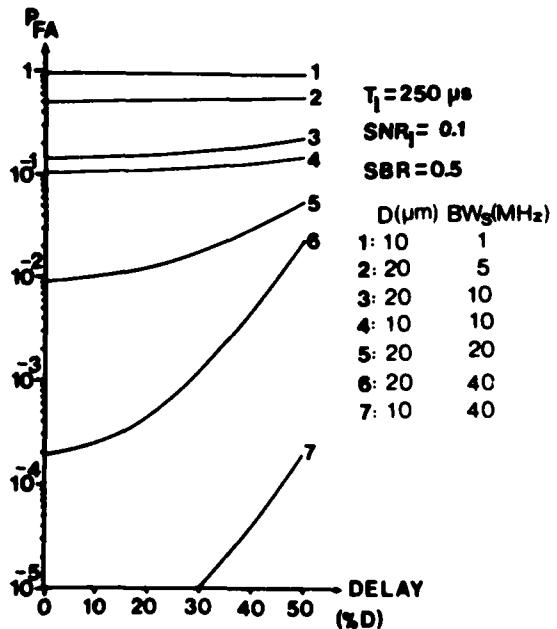


Fig. 6. Effects of the $BW_s$ $D$ product and the location (delay) of the correlation peak within one detector (as a percent of $D$) on $P_{FA}$ (for $P_D = 0.999$).

BW$_s$. Curves 1–7 correspond to systems with increasing BW$_s$ values and different $D$ choices. For the systems considered, the correlation width in seconds for a signal of bandwidths BW$_s$ is 2/BW$_s$. For a TeO$_2$ cell with 1:1 imaging a 10-$\mu$m detector size corresponds to a 16-nsec sampling time per detector. Specifically, a BW$_s$ = 10-MHz signal has a correlation width of 2/10 MHz = 200 nsec. Thus curve 3 corresponds to 200/32 $\cong$ 6 detector samples within the correlation width and curve 4 to $\cong$12 samples. Curves 1 and 2 have considerably more samples. Curve 5 has 100/32 $\cong$ 3 samples. Curve 6 has 1.5 samples, and curve 7 has 3 samples. As BW$_s$ increases (curves 1–7), $P_{FA}$ improves due to the larger signal TBWP. As the detector size becomes less (curves 3 vs 4 and 6 vs 7), the correlation plane sampling is better, and $P_{FA}$ again improves. For larger BW$_s$, the improvement due to smaller $D$ values is larger since the correlation peak is narrower (see Fig. 4). The variation in $P_{FA}$ vs $\tau_0$ (center of detector) to $\tau_0$ = 0.5$D$ (edge of detector) also follows logically. As $\tau_0$ increases for a fixed $D$, the correlation peak power within $D$ decreases (assuming $D$ is less than the width of the correlation peak). As BW$_s$ increases, the sensitivity to the location of the peak within a detector element becomes more important (since the width of the peak is less). The variation in $P_{FA}$ with $\tau_0$ is most severe for signals with a large BW$_s$ (or correspondingly large TBWP). Although $P_{FA}$ is much better for large BW$_s$, this $\tau_0$ effect is still quite significant. For $D$ = 20 $\mu$m and BW$_s$ = 40 MHz (curve 6), $P_{FA}$ varies from 0.0002 (when $\tau_0$ = 0) to 0.02 (when $\tau_0$ = $D/2$). This is a non-negligible factor of 100 loss in performance. Thus, for higher BW$_s$ signals, increased correlation plane sampling is recommended (e.g., curve 6 is curve 7).

The data in Fig. 6 were obtained for intensity modulation (SBR = 0.5). Similar trends are expected for linear amplitude modulation (i.e., SBR = $\infty$), but the actual $P_{FA}$ values will be better (since the effects of finite SBR are absent).

### VII. Effects of Detector Spatial Weighting Function

Thus far, the spatial weighting function for each detector element has assumed a rect function. However, this is not necessarily the case, especially for CCD arrays.[12] In many cases, the response profile for a detector element can be modeled as a trapezoid whose upper-to-lower base ratio $d/D$ depends on the actual array. To consider the effects of such a profile, we assume a detector weighting function:

$$w_n(\tau) = \begin{cases} \dfrac{2\tau}{D-d} - \dfrac{2(n-1/2)D}{D-d} & nD - D/2 < \tau < nD - d/2, \\[2mm] 1 & nD - d/2 < \tau < nD + d/2, \\[2mm] -\dfrac{2\tau}{D-d} + & \\[2mm] \dfrac{2(n+1/2)D}{D-d} & nD + d/2 < \tau < nD + D/2, \end{cases} \quad (21)$$

where $n$ is the detector element number, $D$ is the lower base of the trapezoid, and $d$ is the upper base. We have

chosen to use this weighting function (and the variable delay $\tau_0$) for its versatility in studying the effects of the detector element's profile on the system's performance. For example, for $d = D$, $w_n(\tau)$ describes a rectangular profile, whereas $d = 0$ describes a triangular profile. For any other $d$ and $D$ relationship, $w_n(\tau)$ is a trapezoid.

Let us assume that the correlation peak lies within the $n$ = 0 detector element, i.e., $-D/2 \le \tau_0 \le D/2$, then $w_0(\tau)$ defines the profile of the detector element in which the correlation peak occurs. The average peak power in Eq. (6) then becomes

$$E^2[I(0)] = \left[\int_{-D/2}^{D/2} w_0(\tau)R_s(\tau - \tau_0)d\tau\right]^2. \quad (22)$$

Substituting Eq. (21) into Eq. (22) yields the expression for the average correlation peak power $I_{pp}$. We evaluated this for different $d/D$ ratios and found that a rectangular detector gave the best $I_{pp}$ value. The loss in $I_{pp}$ was 25% when $\tau_0$ = $D/2$ rather than $\tau_0$ = 0. A trapezoidal detector response profile with $d = 0.6D$ gave 30% less $I_{pp}$, and a triangular detector response profile gave 70% less $I_{pp}$. To analyze $w_n(\tau)$ effects on SNR, we assumed $\tau_0$ = 0 (to simplify the analysis), since then the variance of SNR$_1$ is independent of $\tau_0$ (with $w_n$ fixed), and SNR$_1$ and SNR$_2$ vary with $\tau_0$ in the same way that $I_{pp}$ does (for low SNR$_I$). For SBR = $\infty$, we evaluated SNR$_1$ and SNR$_2$ as a function of $D$ and BW$_S$ for $\tau_0$ = 0



SNR$_I$=0.1
T$_I$=50 $\mu$s
SBR=$\infty$
BW$_s$=10MHz
$\tau_0$ = 0

1: RECT
2: 50% TRAPEZOID
3: TRIANGLE
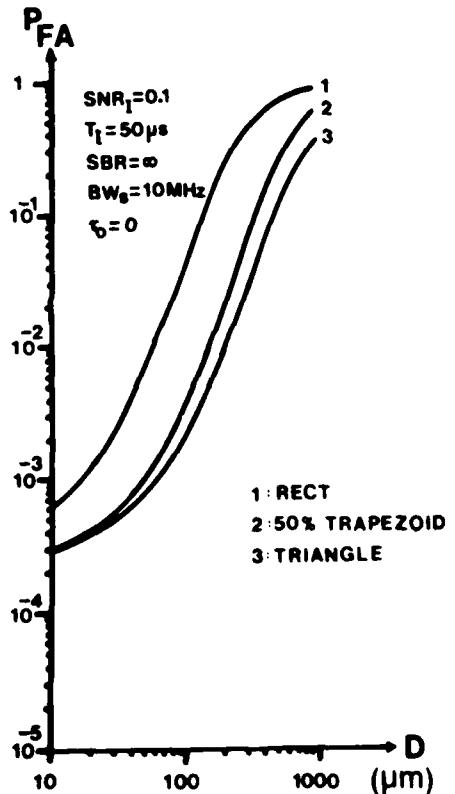
Fig. 7. Effect of detector weighting profiles on $P_{FA}$ (for $P_D$ = 0.999) as a function of the detector element size $D$.

for different $w_n(\tau)$ profiles. We found that both triangular and trapezoidal $w_n(\tau)$ profiles gave better output SNR values than did a rectangular $w_n(\tau)$. This is expected for $\tau_0 = 0$ since these $w_n$ reduce the noise more than the peak value. The SNR improvements increased as $BW_s$ or $D$ increased (since the peak narrows, and the weighting reduces the noise more than the peak value). This is less time if $SNR_I$ is larger.

In Fig. 7 we show $P_{FA}$ vs $D$ for these three detector element profiles for the case $\tau_0 = 0$. We find that the nonrectangular profiles perform best, with only a small improvement in $SNR_2$ (0.5 dB) and in $P_{FA} = 0.00037$ (triangular), $P_{FA} = 0.00044$ (trapezoidal) and $P_{FA} = 0.00125$ (rectangular)].

To compare output SNR values for $\tau_0 \neq 0$ for different $w_n(\tau)$, we consider the worst-case $\tau_0 = D/2$. We compute the $I_{pp}$ loss (due to $\tau_0 = D/2$ compared with $\tau_0 = 0$) and from this subtract the output $SNR_2$ improvement for $\tau_0 = 0$ (due to the use of a nonrectangular vs a rectangular detector element spatial response profile). From this analysis, we find that trapezoidal and triangular profiles give ~0.5 dB better $SNR_2$ than a rectangular profile.

Thus it appears that the simpler rectangular detector response profile model can be used (thus greatly simplifying the analysis) with only small efforts of the SNR or $P_{FA}$ to be expected. The $P_{FA}$ results actually obtained are expected to be slightly better than those predicted by the simplified $w_n(\tau)$ theory.

## VIII. Experimental Verification

Initial experimental results obtained on a laboratory TI AO correlator for signal detection are now reported. A matching pair of $TeO_2$ cells was used for the point modulator and delay line. The center frequency of both cells was 35 MHz. Each cell was operated in the linear intensity mode. The cells were biased at 12 V and operated with a signal level of 6 V (i.e., SBR = 0.5). No additive noise was introduced, and thus $SNR_I = \infty$. Although the cell bandwidth could accommodate 20-MHz data, we could not produce a signal of such bandwidth because of equipment limitations. Thus we used a Gaussian-distributed signal (from a noise generator) with a Gaussian autocorrelation function and a $BW_S = 0.5$ MHz. The width of the correlation peak for this signal is 4 $\mu$sec with the 1:4 imaging system used. To measure $SNR_1$ and $SNR_2$, we used a single detector element with $D = 200$ $\mu$m and an integration time $T_1 = 5$ msec. $D = 200$ $\mu$m corresponds to 1.28 $\mu$sec or about one-third of the width of the correlation peak. No additional system errors were introduced by the bandwidth of the AO cells and the phase response of the transducers over this small $BW_S$. This experimental setup thus allowed detector size effects alone to be studied (with all other error sources reduced to negligible levels). To study the effect of $D$ on system performance, we inserted a variable detector aperture of size $D$ in front of the detector element and varied the aperture (and hence $D$) in one dimension from 50 to 200 $\mu$m. The height of the slit was kept constant (at 100 $\mu$m) as its width was varied.

To measure $SNR_1$, we fixed $D$ and centered the correlation peak at the center of the detector element ($\tau_0 = 0$). Since $D = 50$–$200$ $\mu$m corresponds to 0.32–1.28 $\mu$sec, which is less than the 4-$\mu$sec width of the correlation peak, negligible errors are introduced by slight mispositioning of the detector. Two hundred separate measurements of the detector's output were taken. (The noise or statistical fluctuations were different in each measurement, and thus these data constituted a different sample realization of the random correlation process.) For each choice of $D$, the mean and variance of these 200 samples were calculated and their ratio calculated to provide our desired $SNR_1$.

To obtain $SNR_2$, we measured the value of the correlation at the peak and far from the peak. To achieve this, we moved the detector element far (~15 $\mu$sec) from the peak location. Image plane detector difference errors were negligible, since the same detector was used for measurements both at the peak and far from the peak. To reduce the effects of input light uniformity, AO cell attenuation, and AO cell spatial response variations, we measured the correlator's output with no signal present (i.e., with only the carrier present) and selected two output locations for our $SNR_2$ measurements where the light level was equal within 5%. To facilitate a uniform output (i.e., negligible spatial weighting due to the cell, acoustic attenuation), we used a spatial filter in the frequency plane that reduced AO cell nonuniform response variations. From our 200 measurements of the correlation output far from the peak, we obtained estimates of the correlation noise level and hence $SNR_2$ experimental data. We repeated this procedure for different $D$ values of 50, 100, 150, and 200 $\mu$m, corresponding to samplings of 50 to 12 samples/correlation peak width. For each case we obtained 200 measurement samples. To ensure that the slit was centered in the middle of the detector element, we used a scanning microscope. This also insured us of the exact slit or detector width $D$ used.

Our experimental data and the theoretical results obtained from our theory for Gaussian autocorrelation function signals are shown in Fig. 8. The theoretical $SNR_1$ and $SNR_2$ values were obtained from Eqs. (9) and (10) using Gaussian $R_s(\tau)$ and $R_n(\tau)$ functions. Theoretically, we expect a $SNR_1$ of 24.5 dB (for $D = 10$ $\mu$m) and a $SNR_2$ 22 dB (for $D = 10$ $\mu$m). We expect a constant 2.5-dB difference in these SNR measures with the slight decrease shown for SNR as $D$ is increased. Our experimental data (Fig. 8) are in rather good agreement with theory. Both SNR values are within 2 dB of the theoretical values. $SNR_2$ is larger than $SNR_1$ as predicted by theory, with the difference (2.2 dB) being very close to theory (2.5 dB). Our experimental results show that both SNR data remain approximately constant for $D$ values between 50 and 200 $\mu$m (as predicted by our theoretical analysis). Since both of our experimental SNR values are less than the theoretical ones by 2 dB, further credence is given to our data. Possible reasons for the 2-dB SNR difference (loss) are detector noise and background optical noise.
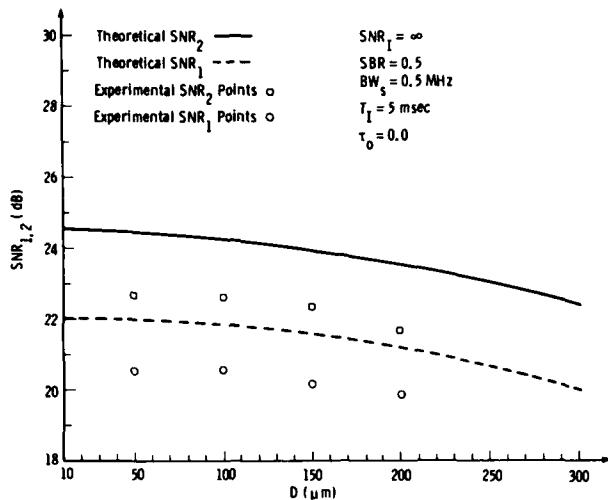
Fig. 8. Theoretical and experimental data on the effect of the detector element size $D$ on $SNR_1$ and $SNR_2$.

Valid $P_D$ and $P_{FA}$ measurements would require more than 1000 samples (to observe $P_{FA} \cong 0.001$ or 1 peak that exceeds the threshold in 100 measurements). Since our results verified the validity of our analysis of the effect of $D$ on $SNR_1$ and $SNR_2$, the conventional relationships between $P_D$ and $P_{FA}$ and SNR should be valid. An advanced experimental verification would require use of a higher $BW_S$ signal, a larger number of samples, more control over the $D$ setting (i.e., smaller and more accurate slit widths), a more sensitive detector, etc. We note that under such conditions one could verify the effects of correlation location and spatial weighting function. In our experiments, verification of the correlation peak location with the available equipment was not possible because of the broad correlation peak obtained with the available signal $BW_S$ equipment.

## IX. Summary and Conclusions

In this paper, we have studied the effects of detector errors on the performance of an acoustooptic time-integrating correlator used for signal detection. In our analysis, we modeled the system's output to include the effects of various detector parameters such as area integration, elemental spatial weighting, and correlation peak location within a detector element. As performance measures we used $P_D$ and $P_{FA}$ and in our theory derived expressions for $P_D$ and $P_{FA}$ in terms of detector parameters and the easily measured $SNR_1$ and $SNR_2$ output correlation parameters.

To study the various detector effects, we performed a general statistical analysis and quantified our results for the case of signals and noise with a Gaussian-Markov autocorrelation function. This provided us with analytical results which fully describe the system's performance as a function of various system, signal, and detector parameters. From these expressions, we

quantified the system performance as a function of the various detector error sources and system and signal parameters in our model.

We found that area integration resulted in a variety of effects such as a degradation in both $P_D$ and $P_{FA}$ as the detector element size $D$ increased. We found that $T_I$ and $SNR_I$ effects were more dominant and that an increase in either (if possible) was more significant than a decrease in the detector area $D$. Thus system design considerations dictate an increase in $T_I$ or $SNR_I$ (if possible) to compensate for losses due to the finite detector element size $D$. Our studies of the signal bandwidth $BW_s$ effects showed that both $P_D$ and $P_{FA}$ are more sensitive to $D$ (when $BW_s$ is large) but that the $P_D$ and $P_{FA}$ values obtained in this case were quite good. Thus such issues appear to be of more concern when the signal time bandwidth product is moderate (i.e., 500–1000). The effect of the SBR and $D$ was quantified and found to be the most important and dominant effect on the choice of $D$ in a system design. It was shown that even for the maximum possible SBR value of 0.5 (for linear intensity modulated AO cells), the system's performance degrades significantly as $D$ increases.

The effect of the location of the correlation peak within a detector element was also studied. From this analysis, we found that the loss encountered as the correlation peak location departed from the center of the detector element depended on both the $D$ and $BW_S$ values. These effects were quantified. In general, the system's performance degraded as either the delay or $D$ increased, with the loss becoming more significant as $BW_S$ increased. The system's designer must select $D$ from the $P_D$ and $P_{FA}$ values obtained when the correlation peak is located at the edge of the detector element. As shown this requires detector element sizes much less than the Nyquist value (for the correlation peak width) to achieve good $P_{FA}$ performance.

To study the effects of the detector element's spatial response, we conducted a simplified but well-approximated statistical analysis. In this analysis, we used a spatial response model that varied to include triangular, trapezoidal, and rectangular detector response functions. Our analysis was conducted under the assumption of a low input SNR. For this case, we found that a triangular profile enhanced performance (since it suppressed the out-of-plane noise more than the signal). In practical cases, the detector element's spatial response is trapezoidal, and the improvement (over a rectangular response) was found to be small (of the order of 0.51 dB in $SNR_1$ or $SNR_2$). Thus future statistical analyses do not seem to require elaborate approximations of the detector's spatial response by a trapezoidal function.

Our experimental work verified several of our theoretical results (specifically the validity of our theoretically predicted difference between $SNR_1$ and $SNR_2$). The observed dependence of $SNR_1$ and $SNR_2$ on $D$ appears to be in very good agreement with our theory. In all cases, our experimental results were in agreement (within 10%) with our theoretically predicted performance.

## References

1. Special Issue on Acousto-Optic Signal Processing, Proc. IEEE 69, No. 1 (1981).
2. Special Issue on Acousto-Optic Bulk Wave Devices, Proc. Soc. Photo-Opt. Instrum. Eng. 214 (1979).
3. R. A. Sprague and C. L. Koliopoulos, "Time Integrating Acoustooptic Correlator," Appl. Opt. 15, 89 (1976).
4. P. Kellman, "Time-Integrating Optical Signal Processors," Ph.D. Thesis, Stanford U. (June 1979).
5. D. Psaltis and B. V. K. Vijaya Kumar, "Acoustooptic Spectral Estimation: A Statistical Analysis," Appl. Opt. 20, 601 (1981).
6. D. Casasent, A. Goutzoulis, and B. V. K. Vijaya Kumar, "Time-Integrating Acoustooptic Correlator: Error Source Modeling," Appl. Opt. 23, 3130 (1984).
7. B. V. K. Vijaya Kumar, D. Casasent, and A. Goutzoulis, "Fine Delay Estimation with Time Integrating Correlators," Appl. Opt. 21, 3855 (1982).
8. H. L. Van Trees, *Detection, Estimation and Modulation Theory: Part 1* (Wiley, New York, 1965).
9. H. Mostafavi and F. Smith, "Image Correlation with Geometric Distortion," IEEE Trans. Aerosp. Electron. Syst. AES-14, 487 (1978).
10. B. V. K. Vijaya Kumar and D. Casasent, "Space-Blur Bandwidth Product in Correlator Performance Evaluation," J. Opt. Soc. Am. 70, 103 (1980).
11. A. Papoulis, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, New York, 1965).
12. R. K. Hopwood, "Design Considerations for a Solid-State Image Sensing System," Proc. Soc. Photo-Opt. Instrum. Eng. 230, 72 (1980).

# 4. AN OVERVIEW AND SUMMARY OF OPTICAL PATTERN RECOGNITION RESEARCH USING FEATURE EXTRACTORS AND CORRELATORS

Ref. 3

# HYBRID OPTICAL/DIGITAL IMAGE PATTERN RECOGNITION: A REVIEW

David Casasent

Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

## ABSTRACT

The parallel processing, high-speed, compact system fabrication possibility, low power dissipation and size, plus weight advantages of optical processors have achieved great strides in recent years. The architectures, algorithms and system fabrication of hybrid pattern recognition processors are reviewed with attention and emphasis to recent results and to techniques appropriate for distortion-invariant multi-class pattern recognition applications.

## 1. INTRODUCTION

The parallel processing advantages of optical pattern recognition (OPR) systems have long been recognized. However, only recently have components, architectures, algorithms and a commitment to fabrication of such systems emerged. As a result, this topic has seen an explosion of conferences and research in recent years. Several recent reviews by the author exist [1-3] and will be summarized in this present paper with attention and emphasis on more recent work than those noted in earlier reviews. Advances in laser diode and detector technology and the commitment of several companies (General Dynamics-Pomona, ERIM, Grumman) and funding agencies, have now made fabrication of such processors and the reduction of research to systems a reality. Spatial light modulator (SLM) technology is summarized in [5] and is not discussed herein. These real-time devices still represent the major obstacles to the widespread low-cost commercial exploitation of OPR systems. However, the future for this aspect of OPR is quite promising. Recent Soviet work in this area has been most significant [81]. These and several U. S. programs have concentrated on practical SLM device technology. Many linear algebra operations are required in OPR [4] and are discussed elsewhere. Thus, the present text assumes a familiarity by the reader with feature extraction and such operations. The availability of two computer generated hologram (CGH) recorders has been a significant adjunct to research and to the fabrication of OPR systems [6]. A general purpose approach to optical computing (presently directed at signal processing rather than image processing) is the use of optical linear algebra processors. These approaches and systems are also summarized elsewhere [7] and are not discussed in this present paper. The various SPIE [8] and IEEE [9] special issues on digital image processing attest to the significant importance of this topic and the growing number of OPR papers in these references signifies the importance of this topic.

In this present review, I restrict attention to OPR algorithms and architectures for pattern recognition rather than image processing (i.e. image enhancement, restoration, etc.). To those authors whose work is not referenced herein, I apologize and plead a lack of time and space bandwidth product. Emphasis will be given to work at the Center for Excellence in Optical Data Processing at CMU, because of my familiarity with it and because of the large scope of its research in the area of optical pattern recognition. To best unify the large volume of research work in OPR, I first review the basic operations achievable in optical systems, two classic OPR architectures, and conventional feature-based pattern recognition (Section 2). Various optical architectures for feature extraction are then reviewed and discussed and results obtained on these system concepts and their present status are then advanced (Section 3). Various new correlator approaches to distortion-invariant OPR are then briefly reviewed together with optical AI/IU research and sub-pixel target identification research (Section 4). SDF techniques to achieve various distortion-invariant 3-D object recognition are then detailed with attention to new results and efficient phase-only and CGH techniques to synthesize such filters (Section 5). Section 6 is devoted to system fabrication issues with attention to new results and to flight-tests on compact architectures and systems for OPR. Our summary and conclusions then follow (Section 7).

## 2. FEATURE-SPACE OPTICAL PATTERN RECOGNITION (OPR)

### 2.1 OPERATIONS ACHIEVABLE

In optical processors, 2-D data (images) are represented by the transmittance of a 2-D data plane. By imaging one such data plane through another, we achieve the point-by-point multiplication of the two 2-D data arrays. A lens can integrate this 2-D product distribution (or any 2-D data distribution) and thus achieve a 2-D data summation (1-D data summations are also possible using cylindrical rather than spherical lenses). With CGHs, random

interconnections between 2-D data arrays, coordinate transformations and other space-variant operations are possible [6]. Thus, we can characterize and summarize the major operations possible on optical systems as 2-D parallel data multiplication and addition. A specific operation that has been the hallmark of coherent OPR is the 2-D Fourier transform (FT). This operation is readily achieved with a simple lens or mirror. In Figure 1, the 2-D light amplitude distribution incident on $P_2$ is the 2-D FT $G(u,v)$ of the input object $g(x,y)$ placed at $F_1$

$$G(u,v) = \iint g(x,y) e^{-j2\pi(ux+vy)} dx dy, \qquad (1)$$

where the spatial frequencies $(u,v)$ of the input object are related to distances $(x_2, y_2)$ in $P_2$ by

$$(u,v) = (x_2/\lambda f_L, y_2/\lambda f_L), \qquad (2)$$

where $\lambda$ is the wavelength of the input light and $f_L$ is the focal length of $L_1$ in Figure 1. If we place a filter function $H^*(u,v)$, i.e. a matched spatial filter (MSF), at $P_2$, then the light distribution leaving $P_2$ is the 2-D data product distribution $G(u,v)H^*(u,v)$ and the $P_3$ output is its FT or

$$u(x_3, y_3) = \mathscr{F}[G(u,v)H^*(u,v)] = g \circledast h. \qquad (3)$$

We represent FT distributions by upper-case letters and corresponding space functions by the corresponding lower-case letters. The symbol $\mathscr{F}$ denotes the FT operator, the superscript * denotes the complex conjugate and $\circledast$ denotes the correlation. The system of Figure 1 is a frequency plane correlator. The optical correlation of two 2-D images can also be realized in a joint transform correlator by forming the FT of the magnitude squared of the FT of the two functions. To synthesize the $H^*$ complex conjugate transmittance function required in (3), holographic techniques are used.
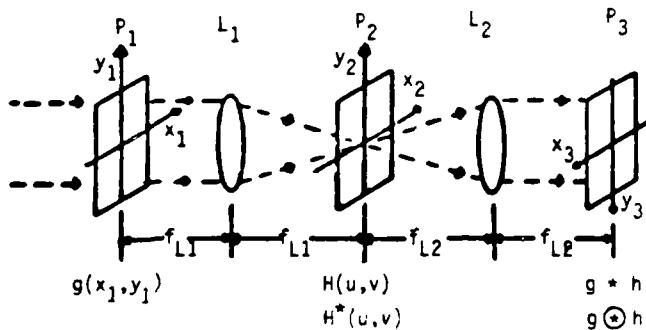


FIGURE 1
Conventional optical Fourier transform and frequency plane correlator

## 2.2  CONVENTIONAL FEATURE-SPACE PATTERN RECOGNITION

The conventional digital and mathematical literature usually considers feature-space pattern recognition. In this method (Figure 2), a set of M image features are calculated and an N x N pixel image is represented as an M-dimensional feature vector $\underline{x}$. The original feature space is often transformed to a new decision space as $\underline{y} = \underline{A}\,\underline{x}$ with independent features and dimensionality reduction. The axes of this space are a set of basis functions $\{c\}$ and the elements of each vector $\underline{y}$ are the projections on the corresponding $\underline{c}$ vectors that define this space. A discriminant vector $\underline{w}_i$ for each class i is chosen such that $\underline{w}_i^T\underline{y} > \underline{w}_j^T\underline{y}$ for all $j \neq i$. From the projection values, the class of the input object is determined. The blocks in Figure 2 are chosen to best define subproblems and thus we need not transform $\underline{x}$ into $\underline{y}$ as the first step, and then project $\underline{y}$ onto $\underline{w}$. Rather, we can project $\underline{x}$ onto a new transformed linear discriminant function vector $\underline{d}_i = \underline{A}\,\underline{w}_i$ for the class i data.

INPUT
OBJECT $\xrightarrow{f(x,y)}$ [ FEATURE GENERATOR ] $\xrightarrow{\underline{x}}$ [ DIMENSIONALITY REDUCTION FEATURE EXTRACTION TRANSFORMATION ] $\xrightarrow{\underline{y}}$ [ CLASSIFIER ] ▷class ◁▷orientation ▷confidence

COMBINED IN PRACTICE
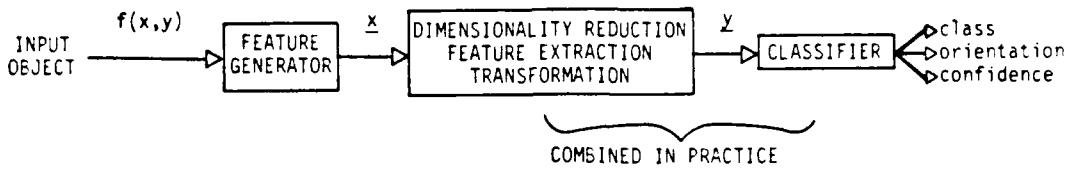
FIGURE 2
Simplified block diagram of a feature-based pattern recognition system

Most intra-class dimensionality reduction techniques are variations of the Karhunen-Loeve (KL) expansion [33] in which the elements of $A_i$ in the transformation matrix noted earlier are the eigenvectors of the correlation matrix for all class i training set images. The use of two two-stage KL transforms, how the means of each data class are handled, the number of dominant eigenvectors used, and how the correlation matrix is calculated are among the different versions of the KL algorithm [34] that can be applied. We will employ K-L, Gram-Schmidt (GS) [35], Fukunaga-Koontz (FK) [36] and Foley-Sammon (FS) [35] techniques in our OPR research. KL methods yield maximum compression of data and an orthogonal basis function set. GS methods are another technique to produce orthonormal basis functions, whereas FK and FS techniques are appropriate for inter-class recognition problems. Various classifiers used include nearest neighbor, nearest mean and the use of a least-square linear discriminant function (LDF) $w_i$. These are among the more popular ones. Our major concern in this present paper is OPR. Such digital PR post-processing algorithms are reviewed elsewhere in this volume [37] and in other OPR references by the author. The important points to emphasize are:
 -the concepts of a feature-space, basis functions, feature vectors, transformations and dimensionality reduction;
 -a training set is used to select $\underline{A}$ and $w_i$ and this operation is off-line;
 -the only required on-line operations are the calculation of the features, a vector inner product and the associated classifier decisions.
The high computational load associated with the feature generation and calculation are the major ones of concern. Thus, we concentrate on the use of the parallelism of optical processors to achieve these functions and relegate the remaining operations in Figure 2 to a general-purpose or dedicated digital hardware post-processor. The resultant hybrid optical/digital system appears to perform properly in each instance. It also appears to be the optimal combination of the parallelism of optics and the flexibility and decision making advantages of digital processors. Optical systems using CGHs can also perform the required transformations and projections directly on the 2-D input image data. The coded-phase processor [38] is one method to achieve this. Examples of its use to realize FK [39], least-squares [40,41] and the hoteling trace [42] operations have also been recently reported. The basic concepts in these optical systems is to determine the linear combination filter desired for each class (this is a linear combination of the training set images). This LDF (linear combination filter) is then encoded on a mask. A separate encoding is required for each input object class. The projection of the input test image onto each discriminant function is then optically produced and the result is summed. The phase of the input data is removed to allow different projections to appear on physically different detectors in the output plane of such a system. The detector with the largest output then denotes the class of the input object. Such a system (as presently described in the literature) is not shift-invariant and is thus best described as a feature-space method. If shifted versions of each input object are included, the space bandwidth product requirements of the associated CGH increase linearly. Such methods are appropriate for achieving shift-invariance of such a system, however such details have yet to be published. The use of optics in this case is thus most attractive when there are a large number of classes to be searched. However, in general, the vector inner product operations required are not computationally intensive unless the number of features used is also quite large.

### 3. FEATURE-SPACE OPTICAL PATTERN RECOGNITION (OPR)

In this section, we briefly discuss nine different optical feature extraction or generation systems and their performance and status.

## 3.1 FOURIER COEFFICIENT FEATURE SPACE

Since the FT operation is automatically performed optically (Figure 1), this is an obvious feature space. It is also attractive because it easily allows for dimensionality reduction. The most attractive optical dimensionality reduction method is to detect and sample the optical FT pattern (plane $P_2$ of Figure 3a) with a detector with wedge and ring shaped detector elements (Figure 3b). This concept was first advanced in [11] and used for screening of aerial images [11], for various production quality inspection tasks [11], and with an image and wedge ring detector (WRD) detection planes for aerial image classification. The commercial version of this device used 32 wedge and 32 ring-shaped detector elements. This achieves dimensionality reduction from $N^2$ to 64 features. Since the intensity of the FT is detected, the system is translation invariant. For real images, the FT is symmetric and no information loss results from the separate use of the two halves of the FT plane. The wedge outputs $F(\theta)$ are scale-invariant and their distribution shifts as the input object rotates. Conversely, the ring outputs $F(r)$ are rotationally-invariant and the distribution shifts as the input object is scaled.



(a)                                                            (b)
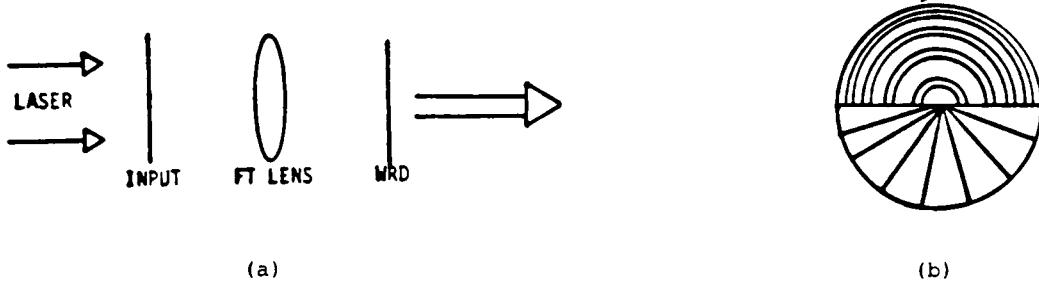
FIGURE 3
Optical Fourier coefficient feature space processor (a) and wedge ring
detector optical pattern recognition concept (b)

The most recent pattern recognition work on this feature space has involved realization of this detector using CGHs [6]. This allows more flexibility and lower cost and size. The optical realization of this unique detector plane sampling appears essential because of the large time required to digitally perform the necessary interpolations. Recent pattern recognition tests on this WRD Fourier coefficient feature space were performed for the purpose of distinguishing letters and different vehicles [12]. Results and details are available elsewhere [12]. The highlights of this work were attention to the use of amplitude versus phase Fourier coefficient features, the effect of noise, and investigation of three different feature extractor algorithms, and demonstration of scale and rotation-invariant object classification and recognition using such a feature space. However, only limited scaled and rotated versions of the input objects were tested.

## 3.2 WIGNER DISTRIBUTION FEATURE SPACE

The Wigner distribution (WD) function

$$W_{fg}(t,w) = \int f(t + \tau/2)g^*(t - \tau/2)e^{-jw\tau}d\tau \qquad (4)$$

is a simultaneous time and frequency display of the signal data. For images, the WD is a 4-D display. Auto and cross WD functions can be defined similar to (4). The WD describes local variations in the frequency control, whereas the FT provides global signal frequency information. Since images are non-stationary, a WD feature space should be most useful. One can optically produce a WD by many different techniques. An attractive method (4) uses the FT of the product of the data in two AO cells at ±45°. A binary mask (using a magneto optic SLM [43]) allows a desired sum of different WD features to be achieved on-line on a single detector, for which subsequent pattern recognition analysis is then greatly simplified. This is essential since the WD of a 1-D function is a 2-D pattern. The most recent review of this work [44] includes an SNR comparison of the optimality of WD features and initial simulation results. For pattern recognition, the auto WD of an input and reference are multiplied and integrated over time and space. In this case, the mask in Figure 4 would contain the WD of the reference(s). This appears to be an attractive approach for many pattern recognition applications. The use of an optical processor and dimensionality reduction technique as in Figure 4 is essential because of the increased dimensionality of the output in a WD feature display. Researchers in Germany [45], Wisconsin [46] and CMU [44] are the most active ones in this pattern recognition research area.
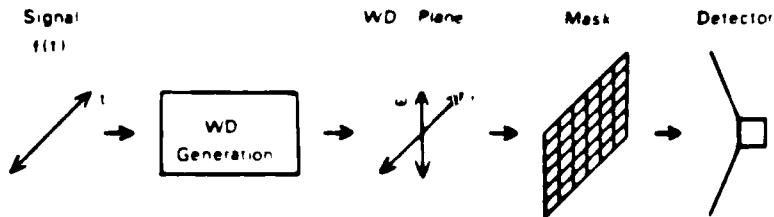
FIGURE 4
Simplified Wigner distribution feature space and detection
pattern recognition processor

### 3.3 CHORD DISTRIBUTION FEATURE SPACE

The chord distribution is defined for a binary boundary object only as the distribution $h(r,\theta)$ of the lengths $r$ and angles $\theta$ for all chords that can be drawn between boundary points on the object. Denoting boundary image points by $b(x,y) = 1$, a chord defined by the polar coordinates $r$ and $\theta$ exist between any two points if

$$g(x,y,r,\theta) = b(x,y)b(x+r\cos\theta, y+r\sin\theta) = 1. \tag{5}$$

The distribution of all chords in the image is the integral of (5) or [21]

$$h'(r,\theta) = \int\int g(x,y,r,\theta)dxdy = b(x,y) \circledS b(x,y) = h(\ell_x, \ell_y), \tag{6}$$

where $(\ell_x, \ell_y) = (r\cos\theta, r\sin\theta)$. From the last expression in (6), we see that the chord distribution can be obtained from the autocorrelation of the boundary of the object. The autocorrelation $h(\ell_x, \ell_y)$ thus contains information from which the conventional chord distribution $h(r,\theta)$ can be obtained, however complicated trigonometric and square-root calculations are required for this transformation. This feature space is still quite useful and attractive [23,24] except for the large computational load required to compute these features. In [20], we first noted that by sampling the autocorrelation of the object with a wedge ring detector, the chord distributions

$$h(r) = \int h(\ell_x, \ell_y)rd\theta, \quad h(\theta) = \int h(\ell_x, \ell_y)rdr \tag{7}$$

could be obtained directly. As before (Section 3.1), the advantages of an optical WRD are again clearly needed to achieve this. Nichols [18] later also noted this and suggested its calculation from a digital or optical FT of the optical power spectrum of the image. The computational load in the interpolation required in (7) can rapidly become excessive however. Thus, WRD sampling techniques [20] and other methods of optically producing the autocorrelation [21] appear preferable. The general block diagram of the hybrid optical/digital chord feature space processor we consider is shown in Figure 5.



FIGURE 5
Block diagram of a hybrid optical/digital chord feature space processor

Several authors [24] have expressed concern over such a feature space and its use for the recognition of complex objects. However, our post-processing algorithm and testing [20-21] have confirmed the usefulness of such a feature space. In [21-22], we extended the technique in (6)-(7) to include a silhouette image of an object with internal gray levels. These generalized chord distributions that result from such a feature space are much more useful object descriptors than the original binary edge chord functions. They also promise better noise performance [22]. In [19], Nichols considered the case when the dynamic range

of the data and more specifically its FT causes a type of edge enhancement to occur in the data for which these features are extracted. In [21], we address the use of this chord feature space for the classification of ships in the presence of out-of-plane rotational distortions. For this, a training set of 12-18 ships per class was used. The 18 best h(θ) and h(r) features were selected using KL and divergence measures and a Fisher LDF was computed from the resultant training set data. Extensive tests [21] showed perfect 100% recognition performance to be possible on separate test set ship imagery. In [22], we further extended this technique to include in-plane scale and rotational distortion invariance and methods to extract these in-plane distortion parameters from the resultant feature space data. Initial demonstrations obtained with this technique were most attractive. By properly weighting the chords of different lengths, global (large r) or local (small r) object features can be emphasized or a weighted combination of both can be used for object identification.

## 3.4 MOMENT FEATURE SPACE

The geometrical moments

$$m_{pq} = \int\int f(x,y)x^p y^q dxdy \qquad (8)$$

are a well-known and attractive feature space. However, the computational load in computing such features is such that present systems are restricted to the calculation of moments for binary objects or for the computation of only a few moments. Various techniques to optically compute the moments of an input object exist. These include the use of computer generated masks [13], a holographic mask [14], acousto-optic (AO) cells [15] and moment calculations from 1-D projections [28]. In the system of Figure 6, the image f(x,y) is imaged through masks g(x,y) at P2 on which the monomials $x^p y^q$ are recorded on different spatial frequency carriers. The products $f(x,y)x^p y^q$ are formed in parallel by optical multiplication, the integration is achieved by the output lens and each moment in (8) is formed at a different detector in P3 (with the location of the detector determined by spatial frequency carriers on the mask). In this way, all 21 moments up to fifth order can be produced optically in parallel. The detector outputs are then fed to a digital post-processor which determines the class of the input object, its orientation and the confidence of these estimates.
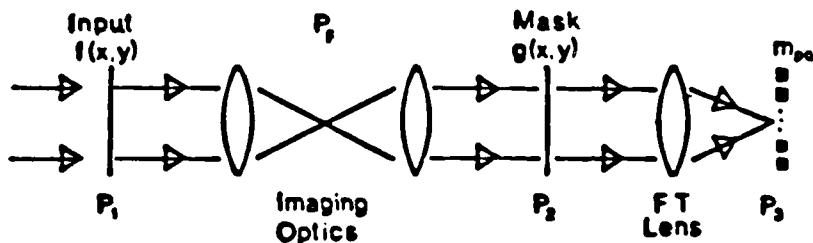


FIGURE 6
Optical system to generate all moments in parallel

This architecture and a moment feature space are attractive because of the ease with which the computed moments can be corrected for different optical system and SLM error sources [13]. A compact version of this system is under design together with alternate ways to optically produce moments. This technique has been successfully demonstrated in the classification of real ship images using very modest digital preprocessing operations [47] and in successfully and accurately estimating the in-plane distortion parameters of ship imagery [48]. The full hybrid pattern recognition system using this feature space is shown in block diagram form in Figure 7. It consists of a first-level estimator that provides class and aspect estimates. The hierarchical tree search used [17] is unique because the classes separated at each node are selected automatically using a multi-class Fisher projection method and because the discriminant vector used at each node is selected automatically from a separate two-class Fisher selection technique. As always, these off-line operations are performed on training set data and the only on-line operations required are the vector inner products (one per node in the tree). Aspect estimates are obtained from the ratio $\hat{u}_{20}/\hat{u}_{02}$ ratio obtained from the computed central moments of the input test object. An iterative nonlinear algorithm is then applied to these classes and aspects are passed from the first-level estimator. The final classification and orientation of the object is then obtained in the second-level classifier. The algorithm used in the second-level classifier is the minimum error Bayesian classification algorithm. This is possible because the

moments are jointly Gaussian random variables. Extensive tests have been conducted on a set
of 180 images of ships [17] in five different classes with 36 different aspect views per
ship and for a data base of 324 pipe parts [16] in five different groups with 36 different
aspect views for each of 9 different pipe objects. These tests showed excellent performance
(86% correct recognition of all ship images, 98% correct recognition of ship views within
50° of broadside, and 97% correct pipe classification). These tests used only 4-9 different
aspect views per class for training, required only 4-6 iterations in the Bayesian classifier
nonlinear algorithm, and showed that the use of the identity matrix as a valid approximation
to the covariance matrix was adequate. These issues greatly reduce the computational load
required on the digital post-processor. Details of this system and these results are avail-
able elsewhere [16,17]. These tests on full 3-D distorted imagery using only a limited
training set, and a large test set have also been applied and verified on real imagery.
This makes such a feature-space pattern recognition technique appear most attractive and
demonstrates a clear role for optical processors in feature extraction based pattern recog-
nition algorithms.
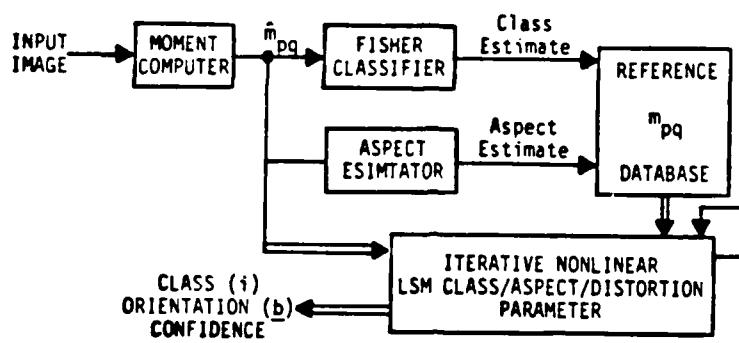


FIGURE 7
Hierarchical hybrid optical/digital moment feature-space
pattern recognition architecture

### 3.5 HOUGH TRANSFORM FEATURE SPACE

The Hough transform (HT) has recently received considerable attention and interest in
digital image processing because of its robustness and the ability to implement it on
pyramid digital architectures using simple histogram and accumulation operations only. Both
coherent [27] and non-coherent [25,26,28] optical architectures to compute the HT have been
detailed and demonstrated. The HT maps each line in an image into a point in a $(\rho,\theta)$
feature space, where $\rho$ is the perpendicular distance from the origin to the line and $\theta$ is
the angle the line makes with the x axis [31]. This technique has been generalized [30],
extended to curve detection [32] and its similarity to an MSF noted [32]. In the optical
realization of this transformation, the equivalence of a radon transform (RT) and HT is used
[29]. The RT is defined as

$$\hat{f}(\rho,\theta) = \int\int_{-\infty}^{\infty} f(x,y)\,\delta(\rho - x\cos\theta - y\sin\theta)\,dxdy. \tag{9}$$

This is equivalent to the projection of $f$ onto a line $\rho$ normal to the angle $\theta$. We denote
$\hat{f}$ at one $\theta$ by $\hat{f}_\theta$ and the full 2-D HT by $\hat{f}$. To provide insight, we note that a point $(x_0,y_0)$
in $f(x,y)$ is a sinusoid in $\hat{f}(\rho,\theta)$ space described by

$$\rho = x_0 \cos\theta + y_0 \sin\theta. \tag{10}$$

The RT is equivalent to the HT with the sinusoid weighted by the value (intensity) of the
$(x_0,y_0)$ pixel point in $f(x,y)$. In this feature space and transformation, a line in $(x,y)$ is
a point in $(\rho,\theta)$, a curve is a set of points in $(\rho,\theta)$, etc. Thus, an object composed of
lines is described by a distribution of points in the $(\rho,\theta)$ Hough space.

A noncoherent architecture [28] to realize the RT or HT is shown in Figure 8. In this
simple system, the 1-D integration (projection) of $f(x,y)$ is performed by a cylindrical lens and
$\hat{f}_\theta$ is produced. The angle $\theta$ is varied different $\hat{f}_\theta$ projections at different $\theta$ produced and the
$\hat{f}(\rho,\theta)$ distribution produced by placing a rotating Dove prism behind the input object.

With a modest 500 rpm rotation rate, one $\hat{f}_\theta(\rho)$ slice of $\hat{f}$ is produced every 60 μsec and a full $\hat{f}$ pattern at TV frame rates. To employ an HT feature space for pattern recognition, the HT of the input test object is compared to the HT of the different reference objects by whichever feature extractor and classification technique (Section 2.2) one desires.
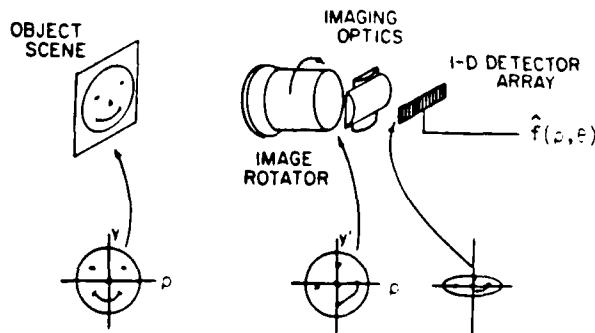


FIGURE 8
Optical system to compute the radon transform or
Hough transform by 1-D image projections [28]

## 3.6  OTHER FEATURES FROM THE RADON TRANSFORM

As noted above, the HT feature space is equivalent to the RT features generated on the system of Figure 8. Smoothing of the input image and converting edges into lines is required and possible by convolution with Gaussian and edge operators. By the central slice theorem, the 1-D FT of $\hat{f}_\theta$ is the 2-D FT of $f$ evaluated along the line at $\theta$. By the filter theorem, the 1-D convolution in $\rho$ at each $\theta$ for $\hat{f}$ and a reference function $\hat{g}$ is a slice through the 2-D convolution of $f$ and $g$ at $\theta$. Thus, all necessary 2-D filtering operations are possible on projection vectors with 1-D operators. Conventional AO FT and convolvers can easily achieve this at the 60 μsec rates needed (or even faster if required).

Many other features can also be obtained from this RT output [28]. If the 1-D $\hat{f}_\theta$ outputs from Figure 8 are fed each $T_A = 60$ μsec to an acousto-optic (AO) spectrum analyzer, their FT is produced. One large area detector covering half of the FT plane produces a wedge sample $F(\theta)$ of the FT of $f$ each $T_A$. A linear detector array with an integration time $NT_A$ in the other half of the FT plane yields the FT ring samples $F(\rho)$ each $NT_A$. Thus, a WRD FT feature space results. The moments $m_n^\theta$ (the n-th moment of $f$ about $\theta$) can be computed and related to the conventional $m_{pq}$. However, one can also simply compute the first ten $m_{pq}$ from only four $\hat{f}_\theta$ projections [28]. At CMU, we often prefer to use the $\hat{f}_\theta$ features directly rather than converting them into $m_{pq}$ features with a loss of information.

From two orthogonal projections 90° apart, the convex hull rectangular boundary of any object can be determined. With N projections, an N-order polygon defining the object boundary can be obtained. The projection widths versus $\theta$ results in a 1-D feature vector $\underline{w}(\theta)$. This or its FT can be used for object identification.

Polar projections of the integral through the centroid $(\bar{x}\ \bar{y})$ of the object as a function of $\theta$ are another useful descriptor of the object shape. If each projection $\hat{f}_\theta$ is evaluated at the one point

$$\bar{\rho}(\theta) = \bar{x}\cos\theta + \bar{y}\sin\theta, \tag{11}$$

then the 1-D feature vector $\underline{s}(\theta)$ results. This is similar to a chord transform, but only for chords through the centroid (the centroid is easily obtained from moments). It is also analogous to older polar space-variant optical transform work without the scale-invariant Mellin transform properties of this earlier optics research [49].

## 3.7  AUTOCORRELATION OBSERVATION SPACE

The shape and distribution of the autocorrelation of an input object contains significant information useful for object recognition. The general architecture for a processor to analyze such an observation space is shown in Figure 9. In recent work [50], Merkle has considered many different sampling methods and features calculated from an autocorrelation observation space. These digitally-calculated features (computed from an optically-produced

autocorrelation pattern) may require extensive time. The features considered include contour features such as chain codes and Fourier descriptors, various histogram operators, moments of the autocorrelation function, etc. A large set of tests on different characters was performed and the results obtained using different features were compared.
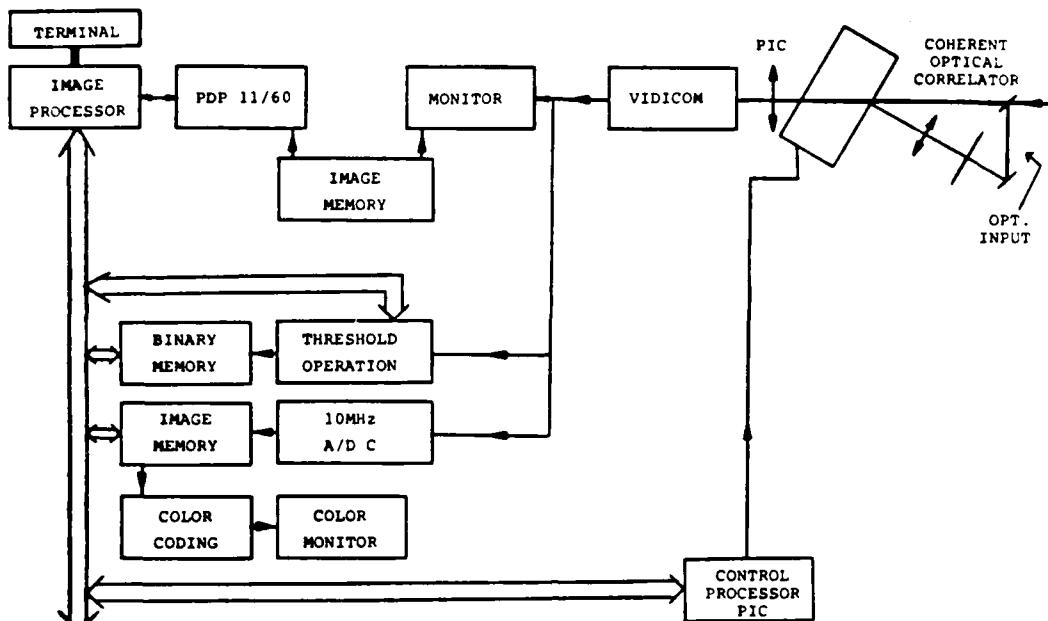


FIGURE 9
Block diagram of a hybrid optical / digital autocorrelation
observation space pattern recognition processor [50]

## 3.8 OTHER FOURIER TRANSFORM FEATURE SPACES

In recent work, the use of CGHs and HOEs to optically realize various special sampling functions (such as wedge ring detection) has been considered and experimentally demonstrated [6]. In other FT observation space research, Duvernoy [84] considered isoenergy contours in the FT plane. He computed the Fourier descriptors for such contours at various levels. Several basis function analyses techniques were used to classify various types of terrain (woods, fields and cities). These isoenergy contours are attractive because they combine spatial frequency as well as directional information. The spatial frequency and directional information is also available from wedge ring detector outputs, however a WRD space provides this information separately, not combined as in an isoenergy contour analysis.

## 3.9 DISCUSSION

As one can easily see, there is significant new research on optically generated features and feature extractors. These new advances allow many different observation spaces to be used. The optical generation and calculation of all major feature spaces has been demonstrated and described. The attractive aspects of this research include:

(1) The same architecture can compute the features for any input object. Thus, a new architecture is not necessary for a new object identification problem.

(2) The parallelism of optics in feature generation and the flexibility of digital feature extractors and classifiers are matched quite well in these architectures.

(3) A different discriminant vector $w$ or feature extractor algorithm can easily be included in the digital post-processor, should a given object identification data base necessitate this.

(4) Dimensionality reduction can and has been employed to reduce the calculations required by the digital post-processor. Alternatively, the projections and transformations can be optically implemented if desired using computer generated holograms.

The shortcomings of these or any digital or analog feature extractor for pattern recognition include:

(1) A higher susceptibility to noise. This is a direct result of dimensionality

reduction.

(2)  The need to segment the input object into interesting candidate regions before
     feature extraction.  We refer to pattern recognition architectures capable of
     handling multiple object simultaneously as shift-invariant (e.g., a correlator).

## 4.  RECENT OPTICAL CORRELATOR ADVANCES

### 4.1  MULTIPLE MSF CORRELATORS
Several advances in optical correlators are briefly reviewed in this section.  The initial
description of Figure 1 assumed a single MSF at $P_2$, however multiple MSFs are also possible
[51-53].  In these cases, spatial and/or frequency-multiplexing of the MSFs are used.  In
the space-multiplexed case, the FT of the input object must be replicated, multiplied by the
different MSFs at different spatial locations in $P_2$ and the correlation of the input with
the different MSFs performed.  Holographic lens arrays and HOEs [52], a fixed screen tech-
nique [53] or a rotating grating [54] can be used to access these multiple filters.  In the
latter case, separate output correlations appear sequentially.  In the other cases, multiple
correlations are available in parallel or can they can all be superimposed (the first choice
requires the analysis of multiple correlation planes whereas the second choice results in
poorer correlation plane SNR).  The best choice depends upon the application.

### 4.2  SPECTRAL CORRELATORS
F.T.S. Yu [55], Ludman [56] and others have actively pursued the use of color or spectral
MSF processors for image processing (image subtraction, deblurring, etc.) and pattern
recognition [55].  These processors have the architecture of Figure 1 with a color input
image, a tricolor grating behind $P_1$ and a white light source at red, green and blue wave-
lengths.  This forms the FT of the portion of the input in each spectral (color) band in a
different spatial location in $P_2$.  Thus, different MSFs can be applied to different spectral
data.  Alternatively, objects in different colors in the input will produce correlation
peaks at $P_3$ in different wavelengths.  The power dissipation and availability of the neces-
sary light sources is a practical problem with such architectures.  The use of color diver-
sity appears to best be utilized as an adjunct to the conventional x,y degrees of freedom
of the system to simplify system fabrication and output data analysis [57].

### 4.3  HYBRID OPTICAL/DIGITAL PATTERN RECOGNITION, IMAGE UNDERSTANDING AND
###       ARTIFICIAL INTELLIGENCE
The use of pattern recognition (PR), image understanding (IU) and artificial intelligence
(AI) techniques in a hybrid optical/digital architecture has recently been addressed in an
interdisciplinary program at CMU.  A general diagram of the architecture is shown in Figure
10.  The optical portion of the system produces features and correlations with generic SDFs
(see Section 5).  Both optically and digitally computed features are considered and the
optical systems are adaptively controlled by feedback from an AI/IU processor that compares
the results obtained to a world model and which uses the results obtained to adaptively
construct and modify the world model.  Such an advanced general architecture appears to be
most attractive for new supercomputers.  Initial tests on aircraft images, an on-line tech-
nique for producing reference objects in any 3-D orientation by synthesis of the object as
polygons, and related Hough transform feature representations for objects appear to make
such an architecture most attractive and realistic for advanced hybrid supercomputers.

### 4.5  SUB-PIXEL TARGET LOCATION, TRACKING AND IDENTIFICATION
Another recent optical correlator application under research at CMU involves the location,
tracking and identification of moving sub-pixel targets from space-based mosaic sensors.
The technique used involves:  (1)  the correlation of two successive image frames, (2)
sampling the central 3 x 3 or 5 x 5 region of the correlation plane, (3) by estimation de-
termining the shift between two successive image frames (this is achieved to sub-pixel
accuracy), and (4) the interpolation and resampling to shift one of the images by this
estimated sub-pixel amount, and (5) registration and the subtraction of these two frames.
The shift and registration must be performed to sub-pixel accuracy to extract the target.
In Figure 11a, we show a typical input with a sub-pixel target 0.2 of a pixel in size.  A
sequence of three such frames was produced with the background shifted by 0.1 pixels frame-
to-frame and with the target shifted by one pixel frame-to-frame.  The result after proces-
sing (Figure 11b) shows the successful location of the sub-pixel target and its relative
position in each frame.  Such a time-history track file provides the necessary information
for target identification and classification.
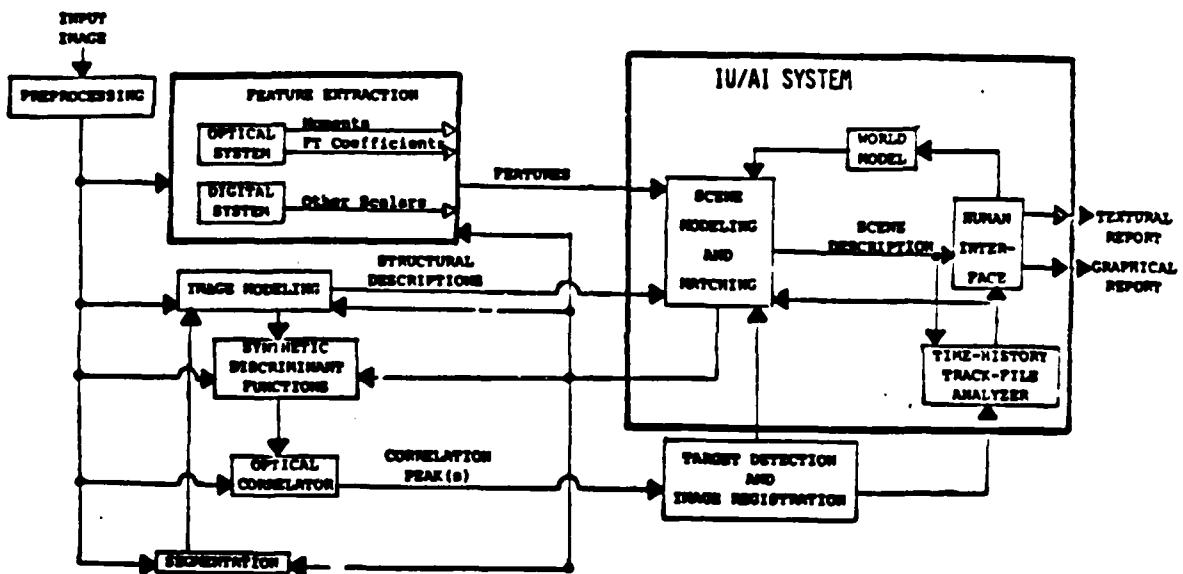
FIGURE 10
Block diagram of a hybrid optical/digital combined
pattern recognition/image understanding/artificial inelligence image processor



(a)                                                    (b)
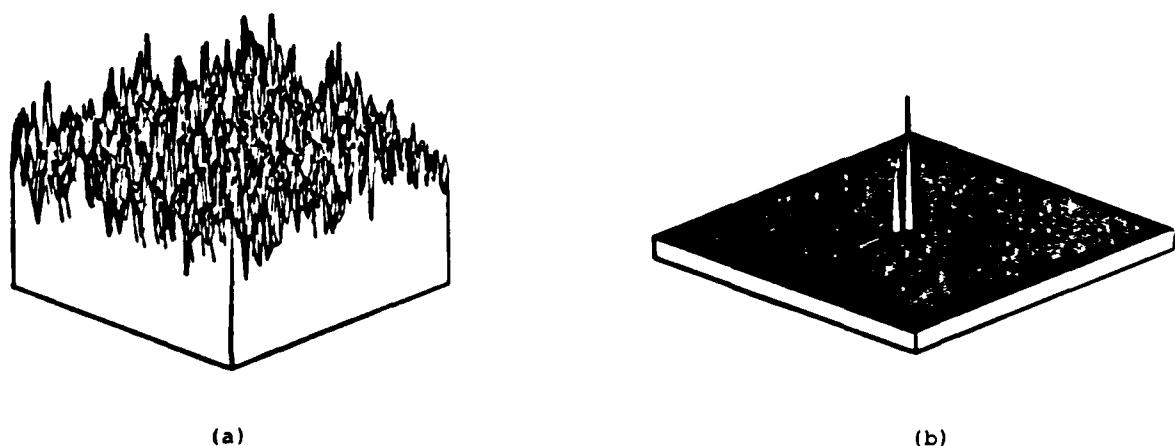
FIGURE 11
Typical imagery (a) and output (b) data plane showing the successful
detection and tracking of a sub-pixel target

Correlators represent quite powerful pattern recognition processors with large processing gain and shift-invariance plus the ability to handle multiple objects simultaneously.  This correlation operation is easily achieved optically (Figure 1).  Although it is the optimum detection scheme only for white Gaussian noise, its performance in practical structural clutter is well-known and has recently been theoretically addressed [58].  Image processing is quite tolerant of the dynamic range requirements of the data.  In fact, binary data performs quite well [59] and is often necessary with some SLMs [60].  The susceptibility of a correlator to distortions between the input image and reference MSF object are its well-known shortcomings.  A basic method to overcome this disadvantage and yet retain advantageous properties of a correlator is shown in Figure 12.  This method involves the synthesis of a synthetic discriminant function (SDF) from a training set of several images of each class, rather than forming a single image representation of an object in one orientation (and using multiple such images in a multi-channel correlator).  The basic technique used is to select a basis function set from the training images (these consist of different distorted views of each object class) and from this synthesize an SDF.  An MSF of this SDF is then produced and used in an optical correlator (Figure 1).  The SDF = h is a linear combination of the basis function set {φ} or the training set images {f}

$$h(x,y) = \Sigma b_n \phi_n(x,y) \ , \ h(x,y) = \Sigma a_n f_n(x,y). \tag{12}$$

This concept was originated by Hester and Casasent [61], applied and demonstrated for intra-class [62] and inter-class [63] recognition.  The filter h and the associated orthonormal basis function set selection by Gram-Schmidt, KL and other techniques have been detailed previously [61-63].  The generalized matched filters (GMFs) of Caulfield [64,65] are a special case of the SDF where the basis function set φ are the exponentials and a Fourier coefficient feature space is used.  No general solution to the $N^2$ coefficients required to be computed in GMFs has been advanced and the system is not necessarily shift-invariant because the full correlation plane response is not specified.  The circular harmonic SDFs of Arsenault [66] use one circular harmonic in the expansion of $f(r,\theta)$ to synthesize the filter.  These filters achieve only rotation-invariance with shift-invariance being a possibility.  In recent work, Stark [83] noted that high $SNR_i$ may be required and that the choice of the center of expansion and the harmonic(s) used is not easy and that for complex objects the resultant processor may not be shift-invariant.  Stark [83] recently offered a vector formulation, used FK techniques and retained several harmonics in an improved version of these circular harmonic filters.  Both these filters and the GMFs require far more extensive noise and discrimination tests on large data bases before they can more properly be assessed.  Since SDFs are more developed, tested, analyzed and have a clear mathematical basis and synthesis algorithm, they will be emphasized.
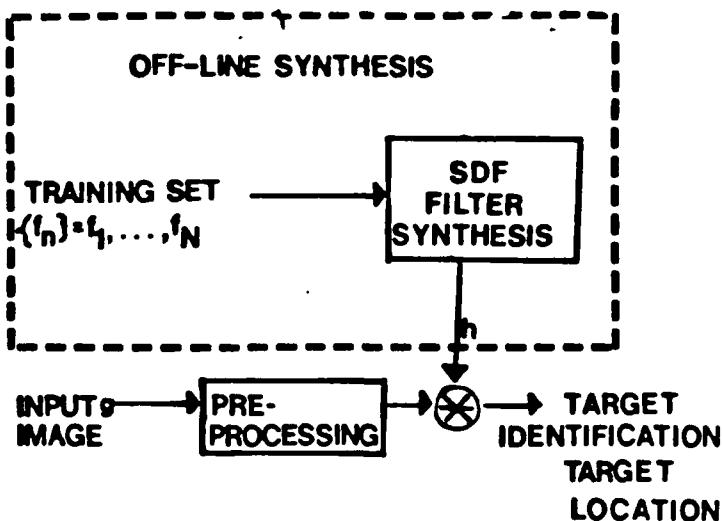


FIGURE 12
General synthetic discriminant function (SDF) matched spatial filter (MSF)
distortion-invariant hybrid correlator concept

SDFs are synthesized from the vector inner product matrix $V$ of the training set images by specifying the desired correlation plane values at different locations (such as the peak of the correlation function). These values are specified by a vector $u$. Depending upon the application, five different SDFs are possible. Each corresponds to a different vector $u$ and matrix $V$. However, in each case, the SDF is defined by

$$\underline{a} = \underline{V}^{-1}\underline{u} \tag{13}$$

which specifies the coefficients $a_n$ in (12) which define the SDF $= h$. If $u$ is all unity, an SDF with the same output correlation peak intensity for all objects of one class (independent of the geometric distortion chosen) results. With alternate choices for $u$, a two-class SDF with unity correlation peak values for objects of one class and zero peak values for objects of class two results. Alternate projection values (1,2,3) allow one SDF to discriminate between three object classes, with the value of the correlation output defining the object class. Use of several SDFs and a truth table of the multiple correlation plane outputs at each location yields the final type of SDF. These SDF synthesis techniques are unified in [67]. Excellent performance on ship imagery has been obtained with these SDFs as summarized in [68] where their performance in noise was also quantified.

We refer to these as projection SDFs. They do not often perform adequately, due mainly to the fact that the synthesis algorithm specifies the correlation peak value at only one point (the center of the correlation function). As a result, nothing prohibits peak values above threshold from occurring for shifted versions of a false target (for which an output below threshold, ideally zero is expected). New correlation SDFs [69] overcome this by specifying the correlation plane values for true and false targets at the central peak and $\pm d_S$ pixels away in x and y. The specified value $\pm d_S$ pixels away is generally zero and the value at the central peak is generally one (for true class objects) and zero (for false class objects). This synthesis algorithm is realized exactly as before with the inclusion of shifted versions of each training image. This results in a well-controlled correlation peak shape (a large central peak and zero or low values $\pm d_S$ pixels away) for true targets and zero values (central and $\pm d_S$ pixels away) for false targets. This also allows the use of both a peak threshold $T_T$ and a peak to sidelobe ratio threshold $C_T$ to be applied to the output correlation plane pattern to determine if a candidate region of the input image contains a target and the class of that target. Recent tests performed with these correlation SDFs considered three automatic target recognition (ATR) objects (Tank 1, Tank 2 and an armored personnel carrier APC). Figure 13 shows representative images of the APC and one of the tanks. For each of these objects, 36 aspect views were available at 10° increments and a given depression angle around the object. The target resolution on these images was degraded to about 50 x 20 pixels. The objective was to form an SDF using only 6 or so different aspect views such that the correlation plane pattern had a peak for one class and no peak for the other object class. Table 1 shows results [69] obtained with three different types of correlation SDFs intended to discriminate Tank 2 from Tank 1 independent of 3-D aspect distortions. As seen, 93-95% correct classification with no missed targets is possible using only 6 aspect views to synthesize the SDF and with the SDF tested against all 72 aspect views of both object classes. Table 2 shows similar data for an SDF to discriminate APCs from tanks. Here, with 12 training set images/class, we find perfect performance to be possible. Figure 14 shows noise test results when four targets (2 tanks and 2 APCs) not present in the training set were placed in a typical scene (Figure 14a) with an input SNR approximately equal to one. The output correlation plane (Figure 14b) shows only two peaks at the correct location of the two tank objects. Clearly, the SDF has discriminated against the APC targets and other structured noise clutter in this scene. These are typical of the excellent results obtained for full correlation plane analyses of the SDF performance for 3-D distortion-invariant multi-class target recognition in clutter.

Recently, considerable attention has been focused on the importance of the efficiency of optical correlators (i.e. the usable optical light in the output correlation plane compared to the energy of the input image) [70] and to the use of phase only MSFs [71] to improve light efficiency. Butler and Riggins [72] distinguish between a phase only MSF in which only the phase of the MSF data is recorded on either an absorption media or on a bleached material. Phase MSFs provide more useful light and initial results indicate that they provide better discrimination (correlation plane SNR). However, only limited simulations on two similar letters were performed [71] and no theoretical basis has yet been advanced for this result. The motivation for this recent attention and an entire conference session [73] on CGH realization of SDFs is the excellent performance of these filters, the availability of several commercial CGH recorders and the use of such CGH filters in the fabrication of a compact SDF-based correlator. Gianino and Horner [75] quantified by simulation the expected worse sensitivity of phase-only MSFs with respect to object distortions. Thus, the use of SDFs with phase-only CGHs is a natural approach to consider. It provides better efficiency and overcomes the distortion sensitive performance of conventional filters. Riggins and Butler [74] recently simulated the original equal correlation peak projection SDF with a CGH. They found 1% light efficiency and good performance on ATR data. However, much more extensive tests are required on the new advanced SDFs and on larger data bases. Kumar et al [76]

have provided the only theoretical basis for the optical efficiency, space bandwidth product and number of filters to be used in an SDF. This work has also only touched the surface of this problem however.
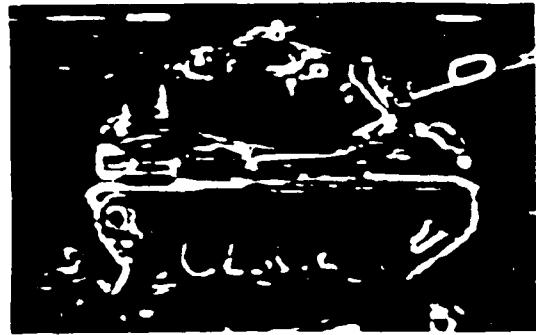


(a)



(b)
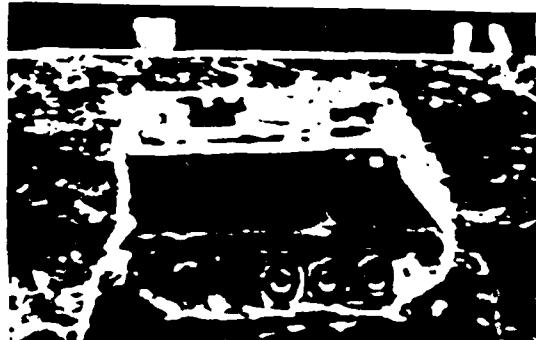


(c)
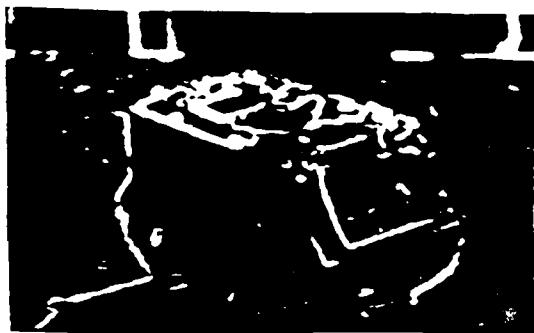


(d)

FIGURE 13
Typical ATR images of a tank (a and b) and an APC (c and d) target
objects used in correlation SDF tests

| SDF | SDF-1 | | SDF-2 | | SDF-3 | |
|---|---|---|---|---|---|---|
| INPUT | TANK 1 | TANK 2 | TANK 1 | TANK 2 | TANK 1 | TANK 2 |
| NUMBER $T_T \gtrsim 0.5$ ERRORS | 3 | 0 | 3 | 0 | - | - |
| $P_C$ ($T_T$ ONLY) | 95.8% | | 95.8% | | ----- | |
| NUMBER $C_T$ ERRORS (LARGEST PEAK ONLY) | 1 | 3 | 3 | 0 | 5 | 0 |
| $P_C$ ($C_T$, LARGEST PEAK) | 94.4% | | 95.8% | | 93% | |
| $P_e$ (MISSED TARGETS) | 0 | | 0 | | 7 | |

TABLE 1
Tank 1 versus Tank 2 correlation SDF test results using 6 training
set images per class. Errors out of 72 test images are listed

| SDF INPUT | SDF-1 | | SDF-2 | | SDF-3 | |
|---|---|---|---|---|---|---|
| | TANK 1 | TANK 2 | TANK 1 | TANK 2 | TANK 1 | TANK 2 |
| NUMBER $T_T \leq 0.5$ ERRORS | 0 | 0 | 0 | 0 | - | - |
| $P_C$ ($T_T$ ONLY) | 100 | | 100 | | ---- | |
| NUMBER $C_T$ ERRORS (LARGEST PEAK ONLY) | 0 | 0 | 0 | 2 | 3 | 2 |
| $P_C$ ($C_T$, LARGEST PEAK) | 100 | | 97.2% | | 93% | |
| $P_e$ (MISSED TARGETS) | 0% | | 0% | | 4.2% | |

TABLE 2
Tank 1 versus APC correlation SDF test results using 6 training set images per class.
Errors out of 72 test images are listed.
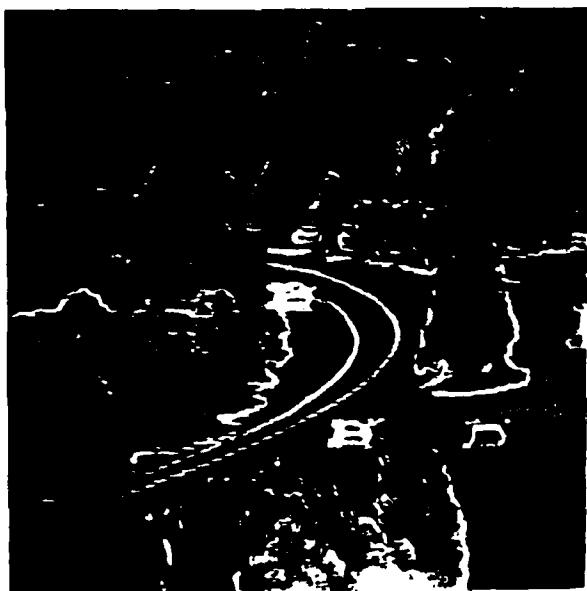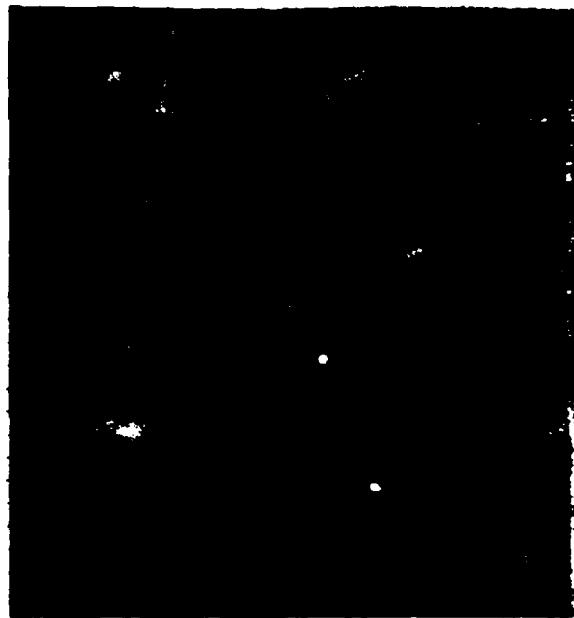


(a)                                                        (b)

FIGURE 14
Multiple non-training set targets in clutter (a) and correlation
output plane with peaks for correlation SDF filtering at true target locations (b)

Since these prior tests did not employ the new correlation SDFs, we synthesized a correla-
tion SDF-3 of the ship and tested it against real ship imagery with different amounts of
amplitude and phase data retained. We found that the phase-only filter performed well, but
that retaining two bits of amplitude MSF data and two bits of phase data gave significantly
better results. Thus, from these recent tests of ours, it appears that a filter with some
amplitude data present is preferable to a phase-only filter. Considerable future work and
results are anticipated in this area.

## 6. SYSTEM FABRICATION

A 2-D real-time SLM is the key element for a successful parallel 2-D OPR system. The
state-of-art of these devices is summarized in [5]. Recent Soviet work has resulted in high
performance PRIZ [81] and liquid crystal SLMs with high sensitivity, resolution and effi-
ciency and with unique properties such as directional spatial filtering, edge enhancement
and the ability to detect and respond only to dynamic removing input objects [81]. Many
real-time optical correlators have been fabricated, described and demonstrated. The General
Motors system for robot inspection [77] is one such system which used a liquid crystal real
time input SLM. Two-dimensional output readout was simplified and rotational invariance
accomplished by use of a rotating prism and a cylindrical optical system using two 1-D detector
arrays, rather than a 2-D readout array. Several multi-channel real-time optical correla-
tors using a liquid crystal input transducer and multiple MSFs have been fabricated at
Huntsville as described in [78]. In these systems, attention was given to filter synthesis
using weighted MSFs [82] to reduce, rather than overcome, scale and angular object sensitiv-
ity and hence the number of multiple filters needed. Multiple MSFs on the same filter were
tested on these systems for light efficiency and spatially-separated MSFs were accessed by
different laser diodes in different input spatial locations. In the first system, all
correlation plane patterns were superimposed. In the second system, different laser diode
sources allowed separate MSFs to be accessed when the difference between the orientation or
scale of the input and reference object caused the correlation peak to drop sufficiently.

A recent magneto optic SLM [43] offering low cost has been developed and demonstrated for
simple white light spatial filtering [80] image processing functions (rather than correla-
tions) and for low space bandwidth product CGH MSF correlations [60]. The binary (rather
than gray scale) response of this SLM and its present low resolution and low transmittance
are limitations that must be overcome before it will see general use.

All of these prior real-time optical correlator systems did not attempt to significantly
reduce the physical size of the optical system. A compact portable version of the Huntsville
optical correlator was recently fabricated by ERIM [79] and is shown in Figure 15. This
system uses four laser diode sources to access one of four spatially-multiplexed MSFs with
the output correlation plane detected by a 2-D charge injection device (CID) detector array.
The correlation unit is 15 x 23 x 42 cm and weighs 8 kg. Electronic support unit for it
is 15 x 26 x 35 cm and weighs 8 kg also. The total power consumption of this portable
compact optical correlator is 55W. It is possible to fabricate far smaller and lower power
dissipation versions of this architecture and several of these are presently being con-
sidered.

All prior well-engineered real-time optical correlators have used only simple or several
simple multiple MSFs and have thus achieved only limited distortion-invariant pattern recog-
nition. While the physical size of these processors is significantly less than the classic
large optical bench processors, they are not yet compact enough for use in a missile. A
more practical optical correlator would be one which employed the advanced SDF MSFs and one
which was significantly smaller in size. The use of SDF filters would reduce the complexity
of the system and extend its practicality and versatility. The system of Figure 16 was
recently fabricated by General Dynamics-Pomona and demonstrated in initial tests using com-
puter generated hologram SDF MSFs. The system is less than 5 inches in diameter and approxi-
mately 12 inches in length. It is intended for use in a 5 inch missile for on-line real-
time ATR pattern recognition. It employs folded optics, mirrors rather than lenses, multi-
ple SDFs, several output correlation planes, and presently a real-time liquid crystal SLM.
Tower tests and captive helicopter tests of this system are expected to be among the high-
lights of OPR work in 1985. This real-time optical correlator of Figure 16 represents the
first such processor suitable for airborne use in a 5 inch missile that has reached hardware.
Further such committments and research support by government and industry are essential to
provide the necessary transfer of technology from OPR research to airborne hardware.

## 7. SUMMARY AND CONCLUSION

A brief review of the major operations achievable and the major OPR architectures has
been provided. This was followed by descriptions of nine different optical feature extractor
hybrid pattern recognition processors. These systems optically produced all of the major
image features using the parallism of optics. Feature extraction and classification on these
optically generated features is then performed in a digital post-processor. The resultant

hybrid optical/digital systems combine the best advantages of optical and digital processors. These techniques are quite noteworthy because the same optical architecture can compute the indicated features for any input object and thus the same system is usable for any object identification application. The discriminant function, feature extractor, transformation and classifier used can be changed as desired by employing the flexibility of the post-processor. In all cases, multi-class 3-D distortion-invariant pattern recognition is the objective considered. Extensive tests have been made on several of these systems on large data bases. These include a large number of related objects in different classes with 36 different aspect views of each object (at every 10° increment) from a 20-40° depression angle. Thus, this represents a multi-class full 3-D distortion problem. Training sets containing only 4-6 different spatial views per object class were found to be adequate to provide excellent 86-98% correct object recognition and classification identification in over 300 test images in one data set and over 175 test images in a second data base.
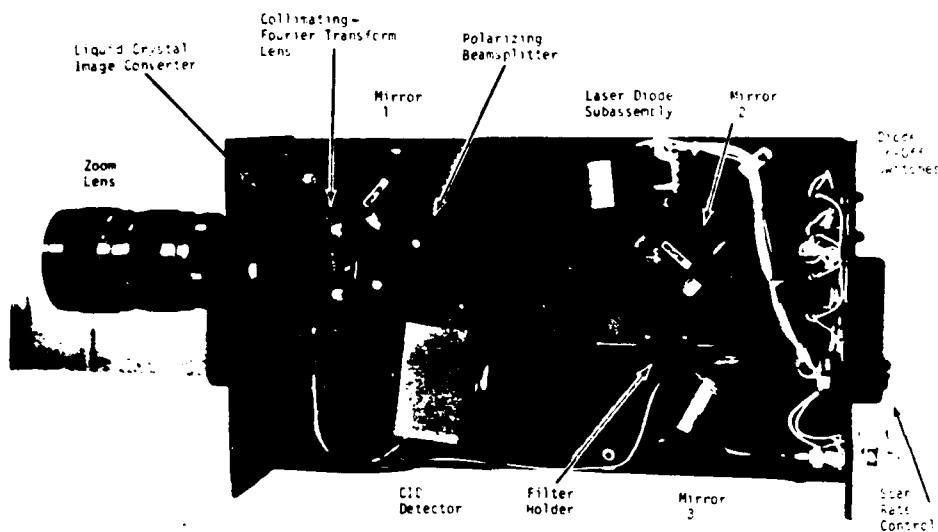


FIGURE 15
Compact portable real-time liquid crystal optical
correlator using 4 matched spatial filters [79]
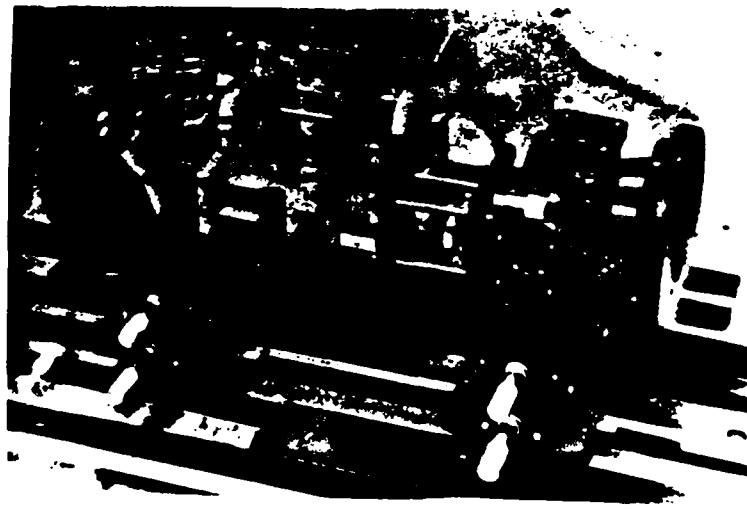


FIGURE 16
Photograph of the General Dynamics-Pomona airborne real-time optical SDF correlator
for packaging in a 5" missile (Photo courtesy of D. Fetterly, General Dynamics-Pomona)

When performance in high clutter and noise is required, a correlator is needed. With SDFs, a 3-D distortion-invariant multi-class pattern recognition correlator is possible with all of the advantages of a correlator retained and with distortion-invariance provided. Excellent initial test results of the full correlation plane data were presented and a typical example of the performance in clutter of the system was included. The final issue in an OPR system is system fabrication. As shown, significant strides have recently been made in this area with an advanced SDF-based real-time optical correlator package for a five inch missile having been fabricated and initially tested. The world of optical pattern recognition has a bright and attractive future in all aspects. Further government and corporate committments are still necessary to insure timely transfer of this technology to hardware however.

## ACKNOWLEDGMENTS

## REFERENCES

1. D. Casasent, _Proc. IEEE_, Vol. 67, 813 (May 1979).
2. D. Casasent and D. Fetterly, _Proc. SPIE_, Vol. 456, 105 (January 1984).
3. D. Casasent, _Optical Engineering_, Vol. 24, 26 (January 1985).
4. D. Casasent, _Proc. SPIE_, Vol. 431, 263 (August 1983).
5. A. Tanguay, _Proc. SPIE_, Vol. 456, 130 (January 1984).
6. D. Casasent, _Proc. SPIE_, Vol. 532 (January 1985).
7. _Proc. IEEE_, Special Issue on Optical Computing, Vol. 72, No. 7 (July 1984).
8. "Applications of Digital Image Processing VII", _Proc. SPIE_, Vol. 504, Ed. A. Tescher (August 1984). (See also all prior Proc. SPIE Digital Image Processing volumes).
9. _Proc. IEEE_, Special Issue on Digital Image Processing, Vol. 67 (May 1979).
10. G.G. Lendaris and G.L. Stanley, _Proc. IEEE_, Vol. 58, 198 (February 1979).
11. H. Kasdan and D. Mead, _Proc. EOSD_, 248 (1975).
12. D. Casasent and V. Sharma, _Optical Engineering_, Vol. 23, 492 (Oct/Nov 1984).
13. D. Casasent, R.L. Cheatham and D. Fetterly, _Applied Optics_, Vol. 21, 3292 (September 1982).
14. J.A. Blodgett, R.A. Athale, C.L. Giles and H.H. Szu, _Optics Letters_, Vol. 7, 7 (January 1982).
15. K. Wagner and D. Psaltis, _Proc. SPIE_, Vol. 352, 82 (August 1982).
16. D. Casasent and R.L. Cheatham, _Proc. ASME_ (August 1984).
17. D. Casasent and R.L. Cheatham, _Proc. SPIE_, Vol. 504, 19 (August 1984).
18. D. Nichols, _Optics Communications_, Vol. 43, 168 (October 1982).
19. D. Nichols, _Optics Communications_, Vol. 48, 242 (December 1983).
20. D. Casasent and W.T. Chang, _Lasers and Electro-Optics_, Vol. 24, 18 (September 1982).
21. D. Casasent and W.T. Chang, _Applied Optics_, Vol. 22, 2087 (July 1983).
22. W.T. Chang and D. Casasent, _Proc. SPIE_, Vol. 521 (November 1984).
23. D. Moore and D. Parker, _Pattern Recognition_, Vol. 6, 149 (1974).
24. S. Smith and A. Jain, _Computer Graphics in Image Processing_, Vol. 20, 1699 (1982).
25. H. Barrett, _Optics Letters_, Vol. 7, 248 (1982).
26. A. Gmitro et al, _Optical Engineering_, Vol. 19, 260 (1980).
27. G. Eichmann and B. Dong, _Applied Optics_, Vol. 22, 830 (March 1983).
28. G. Gindi and A. Gmitro, _Optical Engineering_, Vol. 23, 499 (September 1984).
29. S. Deans, _Proc. IEEE_, Vol. PAMI-2, 185 (1981).
30. K. Sloan, _Proc. IEEE_, Vol. PAMI-4, 87 (1982).
31. R. Duda and P. Hart, _Commun. ACM_, Vol. 15, 11 (1972).
32. J. Sklansky, _Proc. IEEE_, Vol. C-27, 923 (1978).
33. K. Fukunaga, _Introduction to Statistical Pattern Recognition_, Academic Press, New York (1972).
34. J. Kittler and P. Young, _Pattern Recognition_, Vol. 5, 335 (1973).
35. R.O. Duda and P.E. Hart, _Pattern Classification and Scene Analysis_, John Wiley and Sons, New York (1973).
36. K. Fukunaga and W.L. G. Koontz, _IEEE, Trans. Comp._, Vol. C-19, 311 (April 1970).
37. _Proc. SPIE_, Vol.528, Digital Image Processing: Critical Review of Technology, Ed. A. Tescher (January 1985).
38. J. Leger and S.H. Lee, _Applied Optics_, Vol. 21, 274 (January 1982).
39. J. Leger and S.H. Lee, _J. Opt. Soc. Am._, Vol. 72, 556 (1982).
40. Z.H. Gu, J. Leger and S.H. Lee, _J. Opt. Soc. Am._, Vol. 72, 787 (1982).
41. Z.H. Gu and S.H. Lee, _Applied Optics_, Vol. 23, 822 (March 1984).
42. Z.H. Gu and S.H. Lee, _Optical Engineering_, Vol. 23, 727 (November 1984).
43. W. Ross, D. Psaltis and R. Anderson, _Optical Engineering_, Vol. 22, 485 (1983).

44. B.V.K. Vijaya Kumar and C. Carroll, Optical Engineering, Vol. 23, 732 (November 1984).
45. K. Brenner and A. Lohman, Optics Communications, Vol. 42, 312 (1982).
46. L. Jacobson and H. Wechsler, Pattern Recognition Letters, Vol. 1, 61 (1982).
47. D. Casasent and R.L. Cheatham, Optics Communications, Vol. 51, 227 (September 1984).
48. D. Casasent and R.L. Cheatham, "Hierarchical Feature-Based Object Identification", OSA Topical Meeting on Machine Vision (March 1985).
49. D. Casasent and D. Psaltis, "Deformation-Invariant, Space-Variant Optical Pattern Recognition", Chapter in Progress in Optics, Ed. E. Wolf, Vol. XVI, Holland Pub. Co., New York, pp. 291-356 (1979).
50. F. Merkle and T. Lorch, Applied Optics, Vol. 23, 1509 (May 1984).
51. K. Leib, R. Bonburant and M. Wohlers, Optical Engineering, Vol. 19, 414 (1980).
52. J. Mendelsohn, M. Wohlers and K. Lieb, Proc. SPIE, Vol. 186, 190 (1979).
53. H.K. Liu and J.G. Duthie, Applied Optics, Vol. 21, 3278 (1982).
54. F.T.S. Yu and X.J. Lu, Applied Optics, Vol. 23, 3109 (September 1984).
55. F.T.S. Yu, Optical Engineering, Vol. 23, 690 (November 1984).
56. C. Warde et al, Optics Communications, Vol. 49, 241 (March 1984).
57. D. Psaltis and D. Casasent, Proc. SPIE, Vol. 232, 24 (April 1980).
58. B.V.K. Vijaya Kumar and C. Carroll, J. Opt. Soc. Am. - A, Vol. 1, 392 (1984).
59. B.V.K. Vijaya Kumar and D. Casasent, Applied Optics, Vol. 20, 3707 (1981).
60. D. Psaltis, E. Paek and S. Venkatesh, Optical Engineering, Vol. 23, 698 (November 1984).
61. C. Hester and D. Casasent, Applied Optics, Vol. 19, 1758 (June 1980).
62. C. Hester and D. Casasent, Proc. SPIE, Vol. 292, 25 (August 1981).
63. C. Hester and D. Casasent, Proc. SPIE, Vol. 302, 108 (August 1981).
64. H.J. Caulfield, Applied Optics, Vol. 19, 3877 (December 1980).
65. H.J. Caulfield and M.H. Weinberg, Applied Optics, 21, 1699 (May 1982).
66. H. Arsenault, Y.N. Hsu and K.C. Macukow, Optical Engineering, Vol. 23, 705 (November 1984).
67. D. Casasent, Applied Optics, Vol. 23, 1620 (May 1984).
68. D. Casasent, W. Rozzi and D. Fetterly, Optical Engineering, Vol. 23, 716 (November 1984).
69. W.T. Chang, D. Casasent and D. Fetterly, Proc. SPIE, Vol. 507, 9 (August 1984).
70. J. Horner, Applied Optics, Vol. 21, 4511 (1982).
71. J. Horner and T. Gianino, Applied Optics, Vol. 23, 812 (March 1984).
72. S. Butler and J. Riggins, Proc. SPIE, Vol. 519 (November 1984).
73. Proc. SPIE, Vol. 519, International Optical Computing Conference (November 1984).
74. J. Riggins and S. Butler, Optical Engineering, Vol. 23, 721 (November 1984).
75. P. Gianino and J. Horner, Optical Engineering, Vol. 23, 695 (November 1984).
76. B.V.K. Vijaya Kumar, E. Pochapsky and D. Casasent, Proc. SPIE, Vol. 519 (November 1984).
77. A. Gara, Applied Optics, Vol. 18, 172 (1979).
78. B. Guenther, C. Christensen and J. Upatnieks, IEEE, J. Quantum Electron., Vol. QE-15, 1348 (1979).
79. J. Upatnieks, Applied Optics, Vol. 22, 2798 (September 1983).
80. F.T.S. Yu, X. Lu and M. Cao, Applied Optics, Vol. 23, 4100 (November 1984).
81. D. Casasent, F. Caimi and M. Petrov, Applied Optics, Vol. 21, 3846 (November 1982).
82. D. Casasent and A. Furman, Applied Optics, Vol. 16, 1652 (June 1977). See also D. Casasent and A. Furman, Applied Optics, Vol. 16, 1662 (June 1977).
83. A. Wu, and H. Stark, Applied Optics, Vol. 23, 838 (March 1984).
84. J. Duvernoy, Applied Optics, Vol. 23, 828 (March 1984).

# 5. COMPUTER GENERATED HOLOGRAMS FOR WEDGE RING DETECTOR FOURIER TRANSFORM PLANE SAMPLING

Ref. 8

# "A COMPUTER GENERATED HOLOGRAM FOR DIFFRACTION-PATTERN SAMPLING"

David Casasent and Jian-zhong Song*

Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

*Permanent Address: Changchun Institute of Optics & Fine Mechanics, Academia Sinica
P.O. Box 1024, Changchun, People's Republic of China

## ABSTRACT

Diffraction pattern sampling provides a feature space suitable for object classification, orientation and inspection. It allows significant dimensionality reduction. These properties are best achieved by the use of specifically-shaped Fourier transform plane detector elements and this can be realized with considerable flexibility, reduced size and improved performance by the use of computer generated holograms.

## 1. INTRODUCTION

The Fourier transform (FT) or diffraction plane of an object contains a distribution of the spatial frequencies present in the input object. This distribution has many attractive properties. The magnitude of the FT pattern is shift-invariant. Thus, translations of the input image do not effect the magnitude of the Fourier coefficients. Higher horizontal or vertical input spatial frequencies $(u,v)$ lie further from the center (dc or zero spatial frequency) of the FT plane

$$(u,v) = (x_2/\lambda f_L, y_2/\lambda f_L). \tag{1}$$

Input spatial frequencies oriented at an angle in the input plane appear at a radial distance $r = (x_2^2+y_2^2)^{1/2}$ in the FT plane (where $(x_2,y_2)$ are the distance coordinates of the FT plane) and at an angle orthogonal to the orientation of the input data [1]. As the orientation of the input spatial frequencies varies, the angle $\theta$ of the FT distribution also rotates. As the scale of the input object changes, the radial distance at which the frequency peaks are located also scales. Thus, spatial frequency and orientation information are conveniently available in an FT plane representation. Also, such an FT plane representation is most suitable for dimensionality reduction of the data. This issue is of considerable practical importance since the space bandwidth product (SBWP) or number of frequency-plane components required to represent the input object is equal to the input SBWP. Thus, no advantage is obtained by use of a FT plane data representation (in terms of processing requirements), unless dimensionality reduction is employed. Fortunately, an FT plane is well-known to allow considerable data compression, especially for pattern recognition and object identification applications. Hence, an appropriately-sampled FT plane provides a set of features that are most useful for feature extraction based pattern recognition and object identification.

In Section 2, we review the FT properties and prior approaches to efficient FT plane sampling using elements such as the wedge ring detector (WRD). This section provides motivation for our research. In Section 3, we describe our computer generated hologram (CGH) WRD FT plane concept and in Section 4 we detail our synthesis approach for a WRD using CGHs. Section 5 provides initial experimental results obtained using our CGH generated WRD element. Advanced analysis issues and our summary and conclusions associated with this system are then advanced in Section 6.

## 2. WRD PROPERTIES AND FEATURES

For completeness, we first review several common FT properties of use in WRD-sampled FT plane analysis. We consider real input functions $f(x,y)$, i.e. images. Their intensity FT is symmetric, i.e.

$$|F(u,v)|^2 = |F(-u,-v)|^2. \tag{2}$$

The intensity FT is also shift-invariant

$$|\mathcal{F}[f(x-a),y-b)]|^2 = |F(u,v)|^2, \tag{3}$$

i.e. only the phase varies with a shift of the input function. From (2) and (3), we see that an FT plane detector can be placed in the center of the FT plane and that either half of the FT plane can be sampled with no loss of information. The rotational feature of the FT pattern (the FT pattern rotates as the input object rotates) was noted earlier and is described mathematically as

$$|\mathcal{F}[f(x\cos\theta-y\sin\theta,x\sin\theta+y\cos\theta)]|^2 = |F(u\cos\theta-v\sin\theta,u\sin\theta+v\cos\theta)|^2. \qquad (4)$$

The scaling FT pattern property noted earlier is described mathematically as

$$|\mathcal{F}[f(ax,ay)]|^2 = |(1/a)^2 F(u/a,v/a)|^2. \qquad (5)$$
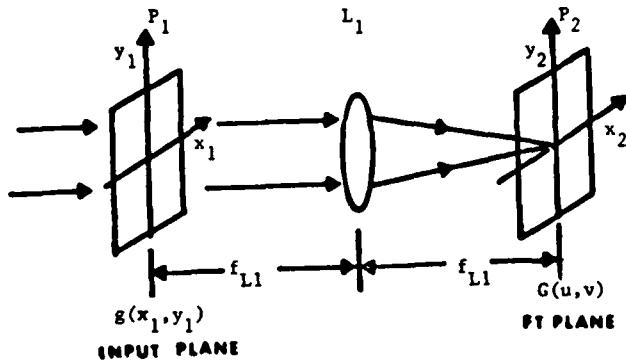


FIGURE 1
Optical wedge ring detector sampled Fourier transform object-recognition processor.



(A)

(B)

(C)

FIGURE 2
(A) Schematic, (B) silicon device and (C) control unit for a silicon-based wedge ring detector

The shift-invariant, rotation and scale properties in (3) - (5) make sampling of the FT pattern intensity with wedge and ring shaped detector elements most attractive. The typical optical arrangement used is shown in Figure 1. The input object is placed in plane $P_1$, its FT pattern is formed at $P_2$ by lens $L_1$, where its intensity is sampled by a WRD. This detector has wedge-shaped elements in one-half of the circular aperture and ring-shaped elements in the other half of the circular aperture. Figure 2A shows this detector schematically. In the version of this device that was fabricated and was commercially available, there were 32 wedge and 32 ring shaped detector elements in each half of a one inch diameter silicon sensor (Figure 2B). The 64 detector outputs are available in parallel and are fed through amplifiers to autoranging amplifiers and potentially into a supporting digital processor for analysis purposes. The output of any detector element can be manually selected and viewed on a digital display or the full (or any partial set of 64 detector outputs) can be selected, automatically scanned and fed to a digital processor. Figure 2C shows the standard control unit. From (4), the FT pattern is seen to rotate as the input object rotates. Since the ring shaped detector outputs integrate over $\theta$, the $f(r)$ ring shaped detector output distribution does not change with input object rotations. From (5), the FT pattern is seen to scale inversely with changes in scale of the input object. Since the wedge shaped detector outputs integrate over $r$, the $f(\theta)$ wedge shaped detector output distribution does not change with input object scale changes. From (1), one can place the wedge shaped detector elements in one-half of the FT plane and the ring shaped detector elements of the other half of the FT plane with no loss of information (beyond that which occurs due to intensity sampling).

## 3.  CGH/Holographic Optical Element (HOE) WRD CONCEPT

These concepts were first introduced by Stanley and Lendaris [1] and later exploited by Recognition Systems Incorporated [2]. They find much use as mission screeners in the identification of the class of different parts of an input scene [1], in object quality inspection [2], line width analysis for ICs, handwriting analysis [3], for producing a generalized chord distribution feature space [6], and in more recent work for object identification and classification [4]. Although these diffraction pattern concepts are attractive, there are several shortcomings with the present silicon detector units. These include: the lack of availability of such silicon detectors, the desire to achieve more compact units of smaller physical size and weight, the attractiveness of often wanting a wider variety of detector shapes, the frequent need for more sensitive and higher speed detectors than one can obtain with the wide area units necessary when fabricated in silicon. One can separate the detection function and the specific sampling shape aspect of the detector elements by sensing the FT pattern using a conventional 2-D grid scan pattern and then digitally implementing various desired detector shape functions. The interpolation required to accurately model the detector shape desired is a significant overhead in a digital realization and often precludes real-time operation. Hence, an optical realization using a CGH to achieve the desired sampling function and a linear array of separate high-performance detectors with parallel outputs is preferable. CGHs and holographic optical elements (HOEs) are presently receiving considerable attention [5] with the availability of several commercial CGH recorders. Thus, this optical approach is also of considerable practical and current interest.

The CGH/HOE-based compact system we envision use of is shown in Figure 3A and in block diagram form in Figure 3B. The FT of the input object is formed at $P_2$ where a CGH and HOE are placed. The CGH has different grating patterns in different regions, with each region having a different shape and location (corresponding to the specific detector shapes required). In each region, the grating is of one spatial frequency and one orientation (the spatial frequency and orientation differ in each region) and determines the location in $P_3$ where the data in each $P_2$ region focuses. An HOE recorded on the CGH plate at $P_2$ achieves the focusing of each $P_2$ region to a separate point in $P_3$.

In practice, the CGH/HOE could be reflective and a folded optical system of reduced size would result. The separate wedge and ring outputs (or other FT plane sampling shapes desired) are produced in spatially-separated regions of $P_3$. Detector arrays or discrete detectors placed at $P_3$ provide parallel outputs corresponding to the wedge and ring sampled FT plane data. This separation of the sampling and detection functions allows high-speed and high-sensitivity detectors to be used. These parallel outputs would then be fed to a dedicated digital processor to perform feature extraction (selection, weighting and combining of the different wedge and ring detector outputs as required for a given application) in this wedge/ring-sampled FT feature space and estimation of the class, orientation and scale of the input object. The classification (for out-of-plane distortions) is performed by projecting the wedge/ring-sampled FT feature vector onto a discriminant vector selected by various pattern recognition techniques [4].

Plane $P_2$ need not be an FT plane. If it is an autocorrelation plane, then the wedge/ring features produced are the chord distributions [6-8]. As noted earlier, with a CGH, one is not restricted to wedge and ring sampling, but any desired sampling-shaped function can be used.

(A)



(B)
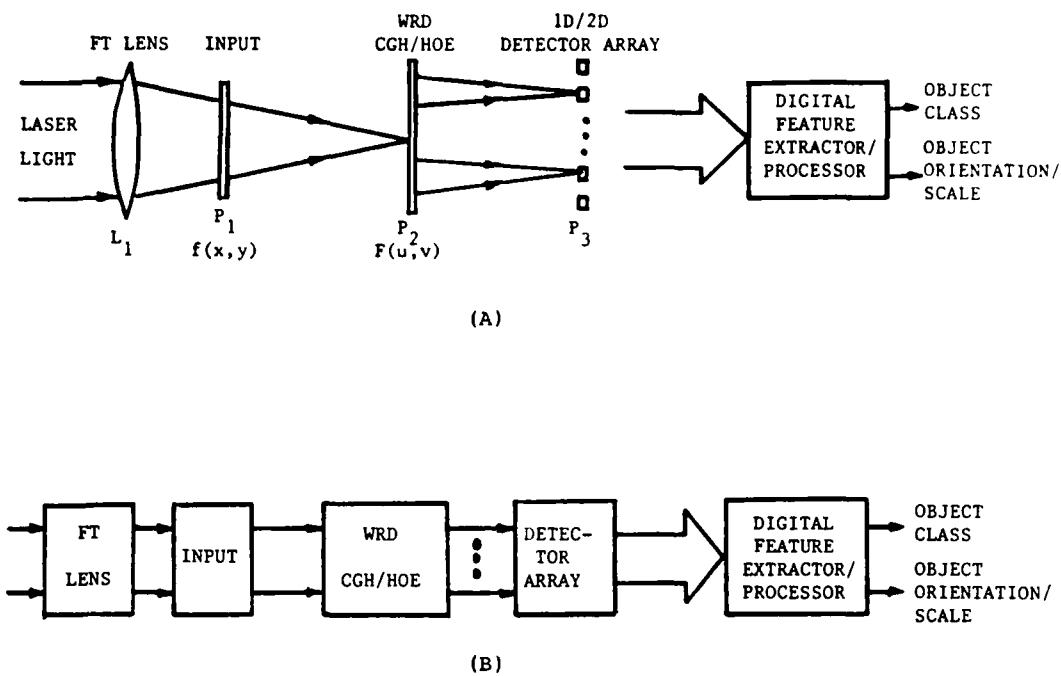
FIGURE 3
Preferred computer generated hologram (CGH)/holographic optical element (HOE)
realization of an optical wedge ring detector (WRD) system

## 4. CGH WRD DESIGN

A schematic and block diagram of the laboratory WRD/CGH system are shown in Figures 4A and 4B respectively. The $P_1$ inputs, $P_2$ CGH and output $P_3$ detector plane coordinates are shown in Figure 4. The CGH at $P_2$ achieves the desired wedge-ring sampling and diffracts all light (incident on each separate wedge and ring sampled $P_2$ region) at a different angle (proportional to the spatial frequency and orientation of the grating present in each $P_2$ region). Lens $L_2$ focuses the parallel light from each wedge-ring $P_2$ region to a different location in $P_3$, where separate high-performance detectors collect this light and provide the desired wedge-ring sampled output data in parallel. This WRD sampling and detection technique using a CGH is preferable to the holographic recording of the necessary pattern in each $P_2$ region as proposed in Ref.[9]. Our proposed CGH technique requires no sophisticated optical system for recording and is thus simpler and cheaper. It allows phase relief CGH recordings to be used and thus has the same high-efficiency advantage of the technique in Ref.[9] when using bleached dichromated gelatin, but with much easier fabrication, with greatly increased flexibility and at a significantly lower cost.

For simplicity, we describe each region of the CGH by a square-wave grating of unit amplitude varying in x only as

$$g(x) = [Rect(\frac{x}{\Delta x}) * Comb(\frac{x}{d})]Rect(\frac{x}{L}), \tag{6}$$

where $\Delta x$ is the width of each bar in the grating, d is the grating spacing, $u_1 = 1/d$ is the grating frequency, L is the grating's extent and $Comb(x/d) = |d|\Sigma\delta(x-nd)$. We could employ a sinewave grating in each $P_2$ plane region. However, a square-wave grating is more easily fabricated using a binary CGH or a binary recorder. Use of a sinewave grating would result in only one diffracted order, a slight increase in useable light and would not require attention to avoiding higher diffracted orders. However, as we show in this paper, the present design with a square-wave grating represents no problem, achieves adequate light budget efficiency, requires lower resolution than a sinewave grating. Primarly the use of a square-wave grating allows simpler binary recording systems to be employed. With g(x) in (6) placed at $P_2$, its FT is formed at $P_3$ in Figure 4 and is

$$G(u) = L \cdot \Delta x \cdot d[\text{Sinc}(u\Delta x)\text{Comb}(ud)] \ast \text{Sinc}(uL), \tag{7}$$

where $u = x_3/\lambda f_{L2}$ relates spatial frequencies u at $P_2$ to distance $x_3$ in $P_3$. Eq.(7) shows that the data from one such 1-D grid produces a $P_2$ pattern containing sinc functions of width $1/L$ replicated every $1/d$ with an overall amplitude weighting across all of the sinc functions given by a sinc function of large width $1/\Delta x$.
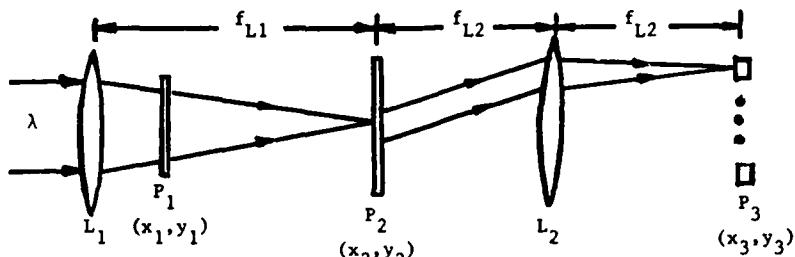


(A)
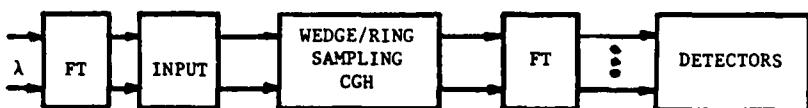


(B)

FIGURE 4

(A) Schematic and (B) block diagram of a laboratory wedge ring detector (WRD) computer generated hologram (CGH) holographic optical element (HOE) system.

The location in $P_3$ of the grating data in the corresponding $P_2$ region is thus

$$x_3 = \lambda f_{L2}/d, \tag{8}$$

where d is the spacing between two gratings square-wave rectangular pulses. The detector plane $P_3$ size, the size of each detector element, and the length of the system $(f_L)$ determine d and the angle $\theta$ for each grating region of $P_2$. We consider a circular CGH of radius R with wedge shaped elements in the upper half and ring shaped elements in the lower half. The highest spatial frequency $u_m$ in the $P_1$ input image determines the radius required for the CGH as

$$R \geq \lambda f_{L1} u_m. \tag{9}$$

We consider two detector formats: a rectangular array (Figure 5A) and two circularly-symmetric detector arrays (Figure 5B). The rectangular detector array offers the use of a simpler commercial detector system with higher CGH requirements. The circular detector arrays require a far simpler CGH but individual detectors in a nonstandard and therefore less commercially available array configuration. Both CGHs have the general form shown in Figure 5C with wedge shaped sampling elements in one half and ring shaped elements in the other half of the plane. Each CGH region contains a grating of spatial frequency $d_{ij}$ and angle $\theta_{ij}$, where the subscripts correspond to the associated detector element. For simplicity, only one wedge and ring grating pattern are shown in Figure 5C.
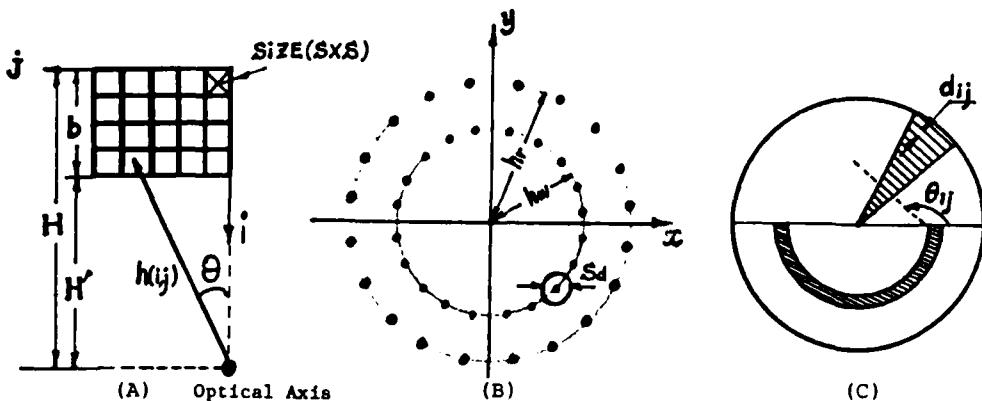
FIGURE 5
Output plane $P_3$ detector geometries (A) rectangular detector,
(B) concentric detector arrays, and (C) basic WRD CGH geometry

## 4.1 RECTANGULAR DETECTOR ARRAY

For the rectangular detector array (with square-wave CGH gratings), we require

$$H \geq 2b \tag{10}$$

to insure that the second-order terms from the CGH do not fall on the detector array and that the position of the detector array in $P_3$ is offset from the optical axis by

$$H' > b \tag{11}$$

for similar reasons. These conditions in (10) and (11) are most easily derived for the vertical detectors $i$ in column $j = 1$. If the grating spacing satisfies

$$d(i,1) = \lambda f_{L2}/h(i,1) = \lambda f_{L2}/[H-(i-1)s] \tag{12}$$

for detector element $(i,1)$ in the first column $(j=1)$, then the second-order is insured to fall outside the detector array for the other detector array elements, where $h(i,j)$ denotes the distance from the origin of $P_3$ associated with detector element $(i,j)$. If the detector size $s$ is fixed, so is $b$ and the minimum grating spacing $d_m$ (maximum spatial frequency $d_m^{-1}$) must satisfy

$$d_m \leq (\lambda f_{L2}/2b)\cos\theta , \tag{13}$$

where the term in parentheses is the grating spacing required for the top right detector element in Fig.5A.

To produce diffracted light focused onto a row of spots at each line in $P_3$, the gratings in each region of $P_2$ must be oriented at an angle $\theta(i,j)$ (Figure 5C) to the horizontal $x$ axis satisfying

$$\theta(i,j) = \text{Arctan}\{(j-1)s/[H-(i-1)s]\}. \tag{14}$$

The grating separation in the $P_2$ region corresponding to detector $(i,j)$ must thus satisfy

$$d(i,j) = \{\lambda f_{L2}/[H-(i-1)s]\}\cos[\theta(i,j)]. \tag{15}$$

Eqs.(14) and (15) define the grating spatial frequency required in each $P_2$ region subject to the grating spacing constraints on $d$ in (12) and (13). This CGH design requires a grating with period $d$ inclined at an angle $\theta$ to be recorded in each $P_2$ region with a different $d$ and $\theta$ for each region (Figure 5C). This requires considerable resolution and accuracy, compared to our concentric detector array system.

## 4.2 CONCENTRIC CIRCULAR DETECTOR ARRAY

For the circular detector array configuration (Figure 5B), the CGH design is far simpler than the case considered in Section 4.1. In this present system, the wedge shaped detector elements lie in the top part of the CGH and the ring shaped elements lie in the bottom portion. The grating spacing $d_W$ is fixed for all wedge regions and only $\theta$ is varied between $P_2$

regions. The grating spacing $d_R$ for all ring regions is also constant and again only the angle $\theta$ of the grating is varied between ring regions. If a Calcomp plotter is used to synthesize this CGH, the end points of each line are specified and the plotter draws the desired line at the necessary angle. In a CGH recorder, the coordinates of each point are generally required and thus sampling effects will be of more concern. This issue is common to both detector array cases, since the grating in each $P_2$ region is at a different angle in both detector cases. The CGH pattern at $P_2$ for the rectangular detector array and the circular detector array are similar as shown in Figure 5C. The first-order diffracted radii $h_W$ and $h_R$ for the wedge and ring gratings of spacing $d_W$ and $d_R$ must satisfy

$$h_W = \lambda f_{L2}/d_W, \quad h_R = \lambda f_{L2}/d_R. \tag{16}$$

We can avoid overlapping of the first and second-orders by selecting

$$2h_W - h_R \geq s_d \quad \text{and} \quad h_R - h_W > s_d, \tag{17}$$

where $s_d$ is the diameter of a detector. Each grating produces + and − diffracted orders. The inner circle of peaks in Figure 5B corresponds to these + and − orders for the wedges and the outer circle corresponds to these for the ring elements. If there are M wedges in the top half of the CGH and M rings in the bottom half, then the bisector for the i-th wedge region is a line

$$y = K(i)x, \quad \text{where } K(i) = \text{Tan}[(\pi/M)(i-0.5)]. \tag{18}$$

The line perpendicular to the bisector is $y = [-1/K(i)]x+C$ and the angle that grating i makes with the +x axis is thus

$$\theta(i) = \arctan \frac{-1}{K(i)}. \tag{19}$$

For simplicity (Figure 5B), the same grating angles are used for both the wedge and ring gratings.

The last design issue we consider is the diffracted spot size $s_2$ on the CGH, its diffracted spot size $s_3'$ on the detector, and the size $s_1$ of an input image region of one uniform spatial frequency, and the size $s_3 = s_d$ of a $P_3$ plane detector. This is a unique issue and requirement for CGH/WRD systems. One spot of diameter $s_2$ at $P_2$ will produce a spot of diameter $s_3' = 2\lambda f_{L2}/s_2$ at the detector plane. This $P_2$ spot diameter is due to a region in $P_1$ of minimum diameter $s_1 = 2\lambda f_{L1}/s_2$. For simplicity, 1.22 factors have been omitted in the above spot size equations. The detector size $s_d = s_3$ must thus satisfy

$$2\lambda f_{L2}/s_2 \leq s_3 \leq \lambda f_{L2}[(1/d_R)-(1/d_W)] \tag{20a}$$

$$s_3 \leq 2\pi h_W/M \tag{20b}$$

where the left side of (20a) insures $s_3 > s_3'$ to collect all diffracted light from $P_2$, the right side of (20a) insures that the wedge and ring detectors do not overlap and (20b) insures that the wedge detectors themselves (lying at a radius $h_W$) do not overlap. We will quantify these values for our experimental system shortly.

## 5. EXPERIMENTAL RESULTS

The experimental results for two WRD CGHs with M = 10 wedge elements and 10 ring elements follows.

### 5.1 RECTANGULAR DETECTOR ARRAY

For this case, a 4 x 5 detector array (Figure 5A) is used with a = 7.5 mm and b = 6.0 mm with s = 1.5 mm detectors on 1.5 mm centers. For the experiment performed, $f_{L2}$ = 815 mm and $\lambda$ = 0.465 μm (an argon laser line). To satisfy H > 2b in (10), we selected H = 19 mm for a corresponding minimum grating spacing $d_m$=0.0186 mm from (13). The grating spacings and angular orientation for each region were selected from (14) and (15). The grating spacing varied from 0.020 to 0.026 mm for the detectors in the first column (J=1) and the grating angle for the detectors in the j=2 column varied from 4.38° to 5.73°. Over the entire detector array, d varied from 0.0186 to 0.0260 mm and $\theta$ varied up to 30°. The general design guidelines for the grating spacings and grating angles for the CGH region corresponding to detector (i,j) = (vertical,horizontal) satisfy (14) and (15) where i = 1 and 2 for wedges and i = 3 and 4 for ring elements. The end points of each grating line were specified and a

line drawn between them using our Calcomp plotter. The full plot was 10" = 254 mm in diameter with the smallest grating interval being 0.635 mm (the Calcomp plotter easily produced lines with spacings of 0.02" = 0.5 mm, i.e. well within our 0.635 mm requirements). This plot was photoreduced by 32.5:1 to 7.8 mm diameter with $d_m$ = 0.02 mm. The 2R = 7.8 mm diameter allows a maximum $P_1$ input spatial frequency $u_m$ = 21 cy/mm (assuming $\lambda$ = 465 μm and $f_{L1}$ = 400 mm), which is more than adequate for realistic imagery.
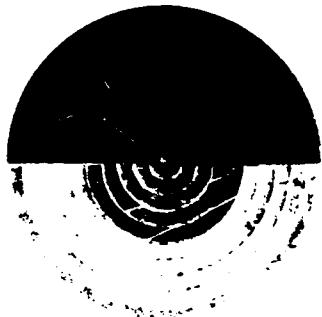


(A)
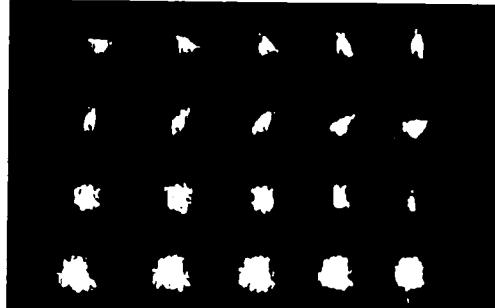


(B)

FIGURE 6
(A) Computer generated hologram and (B) output plane pattern for a wedge ring detector computer generated hologram with an output rectangular detector array

Figure 6A shows the CGH used and Figure 6B shows the 2-D rectangular $P_3$ output diffraction pattern obtained when this CGH is illuminated with a plane wave. The upper two rows correspond to the ten wedge outputs and the bottom two rows to the ten ring outputs. The wedges and the inner rings in Figure 6A are not as easily visible because of their higher grating spatial frequencies. The number of grating lines in the wedge regions varied from 52 to 167 and the number of lines in the ring regions varied from 17 to 227. This accounts for the different intensity (larger spot sizes) in Figure 6B. The spots diffracted by the wedge regions are more uniform. Because the area and number of lines in each ring region varies, their light intensity varies more as is seen. The locations of the diffracted output peaks are in agreement with theory within measurement accuracy.

## 5.2  CONCENTRIC DETECTOR ARRAY

The CGH for this case (Figure 7A) and the resultant $P_3$ diffracted pattern (Figure 7B) again agree with theory. The same $\lambda$ and $f_{L2}$ are used. The original CGH produced by the Calcomp plotter for this case was 8" = 203.6 mm in diameter with the same $d_W$ = 0.042" = 1.07 mm for all wedges and $d_R$ = 0.76 mm for all rings. After photoreducing by 20.7:1, the CGH had: 2R = 9.8 mm, $d_W$ = 0.052 mm and $d_R$ = 0.037 mm. The 2R value allows $u_m$ = 26 cy/mm input spatial frequencies (assuming $f_{L1}$ = 400 mm). These $d_W$ and $d_R$ choices satisfy (16) and (17) with the detector size used $s_d$ = 1.5 mm (since $h_R$ = 5.4 mm and $h_W$ = 3.8 mm). The i-th grating angle is $\theta(i)$ = (i-0.5)18 from (19). These values also satisfy (20) for $P_2$. FT plane spots above $s_2$ = 0.5 mm, corresponding to uniform input spatial frequency regions as small as $\lambda f_{L1}/s_2$ = 0.4 mm for $f_{L1}$ = 400 mm. The light intensity in the two sets of concentric diffracted peaks (Figure 7B) vary (for the outer set of ring elements) due to the area of the rings and the varying number of grating lines per ring (27 to 266 lines).

## 6.  ADVANCED TOPICS

The resolution (0.006" = 0.15 mm) for the Calcomp plotter used in these experiments determines the system's size and the number of wedge and ring elements used. Commercially available recorders with 1 μm resolution [5] allow fabrication of a system of significantly reduced size. For the system of Figure 4 with $\lambda$ = 820 μm (a laser diode source), a 14 x 14 mm detector array and a maximum input spatial frequency of 20 cy/mm, both lenses can have $f_L$ = 4" = 100 mm using a $d_m$ = 2 μm minimum spot size recorder. This represents a considerable reduction in system size.
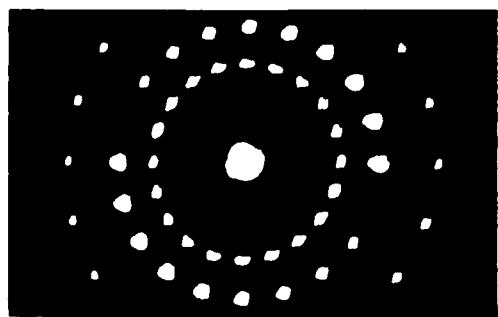
(A)  (B)

FIGURE 7

(A) Computer generated hologram and (B) output plane pattern for a wedge ring detector computer generated hologram with concentric output detector arrays

The light budget for this system is excellent. Assuming lens transmittances of 0.9, a $P_1$ transmittance of 0.5, a transmittance of 0.5/20 for each of the 20 wedge and ring elements in $P_2$ (where a 50% efficiency for a bleached or phase CGH is assumed), a $P_3$ transmittance of 0.5 (each detector intercepts approximately 0.5 of the total light in the sinc function, and 10 different input spatial frequencies, and hence a division of the total light into ten separate regions), then the system's transmittance to one output detector is $0.1(0.9)^2(0.5)$ $(0.1)(0.9)^2(0.5)^3(0.05) = 5 \times 10^{-4} = 0.05\%$. For a typical detector with 0.5 amp/watt sensitivity and 0.3 nA dark current, an input light intensity of $6 \times 10^{-10}$ W corresponds to the dark current and the maximum input light is $6 \times 10^{-4}$ W (for a 6 decade response detector). If the photodiode is biased at 300 nA (much much greater than the dark current), the minimum detector power required is 0.6 μW and hence we require only $0.6 \text{ μW}/(5 \times 10^{-4}) = 1.2$ mW of input light. This is easily achieved by laser diode sources.

A final topic is the converging nature of the light input to $P_2$ in Figures 3 or 4. In a converging beam FT system as is used, the FT is formed on a spherical surface not in a plane [10]. The displacement error to the plane where the CGH is placed is $\Delta z = (x_2^2+y_2^2)/2d =$ $(x_2^2+y_2^2)/f_{L1}$ is only a maximum of 0.25 mm. Thus, the FT spot size is only slightly larger than the theoretical value and the pattern detected at $P_3$ is correct. Phase curvature at $P_3$ is of no concern since the size of the detectors are used (in our design). If (as occurs in practice) only a small part of a wedge or ring is illuminated at $P_2$, the $\Delta z$ effect is of no concern since all of the light still easily falls within a wedge or ring region.

We have concentrated on the use of a WRD CGH in the FT plane. However, as noted earlier, a WRD can also be used in the autocorrelation plane to produce chord distribution functions. As noted earlier, one is not restricted to wedge and ring shaped detector elements but can employ other detector shapes as required. This is attractive both for FT plane sampling and for autocorrelation plane analysis. The use of CGHs clearly allows considerable flexibility in the detection process and it allows separation of the detector shape function from the detection function, thereby allowing more optimized components to be used. In this paper, the general concepts of the CGH detector have been advanced, general calculations and design rules have been advanced and laboratory demonstrations and designs of two different WRD CGHs have been provided.

## ACKNOWLEDGMENT

## REFERENCES

1. G.G. Lendaris and G.L. Stanley, Proc. IEEE, Vol. 58, p. 198 (Feb. 1979).
2. H. Kasden and D. Mead, Proc. Elec. Opt. Sys. Des., Vol. 248 (1976).
3. D. Casasent, "Optical Signal Processing" Chapter 8 in Optical Data Processing, Vol. 23 of Topics in Applied Physics, D. Casasent, Editor, Springer-Verlag (1978).
4. D. Casasent and V. Sharma, "Feature Extractors for Distortion-Invariant Robot Vision", Optical Engineering, Vol. 23, pp. 492-498 (Oct. 1984).
5. "International Conference on Computer Generated Holograms", Proc. Soc. Photo. Opt. Instr. Engrs., Vol. 437, S. Lee, Editor (Aug. 1983).
6. D. Casasent and W.T. Chang, "Generalized Chord Transformation for Distortion-Invariant Optical Pattern Recognition", Applied Optics, Vol. 22, pp. 2087-2094 (July 1983).
7. D. Casasent, "Recent Review of Holography in Coherent Optical Pattern Recognition", Proc. Soc. Photo. Opt. Instr. Engrs., Vol. 532 (January 1985).
8. D. Casasent, "Coherent Optical Pattern Recognition: A Review", Optical Engineering, Vol. 24 (Jan. 1985).
9. M.S. Brown, Optica Acta, Vol. 31, pp. 507-513 (1984).
10. M. Francon, Optique (1972).

# 6. MULTIPLE FEATURE EXTRACTORS AND CLASSIFIERS: AN OPTICAL WEDGE RING DETECTED FOURIER TRANSFORM SPACE CASE STUDY

Ref. 9

# Feature extractors for distortion-invariant robot vision

**David Casasent**
**Vinod Sharma***
Carnegie-Mellon University
Department of Electrical and
    Computer Engineering
Pittsburgh, Pennsylvania 15213

**Abstract.** Various feature extractors/classifiers for a hierarchical feature-space pattern recognition system are described. The system is intended to achieve multiclass distortion-invariant object identification. Although only a Fourier transform feature space is used, our basic hierarchical concepts, our theoretical analysis, and our general conclusions are applicable to other feature spaces. The performance using intensity and phase Fourier transform features and the performance in the presence of noise are studied and quantified for two different two-class pattern recognition data bases.

## CONTENTS

## 1. INTRODUCTION

Distortion-invariant multiclass pattern recognition is considered using a Fourier transform (FT) feature space. Feature extraction, dimensionality reduction, discrimination, and classification are addressed. A simplified block diagram of our hierarchical pattern recognition system is shown in Fig. 1. We begin with a Fourier transform feature space, since such a representation is well known[1] to allow significant data compression. We extract the magnitude, phase, or both from the Fourier transform plane. As the first dimension-ality-reduction technique, we use a wedge-ring detector (WRD) to sample the Fourier transform plane data[2,3] to reduce the dimension-ality of the feature space and retain only the dominant eigenvector for each object class. This reduced subspace is calculated using a Karhunen-Loeve (KL) transformation[4] by new efficient techniques.[5] This completes the dimensionality-reduction step in our system. To

provide discrimination, we employ two nonunitary transformations: the Fukunaga-Koontz (FK)[6] and the Foley-Sammon (FS)[7] trans-formations. Our classifier then selects the best subspace from the KL, FK, and FS feature vectors.

In Sec. 2, we review and highlight our two levels of dimensionality reduction (*WRD Fourier transform sampling and dominant eigenimage calculation*). We then discuss (Sec. 3) how we achieve distortion invariance, and we detail the discrimination algorithms used. Brief theoretical remarks on the use of Fourier transform plane phase or magnitude features and on the noise performance of a feature extractor then are advanced in Sec. 4. The two image data bases used in our experiments and the results of our initial dominant eigenimage feature vector calculations are summarized in Sec. 5. More extensive distortion-invariant image test results are then pre-sented and discussed in Sec. 6. These results include a comparison of the performance of our system for five different discrimination vec-tors, comparison of the performance of amplitude-only and phase-only Fourier transform features, and a comparison of the classifiers and feature extractors in the presence of noise. Our summary and conclusions then are advanced in Sec. 7.

## 2. DIMENSIONALITY REDUCTION AND DISTORTION INVARIANCE

If the input image or object is $256 \times 256$ pixels, its dimensionality is $n = 256^2$. The discrete Fourier transform plane for such an object still has a dimensionality of n. This is quite prohibitive for subse-quent feature extraction, matrix transformations, or other similar operations. Thus, dimensionality-reduction techniques are essential operations that must be applied to such a feature space. A Fourier transform feature space is a most useful representation of structural, resolution, and orientation information on the input object. Such a feature space is also attractive since physical insight about the input object is easily obtained from this feature space. Such a feature space is well known[1] to lend itself easily to dimensionality reduction. These reasons, plus the ease with which such a feature space can be pro-duced optically (using a simple spherical lens) or digitally (by various fast Fourier transform hardware and algorithms), make this an ideal choice for our hierarchical feature-extraction studies.
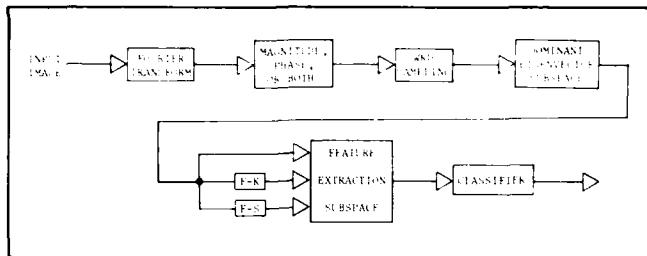
Fig. 1. General Fourier transform (etc.) feature-extraction pattern recognition system block diagram.

As the first level of dimensionality reduction, we sample the Fourier transform plane with a WRD. If an optical system is used to produce the Fourier transform, a commercial WRD device exists.[1] This unit consists of 32 wedge-shaped detector elements in one-half of a circular detector and 32 annular-shaped detector elements in the other half of the detector plane. This device thus provides 64 WRD outputs and hence reduces the dimensionality of the Fourier transform feature space from $n = 256^2$ to 64. One also can digitally model such a device, of course. The ring detector elements provide rotation invariance, whereas the wedge detector elements provide scale invariance (if the values of the wedge-ring detector element readings are properly normalized for object energy).[2,3] To see this, we first recall that the magnitude of the Fourier transform is shift-invariant. Next, we note that as the scale of the input object changes, the two-dimensional Fourier transform distribution changes radially (inversely with the scale change of the input object). Thus, the outputs of the wedge-shaped detector elements will remain invariant to such input object scale changes. Finally, we recall that the orientation of the two-dimensional Fourier pattern rotates as the input object rotates. Thus, the ring-shaped Fourier plane sampling elements have outputs that remain invariant to in-plane rotations of the input object. These remarks follow for the case of a real and positive input function, whose Fourier transform is symmetric. This situation applies for the case of images, and thus the two halves of the Fourier plane can be separately sampled as described with no information loss.

This WRD sampling, plus the training of our system on different distorted images, provides a distortion-invariant pattern recognition algorithm. In this and similar feature space approaches to pattern recognition, one uses $N_1$ images in class 1 and $N_2$ images in class 2 to determine the parameters of the processor. These are referred to as a training set of images. Each of the i images per class is denoted by a vector, with $\{x_i\}$ and $\{y_i\}$ being the set of $i = N_1$ or $N_2$ training set image vectors for a two-class example. The corresponding two-dimensional Fourier transforms are the n-dimensional vector sets $\{x_i''\}$ and $\{y_i''\}$. These are dimensionality-reduced to the WRD-sampled 64-dimensional vector sets $\{x_i'\}$ and $\{y_i'\}$. As the second dimensionality-reduction step, we apply a KL transformation[4] to the autocorrelation matrix formed from the WRD feature vectors for each separate object class. The autocorrelation matrix is formed from the 64 element $x_i'$ vectors for each of the training set images $\{x\}$ in class 1, and a second matrix is formed from the corresponding $y_i'$ vectors of images in class 2. The eigenvalues and eigenvectors of each matrix are calculated and tabulated. One can then retain the dominant $\eta_x$ and $\eta_y$ eigenvectors per class, where $\eta_x$ and $\eta_y$ are typically less than 4. In our experiments, we retained only the dominant eigenvector for classes 1 and 2, which we denote by KL-1 and KL-2. In practice, two or three eigenvectors would be used per class.

## 3. NONUNITARY TRANSFORMATIONS

To use these dominant eigenvectors defined in Sec. 2 for classification, we compute $z'$ for an unknown input object vector $z$, project it onto the eigenvectors KL-1 and KL-2 (for classes 1 and 2, respectively), and select the class for the unknown input based upon which

projection value is larger. The KL, or dominant, eigenvector transformation in Sec. 2 represents a considerable compression of data and simplifies performing the nonunitary transformations discussed below. The dominant eigenvectors represent each class well in an optimal compressed manner. However, there is no assurance that those features that represent each class well will be optimal for discriminating one class from another. Thus, dominant eigenvectors are useful for intraclass pattern recognition (that is, recognizing different versions, i.e., geometrically distorted views of one object), but not necessarily for interclass discrimination (distinguishing one object class from another). In a hyperspace description of a feature vector and a discriminant vector, unitary transformations do not change the distances between vectors. To achieve discrimination or interclass pattern recognition, linear nonunitary transformations represent an attractive approach. These transformations can change interclass distances and hence provide improved discrimination. We pursued this approach rather than utilizing additional eigenvectors per object class. This choice is logical since the use of more eigenvectors would only further increase the dimensionality and computational complexity of the processor. In the next two subsections, we detail two nonunitary transformations that we have employed.

### 3.1. Fukunaga-Koontz transformation

The first nonunitary transformation we consider is the Fukunaga-Koontz (FK) transformation.[6] To describe the steps in this algorithm, we first define $P_i$ as the *a priori* probability for class i and $R_i'$ as the autocorrelation matrix for class i. We form the autocorrelation matrices $R_1$ and $R_2$ for each class, where $R_i = P_i R_i'$, and we form the full autocorrelation matrix $R = R_1 + R_2$. We then determine the transformation matrix $T$ that diagonalizes $R$; i.e.,

$$T R T^T = T(R_1 + R_2)T^T = I ,  \qquad (1)$$

where $I$ is the identity matrix. By this transformation we have orthogonally decomposed the full $R_1 + R_2$ matrix. Next, we apply $T$ to $R_1$ and $R_2$; i.e., we form new matrices for each class given by $T R_1 T^T$ and $T R_2 T^T$.

These new correlation matrices have two attractive features: (a) The eigenvectors $\psi_i^{(1)}$ and $\psi_i^{(2)}$ of $T R_1 T^T$ and $T R_2 T^T$ are the same. (b) The eigenvalues $\lambda_i^{(1)}$ and $\lambda_i^{(2)}$ associated with $\psi_i^{(1)}$ and $\psi_i^{(2)}$ are related by

$$\lambda_i^{(1)} = 1 - \lambda_i^{(2)} .  \qquad (2)$$

From Eq. (2), we see that the dominant eigenvectors of the transformed class 1 matrix are the least-dominant eigenvectors for the transformed class 2 matrix. Thus, those eigenvectors that represent class 1 the least represent class 2 the best (in the new FK transformed feature space). Thus, this transformation has converted the input data into a new space with a common set of basis functions (the $\psi_i$). In this new space, the data in the two classes are now separated. In our two-class problem, we will select two $\psi_i$ with the largest $|\lambda_i^{(1)} - 0.5|$ values.

Since $R$ is formed from the KL vectors (Sec. 2) and since we only retain one KL eigenvector per class, the rank of $R$ is two and there are only two eigenvectors $\psi_i$. We denote these two eigenvectors of the FK transformed data by FK-1 and FK-2. FK-1 and FK-2 are the two vectors that best discriminate class 1 objects from class 2 objects. To use these new discriminant vectors to determine the class of an unknown input image $z$, we form the WRD vector $z'$ and transform it to a new $T z' = z'''$. This transforms the input data to the new FK space. We then project $z'''$ onto an FK discriminant vector $\psi$ by calculating $\psi^T z''' = d$. Depending upon whether d is above or below a threshold, we select class 1 or class 2 for the class of the input object. We normalize the FK vectors and refer to the projections onto the FK directions 1 and 2 (corresponding to FK-1 and FK-2). We note that FK-1 and FK-2 do not refer to discriminant vectors for classes 1 and 2; rather, they refer to the two most dominant eigenvectors of the transformed full autocorrelation matrix of *both* classes.

## 3.2. Foley-Sammon transformation

In the Foley-Sammon (FS) nonunitary transformation,[7] we find a linear discriminant vector **w**, selected to maximize the Fisher ratio[8]:

$$F(\mathbf{w}) = \frac{(\text{difference of means of projections})^2}{\text{sum of variances of projections}} \cdot \qquad (3)$$

In terms of the means $m_1$ and $m_2$ of the projections for class 1 and class 2 training set objects onto **w** and the scatter $s_1^2$ and $s_2^2$ of these projections, we can write

$$F(\mathbf{w}) = \frac{|m_1 - m_2|^2}{s_1^2 + s_2^2} = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \cdot \qquad (4)$$

where $\mathbf{S}_B$ is the between-class scatter matrix and $\mathbf{S}_W$ is the within-class scatter matrix.[8] The solution for **w** that maximizes Eq. (4) is

$$\mathbf{w} = \mathbf{S}_W^{-1}(\mathbf{m}_1 - \mathbf{m}_2) \cdot \qquad (5)$$

where $\mathbf{m}_1$ and $\mathbf{m}_2$ are the vector means of the two classes. To use **w** for an unknown input **z'**, we form $\mathbf{w}^T\mathbf{z}' = d$ and compare the projection value to the threshold T, where

$$T = \frac{(m_1 + m_2)}{2} \cdot \qquad (6)$$

If $d > T$, we select class 1. If $d < T$, we select class 2 for the class of the unknown input image vector **z**.

## 4. INTENSITY-ONLY OR PHASE-ONLY FOURIER TRANSFORM FEATURES

An attractive aspect of a Fourier transform feature space is the fact that its magnitude or phase or both can be used. Considerable work[9,10] exists on the representation of image data by the intensity or phase of the Fourier transform. In general, the conditions under which the Fourier transform phase features are adequate are less restrictive than the conditions under which the Fourier transform magnitude features are adequate. The magnitude of the Fourier transform is adequate if the z-transform does not contain reciprocal pole-zero pairs, poles outside the unit circle, or zeros inside the unit circle.

This prior work has been concerned with aesthetically pleasing image reconstructions from the magnitude or phase of the Fourier transform. However, our present concern is object recognition, not image reconstruction. Little research exists on this topic. In our case studies, we will wedge-ring detect and KL transform the Fourier transform magnitude or phase data (or a combination of both). We will then quantify the pattern recognition performance of magnitude or phase features and their performance in the presence of noise. The Fourier magnitude data are shift-invariant, and thus the location of the object in the input field of view cannot be determined from such data. Conversely, the linear components of the Fourier phase provide data on the location of the input object. Digitally, the computational

complexities in extracting the magnitude or the phase of the Fourier transform are more comparable. Optically, the Fourier magnitude is easily obtained, whereas its phase requires the use of a more complicated heterodyne detection technique.

## 5. DATA BASES

The four image data bases used are summarized in Table I. They include scaled and rotated images of the letters A and B and of hand-drawn images of tanks and trucks. For each of these two object classes, we used a set of five images per class and a set of 25 images per class. Various scaled and rotated views were included in each of these image sets. A scale value of 1.0 is unity scale, and 0.9 corresponds to a 10% scale difference, etc. The specific distorted object views included in each case are detailed in Table I. All images have 16 gray levels, with the 1.0 nominally scaled images having various numbers of pixels: A (584 pixels), B (375 pixels), tank (797 pixels), and truck (292 pixels). For our noise-free tests, these images were present on a zero-valued background. For our noise tests, zero-mean white Gaussian noise was added to all pixels in all images. In our data, we list the standard deviation $\sigma_n$ of the noise. From $\sigma_n$, the total number of pixels N in the image, and the object energy E (the sum of the squares of the pixel values for the object), an input signal-to-noise ratio $SNR_1 = E/N\sigma_n^2$ can be defined. For $N = 10^4$, $E = 10^4$ (400 pixels of average value 5), and $\sigma_n = 0.4$, a small $SNR_1 = 6.25$ results.

**TABLE I. Summary of Experimental Image Data Bases Used**

| Test sets | 5-Image data base | | 25-Image data base | |
|---|---|---|---|---|
| | Scales | Rotations | Scales | Rotations |
| A and B | 0.9, 1.0, 1.1 | 0°, 10° (for 0.9 and 1.1 scales) | 0.8, 0.9, 1.0, 1.1, 1.2 | ±10°, ±5°, 0° (for each scale) |
| Hand-drawn truck and tank | 0.9, 1.0, 1.1 | 0°, 10° (for 0.9 and 1.1 scales) | 0.8, 0.9, 1.0, 1.1, 1.2 | ±10°, ±5°, 0° (for each scale) |

In Table II, we list the five nonzero eigenvalues for the five-image data base for all four object types and for both magnitude and phase Fourier transform features. As seen, the eigenvalue for the dominant eigenvector for magnitude Fourier transform features is approximately 70 times the second dominant (in general). This is more pronounced for the letters A and B. The eigenvalue for the dominant eigenvector for the letter A obtained from Fourier transform phase data is low (0.67). Because of the lower (0.67) eigenvalue, we may expect lower projection values and hence more errors in our pattern recognition of letters using phase features. In general, the dominance of one eigenimage in the magnitude data may be attributed to the fact that the image data base consists of scaled and rotated (in-plane rotation) images rather than different aspect views of each object. In such distorted images, there is no appreciable new information present in each object representation. A possible reason for the lower dominant eigenvalues for phase features may be the reduced accuracy

**TABLE II. Eigenvalues of the Five Nonzero Eigenvectors of the WRD Fourier Transform Data for the Five-Image Data Base**

| No | Truck | | Tank | | A | | B | |
|---|---|---|---|---|---|---|---|---|
| | Magnitude | Phase | Magnitude | Phase | Magnitude | Phase | Magnitude | Phase |
| 1 | 0.98 | 0.99 | 0.98 | 0.89 | 0.99 | 0.67 | 0.99 | 0.95 |
| 2 | $0.17 \times 10^1$ | $0.78 \times 10^2$ | $0.17 \times 10^1$ | $0.98 \times 10^1$ | $0.71 \times 10^2$ | 0.24 | $0.13 \times 10^1$ | $0.43 \times 10^1$ |
| 3 | $0.82 \times 10^4$ | $0.28 \times 10^3$ | $0.21 \times 10^3$ | $0.12 \times 10^1$ | $0.84 \times 10^4$ | $0.72 \times 10^1$ | $0.49 \times 10^3$ | $0.19 \times 10^2$ |
| 4 | $0.81 \times 10^5$ | $0.12 \times 10^3$ | $0.64 \times 10^4$ | $0.24 \times 10^2$ | $0.47 \times 10^4$ | $0.11 \times 10^1$ | $0.29 \times 10^4$ | $0.77 \times 10^3$ |
| 5 | $0.49 \times 10^6$ | $0.11 \times 10^4$ | $0.11 \times 10^5$ | $0.14 \times 10^2$ | $0.65 \times 10^5$ | $0.38 \times 10^2$ | $0.17 \times 10^4$ | $0.70 \times 10^4$ |
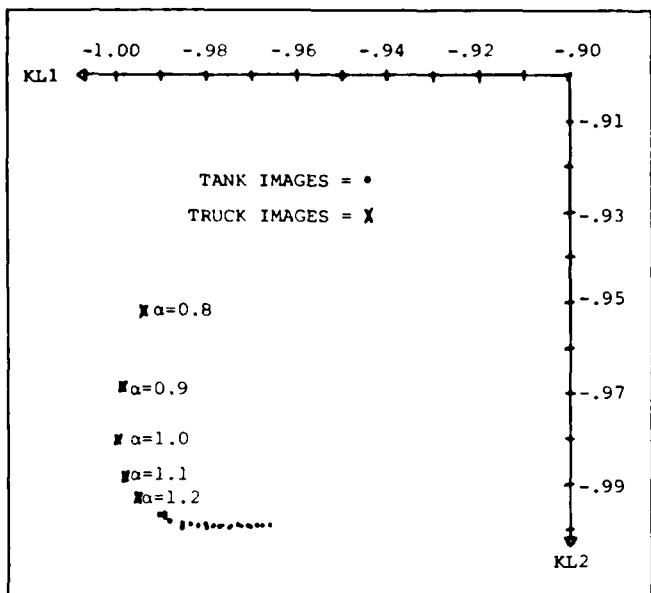
Fig. 2. Magnitude-only WRD Fourier transform features projected onto dominant tank/truck eigenvectors (for 25-image data base). KL1 = dominant truck eigenvector; KL2 = dominant tank eigenvector.



Fig. 3. Magnitude-only WRD Fourier transform feature projections for tank/truck images onto the two FK vectors (25 images/class).

associated with the nonlinear arctangent operation required to compute the phase of the Fourier transform.

The eigenvalue data for the 25-image data base showed comparable results to those in Table II. As noted in Sec. 2, we retained only the dominant eigenvector per class for magnitude-only and phase-only data. For magnitude features, we expect the second dominant eigenvector to provide poor discrimination (this was found to be the case from experiments). For the tank and truck data, phase features may be expected to perform comparably and possibly better than magnitude features. In experiments, phase features (for the tank and truck data) using one dominant eigenvector per class consistently gave larger projection ratios than magnitude features. For the letters A and B, phase features performed poorly (as expected, since the dominant eigenvalue is smaller). Including the second dominant eigenvector for the phase features for our letter recognition tests would be expected to improve performance. However, we included only the most dominant eigenimage per class. Our extensive test results obtained with the 25-image data set are detailed in Sec. 6. They follow the trends noted above, which are expected from the data in Table II.

## 6. INITIAL EXPERIMENTAL RESULTS

### 6.1. KL transformations

All of the results included in this section were obtained on our more extensive data base of 25 object images per class. In Fig. 2, we show the scatter plots for the projections of all tank and truck images onto the dominant eigenvector for tanks (KL-1) and for trucks (KL-2). As seen, all images can be separated and correctly classified from either projection alone. However, all projection values (even those on the dominant eigenvector of the other class) are quite large (all projection values are above 0.95). This might be expected since the KL eigenvectors are useful only for intraclass recognition, not interclass discrimination. Figure 2 shows that the projections of the truck images on the dominant truck eigenvector KL-1 yield essentially invariant values ($\cong$0.993). The tank images projected onto the tank eigenvector KL-2 show a similar invariance with all projection values $\cong$0.995. This intraclass invariance is expected (because of the dominance of the first eigenimage in each class) by the nature of the KL transform. From Figure 2, we can also assess the interclass discrimination of dominant KL eigenvectors. The truck projections on KL-1
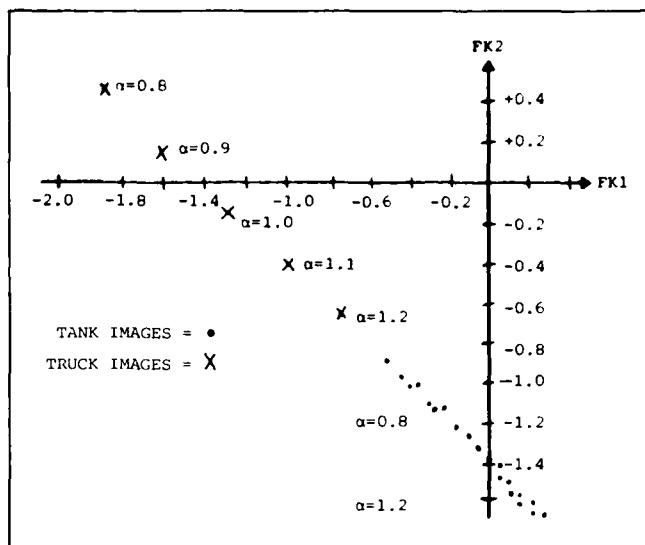
yield lower (0.97 to 0.99) projections (versus 0.995 for projection on KL-2). The truck images show similarly lower projections on KL-2 compared to KL-1. However, the large range for all projections (all are above 0.95) makes the performance of this system in noise suspect. These results thus verify the intraclass recognition ability of the KL transform. If the two classes are sufficiently different, interclass discrimination will be good, but the KL algorithm does not guarantee this. For the case shown, either eigenvector alone is sufficient for discrimination. However, this is not a general conclusion.

An interesting trend from the data of Fig. 2 is that only five points exist for the 25 truck images. These correspond to the five different input image scales (denoted by the values of the parameter $\alpha$, as shown) with all five rotated views per scale giving the same projection. This occurs since rotated images at the same scale have the same energy, whereas different scaled images have different energy. With a different normalization of the image data base, the projection values for different object scales could be made to coincide. This effect is most pronounced for the truck images since they all contain significantly fewer pixels than any scaled tank image used and KL-2 is normalized for the tank images alone.

Similar results were obtained for the projections of the images of the letters A and B onto their dominant eigenvectors (for magnitude-only Fourier data). These results did not exhibit as pronounced a variation with the scale of the input image (since both letters contain a comparable number of pixels). The data still exhibited the five clusters of projection values (one cluster per scale, with only small variations due to rotation) for the reasons advanced above. All projection values for the letters were quite large and even more clustered than in the tank data (all letter projections were above 0.998). More advanced techniques are clearly warranted, and thus we next experimentally considered our nonunitary transformations.

### 6.2. Nonunitary transformations

The projections of the truck-tank data base images on the FK-1 and FK-2 feature vectors are shown in Fig. 3. Comparing these results with the corresponding projection data on the KL-1 and KL-2 eigenimages (Fig. 2), we see that the FK-1 and FK-2 feature vectors separate these two image classes much more than do the KL-1 and KL-2 eigenimages. This verifies our remark that the FK feature vector direction that represents one class best, represents the second class worst, and that FK transformations are preferable for discrimination, whereas KL or dominant eigenvector projections provide

TABLE III. Comparison of Separability Measure S for Magnitude and Phase Features for Different Case Studies and Different Feature Extractors

| FT data images | Magnitude only | | Phase only | | Magnitude and phase | |
|---|---|---|---|---|---|---|
| | Truck − Tank | A − B | Truck − Tank | A − B | Truck − Tank | A − B |
| S for KL-1 | 4.130 | 7.087 | 5.681 | 1.326 | 7.147 | 0.254 |
| S for KL-2 | 2.898 | 5.984 | 4.596 | 8.419 | 8.781 | 5.535 |
| S for FK-1 | 3.908 | 12.135 | 5.450 | 0.201 | 6.285 | 0.226 |
| S for FK-2 | 3.879 | 12.131 | 4.253 | 9.371 | 7.765 | 12.880 |
| S for FS | 4.504 | 11.898 | 7.578 | 9.428 | 8.541 | 12.620 |
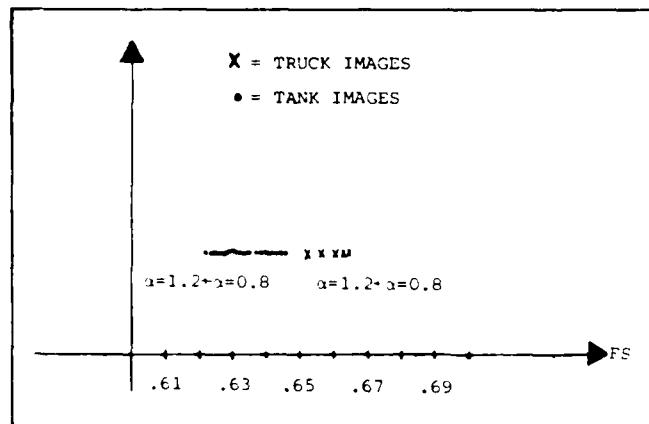


Fig. 4. Magnitude-only WRD Fourier transform feature projections for tank/truck images on the best Foley-Sammon (FS) vector (25 images/class). For clarity of presentation, the projection values are shown displaced from the FS axis.

only intraclass recognition. Now we notice a variation in the projection value along both the FK-1 and FK-2 axes (in Fig. 2, only variations in the projections on the dominant KL eigenvector for the opposite class were observed). Variations in both axes occur here since each FK feature vector is a linear combination of the KL-1 and KL-2 eigenvectors. Figure 4 shows the projection values obtained by projecting the truck-tank image data base onto the FS feature vector. The projection values now appear to be separated more than those for the KL projections, but less than for the FK projections. A quantitative performance measure for comparing these different feature extractors is now advanced.

### 6.3. Performance measure

From Fig. 3 (compared to Fig. 2), the difference in the expected values of the projections of the two classes of data onto the FK feature vectors is larger than for the projections onto the dominant KL eigenvectors per class. However, the variance is also larger in the FK projection case. The same general conclusions also hold for the A-B image recognition data. The scatter plots in Figs. 2 to 4 are useful for visually conveying results. However, they are misleading since they bias one to favor a feature extractor that yields larger differences in the mean values for the projections of different data classes.

To more properly compare different feature extractors, the actual projection values (note the different scales in Figs. 2 to 4) and the variances of the projection values within each class must both be considered. To achieve this and to quantify the performance of our various feature extractors, we use the separation measure

$$S = \frac{\text{difference of means of projections per class}}{\text{average standard deviation per class}} . \tag{7}$$

The denominator in Eq. (7) is $(\sigma_1 + \sigma_2)/2$, where $\sigma_1$ and $\sigma_2$ are the

standard deviations of the class 1 and class 2 projections. This performance measure in Eq. (7) is valid if $\sigma_1$ and $\sigma_2$ are of the same order. For our data, this was found to be generally true. The measure S in Eq. (7) was chosen for its computational ease and because it does not match the measure that any of our feature extractors optimizes. We computed S for all five feature extractor vector subsets (KL-1, KL-2, FK-1, FK-2, and FS) for both magnitude and phase Fourier transform data (and combined magnitude and phase Fourier transform data) for both image data bases (vehicles and letters). The results are shown in Table III and discussed below.

### 6.4. Noise-free performance comparison

A larger S value in Table III indicates better performance. This table includes the S performance measures calculated using magnitude-only, phase-only, and combined magnitude and phase Fourier transform features.

Let us now discuss Table III. The phase features for both image pairs give larger S values than do magnitude features. However, in several cases they perform much worse (for the letters A and B). This occurs because the dominant KL eigenvector for A is small (as in Table II) for these images. Consistent performance improvement with phase features is expected (and such features appear preferable) if more than one dominant phase eigenvector is retained. Ignoring the phase feature data for the letters, consistent trends emerge from Table III. Different results occur for different image pair recognitions. However, FS consistently performs best (or nearly so), with FK always being quite close and, surprisingly, KL-1 being consistently good. If two vector subsets were to be chosen for a given problem, those with the largest S value in one column would be selected. Combined phase and magnitude features perform better than either alone, but the increased complexity in using both magnitude and phase features often yields only a small improvement.

Thus, from such noise-free tests, Fourier transform magnitude data appear to perform well. (Optically, Fourier transform magnitude data are calculated much more easily and hence are preferable if the performance obtained is adequate.) But phase data are preferable (if their largest eigenvector is sufficiently dominant). If two Fourier transform phase eigenvectors are retained, and if Fourier transform phase data can easily be calculated, phase features are preferable. If the object classes being discriminated are sufficiently different, KL is adequate. However, in general, FK or FS is recommended. Clearly, the results are data-dependent. Thus, let us consider the performance of all feature extractors in the presence of noise before advancing a final decision.

### 6.5. Noise performance comparisons

To best assess the performance of our five feature extractors, we consider our two case studies (vehicles and letters separately). In Table IV, we list the calculated S value for the vehicle identification tests for both magnitude-only and phase-only Fourier transform data as a function of the standard deviation $\sigma_n$ of the noise added to the input data. In Table V, similar data for our letter identification case study are provided. In these tables we also include the magnitude of the eigenvalue for the dominant eigenvector for the class 1 and class 2 data (the reason for this will be apparent shortly).

In Table IV, we focus attention on the S performance values

**TABLE IV. Eigenvalues and Separability Measure S for the Truck and Tank Images for Different Noise Standard Deviations and for Magnitude-Only and Phase-Only Fourier Transform Data**

| Noise standard deviation | Truck and tank (magnitude data) | | | | | Truck and tank (phase data) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 |
| Dominant eigenvalue of class 1 | 0.995 | 0.995 | 0.994 | 0.993 | 0.992 | 0.881 | 0.457 | 0.238 | 0.234 | 0.243 |
| Dominant eigenvalue of class 2 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 | 0.617 | 0.599 | 0.558 | 0.504 | 0.450 |
| S for KL-1 | 4.215 | 4.341 | 4.382 | 4.315 | 4.145 | 6.169 | 4.513 | 2.338 | 0.588 | 0.416 |
| S for KL-2 | 2.893 | 2.924 | 2.919 | 2.874 | 2.792 | 4.760 | 4.385 | 4.017 | 3.669 | 3.083 |
| S for FK-1 | 3.923 | 3.955 | 3.952 | 3.915 | 3.847 | 5.624 | 3.787 | 1.668 | 0.436 | 0.429 |
| S for FK-2 | 3.894 | 3.924 | 3.919 | 3.878 | 3.806 | 4.046 | 3.786 | 3.963 | 3.537 | 3.110 |
| S for FS | 4.625 | 4.705 | 4.746 | 4.744 | 4.699 | 7.689 | 5.592 | 3.995 | 3.598 | 3.229 |

**TABLE V. Eigenvalues and Separability Measure S for the Letter A and B Images for Different Noise Standard Deviations and for Magnitude-Only and Phase-Only Fourier Transform Data**

| Noise standard deviation | A and B (magnitude data) | | | | | A and B (phase data) | | |
|---|---|---|---|---|---|---|---|---|
| | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.0 | 0.1 | 0.2 |
| Dominant eigenvalue of class 1 | 0.999 | 0.999 | 0.998 | 0.996 | 0.994 | 0.582 | 0.350 | 0.198 |
| Dominant eigenvalue of class 2 | 0.999 | 0.999 | 0.998 | 0.998 | 0.997 | 0.836 | 0.600 | 0.392 |
| S for KL-1 | 8.515 | 9.321 | 7.748 | 4.580 | 2.744 | 0.744 | 0.382 | 0.198 |
| S for KL-2 | 6.918 | 6.652 | 5.049 | 3.341 | 2.441 | 8.480 | 6.397 | 3.451 |
| S for FK-1 | 13.487 | 16.453 | 12.344 | 6.193 | 3.766 | 0.259 | 0.235 | 0.072 |
| S for FK-2 | 13.475 | 16.402 | 12.284 | 6.172 | 3.753 | 8.547 | 6.143 | 3.548 |
| S for FS | 13.320 | 17.057 | 13.949 | 6.752 | 4.010 | 8.555 | 6.800 | 3.480 |

obtained as $\sigma_n$ increases. Reading the performance measure data horizontally, we see a negligible change in S with $\sigma_n$ for magnitude-only data. Similarly, the maximum eigenvalues $\lambda_{max}$ for both image classes also vary only slightly with $\sigma_n$. Using the phase-only Fourier transform features, we find a quite significant decrease in S as $\sigma_n$ increases. This shows that the performance S for phase features degrades quite significantly as the noise in the data is increased. In this case, $\lambda_{max}$ is also reduced significantly with increasing $\sigma_n$ and thus reflects the trend noted above.

In Table V, similar data are shown for our letter recognition case study. The magnitude feature data show a decrease in S as $\sigma_n$ increases. However, the decrease in S for the phase features is even more appreciable. Thus, from both Table IV and Table V we find that phase features are a less robust feature set than magnitude features in the presence of noise.

Let us now consider the reasons for the observed performance in Tables IV and V. We first note that we expect the Fourier transform magnitude data to be concentrated in several dominant spatial frequencies, whereas the Fourier transform phase data are expected to be more uniformly distributed over the Fourier transform plane. This is logical and is the basis for the success of dimensionality reduction using WRD Fourier transform plane sampling. Thus, with Fourier transform magnitude features, a few wedge or ring detector elements dominate object identification. Conversely, with Fourier transform phase features, all wedge or ring detector elements contribute more equally. Thus, when a given amount of noise is present in the input image, it is evenly distributed over all wedge and ring Fourier transform samples (for white noise). For Fourier transform

magnitude features, the actual noise contribution in the important Fourier transform plane wedge and ring elements is proportionally much less than for Fourier transform phase features. Hence, we might expect (as observed) poorer noise performance using phase features rather than magnitude Fourier transform features. The computational accuracy associated with evaluating the function from which the nonlinear phase features are obtained may be a secondary factor in this observed noise performance for phase features.

## 7. SUMMARY AND CONCLUSIONS

The classic Fourier transform plane has been considered as a feature space for distortion-invariant recognition. The use of wedge- and ring-sampled Fourier transform plane features was employed to reduce the dimensionality of the feature space and to provide scale and rotational insensitivity in our feature extractor. New feature extraction algorithms were applied to these WRD samples of the Fourier transform plane, and the importance of magnitude and phase Fourier transform data for pattern recognition applications was considered. The performance of our pattern recognition system for two different two-class image data bases (vehicles and letters) was quantified for all feature extractors, for phase and magnitude Fourier transform features, and in the presence of noise.

The feature extractors considered were the Karhunen-Loeve dominant eigenvectors for each class, the Fukunaga-Koontz transformed discriminant vectors, and the Foley-Sammon discriminant vector. Extensions of all cases to more than two-class pattern recog-

nition applications follow directly. For the cases considered, the KL-1 vector performed well, but the FK vectors were generally better, and the FS vector was almost always the best. This follows from the fact that the KL technique provides only intraclass recognition, whereas FK and FS techniques provide interclass discrimination. Our study of the use of magnitude or phase Fourier transform features showed that phase features were sometimes better, but that in general the dominance of one eigenvector for phase data was harder to achieve and thus, if such features were used, more eigenvectors must be retained. This, plus the ease with which magnitude Fourier transform features can be optically computed, makes such a feature space preferable. This use of Fourier transform plane magnitude and phase data for pattern recognition differs considerably from its more conventional use in image reconstruction. Lastly, we considered the noise performance and robustness of Fourier transform magnitude and phase features and found magnitude features to be far preferable. An initial heuristic but theoretical basis for this result that appears to be quite plausible was advanced.

## 9. REFERENCES

1. H. C. Andrews, IEEE Trans. Comp. C-20, 1045 (Sept. 1971).
2. G. G. Lendaris and G. L. Stanley, Proc. IEEE 58, 198 (Feb. 1979).
3. H. Kasden and D. Mead, Proc. Elec. Opt. Sys. Des., 248 (1976).
4. S. Watanabe, "Karhunen-Loeve Expansion and Factor Analysis, Theoretical Remarks and Applications," 4th Prague Conf. on Inf. Theory (1965).
5. H. Murakami and B. V. K. Vijaya Kumar, IEEE Trans. Patt. Anal. and Mach. Intell. PAMI-4, 511 (Sept. 1982).
6. K. Fukunaga and W. L. G. Koontz, IEEE Trans. Comp. C-19, 311 (April 1970).
7. D. H. Foley and J. W. Sammon, IEEE Trans. Comp. C-24, 381 (March 1975).
8. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley and Sons, New York (1973).
9. A. V. Oppenheim and J. S. Lim, Proc. IEEE 69, 529 (May 1981).
10. M. H. Hayes, J. S. Lim, and A. V. Oppenheim, IEEE Trans. Acoust., Speech, Signal Proc. ASSP-28, 672 (Dec. 1980).

# 7. <u>DISTORTION-PARAMETER ESTIMATION FROM A CHORD DISTRIBUTION OPTICALLY-GENERATED FEATURE SPACE</u>

Ref. 10

Chord Distributions in Pattern Recognition:
Distortion Invariance and Parameter Estimation

Wen-Thong Chang and David Casasent

Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

## Abstract

The use of chord distributions in pattern recognition is discussed and efficient ways to compute such distributions are noted. New methods to achieve scale and in-plane rotational distortion-invariant multi-class recognition and estimates of the distortion parameters are described. 3-D out-of-plane rotational distortion-invariant methods are reviewed.

## 1. Introduction

Chord distributions are well-known features that describe the shape of an object and that are useful for object identification [1-3]. These features can easily be computed (optically or digitally) from the autocorrelation. In Section 2, we define the chord distribution and discuss different chord pdfs. These include an observation space $h(\ell_x, \ell_y)$ and a feature space $h(r)$ and $h(\theta)$. New insight is provided into the local and global features produced by chord pdfs and the use of silhouette and boundary (profile) imagery. In Section 3, attractive properties of these chord distributions for scale and in-plane rotation invariance are discussed. A new use of such features for distortion-invariant multi-class object recognition and methods to extract the object's scale and orientation are advanced. In Section 4, methods to achieve 3-D object distortion-invariance (to out-of-plane rotations) are reviewed. The resultant feature extractor thus enables multi-class object classification in the presence of a wide variety of geometrical distortions.

## 2. Chord Features and Distributions

**2.1 Definition.** The conventional chord distribution $h(r,\theta)$ is a plot of the distribution of the lengths (r) and directions ($\theta$) of all chords drawn between all pairs of points on the boundary of the object $f(x,y)$. The two chord pdfs of most use are $h(r)$ and $h(\theta)$, the pdfs of chord lengths r and directions $\theta$. To most easily compute the various chord distributions, one can begin by forming the autocorrelation

$$b(x,y) \circledast b(x,y) = \iint b(x,y)\, b(x - \ell_x, y - \ell_y)\, dxdy = R(\ell_x, \ell_y) = h(\ell_x, \ell_y) \tag{1}$$

of the boundary $b(x,y)$ of an object. The autocorrelation describes the number of points of intersection for a given horizontal and vertical shift $(\ell_x, \ell_y)$ between two shifted images of the object. The value of R at a given $(\ell_x, \ell_y)$ thus precisely gives the number of chords with given horizontal and vertical projection lengths $(\ell_x, \ell_y)$ [3-4].

To show this, we write $\ell_1 (\ell_x, \ell_y) = (r \cos\theta, r \sin\theta)$ where $r = (\ell_x^2 + \ell_y^2)^{1/2}$ is the radial chord length and $\theta = \tan^{-1}(\ell_x/\ell_y)$ is the chord's angular orientation. Substituting into (1), we see that $R(\ell_x, \ell_y)$ contains information from which $h(r,\theta)$ can be obtained. From $h(\ell_x, \ell_y)$, the chord distribution $h(r,\theta)$ can be calculated. The chord pdfs $h(r)$ and $h(\theta)$ are more useful and are most easily calculated from $h(\ell_x, \ell_y)$ by appropriately sampling the autocorrelation function. If the autocorrelation is sampled radially, we obtain

$$h(r) = \int h(\ell_x, \ell_y)\, rd\theta \tag{2}$$

If we sample it angularly, we obtain

$$h(\theta) = \int h(\ell_x, \ell_y)\, dr \tag{3}$$

**2.2 Realization.** These $h(r)$ and $h(\theta)$ chord pdfs are the features we will use. To obtain (2) and (3) optically, we form $h(\ell_x, \ell_y)$ optically (typically from the Fourier transform of the power spectrum of the object) and sample this distribution using wedge and ring-shaped detector elements [4]. Such a detector unit exists (Figure 1) with 32 wedges in one-half of a circular plane and 32 rings in the other half of the plane [5]. The autocorrelation function is symmetric and thus no loss of information results by sampling only half of the autocorrelation plane. In terms of chord distributions, the symmetry of the autocorrelation function arises because each chord in the image is counted twice as one traverses the boundary of the object. In one case, one end point of the chord is encountered first and then the other end point is encountered first. The first corresponds to a chord with projections

$(\cdot_x, \cdot_y)$ and a length r. The symmetric case corresponds to a chord with projections $(-\cdot_x, -\cdot_y)$ and a direction $-\theta$ rather than $+\theta$. For similar reasons of symmetry, the orientation of the wedge and ring halves of the detector does not matter. The wedge outputs provide $h(\theta)$ (quantized to 32 $\theta$ values over 180°) and the ring outputs provide $h(r)$ (quantized to 32 r values over the radius of the autocorrelation function). Figure 2 shows the general block diagram of our chord distribution feature generator using a wedge-ring detector (WRD).
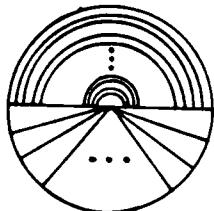


Figure 1. Simplified representa-       Figure 2. Simplified Block Diagram of a chord dis-
          tion of a wedge-ring                   tribution pattern recognition system.
          detector (WRD).

2.3 Boundary, Silhouette and Gray-Level Objects.  Different chord distributions result depending on the type of input object. For a boundary or edge image (case A), the distribution produced is of the number of edge or boundary pixels (i.e., the number of chords). This is the conventional chord distribution. For a silhouette image (binary with all ones on the object and with zeroes on the background), the distribution produced    (case B) is the same as case A, but weighted by the common area of overlap of the two images for the given $(\cdot_x, \cdot_y)$ shift. If the shift is large, corresponding to long chords, the weighting will be small. However, if the shift is small, corresponding to short chords, the weighting will be large. Thus, this weighted chord distribution that results for the case of a silhouette object (case B) emphasizes short chords more than long chords. The chord distribution in case A will be more susceptible to noise in the interior of the object (internal pixels of value 1 result in many new chords being produced in case A, whereas in case B zero internal pixels cause a loss of chords but a much lower percent change results than in case A). When the chord distribution in case A is computed from the autocorrelation or power spectrum (as in Sections 2.1 and 2.2), it is much simpler to calculate than by other methods which have great difficulty when applied to a non-continuous boundary. However, each missing boundary pixel in case A will still result in a loss in the number of chords counted.

The weighted chord distribution (case B) emphasizes short chords. These correspond to local object features (whereas long chords correspond to global object features). Since local object features are useful for discrimination between object classes (inter-class), we expect the weighted chord distributions to provide superior object discrimination. Long chords, corresponding to global object features, are more useful for intra-class object recognition (within one object class, in the face of various object distortions). The performance of weighted chord distribution features in the presence of noise in the input is expected to be superior to the use of conventional chord features. In a boundary image (case A) with N pixels on the boundary, each noise pixel on the object produces N new chords and each missing boundary pixel (due to noise) causes N chords to be removed from the distribution. With $N^2$ total chords, each noise pixel thus changes the total h by a factor 1/N. In case B, each weighting function is on the order of $N^2$ (this is more true for short chords than long chords) and thus each noise pixel produces a change in h by a factor of only $1/N^2$ (this is a considerable improvement, since N is usually quite large). For the same reason that the change in h for short chords is less susceptible to noise, it will also be less susceptible to small differences in the object's shape (due to distortions). but changes due to sufficiently different objects are still retained.

The dynamic range of the chord features in cases A and B appears to be comparable. Since use of the boundary image (case A) whitens the image's spectrum and results in a sharper correlation function compared to the broader correlation pattern that results in case B, wedge-ring detection in case B is much simpler. Case B is clearly preferable from noise considerations, its inter-class discrimination is clearly enhanced and its intra-class recognition should be retained. Since all chords are available (and more easily detectable in case B), one can use the preferable chord features (short or long, local or global) for a given problem.

If the gray-levels of the object and its internal structure are reliable, then the chord distribution for the gray-level image (case C) is most useful. The distribution in case B is one level of a general chord distribution. The distribution in case C is a higher-level of generalized chord distribution [4]. In this case, the chord distribution for all internal chords or internal object points is provided. Algorithms such as (1) with the boundary object $b(x,y)$ replaced by the full object $f(x,y)$ provide such features with no increase in computational load for optical systems (digital systems can achieve simplified correlations when operating on binary imagery).

### 3. Scale and Rotation-Invariant Chord Processor

**3.1 Insight.** The chord pdf $h(r)$ is invariant to in-plane rotation of the object. This is obvious since the in-plane rotation of an object does not alter its radial distribution. The chord pdf $h(\theta)$ simply shifts with in-plane rotations. This follows directly since $h_1(\omega_x, \omega_y) = h(r\cos\theta, r\sin\theta)$ changes to $h_2(\omega_x, \omega_y) = h[r\cos(\cdot + \theta_0), r\sin(\cdot + \theta_0)]$ for rotation of the input object by $\theta_0$, i.e. in $(r,\theta)$ space, $h_2(r,\theta) = h_1(r, \theta + \theta_0)$. Thus in-plane object rotations rotate $h(\omega_x, \omega_y)$ and translate $h(\theta)$. The chord pdf $h(\theta)$ is invariant to scale distortions of the object whereas $h(r)$ scales (rather than shifts) with an input scale change $\alpha$. The invariance of $h(\theta)$ with scale is obvious. For a scale change in the input object, the $h(r)$ distribution scales proportional to $\alpha$ and $h(\alpha r)$ is obtained. As long as half of the correlation plane is sampled in $\theta$ and $r$, the above remarks remain valid [due to the symmetry of the autocorrelation and due to the cyclic shift nature of $h(\theta)$]. Table 1 summarizes these properties.

Table 1. Properties of $R(r)$ and $h(\theta)$ distributions

| PARAMETER | Feature | Distribution Property | Amplitude Effects |
|---|---|---|---|
| Rotation, $\theta_0$ | $h(r)$ | Invariant | None |
| | $h(\theta)$ | Shifts $\propto \theta_0$ | None |
| Scale, | $h(r)$ | Scales $r \rightarrow \alpha r$ | $\alpha^{-3}$ |
| | $h(\theta)$ | Invariant | $\alpha^{-3}$ |
| Translation | $h(r)$ | Invariant | None |
| | $h(\theta)$ | Invariant | None |

Table 1 also notes the effects on the amplitudes of the $h(r)$ and $h(\theta)$ features. We now detail the origin of these variations. We consider first the effect of a scale change (by a factor of $\alpha$) in the input object on the amplitudes of $h(r)$ and $h(\theta)$. First, we consider the observation space $h(\omega_x, \omega_y)$. The image $f(x,y)$ with scale $\alpha = 1$ produces $h_1$. This relates to $h_2$ for $\alpha \neq 1$ as detailed below. From (1),

$$h_1(\omega_x, \omega_y) = \iint f(x,y)\, f(x + \omega_x, y + \omega_y)\, dx\, dy \qquad (4)$$

For the scaled object (scale factor $\alpha$)

$$h_2(\omega_x, \omega_y) = \iint f(\alpha x, \alpha y)\, f[\alpha(x + \omega_x), \alpha(y + \omega_y)]\, dx\, dy \qquad (5)$$

Changing variables $(u,v) = (\alpha x, \alpha y)$, we obtain

$$h_2(\omega_x, \omega_y) = (1/\alpha^2) \iint f(u,v)\, f(u + \alpha \omega_x, v + \omega_y)\, du\, dv = (1/\alpha^2) h_1(\omega_x, \omega_y). \qquad (6)$$

From (6), we see that $h_2$ is a scaled version of $h_1$ with the amplitudes scaled by $(1/\alpha^2)$.

Now we consider the effect of scale changes on the $h(r)$ and $h(\cdot)$ distributions. For the $h(r)$ distribution, we find, from (2),

$$h_1(r) = \int f(r\cos\cdot, r\sin\cdot) r d\cdot . \tag{7}$$

For a scaled object (scale factor $a$),

$$h_2(r) = (1/a^2) \int f(a r\cos\cdot, a r\sin\cdot) r d\cdot = (1/a^3) \int f(a r\cos\theta, a r\sin\cdot) a r d\cdot = (1/a^3) h_1(a r) \tag{8}$$

Thus, from (7), we find a scale change (by $a$) between the $h_1(r)$ and $h_2(r)$ distributions and an amplitude scale factor $(1/a^3)$. For $h(\theta)$, the effect of a scale change $a$ is simply

$$h_2(\theta) = (1/a^3) h_1(\cdot) , \tag{9}$$

i.e., only an $(1/a^3)$ amplitude factor.

The distribution and amplitude effects of $\theta_0$ and $a$ distortions summarized in Table 1 and detailed above are valid for continuous data and continuous $r$ and $\cdot$ sampling. Finite $r$ and $\cdot$ sampling is expected to change the exact results somewhat. Specifically, due to sampling, an exact ratio of $a^3$ is not expected. Furthermore, the scale change from $h(r)$ to $h(a r)$ can be quite difficult to uncover since the distribution for one scale may lie in 11 rings and the distribution for another scale can easily lie in 6 or 8 rings. Thus, the $h(\theta)$ distribution is the most useful one for general ($a$ plus $\theta_0$) distortions. The $h(r)$ scale $r$ changes linearly (to $\cdot$) and is thus not a simple shift. When the effect of a finite number of $r$ samples is included, the $h(r)$ effect with $a$ is nonlinear. If we scale the $h(r)$ distribution in $r$ by $a$, the ratio $a^3$ then exists between the $h(r)$ for a scaled object and the original $h(r)$ scaled in $r$ by $a$. Thus, the distribution and amplitude effects of scale are coupled as just detailed. Specifically, this means that the amplitude ratio is $a^3$, but it is this for different $r$ and $a r$ points in the distribution (not the same $r$ points).

By $g(x,y) = f(a x, a y)$, we describe both the position and value of the pixels. Specifically new pixel $(x,y)$ is old pixel $(a x, a y)$ (i.e. $a > 1$ corresponds to a scale decrease) and the value of the old and new pixel are the same. Our above formulae for amplitude effects proportional to $a^{-3}$ thus apply for binary silhouette images (analogous formulae for gray-scale images can be derived and used if the input data is gray-scale. In such cases, with $\alpha < 1$, we have a larger image with more pixels and more intensity per pixel, since the object is closer and received intensity is proportional to range squared). For binary silhouette images and $\alpha < 1$, the new image is larger. Thus, for a given $(\ell_x, \ell_y)$ shift, we obtain more overlap, larger correlation values, more weighting and more chords. Our new $h_2$ will have larger amplitudes (more chords) than $h_1$ and this agrees with $h_2 = \alpha^{-2} h_1 > h_1$ predicted.

### 3.2 Distortion – Invariant ($\alpha$ and $\theta_0$) Pattern Recognition.

The insight provided in Section 3.1 and the distortion effects summarized in Table 1 are most useful in devising a new pattern recognition feature extractor (invariant to scale $\alpha$ and in-plane rotation $\theta_0$ distortions). We consider 3 distortion cases separately below and summarize our results in Table 2. From Table 1, we note that the $h(\theta_0)$ distribution is the most useful one in general (since it provides invariance to scale automatically and to rotations if shifted versions of $h(\theta)$ are tested; and since the ratio of $h(\theta)$ and a reference $h_R(\theta)$ provides an estimate of $\alpha$, whereas the best shift of $h(\theta)$ provides an estimate of $\theta_0$). For only scale distortions, $h(\theta)$ is best, and for only rotation distortions, $h(r)$ is best for classification (since these features are invariant to the indicated distortions).

#### 3.2.1 In-Plane Rotations.

For the case when $\theta_0$ is the only distortion present, we compare the $h(r)$ distribution $h_R(r)$ for all references $R$. This provides an estimate of the object class $R$. Next, for the best reference $R$ (obtained from the $h(r)$ and $h_R(r)$ comparisons), we compare $h(\theta)$ and $h_R(\theta)$ for various shifts $\theta_0$ in $h_R(\theta)$. From the $h_R(\cdot + \theta_0)$ and $h(\cdot)$ comparisons, we obtain a verification of our initial class estimate $R$ and an estimate of $\theta_0$. A combination of both $h(r)$ and $h(\theta)$ tests thus provides the best class $R$ estimates.

#### 3.2.2 Scale Changes.

For the case of an $\alpha$ distortion alone, we compare $h(\theta)$ for the test input vs. $h_R(\theta)$ for all references $R$. We must compare $h(\theta)/h_R(\cdot)$ for each $\theta$. The reference $R$ for which this ratio is constant for all $\theta$ provides the class estimate $R$. The ratio $h(\theta)/h_R(\theta)$ provides an estimate of $\alpha$ also. To confirm our $R$ and $\alpha$ estimates, we form $h(r)$ and $h_R(a r)$ for the initial $R$ and $a$ estimates. Agreement of $h(r)$ and $h_R(a r)$ confirms our initial estimates. Combining both the $h(\theta)$ and $h(r)$ tests again yields better estimates.

#### 3.2.3 Combined Scale ($a$) and rotation ($\theta_0$) Distortions.

When both $\alpha$ and $\theta_0$ distortions are present (the most general case), analysis relies on $h(\theta)$ and is more complex. We form $h(\theta)/h_R(\cdot + \theta_0)$ for all $R$ and all shifts $\theta_0$. When this ratio is constant for all $\cdot$, the corresponding $R$, $a$ and $\theta_0$ estimates are obtained. The ratio provides the $a$ estimate.

Table 2  Scale $\alpha$ and In-Plane Rotation $\theta_0$ Invariant Multi-Class Pattern Recognition

| CASE | Procedure | Remarks | Results |
|------|-----------|---------|---------|
| **(A)** Rotation $\theta_0$ Only | Compare $h(r)$ and $h_R(r)$ | $h(r)$ is Rotation Invariant | Class R Estimate |
| | Compare $h_R(\theta + \theta_0)$ and $h(\theta)$ | $h(\theta)$ shifts with $\theta_0$ | Confirms R Estimate<br>Provides $\theta_0$ Estimate |
| (B) Scale $\alpha$ only | Compare $h(\theta)/h_R(\theta)$ for each $\theta$ | Constant Ratio Provides R<br>Ratio Provides $\alpha$ Estimate | Class R and Scale $\alpha$ Estimates |
| | Compare $h(r)/h_R(\alpha r)$ | Confirms above estimate | Confirms R and $\alpha$ Estimates |
| (C) Rotation $\theta_0$ and Scale $\alpha$ | Compare $h(\theta)/h_R(\theta + \theta_0)$ for all R and all shifts $\theta_0$ | Constant Ratio Provides R and $\theta_0$. Ratio gives $\alpha$ | Initial Estimates of R, $\theta_0$, $\alpha$ |
| | Compare $h(r)/h_R(\alpha r)$ | Confirm above Estimates | Confirm R and $\alpha$ Estimates |

As a check, we form $h(r)/h_R(\alpha r)$ for the initial R and $\alpha$ estimates. From the constancy of the ratio, we verify our R and $\alpha$ estimates. Forming $h(r)/h_R(\alpha r)$ initially for all $\alpha$, is more computationally intensive and thus the order chosen appears best. This is also the most general case.

#### 4.  Out-Of-Plane Distortions $\phi_0$

For $\alpha$ and $\theta$ distortions, we require one $h(r)$ and $h(\theta)$ distribution per class R for our training set. To accommodate out-of-plane distortions $\phi$, we use several training set images per object class and from all $h(r)$ and $h(\theta)$ features select those with the largest Fisher ratio F (from training set data). We then form a linear discriminant functions $\underline{w}$ that maximizes F for a multi-class feature set. An input test feature vector $\underline{c}$ (chord distribution) is projected onto $\underline{w}$ and the projection value determines the input object class. This algorithm [4] has demonstrated perfect performance in selected image distortion tests.

#### 5.  Summary.

Chord distributions $h(r)$ and $h(\theta)$ have been shown to be easily computed from the autocorrelation of the input object and WRD (radial and angular) sampling. Using the various properties (Table 1) of $h(r)$ and $h(\theta)$, a new multi-class pattern recognition system for scale and in-plane rotational distortions was advanced (Table 2). Combined with our prior out-of-plane rotational distortion work (Section 4), this feature space can provide full 3-D object distortion invariance and estimates of the distortion parameters (orientation and scale) of the object.

#### Acknowledgements

#### References

1.  G. Tenery, IEEE Trans. Mil. Electron. ME-7, 196 (1963).
2.  D. J. H. Moore and D. J. Parker, Pattern Recognition 6, 149 (1974).
3.  S. P. Smith and A. K. Jain, Comput. Graphics Image Process, 20, 1699 (1982).
4.  D. Casasent and Wen-Thong Chang, Applied Optics, Vol. 22, 2087, July 15, 1983.

5.  H. Kasdan and D. Meade, Proc. Electron. Opt. Syst. Des.,248 (1975).

# 8. GENERAL ARCHITECTURE AND INITIAL RESULTS OF A MOMENT-BASED HIERARCHICAL CLASSIFIER

Ref. 11

# HIERARCHICAL FISHER AND MOMENT-BASED PATTERN RECOGNITION

David Casasent and R. Lee Cheatham*

Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

*Present Address:  Battelle Northwest, P.O. Box 999, Richland, WA 99352

## ABSTRACT

A two-level feature extraction classifier using a geometrical-moment feature space is described for multi-class distortion-invariant pattern recognition.  The  first-level classifier provides object class and aspect estimates using multi-class Fisher projections and optimized two-class Fisher projections in a hierarchical classifier.  Aspect estimates are provided from ratios of the computed moments.  The second-level classifier provides the final class estimate, distortion parameter estimates and the confidence of the estimates. Extensive test results on a ship image database are presented.

## 1.  INTRODUCTION

One can efficiently compute the moments of an input object by various methods [1,2]. These features are excellent descriptions of the geometrical aspects of an object.  They are quite unique since they can provide information on the orientation, scale and location of the input object [2] and because they can be corrected for various system computing errors [3].  In this paper, our earlier moment classifier [2] is modified to include a two-level classifier (Section 2).  This provides significantly improved performance.  We earlier [4] described initial results for robotic object parts.  Here, we detail the new two-level classifier design (Section 2), and the performance obtained (Section 4) for an extensive ship image database (Section 3).

## 2.  NEW MOMENT - BASED CLASSIFIER

### 2.1  Moment Statistics

The geometrical moments

$$m_{pq} = \iint f(x,y) x^p y^q dxdy \qquad (1)$$

of an input object $f(x,y)$ are jointly-Gaussian random variables (JGRV) [6] due to the finite spatial sampling of the input image and they are good estimates of the actual moments of an input object.  This JGRV model allows us to use a conventional Bayesian classifier [5] that minimizes the probability of incorrect class estimates (Section 2.4).  The mean $\mu_i$ and covariance $\Sigma_i$ for each object class i must be estimated to use this classifier.  Generally, this requires a training set of imagery.  Because the moment features are JGRVs, we require only one object view per class to achieve such estimates.  Thus, such a classifier using these geometrical moment features does not require a large training set of data.

### 2.2  Aspect Angle Estimator (First-Level Classifier)

The moment features are JGRVs only with respect to scale (a,b), translation $(x_0,y_0)$ and in-plane rotations ($\theta$), but not for out-of-plane rotations ($\phi$).  Thus, we must estimate $\phi$ for the input object.  This is achieved in our first - level classifier,  which thus includes each image aspect view as a separate object class.  We thus distinguish object classes (in our present database tests, Section 3, this refers to different ship classes) from view classes (these include all aspect views of all ship images).  In our first-level classifier, we estimate the aspect angle of the input object from the ratio $\hat{A} = \mu_{20}/\mu_{02}$ of the central moments, where $\mu_{20} = m_{20}-m_{10}^2/m_{00}$ and $\mu_{02} = m_{02}-m_{01}^2/m_{00}$.  For all reference objects in the class being tested, we calculate A and then form $K = \hat{A}/A$.  The aspect view with the K value closest to unity is selected plus all aspect views with $K < T_A$ (the aspect threshold).  In our tests, we use $T_A = 1.5$.  Those aspect views of the class being tested with $K \leq T_A$ are passed to the second-level classifier.

### 2.3  Object Class Estimator (First-Level Classifier)

To further reduce the number of view classes (aspect plus class) passed to the second-level classifier, we use multi-class and two-class Fisher projections [7] on a training set of ship images.  From these scatter plots for the multi-class Fisher projection, we select the two subsets of object classes that are best separated at each node in a tree classifier.  For each node, we then calculate (from training set images) the two-class

Fisher vector that best separates and clusters the two subsets at each node. For example, for node 0, the full set of N-1 multi-class Fisher vectors $F_1$ to $F_{N-1}$ for the N object classes are computed. From examination of the projections of the inputs onto the two most dominant Fisher vectors, we select the two subsets (with possibly several object classes per subset) to be separated at node 0. The two-class Fisher vector for these two subsets is then calculated and the projections of all training set data on this Fisher vector are plotted. From this plot, weighted distances to the two class means were calculated and a class estimation threshold $T_{C1}$ is selected. If the weighted distance for the projection of an input test image exceeds $T_{C1}$, then we proceed down the corresponding branch at that node of the hierarchical tree. On each branch, another node is present at which the classes on that branch are further divided into two smaller subsets. New multi-class Fisher projections are used at each node to determine the two subsets to use and a new two-class Fisher projection vector is calculated for use at each node.

All of these calculations are performed off-line on a limited number of training set images. To account for scale and translational distortions in the input image, the central moments normalized for scale are used in the first-level classifier and the scatter plots are calculated for different aspect views of each class. Details and examples of this organized first-level class estimator are provided in Section 3. This hierarchical procedure is followed until terminal nodes are reached and a decision on the class estimate(s) of the input object is made. For some objects [4], full separation into all classes is not possible. If the calculated weighted-distance measure for the input test image is less than $T_{C1}$, all classes at that node are passed to the next level. Use of alternate nodes is included to allow better separation of subsets at certain nodes for particular databases. The real-time calculations involved in this hierarchical class estimator are quite simple. The test feature vector is simply projected onto several discriminant vectors (each such operation is merely a vector inner product) and from the projected values, class estimate(s) are obtained. For each such class estimate, the aspect class estimator (Section 2.2) is used to determine the total number of view classes to be processed in our second-level classifier.

## 2.4 Bayesian Classifier (Second-Level Estimator)

Because the operations required in the Bayesian classifier are computationally more intense, the first-level estimator is used to reduce the number of view classes to be processed in the second-level estimator. The conventional Bayesian classifier minimizes the probability of an incorrect class i estimate (here i denotes a view class). Using the assumptions of JGRV features, the discriminant function to be minimized is [5]

$$g_i(\underline{x}) = (\underline{x}-\underline{\mu}_i)^T \underline{\Sigma}^{-1} (\underline{x}-\underline{\mu}_i), \tag{2}$$

where $\underline{\mu}_i$ and $\underline{\Sigma}_i = \underline{\Sigma}$ are the mean vector and covariance matrix for class i. For our present case, the feature vector $\underline{x}$ is a moment vector $\underline{m}$ and thus only one object view per class is needed to measure $\underline{\mu}_i$ and $\underline{\Sigma}_i$. Operation of such a classifier thus proceeds by calculating $g_i(x,y)$ for the measured input feature $\underline{x}=\underline{m}$ for all object classes i. The class i that minimizes $g_i(\underline{x})$ is the best class estimate in a Bayesian sense. The discriminant function in (2) is the Mahalanobis distance. If $\underline{\Sigma} = \underline{I}$, it becomes the Euclidean distance measure or a nearest-neighbor classifier. Use of $\underline{\Sigma} = \underline{I}$ assumes that all moments are independent and that the expected variations of all moments are equal.

To utilize (2), i must be a view class. To calculate all object class distortion parameters, i.e., scale (a,b), range (R), translation $(x_0,y_0)$ and in-plane rotations (θ) as well as aspect view angle (:), we let the view class i include the object class and aspect view angle and we include the other parameters in a distortion parameter vector $\underline{b} = (x_0,y_0,a,b,R,\theta)$. We combine the view class and distortion parameters as $\underline{m}_i(\underline{b})$ and thus evaluate (2) for all view classes i and all distortion parameter vectors $\underline{b}$. Since $\underline{m}(\underline{b})$ is a nonlinear function of $\underline{b}$, we use an iterative algorithm of the form

$$\underline{b}^{k+1} = \underline{b}^k + a^k \underline{r}^k, \tag{3}$$

where $\underline{b}^k$ is the $\underline{b}$ estimate at iteration k and $\underline{b}^{k+1}$ is a point in an r-dimensional space at a distance $a^k$ in the direction $\underline{r}^k$ from the present estimate $\underline{b}_k$. To determine the complete form for the iterative algorithm in (3), we expand $\underline{m}_i(\underline{b})$ in a Taylor expansion series about the present $\underline{b}^k$ point as

$$\underline{m}_i(\underline{b}) = \underline{m}_i(\underline{b}^k) + \underline{J}^k(\underline{b}-\underline{b}^k), \tag{4}$$

where $\underline{J}$ is the Jacobian of $\underline{m}_i(\underline{b})$ with respect to $\underline{b}$ at the k-th iteration. For a measured input feature vector $\underline{\hat{m}}$, the error to be minimized is $\underline{e}_i = \underline{\hat{m}} - \underline{m}_i(\underline{b})$ and the square-error measure is $E_i = \underline{e}_i^T \underline{\Sigma}^{-1} \underline{e}_i$, where $\underline{\Sigma}^{-1}$ is the weighting matrix used. Substituting $\underline{e}_i$ and (4)

into the expression for $E_i$, the $\underline{b}$ that minimizes $E_i(\underline{b})$ is found to satisfy

$$\underline{b}^{k+1} = \underline{b}^k + [(\underline{J}^k)^T \underline{\Sigma}^{-1} \underline{J}^k]^{-1} (\underline{J}^k)^T \underline{\Sigma}^{-1} [\hat{\underline{m}} - \underline{m}_i(\underline{b})].$$

(5)

Eq(5) is the nonlinear iterative algorithm used in our second-level classifier to estimate $\underline{b}$. Thus, for each view class i (5) is repeated and new $\underline{b}$ estimates are obtained. For each $\underline{b}^k$, we calculate the normalized difference

$$\Delta g_i = [g^k(i,\underline{b}) - g^{k-1}(i,\underline{b})]/g^k(i,\underline{b})$$

(6)

between two successive $g_i$ estimates, where $g_i(\underline{b}) = E_i$. The iterations in the Gauss-Newton or Newton algorithm in (2) and (5) are continued until $\Delta g_i$ is less than a convergence threshold T.

## 2.5 Parameters and Overview

The full moment-based two-level estimator is shown in block diagram in Figure 1. It consists of an optical moment feature computer, first-level class and aspect estimators, and the second-level Bayesian classifier. The output from the two first-level estimators are used to access those reference moment vectors necessary for the second-level nonlinear iterative classifier. The final outputs are the class estimates i (class and aspect angle :), the target's distortion parameters or orientation information $\underline{b}$ and the confidence $g_i$ of the estimates.
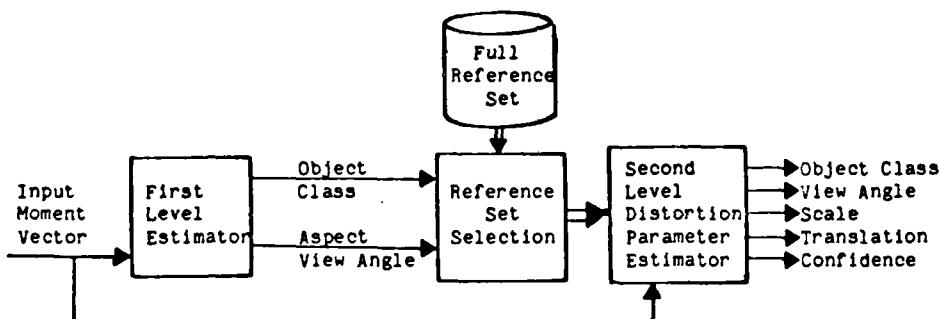


FIGURE 1
Block Diagram of a Two-Level Moment-Based Classifier

To facilitate $\underline{b}$ estimates, $x_0$ and $y_0$ are estimated from $-\hat{m}_{10}/\hat{m}_{00}$ and $-\hat{m}_{01}/\hat{m}_{00}$ and scale is estimated from $\hat{m}^{1/2}_{00}$. To facilitate calculations, $\underline{J}$ is evaluated with $(x_0,y_0,a,b) = (0,0,1,1)$, i.e. assuming that the presently calculated distortions $\underline{b}$ are correct and thus viewing future iterations as updates on the present $\underline{b}^k$ rather than the initial $b^0$ estimates. These and other features of the iterative algorithm allow it to converge in typically less than 15 iterations. Different approximations to $\underline{\Sigma}^{-1}$ were considered in our case study. Such measures were essential since $\underline{\Sigma}$ is ill-conditioned. Approximations considered were: $\underline{\Sigma}^{-1} = \underline{I}$ and $\underline{\Sigma}^{-1} = \underline{W} \underline{W}^T$, where $\underline{W}$ is the multi-class Fisher projection matrix of the reference vector set. The iterative convergence threshold T is typically chosen as 0.01. This corresponds to a 1% difference in successive iterates as in (6). The class estimation threshold $T_{C1} = [1-d_1/d_2]$ is chosen as 0.35, where $d_1$ and $d_2$ are the distances of the projection to the two weighted class boundaries at each node in our first-level class estimator. The class estimation threshold $T_{C2}$ for the second-level Bayesian classifier is defined similarly and is chosen to be 0.35 also.

## 3. SHIP DATABASE

### 3.1 Image Sets

Ships on the open-sea represent an attractive application for feature-space techniques (since one object can often easily be included in field-of-view). The class, orientation and range of the object in this application are necessary for missile guidance and target selection. The set of ship imagery available consisted of five ship models with 36 different aspect views per ship class available from a 90° depression angle (0° attack angle) at 10° intervals (a total of 180 view classes). Figure 2 shows the broadside views of the five ship classes.
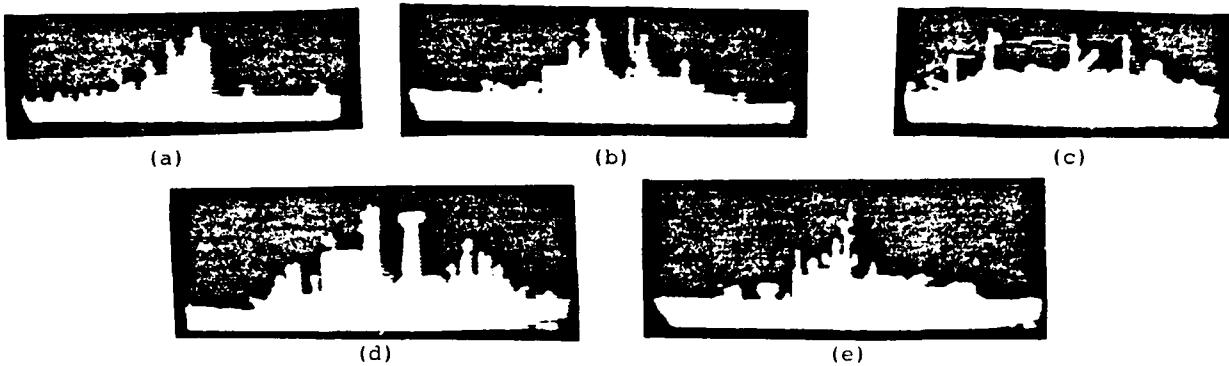
FIGURE 2
Broadside Views of the Five Ship Classes

TABLE 1
Ship Image Database, 36 Images Per Class

| CLASS NUMBER | SHIP NAME | SHIP TYPE |
|---|---|---|
| 0 | Moskva | Soviet Helicopter Cruiser |
| 1 | Leahy | U.S. Guided-Missile Cruiser |
| 2 | Hope | Hospital Ship |
| 3 | Albany | U.S. Guided-Missile Cruiser |
| 4 | Brooke | U.S. Guided-Missile Frigate |

Table 1 lists the names and types of each general ship class. For each ship, the original images were binarized and data sets with and without the hull removed were prepared. All data included were obtained with the hull present. Each image was 128 x 32 pixels with approximately 2000 pixels on the broadside views and less than 200 pixels on the bow and stern views. Several other ship image databases used are noted in Table 2. These include: the standard reference images used in the second-level classifier (these include only four images in the first quadrant, broadside images only and other selected object views).

TABLE 2
Miscellaneous Image Training and Test Sets Used

| DATA SET | SPECIFIC SHIP IMAGES | SYMBOL |
|---|---|---|
| Standard Reference Images | 10°,30°,50°,80° | S |
| Broadside Images | 40°-140°,220°-320° | B |
| Even Views | 0°,20°,etc. | E |
| Odd Views | 10°,30°,etc. | O |
| All Views | 0°,10°,20°,etc. | A |

3.2 Hierarchical Tree

In Figure 3, we show the scatter plot for all views of all five ship classes on the two dominant multi-class Fisher vectors. As seen, ship class two is the most easily separated. Thus, at node 0 we chose to separate the class two ship (the Hope) from the others. This yields a terminal node for one branch from node 0. At node 1, we examined a similar scatter plot for classes 0,1,3 and 4 and chose to separate the class 0 ship (the Moskva) from the three U.S. guided-missile ships. At node 2, we then separated three ships (the Brooke, a Frigate, from the two cruisers) and finally at node 3 we separated the two U.S. cruisers. Figure 4 shows the final hierarchical tree used for our ship image database.
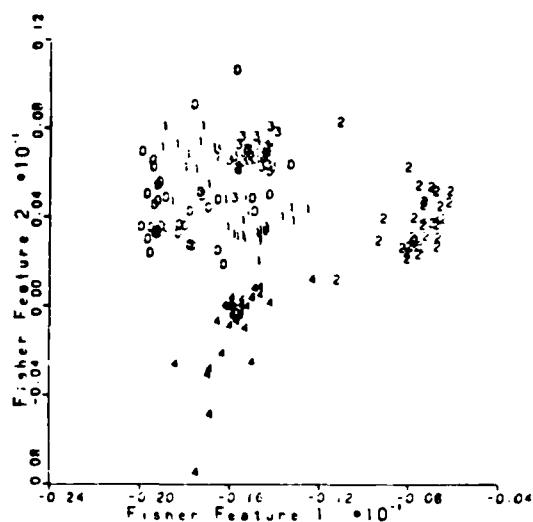
FIGURE 3
Sample Multi-Class Fisher Feature Scatter
Plot of all Ship Images (node 0)
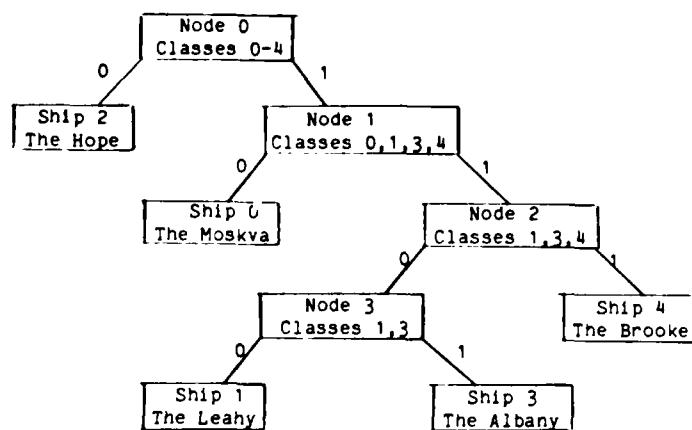


FIGURE 4
Ship Node Tree Constructed from Multi-Class
Fisher Projections and Scatter Analyses

## 3.3 Node Threshold Selection

In Figure 5, we show the projection of the subsets at node 2 in the tree of Figure 4 (a 1 denotes a class four projection and a 0 denotes class one and three projections). The discrimination point D is the point where the weighted distances to the means ($y_0$ and $y_1$) of the two subsets are equal. The lower bounds $D_0$ and $D_1$ for each subset are noted. For less uniform clusters, $D_0$ and $D_1$ are selected at several standard deviations from $y_0$ and $y_1$. The weighted distances $D_0'$ and $D_1'$ (normalized to 1.0) from $D_0$ to $y_0$ and $D_1$ to $y_1$ respectively were calculated. The D' values for all nodes were found to lie in the range from 0.35 to 0.45. Thus, $T_{C1} = 0.35$ was selected. If more noise is expected in the input data, $T_{C1}$ can be lowered. However, if the wrong class estimate is passed from the level-one classifier, this will be quite detrimental to performance. Thus, the use of a lower $T_{C1}$ threshold should be carefully considered. In subsequent tests, we verified that the same hierarchical tree structure of Figure 4 would be chosen from a significantly reduced set of 16 reference images (specifically 4 images in each quadrant). The $T_{C1}$ value was similarly found to be unchanged when this reduced set of training set images was used. This is useful to retain the reduced size training set advantages possible with JGRV features.
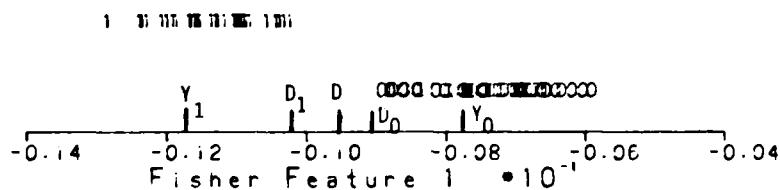


FIGURE 5
Projection of Subsets at Ship Node 2 Showing the Discriminant Point,
Projection Means and Threshold Region

## 4. EXPERIMENTAL RESULTS

Various aspects of the classifier were separately investigated. Each subsection below addresses one major issue of our moment-based classifier for the case of a ship image data-base. In each case, the test number is noted together with the salient conditions and the percentage of ships correctly classified. Each data entry in a table corresponds to 180 test images (case A = all) or 110 test images (case B = broadside views).

### 4.1 Effect of First-Level Estimator

In Tables 3 and 4, we show the results of tests performed with and without the first-level classifier enabled. As seen from the last column, excellent performance (above 98% correct classification) is obtained if the aspect estimator is used. This is expected since the second-level classifier does not provide aspect estimates and without this different ships at different aspect views have similar moments.

TABLE 3
Effect of First-Level Classifier on Performance (Broadside views, Case B)

| TEST NO. | TEST CONDITIONS | AVERAGE NUMBER OF REFERENCE VECTORS PASSED TO SECOND-LEVEL CLASSIFIER | PERCENT CORRECTLY CLASSIFIED |
|---|---|---|---|
| 1 | First-Level Not Used | 20 | 35.5 |
| 2 | First-Level Class Estimator Not Used | 8.12 | 36.3 |
| 3 | Aspect Estimator Not Used | 4.11 | 98.2 |
| 4 | Both Estimators Fully Used | 1.75 | 98.2 |

### 4.2 Computational Load with First-Level Classifier

In column 3 of Table 4, the number of reference vectors for which the second-level classifier must be tested is listed. There are a maximum of four aspect views in each of the five classes. These data correlate well with the percent of objects correctly classified. The fewer view classes passed to the second-level classifier, the better the system performs. In test 1, all 20 view classes are passed to the second-level classifier (i.e. all four aspect views of all five classes, since no first-level estimator was used). In test 2, with only the aspect estimator used, we might expect five view classes to be passed (the number of object classes). The larger average number of 8 view classes passed reflects the indecision in the aspect ratio test with the larger threshold of 1.5 used (versus passing only the best aspect estimate per class). In test 3, the aspect estimator is disabled and thus we might expect four view classes to be passed. This is close to the average number obtained. The data in Table 4 is quite comparable to that in Table 3 with only slightly lower percent correct performance obtained (due to the larger 180 versus 110 number of test images used and the low resolution of the bow and stern views now included).

TABLE 4
Effect of First-Level Classifier on Performance (All Image Views, Case A)

| TEST NO. | TEST CONDITIONS | AVERAGE NUMBER OF REFERENCE VECTORS PASSED TO SECOND-LEVEL CLASSIFIER | PERCENT CORRECTLY CLASSIFIED |
|---|---|---|---|
| 1 | First-Level Not Used | 20 | 36.7 |
| 2 | First-Level Class Estimator Not Used | 7.24 | 37.2 |
| 3 | Aspect Estimator Not Used | 4.78 | 86.7 |
| 4 | Both Estimators Fully Used | 1.72 | 86.7 |

The first-level estimator is thus useful to reduce the number of view classes to be processed by the second-level classifier and hence the computational load on the system. The aspect estimator is the most important part of the first-level classifier, because of the nature of the second-level classifier. In general, if the first-level classifier does not perform well, the second-level classifier cannot improve performance. In the tests performed in Tables 3 and 4, a convergence threshold $T = 10^{-4}$ was used and the reference set was the standard one in Table 2.

### 4.3 Convergence of the Second-Level Classifier

In this test, we consider the number of iterations necessary in the second-level classifier

for convergence for different thresholds T. The results (Table 5) show that effectively the same performance (98.2%) correct results for different convergence thresholds T was obtained. For the case of all ship images (Case A versus Case B), a nearly constant 87% correct class performance was obtained. As expected, the number of iterations required for (5) to converge to the specified T decreases as T increases. In no case are more than 20 iterations necessary however. Several modification details associated with starting the algorithm and choosing the step size were incorporated to insure such convergence. Other refinements in the step size choices in (3) can reduce the number of iterations in the second-level classifier by a factor of two (for the databases tested thusfar).

TABLE 5
Effect of Convergence Threshold T on the Number of Second-Level Class Iterations
(Case B, Broadside Test Images)

| TEST NO. | CONVERGENCE THRESHOLD T | PERCENT CORRECTLY CLASSIFIED OUT OF 110 | NO. OF SECOND-LEVEL ITERATIONS PER VIEW CLASS i |
|---|---|---|---|
| 1 | $10^{-4}$ | 98.2 | 17.04 |
| 2 | $10^{-3}$ | 98.2 | 16.00 |
| 3 | $10^{-2}$ | 98.2 | 14.77 |
| 4 | $10^{-1}$ | 98.2 | 13.30 |
| 5 | 0.5 | 98.2 | 2.0 |
| 6 | 1.0 | 98.2 | 2.0 |

#### 4.4 Number of References in the Second-Level Classifier

In the prior data, only four reference views per class (all in one quadrant) were used and excellent 98% (Case B) or 87% (Case A) correct performance was obtained. In Table 6, we consider the performance obtained when more aspect reference views per class were used in the second-level classifier. Tests 3 and 4 employ all 18 aspect views. The results shown are as expected. The excellent original performance 98% and 86.7% were improved by only 1-4% by increasing the number of aspect reference images per object class by a factor of 4.5 (from 4 to 18). In tests on other images [4] with less symmetry, poorer performance resulted unless reference images in two quadrants were used in the reference set for the second-level classifier. Thus, the exact results obtained depend upon the data and its symmetry. In general, a reduced size reference set can be used. If the number of aspect references is reduced, the accuracy in the aspect angle estimate may also be reduced. For the cases considered, interpolation between different aspect views is possible to provide view angle estimates with 10° accuracy using a reduced reference set. The sign of an odd-order moment can provide quadrant information on the aspect of an unknown test input object.

TABLE 6
Effect of Reference Set Size in the Second-Level Classifier

| TEST NO. | REFERENCE SET | PERCENT CORRECT (OUT OF 110 & 180) |
|---|---|---|
| 1 | 10°,30°,50°,80° | 98.2 |
| 2 | 10°,30°,50°,80° | 86.7 |
| 3 | Even Aspect Views | 99.1 |
| 4 | Even Aspect Views | 91.1 |

#### 4.5 Weighting Matrix Estimates

The final test run concerned the weighting matrix $\underline{\Sigma}$ used in the second-level classifier. The choices considered were $\underline{I}$ and $\underline{W} \, \underline{W}^T$ with $\underline{W}$ calculated from the two dominant Fisher vectors or from the four dominant Fisher vectors for all target views or only the broadside views. The results show that over 90% correct recognition was obtained with only the identity matrix used for the approximation to $\underline{\Sigma}$. Use of the full four Fisher vectors gave only 2% better performance. In all earlier data tests shown, the identity matrix was employed as an approximation to $\underline{\Sigma}$.

## 5. SUMMARY AND CONCLUSION

A new two-level classifier has been described that uses the geometrical moments as the feature set. These features are JGRVs and thus allow use of a Bayesian classifier with only one training set image per view class required. A nonlinear iterative algorithm is used in the second-level classifier to obtain the final class estimate and object distortion parameters. To reduce the number of view classes to be searched, first-level aspect and class estimators are used. The aspect estimator simply employs the ratio $L_{20}/\mu_{02}$ to select only views with a similar aspect ratio. An organized hierarchical tree search is used to obtain class estimates. Multi-class Fisher projections are used to define the nodes in the tree and two-class Fisher vectors are used to determine the subset at each node during testing. In all cases, the computational load is quite low: the first-level classifier requires only several vector inner products, the second-level classifier requires approximately 18000 operations per iteration and fewer than 15 iterations per view class. Thus, a quite efficient and attractive feature-space object classifier results with excellent performance (over 90% correct recognition) for a five-class problem with aspect view object distortions present. All parameters of the classifier have been examined and quantified for a ship image database.

### ACKNOWLEDGMENTS

### REFERENCES

1. A.P. Reeves and R.R. Seban, Proc. of 15th Annual Hawaii International Conference on System Sciences, pp. 388-396 (1982).
2. D. Casasent, R.L. Cheatham and D. Fetterly, Soc. Photo-Opt. Instr. Engrs., Vol. 360, pp. 105-111 (August 1982).
3. D. Casasent, R.L. Cheatham and D. Fetterly, Applied Optics, Vol. 21, pp. 3292-3298 (September 1982).
4. D. Casasent and R.L. Cheatham, Proc. Am. Soc. Mech. Engrs., Las Vegas, NE, August 1984.
5. K. Fukunaga, Introduction to Statistical Pattern Recognition, John Wiley and Sons, NY, (1972).
6. D. Casasent, J. Pauly and D. Fetterly, Soc. Photo-Opt. Instr. Engrs., Vol. 302, pp. 126-133 (August 1981).
7. R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, John Wiley and Sons, New York (1973).

# 9. REAL DATA TESTS AND SEGMENTATION ALGORITHM

Ref. 13

# IMAGE SEGMENTATION AND REAL-IMAGE TESTS
# FOR AN OPTICAL MOMENT-BASED FEATURE EXTRACTOR

David CASASENT and R. Lee CHEATHAM

*Carnegie-Mellon University, Department of Electrical and Computer Engineering,*
*Pittsburgh, PA 15213, USA*

Received 10 May 1984

A two-level classifier has been designed for use in a moment-based hybrid optical/digital processor. The simulation performance of this pattern recognition system using real IR input test images of ships and reference moments obtained from ship models is described with emphasis given to the preprocessing operations required.

## 1. Introduction

The use of optical processors to compute image features for feature-based pattern recognition has recently received renewed interest. The optically-computed image features thus far considered include Fourier coefficients [1–3], chord histogram distributions [4,5], and geometrical moments [6–8]. In this paper, a moment-based feature extractor and classification algorithm for pattern recognition is considered (section 2) and its performance in the classification of ship imagery (section 3) is addressed. Specific attention is given to classification of real input imagery (section 5) and the image preprocessing required (section 4).

## 2. Optical computation of the geometrical moments

The optical system considered to generate the moments of an input object [7] consists of an input plane $P_1$ (in which the input image is placed) imaged onto a moment generating mask at plane $P_2$. The monomials $x^p y^q$ up to fifth-order ($p + q \leqslant 5$) are recorded on the $P_2$ mask each spatially multiplexed using a different spatial frequency for each carrier. The optical Fourier transform of the light distribution leaving $P_2$ is detected on 21 multiple parallel output detectors in the $P_3$ output plane and contains the moments

$$m_{pq} = \iint f(x,y) x^p y^q \, \mathrm{d}x \, \mathrm{d}y \qquad (1)$$

of the $P_1$ input pattern $f(x,y)$ as detailed in [7].

These optically-generated image features are used as inputs to a digital feature-based classifier which then determines the object class and the orientation, scale and aspect of the input object. The details of this classifier are provided elsewhere [8] and are not germaine to our present discussion, however several remarks on the classifier follow for completeness. The optically-calculated input moment vector $\hat{m}$ is projected by the first-level classifier in the digital section onto a multi-dimensional Fisher feature space [9]. From the location of the projection vector, initial estimates of the input object class are made. From the ratio of the normalized second-order moments $\mu_{20}$ and $\mu_{02}$, an estimate of the aspect ratio or aspect angle of the input object is made. These estimates are used to select reference vectors $m_i(\theta)$ for class $i$ and aspect $\theta$ from storage against which $\hat{m}$ is compared. The final decision on the object class and the geometrical location of the input object is made in a second-level classifier implementing a nonlinear least-squares solution as detailed in [8]. Our present concern is the preprocessing required on real images before their moments $\hat{m}$ can be reliably extracted.

## 3. Database

As our reference database we used 180 images of
five types of ships with 36 images available per ship
(at 10° intervals around each ship, from a 90° depres-
sion angle). This reference database was obtained from
ship models under controlled conditions. Each image
contains 128 × 32 pixels with about 2000 pixels cor-
responding to the ship (for the broadside view) and
less than 200 ship pixels (for the bow and stern views).
The moments of 4 images per class (10°, 30°, 50° and
80°, where 0° is the bow view and 90° is the broadside
view) constituted our reference $m_i(\theta)$ database. As test
data, we used various real images of the class 2 ship
(the Leahy). A typical image is shown in fig. 1. It
shows the ship in water with a sky and shoreline back-
ground. We used 256 × 128 pixel images with 8 bits
of gray scale for the real ships in our tests. The hori-
zon (separating the water and the sky background) is
seen and the depression viewing angle for the real
images is 80° (rather than 90°, as in the reference
imagery). The real image (from bottom to top) con-
tains four regions: (1) water, (2) the hull of the ship
and some water, (3) the superstructure of the ship with
a water background, and (4) the sky and shoreline at
the top of the image. In section 4, we detail the pre-
processing used to extract the ship from the back-
ground and in section 5, we discuss the classification
performance obtained on such imagery.

## 4. Image preprocessing

Feature-extraction pattern recognition algorithms
require that one object location within the input field-



Fig. 1. Typical ship test image (the guided-missile cruiser, the
Leahy, ship class 2).
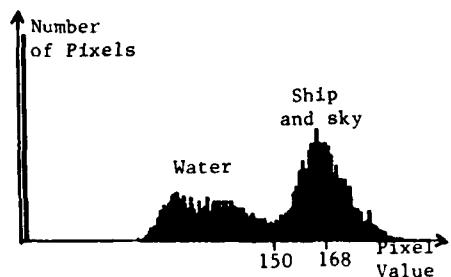
Fig. 2. Bimodal gray-level histogram of fig. 1.

of-view be extracted before the features are computed.
These operations are most commonly referred to as
segmentation and also involve noise removal and filling
in of holes on the object [10]. Care should be taken
to employ only simple image preprocessing operations
that are not computationally expensive. Thus, we used
mainly histogram operations (since they require only
simple tallies of image pixel levels) to aid in threshold
selections. A wealth of such methods exist, but their
specific implementations are quite problem-dependent.
In our case, we used context information (the water is
below the ship, the sky is above the ship and the deck
line and horizon are nearly horizontal due to the sen-
sor system used) to greatly simplify the ship segmenta-
tion. Our approach is quite novel in the techniques
employed to select separate thresholds for the differ-
ent image regions and dynamically select these regions
based on the scene information. Such methods are of
use in feature extractors for diverse applications.

As step 1, we formed the gray-level histogram of
fig. 1 (see fig. 2). It was bimodal as expected extending
from 0 to 255 (8 bits). A broad peak exists at low
pixel values (corresponding to the water and noise,
which is low in intensity in fig. 1) and a sharper peak
is centered at the high 175 pixel level (corresponding
to the ship and the sky, whose pixel values are larger
in fig. 1). A well-defined valley at pixel level 150 exists.
Thus, at step 2, we thresholded the image at 150 (with
all pixel values below 150 set to zero and all pixel val-
ues above 150 set to one). The resultant binary image
is shown in fig. 3.

At step 3, the image in fig. 3 is used to estimate the
location of the four image regions defined in section
3. To achieve this, a horizontal or row-projection
histogram of fig. 3 is formed. This is a graph (fig. 3) of

Fig. 3. Binary version of fig. 1 thresholded from the bimodal gray-level histogram of fig. 2.



Fig. 5. Gray-level histogram of the gray-scale image in fig. 1 after subtraction of the means of the sky and water from the appropriate image rows.

the number of pixels with value equal to one in each row of fig. 3. From fig. 4, the different image regions can be identified. The region to the right of row C (with zero-valued pixels) is the water below the ship. The flatter region just to the left of row C is the hull. The region between row B and where the hull occurs contains the ship's superstructure (plus water background). The sky and shoreline lie in the region to the left of row A. Between rows A and B is a transition region between the sky and water which contains the horizon region with some sky, water and ship superstructure. Row A and C are easily defined and located. Row B was located from the sum of first differences for consecutive row values as the inflection point in the histogram. These procedures are all automated and require only simple computations.

In step 4, the values for rows A, B and C from fig. 4 are used to extract the sky only (top row to row A)

and water only (row A to the bottom row) region of the *original gray-scale image*. Specifically, the average pixel values in these two image regions are calculated. This involves only a simple sum of the pixel levels in the proper rows of fig. 1. In step 5, the mean-value of the sky and shoreline region is subtracted from the rows above A in fig. 1, the mean value of the water region is subtracted from the rows below C in fig. 1, and a linear combination of the mean of the water and sky is subtracted from the rows between A and B. This produces an image with the ship pixels on a positive bias and with the water and sky regions on a zero bias. In step 6, the gray-level histogram of this image is formed. As shown in fig. 5, it has an obvious bimodal structure with a very apparent threshold level or valley point at pixel value $V_T$.

In step 7, all pixels in the image with gray-level values below $V_T$ in fig. 5 are set to zero. This removes the



Fig. 4. Horizontal projection histogram of the binary image of fig. 3. The sky, ship, superstructure and water regions of the image are noted.
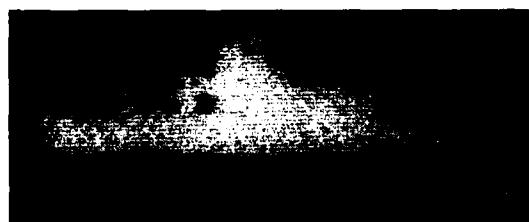


Fig. 6. Segmented ship image produced using the threshold level $V_T$ found from fig. 5.

229

sky, shoreline and water and thus extracts the ship. If the gray-levels above $V_T$ are retained, a gray-scale segmented ship image results. If levels above $V_T$ are set to unity, a binary segmented ship image results (fig. 6). Simple median filtering or other local convolution operations can be used to suppress miscellaneous noise pixels remaining in the background and to fill in holes on the target object.

## 5. Image classification

The moments $\hat{m}$ of the image in fig. 6 were computed and fed to our digital first-level Fisher projection class estimator. This first-level classifier omitted class 1 and 3 ships as possible class matches. The second-level classifier returned class 2 as the most-likely object class. This classifier also provides confidence levels for each possible ship class (classes 2, 4 and 5) passed by the first-level classifier. The class 4 ship, another guided-missile cruiser, had the second-best confidence but it was quite worse than that of the best (and correct) class 2 match. The correct aspect angle (70°) and scale (50%) of the input object are also provided by the classifier.

## 6. Summary and conclusion

A necessary aspect of feature extractors for pattern recognition is the image preprocessing required. A novel digital segmentation preprocessing procedure of

quite general use was detailed for a ship pattern recognition scenario. Such operations are essential if optical or digital feature extraction processors are to achieve good performance. The successful classification of a real input image using moment features and a unique two-level classifier was demonstrated. Similar results were obtained for other real images.

## Acknowledgement

## References

[1] G.G. Lendaris and G.L. Stanley, Proc. Inst. Electrical and Electronic Engrs. 58 (1979) 198.
[2] H. Kasdan and D. Mead, Proc. Elec. Opt. Sys. Des. Conf. 248 (1976).
[3] D. Casasent and V. Sharma, Opt. Eng. 23 (1984).
[4] D. Casasent and W-T. Chang, Appl. Opt. 22 (1983) 2087.
[5] D.G. Nichols, Optics Comm. 43 (1982) 168.
[6] J.A. Blodgett, R.A. Athale, C.L. Giles and H.H. Szu, Optics Lett. 7 (1982) 7.
[7] L. Cheatham, D. Casasent and D. Fetterly, Appl. Opt. 21 (1982) 3292.
[8] D. Casasent, J. Pauly and D. Fetterly, Proc. Soc. Photo-Opt. Instr. Engrs. 302 (1981) 126.
[9] R. Duda and P. Hart, Pattern analysis and scene classification (John Wiley and Sons, New York, 1973).
[10] W.K. Pratt, Digital image processing (John Wiley and Sons, New York, 1978).

# 10. ACCURACY OF DISTORTION-PARAMETER ESTIMATION IN OUR HIERARCHICAL MOMENT-BASED CLASSIFIER

Ref. 14

## Hierarchical Feature-Based Object Identification

David Casasent and R. Lee Cheatham*
Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

*Present Address: Battelle Northwest, Computers & Information Systems
Section, Richland, Washington

### 1. INTRODUCTION

A feature space processor for multi-class distortion-invariant pattern recognition is detailed in Section 2. A moment feature vector space is considered. Test data [1,2] on a robotic database are summarized in Section 3. Results on a ship database, using real input imagery with references from models is presented with attention to preprocessing, distortion parameter estimation, and class identification are advanced in Section 4.

### 2. PROCESSOR

A moment feature space is easily generated optically [3,4,5] or digitally [6]. Its outputs can easily be corrected for processing errors in post-processing [3]. Moments are jointly Gaussian random variables [2] due to sampling with respect to in-plane distortions. Thus, they allow use of a Bayesian classifier and thus can minimize $P_e$. To determine the class i (object class c and aspect view $\phi$) and the object's distortions (described by a distortion parameter $\underline{b}$) for each computed input moment vector $\hat{\underline{m}}$, we calculate

$$g_i = [\hat{\underline{m}} - \underline{m}_i(\underline{b})]^T \underline{\Sigma}^{-1} [\hat{\underline{m}} - \underline{m}_i(\underline{b})], \tag{1}$$

with $\underline{b}$ calculated iteratively (k is the iteration index) using

$$\underline{b}^{k+1} = \underline{b}^k + [(\underline{J}^k)^T \underline{\Sigma}^{-1} \underline{J}^k]^{-1} (\underline{J}^k)^T \underline{\Sigma}^{-1} [\hat{\underline{m}} - \underline{m}_i(\underline{b})]. \tag{2}$$

The class i that minimizes (1) defines c and the out-of-plane rotation angle (aspect) $\phi$ of the input, whereas $\underline{b}$ provides estimates of translations, scales, and in-plane rotations. The number of iterations k can be reduced to 4-6, $\underline{\Sigma} = \underline{I}$ can be used in (1) and (2), and $\underline{J}$ in (2) calculated as an update [1,2]. This significantly reduces the computational load per class/aspect i.

The major problem is the large number of aspect-classes i that need potentially be searched. To relieve this, we use two first-level estimators [1,2] to estimate the aspect (this is achieved by $A = \hat{\mu}_{20}/\hat{\mu}_{02}$) and class (a hierarchical tree is used for this, with the node structure chosen from a multi-class Fisher projection and with a two-class Fisher discriminant vector used per node). As we show in Section 3, this reduces the number of aspect-classes i to be searched and thus makes the processor very computationally efficient. A block diagram of the system is shown in Figure 1.

### 3. PIPE PART TEST RESULTS

Nine different pipe parts (4 classes) viewed from a 50° depression angle
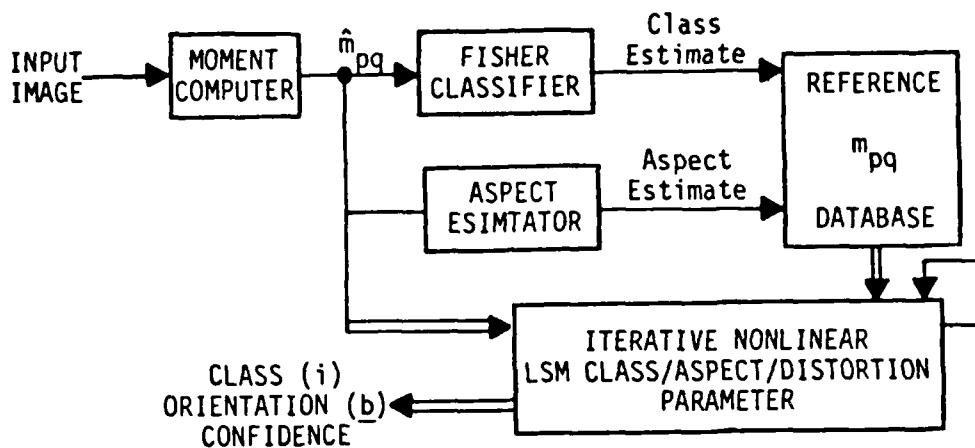
**Figure 1:** Block diagram of a multi-level moment feature-space classifier

were digitized (128 x 128 pixels) with 36 images per part (one image every $10°$ in aspect) and used as our test database. Test results are summarized in Table 1. They show: 9 out of 36 references are adequate (Test 1). Use of the first-level estimator reduces the number of i to be searched in (1) to 10 (Test 2). The number of iterations k in (2) is only 6 over a large $\Delta g_i$ range (Test 3) and $\Sigma = I$ in (1) and (2) is adequate (Test 4). As seen in Table 1, the system of Figure 1 can correctly classify over 97% of the 324 images (using only 9 x 4 = 36 references).

| TEST NUMBER | CONDITIONS | PERCENT CORRECT (OUT OF 324) | REMARKS |
|---|---|---|---|
| 1 | No Aspect Estimator | 97.5% | 9 Aspect Refs each 40° Used 24 View-Class (Avg) Passed |
| 2 | Full First-Level Estimator | 97.5% | 10 View-Class (Avg) Passed |
| 3 | $\Delta g_i = 10^{-4}$ to $10^{-1}$ | 98.2% more refs | 6 Iterations k |
| 4 | Different $\Sigma$ | 90-93.9% | $\Sigma = I$ (90%) Adequate |

**Table 1:** Representative Pipe Part Data (Different Test Conditions)

## 4. DISTORTION PARAMETER ESTIMATION ACCURACY

Related tests on another database [2,7] showed comparable performance and similar operational parameters. In this database, the reference objects were obtained from models and in tests against real-world IR images, excellent recognition was obtained. The preprocessing required [7] used only simple 1D and 2D histogram operations and thresholding (to maintain low computational overhead).

We now consider the class c, aspect $\phi$, scale $\alpha$ and translation $x_0$ estimation

accuracy of the system for a second five-class database (36 images at $10°$ aspect intervals per clas) using only four references per class. The true object was the $80°$ aspect view of the class 1 image. A real IR input image (vs. references obtained from models) at a depression angle $10°$ different from that of the reference set was used with real IR noise present in the input. The tests (Table 2) show perfect class and aspect classification for $\Delta g_i = 10^{-4} - 10^{-1}$ (for $\Delta g_i = 0.5$, errors resulted as expected) and excellent shift ($x_0$ in pixels) and scale factor ($\alpha$) distortion parameter estimation. All distortion parameters were estimated within 5% accuracy, due to the input resolution, noise, etc. factors.

| TEST NUMBER | TRUE SCALE $\alpha$/PIXEL SHIFT $x_0$ | $\alpha/x_0$ ESTIMATE | CLASS/ASPECT ESTIMATE |
|---|---|---|---|
| 1 | 1.0/0 | 1.0/0 | 1/80° |
| 2 | 1.0/15 | 1.016/14.22 | 1/80° |
| 3 | 1.0/25 | 1.023/23.22 | 1/80° |
| 4 | 0.5/0 | 0.499/0.1 | 1/80° |
| 5 | 0.75/0 | 0.750/0.07 | 1/80° |
| 6 | 0.9/0 | 0.90/0.03 | 1/80° |

Table 2:   Results of Class and Distortion Estimation Tests
(True Class 1, Aspect 80°)

## REFERENCES

1. D. Casasent and R.L. Cheatham, Proc. ASME, August 1984.

2. R.L. Cheatham and D. Casasent, Proc. SPIE, 504, August 1984.

3. D. Casasent, R.L. Cheatham and D. Fetterly, Applied Optics, 21, 3292, September 1982.

4. K. Wagner and D. Psaltis, Proc. SPIE, 352, 82, August 1982.

5. J.A. Blodgett, R.A. Athale, C.L. Giles and H.H. Szu, Optics Letters, 7, 7, 1982.

6. A.P. Reeves and R.R. Seban, Proc. of 15th Annual Hawaii International Conference on Systems Sciences, pp. 388-396, 1982.

7. D. Casasent and R.L. Cheatham, "Image Segmentation and Real-Image Tests for an Optical Moment-Based Feature Extractor", Optics Communications, Submitted, April 1984.

# 11. PROJECTION SDF SUMMARY: SHIP DATABASE CASE STUDY

Ref. 15

# Projection synthetic discriminant function performance

**David Cassent**
**William Rozzi**
Carnegie-Mellon University
Department of Electrical
    and Computer Engineering
Pittsburgh, Pennsylvania 15213

**Donald Fetterly**
General Dynamics
Pomona Division
P.O. Box 2507
Pomona, California 91769

**Abstract.** Synthetic discriminant functions (SDFs) allow distortion invariance to be achieved in optical correlators, thus making such systems more practical. The synthesis of four different types of SDFs is reviewed. Their performance in extensive projection tests on 144 images is presented, together with initial performance tests in the presence of noise. This distortion-invariant and shift-invariant pattern recognition algorithm can also be implemented digitally.

## CONTENTS

## 1. INTRODUCTION

The frequency plane optical matched spatial filter (MSF) correlator[1] has been the most studied optical pattern recognition system of the last 20 years. The sensitivity of the MSF to geometrical distortions between the input and reference object is a well-known shortcoming of any correlator. The use of multiple MSFs can reduce such problems at the cost of increased system complexity and reduced light budget efficiency.[2] Special frequency plane weighting and filter synthesis techniques can reduce this sensitivity, but cannot overcome it.[3] Space-variant correlators[4] and coded-phase processors[5] can overcome various distortions, but at the expense of shift invariance and multitarget recognition (although they still retain the processing gain advantages of a correlator).

In this paper, we describe the synthesis and performance of MSFs formed from synthetic discriminant functions (SDFs). These linear combination filters retain the shift invariance and processing gain of correlators while overcoming their sensitivity to geometrical distortions. In Sec. 2, we review the synthesis techniques for four different SDFs.[6] In Sec. 3, we describe the data base used in our projection simulations. Our noise-free results[7] are summarized in Sec. 4, and initial performance in the presence of noise[8] is presented in Sec. 5. Other variants of SDFs also exist[6,9–11]; however, SDFs are the most general and widely tested of such filters.

## 2. SDF SYNTHESIS

A unified SDF synthesis technique was first advanced in Ref. 12 and recently was more fully described.[6] For completeness, we briefly review the different SDF synthesis methods. The basic concept of SDF synthesis is to utilize a training set of images of each object class. From the correlation matrix of the full training set, we synthesize a SDF $h(x, y)$ that is a linear combination of the training set of images. Depending upon the purpose of the filter, different conditions will be placed on h, and different SDFs can be synthesized.

The simplest derivation of a SDF occurs for the case of one filter h that is to yield a constant correlation output $c = 1$ for all versions $\{f_n\}$ of objects f in one class; i.e.,

$$f_n \circledast h = c = 1 , \tag{1}$$

where $\circledast$ denotes correlation. We restrict h to being a linear combination of the $\{f_n\}$; i.e.,

$$h = \sum a_m f_m . \tag{2}$$

For notational simplicity, we do not show the spatial dependence of the functions f and h in Eqs. (1) and (2). We rewrite Eq. (1) for the projection case (i.e., the central value of the correlation output) as $f_n \cdot h = 1$ (where vectors, denoted by boldface type, now describe each function). Substituting Eq. (2) into Eq. (1), we find

$$f_n \cdot h = f_n \cdot \sum a_m f_m = \sum a_m r_{nm} = 1 , \tag{3}$$

where $r_{nm}$ denotes the elements of the correlation matrix R for $\{f_n\}$. In matrix-vector notation, Eq. (3) becomes

$$\mathbf{R a = u} , \tag{4}$$

where u is the unit vector (i.e., u contains all "1" elements). The filter h that satisfies Eq. (1) is thus defined by

$$\mathbf{a} = \mathbf{R}^{-1}\mathbf{u} . \tag{5}$$

We denote the SDF in Eq. (1), defined by Eq. (5), as an equal correlation peak (ECP) SDF. It is only of use in intraclass pattern recognition. The remaining types of SDFs[6] are described quite similarly to Eq. (5), with different training sets, correlation matrices, and exogenous vectors u used.

For both intraclass recognition and interclass discrimination, a SDF $\mathbf{h}_i$ is chosen to recognize objects $\{f_i\}$ in class i with unit output and to yield zero output for objects $\{f_j\}$ in other classes j; i.e.,

$$f_k \circledast h_i = \delta_{ik} . \tag{6}$$

This mutual orthogonal function (MOF) SDF for a two-class problem is synthesized using a training set of $N_1$ images $\{f_{1n}\}$ of class 1 objects and $N_2$ images $\{f_{2n}\}$ of class 2 objects as

$$h_1 = \sum_m a_m f_m ,$$

$$h_2 = \sum_m b_m f_m . \tag{7}$$

where $f_m \cdot h_1 = 1$ for objects in class 1, $f_m \cdot h_1 = 0$ for objects in class 2, and the summations in Eq. (7) are over $1 \leq m \leq N_1 + N_2$. The matrix-vector solutions for the $a_m$ and $b_m$ in Eq. (7) are

$$\mathbf{a} = \mathbf{R}_{1,2}^{-1}\mathbf{u}_1 ,$$

$$\mathbf{b} = \mathbf{R}_{1,2}^{-1}\mathbf{u}_2 . \tag{8}$$

where $\mathbf{R}_{1,2}$ is the $(N_1 + N_2) \times (N_1 + N_2)$ correlation matrix of the full training set of data, and where $\mathbf{u}_1 = [1,\ldots,1,0,\ldots,0]^T$ and $\mathbf{u}_2 = [0,\ldots,0,1,\ldots,1]^T$. In $\mathbf{u}_1$, there are $N_1$ ones (for the $N_1$ class 1 objects) and $N_2$ zeros (for the $N_2$ class 2 objects). Similarly, $\mathbf{u}_2$ has $N_1$ zeros and $N_2$ ones. The extension of Eqs. (6), (7), and (8) to more than two object classes follows directly.[6] These MOF SDFs require one filter per object class and a correlation matrix R of larger order than in Eq. (5).

In some recognition cases, a single multilevel nonredundant filter SDF h (we use the simpler term "multilevel SDF") can be used to recognize multiple object classes. The general requirement for such a filter can be written as

$$f_n \circledast h = n . \tag{9}$$

i.e., the value n of the correlation output defines the class n of the input object. For a three-class intraclass and interclass recognition and discrimination problem, we use $N_1$, $N_2$, and $N_3$ training set images in classes 1, 2, and 3, respectively. The filter is defined by

$$h = \sum_m a_m f_m , \tag{10}$$

where the summation is over $1 \leq m \leq N_1 + N_2 + N_3$, i.e., the full training set. The filter in Eq. (10) satisfying Eq. (9) is defined in matrix-vector notation by

$$\mathbf{a} = \mathbf{R}_3^{-1}\mathbf{u}_3 , \tag{11}$$

where $\mathbf{R}_3$ is the $N_1 + N_2 + N_3$ correlation matrix for the full training set and $\mathbf{u}_3 = [1,\ldots,1,2,\ldots,2,3,\ldots,3]$ has $N_1$ ones, $N_2$ twos, and $N_3$ threes.

The final class of SDFs is the K-tuple two-level nonredundant multiple filter SDF (we use the simpler term "K-tuple SDF"). We describe such filters for the four-class (N = 4) two-filter (K = 2) case (i.e., $2^K = N$). The four object classes are denoted by $\{f_1\}$, $\{f_2\}$, etc., and the two filters by $h_a$ and $h_b$. The object class is determined from the outputs from *both* filters, as in Table 1. For simplicity, binary (0, 1) values are used. Other values are preferable since (0,0) can also

### TABLE I. Truth Table for K-tuple Two-Level Four-Class SDF

| Input/Output | $h_a$ | $h_b$ |
|---|---|---|
| $\{f_1\}$ | 0 | 0 |
| $\{f_2\}$ | 0 | 1 |
| $\{f_3\}$ | 1 | 1 |
| $\{f_4\}$ | 1 | 0 |

indicate no input. In practice, we select K to satisfy $2^K \geq N + 1$ and thus avoid the ambiguity possible in the (0,0) output case.

Synthesis of $h_a$ and $h_b$ to satisfy Table I follows directly. The two filters are linear combinations of the full training set of data:

$$h_a = \sum a_m f_m ,$$

$$h_b = \sum b_m f_m . \tag{12}$$

The coefficients in Eq. (12) are defined by

$$\mathbf{a} = \mathbf{R}_4^{-1}\mathbf{u}_{4a} ,$$

$$\mathbf{b} = \mathbf{R}_4^{-1}\mathbf{u}_{4b} , \tag{13}$$

where R is the correlation matrix of the full training set of images for all four classes ($N_1$, $N_2$, $N_3$, and $N_4$ images in each class, respectively). The vector $\mathbf{u}_{4a}$ in Eq. (13) has $N_1 + N_2$ zeros and $N_3 + N_4$ ones (for $h_a$), with $\mathbf{u}_{4b}$ being similar.

Other obvious combinations of these four basic types of SDFs follow directly. In all cases, the filter function is of the form in Eq. (5), with a different correlation matrix $\mathbf{R}_n$ and exogenous vector $\mathbf{u}_n$. This unified SDF synthesis method significantly simplifies off-line synthesis of the SDF. Since the projection values for the different classes in the different SDFs are fixed by the synthesis algorithm, we refer to such filters as deterministic SDFs.

## 3. DATA BASE

The proper evaluation of the performance of various SDFs described in Sec. 2 in this paper is a key new detail. SDFs require a large data base to properly select training set images that are valid statistical representations of the data in each object class, with a sufficient number of additional test images (not in the training set) remaining to allow sufficiently valid tests on the algorithm. In all experiments performed, the computational load was so large that only correlation plane projection values (i.e., the correlation value at the point of registration) were evaluated as in Eq. (3).

Our most extensive multiclass object data base available consisted of images of four ships from 90° depression angle with 36 views available per ship (at 10° intervals in aspect around the full 360° of the ship). In Fig. 1, we show the broadside views of the four ships. Clearly, other aspect views, such as the bow and stern, contain significantly less object data. The images were each 128 × 32 pixels. For the broadside views, the target contained approximately 1200 pixels out of 4000 pixels in the full frame. For the bow and stern views, about 200 pixels (out of 4000) were present on the target. The classes assigned to each ship and the name and type of each are noted in Table II. The images in class 1 are numbered 1 through 36 (1 is the bow, 18 is the stern, etc.). Class 2 images are numbered 37 through 72, etc. All images were binarized to only "0" and "1" valued pixels, with the threshold selected from simple histogram operations.[13] Two sets of images, one with and one without the hull present, were formed and used. This image data base allows the 3-D aspect distortion invariance of our SDF correlator to be verified and its performance to be quantified.

## 4. NOISE-FREE PROJECTION TEST RESULTS

In Table III, we summarize our digitally simulated SDF projection

performance obtained for the four major types of SDFs described in Sec. 2, using the first four ship image data bases described in Sec. 3. Six different tests using different SDFs were considered. The type of SDF used and the six training set images used per class are noted in the table. The number of errors obtained out of the 36 images in each of the four object classes is noted, together with the percentage correctly recognized.

In test 1, the ECP SDF in Eq. (5) was formed using only six class 1 images and was tested against all 36 aspect views of the class 1 object. All 36 projection values were within 3% of the deterministic value of unity selected in Eq. (5), and thus 100% correct performance was obtained. Test 2 was similar for the class 2 object, and again 100%



Fig. 1. (a)-(d) Broadside views of the data base images in classes 1 to 4, respectively.

**TABLE II. Ship Image Data Base Used**

| Ship class | Ship name | Type |
|---|---|---|
| 1 | Moskva | Soviet helicopter cruiser |
| 2 | Leahy | U.S. guided-missile cruiser |
| 3 | Hope | International hospital ship |
| 4 | Albany | U.S. guided-missile cruiser |

performance was obtained (with all 36 projection values within 5% of 1.0). These test 1 and 2 results show the intraclass recognition performance of SDFs in the face of 3-D aspect distortions. All SDFs require a training set that is a valid statistical representation of the object class for the distortions considered, and in this case 10° to 20° variation in aspect can be tolerated (in agreement with experiments in Refs. 2 and 3). The six training set images used per class in tests 1 and 2 were chosen at approximately 50° intervals around the ship, with three or four images taken from each side of the ship (0° to 180° and 180° to 360° aspect views).

In tests 3 through 6, the discrimination as well as recognition performance of our other SDFs was considered. The two-class MOF SDF defined by Eqs. (6) and (8) was evaluated first (test 3) using the 12 training set images in tests 1 and 2 to form the MOF SDF. The projection values used in the filter synthesis were $(1,0)$. In determining the object class, the projection values $P_j$ of each input object on only the $h_1$ MOF SDF in Eq. (7) were calculated, and the decision on the input object class was made based on whether $P_1 < 0.5$ or whether $P_1 > 0.5$ (where the 0.5 threshold level was chosen as being midway between the original 0 and 1 deterministic projection values). The projection performance obtained was excellent, with 69 of 72 images correctly classified (95.8% correct identification). In test 4, our multilevel SDF defined in Eqs. (9) and (11) was tested on three object classes with deterministic projection values of 0, 1, and 2, respectively, for the three classes. Six training set images per class were used, and excellent performance resulted, as shown, with only five errors obtained out of the 108 test images (95.4% correct performance).

In tests 1 through 4, the hull of the ship was present in the image data base used. Comparable results occur if the hull is not present. In tests 5 and 6, our K-tuple SDF was used for the full four-class recognition and discrimination testing on all 144 images. A new training set of six images per class was used (selected as described in Ref. 8) to synthesize the $h_a$ and $h_b$ filters, and the image data base with the hull of the ship removed was employed (since it yielded better performance due to more discriminatory information in the superstructure of the ship). In test 5, the filters were synthesized with the deterministic projection values noted in Table I. The results were quite good, with only 14 errors out of 144 images (90.3% correct recognition).

However, as shown, 12 of the 14 errors occurred for the class 4 object. The majority of these errors were at the bow and stern, and all of these errors were due to the projection values on the second filter $h_b$ being above the 0.5 threshold (recall from Table I that $h_b$ should force class 4 projections to 0, or to below 0.5). Inspection of Fig. 1 shows that the class 4 object is the largest ship. Since it appears to be more difficult to force the projection values of a large object to zero (compared to the ease of forcing the projection values of smaller objects to zero), we altered the projection value choices in Table I to those shown in Table IV.

As seen in Table IV, the new deterministic projection value choices have reversed the projection values for class 3 and 4 objects. Both filters ($h_a$ and $h_b$) are now designed to yield projection values of $(1,1)$ for the largest (class 4) object. The results for this filter are shown in Table III, test 6. They are excellent, with only two errors out of all 144 test images (98.7% correct classification). Attempts to

**TABLE III. Noise-Free SDF Projection Performance Test Results**

| Test number | Type of SDF | Training set | Errors per class 1 | 2 | 3 | 4 | Percent correct |
|---|---|---|---|---|---|---|---|
| 1 | ECP (class 1) | (1, 6, 10, 15, 20, 25) | 0 | – | – | – | 100 |
| 2 | ECP (class 2) | (38, 45, 50, 55, 60, 65) | – | 0 | – | – | 100 |
| 3 | MOF (1, 0) | (Same as tests 1 and 2) | 1 | 2 | – | – | 95.8 |
| 4 | Multilevel (0, 1, 2) | (Six images per class as above) | 2 | 0 | 3 | – | 95.4 |
| 5 | K-tuple (Table I) | New training set (Six images per class, hull removed) | 0 | 2 | 0 | 12 | 90.3 |
| 6 | K-tuple (Table IV) | New training set (Set images per class, hull removed) | 0 | 2 | 0 | 0 | 98.7 |

obtain comparable performance when the hull was present were not successful for the four-class problem (since the superstructure of the ship clearly contains the major discriminatory information). In general, the amount of the hull that is visible varies considerably with the ship's load, and thus the hull data cannot be reliably assumed to be present.

## 5. SDF PROJECTION PERFORMANCE WITH NOISE

In Table V, we summarize the performance of the K-tuple SDF in test 5 of Table III in the presence of noise. In tests 7 through 11, Gaussian, zero-mean noise was added only to the test set of images, and a noise-free training set of images was used. For the 24 images present in the training and test sets, noise was added only during testing. In tests 12 through 15, noise was present in the training set images also. The noise added was of zero mean, with a standard deviation $\sigma$ as given in the table. The noisy images were then binarized, and these binarized noisy images were used in filter synthesis and testing. For the data in Table V, noise was present everywhere (i.e., in the background and on the target). Because of the binarization, the effect of noise is different when present in the background and on the object. In the background, noise adds +1 valued pixels, and on the target it forces +1 valued pixels to 0. To stabilize the results obtained, we chose not to make a decision on the object class when the projection values were within ±0.03 of the threshold. These "no-decision" cases are indicated in parentheses. The total number of errors, total number of correctly classified objects, and percentage of objects correctly classified (out of the 144 test images used) are given in the table.

The results obtained require discussion. A reduction in the number of class 4 object errors with increasing noise was observed (tests 7 through 10), up to $\sigma = 0.4$ noise. Then, for $\sigma = 0.6$, the total number of errors increases dramatically. To explain this performance, we denote the projection value for filter $h_i$ and object class $j$ by $P_i(j)$. We note that all class 4 errors in test 7 are errors in $P_b(4)$ that are above 0.5 (whereas they should be 0.0). As noise is added to the class 4 object, the zero-valued pixels introduced on the object cause the $P_b(4)$ values to decrease. The +1 valued pixels introduced into the background also cause $P_b(4)$ to decrease (since much of the upper portion of the $h_b$ filter is negative-valued, as needed to force the full

projection to 0). Thus, as noise increases, the $P_b(4)$ projection values decrease. For $\sigma = 0.4$, all $P_b(4)$ projections are now below 0.5 and hence correct, and thus no class 4 errors occur.

However, the $P_b(2)$ and $P_b(3)$ projections for filter $h_b$ on class 2 and 3 objects also decrease as $\sigma$ of the noise increases. These projection values were intended to be 1.0, and initially all were quite close to 1.0. As $\sigma$ increases, their values decrease gradually, and at $\sigma = 0.6$ nearly all filter projections for class 2 objects are below 0.5. The decrease in projection value is quite gradual, and the sharp increase in the number of errors (from tests 10 to 11) occurs because most projections now pass below the threshold level at this $\sigma$ level of noise. The class 2 projection outputs all changed from $(0,1)$ to $(0,0)$; thus, all class 2 errors in test 11 were class 2 objects classified as class 1 objects. All of our test data used the same fixed-decision threshold level. Use of adaptive thresholds can significantly improve performance.

When noise was present in both the test and training sets (all noise is uncorrelated between images), performance remained rather stationary (12 to 14 errors) for $\sigma$ values up to 0.2 (tests 12 through 14). When the training set noise was 0.3 or larger, the number of errors increased significantly (from 14 in test 14 to 52 in test 15). At this noise level, the filters are simply not valid representations of the objects. Recall that the input signal-to-noise ratio (SNR) is different for each image since fewer object pixels exist and more background is present for aspect views further away from broadside views. The performance shown in Table V is still very impressive and should be adequate for most object identification applications over a significant range of sensor noise levels. The tests in Table V were repeated for six training set images per class, chosen at evenly spaced intervals of about every 50°, and comparable results were obtained. The tests in Table V were also repeated for the case when noise was present only in the background of the object (not on the object). Comparable results were obtained, with the number of errors changing slightly less dramatically (since the equivalent input SNR is better for a given $\sigma$ of the noise when noise is present only in the background of the object, rather than on the object and in the background).

As our final noise performance test, the K-tuple SDFs in test 6 (using the projection values in Table IV rather than Table I) were used with varying amounts of noise added to the background only (similar results were obtained when noise was present in the background and on the target) in both the test data and the training data. The results (Table VI) show the excellent performance expected of a correlator in the presence of noise.

## 6. SUMMARY AND CONCLUSIONS

The advantages of a correlator (processing gain, good performance in noise, shift invariance, or multiple object recognition) can be retained and the disadvantages (sensitivity to geometrical distortions) can be overcome by synthesizing the matched spatial filter from SDFs. These SDFs are linear combinations of the training sets

**TABLE IV. New Projection Value Choices Used in Test 6 of Table III**

| Input/Output | $h_a$ | $h_b$ |
|---|---|---|
| $\{f_1\}$ | 0 | 0 |
| $\{f_2\}$ | 0 | 1 |
| $\{f_3\}$ | 1 | 0 |
| $\{f_4\}$ | 1 | 1 |

**TABLE V. K-tuple SDF Performance (Table I Projection Values) in the Presence of Noise in the Training and/or Test Data\***

| Test number | Noise standard deviation ($\sigma$) | | Number of errors per class | | | | Total number of errors | Number correctly classified | Percent correctly classified |
|---|---|---|---|---|---|---|---|---|---|
| | Training | Testing | 1 | 2 | 3 | 4 | | | |
| 7 | 0.0 | 0.0 | 0(0) | 2(0) | 0(0) | 10(2) | 12 | 130 | 90.3 |
| 8 | 0.0 | 0.2 | 0(0) | 1(1) | 0(0) | 8(4) | 9 | 130 | 90.3 |
| 9 | 0.0 | 0.3 | 0(0) | 1(1) | 0(0) | 2(3) | 3 | 137 | 95.1 |
| 10 | 0.0 | 0.4 | 0(0) | 3(9) | 0(3) | 0(0) | 3 | 129 | 89.6 |
| 11 | 0.0 | 0.6 | 0(0) | 27(9) | 36(0) | 0(0) | 63 | 72 | 50.0 |
| 12 | 0.0 | 0.0 | 0(0) | 2(0) | 0(0) | 10(2) | 12 | 130 | 90.3 |
| 13 | 0.2 | 0.0 | 0(0) | 1(1) | 0(0) | 13(2) | 14 | 127 | 88.2 |
| 14 | 0.2 | 0.2 | 0(0) | 2(0) | 0(0) | 12(2) | 14 | 128 | 88.9 |
| 15 | 0.3 | 0.2 | 17(1) | 9(1) | 1(0) | 25(2) | 52 | 88 | 61.1 |

\*No-decision cases with projection values within ±0.03 of the threshold are noted in parentheses; noise everywhere, hull not present.

TABLE VI. K-tuple SDF Performance (Table IV Projection Values) in the Presence of Noise in the Training and/or Test Data*

| Test number | Noise standard deviation (σ) | | Number of errors per class | | | | Total number of errors | Number correctly classified | Percent correctly classified |
|---|---|---|---|---|---|---|---|---|---|
| | Training | Testing | 1 | 2 | 3 | 4 | | | |
| 16 | 0.0 | 0.0 | 0(0) | 2(0) | 0(0) | 0(0) | 2 | 142 | 98.7 |
| 17 | 0.0 | 0.2 | 0(0) | 2(0) | 0(0) | 0(0) | 2 | 142 | 98.7 |
| 18 | 0.0 | 0.3 | 0(0) | 2(1) | 0(0) | 0(0) | 2 | 141 | 97.9 |
| 19 | 0.0 | 0.4 | 0(0) | 3(3) | 0(0) | 0(0) | 3 | 138 | 95.7 |
| 20 | 0.2 | 0.0 | 0(0) | 2(0) | 0(0) | 0(0) | 2 | 142 | 98.7 |
| 21 | 0.2 | 0.2 | 0(0) | 2(0) | 0(0) | 0(0) | 2 | 142 | 98.7 |
| 22 | 0.3 | 0.0 | 0(0) | 0(0) | 1(1) | 6(0) | 7 | 136 | 94.4 |
| 23 | 0.3 | 0.2 | 0(0) | 0(0) | 1(1) | 5(1) | 6 | 136 | 94.4 |
| 24 | 0.3 | 0.3 | 0(0) | 0(0) | 1(1) | 5(1) | 6 | 136 | 94.4 |

*No-decision cases with projection values within ±0.03 of the threshold are noted in parentheses; noise in background only.

of imagery per object class. A unified formulation for four different types of SDFs has been reviewed, and their performance with and without noise has been quantified. Achieving optimum performance for such filters is a complicated data-dependent function of the training set images used and the projection values selected. Initial guidelines for such selections were advanced and verified by simulation experiments.

## 7. ACKNOWLEDGMENTS

## 8. REFERENCES

1. A. VanderLugt, IEEE Trans. Inf. Theory IT-10, 139 (1964).
2. K. Leib, R. Bondurant, S. Hsiao, M. Wohlers, and R. Herold, Appl. Opt. 17, 2892 (1978).
3. D. Casasent and A. Furman, Appl. Opt. 16, 1662 (1977).
4. D. Casasent and D. Psaltis, Proc. IEEE 65, 77 (1977).
5. J. R. Leger and S. H. Lee, Appl. Opt. 21, 274 (1982).
6. D. Casasent, Appl. Opt. 23 (1984).
7. D. Casasent and V. Sharma, in 10th International Optical Computing Conference, Sam Horvitz, ed., Proc. SPIE 422, 193 (1983).
8. D. Casasent and V. Sharma, in Robotics and Robot Sensing Systems, David Casasent and Ernest L. Hall, eds., Proc. SPIE 442, 47 (1983).
9. H. J. Caulfield, Appl. Opt. 19, 3877 (1980).
10. C. F. Hester and D. Casasent, Appl. Opt. 19, 1758 (1980).
11. J. Flenret and H. Maitre, Opt. Commun. 17 (1976).
12. D. Casasent, B. V. K. Vijaya Kumar, and V. Sharma, in Robotics and Industrial Inspection, David P. Casasent, ed., Proc. SPIE 360, 136 (1982).
13. D. Casasent, R. Cheatham, and D. Fetterly, in Computer Vision and Pattern Recognition 1983, p. 171.

# 12. OPTIMAL LINEAR DISCRIMINANT FUNCTION CONCEPTS

Ref. 16

# OPTIMAL LINEAR DISCRIMINANT FUNCTIONS

Vinod Sharma* and David Casasent

Carnegie-Mellon University
Deparment of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

*Present Address: Northeastern University, Department of Electrical and Computer Engineering

and D. Fetterly
General Dynamics - Pomona Division
P.O. Box 2507
Pomona, California 91769

## ABSTRACT

A new class of discriminant filter functions for use in a matched filter correlator for multi-class distortion-invariant pattern recognition is described. Three variations of these optimal linear discriminant functions (OLDFs) that optimize different performance measures are described and initial performance results are presented.

## 1. INTRODUCTION

Correlators represent a powerful class of pattern recognition architectures that allow multiple targets to be located and that provide excellent performance in noise. Optical systems [1,2] easily achieve the correlation operation in real-time and various compact versions of such systems have been fabricated [3] and discussed [4]. Correlators are well known to be quite susceptible to geometrical distortions between the input and reference object. A most attractive technique to achieve distortion-invariant correlation uses synthetic discriminant functions (SDFs) [5], projection SDFs [6,7], correlation SDFs [8], or related methods [9,10]. In general, these prior techniques achieved filter synthesis by forcing fixed projection values for training set images in two classes [5,7,9]. In [8], a least-squares solution and a solution that maximized the peak-to-sidelobe ratio plus the class 1 to class 2 outputs was employed. In [11], class 2 was treated as noise and SNR was maximized.

In this paper, we briefly review the 5 standard SDFs using a new notation (Section 2). We then describe three new optimal linear discriminant function (OLDF) filters that maximize different performance measures (Section 3). These filters differ from the original SDF and other work in that they are optimal (i.e. maximize various performance measures useful in discrimination pattern recognition). They are thus preferable, since predicted $P_D$, $P_{FA}$ and $P_e$ performance and noise effects on them should be able to be analyzed (theoretically and statistically) more easily. Initial simulation results, using only correlation plane projection values are advanced (Section 4) to demonstrate and quantify the intra-class recognition and inter-class discrimination performance of these OLDFs for multi-class cases and in the presence of noise.

## 2. SDFs

The 5 standard SDFs [5] are now reviewed for background and to introduce the new notation that is most appropriate for description of our OLDFs in Section 3. The inner product of vectors $\underline{x}$ and $\underline{y}$ is defined as $<\underline{x},\underline{y}> = \sum_{i=1}^{N} x_i y_i$, where the N elements of $\underline{x}$ and $\underline{y}$ are denoted by $x_i$ and $y_i$ and $\underline{x} = (x_1 \cdots x_N)^T$. The norm $||\underline{x}||$ of $\underline{x}$ is defined by $||\underline{x}||^2 = <\underline{x},\underline{x}>$. Consider four classes of objects with training set images

$$\{I_i\}_{i=1}^{n}, \{J_i\}_{i=1}^{m}, \{K_i\}_{i=1}^{p} \text{ and } \{L_i\}_{i=1}^{q}, \tag{1}$$

where there are n images in $\{I\}$ (class 1), etc.

An intra-class SDF (equal correlation peak SDF) $\underline{F}$ is defined such that

$$<I_i,F> = 1 \text{ for } i = 1 \cdots n. \tag{2}$$

A two-class mutual orthogonal function (MOF) SDF is defined such that

$$\langle I_i, F \rangle = 0 \text{ for } i = 1 \cdots n$$

$$\langle J_i, F \rangle = 1 \text{ for } i = 1 \cdots m, \tag{3}$$

with other projection values possible in (3). Two MOF SDFs $F_1$ and $F_2$ to recognize all versions of class 1 ($I_i$) and reject all versions of class 2 ($J_i$) and vice-versa for $F_2$ are defined such that

$$\langle I_i, F_1 \rangle = 0 \text{ and } \langle I_i, F_2 \rangle = 1 \text{ for } i = 1 \cdots n$$

$$\langle J_i, F_1 \rangle = 1 \text{ and } \langle J_i, F_2 \rangle = 0 \text{ for } i = 1 \cdots m. \tag{4}$$

A multi-level SDF for 4 classes results by choosing different vector-inner product forcing constants, i.e. for the 4 classes in (1), we can require

$$\langle F, I_i \rangle = 0, \quad \langle F, J_i \rangle = 1, \quad \langle F, K_i \rangle = 2, \quad \langle F, L_i \rangle = 3. \tag{5}$$

A K-tuple SDF for M classes uses K SDFs (where $2^K \geq M$) with binary valued projection forcing functions. For the 4 classes in (1), we require $K = 2$ SDFs ($F_1$ and $F_2$) defined such that

$$\langle F_1, I_i \rangle = 0, \quad \langle F_2, I_i \rangle = 0$$

$$\langle F_1, J_i \rangle = 0, \quad \langle F_2, J_i \rangle = 1$$

$$\langle F_1, K_i \rangle = 1, \quad \langle F_2, K_i \rangle = 0 \tag{6}$$

$$\langle F_1, L_i \rangle = 1, \quad \langle F_2, L_i \rangle = 1.$$

To obtain unique solutions F for these SDFs, we require them to lie in the subspace spanned by the training sets. This solution has the additional advantage that the projections of other objects not considered (for classification) will be minimized. Thus, $F_1$ etc. are linear combinations of the 4 training sets, e.g.

$$F_1 = \sum_{i=1}^{n} a_i I_i + \sum_{i=n+1}^{n+m} a_i J_{i-n} + \sum_{i=n+m+1}^{n+m+p} a_i K_{i-n-m} + \sum_{i=n+m+p+1}^{n+m+p+q} a_i L_{i-n-m-p} \tag{7a}$$

$$F_2 = \sum_{i=1}^{n} b_i I_i + \cdots . \tag{7b}$$

Using (7a), we write the vector-inner product of $F_1$ and $I_j$ as

$$\langle F_1, I_j \rangle = \sum_{i=1}^{n} a_i \langle I_i, I_j \rangle + \sum_{i=1}^{m} a_{i+n} \langle J_i, I_j \rangle + \sum_{i=1}^{p} a_{i+n+m} \langle K_i, I_j \rangle + \sum_{i=1}^{q} a_{i+n+m+p} \langle L_i, I_j \rangle . \tag{8}$$

Similar expressions can be obtained for the other vector inner products in (2)-(6). The $a_i$ and $b_i$ coefficients that define $F_1$ and $F_2$, for the SDF in (6), can be obtained by solving the following system of equations

$$
\begin{bmatrix}
I_1^t I_1 & I_2^t I_1 & \cdots & I_n^t I_1 & J_1^t I_1 & \cdots & J_m^t I_1 & K_1^t I_1 & \cdots & K_p^t I_1 & L_1^t I_1 & \cdots & L_q^t I_1 \\
\vdots & & & & & & & & & & & & \\
I_1^t J_1 & & & & & & & & & & & & L_q^t J_1 \\
\vdots & & & & & & & & & & & & \\
I_1^t K_1 & & & & & & & & & & & & L_q^t K_1 \\
\vdots & & & & & & & & & & & & \\
I_1^t L_1 & & & & & & & & & & & & L_q^t L_1 \\
\vdots & & & & & & & & & & & & \\
I_1^t L_q & \cdots & & & & & & & & & & & L_q^t L_q
\end{bmatrix}
\begin{bmatrix}
a_1 & b_1 \\
a_2 & b_2 \\
\cdot & \cdot \\
\cdot & \cdot \\
\cdot & \cdot \\
\cdot & \cdot \\
\cdot & \cdot \\
\cdot & \cdot \\
\cdot & \cdot \\
\cdot & \cdot \\
\cdot & \cdot \\
a_{n+m+p+q} & b_{n+m+p+q}
\end{bmatrix}
=
\begin{bmatrix}
0 & 0 \\
\vdots & \vdots \\
0 & 0 \\
0 & 1 \\
\vdots & \vdots \\
0 & 1 \\
1 & 0 \\
\vdots & \vdots \\
1 & 0 \\
1 & 1 \\
\vdots & \vdots \\
1 & 1
\end{bmatrix}
\begin{array}{l}
\left.\vphantom{\begin{matrix}0\\0\\0\end{matrix}}\right\} n \\
\left.\vphantom{\begin{matrix}0\\0\\0\end{matrix}}\right\} m \\
\left.\vphantom{\begin{matrix}0\\0\\0\end{matrix}}\right\} p \\
\left.\vphantom{\begin{matrix}0\\0\\0\end{matrix}}\right\} q
\end{array}
\tag{9}
$$

In (9), superscript t denotes transpose and $(\ )^t(\ )$ denotes a vector-inner product. If a unique solution does not exist, the least-squares solution (obtained by computing the generalized inverse of the matrix) is used. As noted at the outset, these SDFs are computed using fixed projection values for the various training set classes.

### 3. OLDFs

We can describe the new OLDFs as linear functionals f on the finite dimensional vector space of images. From the Riesz representation theorem [12], we can also describe these OLDFs by a discriminant vector $\underline{u}$, where $f(\underline{x}) = \langle\underline{u},\underline{x}\rangle$ for all $\underline{x}$ in our real linear vector space. The vector-inner product of two functionals $\underline{f}_1$ and $\underline{f}_2$ in the dual vector space and their corresponding $\underline{u}_1$ and $\underline{u}_2$ OLDFs are related simply by $\langle\underline{f}_1,\underline{f}_2\rangle = \langle\underline{u}_1,\underline{u}_2\rangle$. In all of our OLDFs, we consider only two-class problems $\{I_i\}$ and $\{J_i\}$ with n and m training set images respectively.

### 3.1 OLDF-1

As our first OLDF-1, we consider a version of the MOF SDF in which the projection for one input class is 0 whereas the projection for the other class is maximized (rather than being a fixed constant value of 1). Three types of maximizations were considered (corresponding to 3 cases (A,B,C) for OLDF-1). These are defined as finding OLDF-1 $\underline{u}$, such that

$$\underline{\text{CASE A:}} \quad \langle u,J_i\rangle = 0 \text{ for } i = 1\cdots m \tag{10a}$$

$$\langle u,I_i\rangle = \max_{\text{all } x}|\langle x,I_i\rangle| \text{ for } i = 1\cdots n \tag{10b}$$

$$\underline{\text{CASE B:}} \quad \langle u,J_i\rangle = 0 \text{ for } i = 1\cdots m \tag{11a}$$

$$\sum_{i=1}^{n}\langle u,I_i\rangle = \max_{\text{all } x}\sum_{i=1}^{n}|\langle x,I_i\rangle| \tag{11b}$$

$$\underline{\text{CASE C:}} \quad \langle u,J_i\rangle = 0 \text{ for } i = 1\cdots m \tag{12a}$$

$$\sum_{i=1}^{n}\langle u,I_i\rangle^2 = \max_{\text{all } x}\sum_{i=1}^{n}\langle x,I_i\rangle^2 . \tag{12b}$$

In all cases, $||u|| = 1$ and $||x|| = 1$ (i.e. we describe formulation for normalized image and discriminant vectors). This is necessary to insure that physically large objects do not dominate the filter. Of course, all testing is performed on unnormalized images.

Let us discuss the 3 cases in (10)-(12). In (10), we force the projections for one class $\{J_i\}$ to 0 and maximize the absolute value of each of the projection values of all vector images in the other class $\{I_i\}$. There is no general solution to case A [13]. In case B, we maximize the sum of the absolute values of the projections for the first class of images $\{I_i\}$. In case C, we maximize the sum of the squares of the projections on $\{I_i\}$. Case C is an analytically simpler optimization problem. Thus, we form OLDF-1 using (12). In (12), $\underline{u}$ is the $\underline{x}$ for which $\sum\langle x,I\rangle$ is a maximum.

To solve (12), we first denote the subspace spanned by the $\{J_i\}$ as $Sp\{J_i\}_{i=1}^{m}$ (where m is the number of vectors) and the subspace for $\{I_i\}$ by $Sp\{I_i\}_{i=1}^{n}$. In total, there are N = m+n training set vectors and a maximum of N basis functions $\{\phi_n\}$ for this data. We proceed to form a maximal orthonormal set $\{\phi_i\}_{i=1}^{m'}$ from $\{J_i\}$, where $m' \leq m$. Next, we look at the remaining $\{\phi_k\}_{k=m'+1}^{N}$ orthogonal elements in our space. We form a set n' of these $\{\phi_j\}_{j=m'+1}^{m'+n'}$ (where $n' \leq n$) that spans $Sp\{I_i\}$ and is orthogonal to $Sp\{\phi_i\}_{i=1}^{m'}$. Our OLDF-1 $\underline{u}$ is now an element of $Sp\{\phi_j\}_{j=m'+1}^{m'+n'}$. We thus define

$$I_i' = \sum_{j=m'+1}^{m'+n'}\langle I_i,\phi_j\rangle\phi_j, \tag{13}$$

as a weighted sum (with weights given by the vector-inner product $\langle I_i,\phi_j\rangle$) of the $\phi_j$ (which are orthogonal to the $\phi_i$). The optimization problem in (12a) and (12b) thus becomes: find u, such that

$$\sum_{i=1}^{n}\langle u,I_i'\rangle^2 = \max_{\text{all } x}\sum_{i=1}^{n}\langle x,I_i'\rangle^2. \tag{14}$$

We rewrite (14) as

$$\sum_{i=1}^{n} \underline{u}^T \underline{I}_i' \underline{I}_i'^T \underline{u} = \max_{\text{all } x} \sum_{i=1}^{n} \underline{x}^T \underline{I}_i' \underline{I}_i'^T \underline{x}$$

or

$$\underline{u}^T (\sum_i \underline{I}_i' \underline{I}_i'^T) \underline{u} = \max_{\text{all } x} \underline{x}^T (\sum_i \underline{I}_i' \underline{I}_i'^T) x \tag{15}$$

The solution $\underline{u}$ to (15) that defines OLDF-1 is

$$\underline{u} = \text{Dominant eigenvector of } R_{I'}, \tag{16}$$

where

$$\underline{R}_{I'} = \sum_{i=1}^{n} \underline{I}_i' \underline{I}_i'^T = \text{Correlation matrix for } \{I_i'\}. \tag{17}$$

We note that for n' = 1, (16) solves (10) and (11) also [13].

### 3.2  OLDF-2

For OLDF-2, for each $I_i$, we find the $J_i$ image in $\{J_i\}$ that is closest (using the norm distance) to $I_i$. In OLDF-2, we maximize the sum (over all i = 1···n) of the squares of $<u, I_i - J_i>$, i.e. OLDF-2 is $\underline{u}$ such that

$$\sum_{i=1}^{n} <u, I_i - J_i>^2 = \max_{\text{all } x} \sum_{i=1}^{n} <x, I_i - J_i>^2. \tag{18}$$

Following the procedure in Section 3.1, the OLDF-2 solution $\underline{u}$ is

$$\underline{u} = \text{Dominant eigenvector of } \underline{R}_{I_i - J_i}, \tag{19}$$

where

$$\underline{R}_{I_i - J_i} = \sum_{i=1}^{n} (I_i - J_i)(I_i - J_i)^T \tag{20}$$

is the correlation matrix of the $(I_i - J_i)$ vectors, where $J_i$ is the vector image in $\{J_i\}$ that is closest to $I_i$.

### 3.3  OLDF-3

In OLDF-2, we maximized the difference between $I_i$ and $J_i$, the nearest neighbor of $I_i$. In OLDF-3, we maximize the difference between each $I_i$ and all $J_j$, i.e. the overall total separation between both classes. This OLDF-3 filter $\underline{u}$ is defined by

$$\sum_{i=1}^{n} \sum_{j=1}^{m} <u, I_i - J_j>^2 = \max_{\text{all } x} \sum_{i=1}^{n} \sum_{j=1}^{m} <x, I_i - J_j>^2. \tag{21}$$

Following the procedure in Section 3.1, the solution u to OLDF-3 is

$$\underline{u} = \text{Dominant eigenvector of } \underline{R}_{x_{ij}} \tag{22}$$

where

$$\underline{R}_{x_{ij}} = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} x_{ij}^T \tag{23}$$

$$\underline{x}_{ij} = I_i - J_j. \tag{24}$$
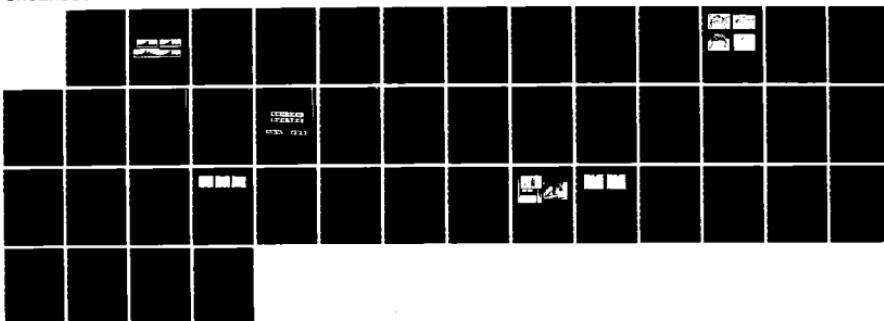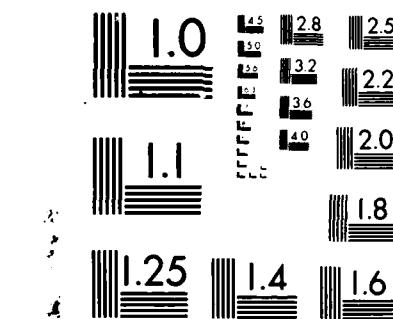
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

## 3.4 EXTENSIONS

In Section 3, we considered OLDF formulation for two-class problems. However, extensions to multi-class recognition can be achieved by extending our OLDFs using the techniques in Section 2. We have described our OLDF solutions as the most dominant eigenvectors. However, one can retain the N most dominant eigenvectors. The number to be retained depends upon the eigenvalues. If the most dominant eigenvalues are close, then we can retain more than one eigenvector. If the two largest eigenvalues are widely separated, keeping the second worst dominant eigenvector will not necessarily improve performance since additional noise is now present in the filter. When N filters are used, the sum of the absolute value of the projections on each is used to compare to a threshold (set from training set data for each class).

## 4. INITIAL TEST RESULTS

### 4.1 DATABASE

To test the performance of our OLDFs, we used two classes (two ships: the Moskva, a Soviet helicopter cruiser, and the Leahy, a U.S. guided-missile cruiser). Each ship was binarized with 128 x 32 pixels. For each ship, 36 views at a 90° depression angle (0° attack angle) were avail ole (every 10° around the ship). The bow is numbered as image 1, broadside as 9 and the stern as 18, etc. For each object class, various sets of 6 images were used for the training set. The OLDF was then tested against all 72 images in the two classes (including the 60 images in both classes that the system had never seen). In Figure 1, there are about 2000 pixels on the broadside ship views and 200 pixels on the bow and stern views. In our tests, we also included noise (in both the training and test set) with $\sigma_n$ listed (SNR is different for each ship aspect view due to the different number of pixels on each aspect view). In Figure 2, we show the views of ship class 1 with $\sigma_n = 0.3$ and $\sigma_n = 0.4$ of noise added.



(a) Class 1 (Moskva)



(a) Class 1 ship ($\sigma_n = 0.3$ noise)



(b) Class 2 (Leahy)



(b) Class 1 ship ($\sigma_n = 0.4$ noise)

FIGURE 1
Broadside views of the two ships

FIGURE 2
Broadside view of ship class 1
with different $\sigma_n$ of noise added

### 4.2 INITIAL TEST RESULTS

Table 1 shows the noise-free performance obtained using the three OLDFs with 6 training set images per class. In general, excellent results are obtained, with no more than 4 misclassifications out of 72 images. In Test 1, the hull was present. Classification is better without the hull present (Tests 2-4), since the ship's superstructure gives good discrimination and the hull is in general common data. For the hull present, two linear functionals had to be used to maintain good performance (thus verifying the above remarks). Test 3 performs worse than Test 2, since maximum separation from one nearest image in the second class is not enough. Test 4 performs better than Test 3 as expected since differences from all images in the two classes was maximized.

The performance of all of the various OLDFs in the presence of noise in the training and testing sets and in both were quantified. The results for OLDF-1 are shown in Table 2. The

results for the other OLDFs are similar. The standard deviation $\sigma_n$ of the noise is listed also. As seen, *performance is excellent in the presence of noise and generally decreases as* $\sigma_n$ increases as expected. From Figure 2 and the amount of deterioration present in the images with $\sigma_n$ = 0.3 and 0.4 noise levels, the initial performance of these OLDFs is quite attractive.

TABLE 1

Test results using various OLDFS (6 training set images/class, binary images)

| TEST | HULL PRESENT | NUMBER OF FUNCTIONALS | OLDF USED | NUMBER OF ERRORS IN | | |
|------|--------------|-----------------------|-----------|---------|---------|-------|
|      |              |                       |           | CLASS 1 | CLASS 2 | TOTAL |
| 1    | Yes          | 2                     | 1         | 0       | 2       | 2     |
| 2    | No           | 1                     | 1         | 1       | 0       | 1     |
| 3    | No           | 1                     | 2         | 3       | 1       | 4     |
| 4    | No           | 1                     | 3         | 0       | 0       | 0     |

TABLE 2

Noise performance of OLDFs
(6 training images/class, binary images, one functional)

| TEST | TRAINING SET $\sigma_n$ | TEST SET $\sigma_n$ | NUMBER OF ERRORS IN | | |
|------|-------------------------|---------------------|---------|---------|-------|
|      |                         |                     | CLASS 1 | CLASS 2 | TOTAL |
| 1    | 0.0                     | 0.0                 | 1       | 0       | 1     |
| 2    | 0.3                     | 0.0                 | 0       | 4       | 4     |
| 3    | 0.3                     | 0.3                 | 1       | 5       | 6     |
| 4    | 0.4                     | 0.4                 | 1       | 8       | 9     |

## 5. SUMMARY

New distortion-invariant correlator filters have been described that maximize various discriminant pattern recognition measures. The theoretical basis and ease of analysis for these new OLDFs is attractive. Initial experimental results are excellent and noise performance is robust. Full correlation tests and further experiments are needed to assess OLDFs more fully.

## ACKNOWLEDGMENTS

## REFERENCES

1. A. Vander Lugt, IEEE, Vol. IT-10, 128 (1964).
2. J. Rau, J. Opt. Soc. Amer., Vol. 56, 1490 (November 1966).
3. J. Upatnieks, Applied Optics, Vol. 22, 2798 (September 1983).
4. J.G. Duthie and J. Upatnieks, Optical Engineering, Vol. 23, 7 (Jan/Feb 1984).
5. D. Casasent, Applied Optics, Vol. 23, 1620 (May 1984).
6. D. Casasent, B.V.K. Vijaya Kumar, and V. Sharma, Proc. SPIE, Vol. 360, 136 (August 1982).
7. D. Casasent, W. Rozzi, and D. Fetterly, Optical Engineering, Vol. 23 (November 1984).
8. W.T. Chang, D. Casasent, and D. Fetterly, Proc. SPIE, Vol. 507 (August 1984).
9. H.J. Caulfield and M.H. Weinberg, Applied Optics, Vol. 21, 1699 (May 1982).
10. Z.H. Gu and S.H. Lee, Applied Optics, Vol. 23, 822 (March 1984).
11. B.V.K. Vijaya Kumar, D. Casasent, and H. Murakami, Optical Engineering, Vol. 21, 43 (Jan/Feb 1982).
12. T. Kato, Perturbation Theory for Linear Operators, Springer Verlag Pub. (1980).
13. V. Sharma, "Design and Analysis of Algorithms for Distortion-Invariant Object Recognition", PhD Thesis, Department of Electrical and Computer Engineering, Carnegie-Mellon University, Pittsburgh, PA (1984).

# 13. CORRELATION SDF CONCEPT AND INITIAL ATR TARGET DATA RESULTS

Ref. 17

# SDF CONTROL OF CORRELATION PLANE STRUCTURE FOR 3-D OBJECT REPRESENTATION AND RECOGNITION

Wen-Thong Chang, David Casasent and Donald Fetterly*

Carnegie-Mellon University
Department of Electrical and Computer Engineering, Pittsburgh, PA 15213

*General Dynamics - Pomona Division, P.O. Box 2507, Pomona, CA 91769

## ABSTRACT

The efficient representation and synthesis of 3-D object information using new synthetic discriminant functions (SDFs) is discussed. The use of SDFs in a correlator for shift-invariant and distortion-invariant discrimination of 3-D objects is detailed and experimental data is provided. The new SDFs described control the peak intensity and the structure and statistics of the correlation plane pattern.

## 1. INTRODUCTION

Correlators represent one of the most powerful techniques for automatic target recognition (ATR). These systems allow multiple objects to be recognized in parallel (by the shift-invariant property of a correlator) in the presence of noise and structured clutter (due to the processing gain achieved by a correlator). The realization of correlators using coherent optical systems is obvious [1,2] and small size and weight real-time coherent optical correlators now exist [3]. Advanced VHSIC chips and architectures may also allow on-line correlations to be implemented digitally. The major shortcomings of any correlator has been their poor performance in the face of geometrical distortions between the input image and reference object from which a matched spatial filter (MSF) is formed [4]. Recently, advanced MSF synthesis techniques have been detailed [5] and demonstrated on ship imagery [6]. These new MSF synthesis algorithms form the MSF from a training set of images of different target objects from different aspects, scales, rotations, etc. These new filter functions are referred to as synthetic discriminant functions (SDFs).

A brief review of conventional SDFs is provided in Section 2 with discussion on their use in the representation of 3-D object information. Three new types of SDFs that control the shape of a correlation plane pattern are then detailed in Section 3 with discussion on their representation of 3-D object data. Initial experimental results are then advanced in Section 4 using a new database of tank, armored personnel carrier (APC) and similar military ATR objects.

## 2. THE SDF-BASED CORRELATOR CONCEPT

To achieve intra-class recognition of different distorted versions of a 3-D input ATR object using a correlator, the MSF $h(x,y)$ can be formed from a linear combination of training set images $\{f_n\}$ that are different 3-D distorted views of the target object, i.e.

$$h(x,y) = \sum_n a_n f_n(x,y). \tag{1}$$

If we restrict the correlation peak value to be unity for $f_n$, then the SDF MSF in (1) is defined by

$$\underline{a} = \underline{R}^{-1}\underline{u} \tag{2}$$

where the elements of the vector $\underline{a}$ define the linear combination coefficients $a_n$, $\underline{u}$ is the unit vector (this forces all correlation plane values to be unity), and $\underline{R}$ is the correlation matrix of the $\{f_n\}$ training set imagery. The SDF in (1) - (2) achieves intra-class recognition. To obtain inter-class discrimination while still retaining intra-class recognition, the training set is expanded to include sets of the distorted objects $\{f_1\}$ and $\{f_2\}$ in two or more classes. A single SDF or several SDFs that are linear combinations of all of the training set imagery can then be formed. The filter synthesis procedure is similar to that in (1) and (2) with larger $\underline{R}$ matrices (for several object classes) and different exogenous vectors $\underline{u}$ as detailed elsewhere [5]. The object class can be determined from the value of the correlation peak or from combinations of different filter output values.

These initial SDFs [5] have performed well in tests on various image databases [6] primarily on ship imagery. In this paper, we consider other ATR targets (tanks and APCs) and we extend the original SDF concept to include control of the shape of the correlation plane

pattern. The original SDFs only control the value at one point in the correlation and thus
we refer to these as projection SDFs. For ATR using the original SDFs, a correlation plane
threshold is set (determined by the filter synthesis algorithm) and from the locations and
peak values of the regions of one or several correlation planes that exceed the threshold,
the object class and object location in the input field-of-view can be determined. This
technique is susceptible to variations in the modulation level of the input data (since the
correlation value varies linearly with the modulation of the input object). From the dc
value of the input Fourier transform (FT) plane pattern, the output threshold can be adjusted.
In Section 2, we detail three new SDFs (correlation SDFs) that automatically control the
shape of the true and false correlation plane locations and thus facilitate correlation
plane analysis by the combination of threshold detection and correlation plane and peak
analysis.
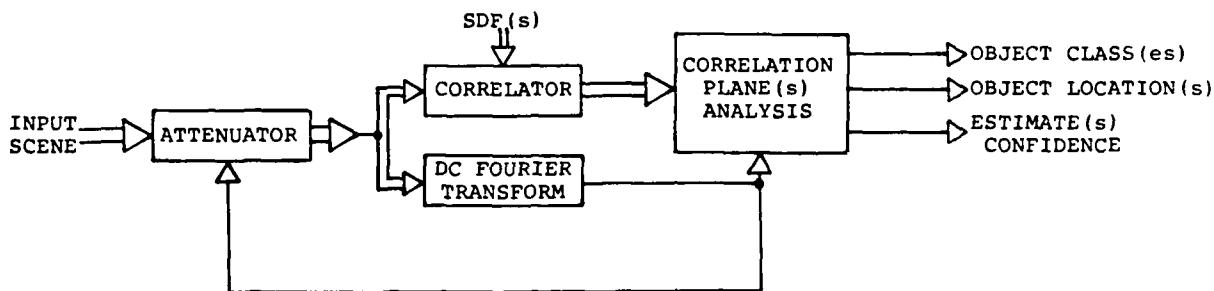


FIGURE 1
Block diagram of an SDF-based correlator

The full correlation system (Figure 1) thus distributes the processing and recognition
load between the filter synthesis, the correlator and the output plane detector. From the
dc value of the FT of the input, an estimate of the input modulation is obtained and used
to adjust the input intensity and the correlation plane threshold. The system's outputs pro-
vide estimates of the object class and location of all objects in the input field-of-view
and the confidence of these estimates.

### 2. CORRELATION SDF SYNTHESIS FOR CORRELATION SHAPE CONTROL

To control the shape of the correlation peak for a true target and to insure suppression
of large correlation plane peaks for shifted versions of false target, we expand the train-
ing set of images to include $N_S$ shifted versions of each object. To describe the filter
synthesis, we consider a two-class pattern recognition problem with $N_1$ and $N_2$ training set
images {f} and {g} per class with $N_S$ shifted versions of each training set image, i.e. a
total of $N_T = N_S(N_1 + N_2)$ training set images. The SDF synthesis algorithm in (2) restricts
only the vector inner product or the vector projection of each object $\underline{f}_i$ or $\underline{g}_i$ onto the
filter function $\underline{h}$, i.e. only the central correlation plane value.

### 2.1 SDF-1 (Exact Correlation SDF)

For the one-filter two-class pattern recognition problem, the new SDF is defined to
satisfy

$$\underline{h} \cdot \underline{f}_i = 1 \quad \text{(central correlation value of 1 for true targets)} \tag{3a}$$

$$\underline{h} \cdot \underline{f}_i' = 0 \quad \text{(0 correlation value away from peak)} \tag{3b}$$

$$\underline{h} \cdot \underline{g}_i' = 0 \quad \text{(0 correlation value away from peak)} \tag{3c}$$

$$\underline{h} \cdot \underline{g}_i = 0 \quad \text{(central correlation value of 0 for false targets)} \tag{3d}$$

where the notation is defined in Table 1. Eqs.(3a) and (3d) are similar to the original SDF
requirements (a central correlation peak value of unity for true targets in class one and
zero projections for false class two targets). Eqs.(3b) and (3c) are the new restrictions

added. They require the correlation plane value $d_S$ pixels from the central value to be zero.

TABLE 1
Notation used to describe synthesis of the new SDFs

| | |
|---|---|
| $\underline{f}_i$ = object in the class to be recognized | $N_1$ = number of training set images $\{f\}$ in class 1 |
| $\underline{f}'_i$ = shifted version of $\underline{f}_i$ | $N_2$ = number of training set images $\{g\}$ in class 2 |
| $\underline{g}_i$ = object in second-class to be rejected | $N_S-1$ = number of shifted versions of each image |
| $\underline{g}'_i$ = shifted version of $\underline{g}_i$ | $N_T = N_S(N_1+N_2)$ = total training set size |
| $d_S$ = amount of shift (in pixels) for ( )' | $\underline{h}$ = MSF SDF filter function |

The filter $h(x,y)$ satisfying (3) will thus have a correlation plane output for a true class-one target with a fixed peak value of unity and a fixed zero-value $d_S$ pixels away (in ±x and ±y) and thus a well-defined correlation plane peak shape. The controlled correlation peak value can allow the use of a fixed correlation plane threshold $T_C$ of 0.5. This selects regions of potential interest in the field-of-view. For each output plane region of interest exceeding the threshold $T_C$, the peak/mean = C ratio is computed. This new classification measure C applied to those correlation plane regions exceeding $T_C$ allows significantly better system performance. To calculate C, the mean is computed over the $(2d_S+1) \times (2d_S+1)$ pixel region around each peak of interest (the mean computation excludes the central peak value). In our specific work, the parameters in Table 2 are used. Many other variations of this basic algorithm are possible such as:

(1) applying the C threshold to the regions of the output with the largest peak values only or just the largest peak location (if only one object is known to be present);
(2) extension of the filter in Eq.(3) to the five different projection SDFs [5];
(3) extension of the requirements in Eq.(3) to include more shifted versions of each input image;
(4) application of a weighted spatial taper to the FT of the SDF to suppress its sidelobe response; and
(5) modification of the mean in C to include only those correlation plane values at the specific $d_S$ pixel shifts from the peak value;
(6) use of input modulation estimates to adjust the $T_C$ threshold.

TABLE 2
Specific filter parameters used

| | |
|---|---|
| $N_1 = N_2 = 6$ | $N_S$ = 5 (centered and 4 shifted) |
| $d_S$ = 5 pixels | $N_T = 5(6+6) = 60$ |

To synthesize the SDF $\underline{h}$ satisfying (3), we restrict $h(x,y)$ to be a linear combination of all training set images

$$h(x,y) = \sum_{N_1} a_{i1}f_i(x,y) + \sum_{N_2} a_{i2}g_i(x,y) + \sum_{(N_S-1)N_1} a_{i3}f'_i(x,y) + \sum_{(N_S-1)N_2} a_{i4}g'_i(x,y), \qquad (4)$$

where the number of images in each summation is indicated under the associated $\Sigma$. Denoting the full $N_T$ set of training set images by $\{z\}$ and individual images by $\underline{z}_n$, the filter $\underline{h}$ in (4) is defined by the coefficient vector $\underline{a}$ that solves the equation

$$\underline{R}\,\underline{a} = \underline{u}_1, \quad \underline{u}_1 = [\underbrace{1\cdots1}_{N_1} \ \underbrace{0\cdots0}_{N_T-N_1}]^T, \qquad (5)$$

where $\underline{R}$ is the full $N_T \times N_T$ correlation matrix of all $\{z\}$ training set images. The choice of $\underline{u}_1$ in (5) satisfies the requirements in (3). Eq.(5) is a simple extension of (2). The SDF in (3) is referred to as a <u>correlation SDF</u> and the specific SDF solution in (5) is denoted as <u>SDF-1</u> or the <u>exact correlation SDF</u>. This terminology refers to the fact that the solution in (5) is an exact solution to (3).

## 2.2 SDF-2 (Least-Square Correlation SDF)

The solution $\underline{a}$ in (5) requires solving the $N_T$ linear algebraic equations (LAEs) defined by (5). As the nature of the problem increases, so will the dimensionality $N_T$ of $\underline{R}$ and computational problems plus ill-conditioned matrices may arise (even though the filter synthesis is performed off-line). The typical solution to such a problem is to reduce the dimensionality of $\underline{R}$, i.e. to reduce the number of training set images. However, a reduction in $N_1$ or $N_2$ will degrade the 3-D object information on each target class and a reduction in $N_S$ will degrade the correlation shape. The SDF synthesis in (4) and (5) is equivalent to describing each object as a d-dimensional vector in hyperspace where $d = N_T$. In these terms, our second realization of SDF-1 retains all $N_T$ training set images but reduces d to D, such that $D < N_T$. This reduction of the dimensionality of our hyperspace rather than the number of training set images is both practical, preferable and new. To select the D basis functions $\phi_d(x,y)$ we use the well-known Karhunen-Loeve (KL) technique [7]. For each unshifted set object, and for each shifted version of each training set object, we compute the dominant KL eigenvectors of the associated correlation matrix. In our experiments, we retain three dominant KL eigenvectors per class (as noted above). Efficient methods of computing the dominant KL eigenvectors of a large matrix and a large database were noted earlier [8].

For the case of $(N_S-1) = 4$ shifted versions of each image (five shifted images, including the central centered image), the three dominant KL eigenvectors of each of the ten correlation matrices $R$ (the matrices for the original object $\underline{f}$, the false target $\underline{g}$, and each of the four sets of shifted images per class, with two shifts in both x and y) were computed and retained. This provides a new $D = 30$ (rather than $d = 5 \times 12 = 60$) basis function set. This new $\{\phi'\}$ basis function set thus represents all of the 3-D information in the training set imagery for the two targets. Retaining more than three KL eigenvectors per class improves the accuracy of this approximate algorithm (at the expense of increased off-line computational complexity). As noted in [11], three eigenvectors are generally adequate to represent over 90% of the 3-D object information. This $\{\phi'\}$ basis function set was then converted to the orthonormal basis function set $\{\phi\}$ using a Gram-Schmidt (GS) orthogonalization technique [12].

In terms of these new $\phi_d$ basis functions, we describe the desired filter function as the linear combination filter

$$h(x,y) = \sum_{d=1}^{D} b_d \phi_d(x,y), \tag{6}$$

Each input image in $\{z\}$ can then be described as a linear combination of the basis functions as

$$z_n(x,y) = \sum_{d=1}^{D} c_{nd} \phi_d(x,y), \tag{7}$$

where n varies from 1 to $N_T$ (i.e. over the full training set of $N_T$ images) and $\phi_d$ is of size $N \times N = N^2$. Hereafter all $z_n(x,y)$ images are represented as vectors $\underline{z}_{nd}$ of length D (with their D elements equal to the projections of $z_n(x,y)$ on the D vectors $\phi_d(x,y)$. According to (3), we require

$$\underline{z}_{nd}^T \underline{h} = 1 \text{ if } \underline{z}_{nd} = \underline{f}_i$$
$$\underline{z}_{nd}^T \underline{h} = 0 \text{ if } \underline{z}_{nd} \neq \underline{f}_i \tag{8}$$

In matrix-vector form, we write (6) and (8) as

$$\underline{C}\,\underline{b} = \underline{u}_1, \tag{9}$$

where $\underline{C}$ contains $N_T$ rows and D columns. Since $N_T > D$, the classic least-square solution is used to determine $h(x,y)$ as

$$\underline{b} = [\underline{C}^T \underline{C}]^{-1} \underline{C}^T \underline{u}_1, \tag{10}$$

where the size of $\underline{C}^T\underline{C}$ is now D x D. Other optical solutions to such an overdetermined least squares problem [9,10] were considered for the case $D > N_T$. These solutions have not been correctly formulated. Specifically, if $D > N_T$, no unique solution exists, since the $[\underline{C}^T\underline{C}]$ matrix is not full and thus cannot be inverted. Thus, a least squares solution is not appropriate.

For our case, $N_T > D$ and a solution exists. We recall that

$$\underline{C}^T\underline{C} = \sum_{n=1}^{N_T} \underline{z}_{nd}\underline{z}_{nd}^T = N_T\underline{R}, \tag{11}$$

where $\underline{z}_{nd}$ is of length D, the matrix multiplication in (11) is formed as a vector-outer-product sum, and $\underline{R}$ in (11) is the D x D correlation matrix of reduced dimension D. We also note that

$$\underline{C}^T\underline{u}_1 = \sum_{n=1}^{N_T} u_n\underline{z}_{nd}, \tag{12}$$

where the matrix-vector product in (12) has been written as a vector-inner product sum over each vector element $\underline{z}_{nd}$ weighted by the elements $u_n$ of $\underline{u}_1$. Using (11) and (12), Eq. (9) becomes

$$N_T\underline{R}\ \underline{b} = \sum_{n=1}^{N_T} u_n\underline{z}_{nd}.$$

The solution for the SDF in terms of the $\underline{z}_{nd}$ of length D and the D x D matrix $\underline{R}$ is

$$\underline{b} = \frac{1}{N_T}\ \underline{R}^{-1} \sum_{n=1}^{N_T} u_n\underline{z}_{nd}. \tag{14}$$

From the D elements of $\underline{b}$, the D vectors $\underline{\phi}_d$, and (6), the SDF h(x,y) is defined.

Solving (14) involves the solution of D rather than d simultaneous LAEs. Thus, the solution (14) is computationally simpler and faster, although the final result is more approximate. The use of (14) lies primarily in its computational ease for cases when $N_T$ is large. It is also useful in cases when $\underline{h}$ must be updated in real-time by an on-line processor. This corresponds to a Kalman filter update of the SDF function. This situation arises when the projection value for an input object is near threshold. In this case, we can update the filter with subsequent views of the input object and thus improve its original projection value. Such cases occur when the scale of the target or its depression angle, etc. differ from that of the training set images used. The solution in (14) minimizes the mean square error

$$J = \sum_{n=1}^{N_T} ||\underline{z}_{nd}^T\underline{h} - \underline{u}_n||, \tag{15}$$

where $u_n = 1$ for $\underline{z} = \{f_i\}$ and is zero otherwise. Setting $\partial J/\partial h = 0$ in (15) and solving for $\underline{h}$, we obtain the coefficient solution in (14). Since $\underline{R}$ in (14) is of reduced dimensionality $\overline{D} = 30$ (rather than d = 60), the solution in (14) is far simpler and more accurately computed.

The least-square correlation SDF filter solution h in (14) is an approximate solution to the exact filter function problem in Section 2.1. Hence, the associated name for this filter function noted earlier is employed. The accuracy of this solution depends upon the accuracy to which the several dominant KL eigenvectors per correlation matrix adequately represent the data in the full correlation matrix. The summation of the eigenvalues associated with these eigenvectors quantifies this accuracy. This SDF-2 filter function thus attempts to select $\underline{h}$ such that the desired peak values are as close to unity as possible and that the false target peak values at all shifted image correlation values are as close to zero as possible (in a least-square sense). A correlation peak threshold T = 0.5 can be used as before or one can simply calculate C for all correlation plane regions with large peak values. Experimental data on these methods using such a SDF are advanced in Section 3.

### 2.3 SDF-3 (Generalized Correlation SDF)

SDF-2 is somewhat statistical since the $\underline{h}$ choice minimizes J in (8). The final type of correlation SDF is also statistical. Rather than selecting h to cause the desired correlation plane values to be as close to 1.0 and 0.0 as possible (as in SDF-1 and SDF-2), SDF-3 is chosen to maximize the peak-to-mean ratio C. In SDF-1 and SDF-2, the C test will provide target discrimination and recognition plus invariance to object modulation (C is invariant to object modulation or contrast). However, if the peak value for the target is not above the threshold T, the C test will never be applied to the proper correlation plane region. Assuming that maximizing C (or the correlation plane SDF) maximizes the peak value $I_p$ of the correlation, then SDF - 3 will produce both large peak values and large C values. Specifically, the correlation plane regions with the largest peak values

are expected to include the correct target objects and the correlation plane regions corresponding to the correct targets will also have large C values [4].

The central correlation plane value for an $\underline{f}_i$ input image vector is $\underline{h}^T \cdot \underline{f}_i$. The mean-square-value of this correlation point value for all $\{\underline{f}_i\}$ is $E[(\underline{h}^T \underline{f}_i)^2] = \underline{h}^T \underline{R} \underline{h}$, where $\underline{R}$ is the correlation matrix of $\{\underline{f}_i\}$. Maximizing the correlation plane SNR for correct targets thus requires maximization of

$$J = \frac{\text{mean-square value of central correlation point for } \{f_i\}}{\text{mean-square value of central point for } \{f'_i\}, \{g_i\} \text{ and } \{g'_i\}}$$

$$= \frac{\underline{h}^T R_f \underline{h}}{\underline{h}^T [\underline{R}_{fs} + \underline{R}_g + \underline{R}_{gs}] \underline{h}} , \tag{16}$$

where $\underline{R}_{fs}$ is the sum of the four correlation matrices $\underline{R}_{f1}$, $\underline{R}_{f2}$ etc. for the objects $f_i$ shifted in $\pm x$ and $\pm y$ ($\underline{R}_{gs}$ is similarly defined). The solution that maximizes J in (16) is the solution of

$$\underline{R}_f \underline{h} = \lambda [\underline{R}_{fs} + \underline{R}_g + \underline{R}_{gs}] \underline{h} \tag{17}$$

where $\lambda$ is the generalized eigenvalue of the matrices. The problem defined by (17) is the well-known generalized eigenvalue problem and thus we refer to the SDF $\underline{h}$ that maximizes the SNR defined by J as a <u>generalized correlation SDF</u> (SDF-3). The same $\{\phi\}$ orthonormal basis function set used in the least-square SDF is again employed here with each matrix in (17) being D x D = 30 x 30 (for our cases).

For SDF-1 and SDF-2, regions of the correlation plane above T $\approx$ 0.5 as well as the largest peaks anywhere in the correlation plane are classified as interesting regions of potential interest. For each of these regions, we calculate the peak-to-mean ratio

$$C = \frac{\text{central peak intensity}}{\text{mean in 11 x 11 window}}, \tag{18}$$

where an 11 x 11 = $(2d_s+1) \times (2d_s+1)$ window was chosen to agree with the $d_s = 5$ pixel image shifts used in our data and where the correlation peak value is not included in the mean calculation. For each potential region of interest, C is compared to a threshold $C_T$ determined from the C values calculated for the $N_1$ and $N_2$ centered training set images. Since SDF-3 does not fix a correlation plane peak intensity, the largest correlation plane peaks are selected, C is calculated for these points and compared to $C_T$.

## 3. INITIAL EXPERIMENTAL RESULTS

### 3.1 Database

The ATR data base used in our initial experimental results reported herein consisted of t'ree different objects (two tanks denoted as tank 1 and tank2 and an APC). High resolution i ages of these objects were obtained and decimated to produce 56 x 22 target pixel images typical of data from a FLIR at the typical ATR acquisition range of interest. For each object, 36 images from a 20° depression angle were available at 10° aspect intervals. The pixel values of the images varied from 0 to 255 with most target pixels having values near 0 and 150. In Figure 2, we show two images of tank 1 (M60) and the APC at two different aspect views. Denoting the front of the tank as image 1, tank 2 images 11-15 were much dimmer and tank 1 images 30-34 were much brighter.

Six or twelve image aspects per class were selected for the training set (all approximately evenly spaced in aspect angle). The centered and four shifted images of each object were used for the training set. Each SDF was designed to recognized tank 1 and to reject either the APC or tank 2. Intra-class recognition and inter-class discrimination were always tested using all 72 images in the two classes. Each test image was centered but all points in the correlation plane were tested for the threshold T = 0.5. The correct central correlation peak value was measured and compared to T = 0.5. The largest peak value anywhere in the correlation plane was also measured. C in (18) was calculated only for this largest peak point regardless of whether it was >T. Regardless of whether the central point was above threshold, C was calculated only for the largest correlation plane point. Errors in the peak intensity are expected due to aspect views not in the training set, due to
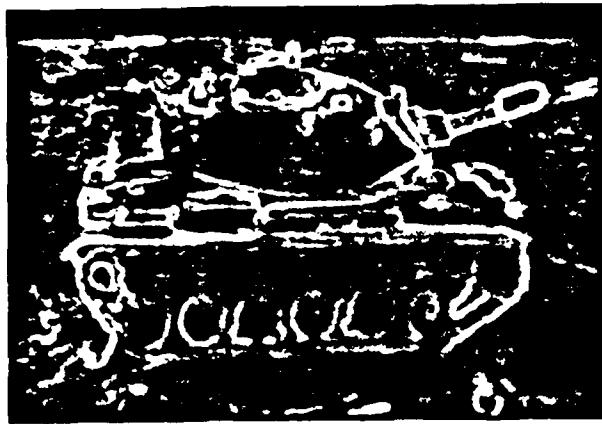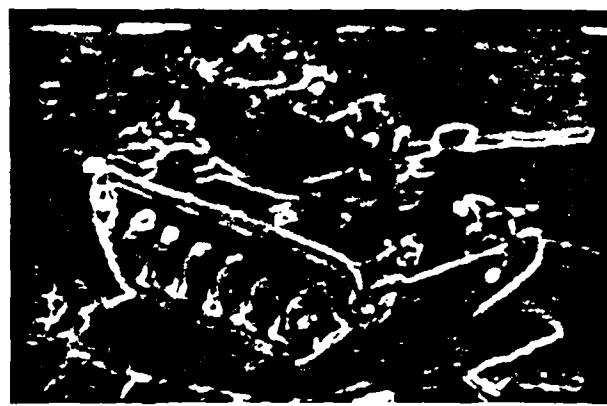
(a) Tank 1, View 1            (b) Tank 1, View 2
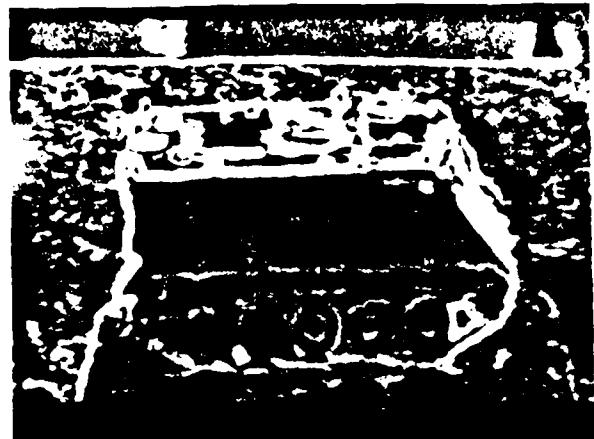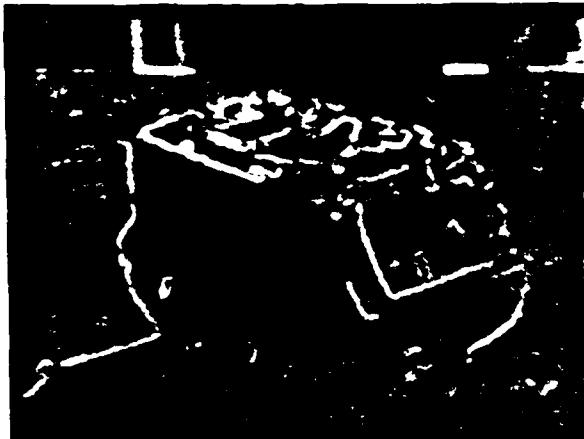
(c) APC, View 1            (d) APC, View 2

FIGURE 2

Representative images for 3-D ATR testing

different aspect views of different objects being similar and due to variations in the modulation level of the training set and test set images. Our data represents very worse-case results. For true targets, by evaluation only at one point, we often miss a target, since C at the wrong point never exceeds $C_T$ (for a true target). For false targets by calculating C at the wrong point even though the peak intensity there is below T, we can often misclassify a ta.get. Thus, the data presented is quite worse-case and significant improvement is possible.

In Table 3, we show the data for tank 1 - APC for six aspect views per class. From row one, we see that all correlation plane values are correct (below T) for false targets and only three or five images have central peak values below T (for true targets). This was found to be due to low modulation of the imagery and to aspect views quite different from those in the reference data set). With T = 0.4 (0.35) for SDF-1 (SDF-2) all correct peaks exceeded T. No peak threshold was used with SDF-3, rather the largest several correlation plane peaks would be investigated. Similar remarks apply to the practical realization of SDF-1 and SDF-2. In row two, we see that the largest correlation peak always occurs in the wrong place for a false target (as expected), but from row three, at the most only six of these peaks have C > $C_T$. In row two, the largest peak is always in the correct location (for SDF-1 and SDF-2). The number of C errors (C < $C_T$ for a tank input and C > $C_T$ for an APC input) are listed in the table. Errors in the first case are missed targets. Errors in the second class are misclassified objects. Most errors occurred for the same images (four of which were very bright) and three of which had aspect views significantly different from those in the training set). In general, least-squares SDF-2 performs comparable to SDF-1. The projection values were in general lower for SDF-2 (especially for the central (correct) peak value). This is expected, since this is an approximate image solution and since only three eigenvectors are used to represent each set of training set images for each shift. SDF-3 performed worst. We might expect it to perform better, since it maximizes C. Modulation variations in the training set appear to be the cause for its poorer performance. The percent of all 72 images correctly classified is noted and the percent of the objects with T ≥ 0.5 misclassified is noted in the tables.

TABLE 3
Worst-case performance of the three SDFs for Tank 1/APC data (72 images)
(6 training set images per class, approximately every 60°, 5 shifted versions of each)

| SDF | SDF-1 | | SDF-2 | | SDF-3 | |
|---|---|---|---|---|---|---|
| INPUT | TANK 1 | APC | TANK 1 | APC | TANK 1 | APC |
| No. of Central Peak Errors T ≤ 0.5 | 3 | 0 | 5 | 0 | - | - |
| No. of Largest Peaks in Wrong Location | 0 | All | 0 | All | 7 | All |
| No. of C Errors | 3 | 1 | 2 | 1 | 7 | 6 |
| Percent Correct | 94.4% | | 95.8% | | 81.9% | |
| Percent Wrong | 0% | | 0% | | 8.3% | |
| $C_T$ Threshold | 5.0 | | 4.3 | | 3.5 | |

Table 4 shows similar data for tank 1 versus tank 2. The trends are quite similar. Table 5 shows data for the case of 12 training set aspect images per class. The significant reduction in the number of errors observed is due to the fact that the largest correlation plane value is now in the correct location (for true targets). As in Tables 3 and 4, since C is calculated only at the largest correlation plane point, if this point is wrong (for a correct target), then C never exceeds $C_T$ and a target is missed. In Tables 3 and 4, $C_T$ was set at 1.5 below the average C value for the training set images in both classes. In Table 5, $C_T$ was set higher at 0.5 below the average (since with more training set images, our confidence is higher).

**TABLE 4**
Worst-case performance of three SDFs for Tank 1/Tank 2 data (72 images)
(6 training set images per class, every 60°, 5 shifted versions of each)

| SDF | SDF-1 | | SDF-2 | | SDF-3 | |
|---|---|---|---|---|---|---|
| INPUT | TANK 1 | TANK 2 | TANK 1 | TANK 2 | TANK 1 | TANK 2 |
| No. of Central Peak Errors $T \gtrless 0.5$ | 3 | 0 | 3 | 0 | - | - |
| No. of Largest Peaks in Wrong Location | 1 | All | 0 | All | 4 | All |
| No. of C Errors | 1 | 3 | 3 | 0 | 5 | 0 |
| Percent Correct | 94.4% | | 95.8% | | 93.0% | |
| Percent Wrong | 0% | | 0% | | 7% | |
| $C_T$ Threshold | 4.1 | | 4.6 | | 3.9 | |

**TABLE 5**
Worst-case performance of three SDFs for Tank 1/APC data (72 images)
(12 training set images, every 30°, 5 shifted versions of each)

| SDF | SDF-1 | | SDF-2 | | SDF-3 | |
|---|---|---|---|---|---|---|
| INPUT | TANK 1 | APC | TANK 1 | APC | TANK 1 | APC |
| No. of Central Peak Errors $T \gtrless 0.5$ | 0 | 0 | 0 | 0 | 0 | 0 |
| No. of Largest Peaks in Wrong Location | 0 | All | 0 | All | 0 | All |
| No. of C Errors | 0 | 0 | 0 | 2 | 3 | 2 |
| Percent Correct | 100% | | 97.2% | | 93.0% | |
| Percent Wrong | 0% | | 0% | | 4.2% | |
| $C_T$ Threshold | 5.5 | | 4.8 | | 4.4 | |

## 4. SUMMARY AND CONCLUSION

The three new SDFs described represent 3-D object information and discrimination information between 3-D objects quite well. Initial tests show excellent results. As noted, the test performed is quite worst-case, since the largest correlation plane point only was used and because of fluctuations in the modulation of the training set imagery.

## ACKNOWLEDGMENTS

## REFERENCES

1.  A. VanderLugt, IEEE, Vol. IT-10, 128 (1964).
2.  D. Casasent, "Coherent Optical Pattern Recognition: A Review", Optical Engineering, Special Issue, Vol. 24 (January 1985).
3.  J. Upatnieks, Applied Optics, Vol. 22, 2798 (September 1983).
4.  D. Casasent and A. Furman, Applied Optics, Vol. 16, 1652 (June 1977). See also D. Casasent and A. Furman, Applied Optics, Vol. 16, 1662 (June 1977).

5.  D. Casasent, _Applied Optics_, Vol. 23, 1620 (May 1984).

6.  D. Casasent, W. Rozzi and D. Fetterly, _Optical Engineering_ (November 1984).

7.  K. Fukunaga and W.L.G. Koontz, _IEEE Trans. Comp._, Vol. C-19, 311 (April 1970).

8.  H. Murakami and B.V.K. Vijaya Kumar, _IEEE Trans. on Patt. Ana. and Mach. Intell._, Vol. PAMI-4, 511 (1982).

9.  Zu-Han Gu, J.R. Leger and S.H. Lee, _J. Opt. Soc. Am._, Vol. 72, 787 (1982).

10. Zu-Han Gu and S.H. Lee, _Applied Optics_, Vol. 23, 822 (March 1984).

11. S. Watanabe, "Karhunen-Loeve Expansion and Factor Analysis, Theoretical Remarks and Applications", _4th Prague Conf. on Inf. Theory_, 1965.

12. J.M. Wozencraft and I.M. Jacobs, _Principles of Communication Engineering_, John Wiley and Sons, New Jersey (1965).

# 14. CORRELATION SDF TESTS ON AIRCRAFT IMAGERY

Ref. 18

## Correlation Filters for Distortion-Invariance and Discrimination

David Casasent and Abhijit Mahalanobis
Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

### 1.   INTRODUCTION

Correlators are powerful shift-invariant object recognition systems that perform well in noise.   However, they are quite sensitive to distortions between the input and reference object.   Synthetic discriminant functions (SDFs) [1] accommodate intra-class distortions and provide inter-class discrimination.   In Section 2, we review these projection SDFs and note that they restrict only the peak point in the correlation plane.   In Section 3, new correlation SDFs [2] are described. They control both the peak and sidelobe response and thus exhibit superior performance.   Initial test data on these SDFs are presented in Section 4.

### 2.   PROJECTION SDFs

In the synthesis of projection SDFs, the SDF $h$ is a linear combination of the training set images $\{f\}$ in classes 1, 2, etc., i.e.

$$h(x,y) = \sum_n a_n f_n(x,y). \tag{1}$$

The coefficient vector $\underline{a}$ that defines $\underline{h}$ is given by

$$\underline{a} = \underline{R}^{-1}\underline{u}, \tag{2}$$

where $\underline{R}$ is the vector inner product matrix of all $\{f\}$ with class one data $\{f_1\}$ being the first $N_1$ images and class 2 data the next $N_2$ images, etc.   The elements of the deterministic vector $\underline{u}$ define the filter's desired response for the $\{f\}$ data.   With the first $N_1$ elements of $\underline{u}$ unity and the next $N_2$ elements zero, the SDF provides a "1" output for all $\{f_1\}$ and a "0" for all $\{f_2\}$.   Many other choices for $\underline{u}$ exist and correspond to the various types of SDFs [1].

However, this filter synthesis only restricts the central peak or correct projection value in the correlation output.   There is no guarantee that the value at other locations in the correlation plane will not exceed the value at the point of registration (we refer to this as the central value, with no loss of generality and denote this value by $I_p$).   This problem is particularly severe when the input is a false target (one in class 2) for which a "0" output is desired.   Another shortcoming of projection SDFs is that only a simple correlation plane threshold ($T = 0.5$ or other levels) is used to achieve object detection.

### 3.   CORRELATION SDFs

Projection SDFs adequately control $I_p$.   To control the sidelobes, we increase the training set size to include $N_s$ shifted versions of each training set imgae, the centered image and $(N_s-1)$ shifted versions.   We select $N_s = 5$ and the shifted images symmetrically to be $d_s$ pixels in both $\pm x$ and $\pm y$.   Correlation SDF synthesis still uses (1) and (2) with $\{f\}$ and $\underline{R}$ being larger, i.e. with $N_s(N_1+N_2) = N_T$ training images (for a two-class problem).   The control vector $\underline{u}$ has zero-

values for elements corresponding to all shifted versions of all images. Denoting class 1 (true) objects by $\underline{f}_i$, class 2 (false) objects by $\underline{g}_i$ and shifted versions of each by primes, the filter requirements are

$$\underline{h} \bullet \underline{f}_i = 1, \quad \underline{h} \bullet \underline{f'}_i = 0$$
$$\underline{h} \bullet \underline{g}_i = 0, \quad \underline{h} \bullet \underline{g'}_i = 0. \tag{3}$$

The linear combination correlation SDF = $\underline{h}$ is

$$h(x,y) = \underset{N_1}{\Sigma} a_{i1}f_i(x,y) + \underset{N_2}{\Sigma} a_{i2}g_i(x,y) + \Sigma a_{i3}f'_i(x,y) + \Sigma a_{i4}g'_i(x,y), \tag{4}$$

where the last two summations are over $(N_S-1)N_1$ and $(N_S-1)N_2$. The vector inner product matrix $\underline{R}$ is $N_T \times N_T$ with the first $N_1$ images being $\underline{f}_i$. The SDF-$\underline{h}$ is now defined by the solution $\underline{a}$ to

$$\underline{R} \, \underline{a} = \underline{u}_1 = [\underbrace{1 \bullet\bullet\bullet 1}_{N_1} \underbrace{0 \bullet\bullet\bullet 0}_{N_T-N_1}]^T. \tag{5}$$

This correlation SDF thus forces the true-class peak to 1, the false class peak to 0 and the sidelobes ($d_S$ pixels from the peak) to 0 for both $\underline{f}_i$ and $\underline{g}_i$. Thus, the true correlation peak will have a well-defined shape. False targets will have low response over most of the central correlation region. Use of more training set images with shifts $2d_s$, etc. can control the full correlation plane response. This correlation SDF synthesis concept first introduced in [2] is a refinement of the decorrelation SDF in [3]. Other variations follow directly [2] such as: a least square solution (to reduce the dimensionality of the data), an SDF that maximizes the peak to sidelobe ratio (PSR) (rather than forcing the peak and sidelobes to specific values), etc. In Section 4, we present new test data on the performance of these correlation SDFs.

## 4. TEST RESULTS

To test the performance of these correlation SDFs, available software that produced images of different aircraft at different in-plane rotations $\theta$ and scales and from different viewing angles $\phi$ (out-of-plane rotations) was used. We selected two aircraft (Set A: Class 1 = Mig, Class 2 = DC10; and Set B: Class 1 = Mig, Class 2 = F105), $d_s$ = 5 and 8 pixels, $N_s$ = 5 (and thus $N_T$ = 10N, where N is the number of training set images per class). We generated 36 images ($10^\circ$ in-plane rotation increments $\Delta\theta$ = $10^\circ$) per class and thus desire N < 36. The image resolution used was 128 x 128.

With $\phi$ = 0, $d_s$ = 5, N = 6 ($\Delta\theta$ = $60^\circ$) was used for Set A. The 6 training set images per class are shown in Figure 1. Three intermediate images per class ($\theta$ = $15^\circ$, $30^\circ$, $45^\circ$) are shown in Figure 2. The correlation SDF was formed (Section 3). Tests of the full correlation plane data for the training set data are shown in Table 1. The value of the correct peak $I_p$, the largest peak $I_p$ and PSR at both peaks are listed. For true targets (class 1), the correct and largest peak coincide. All $I_p$ are 1.0 as expected and PSR is large ($\approx 3.88$) and rather constant. Data for the false targets (right side of Table 1) show the expected values (0.0) at the central peak and large ($\approx 0.52$) but less than 1.0 peak values

at other locations. As expected, PSR at these points is less (2.00 max) than for true targets. Data for test set imagery (Table 2) shows larger peaks ($I_p \geq 0.53$) for true targets than for false targets ($I_p \leq 0.50$) and larger PSR for true targets ($\geq 2.21$) than for most false targets ($\leq 2.1$). The PSR = 2.4 value for one false target corresponds to an $I_p$ = 0.35 and is thus easily distinguished.

From these data, we see that an $I_{pT}$ = 0.5 threshold alone provides 100% correct recognition. The combination of $I_{pT}$ = 0.5 and $PSR_T$ = 2.3 insures even more reliable performance. Because of symmetry, the three test data in Table 2 typify all results. Tests of $\phi$ rotation effects were conducted. They are more severe conditions and require more training set images, different $d_s$ and tighter $I_{pT}$ and $PSR_T$ thresholds.



(a) Class 1 Mig Training Set Images, $\phi=0$, $\Delta\theta=60$



(b) Class 2 DC10 Training Set Images, $\phi=0$, $\Delta\theta=60$
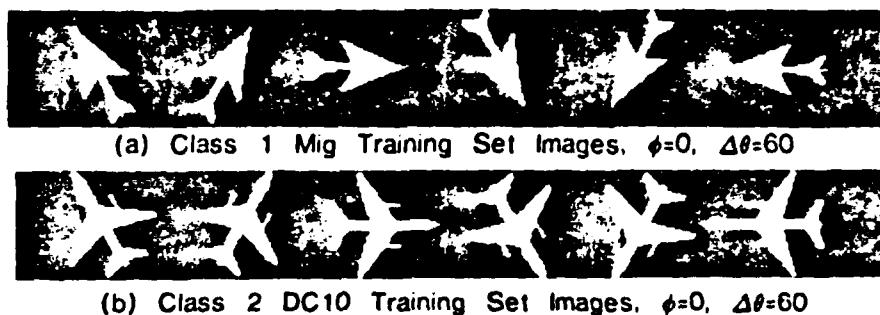
Figure 1:     Training Set Images Used



(a) Class 1



(b) Class 2

Figure 2:     Three Typical Test Images Per Class

| θ | CLASS 1 (MIG) TRUE CLASS | | | CLASS 2 (DC10) FALSE CLASS | | | |
|---|---|---|---|---|---|---|---|
| | LOCATION | $I_p = \overline{I}_p$ | PSR | $I_p$ | $\overline{I}_p$ | LOCATION | 'PSR |
| 60° | (65,65) | 1.00 | 3.90 | 0.0 | 0.63 | (60,61) | 1.14 |
| 120° | (65,65) | 1.00 | 3.86 | 0.0 | 0.51 | (61,69) | 1.01 |
| 180° | (65,65) | 1.00 | 3.94 | 0.0 | 0.61 | (62,50) | 1.92 |
| 240° | (65,65) | 1.00 | 3.88 | 0.0 | 0.54 | (60,61) | 2.00 |
| 300° | (65,65) | 1.00 | 3.81 | 0.0 | 0.43 | (80,61) | 1.57 |
| 360° | (65,65) | 1.00 | 3.84 | 0.0 | 0.51 | (66,59) | 1.35 |

Table 1:     Peak Intensity $I_p$, Largest $\overline{I}_p$, and PSR (Training Set Data)

| θ | CLASS 1 (TRUE CLASS) INPUT | | | | | CLASS 2 (FALSE CLASS) INPUT | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\bar{I}_p$ | PSR | LOCATION | $I_p$ | PSR | $\bar{I}_p$ | PSR | LOCATION | $I_p$ |
| 15° | 0.71 | 2.4 | (69,70) | 0.43 | 3.1 | 0.43 | 2.1 | (44,66) | None |
| 30° | 0.82 | 3.3 | (69,53) | 0.51 | 2.2 | 0.50 | 1.6 | (52,64) | None |
| 45° | 0.53 | 2.4 | (64,71) | 0.48 | 3.0 | 0.35 | 2.4 | (60,61) | None |

Table 2:   Typical Test Image Data

## REFERENCES

1. D. Casasent, Applied Optics, 23, 1620, May 1984.

2. W.T. Chang, D. Casasent and D. Fetterly, Proc. SPIE, 507, August 1984.

3. C. Hester and D. Casasent, Proc. SPIE, 302, 108, August 1981.

# 15. ERROR SOURCE MODELS FOR OPTICAL LINEAR ALGEBRA PROCESSORS

Ref. 19

# Optical linear algebra processors:
# noise and error-source modeling

**David Casasent and Anjan Ghosh***

*Department of Electrical and Computer Engineering, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213*

The modeling of system and component noise and error sources in optical linear algebra processors (OLAP's) are considered, with attention to the frequency-multiplexed OLAP. General expressions are obtained for the output produced as a function of various component errors and noise. A digital simulator for this model is discussed.

Optical linear algebra processors (OLAP's) represent a most attractive class of general-purpose optical processors with parallel and real-time features.[1] The frequency-multiplexed OLAP[2] is easily fabricated, permits a competitive high computation rate, and with different data-encoding schemes allows all the basic operations of linear algebra functions to be performed with excellent pipelining and flow of data.[3] We thus emphasize this architecture in our present study. Many OLAP's that operate on digital data have also been suggested.[1] These systems achieve the accuracy of a digital processor together with the speed and parallel-processing advantages of optical systems. Despite this widespread interest, little attention[4] has been given to an analysis and modeling of the various noise and error sources in such optical architectures. We briefly review the frequency-multiplexed OLAP and the basic linear algebra operations required. Then we detail the types of errors possible in such a processor and derive our model for noise- and error-source effects in OLAP's and the expression for the output obtained as a function of the various system-component noise and errors. We discuss digital simulation of this model and its use. The modeling, simulation procedure, and general approach that we use are valid for most OLAP's, including digital-optical linear algebra processors.

A simplified diagram of the frequency-multiplexed OLAP is shown in Fig. 1. This architecture consists of $N$ input point modulators imaged through $N$ separate regions of an acousto-optic (AO) cell (with each region separated by a bit time $T_B$). The AO cell is fed with $N$ 1-D input signals, each on a different temporal-frequency carrier. We view these signals as $N$ vectors, each on a spatial carrier. The light intensity distribution leaving the cell is then the products of the input vector (from the point modulators) and the $N$ vectors in the cell, with each such product leaving the cell at an angle proportional to the input frequency to the cell. The Fourier-transform (FT) lens sums the elements of each vector product (by space integration) and forms each of the $N$-vector inner products on a separate output detector. The detector output voltages (or currents) are thus proportional to the ($N \times N$) matrix–vector product, with one matrix–vector multiplication performed each $T_B$.

If intensity-mode operation is used, the signals to be

processed are present on a bias. The effects of these bias terms in the output data must be removed and corrected for. The necessary correction signals can be easily obtained with a separate adjunct processor channel similar to the way in which bias was corrected in the initial optical matrix–vector processors using two-dimensional masks. Amplitude-mode operation of the AO cells and the system is also possible and in some cases preferable. In the conventional system, the detected output intensity will be the square of an amplitude product, and thus the square root of the input (or output) data must be produced. Methods to achieve this exist, but coherent detection at the output is preferable. In this case, the detector output voltage is proportional to the desired amplitude product. Either mode of operation requires attention to the choice of frequencies and their separation to ensure linearity and suppression of cross talk. The effects of intermodulation-induced cross talk require further examination.

No delays exist in this processor since data flow continuously, as detailed elsewhere,[3] even though the same matrix remains in the AO cell for $NT_B$. With different space ($x$), time ($t$), and frequency ($f$) encoding, matrix data can be processed by the system, and various matrix–vector, matrix–matrix, and matrix–matrix–matrix multiplications and iterative and direct solutions of systems of linear algebra equations can be realized.[3] The basic operation performed by the system is thus a matrix–vector product each $T_B$. This is the basic building block of all other matrix operations and direct and indirect solutions of linear and nonlinear algebraic equations.[3] In this Letter, we describe our noise- and error-source modeling of the frequency-multiplexed OLAP in terms of this basic **Ab** = **c** system operation.
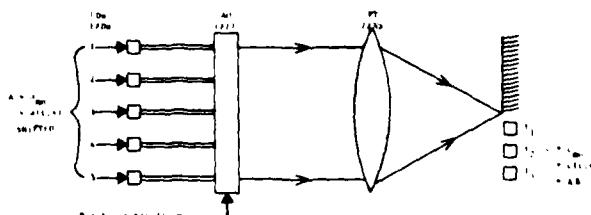


Fig. 1. Simplified schematic of a frequency-multiplexed optical linear algebra processor. (After Ref. 2.)

The basic architecture of most OLAP's consists of a linear array of input point modulato.s, an AO cell, and a detector array. Thus limiting our modeling to the system of Fig. 1 is not overly restrictive. In the initial modeling, we assume ideal lenses, no dispersion, and no cross talk. This yields a useful closed-form expression for the effect of errors, which provides useful insight. Accommodating other effects and more advanced component errors in the simulator is discussed below. In the system of Fig. 1, various input-plane (point-modulator) errors are possible. These include variations in the bias level or level of lasing for the input modulators and variations in the response of each input point modulator. Acoustic attention of the signal in the cell produces a deterministic taper $\exp(-\alpha x)$ across the length $x$ of the AO cell, where $\alpha$ is the attenuation constant of the AO cell material used. For now, we assume that $\alpha$ is small and nondispersive. These spatial errors plus variations in the spatial response of the AO cell owing to imperfections in the AO material or the transducer used can also be modeled as input-plane errors. These spatial errors are correctable and can be reduced to low residual levels by adjusting the input signals to the point modulators and the AO cell or by use of a correction mask in front of the AO cell (the $\alpha$ error

$$\hat{b}_i = b_i[1 + \delta_{i1}^{(1)} + \delta_{i2}^{(1)} + \delta_{i3}^{(1)}]. \tag{1}$$

Similarly, the actual and ideal transmittance of the matrix data in the AO cell for element $j, i$ ($i$ denotes space and $j$ denotes frequency) are related by

$$\hat{a}_{ji} = a_{ji}[1 + \delta^{(2)}]H(f_j)\exp(-\alpha x_i), \tag{2}$$

where $x_i$ denotes the distance of the $i$th data block from the AO cell's transducer. Likewise, the elements of the observed and ideal detector plane outputs $\hat{s}$ and $s$ are

$$\hat{s}_j = s_j[1 + \delta_j^{(3)}] + d_j + n_j(t). \tag{3}$$

We combine all spatial errors (subscript $i$) into the single variable

$$\psi_i = \delta_{i1}^{(1)} + \delta_{i2}^{(1)} + \delta_{i3}^{(1)} + \delta^{(2)} + \delta_i^{(2)}. \tag{4}$$

Combining Eqs. (1)–(4) and assuming all error sources to be small, the elements $\hat{c}_j$ of $\hat{c}$ are

$$\hat{c}_j = \sum_i a_{ji}b_i(1 + \psi_i)(1 + \delta_j)H(f_i)\exp(-\alpha x_i) + d_j + n_j(t). \tag{5}$$

To provide a more vivid description relating $\hat{c}$ to c and the various system and component noise and errors, we detail Eq. (5) for a $2 \times 2$ matrix as

$$\begin{bmatrix} \hat{c}_1 \\ \hat{c}_2 \end{bmatrix} = \begin{bmatrix} 1 + \delta_1^{(3)} & 0 \\ 0 & 1 + \delta_2^{(3)} \end{bmatrix} \begin{bmatrix} H(f_1) & 0 \\ 0 & H(f_2) \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\begin{bmatrix} 1 + \psi_2 & 0 \\ 0 & 1 + \psi_1 \end{bmatrix} \begin{bmatrix} \exp(-\alpha x_2) & 0 \\ 0 & \exp(-\alpha x_1) \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix}. \tag{6}$$

To provide further insight, we explicitly describe each error-matrix term in Eq. (6) by its associated origin, i.e.,

$$\begin{bmatrix} \hat{c} \\ \text{Observed} \end{bmatrix} = \begin{bmatrix} \text{Detector} \\ \text{Spatial-} \\ \text{Response} \\ \text{Variations} \end{bmatrix} \begin{bmatrix} \text{AO} \\ \text{Frequency-} \\ \text{Response} \\ \text{Variations} \end{bmatrix} \begin{bmatrix} \text{A} \\ \text{Exact} \\ \text{Matrix} \end{bmatrix} \begin{bmatrix} \text{Input} \\ \text{Spatial-} \\ \text{Response} \\ \text{Variations} \end{bmatrix}$$

$$\begin{bmatrix} \text{Acoustic} \\ \text{Attenuation} \end{bmatrix} \begin{bmatrix} \text{b} \\ \text{Exact} \\ \text{Vector} \end{bmatrix} + \begin{bmatrix} \text{Detector} \\ \text{Dark} \\ \text{Current} \end{bmatrix} + \begin{bmatrix} \text{Detector} \\ \text{Noise} \end{bmatrix}. \tag{7}$$

effect can be corrected only at one frequency, however). As noted above, all spatial errors in the AO cell can be mapped into spatial input-plane errors. Similarly, any frequency-dependent AO cell errors can be mapped to the output plane (since the FT lens converts frequency in the AO plane into position in the detector plane). The output detector plane errors thus include variations in the frequency response $H(f)$ of the AO cell, variations in the spatial response of the individual output detectors, variations in the dark current of the individual detectors, and time-varying detector noise. The effect of these last two detector plane errors on the system output is additive rather than multiplicative, as we will shortly demonstrate.

In Table 1, we summarize the notation used and the various input, AO, and detector plane errors. We also include errors that describe spatial variations in the coupling between the inputs and the AO cell. With this formulation and notation, the elements $\hat{b}_i$ of the actual input vector are related to the elements $b_i$ of the ideal input vector by

When acoustic attenuation is small, Eq. (6) is

$$\begin{bmatrix} \hat{c}_1 \\ \hat{c}_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix}, \tag{8}$$

where the spatial and temporal errors are now additive. From Eqs. (7) and (8), and the fact that all OLAP spatial errors can be reduced to the desired residual levels by correction, detector noise and the dispersive nature of $\alpha$ are potentially the most dominant error sources. If $\alpha$ effects are not small, then the decoupling in Eq. (8) does not occur, the various spatial and detector plane errors can still be grouped and combined as in Eq. (7), but the simplified form in Eq. (8) does not result.

To quantify system performance and the effect of each noise- and error-source component in the OLAP for a given operation, computer simulation is required. The error sources are quite different from those typically treated in analysis of conventional linear algebra processors.[5] We now briefly discuss how we digitally model the various error sources in Eqs. (6) and (7).

**Table 1.   SAOP Error Source Model**

| Error Source | Notation |
| --- | --- |
| Spatial errors | Subscript $i$ |
| Frequency errors | Subscript $j$ |
| Input plane errors | Superscript 1 |
| AO cell errors | Superscript 2 |
| Detector-plane errors | Superscript 3 |
| **Input Plane Errors** | **Notation** |
| Point modulator | |
|   Spatial gain | $1 + \delta_{i1}^{(1)}$ |
|   Bias nonuniformity | $1 + \delta_{i2}^{(1)}$ |
| Coupling (spatial) | $1 + \delta_{i3}^{(1)}$ |
| **AO Cell Plane Errors** | **Notation** |
| Amplifier errors | $1 + \delta^{(2)}$ |
| Spatial response | $1 + \delta_i^{(2)}$ |
| AO transfer function | $H(f_j)$ |
| Acoustic attenuation | $\exp(-\alpha x_i)$ |
| **Detector Plane Errors** | **Notation** |
| Spatial response | $1 + \delta_j^{(3)}$ |
| Dark current | $d_j$ |
| Time-varying noise | $n_j(t)$ |

From experiments on our laboratory OLAP systems, we found that the residual spatial errors and the detector noise can be modeled as zero-mean Gaussian random numbers and that signal-dependent (quantum) noise is not present. The frequency response $H(f)$ and the acoustic attenuation can be modeled as deterministic errors that are quantified by measurements on the OLAP. This deterministic function multiplies the matrix data in the cell as in Eq. (6). Since the spatial errors are independent of time, the random numbers representing each such error are generated once by standard IMSL[6] or other software and stored. The $3\sigma$ standard deviation of each random number is chosen to equal the percentage error to be modeled. For input and AO cell spatial errors, the random numbers are included in each input vector datum $\mathbf{b}$ each $T_B$, and for detector spatial errors the associated random numbers are added to the computed output vector each $T_B$ as in Eq. (6). For fixed or spatial errors, the same set of random numbers is used at each $T_B$. To simulate detector noise, a new set of uncorrelated variables with Gaussian probability distribution is generated each $T_B$.

The model above and the form of the result in Eqs. (6)–(8) are useful for conveying error effects in closed form, for showing how various error sources can be grouped, and for noting which error sources are correctable, multiplicative, and additive. Other error sources and other models for the various components can be included directly in the simulator [but do not lend themselves to convenient diagonal matrices as in Eq. (6) and to a closed-form expression for the system]. Variations in the bias level of the point modulators and all errors are assumed to be small *residual* errors (after correction). Thus bias-level variations are included in $\psi_i$. If such individual errors are not small, performance will be too poor to consider. A primary purpose of our initial model and its simulator is to quantify the dominant error sources and the magnitude allowed for each (i.e., the level to which fixed spatial errors must be corrected and the amount of noncorrectable errors allowed).

For quantitative performance data, other advanced models can be used. Exact transfer curves (after correction) for each point modulator and detector can be measured and used in the actual simulator. We have done this and found the results (for the small residual errors present in practice) to be the same as those obtained using our random variable modeling. To include the dispersive nature of $\alpha$, a different $\exp(-\alpha x)$ factor is used for each signal in the AO cell. This is a fixed factor (different for each frequency signal) that multiplies the present spatial contents of the cell each $T_B$. Our simulator includes this feature, but it is not conveniently included in the equation formulations above. Similar remarks apply to cross-talk effects in the AO cell and to the electronic circuit models.

From detailed simulations and analyses with the model in Eq. (6), we found that acoustic attenuation and detector noise are the dominant error sources. In initial simulations,[4] we found that $\alpha$ effects are dominant in iterative algorithms and detector noise is dominant in direct algorithms. We also found that the effects of small multiple-error sources are additive as in Eq. (8).

The various error sources that arise in an OLAP have been tabulated and grouped into two classes (correctable or fixed and time-varying) and classified according to the plane (input, AO cell, output detectors) in which they originate. Combining these separate error sources, we find that error matrices in systolic processors are multiplicative and that acoustic attenuation is an important error source in OLAP's employing AO cells. The model and simulation technique advanced can and should be applied to other OLAP's to quantify the dominant error sources, the effect of multiple errors, and the performance to be expected from each system for each application and algorithm.

  * Present address, AT&T Bell Laboratories, Allentown, Pennsylvania 18103.

### References

1. Special issue on optical computing, Proc. IEEE 72 (July 1984).
2. D. Casasent, J. Jackson, and C. P. Neuman, Appl. Opt. 22, 115 (1983).
3. D. Casasent, Proc. IEEE 72, 831 (1984).
4. D. Casasent, A. Ghosh, and C. P. Neuman, Proc. Soc. Photo-Opt. Instrum. Eng. 431, 201 (1983).
5. J. Wilkinson, *Rounding Errors in Algebraic Processes* (Wiley, New York, 1963).
6. *International Mathematics and Statistics Library Reference Manual*, 8th ed. (IMSL, Houston, Tex., 1980).

# 16. A QUADRATIC NONLINEAR MATRIX ALGORITHM

Ref. 20

# ITERATIVE SOLUTIONS TO NONLINEAR MATRIX EQUATIONS USING A FIXED NUMBER OF STEPS

D. Casasent, A. Ghosh* and C.P. Neuman

Carnegie-Mellon University
Department of Electrical and Computer Engineering, Pittsburgh, Pennsylvania 15213

*Present Address: AT&T Bell Laboratories, Allentown, Pennsylvania 18103

## ABSTRACT

An iterative algorithm for the solution of a quadratic matrix equation (the algebraic Ricatti equation) is detailed. This algorithm is unique in that it allows the solution of a nonlinear matrix equation in a finite number of iterations to a desired accuracy. Theoretical rules for selection of the operation parameters and number of iterations required are advanced and simulation verification and quantitative performance on an error-free processor are provided. An error source model for an optical linear algebra processor is then advanced, analyzed and simulated to verify and quantify our performance guidelines. A comparison of iterative and direct solutions of linear algebraic equations is then provided. Experimental demonstrations on a laboratory optical linear algebra processor are included for final confirmation. Our theoretical results, error source treatment and guidelines are appropriate for digital systolic processor implementation and for digital-optical processor analysis.

## 1. INTRODUCTION

Optical linear algebra processors (OLAPs) represent a most general and attractive use of the parallelism and real-time processing features of optical systems [1]. The frequency-multiplexed acousto-optic (AO) processor [2,3] of Figure 1 represents a most general-purpose OLAP architecture with ease of fabrication [4] and competitive computational rates [2,4]. In this architecture (Figure 1), N point modulator inputs are imaged through N separate regions of an AO cell. These individual regions are separated by $T_B$ of time (for propagation of the acoustic wave) and by a physical distance $d_B$. In [2], the use of this processor in iterative algorithms, direct LU and QR matrix decomposition algorithms, and triangular system solutions was detailed.
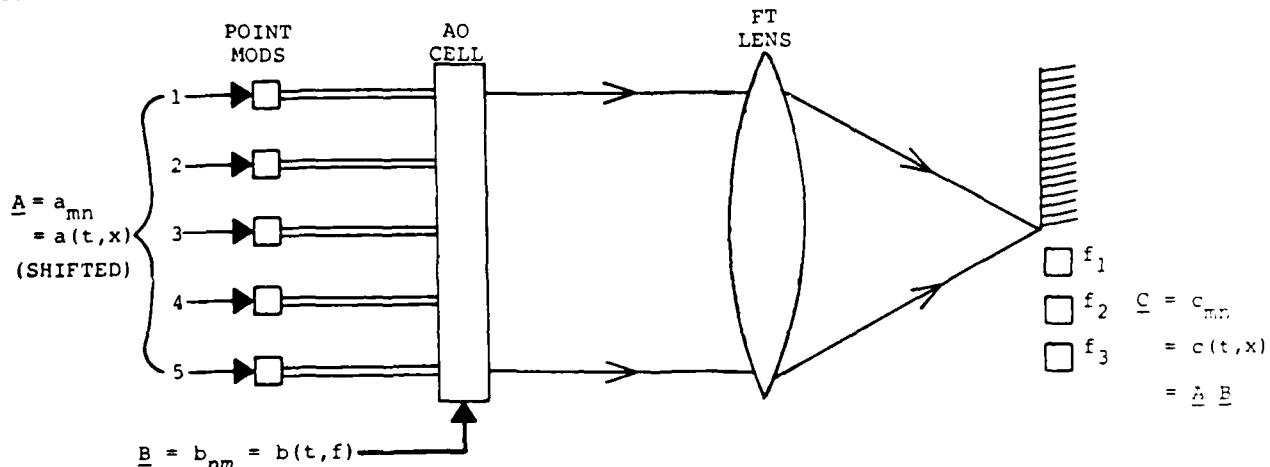


FIGURE 1
Simplified schematic of a frequency-multiplexed optical linear
algebra processor [3]

In this paper, we consider the use of this processor for the solution of a nonlinear matrix equation (Section 2). The specific application chosen is the solution of the algebraic Ricatti equation (ARE). This nonlinear equation is similar to the expressions to be solved in Kalman filtering and other advanced modern signal processing algorithms. An iterative solution is necessary for such problems and for eigensystem solutions. Our proposed nonlinear ARE solution is quite unique since it requires a finite number of steps to achieve a specific accuracy and performance. In Section 3, we summarize selection of the operational parameters for such an iterative algorithm and the theoretical basis for our choice of

the fixed number of iterations to be used. Section 4 presents initial error-free simulation data. In Section 5, we advance our error source model. In Section 6, we review our iterative and direct solutions to systems of linear algebraic equations (LAEs). This represents the fundamental operation required in advanced linear algebra algorithms. Section 7 contains simulation data to quantify the dominant error sources and the accuracy expected from such algorithms. We conclude in Section 8 with the experimental verification and quantification of our theoretical results. Our summary and conclusions are then advanced in Section 9.

## 2. NONLINEAR MATRIX SOLUTION

In reference [5], we detailed a solution to the linear quadratic regulator control problem to minimize a quadratic performance index for a linear system. Computation of the regulator feedback gain matrix $\underline{K}$ that defines the optimal controls $\underline{u}$ involves the solution of the ARE

$$\underline{S} \; \underline{F} + \underline{F}^T \underline{S} - \underline{S} \; \underline{L} \; \underline{S} + \underline{Q} = \underline{0} \tag{1}$$

for $\underline{S}$. To achieve this, we used the Kleinman algorithm [5] and the solution of the vectorized Lyapanov equation to format the solution of (1) as a solution of the set of LAEs

$$\underline{H}(k)\underline{s}(k) = \underline{y}(k), \tag{2}$$

where $\underline{s}$ and $\underline{y}$ are the vectorizations of $\underline{S}$ and $\underline{S} \, \underline{L} \, \underline{S} - \underline{Q}$ respectively and $\underline{H}$ is a Kronecker formatted matrix. This system of LAEs must be solved successively with different matrices $\underline{H}$ and vectors $\underline{y}$ with the results of one cycle used to compute the matrix $\underline{H}$ and vector $\underline{y}$ for the next cycle. To achieve this, we employ a two-loop iterative algorithm described by

$$\underline{s}(r+1,k) = [\underline{I} - \_(k)\underline{H}(k)]\underline{s}(r,k) + \_(k)\underline{y}(k). \tag{3}$$

In solving (2) using (3), we solve (2) for one outer loop iteration k, update $\underline{H}$ and $\underline{y}$ and solve the next LAE. This procedure continues until $\underline{s}$ is of sufficient accuracy. The algorithm in (3) implies an iterative solution for each LAE. Direct solutions are also possible as we discuss in Sections 6 and 7. The indices r and k in (3) refer to Richardson (inner) and Kleinman (outer) loop iterations respectively.

## 3. OPERATIONAL PARAMETER SELECTION

In an iterative algorithm such as (3), various operational parameters must be selected. The initial selection $\underline{s}(0,0)$ for $\underline{S}$ and the choice $\underline{s}(0,k)$ for each LAE solution are required. For $\underline{s}(0,0)$, we use $\underline{0}$ to insure outer loop convergence (a stability matrix). For $\underline{s}(0,k)$, we use the obvious choice of the prior $\underline{s}(0,k-1)$ estimate. The acceleration parameter $\_$ in (3) is chosen to be $\_ = n/\`_{max} = 3/ \| \underline{H}(k) \|$ . This insures inner loop convergence [2,5]. Stopping the inner loop ierations (index r) for each LAE solution and stopping the number of outer loop iterations (index k) is a major decision.

In reference [5], we derived bounds for the inner loop error, the outer loop error and their coupling. From this analysis, we derived the selection of a fixed number of inner loop iterations R to solve each LAE given by

$$R = nC = C \log \alpha \approx 1.5C \text{ to } 3.0C. \tag{4}$$

where $\| \underline{x}^*(0) - \underline{x}^*(1) \| \leq \alpha$ and $[1 - 1/C]^R \approx \exp(-n) < 1/\lambda$ is chosen. This follows from our analysis of the error in an iterative solution (due to a fixed number of iterations R), which showed that the norm of such an error is

$$\| \underline{s}(r,k) - \underline{s}^* \| = \| \underline{I} - \_ \underline{H}(k) \|^r = (1 - 1/C(k))^r, \tag{5}$$

where C is the condition number of $\underline{H}$. Since r is expected to increase with C, we set $r = nC$ and thus select n such that the error between the computed solution s and the exact solution $s^*$ in (5) is as small as is required. For the fixed number of outer loop iterations K, we use K = 5 or 6, which can be theoretically derived (and appropriately modified) for other applications with matrices with specific features. These iterative operational parameter selections are summarized in Table 1.

## 4. ERROR-FREE SIMULATION RESULTS

The performance measures we adopted to assess performance of the algorithm in Section 2 implemented using the operational parameters in Table 1 are the maximum percent error in any element of the matrix $\underline{K}$ (i.e. $\Delta K_{max}\%$) and the maximum error in the location of the closed-loop poles of the system ($\Delta\`_{max}\%$). We expect $\Delta K >> \Delta\`$ and note that $\Delta\`$ is the more

appropriate error measure for this specific application and that similar error measures should be used to evaluate the performance of other specific case studies. In Figures 2 and 3, we show the variation of these two error measures with the number of outer loop iterations k for a fixed number of inner loop iterations for two case studies. These case studies are the fifth (Figure 2) and third (Figure 3) order models of an F100 engine. As seen from the data for these two case studies, the use of a fixed number of iterations results in a monotonic decrease in the solution error with the $\Delta K$ error being approximately ten times that of the $\Delta\lambda$ error. From these results, we conclude that the use of a fixed number of iterations can yield adequate results when the number of iterations is properly chosen. Our parameter selection guidelines in Table 1 have thus all been verified and discussed.

TABLE 1
Operational Parameter Selection Guidelines [5]

| SYMBOL | PARAMETER | PREFERRED CHOICE |
|--------|-----------|------------------|
| $\underline{s}(0,0)$ | Initial Initialization | $\underline{s}(0,0) = \underline{0}$ |
| $\underline{s}(0,k)$ | k-th Kleinman Loop Initialization | $\underline{s}(0,k) = \underline{s}(0,k-1)$ |
| R | Number of Inner Loop Iterations | $R = 1.5C$ to $3.0C$ |
| K | Number of Outer Loop Iterations | $K = 5 - 6$ |
| $\omega(k)$ | Acceleration Parameter | $\omega(k) = 3/\|\underline{H}(k)\|$ |
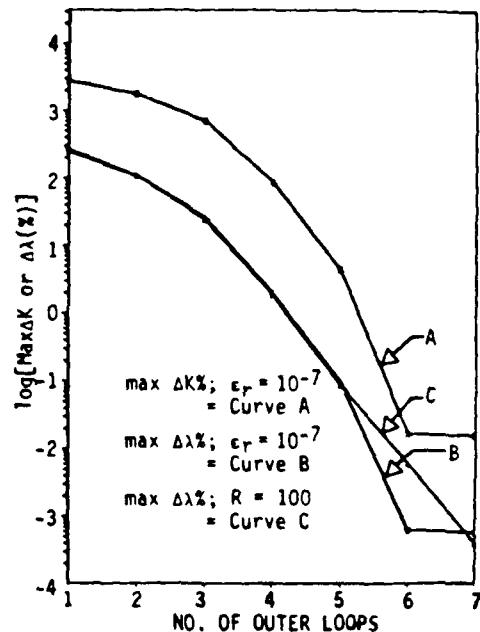


FIGURE 2
Variation of the error measures $\Delta K_{max}(\%)$ and $\Delta\lambda_{max}(\%)$ with the number of outer-loop iterations K for different inner-loop iteration stopping criteria for the fifth-order HPG3 F100 model
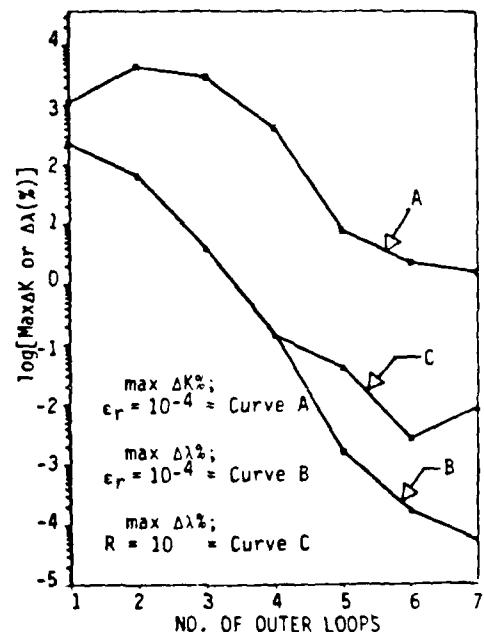
FIGURE 3
Variation of the error measures $\Delta K_{max}(\%)$ and $\Delta\lambda_{max}(\%)$ with the number of outer-loop iterations K for different inner-loop iteration stopping criteria for the third-order HPG3 F100 model

## 5. ERROR SOURCE MODEL

In earlier publications [7,8] we detailed the first system and component error source model for an OLAP and the general issue of errors in such an architecture. In this section, we review this OLAP error source model. In this model, we distinguish input, AO cell and detector plane errors separately. Spatial errors include: input and detector response variations and errors in the interconnections between the input modulators and the AO cell, and detector dark current. The spatial variations are fixed (time-independent) and are

correctable to small residual levels as required (by adjusting the gain of the input point modulators, the detector amplifiers, and the input matrix and vector data). Detector noise is the only time-varying error source considered. Acoustic attenuation produces a deterministic exponential variation of the data in the AO cell. This effect is dispersive, but its frequency dependence is not included in our present model. Acoustic attenuation can be corrected at one frequency and is thus an input spatial error. The product of an input matrix $\underline{A}$ and vector $\underline{b}$ thus yields a final output $\underline{d}$ given by

$$\hat{\underline{d}} = \begin{bmatrix} \text{Detector} \\ \text{Spatial} \\ \text{Response} \\ \text{Variations} \end{bmatrix} \begin{bmatrix} \text{AO Cell} \\ \text{Frequency} \\ \text{Response} \\ \text{Variations} \end{bmatrix} \begin{bmatrix} \text{Data} \\ \text{Matrix} \\ \underline{A} \end{bmatrix} \begin{bmatrix} \text{AO Cell} \\ \text{Attenuation} \end{bmatrix} \begin{bmatrix} \text{Point Mod} \\ \text{Response and} \\ \text{Interconnection} \\ \text{Variations} \end{bmatrix} \begin{bmatrix} \text{Data} \\ \text{Vector} \\ \underline{b} \end{bmatrix} + \begin{bmatrix} \text{Detector} \\ \text{Dark} \\ \text{Current} \end{bmatrix} + \begin{bmatrix} \text{Time-} \\ \text{Varying} \\ \text{Detector} \\ \text{Noise} \end{bmatrix} . \quad (6)$$

As seen, the different types of system and component variations are described by error matrices that multiply the input data vector or input matrix data. Thus, the system errors are described by the corresponding variations in the data matrix and vector. The detector dark current and noise appear additively in the output vector as shown in Eq. (6).

## 6.  DIRECT AND INDIRECT SOLUTIONS OF LAEs

The solution of a system of LAEs, $\underline{A}\,\underline{x} = \underline{b}$ is the fundamental operation required in most linear algebra processors and signal processing applications. Thus, we concentrate on this function. The two major types of LAE solutions are direct or matrix decomposition solution and an iterative or indirect solution.

The preferable iterative algorithm is [2,9]

$$\underline{x}(r+1) = \underline{x}(r) + \omega[\underline{b} - \underline{A}\,\underline{x}(r)], \quad (7)$$

where $\omega$ is an acceleration parameter chosen to insure convergence. The iterations (described by the iterative index r) continue until $\underline{x}(r) = \underline{x}(r+1)$. Then, (7) reduced to $\underline{A}\,\underline{x} = \underline{b}$ and the system's output $\underline{x}$ is the desired solution. To implement (7) on the system of Figure 1, the matrix data is fed to the AO cell one column at a time in parallel with the rows of the matrix frequency-multiplexed, i.e. with the matrix elements $a_{mn}$ encoded in time and frequency as $a(f,t)$ and with the vector data $\underline{x}$ spatially-multiplexed as $x(x)$ and fed in parallel to the input point modulators. The matrix-vector product $\underline{A}\,\underline{x}$ is formed, operated upon in analog or digital post-processing electronics to produce the right-hand side of (7) and hence the new $\underline{x}$ iterate input to the point modulators. Thus, the detector output is fed back to the input point modulators. The length of the AO cell $NT_B$ is chosen to be just as sufficient to accommodate the matrix data. Each $T_B$, as one column of the matrix leaves the AO cell, it is reintroduced into the bottom of the cell. This recycling of the matrix data is more efficient for system fabrication and reduces the effects of acoustic attenuation.

In direct solutions, the matrix $\underline{A}$ and the vector $\underline{b}$ are multiplied by a decomposition matrix $\underline{P}_1$ to generate new $\underline{A}_1$ and $\underline{b}_1$. Each such matrix-matrix and matrix-vector multiplication generates one row of the final $\underline{A}'$ matrix and one element of the final $\underline{b}'$ vector. After each matrix-matrix multiplication, the order of the matrix and vector are reduced by one and the new reduced order $\underline{A}_1$ and $\underline{b}_1$ are multiplied by a new $\underline{P}_2$. This procedure is repeated N-1 times (for an N x N matrix) and yields a new upper-triangular matrix $\underline{U}$ and a new vector $\underline{b}'$. This simplified upper-triangular system of equations $\underline{U}\,\underline{x} = \underline{b}'$ is then easily solved by back-substitution. The matrix-decomposition can be realized either as an LU decomposition (this is the technique we use when the matrix is positive-definite or diagonally-dominant, as is the case here, since pivoting is then not required) or as a QR orthogonal decomposition (this technique is more general and stable, but is more difficult to realize). The detailed implementation of LU [2,10] and QR [2,11] decomposition and back-substitution [2,12] have been described elsewhere. To implement the Gaussian-elimination algorithm (LU) used in the present application on the system of Figure 1, we feed one row of the matrix $\underline{A}$ to the AO cell in parallel (with the columns of $\underline{A}$ frequency-multiplexed, i.e. with the elements $a_{mn}$ of $\underline{A}$ frequency and time encoded as $a(t,f)$) and with one row of the decomposition matrix $\underline{P}_j$ fed to the input point modulators in parallel (with the elements $p_{mn}$ of $\underline{P}$ time and space encoded as $p(t,x)$). To facilitate data flow and for speed, we simultaneously operate on $\underline{A}$ and $\underline{b}$ by using an augmented matrix. One row of the augmented matrix $\underline{A}'$ is produced in parallel as $a'(t,x)$ on the output detector during each of the N cycles. The new $\underline{P}_j$ matrix is easily calculated from the elements of the j-th column of the augmented matrix in dedicated electronics.

## 7.  SYSTEM ERROR EFFECTS ON THE SOLUTION OF LAEs

The direct solution requires an AO cell of twice the length of the matrix, but achieves the decomposition in a fixed number of steps. However, as noted in Section 3, iterative

algorithms can be operated with a fixed number of iterations to achieve a given desired accuracy and iterative algorithms are essential [2] for eigen-systems solutions and the solution of nonlinear matrix equations such as the ARE [5] and in Kalman filtering [13]. In our new results (Sections 7 and 8),we compare [6] the performance of direct and iterative algorithms in the solution of the LAEs that arise in a specific ARE solution for the F100 engine. The two cases considered are third and fifth-order F100 models. These give rise to $9 \times 9$ and $25 \times 25$ matrices. Bipolar data is handled by space-multiplexing [3] and this doubles the size of the matrices and vectors required. For the third-order problem, $C = 2.48$, the dynamic range is 47.7 and from (5), $j = 10$ iterations are required to solve each LAE. For the fifth-order problem, $C = 56.9$, the matrix dynamic range is 1117 and from (5), $j = 100$ iterations are required to solve each LAE. We consider three solutions: an iterative algorithm, direct LU Gaussian-elimination with the back-substitution performed optically and direct Gaussian-elimination with the back-substitution performed digitally with high accuracy. We consider two problems: the solution of $A_5 x_5 = b_5$ for the fifth and last outer loop in (2) and (3) for the solution of the ARE in (1) with $A_5$ and $b_5$ digitally calculated exactly, and the solution of all five LAEs for all outer loop iterations.

TABLE 2

Performance of Three Algorithms for Two Data Sets in the Solution of One System of LAEs

| ALGORITHM | TEST NO. | F100 DATA SET | RESP. VARIATIONS Point Mods(%) | Dets(%) | ACOUSTIC ATTEN. (dB / cm) | DET RMS NOISE(%) | $\Delta x$ (%) | $\Delta_{max}$ (%) |
|---|---|---|---|---|---|---|---|---|
| (I) Iterative | 1 | 3 | 1 | 1 | 0.1 | 0.6 | 2.49 | $0.2 \times 10^{-3}$ |
|  | 2 | 5 | 1 | 1 | 0.001 | 0.06 | 4.31 | 1.3 |
| (II) LU and Optical Back-Substitution | 3 | 3 | 1 | 1 | 0.1 | 0.6 | 2.39 | 0.52 |
|  | 4 | 5 | 1 | 1 | 0.1 | 0.006 | 9.77 | 0.93 |
| (III) LU and Digital Back-Substitution | 5 | 3 | 1 | 1 | 0.1 | 0.6 | 3.04 | 0.33 |
|  | 6 | 5 | 1 | 1 | 0.1 | 0.006 | 6.78 | 0.71 |

TABLE 3

Performance of Three Algorithms for Two Data Sets in the Solution of the Nonlinear ARE

| ALGORITHM | TEST NO. | F100 DATA SET | RESP. VARIATIONS Point Mods(%) | Dets(%) | ACOUSTIC ATTEN. (dB / cm) | DET RMS NOISE(%) | $\Delta x$ (%) | $\Delta_{max}$ (%) |
|---|---|---|---|---|---|---|---|---|
| (I) Iterative | 7 | 3 | 1 | 1 | 0.1 | 0.6 | 2.98 | 0.77 |
|  | 8 | 5 | 1 | 1 | 0.001 | 0.06 | 5.24 | 1.62 |
| (II) LU and Optical Back-Substitution | 9 | 3 | 1 | 1 | 0.1 | 0.6 | 4.56 | 0.72 |
|  | 10 | 5 | 1 | 1 | 0.1 | $6 \times 10^{-4}$ | 11.34 | 1.44 |
| (III) LU and Digital Back-Substitution | 11 | 3 | 1 | 1 | 0.1 | 0.6 | 4.12 | 0.5 |
|  | 12 | 5 | 1 | 1 | 0.1 | $6 \times 10^{-4}$ | 10.1 | 1.1 |

In Table 2, we show the results for the solution of the single fifth set of LAEs. Our results for the full set of five LAEs, i.e. the full ARE solutions are included in Table 3. Data sets 3 and 5 refer to the third and fifth-order F100 matrix problems respectively. The performance measures used in evaluating each approach are the average norm $\Delta x$ of the error in the calculated vector $x$ and the maximum error $\Delta_{max}$ in the location of the closed-loop poles of the final system. The spatial, detector noise, and acoustic attenuation errors noted earlier were selected to produce approximately equal output errors for each error source treated separately.

In Tests 1 and 2, we see that our theoretical operational parameters (Table 1) are also valid when noise and system errors are present. Comparing the results for Algorithm I and II, we see that acoustic attenuation is the dominant error source for an iterative algorithm and detector noise dominates the performance of a direct algorithm. This is expected

because of the cyclic data flow of the matrix in the AO cell during the iterative algorithm. This alters C for the matrix. In the direct algorithm, detector noise on one cycle is fed back to both the inputs and the AO cell and thus changes the noise distribution and its effects accumulate. Also, detector noise affects the small vector elements and this effect also compounds on successive cycles. From the results of Algorithms II and III, we see that optical back-substitution yields comparable performance to digital back-substitution. This is expected, since the operations required in back-substitution are only vector inner products and only N-1 of these are required. This is a substantially lower computationally intensive set of operations than those required in the matrix decomposition. Thus, the accuracy of the matrix decomposition determines the final accuracy in our results. Comparing the results for data sets 3 and 5 and the corresponding data in Tables 2 and 3, we see that the larger matrix size and the increased number of steps required in the ARE versus the LAE solution causes the required accuracy to increase for direct algorithms more than for iterative algorithms (e.g. a lower acoustic attenuation constant $\alpha$ is noted to be required for the iterative ARE solution than for a direct LAE solution). We have derived a theoretical expression [6]

$$\alpha < (1/2.3LC) \qquad (7)$$

for the amount of acoustic attenuation $\alpha$ in dB/cm allowed for convergence of an iterative algorithm, where L is the length of the AO cell in cm. From the last two columns in both tables, we see that $\Delta\lambda_{max}$ errors are significantly less than $\Delta x$ errors as expected. The results in Tables 2 and 3 are in agreement with the theoretical guidelines in (7). From Test 1 and all other tests, we find that spatial errors are additive and that for small errors the percent performance scaled with the magnitude of the error. In Tables 2 and 3 and in (7), we assume that each $T_B$ of the AO cell corresponded to 1mm and we assumed new input data to the point modulators in the AO cell to be introduced every $T_B$. To achieve more practical levels, closer spacing of data packets in the cell is necessary. This can easily be obtained by scaling the values given in Tables 2 and 3. Operation of the input point modulators at a higher rate than the AO cell data [2] can also improve the $\alpha$ and detector noise values found in Tables 2 and 3. These initial test results are intended to provide guidelines for the efficient use of various algorithms, efficient solutions to linear and non-linear matrix equations, and quantitative data on performance expected. Our theory, guidelines, and modeling are also appropriate for digital-optical linear algebra architectures.

## 8. REAL-TIME LABORATORY EXPERIMENTS

In Figure 2, we show the nine outputs from a laboratory system to iteratively solve the fifth set of LAEs for the third-order F100 model (Test 1, Table 2). The outputs are shown after 80, 400 and 640 iterations. The laboratory system used a fixed 2-D photographic mask for the matrix in place of the AO cell and 2-D space-multiplexing in place of frequency-multiplexing. To accomodate bipolar data, the matrix and vector were biased positive. This increased C to 120. The laboratory system was operated at a 10MHz data rate per channel. To facilitate easy monitoring of the system, we used $\omega = -0.125$. The number of iterations $j = nC$ required for 0.6% accuracy was calculated from (3) to be 613 iterations. Our experimental value of 640 iterations at which convergence occurred is thus in excellent agreement with theory. In the laboratory system, the mask errors were ±7.2% and these dominated other spatial system errors. The detector noise was measured as 0.4%. With these errors included in our simulator, the solution vector $\underline{x}$ was calculated, compared to the ideal theoretical $\underline{x}^*$ value and to the $\underline{x}$ vector calculated on the laboratory system. The locations of the closed-loop poles of the system in each case were calculated and compared. The results in Table 4 show excellent agreement (0.5% accuracy or better) in the location of the poles and with the nature of the poles preserved (e.g. complex-conjugate pole pairs).

TABLE 4
Comparison of the Closed-Loop Poles Computed Theoretically and Using
the Optical Laboratory System

| THEORETICAL POLE LOCATIONS | OPTICAL LABORATORY COMPUTED POLES | % ERROR |
|---|---|---|
| -20.45 + j6.26 | -20.74 + j5.88 | 0.5 |
| -20.45 - j6.26 | -20.74 - j5.88 | 0.5 |
| -4.53 | -4.53 | $10^{-3}$ |

| (a) 80 ITERATIONS | (b) 400 ITERATIONS | (c) 640 ITERATIONS |
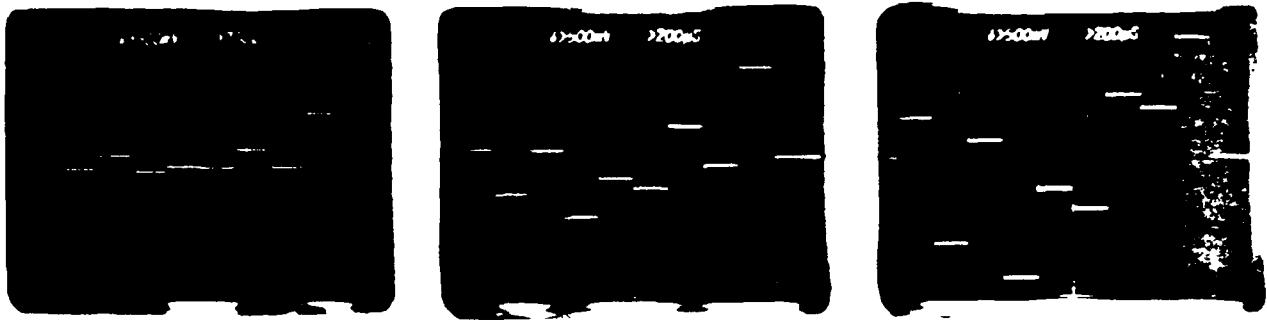
FIGURE 2

The nine photo-detectors outputs from a fixed mask OLAP at selected cycles in the iterative
solution of the system of LAEs $\underline{A}_5\underline{x}_5 = \underline{b}_5$ that arise in the final loop of
the solution of the nonlinear ARE

## 9. SUMMARY AND CONCLUSION

We have detailed a two-loop solution to the nonlinear ARE. In the iterative solution, a
fixed number of iterations can be employed to achieve a given performance accuracy. A
direct solution of each LAE can also be employed, however the iterative solution is faster
($100T_B$ vs. $975T_B$). Selection of the operational parameters for the two-loop algorithm were
theoretically derived, verified by noise-free simulations and shown to be appropriate when
system noise and errors were present. The implementation of direct and iterative solutions
of LAEs on a frequency-multiplexed OLAP was detailed. A theoretical analysis of both algo-
rithms showed that acoustic attenuation was the dominant error source in iterative algorithms
and detector noise dominated direct algorithms. Our simulations verified these theoretical
predictions and quantified the performance obtained with each. Our theoretical values for
the amount of acoustic attenuation allowed to permit convergence of an iterative algorithm
was verified by simulations. We confirmed and quantified by simulations that optical back-
substitution yields comparable performance to its digital realization. Experimental verifi-
cation on a laboratory system was obtained. The guidelines, and theory provided are appro-
priate for various other system processors (optical and digital) and for high-accuracy
digital-optical linear algebra processors. Our nonlinear matrix solution using a fixed num-
ber of iterations is appropriate for realization on any linear algebra processor.

## ACKNOWLEDGMENTS

## REFERENCES

1.  Leith, H.J., Special Issue on Optical Computing, Vol. 72, July 1984.
2.  D. Casasent, "Acousto-Optic Linear Algebra Processors: Architectures, Algorithms and
    Applications", Proc. IEEE, Special Issue on Optical Computing, Vol. 72, pp. 831-849,
    July 1984.
3.  D. Casasent, J. Jackson, C.P. Neuman, Applied Optics, Vol. 22, pp. 115-124, January
    1983.
4.  D. Casasent, J. Jackson, Proc. SPIE, Vol. 431, January 1984.
5.  D. Casasent, A. Ghosh, C.P. Neuman, "A Quadratic Matrix Algorithm for Linear Algebra
    Processors", Submitted, J. of the Opt. Soc. of Am. A, April 1984.
6.  A. Ghosh, D. Casasent, C.P. Neuman, "Performance of Direct and Iterative Algorithms on
    an Optical Systolic Processor", Submitted, Applied Optics, September 1984.
7.  D. Casasent, A. Ghosh, C.P. Neuman, Proc. SPIE, Vol. 431, pp. 201-208, August 1983.
8.  D. Casasent, A. Ghosh, "Optical Linear Algebra Processors: Noise and Error Source
    Modeling", To be submitted, Optics Letters, August 1984.
9.  M. Carlotto, D. Casasent, Applied Optics, Vol. 21, pp. 147-152, January 1982.
10. D. Casasent, A. Ghosh, Optics Communications, Vol. 46, pp. 270-273, July 1983.
11. D. Casasent, A. Ghosh, Applied Optics, Vol. 22, pp. 3572-3578, November 1983.
12. A. Ghosh, D. Casasent, Applied Optics, Vol. 22, pp. 1795-1796, June 1983.
13. D. Casasent, C.P. Neuman, J. Lycas, Applied Optics, Vol. 23, pp. 1960-1966, July 1984.

# 17. OPTICAL LABORATORY SYSTEM AND ELECTRONIC SUPPORT FOR AN OPTICAL LINEAR ALGEBRA MULTIPLEXED PROCESSOR

Ref. 21

# Fabrication and Testing of a Space and Frequency-Multiplexed Optical Linear Algebra Processor

David Casasent

Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

ABSTRACT. A new space/frequency-multiplexed optical linear algebra processor

is described. The electronic support system, fabrication of the processor

and initial performance data are presented.

# Fabrication and Testing of a Space and Frequency-Multiplexed Optical Linear Algebra Processor

David Casasent

Carnegie-Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, Pennsylvania 15213

## 1. INTRODUCTION

Optical linear algebra processors (OLAPs) represent a most flexible and general-purpose class of optical system. In Section 2, we describe the architecture for a space and frequency-multiplexed OLAP. We detail (Section 2) how this system accommodates bipolar and complex-valued data and its use in matrix-vector processing. The electronic support system is described in Section 3. The optical system and initial experimental results obtained on it are detailed in Section 4.

## 2. COMPLEX AND BIPOLAR PROCESSOR ARCHITECTURE

The optical schematic for a new OLAP architecture [1] to accommodate bipolar and/or complex-valued matrix and vector data is shown in Figure 1. For the case shown, the matrix $A$ has bipolar-valued elements and the vector $b$ has complex-valued elements. The bipolar-valued elements of one row of $A$ are spatially-multiplexed on two linear point modulator input arrays at $P_1$ and the complex-valued elements of $b$ are encoded in the conventional three-tuple representation [2] frequency-multiplexed [4] to the acousto-optic (AO) cell at $P_2$. This architecture uses input space-multiplexing (rather than time-multiplexing as in reference [3]) together with frequency-multiplexing [4] to accommodate bipolar and complex-valued matrix and vector data. If both the matrix and vector elements are complex-valued, three linear input arrays are used at $P_1$. If both the matrix and vector data are bipolar, Figure 1 can be used.

The N point modulators per row at $P_1$ are imaged through separate regions of $P_2$, with the different regions of $P_2$ separated in time by $T_B$ (the propagation time of the acoustic wave between the different portions of the AO cell at $P_2$). Each $T_B$, new input data is entered at $P_1$ and a shifted version of the $P_2$ vector is produced (with the vector-shift provided by the motion of the acoustic wave with time). Thus, an N-element vector inner product is produced each $T_B$ and a matrix-vector product is computed each $NT_B$ (for an N x N matrix). This basic OLAP architecture can solve linear and nonlinear matrix equations. The basic linear algebra operation of concern is the solution of a system of linear algebraic equations. Various algorithms to achieve this on such a processor have been detailed elsewhere [7].

The frequency-multiplexing requirements for such a system were detailed in Ref.[1]. For the M=3 frequency case and the system of Figure 1, these requirements are quite modest ($\Delta f=70MHz$). For the case of a banded matrix with bandwidth B=M, the number of input point modulators per row is also quite modest. When B exceeds the number of input point modulators, partitioning is easily achieved as detailed elsewhere [5]. With a multi-channel AO cell at $P_2$, and the appropriate data encoding and time-integration of the output, the same

architecture can achieve high accuracy as detailed in Ref [5]. Thus, this is a most attractive, powerful and flexible OLAP architecture.
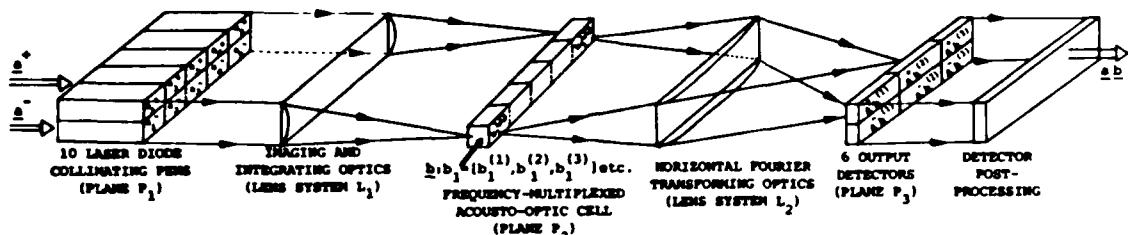


FIGURE 1

New analog matrix-vector space and frequency-multiplexed architecture for complex and bipolar-valued matrix and vector elements.

### 3. ELECTRONIC SUPPORT SYSTEM

Any optical or digital linear algebra processor or systolic system must provide parallel input data to $P_1$ (N words) and to $P_2$ (M words) each $T_B$, plus provide acquisition and analysis of the parallel output $P_3$ data each $T_B$. To achieve this with flexibility and programability and to allow input data for any application to be processed from a digital database, a dedicated high-speed microprocessor system was assembled. This electronic support system (Figure 2) contains many special-purpose boards, a hard disk (10M bytes), magnetic tape, on-line multibus memory (512K bytes) and processor memory (512K bytes) with 600nsec memory access, and video (Matrox interface) and graphics processor outputs. The microprocessor used at present is an M68000 with an Intel 86/380 RMX version also under present evaluation. The general philosophy of this support system is to download digital data from a VAX, magnetic tape, etc. into high-speed parallel output buffer memory which drives parallel D/As to the $P_1$ and $P_2$ inputs. Output data is similarly A/D converted in parallel and buffered in an output memory. The disk system provides storage of the input and output data. The microprocessor provides control, formatting, etc.

To provide the parallel $P_1$ and $P_2$ analog inputs, three special-purpose cards with five parallel output D/As (12 bits at 10MHz) and drivers per card were fabricated. With 10 inputs to $P_1$ and three inputs to $P_2$, the system processes 130 M byte (12 bit bytes) input data (1.5 G-bit per second data) with $T_B=0.1\mu sec$. This represents a reasonable compromise between available D/A converters and other hardware and system performance. Each D/A input is obtained from a separate high-speed parallel buffer memory channel, each 4K words deep (12 bit words). Three special-purpose buffer memory boards with 8 memory channels per board have been fabricated and are used for input and output buffering. The $P_3$ outputs are detected (with special-purpose 20MHz, low-noise circuits), A/D converted (using special-purpose circuitry with one A/D per board with 12 bit accuracy and 10MHz bandwidth), and fed to a parallel input buffer memory. The system's inputs settle to 0.2% in 100nsec, thus allowing 10MHz data rate (analog, 12 bits) per channel. The necessary spatial corrections [6] for the $P_1$ and $P_2$ transducers are obtained off-line and applied to the input data. These corrections, plus input and output bipolar and complex data normalization and encoding, are performed in software (with their on-line hardware realization straightforward). An interface board to control the system, and an RF modulator drive board for the AO cell complete the electronic support system.

FIGURE 2
Photograph of the electronic
support system
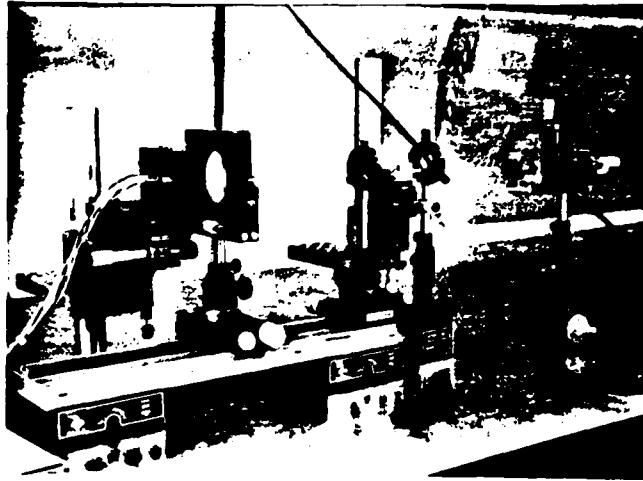
FIGURE 3
Photograph of the laboratory
optical matrix-vector system

## 4. OPTICAL SYSTEM FABRICATION AND INITIAL RESULTS

The optical system of Figure 1 was assembled (Figure 3) using a laser diode (LD) input array at $P_1$ with individual collimating optics integrated with each LD source. The $P_1$ outputs had a 50% fill-factor and the full $P_1$ input was reduced by a two lens system to be compatible with the size of $P_2$ and the 0.1μsec data packet spacings. A special input $P_1$ mount was fabricated to allow each $P_1$ source to be separately aligned within 0.3mrad to illuminate the correct region of the AO cell at $P_2$ with the necessary beam divergence. For the initial laboratory system, the beam reducing optics from $P_1$ to $P_2$ occupied 600+20mm and the optics from $P_2$ to $P_3$ required 160mm. An even more compact system with folded optics is easily possible.

In Figures 4 and 5, several examples of the performance of the system of Figure 3 are provided. The laboratory system is fully automated with data loading and output display under control of a dedicated terminal through the M68000 system. The inputs to 3 of the $P_1$ laser diodes, the AO cell and the output vector inner product from one detector are shown as functions of time. The results obtained are as expected and verify the digital control and performance of the full hybrid optical/digital system.
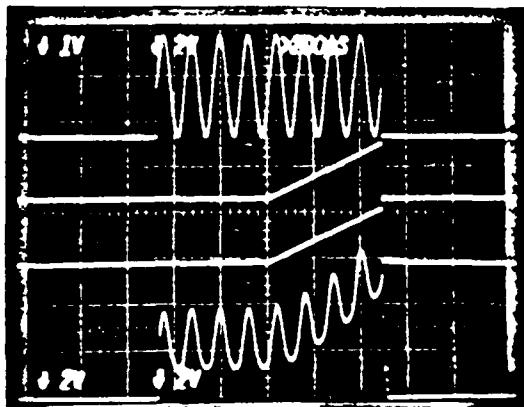
FIGURE 4
Three LD inputs (top 3 traces,
sinewave and 2 biased ramps) and
output (bottom trace) with a
constant RF AO input
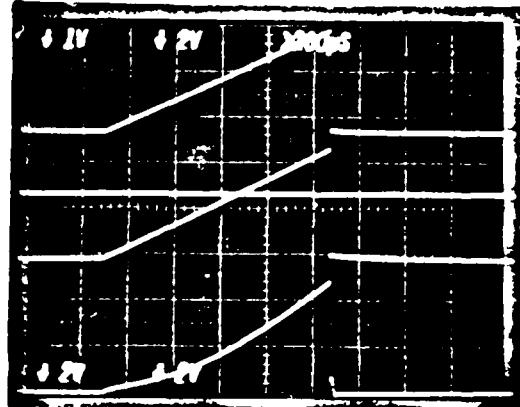
FIGURE 5
Three LD inputs (top 3 traces, 2
linear ramps and a 0 input) and output
(bottom trace, quadratic as expected)
with the RF AO input varying linearly in
power with time over the duration of the
input ramps

REFERENCES

1. D. Casasent, J. Jackson, "Fabrication Considerations for Acousto-Optic Systolic Processors", Proc. SPIE, Vol. 465, pp. 104-112, January 1984.
2. J.W. Goodman, L.M. Woody, Applied Optics, Vol. 16, p. 2611, 1977.
3. M. Carlotto, D. Casasent, "A Microprocessor-Based Fiber-Optic Iterative Optical Processor", Applied Optics, Vol. 21, pp. 147-152, January 1982.
4. D. Casasent, J. Jackson, C.P. Neuman, "Frequency-Multiplexed and Pipelined Iterative Optical Systolic Array Processors", Applied Optics, Vol. 22, pp. 115-124, January 1983.
5. D. Casasent, B. Taylor, "High Performance Banded Matrix Algorithms and Architectures", Submitted, Applied Optics, October 1984.
6. D. Casasent, A. Ghosh, "Optical Linear Algebra Processors: Noise and Error Source Modeling", Optics Letters, Submitted, August 1984.
7. D. Casasent, "Acousto-Optic Linear Algebra Processors: Architectures, Algorithms and Applications", Proc. IEEE, Special Issue on Optical Computing, Vol. 72, pp. 831-849, July 1984.

# 18. PUBLICATIONS, PRESENTATIONS AND THESES PRODUCED

Publications from the start of this new grant are listed in Section 18.1.1. A list of presentations at conferences, companies and seminars on our *AFOSR* research during the past year follows in Section 18.2. The theses supported by this AFOSR grant are noted in Section 18.3.

## 18.1 PUBLICATIONS (AFOSR SUPPORTED, SEPTEMBER 1984-DATE

### 18.1.1 PAPERS PUBLISHED UNDER AFOSR SUPPORT (SEPTEMBER 1984-

#### SEPTEMBER 1985

1. D. Casasent and R.L. Cheatham, "Hierarchical Pattern Recognition Using Parallel Feature Extraction", *Proc. ASME, Computers in Engineering 1984*, Vol. 1, pp. 1-6, August 1984.

2. D. Casasent, A. Ghosh and C.P. Neuman, "Iterative Solutions to Nonlinear Matrix Equations Using a Fixed Number of Steps", *Proc. SPIE*, Vol. 495, pp. 102-108, August 1984.

3. R.L. Cheatham and D. Casasent, "Hierarchical Fisher and Moment-Based Pattern Recognition", *Proc. SPIE*, Vol. 504, pp. 19-26, August 1984.

4. W.T. Chang, D. Casasent and D. Fetterly, "SDF Control of Correlation Plane Structure for 3-D Object Representation and Recognition", *Proc. SPIE*, Vol. 507, pp. 9-18, August 1984.

5. D. Casasent, A. Goutzoulis and B.V.K. Vijaya Kumar, "Time-Integrating Acousto-Optic Correlator: Error Source Modeling", *Applied Optics*, Vol. 23, pp. 3230-3237, September 1984.

6. D. Casasent and R.L. Cheatham, "Image Segmentation and Real-Image Tests for an Optical Moment-Based Feature Extractor", *Optics Communications*, Vol. 51, pp. 227-230, September 1984.

7. D. Casasent and V. Sharma, "Feature Extractors for Distortion-Invariant Robot Vision", *Optical Engineering*, Vol. 23, pp. 492-498, September/October 1984.

8. V. Sharma and D. Casasent, "Optimal Linear Discriminant Functions", *Proc. SPIE*, Vol. 519, pp. 50-55, October 1984.

9. D. Casasent, W. Rozzi and D. Fetterly, "Projection Synthetic Discriminant Function Performance", Submitted, *Optical Engineering*, Vol. 23, pp. 716-720, November 1984.

10. W.T. Chang and D. Casasent, "Chord Distributions in Pattern Recognition: Distortion-Invariance and Parameter Estimation", *Proc. SPIE*, Vol. 521, pp. 2-6, November 1984.

11. A. Goutzoulis, D. Casasent and B.V.K. Vijaya Kumar, "Acousto-Optic Processor for Adaptive Radar Noise Environment Characterization", *Applied Optics*, Vol. 23, pp. 4303-4308, December 1984.

12. D. Casasent, "Optical Processing Research Making Significant Advancements", *Laser Focus*, pp. 150, October 1984.

13. D. Casasent, "Coherent Optical Pattern Recognition: A Review", *Optical Engineering*, Vol. 24, Special Issue, pp. 26-32, January 1985.

14. D. Casasent and J.Z. Song, "A Computer Generated Hologram for Diffraction-Pattern Sampling", *Proc. SPIE*, Vol. 523, January 1985.

15. D. Casasent, "Hybrid Optical/Digital Image Pattern Recognition: A Review", *Proc. SPIE*, Vol. 528, pp. 64-82, January 1985.

16. D. Casasent, "Computer Generated Holograms in Pattern Recognition: A Review", *Proc. SPIE*, Vol. 532, pp. 106-118, January 1985.

17. D. Casasent, "Parallel Coherent Optical Processor Architectures and Algorithms for ATR", *Proc. of the Workshop on Algorithm-Guided Parallel Architectures for Automatic Target Recognition*, Leesburg, Virginia, July 1984, Published February 1985, pp. 33-49.

18. D. Casasent, "Frequency-Multiplexed Acousto-Optic Architectures and Applications", *Applied Optics*, Vol. 24, March 1985.

19. D. Casasent, "Fabrication and Testing of a Space and Frequency-Multiplexed Optical Linear Algebra Processor", *OSA Topical Meeting on Optical Computing*, pp. TuD7-1 - TuD7-4, March 1985.

20. D. Casasent and R.L. Cheatham, "Hierarchical Feature-Based Object Identification", *OSA Topical Meeting on Machine Vision*, pp. ThD4-1 - ThD4-4, March 1985.

21. D. Casasent and A. Mahalanobis, "Correlation Filters for Distortion-Invariance and Discrimination", *OSA Topical Meeting on Machine Vision*, pp. FB5-1 - FB5-3, March 1985.

22. D. Casasent and B.T. Taylor, "Banded-Matrix High-Performance Algorithm and Architecture", *Applied Optics*, Vol. 24, pp. 1476-1480, 15 May 1985.

23. D. Casasent and A. Ghosh, "Optical Linear Algebra Processors: Noise and Error-Source Modeling", *Optics Letters*, Vol. 10, pp. 252-254, June 1985.

24. D. Casasent, "Computer Generated Holograms in Pattern Recognition: A Review", *Optical Engineering*, Vol. 24, pp. 724-730, September/October 1985.

## 18.1.2 PAPERS SUBMITTED UNDER AFOSR SUPPORT

25. D. Casasent, A. Ghosh and C.P. Neuman, "A Quadratic Matrix Algorithm for Linear Algebra Processors", *J. Large-Scale Systems*, Accepted.

# 18.2 SEMINARS AND CONFERENCE PRESENTATIONS OF AFOSR RESEARCH (SEPTEMBER 1984 - SEPTEMBER 1985)

### September 1984

1. Philips Research Laboratories - Briarcliff, NY - "Optics and Pattern Recognition in Robotics".

2. Optical Society of America - Pittsburgh, PA, "CMU Center for Excellence in Optical Data Processing".

3. Carnegie-Mellon University, ECE Graduate Seminar - Pittsburgh, PA, "Optical Processing Research in the Center for Excellence in Optical Data Processing".

4. Westinghouse Corporation - Baltimore, MD, "Research and Facilities in the Center for Excellence in Optical Data Processing".

### October 1984

5. Washington, D.C., "Optical Pattern Recognition: Feature Extraction".

6. Washington, D.C., "Optical Pattern Recognition: Correlators".

7. Washington, D.C., "Synthetic Discriminant Function Case Studies".

8. Washington, D.C., "Basic Optical Signal Processing Architectures and Algorithms".

9. Washington, D.C., "Advanced Optical Signal Processing Architectures and Algorithms".

10. Washington, D.C., "Optical Linear Algebra Processor Algorithms and Architectures".

11. Washington, D.C., "Optical Linear Algebra Processor Applications and High-Accuracy Architectures".

12. Carnegie-Mellon University, ECE Sophomore Seminar - Pittsburgh, Pennsylvania, "Research in the Center for Excellence in Optical Data Processing".

13. University of Pittsburgh, Center for Multivariate Analysis - Pittsburgh, PA, "Advanced Multi-Class Distortion-Invariant Pattern Recognition".

14. Wright Patterson Air Force Base - Ohio, "Multi-Functional Optical Signal Processor for Electronic Warfare".

15. George Mason University - Washington, D.C., "Optical Information Processing".

16. SPIE (IOCC) Conference - Boston, Massachusetts, "Optimal Linear Discriminant Functions".

### November 1934

17. SPIE Robotics Conference - Boston, MA, "Chord Distributions in Pattern Recognition".

18. University of Maryland - "Optical Processing for Autonomous Land Vehicle Navigation".

### January 1985

19. Fairchild Weston - Long Island, NY, "Optical Pattern Recognition and Optical Processing".

20. SPIE Conference - Los Angeles, CA, "Hybrid Optical/Digital Image Pattern Recognition: A Review".

21. SPIE Conference - Los Angeles, CA, "A Computer Generated Hologram for Diffraction-Pattern Sampling".

22. SPIE Conference - Los Angeles, CA, "A Recent Review of Holography in Coherent Optical Pattern Recognition".

23. Sandia National Laboratories - Albuquerque, NM, "Optical Pattern Recognition and Optical Processing".

**February 1985**

24. NASA Lewis - Cleveland, OH, "Optical Linear Algebra Processors (Systolic)".

**March 1985**

25. George Washington University, - Washington, D.C., "Optical Linear Algebra for SDI".

26. Lockheed Missiles & Space Co. - Sunnyvale, CA, "Advanced Hybrid Optical/Digital Pattern Recognition"

27. OSA Topical Meeting on Optical Computing - Lake Tahoe, NV, "Fabrication and Testing of a Space and Frequency-Multiplexed Optical Linear Algebra Processor".

28. OSA Topical Meeting on Machine Vision - Lake Tahoe, NV, "Hierarchical Feature-Based Object Identification".

29. OSA Topical Meeting on Machine Vision - Lake Tahoe, NV, "Correlation Filters for Distortion-Invariance and Discrimination".

30. Texas Instruments - Dallas, TX, "Optical Pattern Recognition".

**April 1985**

31. Electro-Com Automation, Inc. - Dallas, TX, "Optical Pattern Recognition".

32. Eglin Air Force Base - Ft. Walton Beach, FL, "Optical Pattern Recognition and Kalman Filtering".

**May 1985**

33. Carnegie-Mellon University - Board of Trustees, "Optical Data Processing".

**August 1985**

34. SPIE - San Diego, CA, "Correlation Synthetic Discriminant Functions for Object Recognition and Classification in High Clutter".

35. SPIE - San Diego, CA, "A Factorized Extended Kalman Filter".

36. SPIE - San Diego, CA, "Optical Finite-Element Processor".

**September 1985**

37. SPIE - Cambridge, MA, "Parameter Estimation and In-Plane Distortion Invariant Chord Processing".

38. SPIE - Cambridge, MA, "Optical Processing Techniques for Advanced Intelligent Robots and Computer Vision".

39. SPIE - Cambridge, MA, "High-Dimensionality Feature-Space Processing with Computer Generated Holograms".

## 18.2.1 THESES SUPPORTED BY AFOSR FUNDING (SEPTEMBER 1984 - SEPTEMBER 1985

1. Eugene Pochapsky, M.S. Dissertation, "The Simulation of Optical Pattern Recognition Systems", September 1984.

2. William Rozzi, M.S. Dissertation, "Advanced Quantitative Synthetic Discriminant Function Tests on Ship Imagery", December 1984.

3. James Fisher, M.S. Dissertation, "Extended Kalman Filter Algorithms for Implementation on a High-Accuracy Optical Processor", December 1984.

4. W.T. Chang, Ph.D. Dissertation, "Chord Distributions and Correlation SDFs in Pattern Recognition", March 1985.

# 19. REFERENCES

1. D. Casasent, A. Goutzoulis and B.V.K. Vijaya Kumar, "Time-Integrating Acousto-Optic Correlator: Error Source Modeling", *Applied Optics*, Vol. 23, pp. 3230-3237, September 1984.

2. A. Goutzoulis, D. Casasent and B.V.K. Vijaya Kumar, "Detector Effects on Time-Integrating Correlator Performance", *Applied Optics*, Vol. 24, pp. 1224-1233, 15 April 1985.

3. D. Casasent, "Hybrid Optical/Digital Image Pattern Recognition: A Review", *Proc. SPIE*, Vol. 528, pp. 64-82, January 1985.

4. D. Casasent, "Parallel Coherent Optical Processor Architectures and Algorithms for ATR", *Proc. of the Workshop on Algorithm-Guided Parallel Architectures for Automatic Target Recognition*, Leesburg, Virginia, July 1984, Published February 1985, pp. 33-49.

5. D. Casasent, "Coherent Optical Pattern Recognition: A Review", *Optical Engineering*, Vol. 24, Special Issue, pp. 26-32, January 1985.

6. D. Casasent, "Computer Generated Holograms in Pattern Recognition: A Review", *Proc. SPIE*, Vol. 532, pp. 106-118, January 1985.

7. D. Casasent, "Computer Generated Holograms in Pattern Recognition: A Review", *Optical Engineering, 24*, pp. 724-730, September/October 1985.

8. D. Casasent and J.Z. Song, "A Computer Generated Hologram for Diffraction-Pattern Sampling", *Proc. SPIE*, Vol. 523, January 1985.

9. D. Casasent and V. Sharma, "Feature Extractors for Distortion-Invariant Robot Vision", *Optical Engineering*, Vol. 23, pp. 492-498, September/October 1984.

10. W.T. Chang and D. Casasent, "Chord Distributions in Pattern Recognition: Distortion-Invariance and Parameter Estimation", *Proc. SPIE*, Vol. 521, pp. 2-6, November 1984.

11. R.L. Cheatham and D. Casasent, "Hierarchical Fisher and Moment-Based Pattern Recognition", *Proc. SPIE*, Vol. 504, pp. 19-26, August 1984.

12. D. Casasent and R.L. Cheatham, "Hierarchical Pattern Recognition Using Parallel Feature Extraction", *Proc. ASME, Computers in Engineering 1984*, Vol. 1, pp. 1-6, August 1984.

13. D. Casasent and R.L. Cheatham, "Image Segmentation and Real-Image Tests for an Optical Moment-Based Feature Extractor", *Optics Communications*, Vol. 51, pp. 227-230, September 1984.

14. D. Casasent and R.L. Cheatham, "Hierarchical Feature-Based Object Identification", *OSA Topical Meeting on Machine Vision*, pp. ThD4-1 - ThD4-4, March 1985.

15. D. Casasent, W. Rozzi and D. Fetterly, "Projection Synthetic Discriminant Function Performance", Submitted, *Optical Engineering*, Vol. 23, pp. 716-720, November 1984.

16. V. Sharma and D. Casasent, "Optimal Linear Discriminant Functions", *Proc. SPIE*, Vol. 519, pp. 50-55, October 1984.

17. W.T. Chang, D. Casasent and D. Fetterly, "SDF Control of Correlation Plane Structure for 3-D Object Representation and Recognition", *Proc. SPIE*, <u>Vol. 507</u>, pp. 9-18, August 1984.

18. D. Casasent and A. Mahalanobis, "Correlation Filters for Distortion-Invariance and Discrimination", *OSA Topical Meeting on Machine Vision*, pp. FB5-1 - FB5-3, March 1985.

19. D. Casasent and A. Ghosh, "Optical Linear Algebra Processors: Noise and Error-Source Modeling", *Optics Letters*, <u>10</u>, pp. 252-254, June 1985.

20. D. Casasent, A. Ghosh and C.P. Neuman, "Iterative Solutions to Nonlinear Matrix Equations Using a Fixed Number of Steps", *Proc. SPIE*, <u>Vol. 495</u>, pp. 102-108, August 1984.

21. D. Casasent, "Fabrication and Testing of a Space and Frequency-Multiplexed Optical Linear Algebra Processor", *OSA Topical Meeting on Optical Computing*, pp. TuD7-1 - TuD7-4, March 1985.

END

12-86

DTIC