

AD-A086 208

UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES IMAGE --ETC F/G 17/7
INVARIANT FEATURES FOR QUANTITATIVE SCENE ANALYSIS.(U)
JUN 76 E L HALL, W FREI

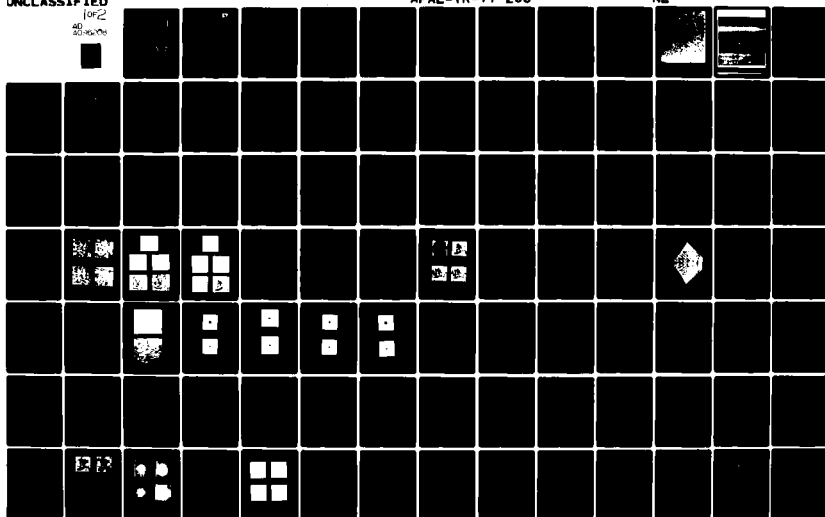
F08606-72-C-0008

UNCLASSIFIED

AFAL-TR-77-253

NL

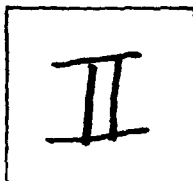
for 2
40
20-sec



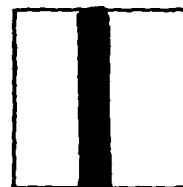
PHOTOGRAPH THIS SHEET

ADA 086208

DTIC ACCESSION NUMBER



LEVEL



INVENTORY

Univ. of Southern Calif.,
Los Angeles, CA. Image Processing Inst.

AFAL-TR-77-253, Final, 30 Jun. 76

DOCUMENT IDENTIFICATION

Contract No. F08606-72-C-0008

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

DISTRIBUTION STATEMENT

ACCESSION FOR	
NTIS	GRA&I
DTIC	TAB
UNANNOUNCED	
JUSTIFICATION	
BY	
DISTRIBUTION /	
AVAILABILITY CODES	
DIST	AVAIL AND/OR SPECIAL
A	

DISTRIBUTION STAMP

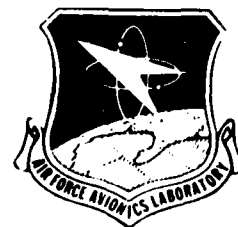
DTIC	
ELECTE	
S	D
JUL 8 1980	
D	

DATE ACCESSIONED

80 7 3 091

DATE RECEIVED IN DTIC

PHOTOGRAPH THIS SHEET AND RETURN TO DTIC-DDA-2



AFAL-TR-77-253

INVARIANT FEATURES FOR QUANTITATIVE SCENE ANALYSIS

Image Processing Institute
Department of Electrical Engineering
University of Southern California
Los Angeles, California 90007

30 June 1976

TECHNICAL REPORT AFAL-TR-77-253

Final Report for Period 1 September 1975-30 June 1976

Approved for public release; distribution unlimited.

AIR FORCE AVIONICS LABORATORY
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

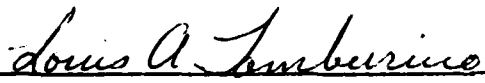
ADA 086208

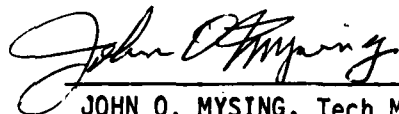
NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

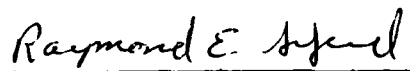
This report has been reviewed by the Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.


LOUIS A. TAMBURINO
Project Engineer


JOHN O. MYSING, Tech Mgr
Information Presentation &
Control Group
System Technology Branch

FOR THE COMMANDER


RAYMOND E. SIFERD, Colonel, USAF
Chief, System Avionics Division
Avionics Laboratory

"If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization please notify AFWAL/AAAT, W-PAFB, OH 45433 to help us maintain a current mailing list".

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFAL-TR-77-253	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) INVARIANT FEATURES FOR QUANTITATIVE SCENE ANALYSIS		5. TYPE OF REPORT & PERIOD COVERED Final Report 1 Sep 75 to 30 Jun 76
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ernest L. Hall and Werner Frei (Principal Investigators)		8. CONTRACT OR GRANT NUMBER(s) F08606-72-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Image Processing Institute University of Southern California, University Park Los Angeles, California 90007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 1706-00-01
11. CONTROLLING OFFICE NAME AND ADDRESS Hq Space & Missile Systems Orgn. (USAF/AFSC) P.O. Box 92960 Worldway Postal Center Los Angeles CA 90009		12. REPORT DATE 30 June 1976
		13. NUMBER OF PAGES 170
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) AFAL/AAT Wright-Patterson AFB OH 45433		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Image Processing, Pictorial Recognition, Map Matching, Scene Analysis, Image Registration, Invariant Measurements		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This report summarizes a study of map matching technique for precision guided reentry vehicles using invariant measurements and scene analysis techniques. A particular case of matching optical and side looking radar images is considered in detail. Both geometric and sensor transformations are developed. Edge registration including a new vector representation of scenes which may be used to derive edge operators and unifies most edge operators currently in use, is presented. Invariant measurements depend on algebraic invariant theory and perceptual experience. The set of moment invariants are considered which are		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Block 20. ABSTRACT--Cont'd.

invariant to translation, rotation, size and mirror images, for computer use. Finally, hierarchical search techniques which are logarithmically efficient are presented for map matching. These techniques are illustrated which correlation measures of picture values, edge measurements and moment invariants. These techniques appear very effective for map matching.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

TABLE OF CONTENTS

<u>SECTION</u>	<u>Page</u>
I. INTRODUCTION.	1
1.1 Problem Statement and Background of Previous Work.	6
1.2 Outline of Report.	18
II. TRANSFORMATIONS FOR MAP MATCHING.	20
2.1 Geometric Transformations.	20
2.2 Sensor Transformations	29
2.3 Fourier Transforms	44
III. EDGE EXTRACTION AND REGISTRATION.	61
3.1 Introduction	61
3.2 Fast Edge Detection.	64
3.3 Edge Registration.	71
IV. INVARIANT MEASUREMENTS.	94
4.1 Algebraic Invariants	94
4.2 Perceptual Invariants.	103
V. HIERARCHICAL SEARCH TECHNIQUES.	127
5.1 Introduction	127
5.2 Experimental Results	136
VI. SCENE CONTENT MEASUREMENTS AND RECOMMENDATIONS FOR FUTURE WORK. .	163

SECTION I.

INTRODUCTION

The need for analyzing images of the earth taken from high and low altitude aircraft is greatly increasing. The purposes of the photographs are diverse and include:

- . Navigational aid
- . Geographic map making
- . Natural resources analysis
- . Weather prediction

Typically, a sequence of images is taken at different times with different sensors. A basic problem in the analysis of these images arises when it is desired to locate the same region in a pair of dissimilar images of the same field of view. For example, one may wish to locate a feature in both optical and radar images of a scene in order to verify the accuracy of the radar sensor or locate a feature in two optical images taken under different weather conditions in order to study the changes. This general problem is called map matching. Special cases of this problem are image registration, change detection, stereometric analysis, and multispectral image analysis.

Several methods including photographic techniques such as image subtraction, optical techniques such as optical correlation and digital computer techniques may be used to approach the map matching problem. In this report only computer imaging techniques will be considered. The versatility, reliability, and economical feasibility of these techniques have produced the trend of these methods for all application areas.

Several noteworthy approaches to the map matching problem have been proposed. These include the standard correlation detector as well as highly efficient sequential techniques. These methods and techniques together with an analysis of their performance are described in the next section. The results of this evaluation indicate that although some of the methods are partially successful, further research is needed to accomplish the difficult task of map matching in an efficient manner.

The type imagery encountered in map matching is varied and may include optical, radar, radiometric or multispectral images. A typical optical image is shown in Figure 1, and the corresponding side looking radar image is shown in Figure 2. Note that although each of these images shows significant scene detail, the two sensors respond differently to the same feature. Most noticeable is the intensity reversal of roads which appear white in the optical and dark in the radar image. Upon careful study one may also note a nonlinear geometric distortion between the two images as well as scale changes. The appearance of bright "glint" features in the radar image may also be noted. These features are not visible in the optical images. Forest regions present texture patterns in both optical and radar images with a shift or reversal of intensity. Urban structures such as buildings do not appear in great detail in the radar images but shadowing effects are apparent in the radar images. The predominant similar features in both optical and radar images are shapes of certain features such as roads, forest boundaries and land-water boundaries. Furthermore, these shapes remain recognizable

even after some degree of degradations in resolution, intensity, geometry or weather. Thus, it may be expected that structural features will be important in map matching.



Figure 1. Typical optical image used in this study.

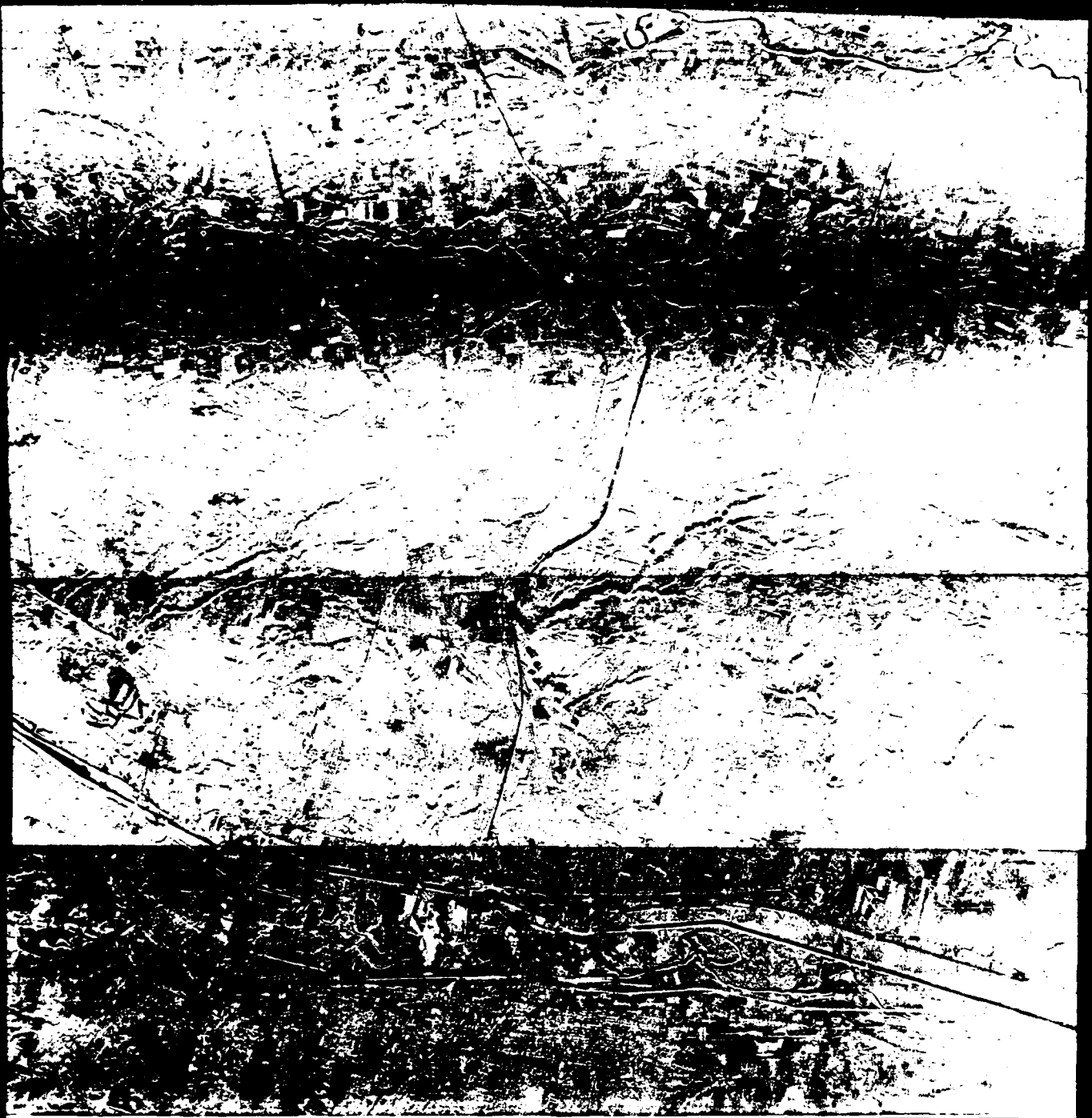


Figure 2. Typical radar image considered in this study.

1.1 PROBLEM STATEMENT AND BACKGROUND OF PREVIOUS WORK

Let two images, S the search, and W the window, be as shown in Figure 3. S is taken as an MxN array of digital picture elements which assume one of K gray levels:

$$0 \leq S(m,n) \leq K - 1$$

$$1 \leq m \leq M$$

$$1 \leq n \leq N.$$

W is taken as a JxK array of digital picture elements (with $J < M$ and $K < N$) having one of Q gray levels:

$$0 \leq W(j,k) \leq Q - 1$$

$$1 \leq j \leq J$$

$$1 \leq k \leq K.$$

When W is superimposed on S, the position of the upper left hand corner of W is used as a reference point in identifying the position of W in the search area coordinates. Therefore, when W is located at (u,v), a picture element S(i,j) in S is superimposed by a picture element W(i-u,j-v) in W.

In the physical world, S may represent a high-resolution optical picture taken by an air-borne camera at high altitude. This is as depicted in Figure 4. W may represent an image taken by:

- . The same sensor at a different time and look angle
- . A difference sensor such as a side-looking radar at a lower altitude

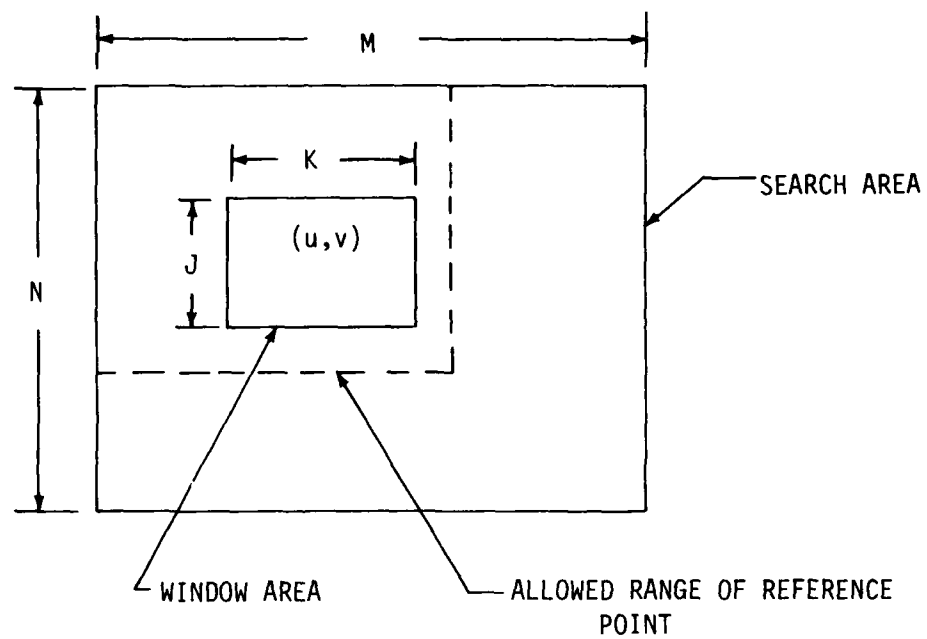


FIGURE 3. SEARCH SPACE

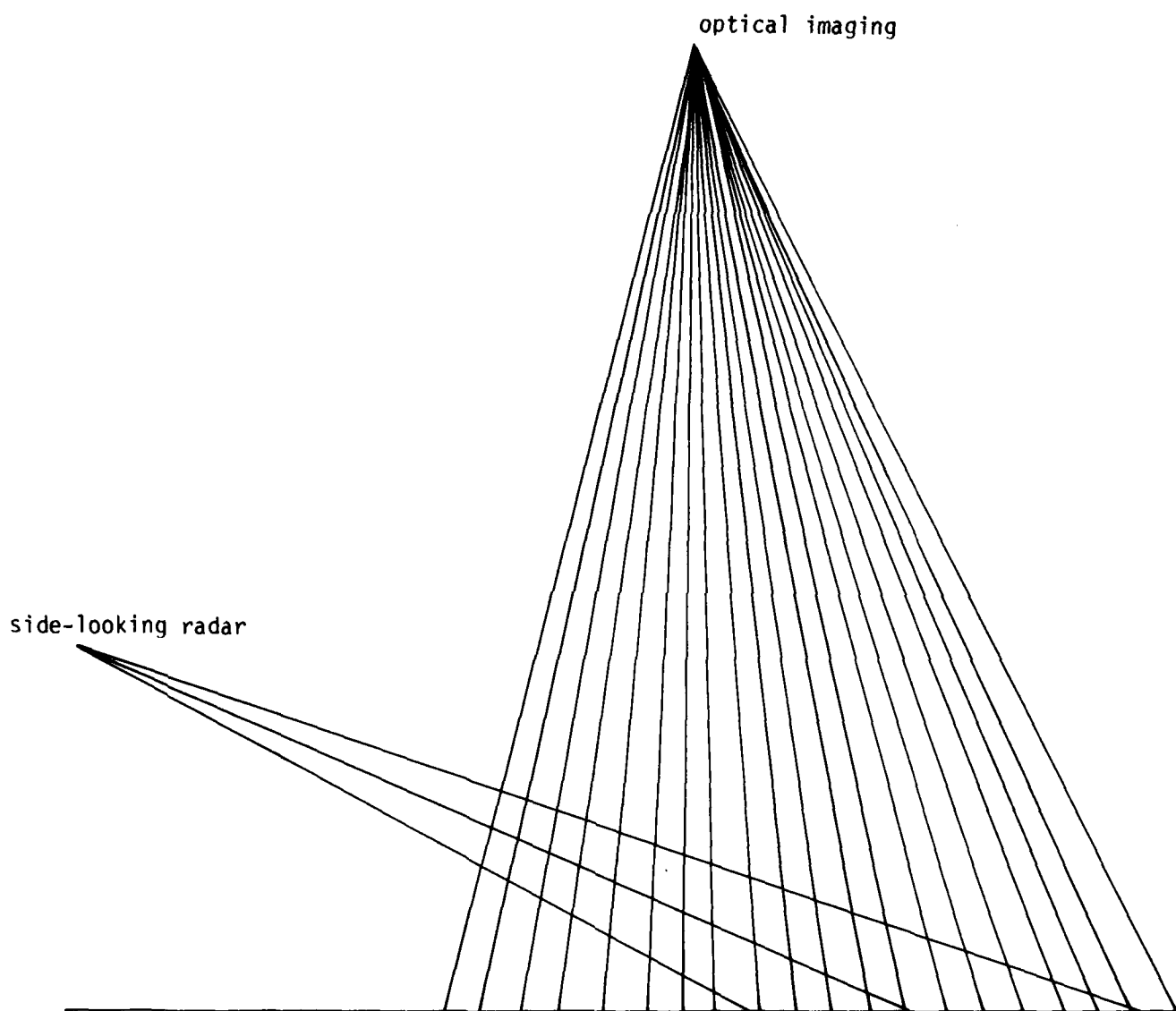


FIGURE 4. GENERAL GEOMETRY OF THE OPTICAL AND RADAR IMAGING SYSTEMS

The basic problem is that given an object terrain as represented by W , we wish to determine by the use of a digital computer whether the same object terrain appears in S .

A number of researches have been made to develop methods of detecting local similarity and perform image matching. Some of the more promising ones are:

- . Basic correlator
- . Statistical correlator
- . Sequential correlator
- . Sequential template matching

In the following sections, each of these methods is briefly described and their relative performances are analyzed.

1.1.1 Basic Correlator

The basic correlation is a method used to form a correlation measure between two picture functions and to determine the location of the maximum correlation [1], [2]. In applying this technique, the correlation measure $R(u,v)$ at the reference location (u,v) is defined as:

$$R(u,v) = \frac{\sum_{j=1}^K \sum_{i=1}^J S(i,j)W(i-u,j-v)}{\left\{ \left[\sum_{j=1}^K \sum_{i=1}^J S^2(i,j) \right] \left[\sum_{j=1}^K \sum_{i=1}^J W^2(i-u,j-v) \right] \right\}^{\frac{1}{2}}} \dots \quad (1)$$

For a given window W , the term $\sum_{j=1}^K \sum_{i=1}^J W^2(i-u,j-v)$ in Equation (1) is

a constant for any reference point u, v . Therefore Equation (1) can be reduced to:

$$R(u, v) = \frac{\sum_{j=1}^K \sum_{i=1}^J S(i, j) W(i-u, j-v)}{\left[\sum_{j=1}^K \sum_{i=1}^J S^2(i, j) \right]^{\frac{1}{2}}} \dots \quad (2)$$

To determine the location of maximum correlation, $R(u, v)$ must be computed at each location u, v ; $1 \leq U \leq (M-J+1)$, $1 \leq V \leq (N-K-1)$. This is because no decision can be made until the correlation array $R(u, v)$ is computed for all u, v . The performances of this correlator can be described as follows:

(1) This method is relatively sensitive to image noise [3]. In the presence of image noise, the correlation function produces a relatively broad peak, thus making a selection of a correlation peak difficult.

(2) A great amount of computation must be performed since the window and search areas are usually large in an actual photograph. With this technique no decision can be made until the correlation array $R(u, v)$ is computed for all u, v .

(3) With the exception to the matching of every simple pictures this method and the variations of this method [4], [5] does not provide satisfactory performance in image matching.

1.1.2 Statistical Correlation

To overcome some of the difficulties mentioned in the previous method, the statistical knowledge of the spatial relation of picture elements within each image were used in this statistical

correlation [6]. The statistical correlation measure $R_s(u,v)$ is:

$$R_s(u,v) = \frac{\sum_{j=1}^K \sum_{i=1}^J F_s(i,j) F_w(i-u, j-v)}{\left[\sum_{j=1}^K \sum_{i=1}^J F_s^2(i,j) \right]^{\frac{1}{2}} \left[\sum_{j=1}^K \sum_{i=1}^J F_w^2(i,j) \right]^{\frac{1}{2}}} \dots \quad (3)$$

where $F_s(i,j)$ and $F_w(i,j)$ are obtained by spatially convolving the images $S(i,j)$ and $W(i,j)$ with spatial filter functions $D_s(i,j)$ and $D_w(i,j)$:

$$F_s(i,j) = S(i,j) \otimes D_s(i,j) \dots \quad (4)$$

$$F_w(i,j) = W(i,j) \otimes D_w(i,j) \dots \quad (5)$$

The spatial filter functions $D_s(i,j)$ and $D_w(i,j)$ are chosen to maximize the correlation peak. The first step in the spatial filter design is to decorrelate or whiten the images as follows:

$$A = [H_s]^{-1} S \quad (6)$$

$$B = [H_w]^{-1} W \quad (7)$$

S and W are column vector representations of the images $S(i,j)$ and $W(i,j)$ obtained by column scanning the images. H_s and H_w are obtained by a factorization of the image covariance matrices

$$K_s = H_s H_s^T \quad (8)$$

$$K_w = H_w H_w^T \quad (9)$$

H_s and H_w may be formulated in terms of the eigenvector and eigenvalues of K_s and K_w as follows:

$$\begin{aligned} K_s &= E_s \Lambda_s E_s^T = (E_s \Lambda_s)^{\frac{1}{2}} (E_s \Lambda_s)^{\frac{1}{2}} \\ &= H_s H_s^T \end{aligned} \quad (10)$$

$$\begin{aligned}
 K_W &= E_W \Lambda_W E_W^T = (E_W \Lambda_W)^{\frac{1}{2}} (E_W \Lambda_W)^{\frac{1}{2}} \\
 &= H_W H_W^T.
 \end{aligned} \tag{11}$$

The correlation operation is performed on the whitened vectors A and B to yield the statistical correlation measure.

$$R_S(u, v) = \frac{A^T B}{(A^T A)^{\frac{1}{2}} (B^T B)^{\frac{1}{2}}} \tag{12}$$

or

$$R_S(u, v) = \frac{(K^T)^{-1} S^T W}{\left\{ [(K^T)^{-1} S]^T [(K^T)^{-1} S] W^T W \right\}^{\frac{1}{2}}} \tag{13}$$

where

$$(K^T)^{-1} = (H_S H_W^T)^{-1}. \tag{14}$$

If the image elements are assumed to be samples of Markov process then:

$$(K^T)^{-1} = K^{-1} = 1/(1-\rho^2) \begin{bmatrix} -1 & -\rho \Sigma^{-1} & 0 & 0 & \dots 0 \\ -\rho \Sigma^{-1} & (1+\rho^2) \Sigma^{-1} & -\rho \Sigma^{-1} & 0 & \dots 0 \\ 0 & -\rho \Sigma^{-1} & (1+\rho^2) \Sigma^{-1} & -\rho \Sigma^{-1} & \dots 0 \\ 0 & & & & 1 \end{bmatrix} \tag{15}$$

where

ρ = correlation between adjacent image elements

$$\Sigma^{-1} = 1/(1-\rho^2) \begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 \\ -\rho & (1+\rho^2) & -\rho & 0 & \cdots & 0 \\ 0 & -\rho & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} \quad (16)$$

Multiplying S by the $(K^T)^{-1}$ is equivalent to convolving the image $S(i,j)$ with the spatial filter function $D(i,j)$.

$$D(i,j) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \rho^2 & -\rho(1+\rho^2) & \rho^2 & 0 \\ 0 & -\rho(1+\rho^2) & (1+\rho^2)^2 & -\rho(1+\rho^2) & 0 \\ 0 & \rho^2 & -\rho(1+\rho^2) & \rho^2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (17)$$

and Equation (4) and (5) become

$$F_S(i,j) = S(i,j) \otimes D(i,j) \dots \quad (18)$$

$$F_W(i,j) = W(i,j) \otimes D(i,j) \dots \quad (19)$$

Performance of this correlator can be described as follows:

(1) Experiments performed on selected images indicated that the statistical method does provide better performance in terms of providing a sharper peak at the location of image matching. In order for the method to work well, prior knowledge of the picture statistics is required. In an actual situation, this information is

either not available or extensive computations are required so that the correlator can be re-designed to tailor to the input data.

(2) It is anticipated that performance of this correlator would be worse than that of the basic correlator if the statistics of the input data differ from statistics used in the design of the correlator.

(3) Because of extra convolution steps needed in this method, the amount of computations needed by this method is even more than that needed for the basic correlator. It remains to show that the improvement in performance is worth the extra cost in computations.

1.1.3 Sequential Correlator

A common criticism of both the basic and statistical correlators is the great amount of computations that must be performed. A method of sequential correlation has been proposed [7] to reduce the computation time. The basic form of this algorithm is simple. An error function is defined as follows:

$$\delta(u, v) = \sum_{j=1}^K \sum_{i=1}^J |S(i, j) - W(i-u, j-v)| \quad \dots \quad (20)$$

Instead of testing each of the elements in the window area, elements of the area are selected at random. The error is accumulated for as each of the elements is compared. If the error exceeds a predetermined threshold value before all the elements in the window area are tested, the test is considered failed for the window (u,v) and a new window is tested. The test procedure is depicted in Figure 5. Curves A, B, and C depict the cumulative errors for three different reference points. A and B accumulate errors rapidly

and the tests terminate early. Curve C, however, accumulates error more slowly. It is therefore, much more likely to be a candidate as a matching point. Theoretical analyses and simulation tests [7] indicated that a saving of computation time of at least two orders of magnitude is possible. As with the other two correlation methods mentioned earlier, only very simply structured images were processed successfully by this new approach.

1.1.4 Sequential Template Matching

Extending the basic concept described in 1.2.3 a more complicated sequential detecting method was proposed [8]. Local similarity between a set of templates is matched to a given image. Instead of matching each template of a set to an image at every location, the templates are partitioned and a representative template is defined for each of the partitions. Several levels of partitions are defined. Elimination of mismatching locations and termination of computation can take place at each level of detection. Each level of testing is over a more restrictive subset of template class than the previous level. Matching process terminates when the accumulative template matching error exceeds a threshold level. A location which has gone through successive levels of matching without rejection is declared a likely candidate. The performance of the sequential template matching method can be described as follows:

1. Computation time is reduced due to the sequential method of testing. In matching a real image, several hundreds of templates are needed, thus making the task of template partition difficult.
2. In a sequential testing, the ordering of the features is important. Essential features (such as roads, rivers, etc.) must

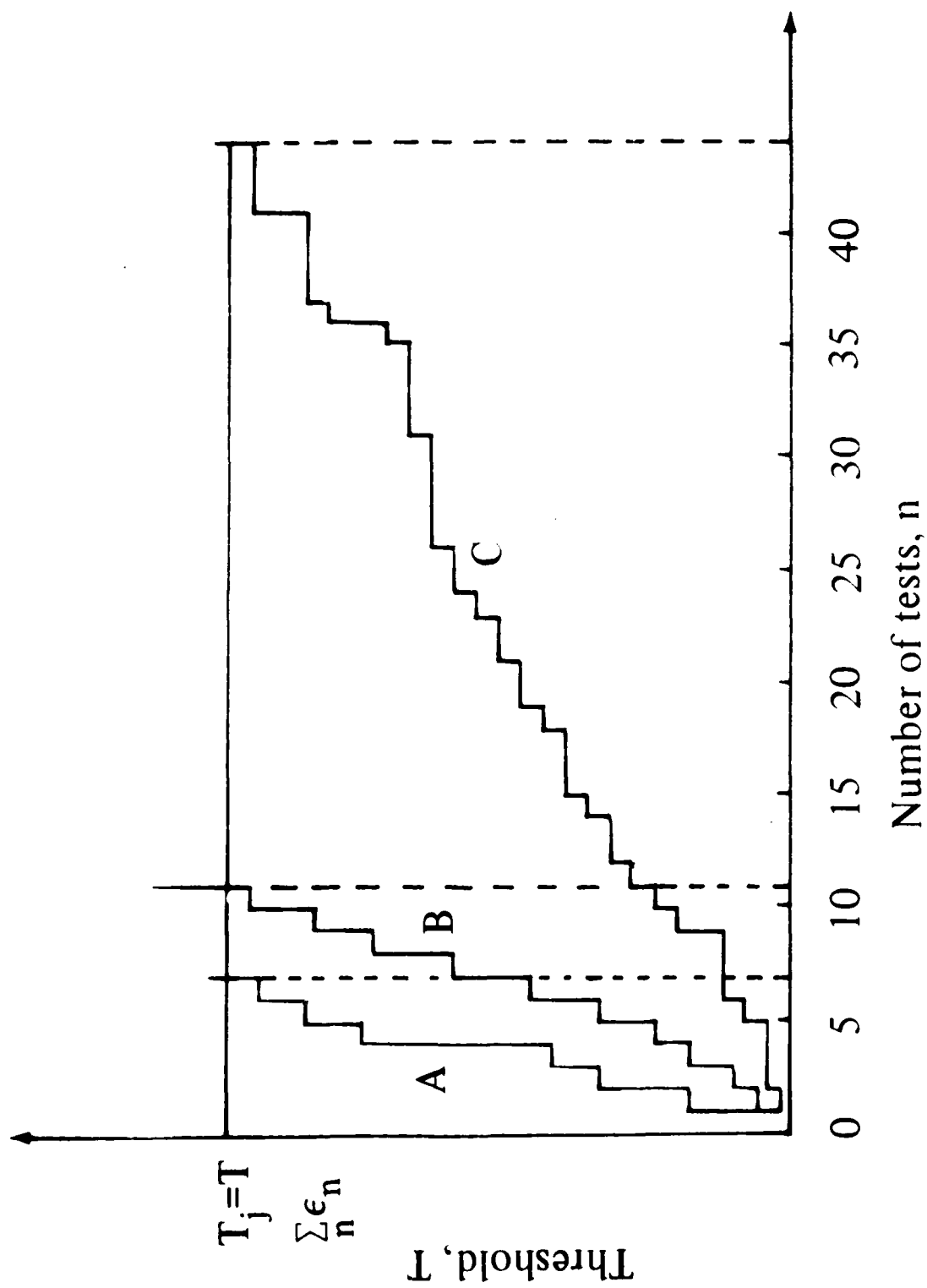


FIGURE 5. GROWTH CURVES FOR ALGORITHM A WITH $T = \text{CONST.}$

appear in both images in order to be considered as a match.

Methods of ordering features have not yet been developed.

1.2 OUTLINE OF REPORT

Two general approaches to map matching may be differentiated. The first approach may be called a pictorial method and generally involves a transformation of one image into a new image which corresponds to the reference image. The second approach may be called a feature method of the emphasis is placed on the second more robust method.

The pictorial method is developed in Section II. Transformations for matching maps from different sensors and geometrics are analyzed. Examples of perspective transformations are given for optical, radiometric, and radar images. A sensor transformation based upon a statistical matching of corresponding image points is also presented.

The use of edge features for map matching is described in Section III. A new orthogonal decomposition in terms of point, edge, line, and other basis vectors is presented. The advantage of this representation in a local region. Map matching using edge features also permits high speed logical similarity measurement. The number of computations using this technique is especially attractive for real time implementations. The use of edge features for map matching is also demonstrated. In the absence of geometric and sensor distortions, the method works well as indicated by the results of the experiments on matching an edge image to a noise-corrupted version of the same image. For optical to radar matching, the performance is somewhat degraded but can be improved using geometric and sensor intensity corrections as well as data pre-processing.

The theory of invariants is reviewed in Section. The study of algebraic invariants is an important field of mathematics and several results of this theory are summarized. The concept of perceptual invariants has also been studied for the visual system. Several invariant properties of the visual system are presented and a simple device which produces invariant measurements is described. The use of spatial moments for invariant measurements is described in Section 4.3. A powerful theorem which relates moment invariants to algebraic invariants for continuous functions is stated. New results indicating a degree of variance in the invariant measurements for discrete functions are presented. Examples are given which also illustrate that the magnitude of the variance can be controlled by careful processing.

Sequential techniques for searching for matching image locations are presented in Section V. These methods promise logarithmic efficiency over conventional correlation techniques. Zoom techniques are an important special case of the sequential techniques and are considered in detail. The effects of region size and resolution limitations are considered and demonstrated by several examples. Finally, recommendations for future work in map matching are given in Section VI.

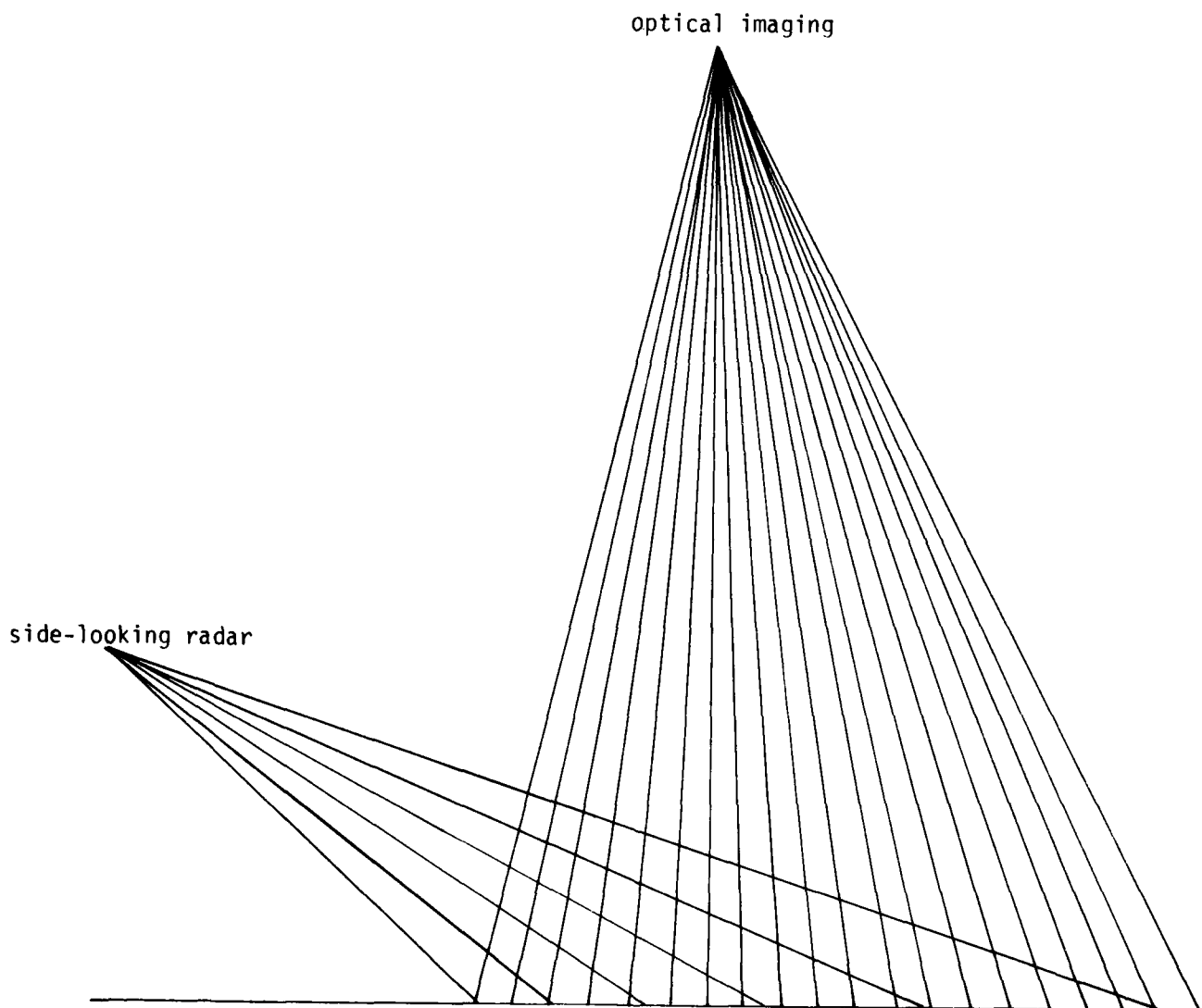


FIGURE 6. GENERAL GEOMETRY OF THE OPTICAL AND RADAR IMAGING SYSTEMS

SECTION II. TRANSFORMATIONS FOR MAP MATCHING

2.1 GEOMETRIC TRANSFORMATIONS

In this section, methods of registration of the optical and radar images are described. Figure 6 shows the general geometry of the optical and radar imaging systems together with an object terrain of interest. In this figure, the optical system is assumed pointing at the nadir and the radar system is scanning the same object terrain at a slant range. Figure 7(a) depicts a rectangular array of digitized pixels of the optical image. Figure 7(b) represents the rectilinearly digitized radar image plotted in the optical image coordinates. Point (x,y) in the optical image and point (x,y) in the radar image represent an image of the same object point in the terrain of interest. The basic transformation consists of deforming the radar image so that its features correspond as closely as possible to those in the optical image.

Three methods of transformations were studied. These are:

- Perspective Transformation

- Polynomial Estimate

- Interactive Nonlinear Transformation

2.1.1 Perspective Transformation

The general geometry of the optical imaging system is shown in Figure 8. In this figure, the object (ground) plane is assumed to be parallel to the image plane. The

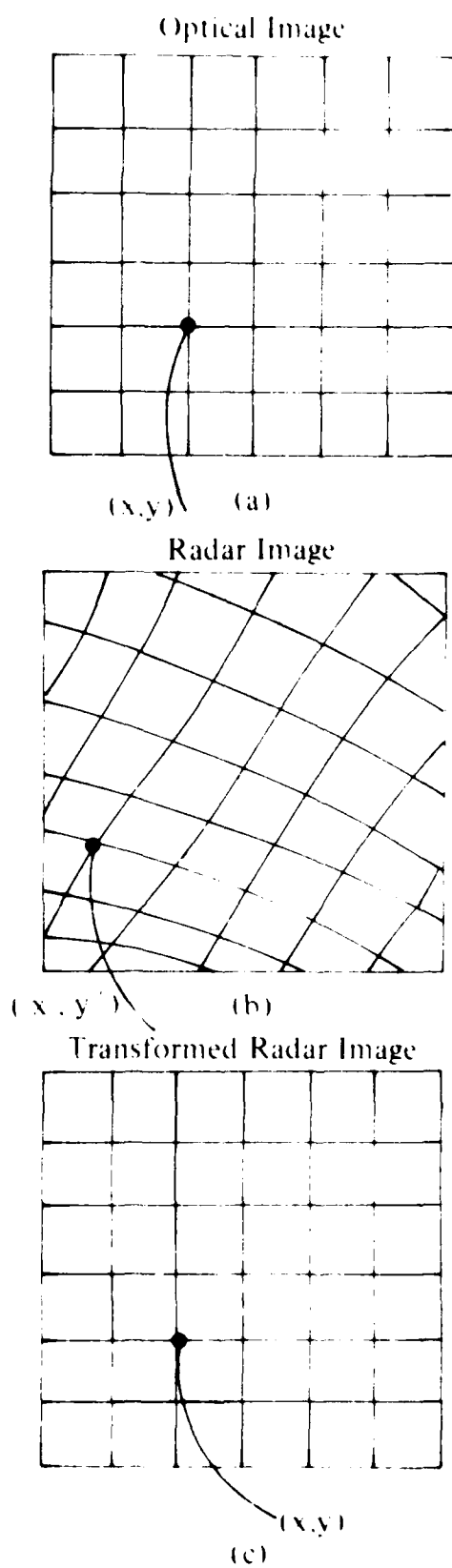


FIGURE 7. IMAGE TRANSFORMATION

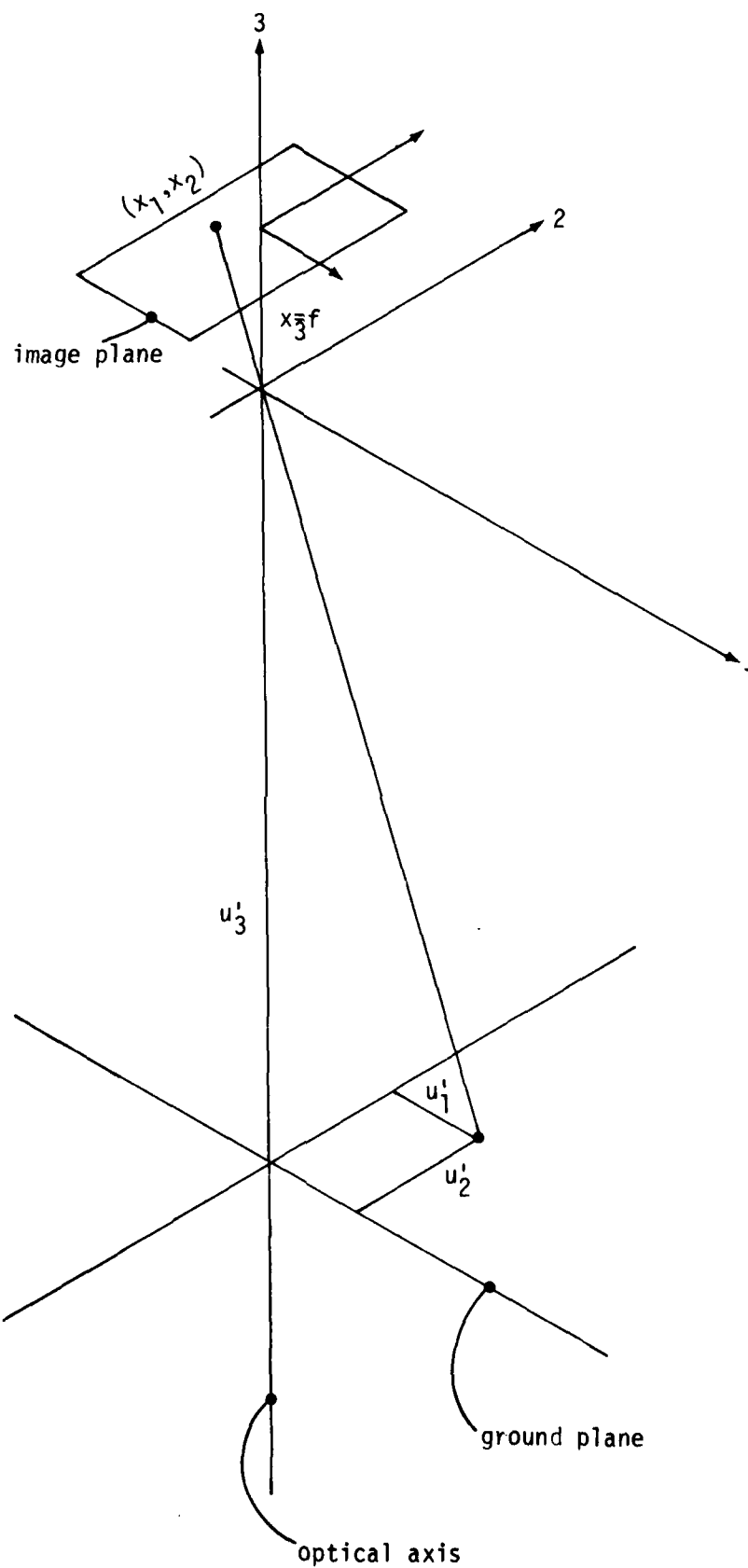


FIGURE 8. GEOMETRY OF OPTICAL IMAGING SYSTEM

optical axes is perpendicular to both planes and the entrance pupil of the camera is located at the origin of a rectangular coordinate system. From the lens law of geometrical optics, the object distance u_3' and image distance x_3 are related by:

$$\frac{1}{u_3'} + \frac{1}{x_3} = \frac{1}{f} \quad (1)$$

where f is the optical system focal length. Since the object distance is large compared to f , the image plane is the focal plane and $x_3 = f$. Any object point $u' = (u_1', u_2', u_3')$ is imaged into a point $x = (x_1, x_2, x_3)$ in the focal planes. The components of x can be computed by the following equations:

$$x_1 = \frac{f \cdot u_1'}{u_3'} \quad (2)$$

$$x_2 = \frac{f \cdot u_2'}{u_3'} \quad (3)$$

$$x_3 = f \quad (4)$$

The geometry of the radar imaging system is shown in Figure 9. In this geometry the radar image plane is assumed to be rotated β , θ and ϕ in the pitch, roll and yaw axes. Therefore, the orientation of the radar system is specified by the angles of rotation β , θ and ϕ . An object with a ground coordinates $u = (u_1, u_2, u_3)$ in the radar

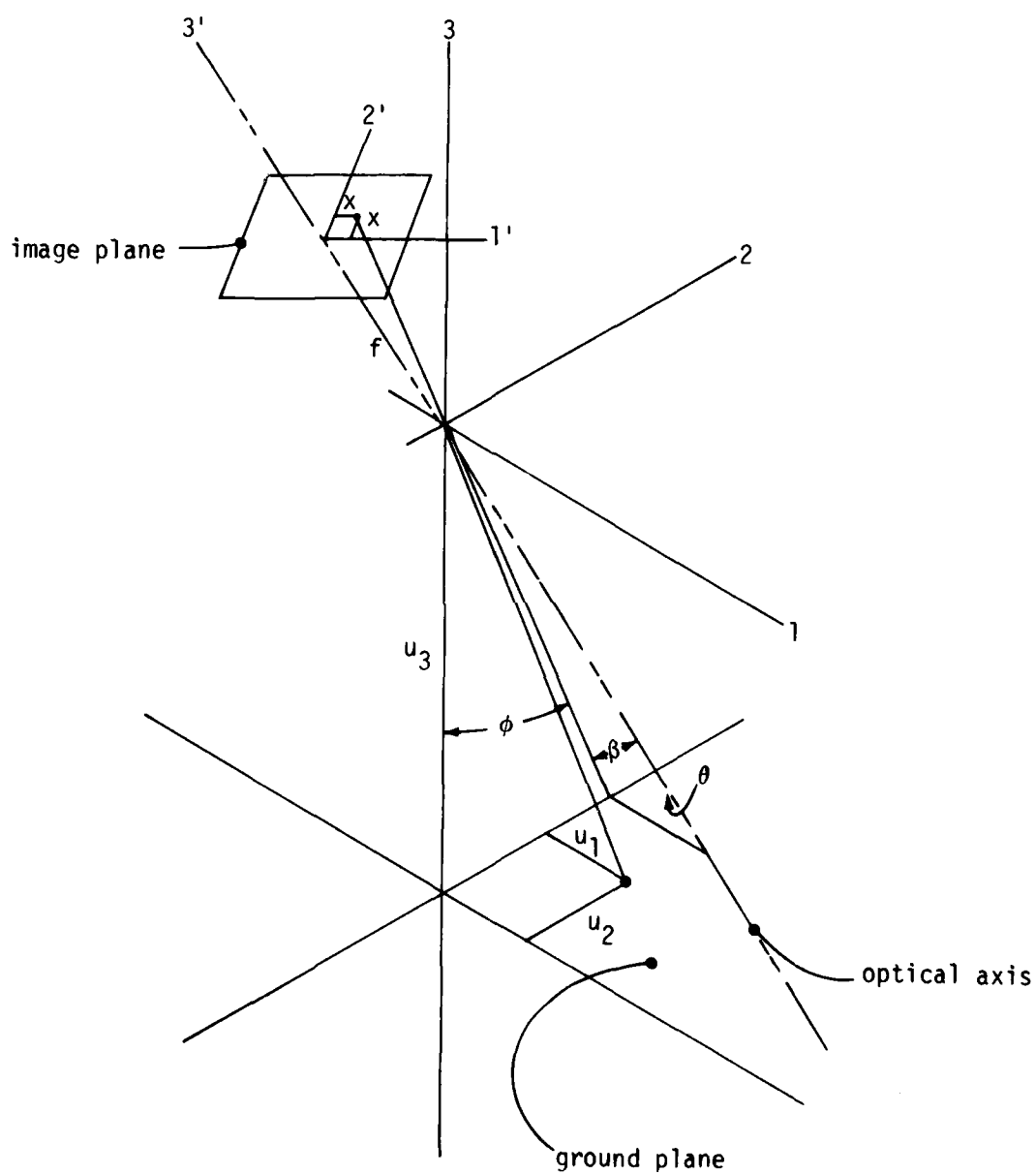


FIGURE 9. GEOMETRY OF SIDE-LOOKING RADAR

imaging system can be related to the old coordinates $u' = (u'_1, u'_2, u'_3)$ in the optical imaging system by the matrix equation

$$u' = Mu \quad (5)$$

where M is the product of three rotational transformation matrices:

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi \\ 1 & -\sin \varphi & \cos \varphi \end{bmatrix} \quad (6)$$

$$M_2 = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \quad (7)$$

$$M_3 = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (8)$$

Successive rotations in a pointing sequence can be defined by M which is the product of the three matrices (6), (7), and (8). A list of these product matrices for all six permutations of 1-, 2-, and 3- axis are:

$$m_{123} = \begin{bmatrix} (\cos \theta \cos \beta) & (\cos \theta \sin \beta \sin \varphi + \sin \theta \cos \varphi) & (\cos \theta \sin \beta \cos \varphi + \sin \theta \sin \varphi) \\ (\sin \theta \cos \beta) & (\cos \theta \cos \varphi + \sin \theta \sin \beta \sin \varphi) & (\sin \theta \sin \beta \cos \varphi - \cos \theta \sin \varphi) \\ (-\sin \beta) & (\cos \beta \sin \varphi) & (\cos \beta \cos \varphi) \end{bmatrix}$$

$$m_{132} = \begin{bmatrix} (\cos \theta \cos \beta) & (\sin \beta \sin \varphi - \sin \theta \cos \beta \cos \varphi) & (\sin \theta \cos \beta \sin \varphi + \sin \beta \cos \varphi) \\ (\sin \theta) & (\cos \theta \cos \varphi) & (-\cos \theta \sin \varphi) \\ (-\cos \theta \sin \beta) & (\sin \theta \sin \beta \cos \varphi + \cos \beta \sin \varphi) & (\cos \beta \cos \varphi - \sin \theta \sin \beta \sin \varphi) \end{bmatrix}$$

$$\begin{aligned}
m_{213} &= \begin{bmatrix} (\cos \beta \cos \theta - \sin \beta \sin \theta \sin \varphi) & (-\sin \theta \cos \varphi) & (\sin \beta \cos \theta + \cos \beta \sin \theta \sin \varphi) \\ (\sin \beta \cos \theta \sin \varphi + \cos \beta \sin \theta) & (\cos \theta \cos \varphi) & (\sin \beta \sin \theta - \cos \beta \cos \theta \sin \varphi) \\ (-\sin \beta \cos \varphi) & (\sin \theta) & (\cos \theta \cos \varphi) \end{bmatrix} \\
m_{231} &= \begin{bmatrix} (\cos \theta \cos \beta) & (-\sin \theta) & (\cos \theta \sin \beta) \\ (\sin \theta \cos \beta \cos \varphi + \sin \beta \sin \varphi) & (\cos \theta \cos \varphi) & (\sin \theta \sin \beta \cos \varphi - \cos \beta \sin \varphi) \\ (\sin \theta \cos \beta \sin \varphi - \sin \beta \cos \varphi) & (\cos \theta \sin \varphi) & (\cos \beta \cos \varphi + \sin \theta \sin \beta \sin \varphi) \end{bmatrix} \\
m_{312} &= \begin{bmatrix} (\cos \theta \cos \beta + \sin \theta \sin \beta \sin \varphi) & (\cos \theta \sin \beta \sin \varphi + \sin \theta \sin \beta) & (\sin \beta \cos \varphi) \\ (\sin \theta \cos \varphi) & (\cos \theta \cos \varphi) & (-\sin \varphi)(-\sin \varphi) \\ (\sin \theta \cos \beta \sin \varphi - \cos \theta \sin \beta) & (\cos \theta \cos \beta \sin \varphi + \sin \theta \sin \beta) & (\cos \beta \cos \varphi) \end{bmatrix} \\
m_{321} &= \begin{bmatrix} (\cos \theta \cos \beta) & (-\sin \theta \cos \beta) & (\sin \beta) \\ (\sin \theta \cos \varphi + \cos \theta \sin \beta \sin \varphi) & (\cos \theta \cos \varphi - \sin \theta \sin \beta \sin \varphi) & (-\cos \beta \sin \varphi) \\ (\sin \theta \sin \varphi - \cos \theta \sin \beta \cos \varphi) & (\sin \theta \sin \beta \cos \varphi + \cos \theta \sin \varphi) & (\cos \beta \cos \varphi) \end{bmatrix}
\end{aligned}$$

The matrices (6), (7) and (8) and their products M (in any order) are orthonormal, nonsingular, and have unity determinates. These properties imply that the transform in (5) can be inverted as:

$$u = M^{-1} u^1 \quad (9)$$

Substituting (5) into (2), (3) and (4) yields:

$$x_1 = C_1^{-1} (u_1, u_2, u_3) = f \cdot \frac{m_{11} u_1 + m_{12} u_2 + m_{13} u_3}{m_{31} u_1 + m_{32} u_2 + m_{33} u_3} \quad (10)$$

$$x_2 = C_2^{-1} (u_1, u_2, u_3) = f \cdot \frac{m_{21} u_1 + m_{22} u_2 + m_{23} u_3}{m_{31} u_1 + m_{32} u_2 + m_{33} u_3} \quad (11)$$

$$x_3 = C_3^{-1} (u_1, u_2, u_3) = f \quad (12)$$

where m_{ij} are the elements of the product matrix M.

Let the radar imaging system be located at an altitude h . Substituting h for u_3 in Equations (10, (11) and (12), we have:

$$x_1 = C_1^{-1} (u_1, u_2) = \frac{m_{11} u_1 + m_{12} u_2 + m_{13} h}{m_{31} u_1 + m_{32} u_2 + m_{33} h} \quad (13)$$

$$x_2 = C_2^{-1} (u_1, u_2) = \frac{m_{21} u_1 + m_{22} u_2 + m_{23} h}{m_{31} u_1 + m_{32} u_2 + m_{33} h} \quad (14)$$

$$x_3 = C_3^{-1} (u_1, u_2) = f \quad (15)$$

In the case that the side looking radar is tilted at an angle β and $\phi = \theta = 0$, equations (13) and (14) become:

$$x_1 = f \frac{-u_1 \cos \beta + h \sin \beta}{u_1 \sin \beta + h \cos \beta} \quad (16)$$

$$x_2 = \frac{-fu_2}{u_1 \sin \beta + h \cos \beta} \quad (17)$$

Given the grid points (u_1, u_2) in the object plane coordinates, Equations (13) - (17) can be used to compute the corresponding points (x_1, x_2) in the radar imaging system coordinates. In a quantized system, the computed coordinate points do not generally fall on the grid points in the radar image. So it is necessary to adapt rules for picking the appropriate element out of the radar image. The easiest method is to simply pick out the value of the nearest neighbor to the computed grid point. This method works well when the geometric distortion is not severe and when the x and u grids are about the same size. A second method is to take the average (possibly weighted average) of the neighbor elements. A third and more complicated method is to use interpolating polynomials as a means of computing values of the elements between points of known values.

2.1.2 Polynomial Estimate

Let the optical image intensity be given by $f(u_1, u_2)$ and the radar image intensity be given by $f(x_1, x_2)$. A technique can be used to map the radar image into the optical image using the following equations:

$$x_1 = g_1(u_1, u_2) \quad (18)$$

$$x_2 = g_2(u_1, u) \quad (19)$$

g_1 and g_2 can be approximated by a polynomial in u_1 and u_2 of degree N as:

$$x_1 = C_1(u_1, u_2) = \sum_{i=0}^N \sum_{j=0}^{N-i} k_{1ij} u_1^i u_2^j \quad (20)$$

$$x_2 = C_2(u_1, u_2) = \sum_{i=0}^N \sum_{j=0}^{N-i} k_{2ij} u_1^i u_2^j \quad (21)$$

where k_{1ij} and k_{2ij} are the constant polynomial coefficients. For most practical cases, a second degree ($N = 2$) approximation is adequate. This has been verified by experiments[9] (on data from side look radar images) in which it was found that the decrease in rms error as a function of increasing polynomial terms is not significant for N greater than 2.

The coefficients of Equations (20) and (21) can be computed by fitting the two dimensional functions to a set of $f(u_1, u_2)$ and $f(x_1, x_2)$ values. The linear - least - squares estimate procedure can be used to express $C_1(u_1, u_2)$ surfaces by polynomials whose squared distances from the true surface is a minimum. To obtain the coefficients, prominent features that appeared in both the optical and radar images are selected. These prominent features can be:

- End points of a long edge
- Intersection of two lines
- Corners of detected squares
- Distinguishable points on the boundary of texture region

For a second degree ($N = 2$) approximation, at least six pairs of conjugate points are needed. The values of these

points can be arranged in the following manner:

$$\begin{bmatrix} 1 \\ x_1 \\ 2 \\ x_1 \\ 3 \\ x_1 \\ \vdots \\ \vdots \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 & u_1 & u_2 & u_1^2 & u_2^2 & u_1 u_2 \\ & & & \cdot & & \\ & & & \cdot & & \\ & & & \cdot & & \\ & & & \cdot & & \\ & & & \cdot & & \\ & & & \cdot & & \end{bmatrix} \begin{bmatrix} k_{100} \\ k_{110} \\ k_{101} \\ k_{120} \\ k_{102} \\ k_{111} \end{bmatrix} \quad (22)$$

Where $R \geq 6$.

A similar equation exists for x_2 and k_{2ij} . Both of these may be written in matrix notation as:

$$x_1 = U K_1 \quad (23)$$

$$x_2 = U K_2 \quad (24)$$

Using the linear - least - squares estimate theory, the best estimate for K_1 and K_2 are given by the pseudo inverse solution:

$$K_1 = (U^T U)^{-1} U^T x_1 \quad (25)$$

$$K_2 = (U^T U)^{-1} U^T x_2 \quad (26)$$

2.1.3 Interactive Nonlinear Transformation

In this section, a method of removing spatial distortion in the radar image with respect to the optical image is described. This method, which was first described by

Ulstad [10] is currently being studied as a possible method of removing geometric distortion by a nonlinear spatial transformation.

The spatial registration system is shown in Figure 10. The two sets of image data are sequentially available as columns of matrices X and U each with L rows. The L -element column vectors of X and U are processed by the system in order to produce a third L -element column vector R . The L elements of the column of U map into L corresponding points in X based on the following equation.

$$u(i,k) = x[i + \beta(i,k), k + \alpha(i,k)] \quad (27)$$

The quantities $\beta(i,k)$ and $\alpha(i,k)$ are real scalar correction factors are functions of the indices N and k .

As shown in Figure 11, use is made of N submatrices defined in U . Submatrix U_n consists of $2p$ rows labeled $s_n - p$ through $s_n + p$ where $1 \leq n \leq N$, p is an integer and

$$s = \frac{L}{N+L} \quad (28)$$

Let the N index points (a_n, b_n) in Matrix X be such that:

$$U(s_n, k) = X(a_n, b_n), \quad (29)$$

The points (a_n, b_n) in X correspond to the middle row of the n th submatrix U at the k th column. The points will be used as the corner points of a piecewise linear synthetic scan line in the transformation. In determining (a_n, b_n) during the processing of the k th column, correlations of

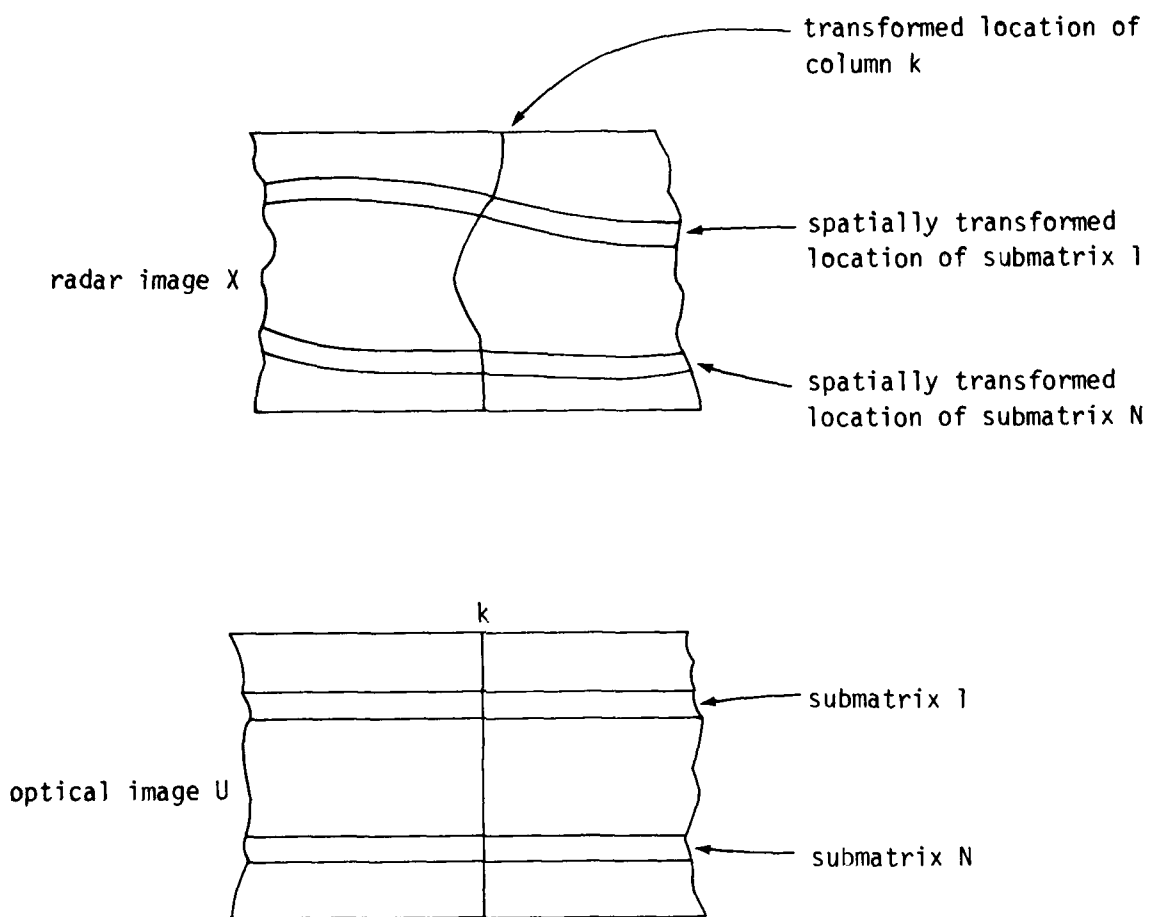


FIGURE 10. SPATIAL REGISTRATION

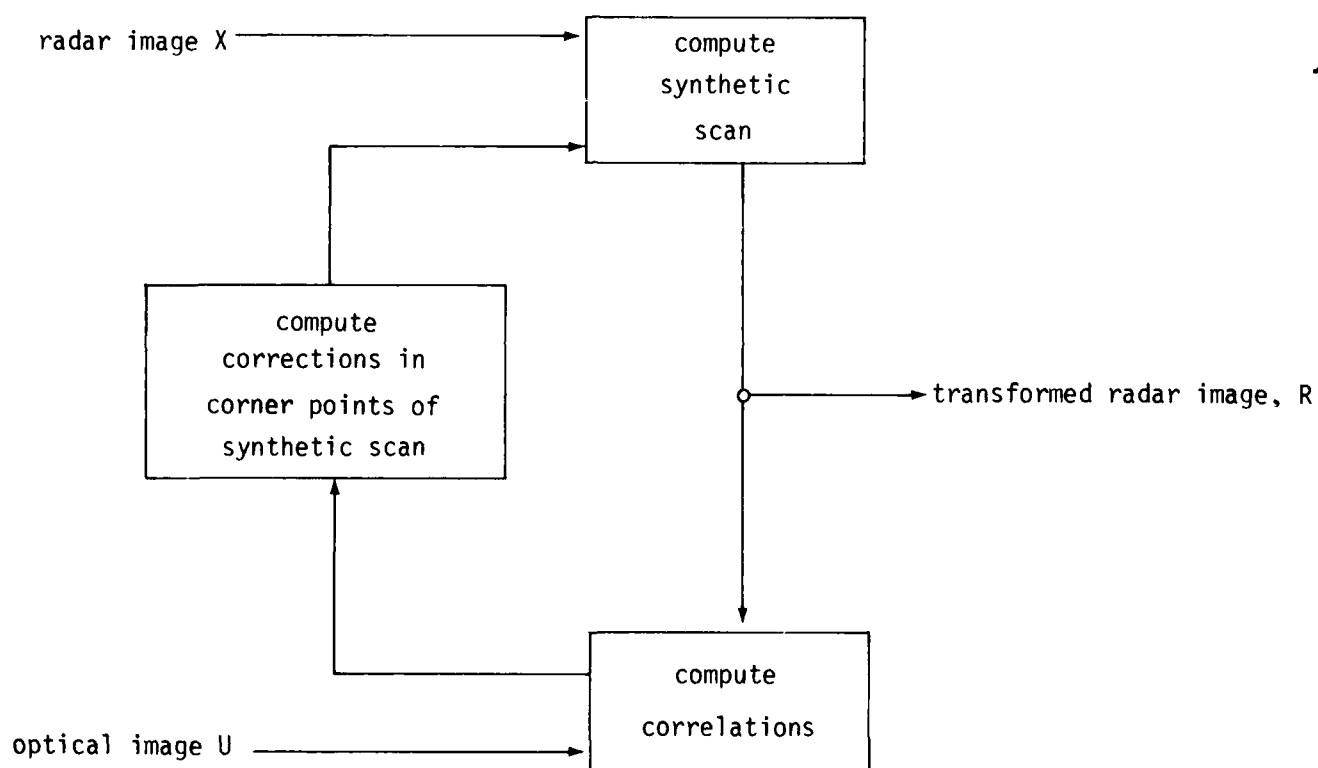


FIGURE 11. INTERACTIVE TRANSFORM SYSTEM

$R(i,k)$ of matrix R and $U(i,k)$ of matrix U are made for i , in ranging from $sn-p$ to $sn+p$. The results of these correlations is then used to compute a new corner point $(a_n + \Delta a_n, b_n + \Delta b_n)$. The corner point with the updated value can then be used to produce the $(k+1)$ th column of the R matrix. This process continues until an entire matrix R is produced.

2.1.4 Results

The polynomial method is preferred over the perspective transformation. Perspective transformation requires tracking and ephemeris data to determine the positions of the sensors at the time of exposure. These data are often not available. In the polynomial method, data needed for the transformation are entirely contained in the images.

Figures 12, 13 and 14 show the results of geometric distortion corrections by the method discussed in Section 2.1.2. Figure 12 shows the various possible degrees of corrections. Figure 13 shows a radiometric and an optical picture. Geometric transformation was performed on the optical picture in order to match the radiometric picture. Corrections were concentrated on matching the three straight roads. A total of eight reference points from each of the two original pictures was used for the transformation. Figure 14 shows two pictures taken by two different sensors, an optical camera and a radar, at different look angles. The transformation is made on the optical image so that it can be geometrically in registration with the radar image.



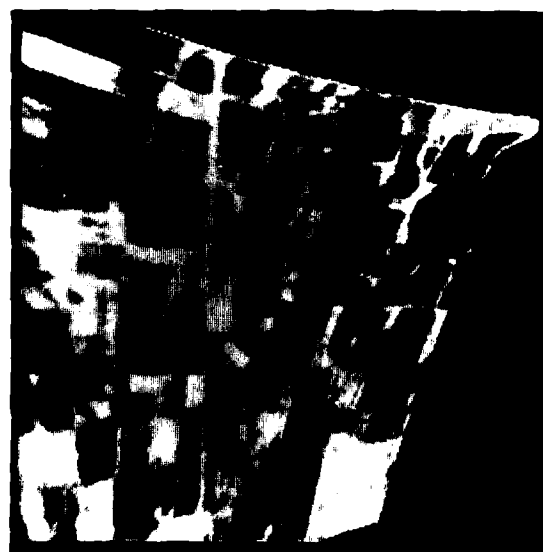
Radar



Micrad



Transformed Micrad



Transformed Micrad

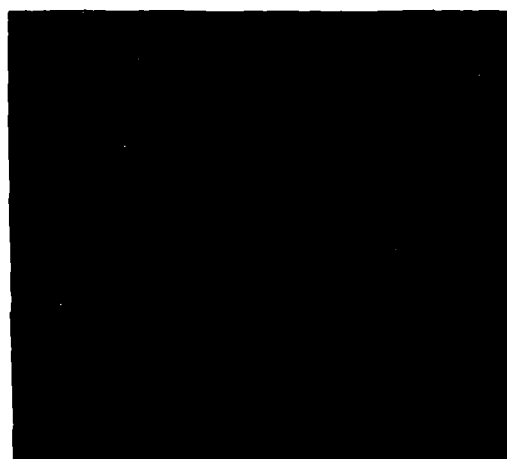
Figure 12. Polynomial Geometric Corrections



Original Micrad
(a)



Original Optical
(b)



Transformed Optical
(c)

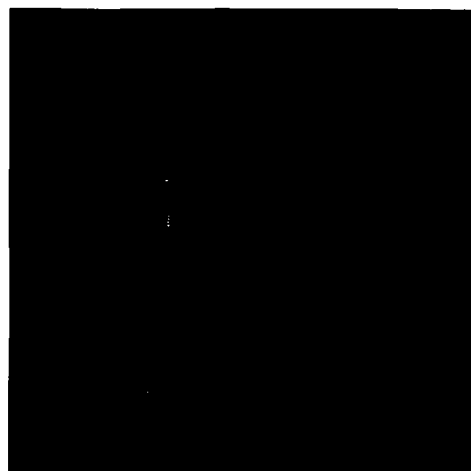


Superimposed Micrad
and Transformed Optical
(d)

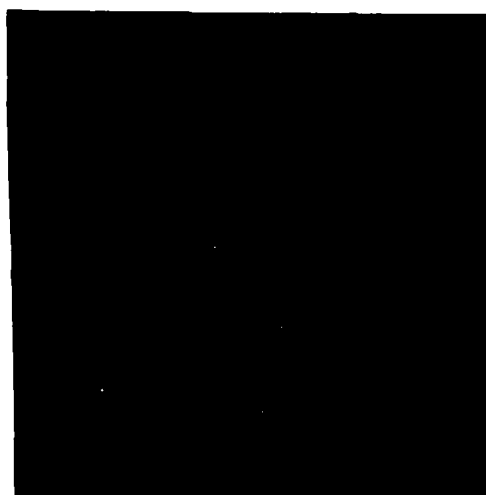


Superimposed
(e)

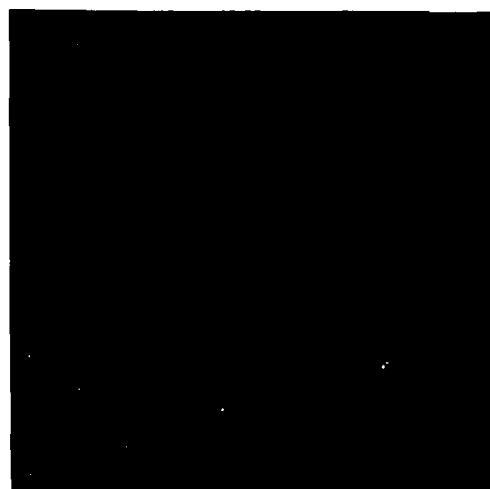
Figure 13. Geometrical Transformation



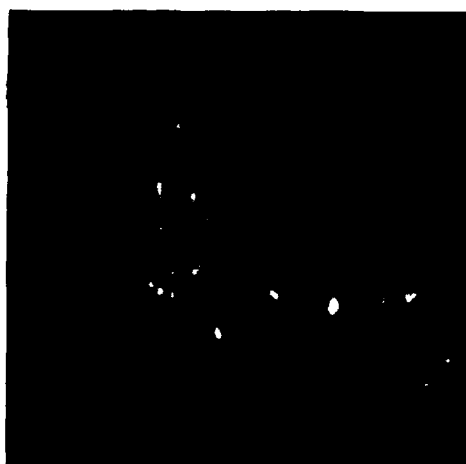
Original Radar
(a)



Original Optical
(b)



Transformed Optical
(c)



Superimposed Radar and
Transformed Optical
(d)



Superimposed
(e)

For this transformation, a total of twelve reference points from each of the two original images is used to form based for a pseudo-inverse transformation. As shown in Figure 14 (d), a nearly perfect registration was obtained.

2.2 Sensor Transformation

Because of the differences in operating characteristics of the two sensors, images of the same object taken by the optical system and the radar system will have different intensity profiles. The most prominent difference is that the optical image is positive while the radar image is negative. In this section, methods of intensity transformation are described. These methods consist of transforming the radar image so that its intensity profile matches as closely as possible to that of the optical image. The two methods studied are:

- Intensity Equalization using Karhunen-Loeve Transform
- Equalization using Intensity Averaging

2.2.1 Intensity Reversal

As the first step in the intensity transformation, the negative radar image is transformed into a positive image. This is done by replacing the amplitude of each picture element e_{ij} by its complementing value $\bar{e}_{ij}$. For an image which has been digitized with n bits:

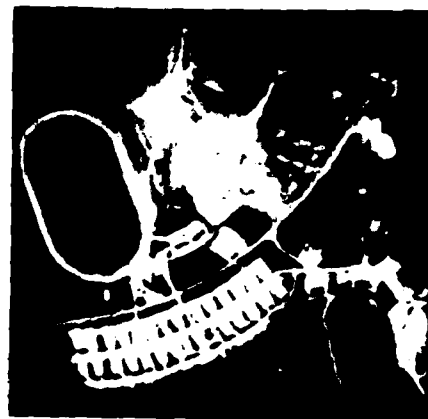
$$\bar{e}_{ij} = 2^n - e_{ij} - 1 \quad (30)$$

where 2^n is the number of quantization level. Figure 14(a) shows the original radar image. An intensity reversal is performed on this image by applying Equation(30). The result is shown in Figure 15(a). It can be seen that the contrast of the image is low and intensity modification must be performed so that its intensity profile matches as closely as

possible to that of the optical image shown in Figure 15(b).



(a) Radar Image
Intensity Reversed



(b) Optical Image



(c) Intensity Transformation
(Karhunen-Loeve Transform)



(d) Intensity Transformation
(Averaging)

Figure 15. Intensity Transformation

2.2.2 Intensity Equalization Using Karhunen-Loeve Transform

A two-dimensional image intensity function $f(x_1, x_2)$ can be constructed as follows:

$$f(x_1, x_2) = n \quad (31)$$

where

x_1 = Intensity level of the picture element in the radar image,

$$0 \leq x_1 \leq 2^n - 1$$

x_2 = Intensity level of the picture element in the optical image,

$$0 \leq x_2 \leq 2^n - 1$$

n = The numbers of picture elements which have intensity level of x_1 in the radar image and the intensity level of x_2 in the optical image.

A covariance matrix Σ_x is then formed to indicate the relative intensity correlation between the two images as follows:

$$\Sigma_x = E[(x_1 - M_{x1})(x_2 - M_{x2})^T] \quad (32)$$

where

M_{x1} = Expected values of x_1

M_{x2} = Expected values of x_2

A Karhunen-Loeve transform can then be performed to obtain a new set of coordinates ϕ_1 and ϕ_2 as follows:

$$\Sigma_\lambda = \phi \Sigma_x \phi^T = \Lambda \quad (33)$$

where Λ = Diagonal matrix with the eigenvalues λ_i of Σ_x as the diagonal elements

φ = Matrix whose columns are the ordered eigenvectors of Σ_x
 $\varphi = [\varphi_1, \varphi_2]$ for $\lambda_1 \geq \lambda_2$ (34)

The transformed coordinates ϕ_1 and ϕ_2 are shown in Figure 16. Based on ϕ_1 and ϕ_2 intensity corrections can now be made by scaling the intensity profile of the radar image such that the resulting coordinates ϕ_1 and ϕ_2 are orthogonal and the angle between ϕ_1 and the f_1 - axis is approximately 45 degrees.

Figure 17 shows a three-dimensional plot of the resulting intensity transformation. It can be seen from this figure that both images have approximately the same numbers of picture elements at each intensity level. Figure 15(c) shows the resulting transformed radar image. In this figure, considerable details which are not visible in the original radar image have been made visible by the intensity transformation.

2.2.3 Intensity Equation Equalization by Intensity Averaging

The intensity transformation method described in Section 2.2.2 can be modified such that the average intensity of the transformed radar image is made approximately equal to the average intensity of the optical image. The resulting transformation based on this method is shown in Figure 15(d). In this method, a compromise is made to match the overall intensity of the two images whereas image transformation using

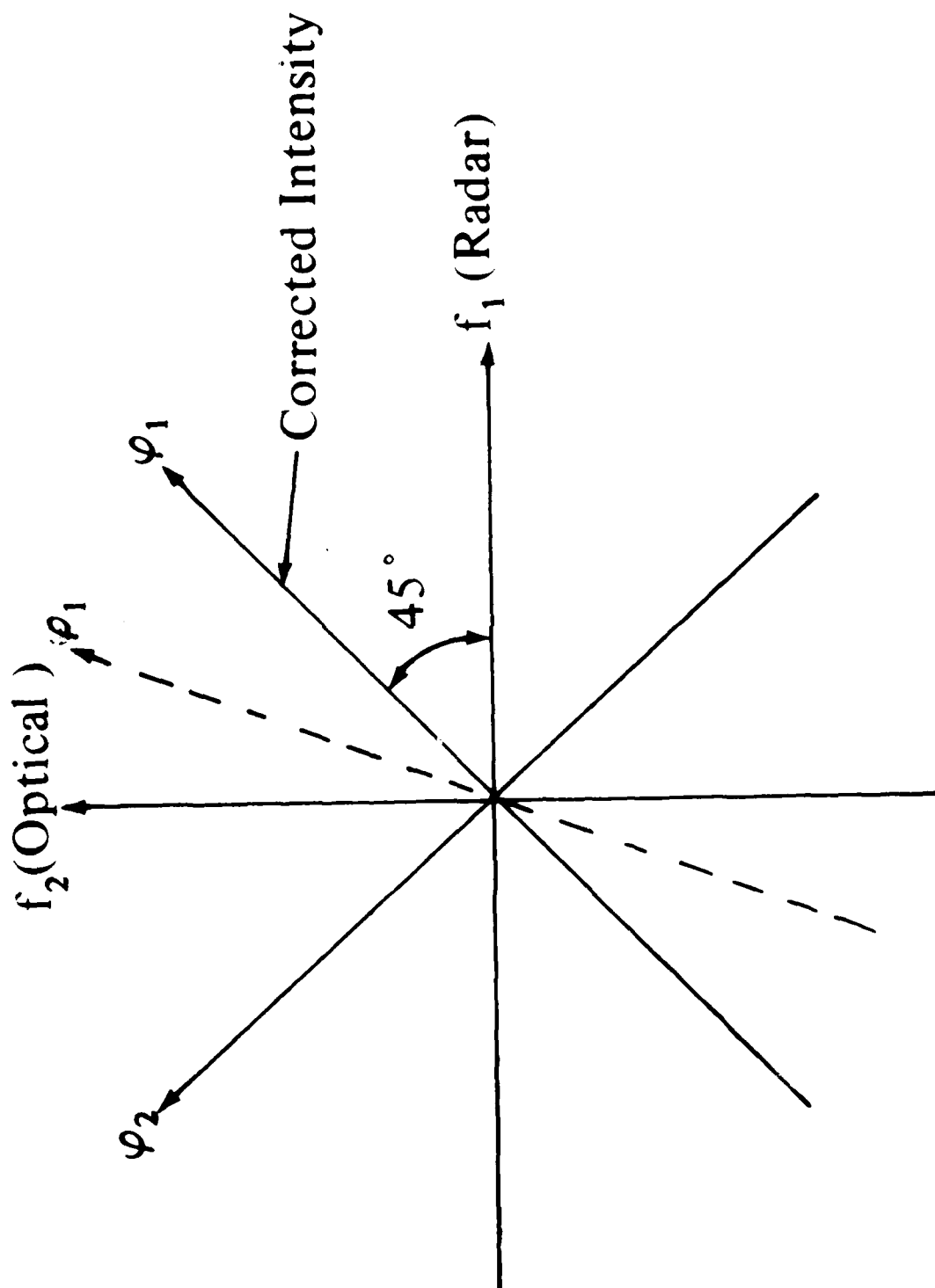


FIGURE 16. INTENSITY COORDINATE TRANSFORMATION

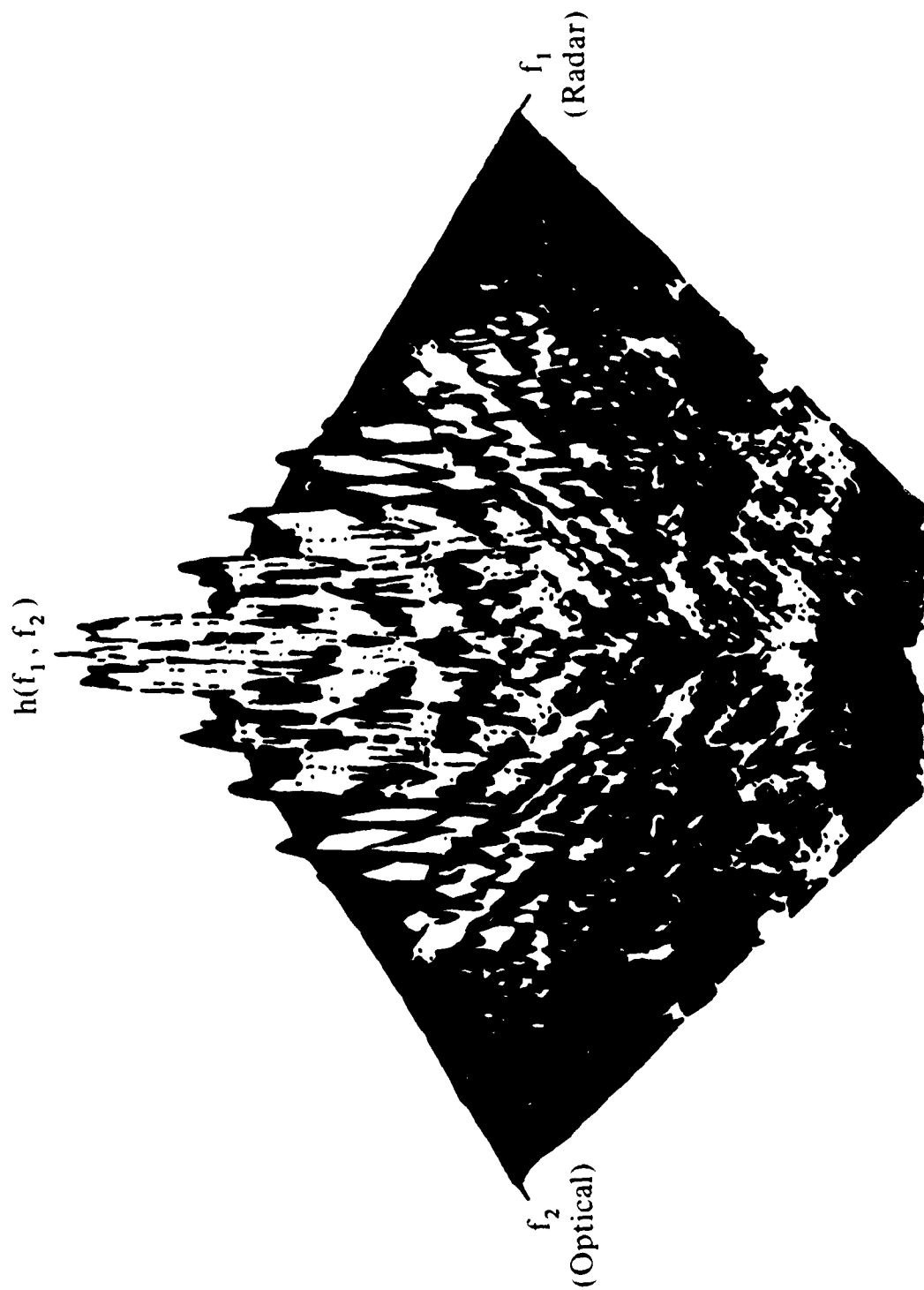


Figure 17. IMAGE INTENSITY FUNCTION, $h(x_1, x_2)$

the method described in Section 2.2.2 products a better match in tracking the variations of local intensity. Matching the overall intensity is most important whenever a non-normalized similarity measure is used.

2.3 Fourier Transform

The most salient of the optical systems was developed by Lendaris and Stanley for terrain classification. The sensor or feature extractor measured aspects of the optical Fourier transform in specified geometric patterns. The initial system utilized coherent optical techniques to extract measurements (features) of the Fraunhofer diffraction pattern which were classified via an interactive nonstatistical classification method. In the present application the image $f(x, y)$ is transformed optically into its Fourier space $F(v_x, v_y)$ via the well-known property that coherent light passed through an image $f(x, y)$ and then through a thin convex lens produces the diffraction pattern in the back focal plane. The coordinates (u, v) of this transform plane are directly related to spatial frequencies and of the input image. This is shown in Eqs. (35) and (36).

$$\hat{F}(u, v) = C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{2\pi f(ux + vy)}{\lambda f}} f(x, y) dx dy \quad (35)$$

$$v_x = \frac{u}{\lambda f}, \quad v_y = \frac{v}{\lambda f} \quad (36)$$

The constant λ is the wavelength of the incident light, f is the focal length of the transform lens, and C is the amplitude transmission factor. The sensor is composed of 32 angular wedges which sample the squared modulus of the diffraction pattern (power spectrum $I(u, v)$) and produces

appropriate signatures

$$I(u, v) = \hat{F}(u, v) \hat{F}^*(u, v) = \left| \hat{F}(u, v) \right|^2, \quad (37)$$

where $\hat{F}$ denotes a complex function and $*$ denotes conjugate. The annular ring sample signature is expressed in polar coordinates as

$$r_j = \int_0^\pi \int_{\rho_j}^{\rho_j + \Delta\rho} I(\rho, \theta) \rho d\rho d\theta, \quad (38)$$

$$j = 1, 2, \dots, 32$$

where $\rho = (u^2 + v^2)^{1/2}$ and $\theta = \tan^{-1} v/u$. Each r_j signature value represents the total power in the annular region $(\rho_j, \rho_j + \Delta\rho)$. The wedge signature values may be similarly described by

$$w_j = \int_{\rho_{\min}}^{\rho_{\max}} \int_{\theta_j}^{\theta_j + \Delta\theta_j} I(\rho, \theta) \rho d\rho d\theta \quad (39)$$

$$j = 1, \dots, 32$$

in which each w_j represents the total power in an angular frequency band from $(\theta_j, \theta_j + \Delta\theta)$.

The focal length f of the transform lens is adjusted so that the outer ring of the detector images 8 $\lambda\rho/\text{mm}$. The annular ring signature provides a rotationally insensitive

means of detecting one- or two-dimensional spatial periodicity within apertured areas. In contrast, the wedge signatures are insensitive to periodic structure but sensitive to direction. Both frequency signatures subsequently have yielded valuable features for object detection.

To initially evaluate the effectiveness of Fourier features for map matching, the following experiment was performed. Four corresponding regions in an optical and radar scene were selected as illustrated in Figure 18. The Fourier transform of the corresponding regions were computed using a coherent optical system. Comparative photographs of the magnitudes of the Fourier transforms of the regions are shown in Figures 19 to 22. Note that the most obvious pattern is that the optical image contains more high frequency energy corresponding to greater detail in the optical images. Also, a correlation between line structures in the images and transforms is visible in several of the patterns. To investigate the transforms with a more sensitive detector, ring and wedge measurements were made using the RSI detector. Graphs of these results are shown in Figures 23 to 30. Note that these measurements indicate a correlation in both ring and wedge measurements. Thus, more quantitative studies such as map matching with these measurements are required.

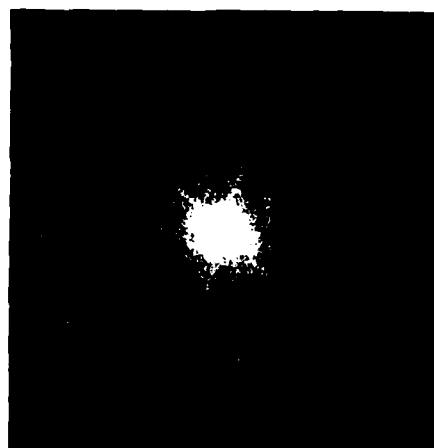


(a) Optical Image



(b) Radar Image

Figure 14. Corresponding scene selected for Fourier transforms measurements

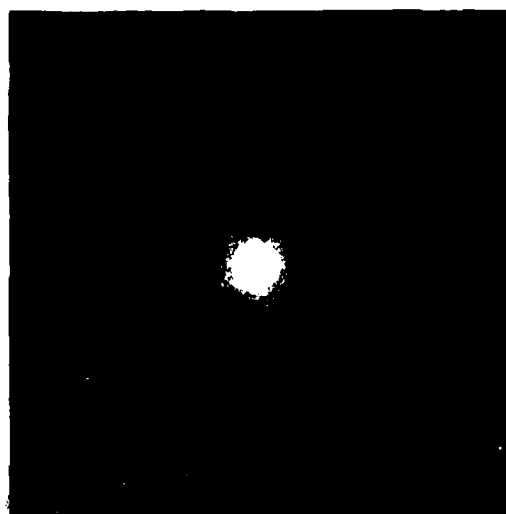


(a) Fourier transform of
optical image of the
rural field region

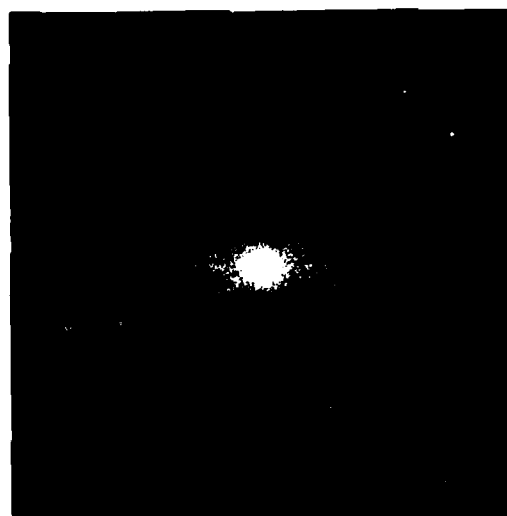


(b) Fourier transform of
radar image of rural
field region

Figure 19. Fourier transform of rural field region

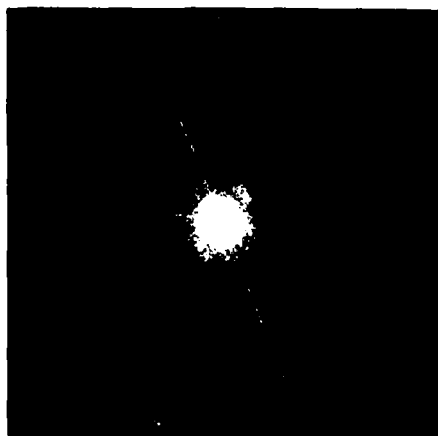


(a) Fourier transform of optical
image of road intersection
region

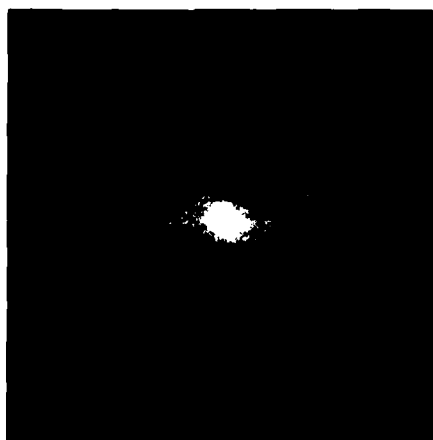


(b) Fourier transform of radar
image of road intersection
region

Figure 20. Fourier transform of road intersection region

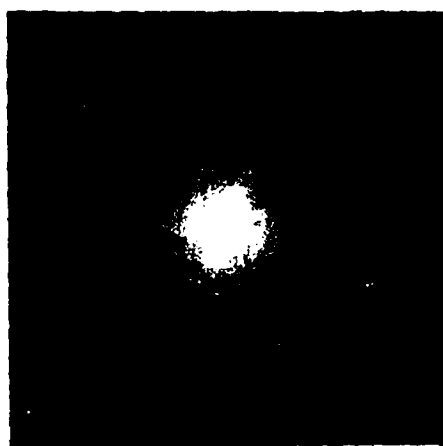


(a) Fourier transform of optical
image of curving road region



(b) Fourier transform of radar
image of curving road region

Figure 21. Fourier transform of curving road region



(a) Fourier transform of optical
image of stadium region

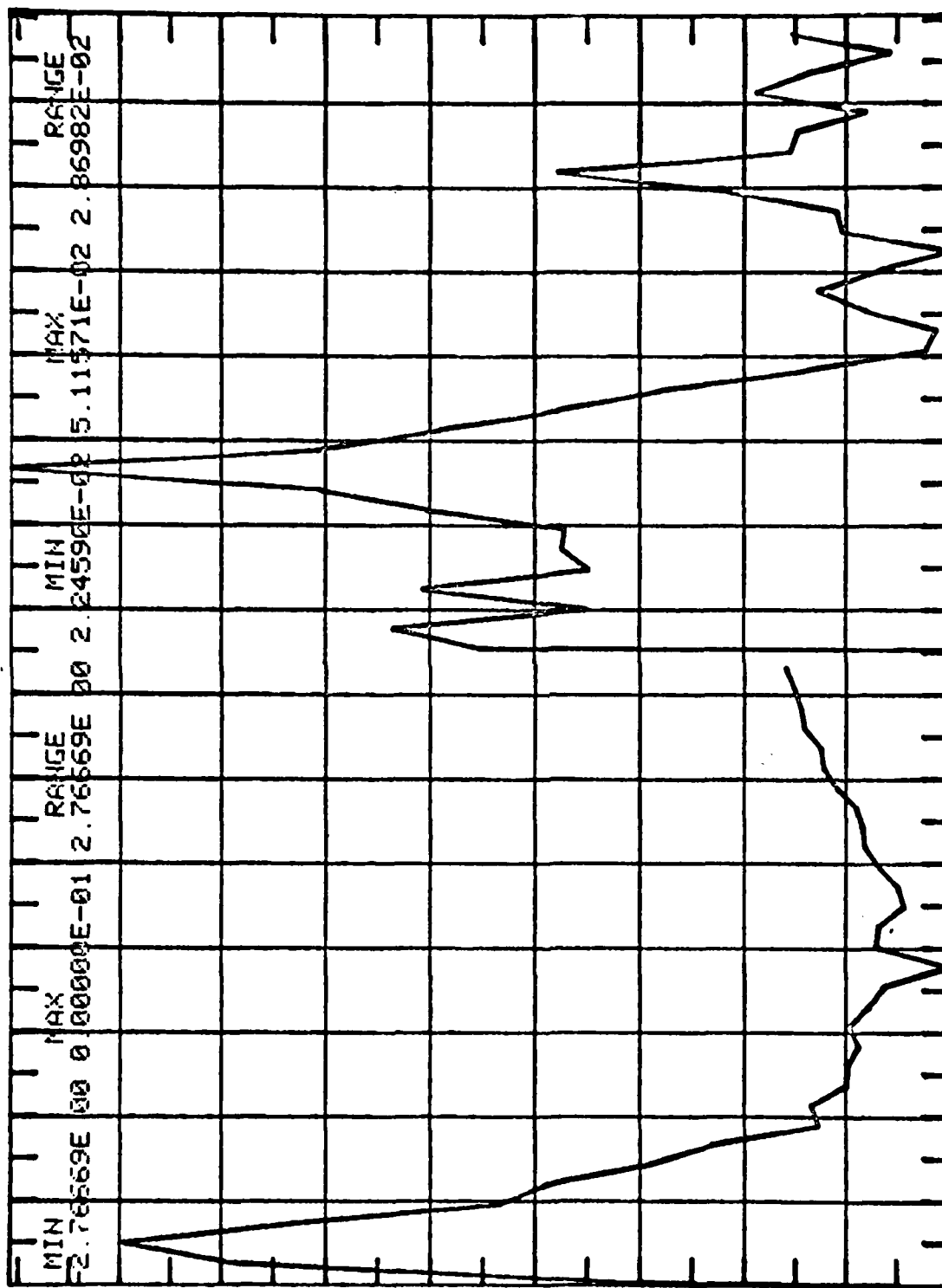


(b) Fourier transform of radar
image of stadium region

Figure 22. Fourier transform of the stadium region

Figure 23.

FOURIER TRANSFORM SIGNATURE OF THE OPTICAL IMAGE OF RURAL FIELD REGION



Sensor No. 1

Spatial

Frequency

(lp/mm): 0.0

1.21

8.0

Angle: 0°

90°

180°

RINGS

32 33

WEDGES

64

Figure 24.

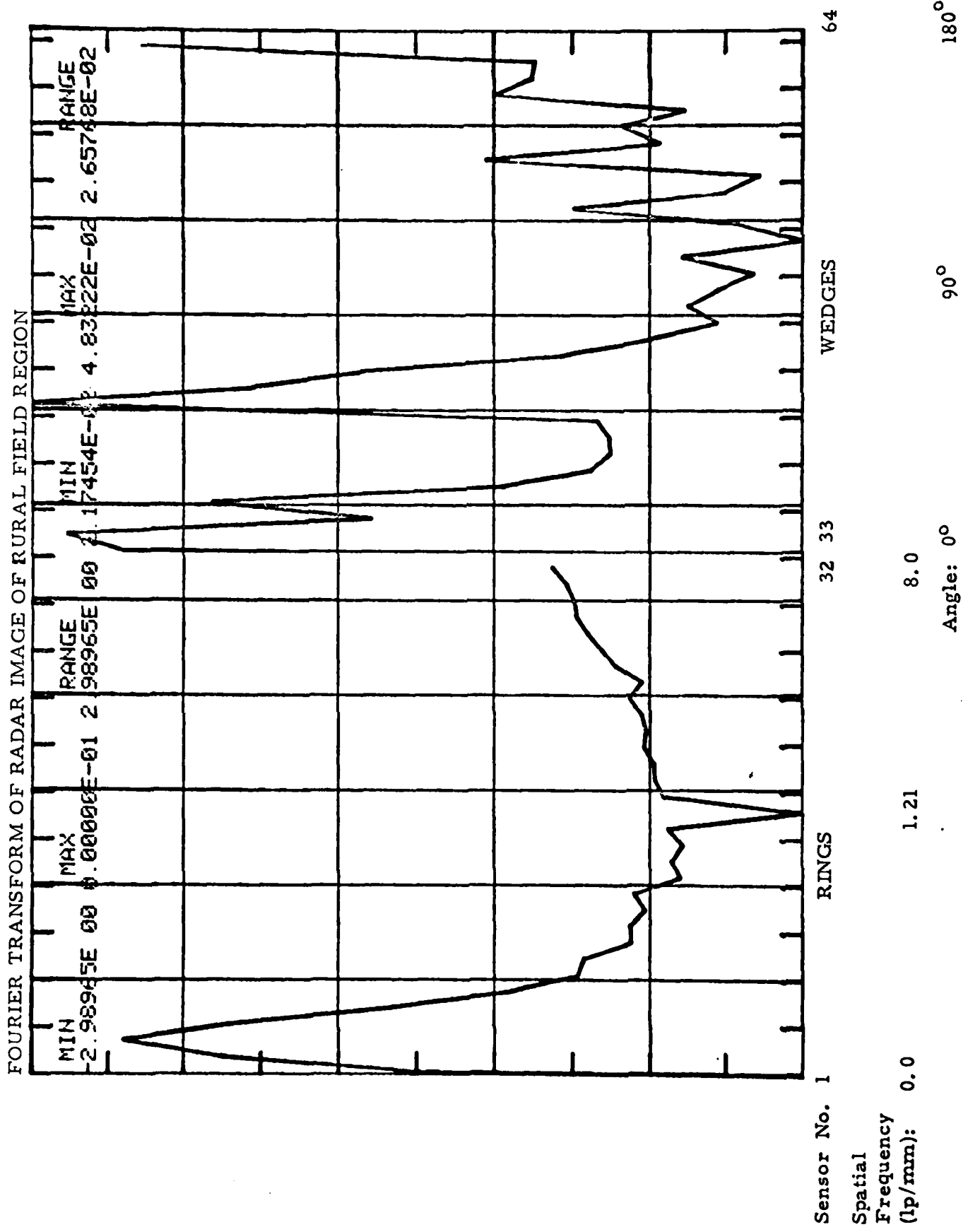
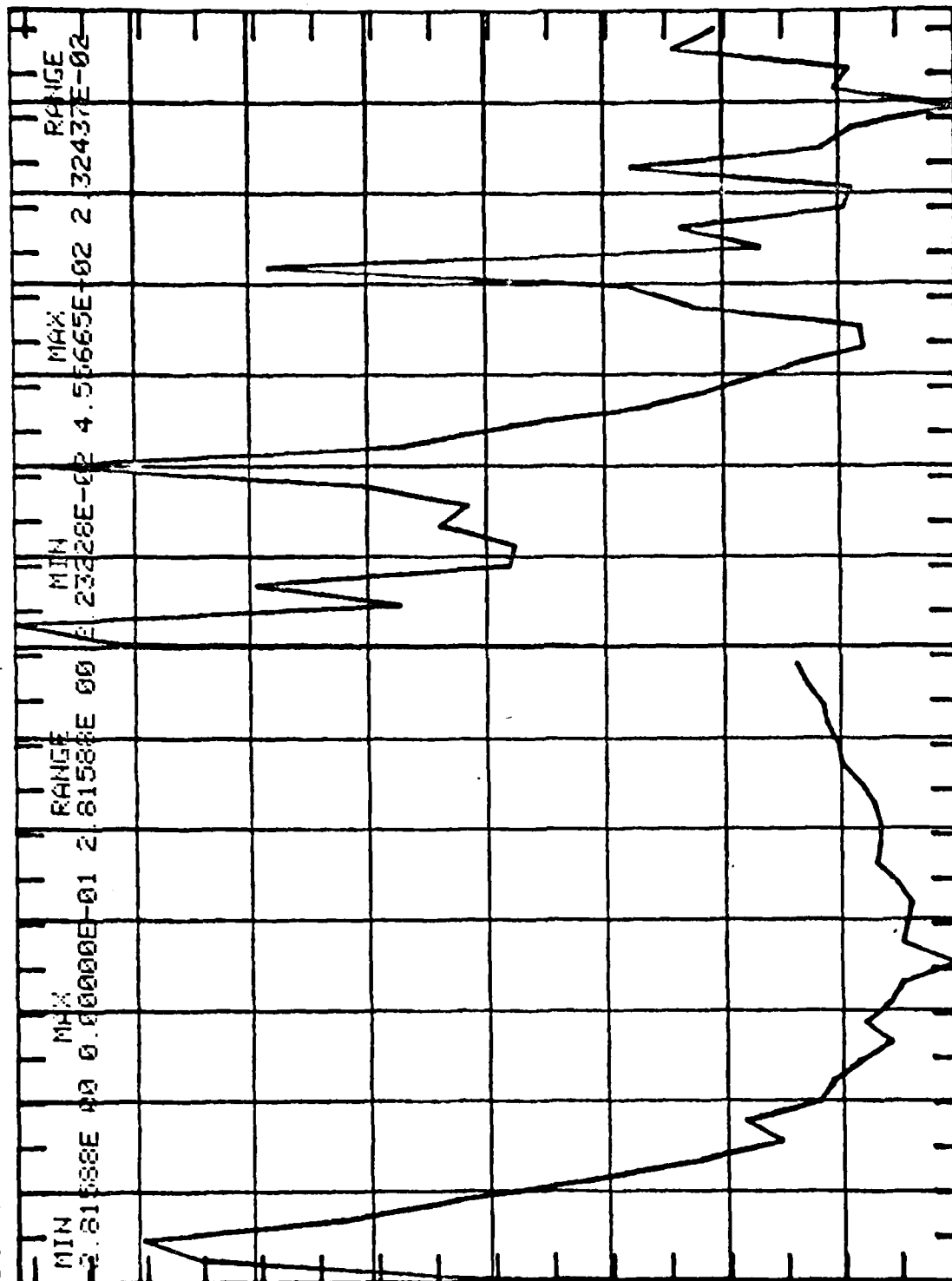


Figure 25.

FOURIER TRANSFORM OF OPTICAL IMAGE OF ROAD INTERSECTION REGION



Sensor No. 1

Spatial

Frequency

(lp/mm): 0.0

1.21

8.0

Angle: 0°

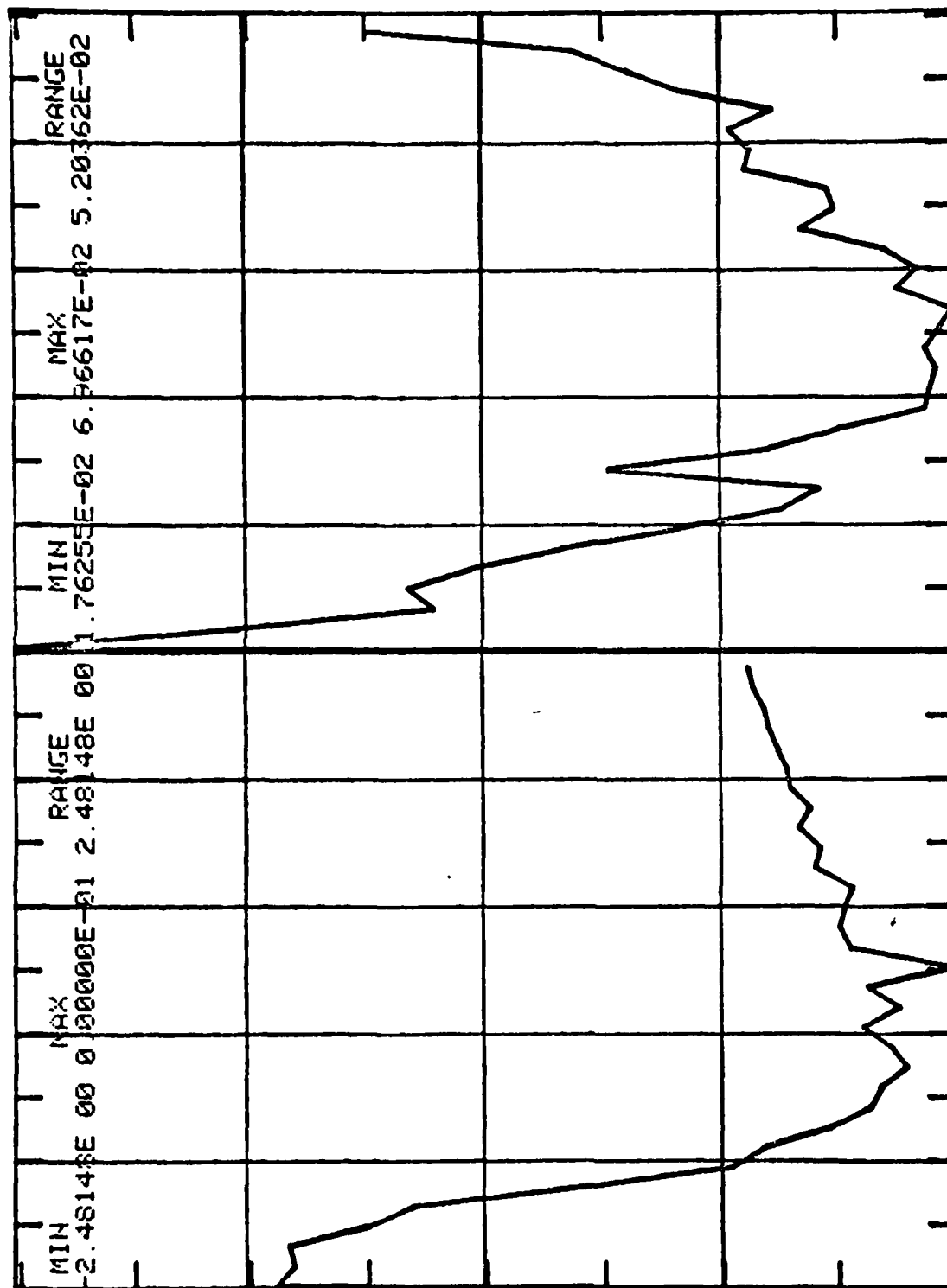
90°

180°

64

Figure 26.

FOURIER TRANSFORM OF RADAR IMAGE OF ROAD INTERSECTION



Sensor No. 1
Spatial
Frequency
(lp/mm): 0.0

RINGS

32 33

WEDGES

64

1.21

8.0

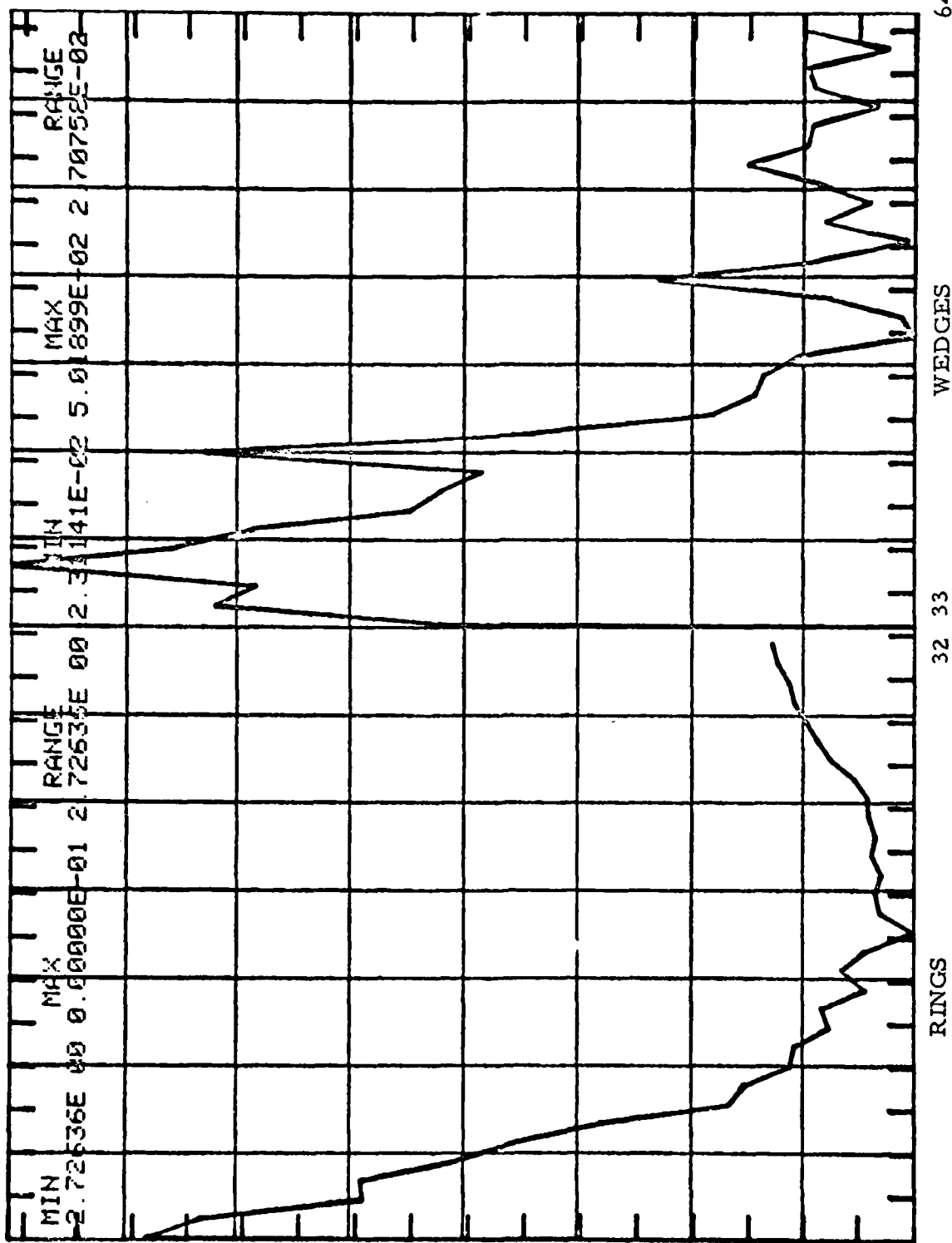
Angle: 0°

90°

180°

Figure 27.

FOURIER TRANSFORM SIGNATURE OF THE OPTICAL IMAGE OF THE CURVING ROAD REGION.



Sensor No. 1

Spatial

Frequency

(lp/mm): 0.0

1.21

8.0

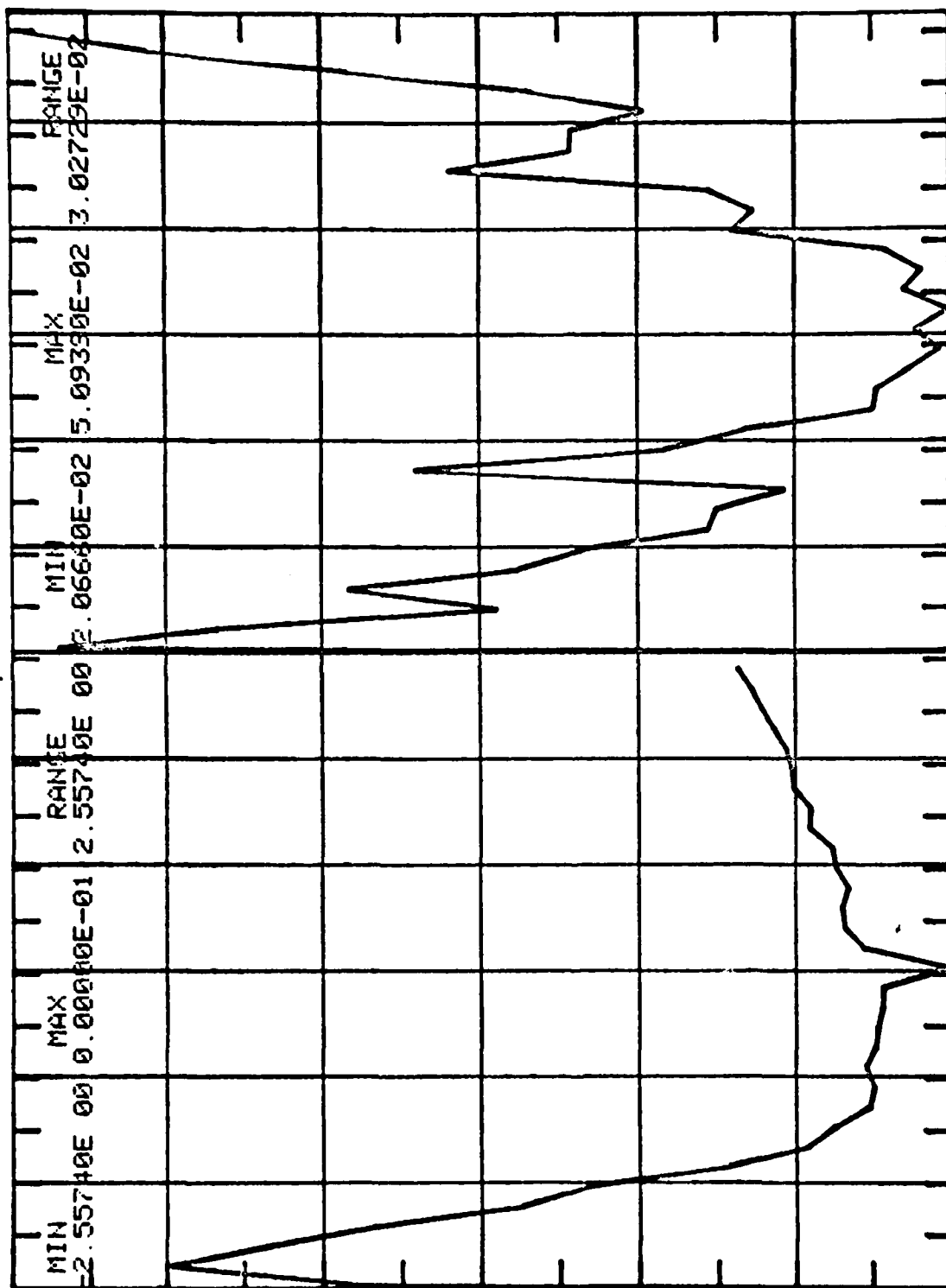
Angle: 0°

90°

180°

Figure 28.

FOURIER TRANSFORM SIGNATURE OF THE RADAR IMAGE OF THE CURVING ROAD REGION



Sensor No. 1

Spatial

Frequency

(lp/mm): 0.0

1.21

8.0

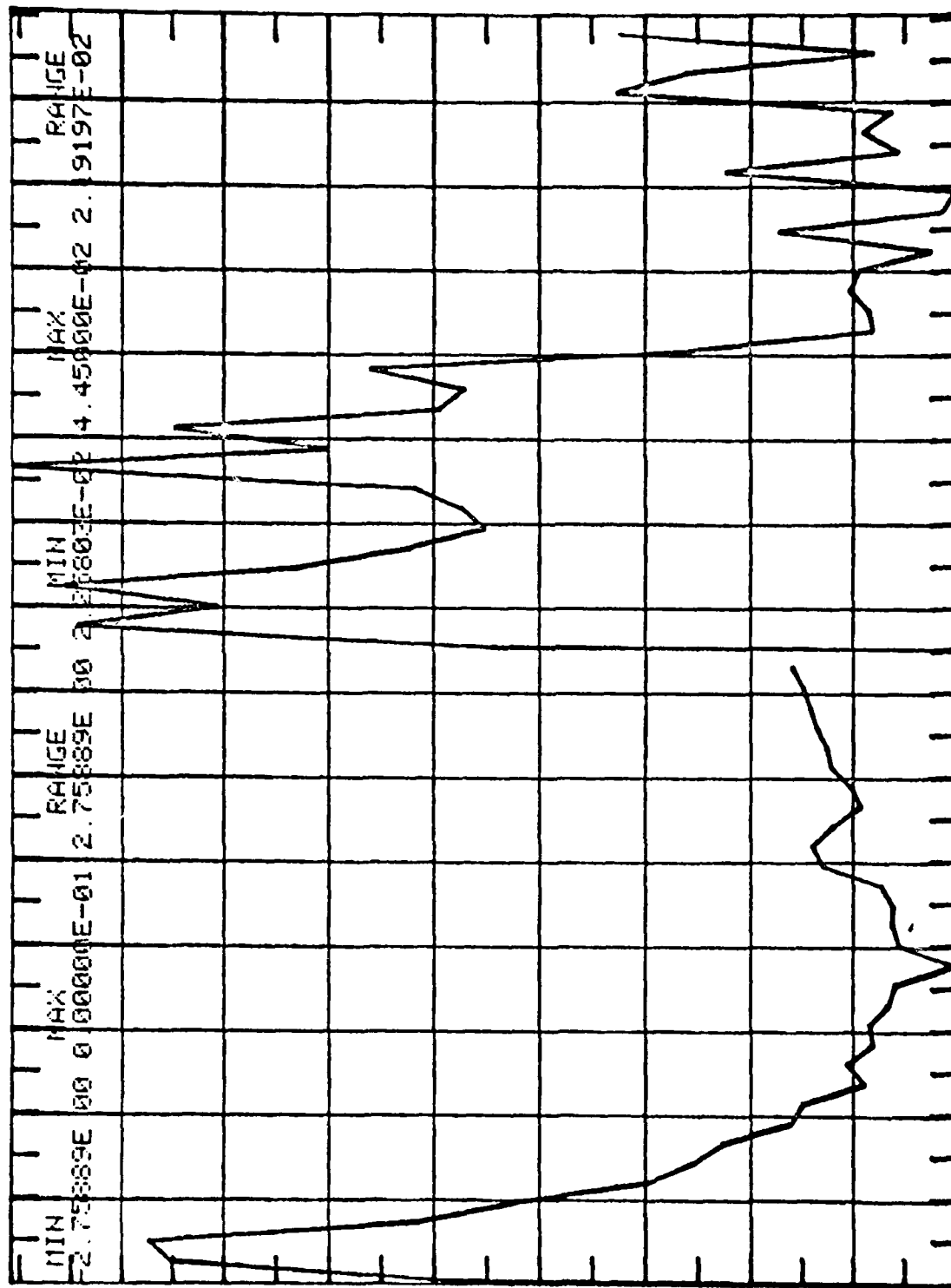
Angle: 0°

90°

180°

64

Figure 29.
FOUIER TRANSFORM OF OPTICAL IMAGE OF STADIUM REGION



Sensor No. 1

Spatial

Frequency

(lp/mm): 0.0

RINGS

1.21

32 33

8.0

WEDGES

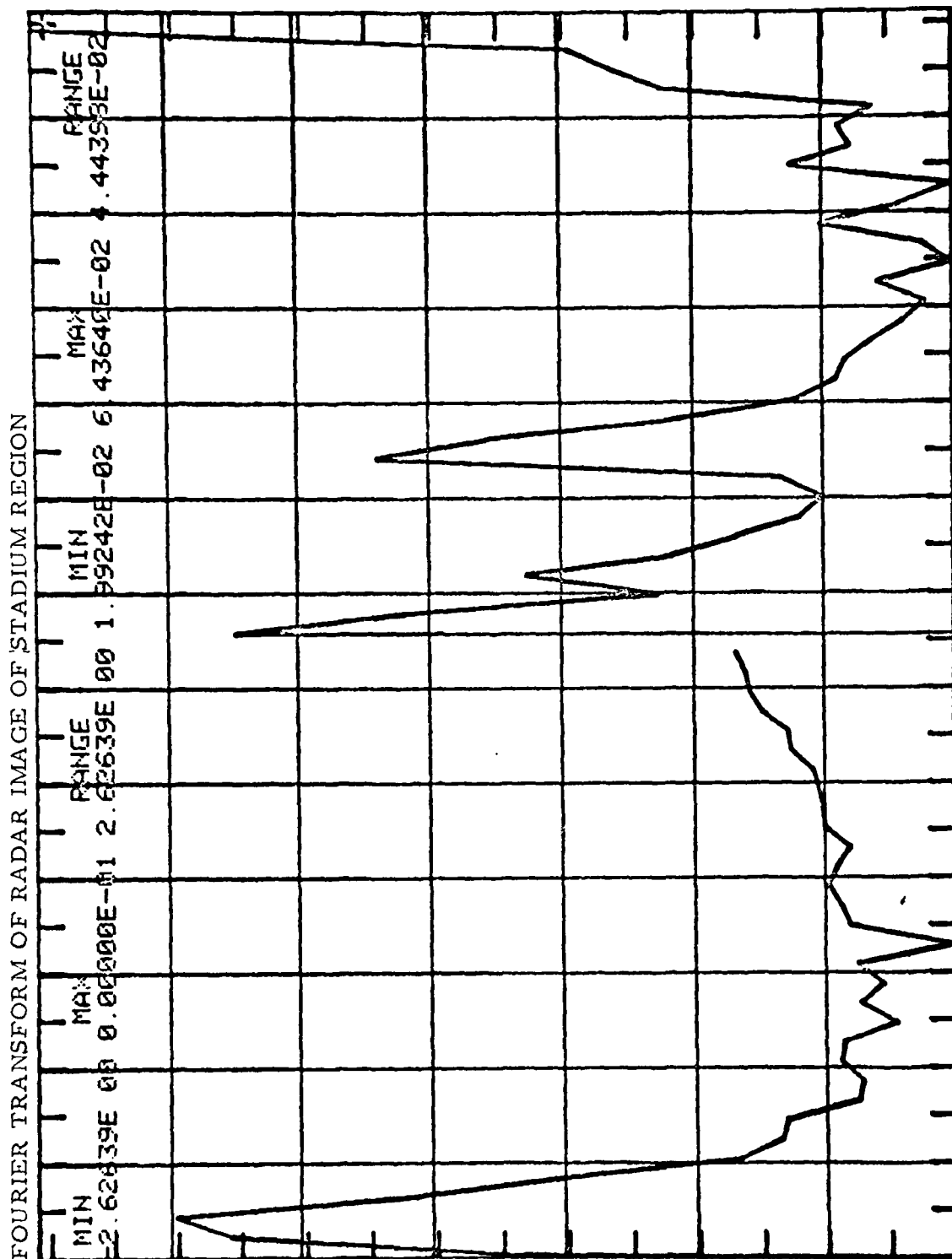
Angle: 0°

90°

180°

64

Figure 30.



Sensor No. 1

Spatial

Frequency

(lp/mm): 0.0

1.21

8.0

Angle: 0°

90°

180°

64

REFERENCES

1. P.E. Anuta, "Spatial Registration of Multispectral and Multitemporal Digital Imagery using Fast Fourier Techniques," IEEE Trans. Geosci. Electronics, Vol. GE-8, pp. 353-368, October 1970.
2. J.A. Leese, C.S. Novak, and B.B. Clark, "An Automatic Technique for Obtaining Cloud Motion from Geosynchronous Satellite Data using Cross-Correlation," J. Applied Meteorology Vol. 10, pp. 110-132, February 1971.
3. G.S. Garrett, "Detection Threshold Estimate for Digital Area Correlation," IEEE Trans. on Systems, Man, and Cybernetics, pp. 65-70, January 1976.
4. I.N. Durbaraw, "Optimal Data Processing for Evaluation of Correlation Guidance Techniques," AIAA Guidance and Control Conference, Stanford, California, August 1972.
5. G. Nagz, "Digital Image-Processing Activities in Remote Sensing for Earth Resources," Proceedings of the IEEE, Vol. 60, No. 10, pp. 1177-1199, October, 1972.
6. W.K. Pratt, "Correlation Techniques of Image Registration," IEEE Transaction on Aerospace and Electronic Systems, Vol. AES-10, No. 3, May 1974.
7. D.I. Barnea and H.E. Silverman, "A Class of Algorithms for Fast Digital Image Registration," IEEE Trans. on Computers, Vol. C-21, No. 2, February 1972.
8. H.K. Ramapriyan, "A Multilevel Approach to Sequential Detection of Pictorial Features," IEEE Transaction on Computers, pp. 66-78, January 1976.
9. R.L. Lillestrand, "Techniques for Change Detection," IEEE Transactions on Computers, Vol. C-21, No. 7, pp. 654-659, July, 1972.
10. M.S. Ulstad, "An Algorithm for Estimating Small Scale Differences between Two Digital Images," Pattern Recognition, Vol. 5, pp. 323-333, 1973.

SECTION III.

EDGE EXTRACTION AND REGISTRATION

3.1 Introduction

An important technique used in automatic scene analysis is image segmentation. This technique involves separating the image into various regions corresponding to individual objects. Assuming that these regions have some homogeneous characteristic, for example, luminance, color, texture, etc., one segmentation technique is to detect sharp transitions called edges, which tend to outline the desired boundaries. The opposite alternative is to "grow" regions by connecting small adjacent areas of similar characteristics. Of interest here is the detection of edges that separate regions of different constant luminances, and lines which can be regarded as a degenerate pair of edges. This operation requires the examination of several picture elements within contiguous or overlapping sub-areas of the image, followed by a decision as to whether an edge or the line segment is present or not within each sub-area.

The segments can be characterized by variable such as amplitude, orientation, position within the sub-area, etc., and possibly a measure of confidence.

Upon examination of the whole picture, the object boundaries are constructed by connecting the edge and line elements detected previously. This operation can be directed by simple syntactic rules, for example connect neighbor edge elements that line up approximately, and delete isolated parallel elements.

Some of the difficulties of edge detection are caused by noise, but much more so by the fact that visually distinct edges sometimes cannot be discriminated within a small image sub-area or, conversely that what appears to be an edge within the sub-area could belong to a homogeneously textured domain of the picture. Increasing the size of the sub-area apparently solves the problem, but is limited by computational costs and the complexity of large segment description.

It is generally recognized that boundary detection is therefore best done by the combination of a relatively simple edge and/or line segment detector, followed by algorithms that thin and link the segments obtained into continuous boundaries. Several fast numerical techniques for luminance edge extraction have been published, for example, Robert's "gradient" [1], Kirsch's [2], Sobel's [3], Prewitt's [4], Robinson's [5], and the so-called "smoothed gradient" [6] operators. Comparison of the above algorithms reveals similarities that suggest underlying general principles.

From these we develop a set of orthogonal functions which are closely related to distinctive image features. The properties of these functions suggest ways to minimize the amount of computations as well as an improved decision criterion. Considerable improvements are obtained in terms of boundary "thickness" and sensitivity to faint edges. The edge detection methods are described in Section 3.2.

The simplicity of edge data with its low computer memory requirement suggests several ultra-fast image registration methods through edge correlation. Theories of edge

correlation have been investigated and several experiments have been performed to test their applicability to real images. Section 3.3 describes edge registration as a potential method of image matching. Experiments were performed to register:

- Optical-to-optical images
- Radar-to-radar images
- Radar-to-optical images

The results of these experiments are discussed in Section 3.3.

3.2 Fast Edge Detection

3.2.1 Definitions

The problem of boundary element detection can be formulated as follows: given a set of n^2 luminance samples from an image sub-area, determine whether the sub-area contains a boundary element between two regions of different homogeneous luminances (edge). It may also be of interest to determine whether the area contains a line or a pair of degenerate edges enclosing an object too thin to be resolved. To this end, we define the following models of "ideal boundary elements".

Consider an image sub-area A of size $n \times n$ sampling intervals shown in Figure 31a. In the continuous image domain, we define an "ideal edge element" as a straight boundary, passing through the center of A , and which separates two regions of different, constant luminances b_1 and b_2 . Adopting the convention $b_1 > b_2$, the direction ϕ_e of the edge element is uniquely determined with respect to any arbitrary, fixed direction as shown in Figure 31b. The ideal edge element is characterized by its "magnitude" = $|b_1 - b_2|$ and orientation ϕ_e , $0 \leq \phi_e \leq 2\pi$.

Next we define an "ideal line element" (in the continuous image domain) as a straight strip of width approximately equal to one sampling interval, passing through the center of A , and of different luminance b_1 than its surrounding b_2 illustrated in Figure 31c. The ideal line element is characterized by its "magnitude" = $|b_1 - b_2|$, its orientation ϕ_ℓ , $0 \leq \phi_\ell \leq \pi$ and polarity sign $(b_1 - b_2)$.

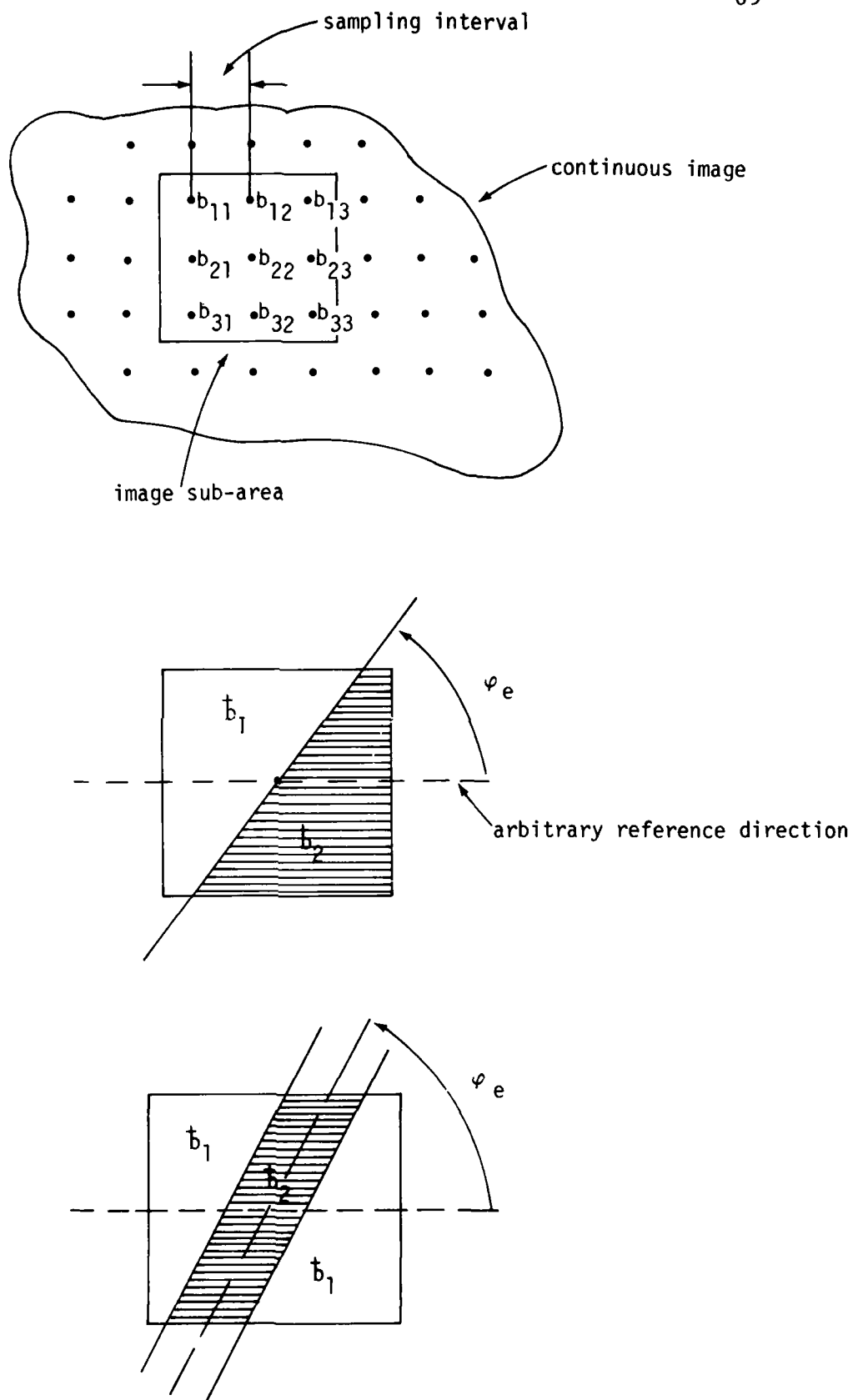


FIGURE 31.

Finally, an "ideal point" is defined as a point of brightness b_1 different than a constant brightness b_2 of the surround. The ideal point is characterized by its "magnitude" = $|b_1 - b_2|$ and polarity $\text{sign}(b_1 - b_2)$.

For the discrete case, we define the following notation. Consider the set of n^2 luminance samples b_{ij} of the image sub-area as an element of an n^2 -dimensional vector space B . The elements of B can be represented by a matrix B or a column vector $\underline{b}$, for example ($n=3$)

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}, \text{ or } \underline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n^2} \end{bmatrix} \quad (1)$$

Finally we define an inner- (or dot) product $(\cdot, \cdot)$ on B as

$$(B, C) = \sum_{i=1}^n \sum_{j=1}^n b_{ij} c_{ij} \text{ or } (\underline{b}, \underline{c}) = \sum_{i=1}^{n^2} b_i c_i. \quad (2)$$

3.2.2 Review of Previous Work

Previous fast edge detection algorithms [1-5] fall essentially into two categories:

a) Evaluate the maximum average gradient AG (or "smoothed gradient") present in each image sub-area. The average is estimated in a direction perpendicular to the (unknown) edge element orientation and the maximum is approximately obtained by

$$AG \approx [(B, W_1)^2 + (B, W_2)^2]^{\frac{1}{2}} \quad (3)$$

where B is the vector of luminance samples and W_1, W_2 are weighting functions shown in Figure 32 a, b, c[1,3,6]. When the average gradient exceeds an arbitrary threshold, the image sub-area is considered to contain an edge element. Orientation is then obtained approximately as

$$\varphi = \text{atan}[(B, W_1)/(B, W_2)]. \quad (4)$$

In order to simplify computations, the sum of squares of eq. 3 is sometimes replaced by a sum of absolute values. It is pointed out that the above measures are not isotropic, e.g. certain edge orientations are favored over other ones [5]. A pair of isotropic weighting functions is shown in Figure 32d.

b) The second approach is to form inner products of the luminance vector B with a set of discrete edge templates or

1	0
0	-1

0	1
-1	0

a) Roberts "gradient"

1	1	1
0	0	0
-1	-1	-1

1	0	-1
1	0	-1
1	0	-1

b) "smoothed gradient"

1	2	1
0	0	0
-1	-2	-1

1	0	-1
2	0	-2
1	0	-1

c) Sobel weighting functions

1	$\sqrt{2}$	1
0	0	0
-1	$-\sqrt{2}$	-1

1	0	-1
$\sqrt{2}$	0	$-\sqrt{2}$
1	0	-1

d) Isotropic weighting functions

FIGURE 32.

masks T , of different orientations as shown in Figures 33a, b, c, d[2], [4], [5]), and retain the largest value

$$\max \{ (B, T_i) \}.$$

When this value exceeds an arbitrary threshold, the sub-area B is considered to contain an edge element. The direction is approximately equal ($\pm\pi/4$) to the orientation of the template giving the largest inner product.

This second concept can be immediately extended to line and point detection with the template vectors shown in Figure 34a and b.

(a) Prewitt Masks	(b) Kirsch Masks	(c) Three-Level Simple Masks	(d) Five-Level Simple Masks																																				
(Robinson)																																							
<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>1</td><td>-2</td><td>1</td></tr><tr><td>-1</td><td>-1</td><td>-1</td></tr></table>	1	1	1	1	-2	1	-1	-1	-1	<table><tr><td>5</td><td>5</td><td>5</td></tr><tr><td>-3</td><td>0</td><td>-3</td></tr><tr><td>-3</td><td>-3</td><td>-3</td></tr></table>	5	5	5	-3	0	-3	-3	-3	-3	<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>-1</td><td>-1</td><td>-1</td></tr></table>	1	1	1	0	0	0	-1	-1	-1	<table><tr><td>1</td><td>2</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>-1</td><td>-2</td><td>-1</td></tr></table>	1	2	1	0	0	0	-1	-2	-1
1	1	1																																					
1	-2	1																																					
-1	-1	-1																																					
5	5	5																																					
-3	0	-3																																					
-3	-3	-3																																					
1	1	1																																					
0	0	0																																					
-1	-1	-1																																					
1	2	1																																					
0	0	0																																					
-1	-2	-1																																					
<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>1</td><td>-2</td><td>-1</td></tr><tr><td>1</td><td>-1</td><td>-1</td></tr></table>	1	1	1	1	-2	-1	1	-1	-1	<table><tr><td>5</td><td>5</td><td>-3</td></tr><tr><td>5</td><td>0</td><td>-3</td></tr><tr><td>-3</td><td>-3</td><td>-3</td></tr></table>	5	5	-3	5	0	-3	-3	-3	-3	<table><tr><td>1</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>-1</td></tr></table>	1	1	0	1	0	-1	0	-1	-1	<table><tr><td>2</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>-2</td></tr></table>	2	1	0	1	0	-1	0	-1	-2
1	1	1																																					
1	-2	-1																																					
1	-1	-1																																					
5	5	-3																																					
5	0	-3																																					
-3	-3	-3																																					
1	1	0																																					
1	0	-1																																					
0	-1	-1																																					
2	1	0																																					
1	0	-1																																					
0	-1	-2																																					
<table><tr><td>1</td><td>1</td><td>-1</td></tr><tr><td>1</td><td>-2</td><td>-1</td></tr><tr><td>1</td><td>1</td><td>-1</td></tr></table>	1	1	-1	1	-2	-1	1	1	-1	<table><tr><td>5</td><td>-3</td><td>-3</td></tr><tr><td>5</td><td>0</td><td>-3</td></tr><tr><td>5</td><td>-3</td><td>-3</td></tr></table>	5	-3	-3	5	0	-3	5	-3	-3	<table><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr></table>	1	0	-1	1	0	-1	1	0	-1	<table><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>2</td><td>0</td><td>-2</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr></table>	1	0	-1	2	0	-2	1	0	-1
1	1	-1																																					
1	-2	-1																																					
1	1	-1																																					
5	-3	-3																																					
5	0	-3																																					
5	-3	-3																																					
1	0	-1																																					
1	0	-1																																					
1	0	-1																																					
1	0	-1																																					
2	0	-2																																					
1	0	-1																																					
<table><tr><td>1</td><td>-1</td><td>-1</td></tr><tr><td>1</td><td>-2</td><td>-1</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table>	1	-1	-1	1	-2	-1	1	1	1	<table><tr><td>-3</td><td>-3</td><td>-3</td></tr><tr><td>5</td><td>0</td><td>-3</td></tr><tr><td>5</td><td>5</td><td>-3</td></tr></table>	-3	-3	-3	5	0	-3	5	5	-3	<table><tr><td>0</td><td>-1</td><td>-1</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>1</td><td>1</td><td>0</td></tr></table>	0	-1	-1	1	0	-1	1	1	0	<table><tr><td>0</td><td>-1</td><td>-2</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>2</td><td>1</td><td>0</td></tr></table>	0	-1	-2	1	0	-1	2	1	0
1	-1	-1																																					
1	-2	-1																																					
1	1	1																																					
-3	-3	-3																																					
5	0	-3																																					
5	5	-3																																					
0	-1	-1																																					
1	0	-1																																					
1	1	0																																					
0	-1	-2																																					
1	0	-1																																					
2	1	0																																					
<table><tr><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>1</td><td>-2</td><td>1</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table>	-1	-1	-1	1	-2	1	1	1	1	<table><tr><td>-3</td><td>-3</td><td>-3</td></tr><tr><td>-3</td><td>0</td><td>-3</td></tr><tr><td>5</td><td>5</td><td>5</td></tr></table>	-3	-3	-3	-3	0	-3	5	5	5	<table><tr><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table>	-1	-1	-1	0	0	0	1	1	1	<table><tr><td>-1</td><td>-2</td><td>-1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>2</td><td>1</td></tr></table>	-1	-2	-1	0	0	0	1	2	1
-1	-1	-1																																					
1	-2	1																																					
1	1	1																																					
-3	-3	-3																																					
-3	0	-3																																					
5	5	5																																					
-1	-1	-1																																					
0	0	0																																					
1	1	1																																					
-1	-2	-1																																					
0	0	0																																					
1	2	1																																					
<table><tr><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>-2</td><td>1</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table>	-1	-1	-1	-1	-2	1	1	1	1	<table><tr><td>-3</td><td>-3</td><td>-3</td></tr><tr><td>-3</td><td>0</td><td>5</td></tr><tr><td>-3</td><td>5</td><td>5</td></tr></table>	-3	-3	-3	-3	0	5	-3	5	5	<table><tr><td>-1</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>1</td></tr></table>	-1	-1	0	-1	0	1	0	1	1	<table><tr><td>-2</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>2</td></tr></table>	-2	-1	0	-1	0	1	0	1	2
-1	-1	-1																																					
-1	-2	1																																					
1	1	1																																					
-3	-3	-3																																					
-3	0	5																																					
-3	5	5																																					
-1	-1	0																																					
-1	0	1																																					
0	1	1																																					
-2	-1	0																																					
-1	0	1																																					
0	1	2																																					
<table><tr><td>-1</td><td>1</td><td>1</td></tr><tr><td>-1</td><td>-2</td><td>1</td></tr><tr><td>-1</td><td>1</td><td>1</td></tr></table>	-1	1	1	-1	-2	1	-1	1	1	<table><tr><td>-3</td><td>-3</td><td>5</td></tr><tr><td>-3</td><td>0</td><td>5</td></tr><tr><td>-3</td><td>-3</td><td>5</td></tr></table>	-3	-3	5	-3	0	5	-3	-3	5	<table><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr></table>	-1	0	1	-1	0	1	-1	0	1	<table><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-2</td><td>0</td><td>2</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr></table>	-1	0	1	-2	0	2	-1	0	1
-1	1	1																																					
-1	-2	1																																					
-1	1	1																																					
-3	-3	5																																					
-3	0	5																																					
-3	-3	5																																					
-1	0	1																																					
-1	0	1																																					
-1	0	1																																					
-1	0	1																																					
-2	0	2																																					
-1	0	1																																					
<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>-1</td><td>-2</td><td>1</td></tr><tr><td>-1</td><td>-1</td><td>1</td></tr></table>	1	1	1	-1	-2	1	-1	-1	1	<table><tr><td>-3</td><td>5</td><td>5</td></tr><tr><td>-3</td><td>0</td><td>5</td></tr><tr><td>-3</td><td>-3</td><td>-3</td></tr></table>	-3	5	5	-3	0	5	-3	-3	-3	<table><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-1</td><td>-1</td><td>0</td></tr></table>	0	1	1	-1	0	1	-1	-1	0	<table><tr><td>0</td><td>1</td><td>2</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-2</td><td>-1</td><td>0</td></tr></table>	0	1	2	-1	0	1	-2	-1	0
1	1	1																																					
-1	-2	1																																					
-1	-1	1																																					
-3	5	5																																					
-3	0	5																																					
-3	-3	-3																																					
0	1	1																																					
-1	0	1																																					
-1	-1	0																																					
0	1	2																																					
-1	0	1																																					
-2	-1	0																																					

FIGURE 33. "TEMPLATES" OR "MASKS" FOR EDGE DETECTION [4],[2],[5]

-1	-1	-1
2	2	2
-1	-1	-1

 T_1

-1	-1	2
-1	2	-1
2	-1	-1

 T_2

-1	2	-1
-1	2	-1
-1	2	-1

 T_3

2	-1	-1
-1	2	-1
-1	-1	2

 T_4

a) line "templates"

-1	-1	-1
-1	8	-1
-1	-1	-1

b) point "template"

FIGURE 34.

3.2.3 Derivation of Orthogonal Feature Basis

We now seek an appropriate basis for B . Because the templates of Fig. 33c and 34a represent samples of ideal edge and line elements positioned in eight equidistant orientations, we assume "edge" and "line" subspaces of B spanned by these vectors. Of all possible orthogonal bases for these subspaces, we choose the one shown in Fig. 35 because of the following properties: a) the first pair of basis vectors W_1 and W_2 represents the isotropic smoothed gradient weighting function. This pair, taken together with the second pair spans the above "edge" subspace. b) The second pair of basis vectors W_3 and W_4 has a distinctive higher order aspect (three zero crossings instead of one) and will be shown to contribute little to the magnitude of the edge subspace component. c) The "line" vectors were decomposed into pair of vectors W_5 , W_6 with directional preference and a pair W_7 , W_8 without directional preference. Note that the point basis vector of Fig. 34b is equal to the sum of the latter pair, which, incidentally, span all possible discrete realizations of the discrete Laplacian [7]. Finally, the vector W_9 was added to complete the basis. Observe that linear combinations of each pair of vectors produce similar distinctive patterns, which we call "average gradient," "ripple," "line," and "Laplacian" respectively.

Fig. 36 and 37 illustrate the above discussions. An original image of size 256 x 256 pixels was projected onto each one of the nine orthogonal basis vectors of Fig. 35.

Isotropic
"average
gradient"
subspace

1	$\sqrt{2}$	1
0	0	0
-1	$-\sqrt{2}$	-1

w_1

1	0	-1
$\sqrt{2}$	0	$-\sqrt{2}$
1	0	-1

w_2

Basis of edge
subspace

"Ripple"
subspace

0	-1	$\sqrt{2}$
1	0	-1
$-\sqrt{2}$	1	0

w_3

$\sqrt{2}$	-1	0
-1	0	1
0	1	$-\sqrt{2}$

w_4

"line"
subspace

0	1	0
-1	0	-1
0	1	0

w_5

-1	0	1
0	0	0
1	0	-1

w_6

Basis of line
subspace

"Discrete
Laplacian"
subspace

1	-2	1
-2	4	-2
1	-2	1

w_7

-2	1	-2
1	4	1
-2	1	-2

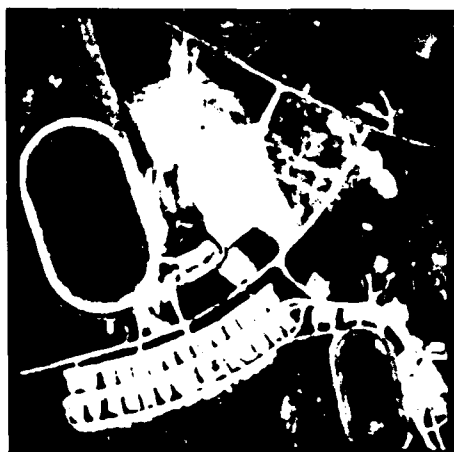
w_8

"Average"
subspace

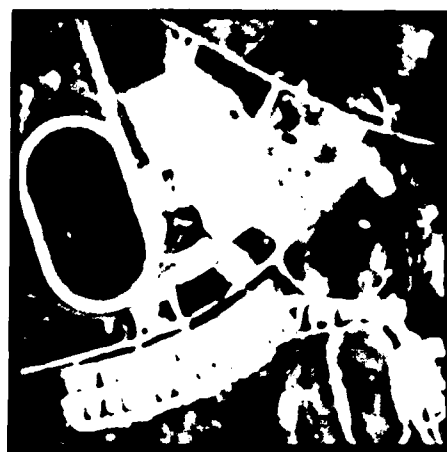
1	1	1
1	1	1
1	1	1

w_9

FIGURE 35.



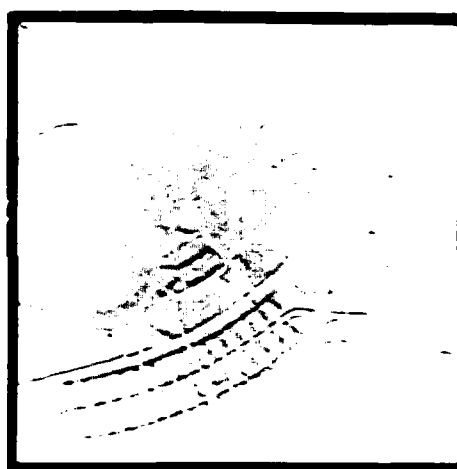
Original Images



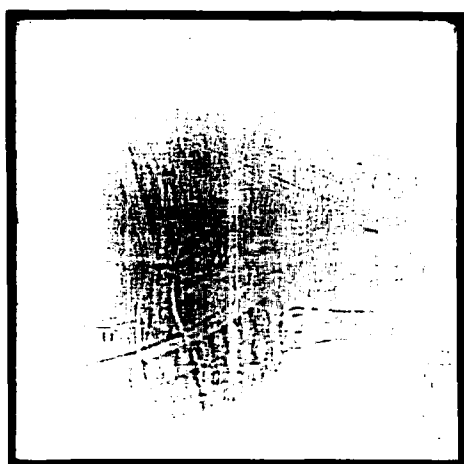
W_q



W_1 (edge)



W_2 (edge)



W_5 (line)



W_6 (line)

Figure 36.

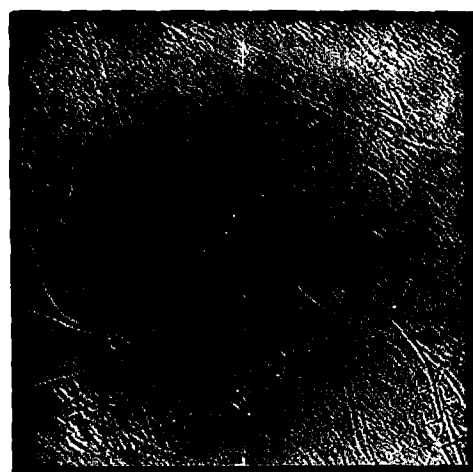
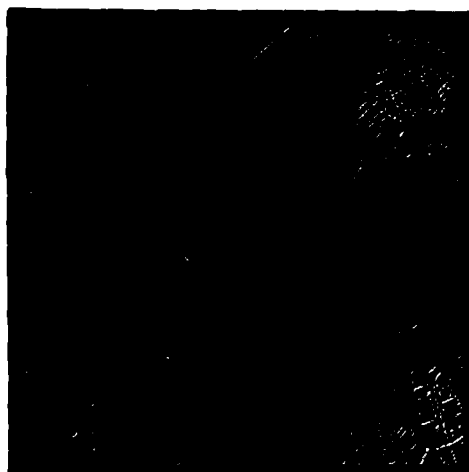
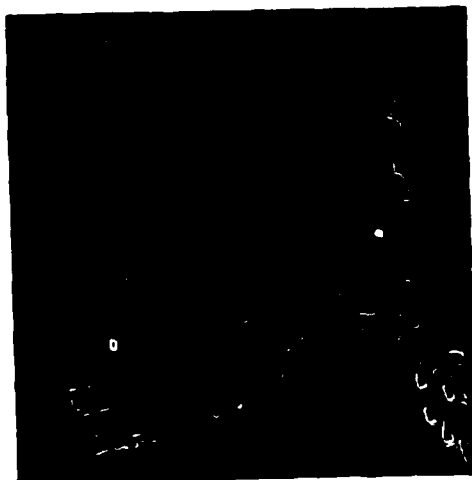
 W_7 (laplacian) W_8 (laplacian) W_3 (ripple) W_4 (ripple)

Figure 37.

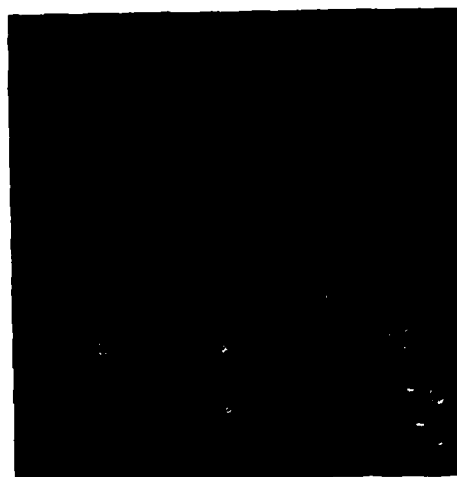
Since the projections are bipolar, a constant value was added for display, and the images were scaled for better display. The complementary nature of (edge, ripple) and (line, point) spaces is clearly visible and can be attributed to the fact that the former basis vectors are odd with respect to one axis of symmetry whereas the latter basis vectors are even (see Figures 33c and 34a).

In order to reduce computation, it would be desirable to reduce the dimension of the feature subspaces. Figure 38 shows the magnitudes of the projections onto "average gradient," "ripple," "line," and "point" respectively. While the last pair of projections appear similar (any line is composed of points!), the "average gradient" and "ripple" projections are quite different. It is easy to see that the "ripple" subspace adds little to the "edge" subspace and may be ignored to save computation.

An improved edge picture can be obtained by a simple use of the complementary property of projected components. If we subtract the edge magnitude from the sum of line and point magnitudes, the resultant picture will have diminishing values of those "edge points" with impure edgeness, making real edges sharp and obvious. Depending on the noise level of the original image an alternative edge decision criterion could be a relative edgeness thresholding in stead of absolute one. In this way faint edges can be detected with the same opportunities as those of strong edges. However, care must be taken as deciding faint edges because they are more subject to noise.



Magnitude in E-Subspace



Magnitude in L-Subspace



Magnitude in P-Subspace



Magnitude in R-Subspace

Figure 38.

Following this, the edge picture can even be improved by running the convolution (Fig. 34b) on it. This "point operation" effectively extracts peak values of the source array, making the edges even thinner and prominent. However, this follow-up process will raise the noise level somewhat and it has to be accompanied by noise cleaning. This is similar to peak detection in [9].

3.3 Edge Registration

3.3.1 Edge Correlation Methods

For the purpose of fast feature mapping the edge information of both the search region and searching windows is extracted. Correlation or similarity algorithms are then run on them. There are several advantages to edge feature registration instead of original scene magnitude registration. Firstly, edges are more likely to be sensor independent (optical vs. radar). Secondly, they are invariant to illumination change or even to a local illumination reversal. Thirdly, the storage of edge map requires up to 98% less storage, because of the low percentage of edge points among a picture in general. If bit operation is available the storage can even go down for several other orders. Finally, search computations can be reduced to logical functions instead of fixed point multiplications.

Let two images, A the search region and B the window be defined as $M \times N$ and $m \times n$ pixels respectively and both are represented by integers. The task is to find the sub-area of A which registers or best fits B (Appendix A). Either a sequential or random search strategy can be implemented. Besides this, there are several ways to define similarity and their computational requirements differ considerably[10].

The elements of unnormalized cross correlation surface $R(i, j)$ are defined as

$$R(i, j) = \sum_{k=1}^n \sum_{l=1}^m A(k, l) \cdot B(i+k, j+l) \quad (5)$$

A search for the maximum on the surface is then initiated. However, the maximum does not necessarily yield the registration point or yields a bad fix. Therefore, the correlation must be normalized as

$$R_z(i, j) = \frac{\sum_{k=1}^n \sum_{l=1}^m A(k, l) \cdot B(i+k, j+l)}{\left\{ \left[\sum_{k=1}^n \sum_{l=1}^m A^2(k, l) \right] \left[\sum_{k=1}^n \sum_{l=1}^m B^2(i+k, j+l) \right] \right\}^{\frac{1}{2}}} \quad (6)$$

The resultant fix is much better than the unnormalized correlation and the normalized correlation does give excellent results as to the self correlation of both optical and radar scenes. It is sensitive to background level and it requires extensive computations.

As a remedy, the following new similarity measures are proposed. Further details of which are described in Appendix B. These measures represent the edge pictures with binary numbers 0's and 1's for edge points and non-edge points respectively. Define a logical similarity measure as

$$R_l(i, j) = \frac{\sum_{k=1}^n \sum_{l=1}^m A(k, l) \cdot B(k+i, l+j)}{\left[\sum_{k=1}^n \sum_{l=1}^m A(k, l) \times \sum_{k=1}^n \sum_{l=1}^m B(k+i, l+j) \right]^{\frac{1}{2}}} \quad (7)$$

formula is for use in the case where only edge points are to be considered and can be further approximated as

$$R_{\text{AND}}(i, j) = \frac{1}{m \times n} \sum_{k=1}^m \sum_{l=1}^n A(k, l) \cdot B(k+i, l+j) \quad (8)$$

Alternatively, if all points are to be considered (edge and non-edge points) we define the NOT-EX-OR correlation measure as

$$R_{\text{EXOR}} = \frac{1}{m \times n} \sum_{k=1}^n \sum_{l=1}^m \frac{A(k, l) \text{ EX-OR } B(k+i, l+j)}{1} \quad (9)$$

Note that $R_{\ell}(i, j)$, $R_{\text{AND}}(i, j)$ and $R_{\overline{\text{EXOR}}}$ have values between 0 and 1; they are normalized quantities. A comparison of the different correlation computations is shown in Figure 39.

COMPARISON OF HARDWARE COMPLEXITY
AND SPEED FOR EXHAUSTIVE SEARCH EXAMPLE

NUMBER OF OPERATIONS	CONVENTIONAL NORMALIZED CROSS-CORRELATION	LOGICAL AND-CORRELATION MEASURE	LOGICAL NOT-EX-OR CORRELATION MEASURE
NUMERATOR			
(n, m) (N-n) (M-m)	integer multiply	1 bit AND	1 bit NOT-EX-OR
(n, m) (N-n) (M-m)	integer add	incremental count	incremental count
DENOMINATOR			
(n, m+1) (N-n) (M-m)	integer multiply	none	none
(n, m+1) (N-n) (M-m)	integer add	incremental count	none
(N-n) (M-m)	integer multiply	integer multiply	none
(N-n) (M-m)	floating square root	floating square root	none
(N-n) (M-m)	floating divide	floating divide	floating divide

Figure 39.

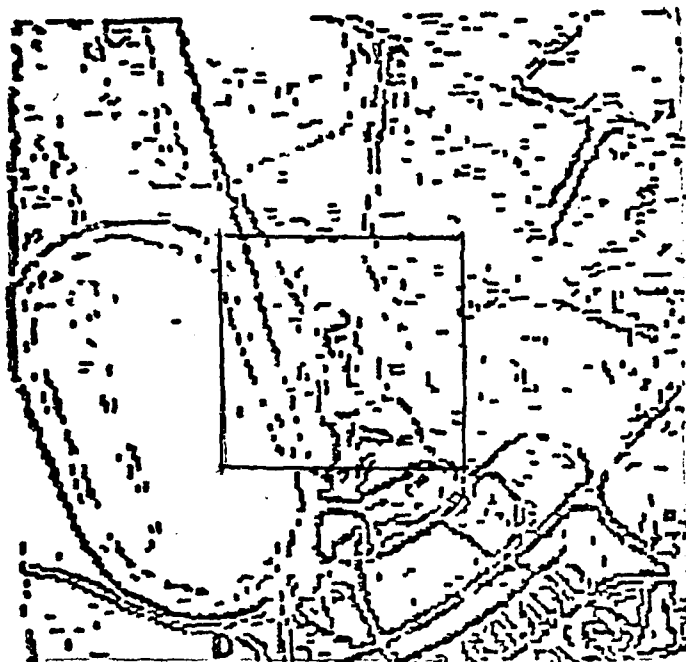
3.3.2 Experimental Results

A sequence of experiments were performed using the edge correlator method (Eq. (8)) to register:

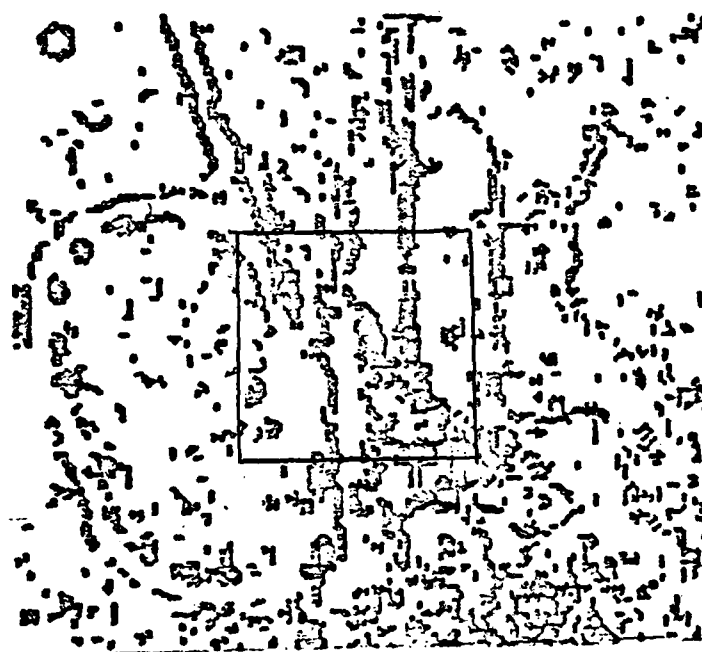
- Optical-to-optical images
- Radar-to-radar images
- Radar-to-optical images

Figure 40 shows the edge pictures extracted from the optical and radar images. The size of both images is 180 x 180. A window of size 64 x 64 located at the coordinates of (57, 57) was selected from the optical image as shown by solid lines in Figure 40(a). This window was correlated with the 180 x 180 optical image at every fourth location to produce a correlation map of size 30 x 30. The result is shown in Figure 41a. Similarly, the radar-to-radar correlation was performed and the results are in Figure 41b. Both figures show sharp correlation peaks located at the points of true registrations, thus providing a high confidence level that the correct matches have been found. These results indicated that when two images are relatively free from geometric and sensor distortions, the implementation of edge correlator using the fast algorithm with the AND operator as in equation (8) is especially attractive in real time computation. It reduces computations and storage requirements since the fixed point multiplications have been replaced by the simple AND operations.

To see how the edge correlator can be applied to radar-to-optical registration, two windows of different sizes and



a. Optical with window no. 1

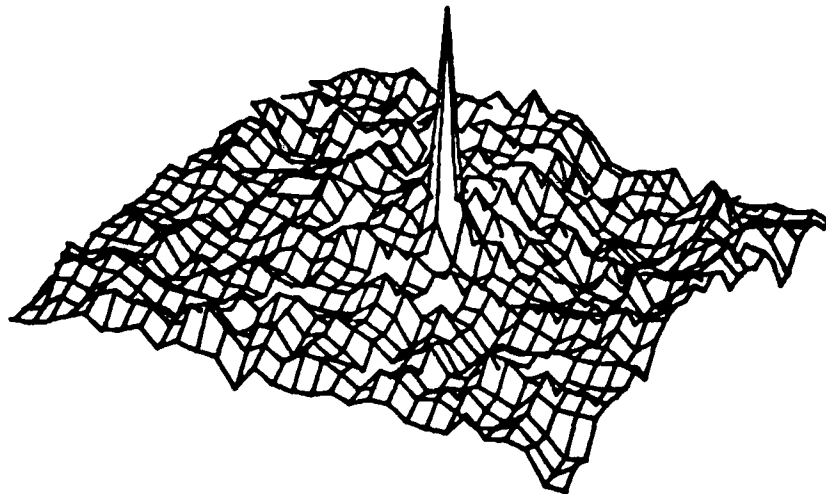


b. Radar

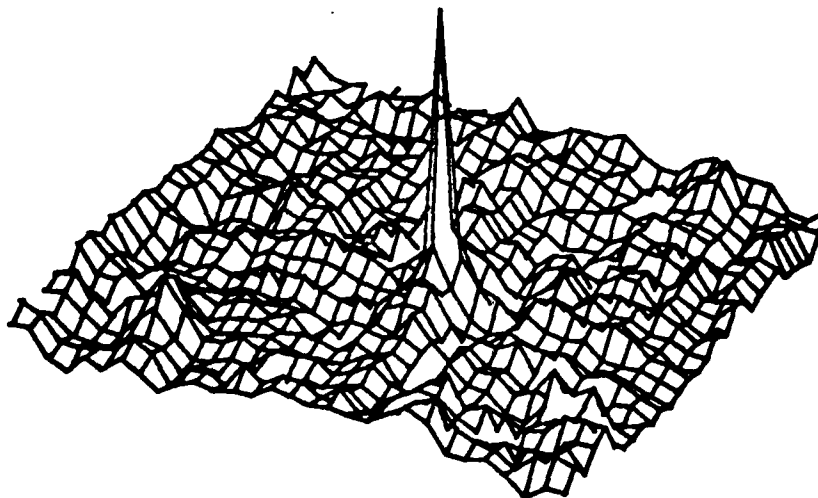
Figure 40. Optical and radar edge pictures (180 x 180)

Window: 64 x 64
Reference: 180 x 180
Window Coord.: (57, 57)
Pixels per Shift: 4
Algo.: Correlation

89



a. Optical-to-optical correlation



b. Radar-to-radar correlation

Figure 41.

locations were chosen from optical images. The two windows were then correlated with the radar image.

The window sizes and locations as shown in Figures 40 and 43 can be summarized as follows

<u>Window No.</u>	<u>Window Size</u>	<u>Location</u>	<u>Reference Size</u>
1	(64 x 64)	(57, 57)	(180 x 180)
2	(32 x 64)	(49, 21)	(90 x 122)

Window No. 1 was correlated with the reference image at every 4th location. Window No. 2 was correlated with the reference image at every 2nd location. Figure 42 shows the correlation of window No. 1. For clarity a threshold was selected so that only those points with relatively high correlation were shown. Similarly the correlations of window No. 2 are shown in Figure 44. In the case of window No. 1, although a relatively high peak was found to be located at the true registration location, other peaks of equal or higher amplitude were also obtained. This provided only a relatively low confidence in locating the true registration point. In the case of window No. 2, a high peak was obtained at a location which is four pixels away from the true registration location. An examination of the optical window and the reference radar map indicates that a slight geometric misregistration exists between the two images. Thus, the correct matching location was located more accurately by the automated technique.

AD-A086 208

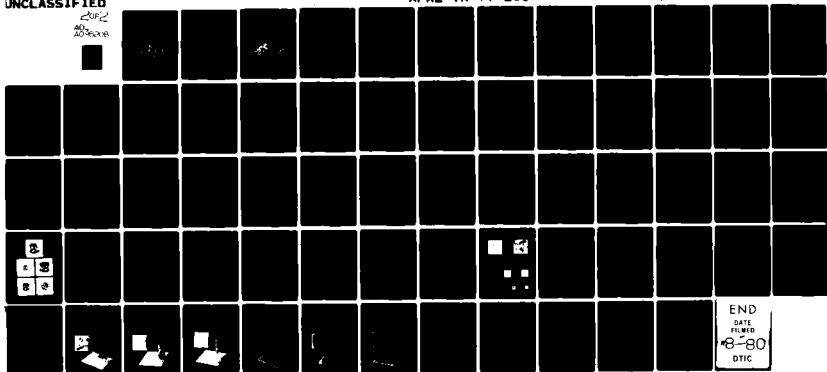
UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES IMAGE --ETC F/6 17/7
INVARIANT FEATURES FOR QUANTITATIVE SCENE ANALYSIS.(U)
JUN 76 E L HALL, W FREI F08606-72-C-0008

UNCLASSIFIED

AFAL-TR-77-253

NL

20p2
20p2



END
DATE
FILMED
8-80
DTIC

Window: Optical (64 x 64)
Reference: Radar (180 x 180)
Window Coord.: (57, 57)
Peak Coord.: (81, 69)
Peak Amplitude: 0.470
Pixels per Shift: 4

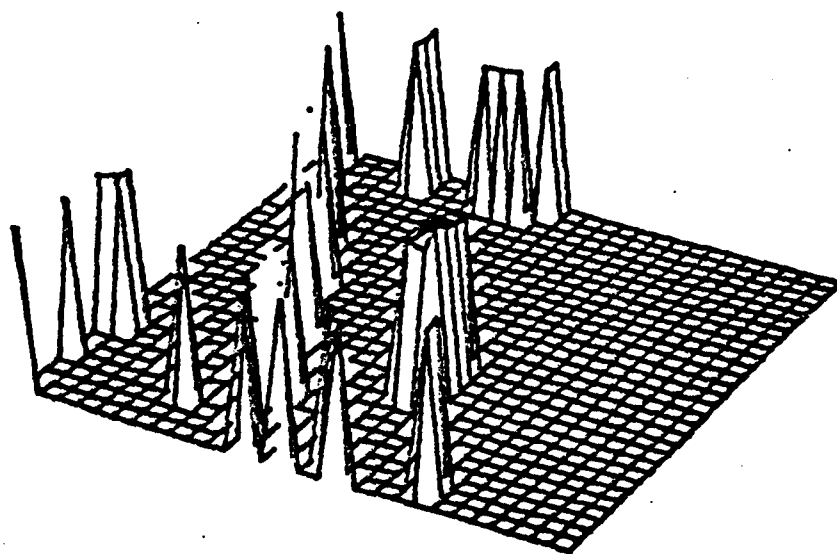
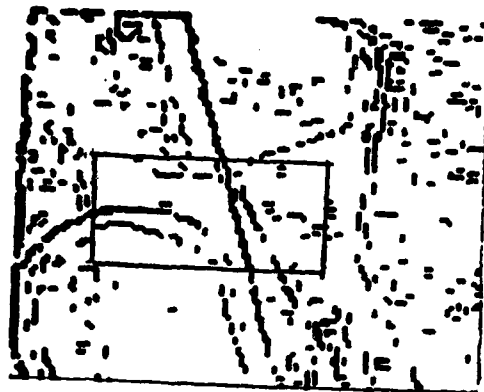


Figure 42. Radar to optical correlation thresholded at $T = 0.420$ to emphasize peak structure



a. Optical with window no. 2



b. Radar

Figure 43. Optical and radar edge pictures (90 x 122)

Window: Optical (32 x 64)
Reference: Radar (90, 122)
Window Coord.: (49, 21)
Peak Coord.: (53, 25)
Peak Amplitude: 0369
Pixels per Shift: 2

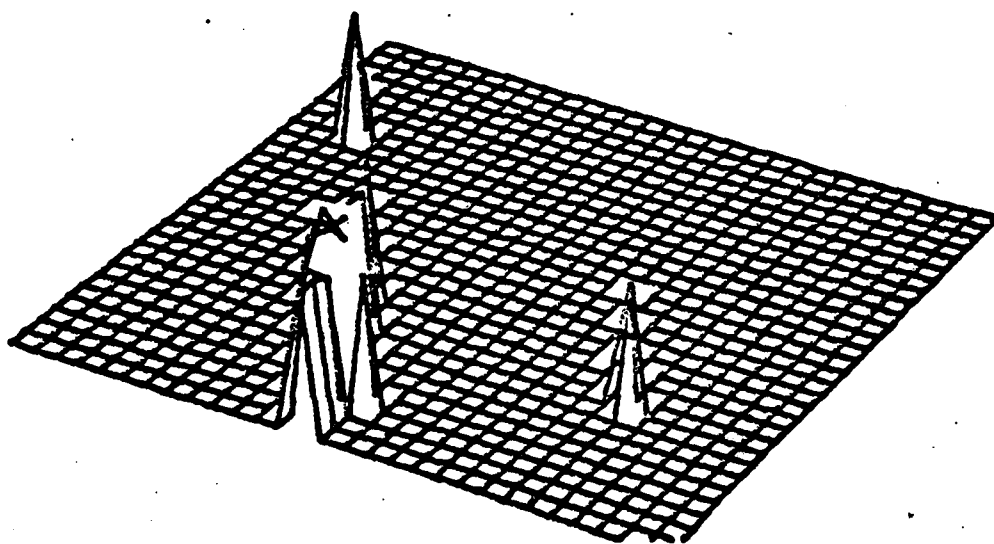


Figure 44. Radar to optical correlation thresholded at $T = 0.280$ to emphasize peak structure.

3.4 Conclusions

Several significant techniques on edge extraction and registration have been developed in this section. These are:

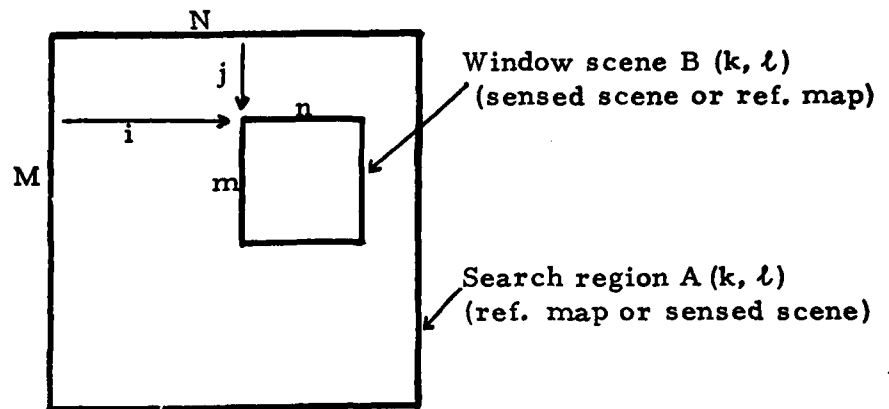
1. A set of edge operators extracted from the derived orthogonal basis were identified. These operators encompass nearly all of the previously derived operators by other researchers on a 3 by 3 image window. Therefore, the new edge operators can be applied to extract edges as well as other important image features such as points and lines.

2. A correlator with AND-operations was developed. This correlator is far more efficient than the existing correlators in terms of computational efficiencies and storage requirements. Theoretical and experimental results derived from this section indicate that edge feature can be used as a useful tool in image registration. These results indicate that the edge correlator is efficient and works very well on most regions of a scene in which only a few salient edges are present and no significant geometric or sensor distortion is present. Whenever, geometric or sensor distortions are present, the edge correlator performance is heavily dependent upon the scene content as well as the degree of distortion. Further research is needed to predict the degree of sensitivity to these distortions as well as on other edge detection operators and techniques.

REFERENCES

1. L.G. Roberts, "Machine Perception of Three-Dimensional Solids," in Optical and Electro-Optical Information Processing, J.T. Tippet, et al., eds., MIT Press, 1965.
2. R. Kirsch, "Computer Determination of the Constituent Structure of Biological Image," Computers and Biomedical Research, Vol. 4, 1971, pp. 315-328.
3. R. O. Duda and P.E. Hart, Pattern Classification and Analysis, Willey, New York, 1971.
4. J.M.S. Prewitt, "Object Enhancement and Extraction," in Picture Processing and Psychopictorics, B.S. Lipkin and A. Rosenfeld, eds., Academic Press, New York, 1970, pp. 75-149.
5. G.S. Robinson, "Detection and Coding of Edges using Directional Masks," USC-IPL Report No. 660, March 1976, pp. 40-57.
6. T.S. Huang, Purdue University, private communication, 1975.
7. W. Frei, "A New Class of Edge and Feature Detection Operators," USC-IPL Report No. 660, March 1976, pp. 22-40.
8. Z. Meiri, Bell Laboratories, private communication, January 1976.
9. A. Rosenfeld and M. Thurston, "Edge and Curve Detection for Visual Scene Analysis," IEEE Transactions on Computers, May 1971.
10. D. I. Barnea and H.F. Silverman, "A Class of Algorithms for Fast Digital Image Registration," IEEE Transactions on Computers, February 1972.

CONVENTIONAL SIMILARITY MEASURES FOR CORRELATION SEARCH



- Search region A is $M \cdot N$ pixels
- Window scene B is $m \cdot n$ pixels
- A and B are represented by integers

Cross-Correlation

$$R(i, j) = \sum_{k=1}^n \sum_{l=1}^m A(k, l) \cdot B(i+k, j+l)$$

Problems: $R(i, j)$ is a function of energy content of B within window at position (i, j) . Fix is very bad (see 2-D plot)

Normalized Cross-Correlation

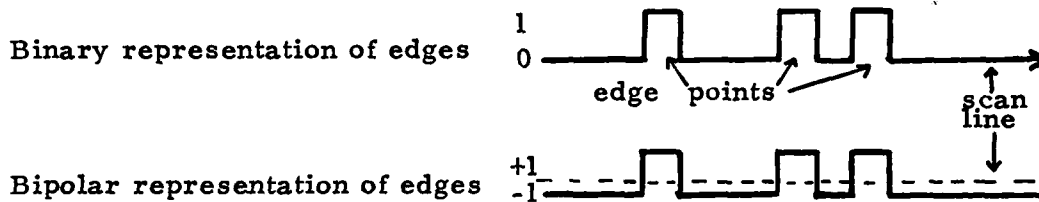
$$R_N(i, j) = \frac{\sum_{k=1}^n \sum_{l=1}^m A(k, l) B(i+k, j+l)}{\left[\sum_{k=1}^n \sum_{l=1}^m [A(k, l)]^2 \cdot \sum_{k=1}^n \sum_{l=1}^m [B(i+k, j+l)]^2 \right]^{\frac{1}{2}}}$$

Remarks: Fix is greatly improved (see 2-D plot)

THE LOGICAL NORMALIZED CROSS-CORRELATION MEASURE

A NEW SIMILARITY MEASURE FOR
CORRELATION SEARCH OF EDGE MAPS

Problem: The conventional cross-correlation or normalized cross-correlation measures cannot be used for binary representation; bipolar representation must be used to account for points where $A=0$, $B=1$ or $A=1$, $B=0$. Failure to do so would bring in only points where $A=B=1$.



This is achieved at great computational savings by the LOGICAL NORMALIZED CROSS-CORRELATION MEASURE

$$R_L(i, j) = \frac{\sum_{k=1}^n \sum_{l=1}^m A(k, l) \oplus B(k+i, l+j)}{\left[\sum_{k=1}^n \sum_{l=1}^m A(k, l) \cdot \sum_{k=1}^n \sum_{l=1}^m B(k+i, l+j) \right]^{+\frac{1}{2}}}$$

- The sums are simple incremental counters.
- The integer multiplications of the numerator are replaced by a 1 bit EX-OR function.
- The integer squares of the denominator are eliminated.
- Sequential similarity search can be implemented by randomizing the memory addresses corresponding to the indices $k, l, k+i, l+j$.

SECTION IV. INVARIANT MEASUREMENTS

4.1 INTRODUCTION

The theory of algebraic invariants from which the various facets of invariants developed is based upon a simple idea which first appeared in Lagrange's papers and was subsequently developed by Gauss [1,2].

This simple idea was: given a quadratic equation

$$ax^2 + 2bx + c = 0$$

lets us transform x to y by:

$$y = \frac{px+q}{rx+s} \quad \text{or} \quad x = \frac{q-sy}{ry-p}$$

then transformed equation will be

$$Ay^2 + 2By + C = 0 \text{ where}$$

$$A = as^2 + cr^2 - 2bsr$$

$$B = -aqs + b(qr+sp) - cpr$$

$$C = aq^2 - 2bqp - cp^2$$

then let us form the discriminant:

$$B^2 - AC = (ps-qr)^2(b^2 - ac).$$

Hence the discriminant of the equation in y is equal to the discriminant of the x equation, times the factor $(ps-qr)^2$ which depends only upon the coefficients p, q, r, s in the transformations $y = \frac{px+q}{rx+s}$.

Boole was the first to observe that for every equation the discriminant remains unchanged (except for a factor) if x is transformed to y by some transformations.

A uniform method which would give all the invariant expressions was given by Cayley in 1845 by his memoir, "On the Theory of Linear Transformations." Then Sylvester joined him in the construction of the general theory of algebraic invariants.

The development of invariant algebra continued through the first half of the present century by efforts of E. Elliot, Salmon, Grace and Young, et. al.

Algebraic invariant theory contributes the derivation of algebraic expressions which will remain invariant under certain transformations. The first use of algebraic invariants and moment invariant measurements for the purpose of picture recognition was advanced by Hu [2]. He used spatial moments as a means of representing characters and normalized them by developing a theorem relating moment measurements to algebraic invariants. Hall, et al. [3,4] has used moment invariants to describe both the spatial distributions of objects and edge structures of texture patterns for chest X-ray recognition.

A major problem encountered in the computer representation and measurement of a scene is the fact that by changing the viewing angle, scale, or rotation, most numerical measurements will change substantially. Therefore, it is desirable to consider normalization of the derived measurements in such a way that invariant properties be established.

Any discrete two dimensional scene may be represented by $f(x_i, y_j)$; $i = 1, 2, \dots, N$, $j = 1, \dots, M$. Certain measurements from this two dimensional array may be normalized by the utilization of algebraic invariant techniques.

The purpose of this section is to explore more deeply the use of moment invariants for the analysis and representation of complex scenes, present an analysis of a variance of the "invariant measurements" when made on computer images, and present pictorial examples for scene analysis.

4.1.2 QUANTICS AND INVARIANTS

Quantic

A function of several variables $x, y, z, \dots$ which is rational, integral and homogeneous is called a quantic.

If there are only two variables, such a function is called a binary quantic, and for a general q variable we call it q -ary quantic.

The order of a quantic is the degree of $x, y, \dots$.

The quantics of 1st, 2nd, 3rd, ..., p th are called linear, quadratic, cubic, ..., p -ics respectively. The binary p -ic will be denoted by the form: $a_0 x^p + p a_1 x^{p-1} y + \dots + \binom{p-r+1}{r} a_r x^{p-r+1} y^r$ which is usually shown in short as:

$$(a_0, a_1, a_2, \dots, a_p) (x, y)^p.$$

Note: When we talk about the coefficient of a binary quantic we mean $a_0, a_1, \dots, a_p$ and not $a_0, p a_1, \dots, \binom{p-r+1}{r} a_r, \dots$.

Invariants

An invariant of a single quantic is a function of the coefficients of that quantic which, after a linear transformation, remains constant except for a multiplicative factor which is a function only of the coefficients of the transformation.

Ex. Consider a binary quantic of order p :

$$Q = (a_0, a_1, \dots, a_p)(x, y)^p$$

in which we transform x, y by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} x^1 \\ y^1 \end{pmatrix}$$

then the same quantic in the transformed domain would be

$$Q = (a_0^1, a_1^1, \dots, a_p^1)(x^1, y^1)^p$$

where the values of $a_0^1, a_1^1, \dots, a_p^1$ are evaluated by the identity:

$$(a_0, a_1, \dots, a_p)(x, y)^p \equiv (a_0^1, a_1^1, \dots, a_p^1)(x^1, y^1)^p.$$

Now we say $f(a_0, \dots, a_p)$ is invariant under the above transformation if the following identity holds:

$$f(a_0^1, a_1^1, \dots, a_p^1) = \theta(l_{11}, l_{12}, l_{21}, l_{22})f(a_0, \dots, a_p)$$

where $\varnothing$ is a power of the determinant of the transformation:

$$\varnothing = \Delta^{\omega}$$

Δ being the determinant of the transformation, i.e.

$\Delta = l_{11}l_{22} - l_{12}l_{21}$ for the case of $\omega = 0$ we have absolute invariants.

Two sets of $(x, y, \dots)$ and $(x', y', \dots)$ are called co-gradient if a linear transformation on one set compells the other set to the same linear transformation.

Two sets $(x_1, x_2, \dots, x_p)$ and $(y_1, y_2, \dots, y_p)$ are called contragradient if we transform the first the first set by

$$\underline{x}' = L\underline{x}$$

where

$$L = \begin{bmatrix} l_{rj} \end{bmatrix} \quad \det L \neq 0$$

$$\underline{x}' = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_p \end{bmatrix}$$

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

then the transformation for the second set has to be

$$\underline{y}' = \underline{L}^{-T} \underline{y}$$

where

L^{-T} is the transpose of the inverse of L

Theorem: In contragradient quantities $\underline{x}^T \underline{y}$ remains invariant.

Proof: $\underline{x}' = L \underline{x}$

$$\underline{y}' = L^{-T} \underline{y}$$

$$\underline{x}'^T \underline{y}' = \underline{x}^T L^T L^{-T} \underline{y} = \underline{x}^T \underline{y}$$

4.1.4. Invariants of Orthogonal Transformations

Suppose we have a binary quantic of order p :

$$(a_{p0}, a_{p-1,1}, \dots, a_{0p})(x, y)^p$$

consider the transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Now we transform (x, y) and (x', y') as follows

$$\begin{pmatrix} U \\ V \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and

$$\begin{pmatrix} U' \\ V' \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

then we would have

$$U' = \frac{1}{2}(x' + iy')$$

$$V' = \frac{1}{2}(x' - iy')$$

$$U = \frac{1}{2}(x + iy)$$

$$V = \frac{1}{2}(x - iy).$$

Now substituting for x and y

$$\begin{aligned} U &= \frac{1}{2}[x'\cos\theta - y'\sin\theta] + \frac{1}{2}i[x'\sin\theta + y'\cos\theta] \\ &= \frac{1}{2}x'e^{i\theta} + \frac{1}{2}iy'e^{i\theta} = e^{i\theta}U' \end{aligned}$$

Hence

$$\boxed{U' = Ue^{-i\theta}}$$

$$\begin{aligned} V &= \frac{1}{2}(x - iy) = \frac{1}{2}[x'\cos\theta - y'\sin\theta] - \frac{i}{2}[x'\sin\theta + y'\cos\theta] \\ &= \frac{1}{2}x'e^{-i\theta} - \frac{i}{2}y'e^{-i\theta} = e^{-i\theta}V'. \end{aligned}$$

Hence

$$\boxed{V' = e^{i\theta}V}.$$

Then we will have the following identities

$$\begin{aligned} (I_{p0}, \dots, I_{op})(U, V)^p &\equiv (a_{p0}, \dots, a_{op})(x, y)^p \\ &\equiv (a'_{p0}, \dots, a'_{op})(x', y')^p \\ &\equiv (I'_{p0}, \dots, I'_{op})(Ue^{-i\theta}, Ve^{i\theta})^p. \end{aligned}$$

Then the invariants of which there are $p+1$ linearly independent ones, are obtained from the following equations.

$$\begin{aligned}
 I'_{p0} &= e^{ip\theta} I_{p0} \\
 &\vdots \\
 I'_{p-1,1} &= e^{i(p-2)\theta} I_{p-1,1} \\
 &\vdots \\
 I'_{1,p-1} &= e^{-i(p-2)\theta} I_{1,p-1} \\
 I'_{0p} &= e^{-ip\theta} I_{0p}.
 \end{aligned}$$

From the first two identities we have

$$\begin{aligned}
 I_{p0} &= a_{p0} - i \binom{p}{1} a_{p-1,1} - \binom{p}{2} a_{p-2,2} \\
 &\quad + i \binom{p}{3} a_{p-3,3} + \dots + (-i)^p a_{0p}
 \end{aligned}$$

$$\begin{aligned}
 I_{p-1,1} &= (a_{p0} + a_{p-2,2}) - i(p-2)(a_{p-1,1} + a_{p-3,3}) \\
 &\quad + \dots + (-i)^{p-2}(a_{2,p-2} + a_{0p})
 \end{aligned}$$

$$\begin{aligned}
 I_{p-2,2} &= (a_{p0} + 2a_{p-2,2} + a_{p-4,4}) - i(p-4)(a_{p-1,1} + 2a_{p-3,3} + a_{p-5,5}) \\
 &\quad + \dots + (-i)^{p-4}(a_{4,p-4} + 2a_{2,p-2} + a_{0p})
 \end{aligned}$$

$$\begin{aligned}
 I_{p-r,r} &= [(a_{p0}; a_{p-2,2}; \dots; a_{p-2r,2r})(1,1)^r; \\
 &\quad (a_{p-1,1}; a_{p-3,3}; \dots; a_{p-2r-1,2r+1})(1,1)^r; \dots \\
 &\quad (a_{2r,p-2r}; a_{2r+2,p-2r-2}; \dots; a_{0p})(1,1)^r](1,-i)^{p-2r} p_{p-2r} > 0
 \end{aligned}$$

and

$$I_{p/2, p/2} = a_{p0} + \binom{p/2}{1} a_{p-2, 2} + \binom{p/2}{2} a_{p-4, 4} + \dots + a_{0p}$$

$p = \text{even}.$

4.2 Perceptual Invariants

In a discussion of shape recognition, Deutsch[3] described four invariance properties of human perception.

1. Shapes may be recognized independent of their location in the visual field. It is not necessary to fixate on the center of a figure in order to recognize it nor need the eyes be moved around the contours of a figure. Thus, shape recognition is invariant to translation provided the object remains in the visual field.

2. Recognition can be effected independently of the angle of inclination of a figure in the visual field. The angle refers to the angular orientation of the figure in two dimensions although some invariance to the depth angle is also observed. Therefore, the recognition of a figure is invariant to rotation in two dimensions, although the angle of rotation may also be recognized.

3. The size of a figure does not interfere with the recognition of its shape provided the entire figure is within the visual field. Thus, visual recognition exhibits an invariance to scale.

4. Mirror images appear alike. Both humans and animals tend to confuse these. This observation rules out a template matching theory of shape recognition because a template superposition cannot

take place for mirror images and thus the confusion cannot be predicted.

Deutsch also points out that certain primitive organisms find it more difficult to distinguish between squares and circles than between rectangles and squares, which casts some doubt on recognition theories based on angular properties of figures. He also argues that the invariance properties are common to all parts of the visual cortex, which rules out a scanning process for shape recognition. To illustrate that the invariance properties could be realized easily, he proposed the following system.

The receptor surface is connected to a two dimensional coding array which is excited by a contour falling on the corresponding receptors. For example, if a spherical object is focused on the retina, the circular contour is projected on the coding array. The processing of the edge contour is described by the following four rules.

- (i) Every figure excites the retina only by its contour i.e., only the contour of the figure will appear on the two dimensional array.
- (ii) Every array point will transmit a pulse down the common cable as soon as an excitation reaches it from the retina. The integration of these pulses gives a measure of the total number of edge points.
- (iii) Each contour point on the array will also excite all the units which are at right angles to it on the array. These excitations will not pass to the common cable but only to the

neighboring cells. It is also assumed that these lateral excitations move at the same speed in all directions.

- (iv) As such lateral excitation from a point in a contour advances another message will be sent down the common cable as soon as it coincides with another contour point.

The recognition of shape is made from the signals on the common cable. For each pattern, two sets of pulses, separate in time, are transmitted. The first set relates to the number of contour points. The second provides shape information. This system can be shown to be invariant to translation, rotation, scale and mirror image as previously described. The invariance to translation, rotation and mirror image transformations would affect neither the number of contour points nor the distance between edges in the figure. The size invariance may be obtained by normalizing the second signal by the first. That is, the lateral excitation signal would be used to normalize the occurrence times of the second signal.

It is interesting to note that the proposed system would be more easily confused in differentiating circles and squares than rectangles and squares.

Both circles and squares would produce two sharp pulses which could be of equal magnitude if the same number of edge points were detected. However, the rectangle would produce three pulses and could easily be distinguished from

either a square or circle.

The limitations of this simple model are many; however, it does illustrate a logical approach to recognition. First, the desired recognition properties are determined, then a system which satisfies these properties is designed. The particular system described was devoted to shape recognition. Other systems need to be developed which are invariant to intensity changes and more complex images. The power and simplicity of Deutsch's approach was a major contribution from psychology to the understanding of recognition systems.

4.3.1 Moment Invariants

Any two-dimensional pattern can be represented by a picture function, $f(x,y)$, with respect to a pair of axes fixed in the visual field. It is also true that by means of a uniqueness theorem[3], two dimensional moments, m_{pq} , which are defined as:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy$$

$$p, q = 0, 1, 2, \dots$$

can be used to uniquely represent $f(x,y)$. Hence one may use m_{pq} as a means of representing any two-dimensional pattern.

The computations of m_{pq} consist of multiplying the functions $f(x,y)$ by a monomial $x^p y^q$ and integrating the result. The monomials of order 3 or less, $x^0 y^0$, $x^0 y^1$, $x^0 y^2$, $x^1 y^0$, $x^1 y^1$, $x^2 y^0$, $x^2 y^1$, $x^3 y^0$, and $x^1 y^2$ are shown in Figure 45.

The moments of order $p+q$ may also be interpreted as the response of an imaging system with the transfer function of, $x^p y^q$, and the input, $f(x,y)$. Low order moments have intuitive relations to objects. For example, m_{00} is related to mass m_{10} and m_{01} to center of mass and m_{11} , m_{02} , and m_{20} to principal directions.

Another useful concept is the moment generating functions which is defined as:

$$M(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[ux + vy] f(x, y) dx dy .$$

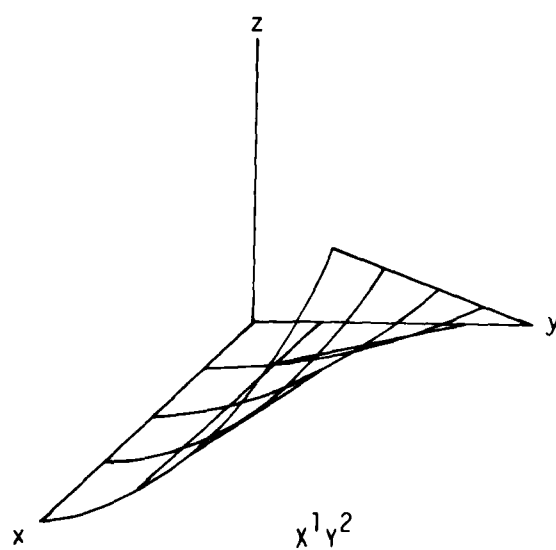
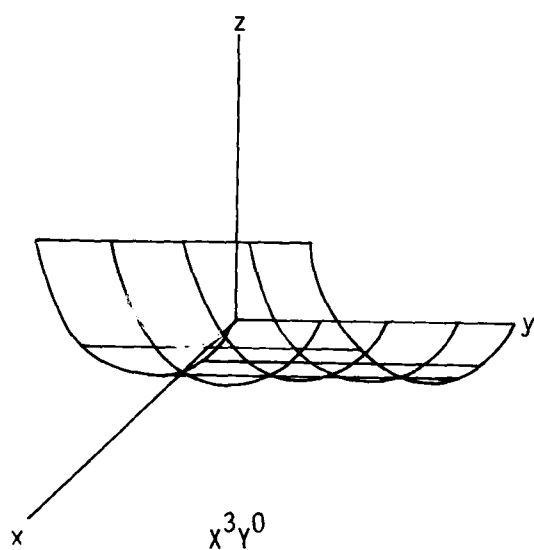
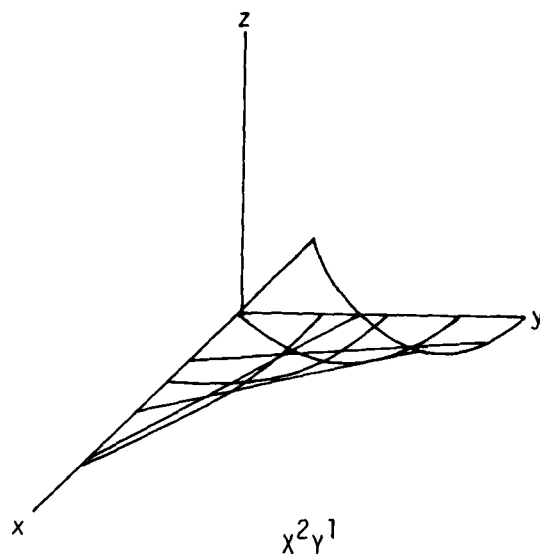
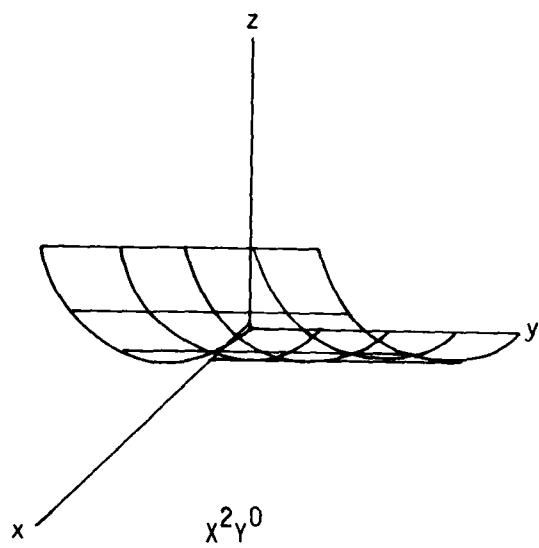


FIGURE 46. CONTINUED

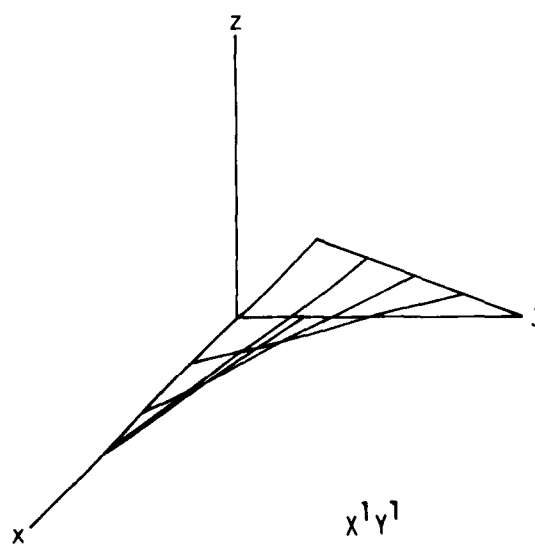
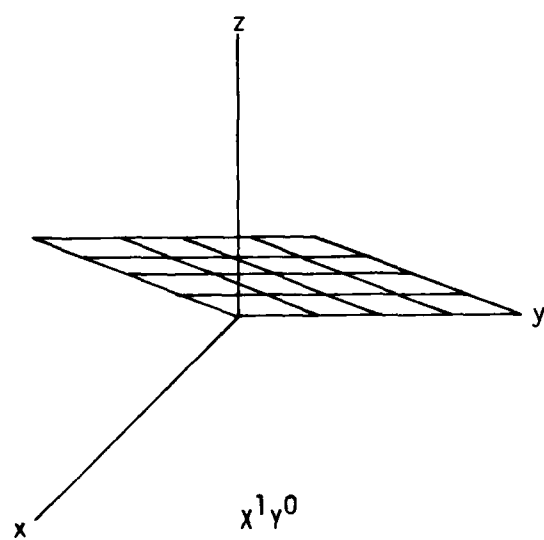
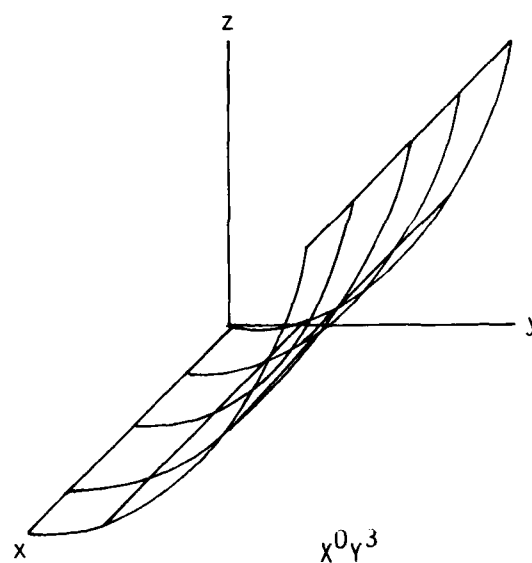
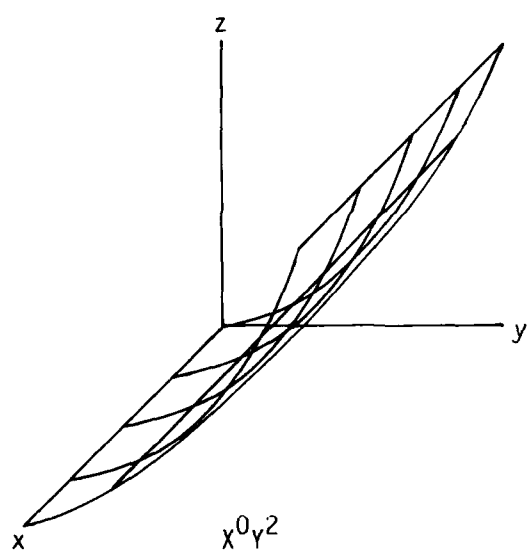
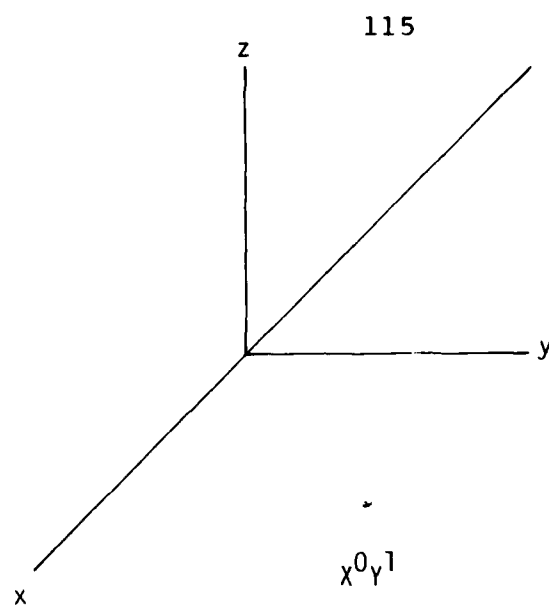
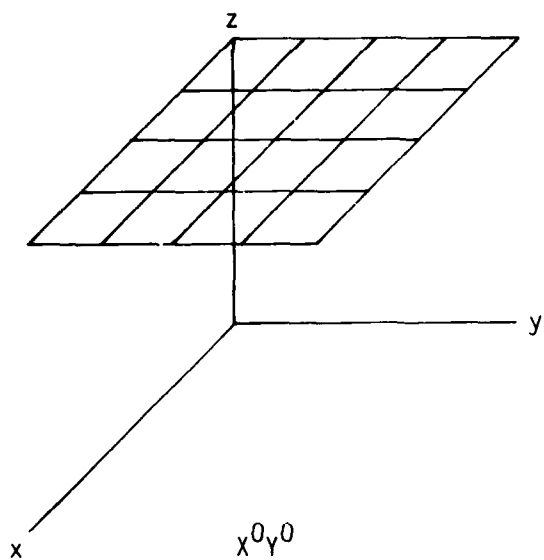


FIGURE 46.

If μ and ν are considered as complex variables, this expression is a two sided Laplace transform. For the invariant development both μ and ν are assumed to be real. This function can also be written in the form of

$$M(\mu, \nu) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} m_{pq} \frac{(\mu)^p}{p!} \frac{(\nu)^q}{q!}$$

where the exponential has been expanded in its Taylor series equivalent, assuming that moments of all orders exist. This equation shows that the moments may be determined from the derivatives of the moment generating functions evaluated at the origin.

Central moments are defined as

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\bar{x})^p (y-\bar{y})^q \rho(x, y) dx dy$$

where

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}.$$

Central moments may be easily shown to be invariant under translation and from here for the sake of simplicity we will assume to be using central moments.

To relate the moments to algebraic invariants, one may first expand the exponential term in the moment generating function to obtain

$$M(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=0}^{\infty} \frac{1}{p!} (ux + vy)^p \rho(x, y) dx dy.$$

Now after using the binomial expansion and carrying out the integration:

$$M(u, v) = \sum_{p=0}^{\infty} \frac{1}{p!} \mu_{p-r, r} \binom{p}{r} \binom{p-r}{u} \binom{r}{v}.$$

The functions of the form

$$\sum_{r=0}^p k_{p-r, r} \binom{p}{r} \binom{p-r}{u} \binom{r}{v}$$

are called qualtics in the study of algebraic invariants[1] and play a major role in the computation of moment invariants. The fundamental definitions of this theory are described in Section 4.2.1.

4.3.2 Fundamental Theorem of Moment Invariants

If the binary p-ic has an invariant

$$f(a'_{p0}, \dots, a'_{0p}) = \Delta^w f(a_{p0}, \dots, a_{0p})$$

then the moment of order p has an algebraic invariant

$$f(\mu'_{p0}, \dots, \mu'_{0p}) = |J| \Delta^w f(\mu_{p0}, \dots, \mu_{0p})$$

where J is the Jacobian of the transformation.

The importance of this theorem lies in the fact one may find an invariant function of moments once we have a corresponding algebraic invariant function.

A point which should be emphasized is the generality involved in linear transformations. The only restriction was $\Delta \neq 0$. Hence we can use rotation, reflection, magnitude change, and correspondingly make our normalization with respect to magnitude, rotation, orientation, etc.

Example: Similitude Moment Invariants

Consider the transformation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \alpha \text{ constant.}$$

we can write an algebraic invariant simply as

$$a'_{pq} = \alpha^{p+q} a_{pq}$$

where $\alpha \beta$ is the determinant. By using the fundamental theorem the moment invariant is:

$$u'_{pq} = \alpha^{p+q+2} u_{pq}. \quad (1)$$

Since $J = \alpha^2$.

For zero order ($p+q = 0$) we have:

$$\mu' = \alpha^2 \mu \quad (2)$$

now by eliminating α between (2) and (1) we have the following absolute moment invariants:

$$\frac{\mu'_{pq}}{\frac{p+q}{2}+1} = \frac{\mu_{pq}}{\frac{p+q}{2}+1}, \quad p+q = 2, 3, \dots$$

$$\mu'_{10} = \mu'_{01} = 0$$

4.3.3 Derivation of Moment Invariants

In the following we attempt to use the concept and tools of invariant algebra which is discussed in Section 4.1 to derive the seven invariant moments which are simultaneously invariant with respect to size, orientation and location of the object in the scene. The detail theory behind these derivations is given in Section 4.1. So in the rest of this sub section we proceed to use those tools in calculating the invariant moments.

$$\begin{aligned}\mu_{10} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^0 f(x, y) \\ &= m_{10} - \frac{m_{10}}{m_{00}} (m_{00}) \\ &= 0\end{aligned}$$

$$\begin{aligned}\mu_{11} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^1 f(x, y) \\ &= m_{11} - \frac{m_{10} m_{01}}{m_{00}}\end{aligned}$$

$$\begin{aligned}\mu_{20} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^0 f(x, y) \\ &= m_{20} - \frac{2m_{10}^2}{m_{00}} + \frac{m_{10}^2}{m_{00}} = m_{20} - \frac{m_{10}^2}{m_{00}}\end{aligned}$$

$$\begin{aligned}\mu_{02} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^2 f(x, y) \\ &= m_{02} - \frac{m_{01}^2}{m_{00}}\end{aligned}$$

$$\begin{aligned}\mu_{30} &= \sum_x \sum_y (x-\bar{x})^3 (y-\bar{y})^0 f(x, y) \\ &= m_{30} - 3\bar{x}m_{20} + 2m_{10}\bar{x}^2\end{aligned}$$

$$\begin{aligned}\mu_{12} &= \sum_x \sum_y (x-\bar{x})(y-\bar{y})^2 f(x, y) \\ &= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2 m_{10}\end{aligned}$$

$$\begin{aligned}\mu_{21} &= \sum_x \sum_y (x-\bar{x})^2 (y-\bar{y})^1 f(x, y) \\ &\vdots \\ &= m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2 m_{01}\end{aligned}$$

$$\begin{aligned}\mu_{03} &= \sum_x \sum_y (x-\bar{x})(y-\bar{y})^3 f(x, y) \\ &= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^3 m_{01}.\end{aligned}$$

In summation:

$$\begin{array}{ll}\mu_{00} = m_{00}, & \mu_{11} = m_{11} - \bar{y}m_{10} \\ \mu_{10} = 0, & \mu_{30} = m_{30} - 3\bar{x}m_{20} + 2m_{10}\bar{x}^2 \\ \mu_{01} = 0, & \mu_{12} = m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2 m_{10} \\ \mu_{20} = m_{20} - \bar{x}m_{10} & \mu_{21} = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2 m_{01} \\ \mu_{02} = m_{02} - \bar{y}m_{01} & \mu_{03} = m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2 m_{01}.\end{array}$$

The normalized central moments, denoted by η_{pq} are defined as:

$$\eta_{pq} = \frac{\mu_{pq}}{\left(\frac{p+q}{2} + 1 \right) \mu_{00}}$$

$$p+q = 2, 3, \dots$$

The normalized central moments are invariant to size changes as well as translation.

From the second and third order moments a last set of measurements can be derived which are invariant to proper and improper orthogonal rotation, and mirror images as well as invariant to translation, size change.

$$\varphi_1 = \eta_{20} + \eta_{02}$$

$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} + \eta_{03})^2$$

$$\varphi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\begin{aligned} \varphi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

$$\begin{aligned} \varphi_6 = & (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ & + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}). \end{aligned}$$

A seventh "metric" can be added which will change sign under improper rotation.

$$\begin{aligned} \varphi_7 = & (3\eta_{12} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]. \end{aligned}$$

4.3.4 Numerical Computation of Moment Invariants

In the previous section, the moments were computed for a continuous function of continuous variables. For computer implementation, the moments must be computed for a discrete function of discrete variables. Also, the range of the moments must be considered. In this section, the computer representation and numerical computation of moment invariants will be described.

Any finite set of moments may be computed and stored as a matrix. For example, the matrix, M , of non-central moments with p, q less than or equal to 3 is given by:

$$M = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix}$$

If an order, $n = p+1$, is selected, and all moments with n less than or equal to fixed value is selected, then an upper triangle matrix is produced. This condition is illustrated

by the elements above the line in the matrix, M . Similar matrices may be defined for central and normalized moments.

An important question for the computer implementation of moment invariants is the range of the moments since this determines the computer storage requirements. The maximum value of the noncentral moments is easily determined by the following procedure. Let the objective function, $J(x,y)$, be given by

$$J(x,y) = m_{pq}.$$

To maximize J subject to the constraint

$$\iint f(x,y) dx dy = 1$$

one may use the method of Lagrangian multipliers:

$$H = J + \lambda (\iint f(x,y) dx dy - 1).$$

Equating the partial derivatives of H with respect to x, y and λ to zero, give the conditions

$$pf(x,y) + xf_x(x,y) = 0$$

$$qf(x,y) + yf_y(x,y) = 0.$$

Thus, the maximum value of m_{pq} may be expressed as

$$J_{\max} = \iint \frac{-pf}{f_x} \frac{-qf}{f_y} f(x,y) dx dy.$$

This expression may be used to determine the maximum value of a known function, $f(x,y)$. For example, suppose that

$$f(x,y) = K \exp\{-x-y\}.$$

Application of the above formula gives

$$J_{MAX} = p^p q^q.$$

Although the previous expression gives the maximum value of the moments, it depends on the function. It is sufficient for computer implementation to use upper bounds for the maximum value. For a finite rectangular region, easily computed bounds may be based on the following relationship:

$$m_{pq} = \int_a^b \int_c^d x^p y^q \rho(x, y) dx dy \leq f_{max} \int_a^b \int_c^d x^p y^q dx dy$$

which gives

$$m_{pq} \leq f_{max} \left[\frac{b^{p+1} - a^{p+1}}{p+1} \right] \left[\frac{d^{q+1} - c^{q+1}}{q+1} \right].$$

A similar upper bound for the central moments is easily shown to be

$$\mu_{pq} \leq f_{max} \left[\frac{(b-a)^{p+1}}{p+1} \right] \left[\frac{(d-c)^{q+1}}{q+1} \right].$$

Lower bounds may be computed in a similar manner.

The computed bounds on the moments may be used to scale the computer operations, select register lengths, and normalize computed feature values.

Another factor to be considered in the numerical computation of moment invariants is that since the invariance properties were proven only for continuous functions, certain variations should be expected in the discrete case. The following development shows that these variations do occur but can be controlled by careful techniques.

Lemma 1. The moments are invariant under translation of coordinates provided $x_0 = k\Delta x$, $y_0 = l\Delta y$.

Proof: Suppose the density with respect to coordinates x, y is $f(x, y)$. Let x', y' be a translation of x, y :

$$x' = x - x_0, \quad y' = y - y_0$$

for some x_0, y_0 . Then with respect to the x', y' coordinates, the density must be $f'(x', y') = f(x' + x_0, y' + y_0)$. With respect to the x, y coordinates, the p - q th central moment is

$$\mu_{x,y}(f, p, q) = \sum_m \sum_n \left(x_n - \bar{x} \right)^p \left(y_m - \bar{y} \right)^q f(x_n, y_m) \Delta x \Delta y$$

$$\left(x_n, y_m \right) \in \Gamma$$

with respect to the x', y' coordinates, the p - q th central moment is

$$\mu_{x',y'}(f', p, q) = \sum_m \sum_n \left(x'_n - \bar{x}' \right)^p \left(y'_m - \bar{y}' \right)^q f'(x'_n, y'_m) \Delta x' \Delta y'$$

$$\left(x'_n, y'_m \right) \in \Gamma$$

where of course, $\Delta x' = \Delta x$, $\Delta y' = \Delta y$.

Note that we can index x'_n such that $x'_n = x_n - x_0$ iff $x_0 = k\Delta x$ for some integer k . Similarly for y'_m . Hence, provided that $x_0 = k\Delta x$, $y_0 = l\Delta y$, then it is readily verified that

$$\bar{x}' = \bar{x} - x_0, \quad \bar{y}' = \bar{y} - y_0 \quad (1)$$

$$x'_n - \bar{x}' = x_n - \bar{x}, \quad y'_m - \bar{y}' = y_m - \bar{y} \quad (2)$$

and

$$f'(x'_n, y'_m) = f(x'_n + x_0, y'_m + y_0) = f(x_n, y_m) \quad (3)$$

and

$$\mu_{x', y'}(f', p, q) = \mu_{x, y}(f, p, q) \quad ; \quad (4)$$

i.e., provided that $x_0 = k\Delta x$, $y_0 = l\Delta y$, then $\mu_{x, y}(f, p, q)$ is invariant under translation of coordinates.

Note: $x_0 = k\Delta x$, $y_0 = l\Delta y$ is required if (1), (2), (3), (4) are to hold exactly. If $x_0 \neq k\Delta x$ and/or $y_0 \neq l\Delta y$, then (1), (2), (3), (4) become approximation and the conclusion is only approximately valid.

Lemma 2. The moments are approximately invariant under similarity transformation.

Proof: We write the moments in the form of

$$\mu_{pq} = \sum_m \sum_n \left(x_n - \bar{x} \right)^p \left(y_m - \bar{y} \right)^q f(x_n, y_m) \Delta x \Delta y$$

$$(x_n, y_m) \in \Gamma$$

$$\mu'_{pq} = \sum_m \sum_n \left(x'_n - \bar{x}' \right)^p \left(y'_m - \bar{y}' \right)^q f(x'_n, y'_m) \Delta x \Delta y$$

$$(x'_n, y'_m) \in \Gamma$$

$$\bar{x}' = \bar{x}, \quad \bar{y}' = \bar{y}$$

$$\mu'_{pq} = \sum_m \sum_n \left(x_n - \bar{x} + x'_n - x_n \right)^p \left(y_m - \bar{y} + y'_m - y_m \right)^q f(x_n, y_m) \Delta x \Delta y$$

$$(x'_n, y'_m) \in \Gamma$$

$$= \sum_m \sum_n (x_n - \bar{x})^p (y_m - \bar{y})^q \left(1 + \frac{x'_n - x_n}{x_n - \bar{x}}\right)^p \cdot \left(1 + \frac{y'_m - y_m}{y_m - \bar{y}}\right)^q f(x_n, y_m) \Delta x \Delta y$$

for small variations of x' from x_n and y'_m from y_m we will have

$$\mu'_{pq} = \mu_{pq} + \sum_m \sum_n \left[p \frac{x'_m - x_n}{x_n - \bar{x}} + q \frac{y'_m - y_m}{y_m - \bar{y}} \right] (x_n - \bar{x})^p (y_m - \bar{y})^q f(x_n, y_m) \Delta x \Delta y$$

$$\mu'_{00} = \mu_{00}$$

So

$$\frac{\mu'_{pq}}{\frac{\frac{p+q}{2} + 1}{(\mu'_{00})}} = \frac{\mu_{pq}}{\frac{\frac{p+q}{2} + 1}{(\mu_{00})}} + R$$

where

$$R = \frac{1}{\frac{\frac{p+q}{2} + 1}{(\mu_{00})}} \left[\sum_m \sum_n \left(p \frac{x'_n - x_n}{x_n - \bar{x}} + q \frac{y'_m - y_m}{y_m - \bar{y}} \right) (x_n - \bar{x})^p (y_m - \bar{y})^q f(x_n, y_m) \Delta x \Delta y \right]$$

The moments are also approximately invariant under digital rotation with the variation dependent upon the size of rotation and the interpolation method used.

4.3.5 Experimental Results for Scene Analysis

To demonstrate the advantages and limitations of moment invariants for scene analysis the following experiment was performed. A general segment of an aerial scene was digitized and imbedded into a zero array since the moment computation depends upon the selected frame size which must be finite. This image was then reduced in size by a factor of two by averaging four points into one; inverted in magnitude modulo 255; mirror imaged, rotated 45° and 2° using nearest neighbor interpolation. These images are shown in Figure 47.

The corresponding invariant measurements are shown in Figure 46. The logarithms have been taken of the invariants to compress the dynamic range. Note that the invariants are close but not identical for the different transformations of the original scene.

The main conclusion which can be deduced from these experimental results is the great conformity of the theory with the empirical results. This suggests moment invariants as a practical and convenient tool in shape representation and image registration. At the end point of the calculation one obtains a set of measurements which are invariant to any translation, rotation or size change which might occur in any picture of a scene due to various factors. The method may be expanded beyond this extent, i.e. for order p moments, one has $p+1$ invariant measurements (refer to Section 4.1). Hence, a large supply of these invariant measurements are available.

EXPERIMENTAL RESULTS:

In this part we verified empirically the seven invariant moments which have been derived from a given scene. The measurements, shown in Fig. 46 correspond to the pictures in the following pages respectively.

FINAL INVARIANTS-

6.24993
17.18015
22.65516
22.91954
45.74918
31.83071
45.58951

Optical Window

FINAL INVARIANTS-

6.22637
16.95439
23.53142
24.25687
48.34990
32.91619
48.34356

Optical Reduced 2

FINAL INVARIANTS-

6.21612
16.58215
21.79714
22.69471
44.92849
31.22309
44.90914

Optical Inverse

FINAL INVARIANTS-

6.91980
19.95532
26.68924
26.90140
53.72453
37.13457
53.59021

Optical Mirror

FINAL INVARIANTS-

6.31823
16.80396
19.72426
20.43774
40.52568
29.31589
40.47074

Optical Rotation 45°

FINAL INVARIANTS-

6.25346
17.27091
22.83652
23.13025
46.13627
32.06803
46.01707

Optical Rotation

Figure 46.



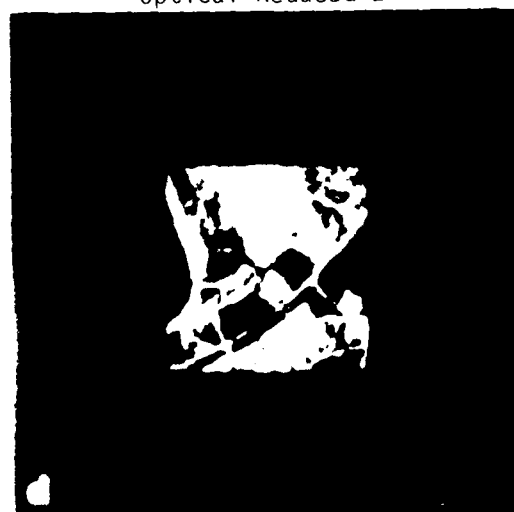
Optical Window



Optical Reduced 2



Optical Mirror



Optical Rotation 2



Optical Rotation 45

FIGURE 48.

REFERENCES

1. E.T. Bell, Men of Mathematics, Simon and Schuster, New York, 1965.
2. M.K. Hu, "Visual Pattern Recognition by Moment Invariants," IRE Trans Inform. Theory, Vol. IT-8, pp. 179-187, 1962.
3. A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
4. E.L. Hall, W.O. Crawford, and F.E. Roberts, "Computer Classification of Pneumonconioses from Radiographs of Coal Workers," IEEE Transactions on Biomedical Engineering, Vol. BME 22, No. 6, November 195, pp. 518-527.
5. E.L. Hall, R.P. Kruger, and A.F. Turner, "An Optical-Digital System for Automatic Processing of Chest X-Rays," Optical Engineering, Vol. 13, No. 3, June 1974, pp. 250-257.

SECTION V.

HIERARCHICAL SEARCH TECHNIQUES

5.1 Introduction

The problem of matching two images of the same scene taken by different sensors under different conditions is a challenging problem in scene analysis. The two scenes called the reference scene and sensed scene may be transformed so drastically by different data collection geometrics or sensors that it is not possible to match the scenes. However, in most interesting cases, the scenes can be matched by a photointerpreter. Only this class of scenes will be considered. A general approach to matching two scenes involves locating corresponding regions in the scenes. If several corresponding regions can be located, then geometric and intensity mappings may be developed to match the scenes. Therefore, a basic problem is: given a region of the reference (sensed) image, determine its location in the sensed (reference) image (Fig. 48). The standard computer approach to this problem is to use the gray levels or a derived image of the reference region as a template and select the position of maximum cross correlation between the template at each possible shift position of the sensed image as the match location. Since a template of size $M \times M$ can be shifted into $(N-M)^2$ possible positions in an $N \times N$ image as shown in Figure 1, the number of correlation computations can be extremely large. Fast correlation and edge correlation techniques

decrease the correlation time at each shift position; however, these methods still require a computation at each of the $(N-M)^2$ positions. The purpose of this section is to describe hierarchical search techniques which are logarithmically efficient, i.e. reduce the number of search positions to $K \log (N-M)$.

Although no previous work on hierarchical map matching has been discovered, several previous experiments on hierarchical decompositions of pictures have motivated this study. Klinger and Dyer[1] developed a regular decomposition of picture areas into successively smaller quadrants resulting in a logarithmic search. Tanimoto and Pavlidis [2] also considered a hierarchical data structure for picture processing in order to speed up edge detection operations. Ramaperian [3] also used multilevel search techniques for edge detection.

Several advantages of a structured approach are apparent. It is not necessary to examine each pixel in a high resolution image to locate a region at high resolution. The selectivity of the hierarchical techniques, especially coarse-fine search methods, is similar to the perceptual operation of the efficient human visual system. The proposed method in which a match region is obtained at different levels is extendable to other problems such as edge or object location. The method provides an arbitrary degree of precision in locating a region which is limited only by the highest resolution size and the uniqueness of the match region. Finally, the method permits

an efficient decomposition of the sensed scene into "informative" and "irrelevant" regions

II. Coarse-Fine Search Technique

In the hierarchical technique a structured set of pictures at different resolutions will be used. The high resolution sensed scene will be denoted, $P_L(i,j)$ where for simplicity

$$i, j = 0, 1, 2, \dots, 2^L - 1$$

The index L is called the level of the search. The reference region at level L , $Q_L(i,j)$ will be assumed to be smaller than the sensed scene. The agglomerative rule by which the level K scene is reduced to a level $K-1$ scene is simple four point averaging, i.e.

$$P_{K-1}(i,j) = \frac{1}{4} \left\{ P_K(2i, 2j) + P_K(2i, 2j+1) + P_K(2i+1, 2j) + P_K(2i+1, 2j+1) \right\}$$

$K-1$

for $i, j = 0, 1, 2, \dots, 2^{K-1} - 1$

The reference image region size varies with the level and is an important consideration. Obviously, objects present at one level may not be recognizable at a lower level. However, allowing the reference image size to change alleviates this problem. When looking for a forest one need not look at the leaves. A reference match region must be selected for each level and the performance of the algorithm depends upon the uniqueness of these reference regions.

A matching rule must also be specified to guide the search from level $K-1$ to level K . Several possible match function such as correlation may be used. The match function $\underline{x} \sim \underline{y}$, for $\underline{x}$ and $\underline{y}$ should have certain properties, such as

1. Identity, $\underline{x} \sim \underline{x}$
2. Symmetry, if $\underline{x} \sim \underline{y}$ then $\underline{y} \sim \underline{x}$
3. Shift invariance, if $\underline{x} \sim \underline{y}$ then $\underline{x} + \underline{a} \sim \underline{y} + \underline{a}$ where $\underline{a}$ is a constant vector.
4. Scale invariance, if $\underline{x} \sim \underline{y}$ then $\lambda \underline{y}$ where λ is a scalar.
5. Rotation invariance, if $\underline{x} \sim \underline{y}$ then $\underline{T} \underline{x} \sim \underline{y}$ where $\underline{T}$ is a transformation matrix.
6. Sensor invariance, if $\underline{x} \sim \underline{y}$ then $t(\underline{x}) \sim \underline{y}$ where t is the sensor transformation.

The vectors $\underline{x}$ and $\underline{y}$ may be the gray level values or derived measurements from corresponding sensed and reference regions. Note that normalized correlation satisfies properties 1-4 but not 5 and 6. Invariant measurements such as moment invariants used with normalized correlation satisfy properties 1-6. Certain derived functions such as edges or sensor corrected images used with invariant measurements and normalized correlation may satisfy all the properties. Also, note that other properties such as invariance to weather or even man made changes may be desirable.

A specific example which illustrates a coarse-fine search will now be considered. A quadrant tree will be used for decomposition of the sensed scene as shown in Figure 49.



FIGURE 49. SEQUENTIAL ZOOM TECHNIQUE

The efficiency of a quadrant subdivision algorithm for a coarse-fine search is illustrated in Figure 50. Location of an element in an $N \times N$ array, requires only $\log_2 N$ steps.

One of the prominent features in cross-correlation graphs is the differences in the location of the peak when one obtains cross-correlation of radar picture against the corresponding optical picture, and the cross-correlation of the optical picture against the corresponding radar picture. This phenomenon can be explained fully by noting the existing misregistration of the two images.

Consider two functions $f(x,y)$ and $h(x,y)$. The correlation function of these two can be defined as:

$$C(\alpha, \beta) = \iint_{-\infty}^{\infty} f(x, y) h(x + \alpha, y + \beta) dx dy$$

Now suppose there is another function $h(x,y)$, which is the translated version of $h(x,y)$. Then:

$$h'(x,y) = h(x-a, y-b)$$

where a, b are relative distances that has been translated.

The value for correlation in this case would be:

$$\begin{aligned} C'(\alpha, \beta) &= \iint_{-\infty}^{\infty} f(x, y) h'(x, y) dx dy \\ C'(\alpha, \beta) &= \iint_{-\infty}^{\infty} f(x, y) h(x + \alpha - a, y + \beta - b) dx dy \\ &= C(\alpha - a, \beta - b) \end{aligned}$$

Hence one concludes that the new correlation function is just the translated version of the original correlation. The radial dislocation is then

$$d = \sqrt{a^2 + b^2}$$

Now if $f'(x,y)$ is the translated version of $f(x,y)$, i.e.

$$f'(x,y) = f(x-a', y-b')$$

then

$$C''(\alpha, \beta) = \iint_{-\infty}^{\infty} f(x-a', y-b') h(x+\alpha-a) (y+\beta-b) dx dy$$

Now

$$x = a' = x' \quad , \quad y - b = y'$$

then

$$x = x' + a' \quad , \quad y = y' + b'$$

$$C''(\alpha, \beta) = \iint_{-\infty}^{\infty} f(x', y') h(x'+a'-a+\alpha, y'+b'-b+\beta) dx' dy'$$

or

$$C''(\alpha, \beta) = C(\alpha+a'-a, \beta+b'-b)$$

Hence the relationship of this new correlation when both $f(x,y)$ and $h(x,y)$ are translated is just the translated version of the original correlation.

Conclusion:

In the situation of the correlation graphs that was referred before we have

$$a' = -a, \quad \text{and} \quad b' = -b$$

hence

$$C'(\alpha, \beta) = C(\alpha - 2a, \beta - 2b)$$

then the radial distance of C with respect to C" is

$$d = 2d$$

5.2 EXPERIMENTAL RESULTS

A sequence of experiments were performed to investigate the zoom technique for actual images. The simple 4 for 1 averaging was used for the agglomerative step. An investigation was made of the size reduction allowable for matching of the large and small stadium regions for both the optical and radar images. Correlation was used as the measure of similarity and a resolution which permitted a significant correlation peak was taken as adequate. Another experiment using a derived edge picture and moment invariants was also performed. Finally, an example is presented of selecting moment measurement from the optical regions for location of the match position in the radar image at various resolutions.

The first sequence of experiments were performed to verify the basic concept of the zoom technique. The geometrically corrected optical image scanned at 256×256 resolution was reduced to 128×128 , 64×64 and 32×32 . A similar sequence was produced for the intensity corrected radar image. The following experiments were done:

Step 1:

A 64×64 optical picture was divided into four equal major regions. Then the invariant moments of the picture were used as a means of representing the picture. The top right window with (1,1) coordinates was then used as a referenced region and the correlation function of this sensed region with the whole picture was obtained. The reference region was moved pixel by pixel across the whole

sensed picture and thus the shown correlation function was obtained. The window size was 32×32 .

Step 2:

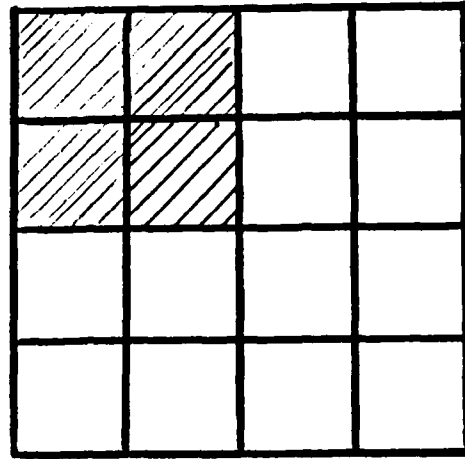
The selected region in step one was enlarged to 64×64 , then this picture again was divided into four equal regions. The region with the coordinate (33, 33) was selected and, using a similar procedure as in step one, the correlation function was obtained.

Step 3:

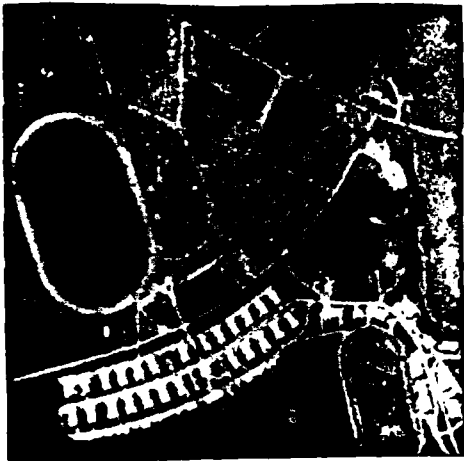
A similar procedure as in step 2 was performed, with the exception of the selection of the region, which in this case was with the coordinate (1,1). The obtained correlation function is shown.

In all the pictures shown, the photograph corresponds to the next largest square, which contains the shaded region. For example, in the Figure 51, the photograph corresponds to the square which is labeled c' in the large square.

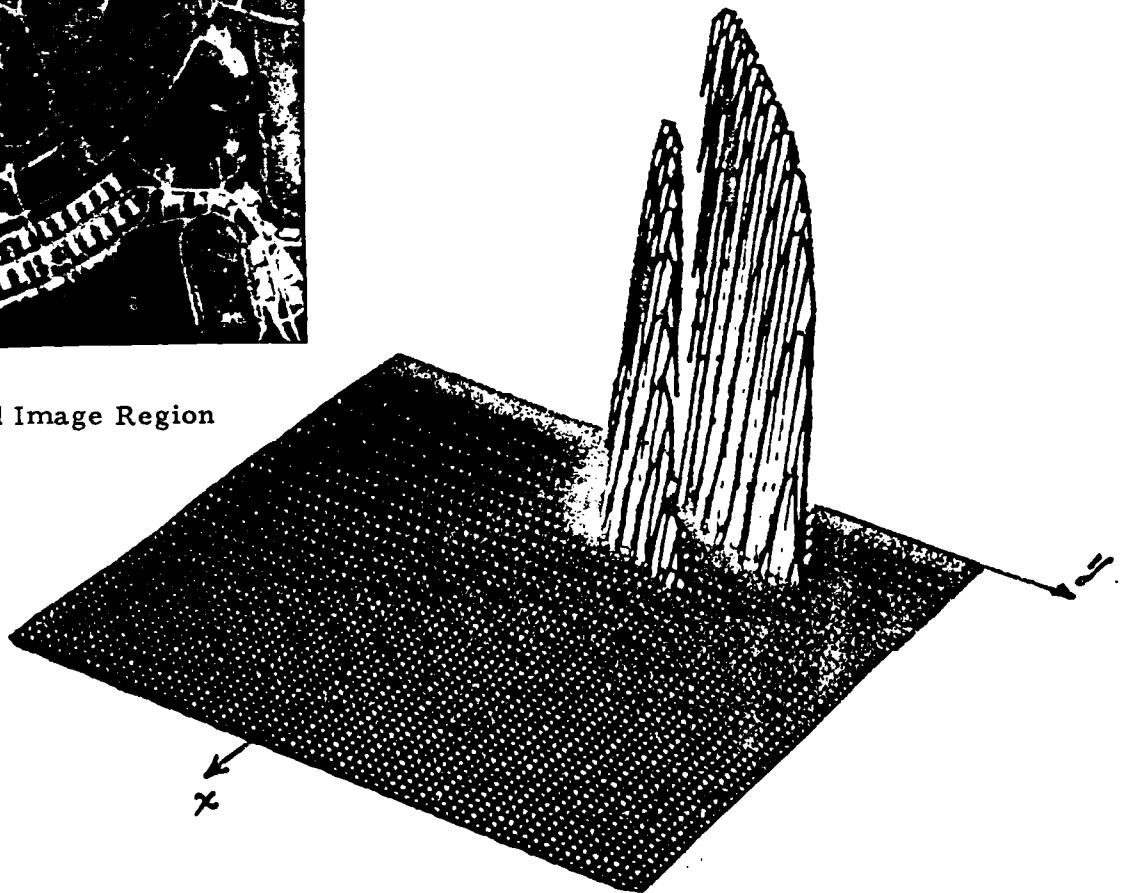
The resulting correlations for the optical image are shown in Figures 51, 52 and 53. The corresponding experiment for the radar images are shown in Figures 54, 55 and 56.



Region Selection

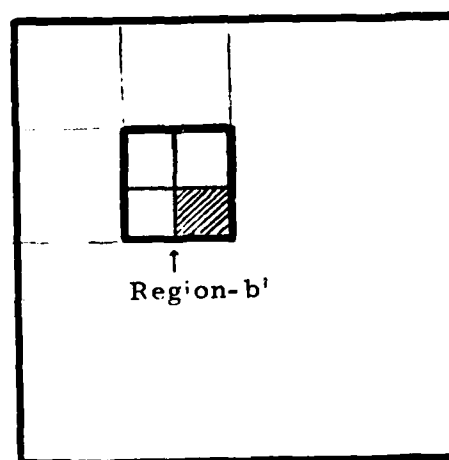


Selected Image Region

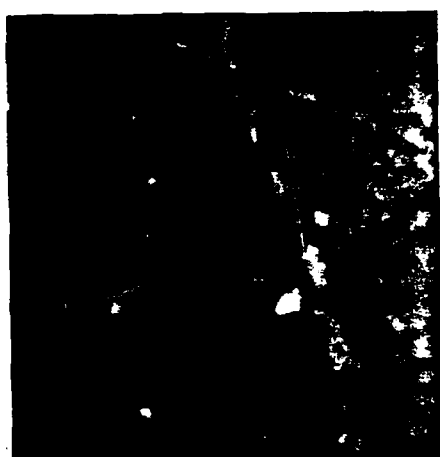


Correlation Function

Figure 51.



Region Selection



Selected Image Region-b'

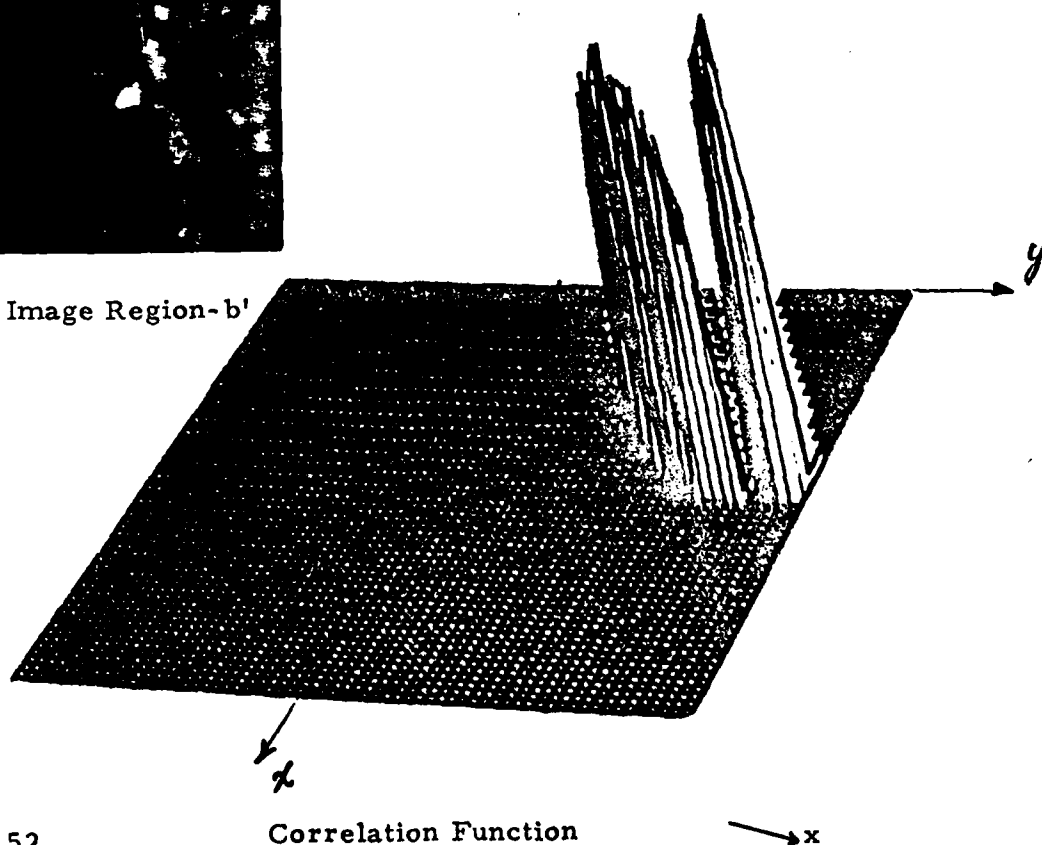


Figure 52 .

Correlation Function

OPTICAL DATA

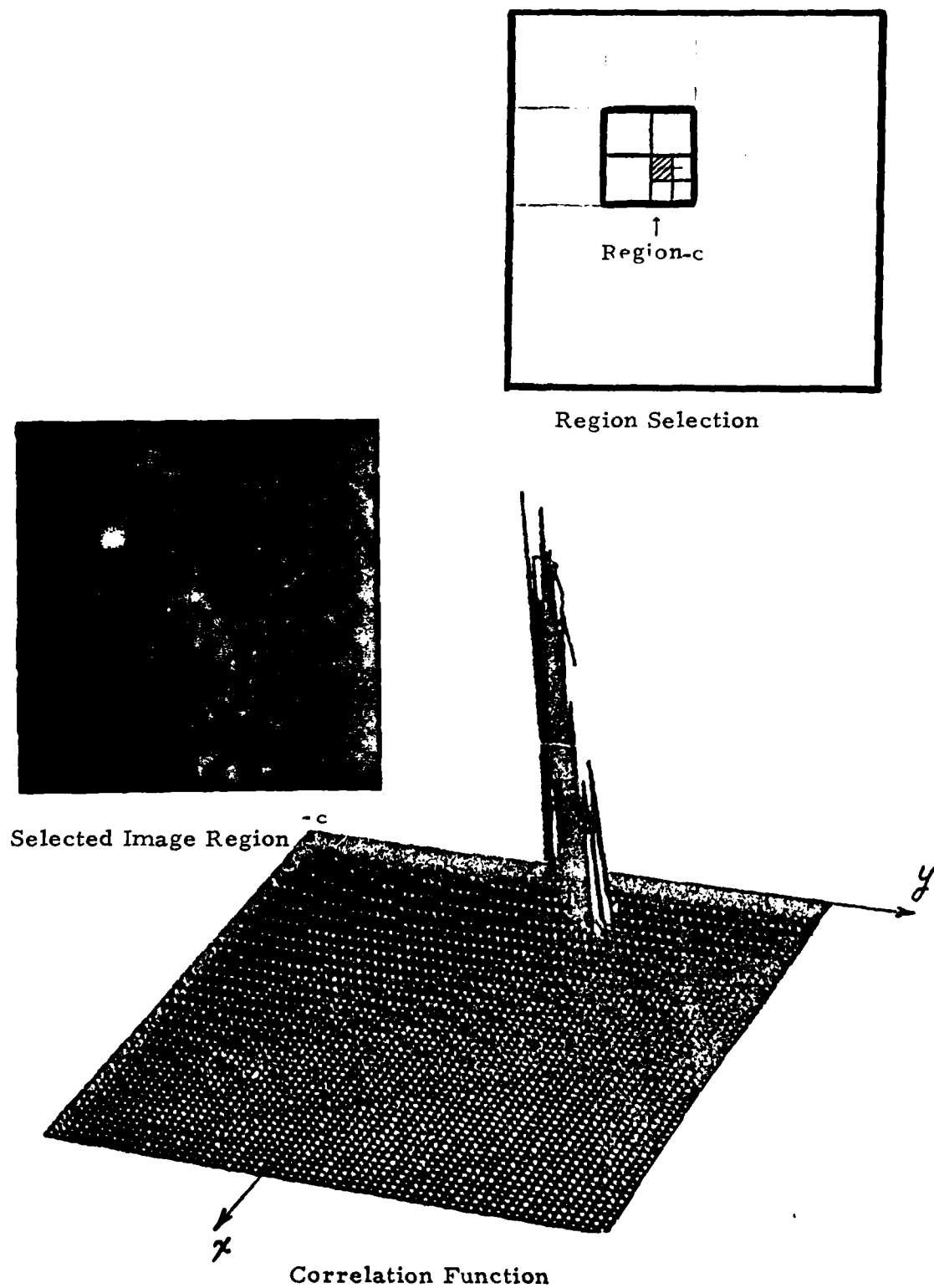
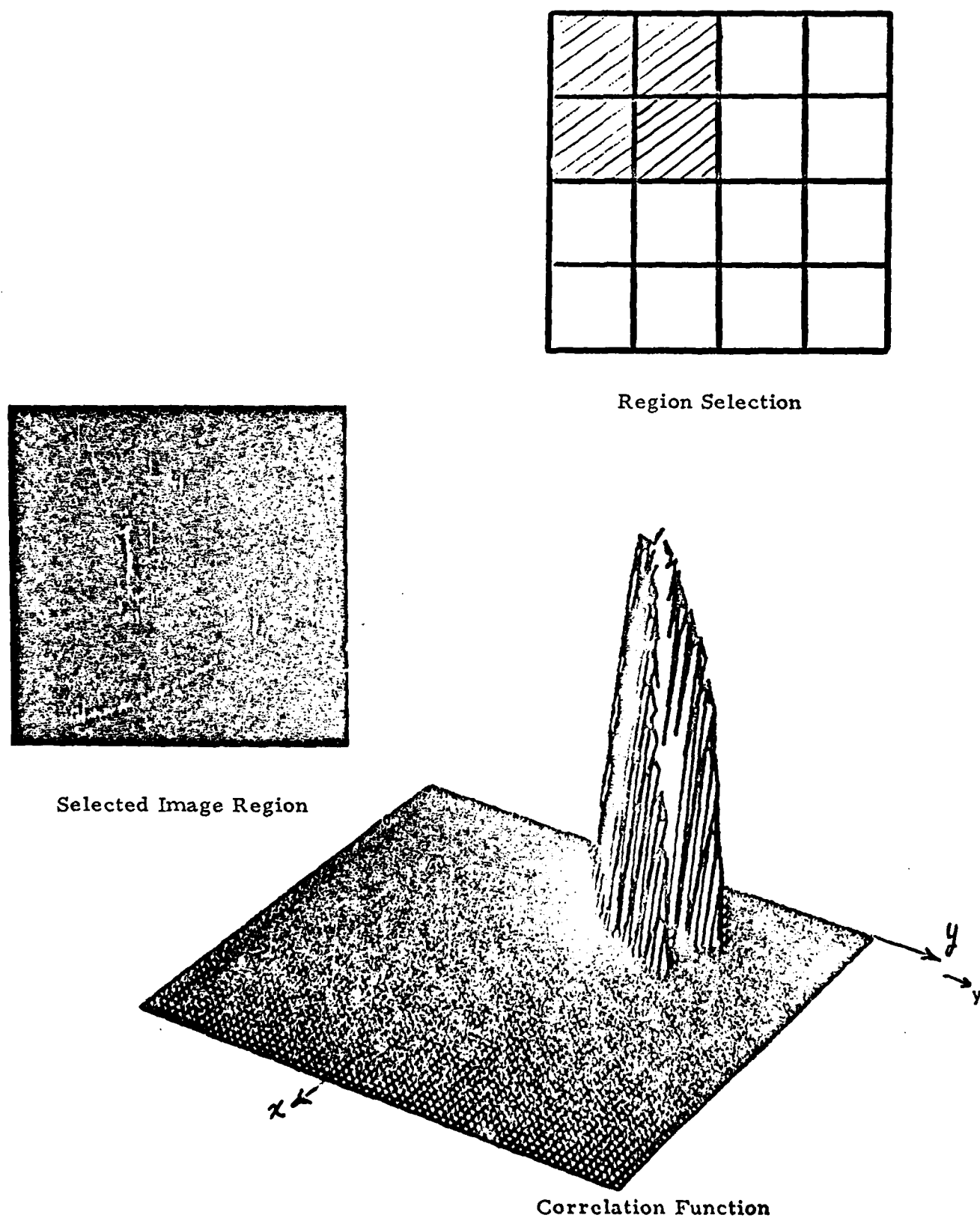
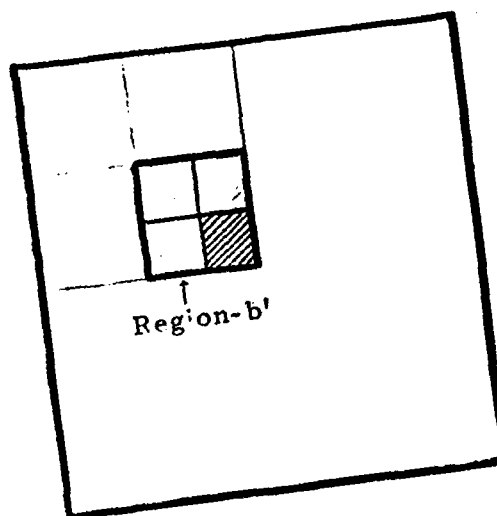


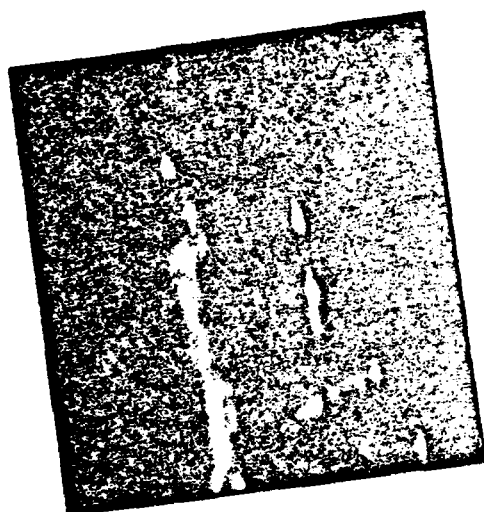
Figure 53.

Figure 54. Cross correlation of optical and radar measurements

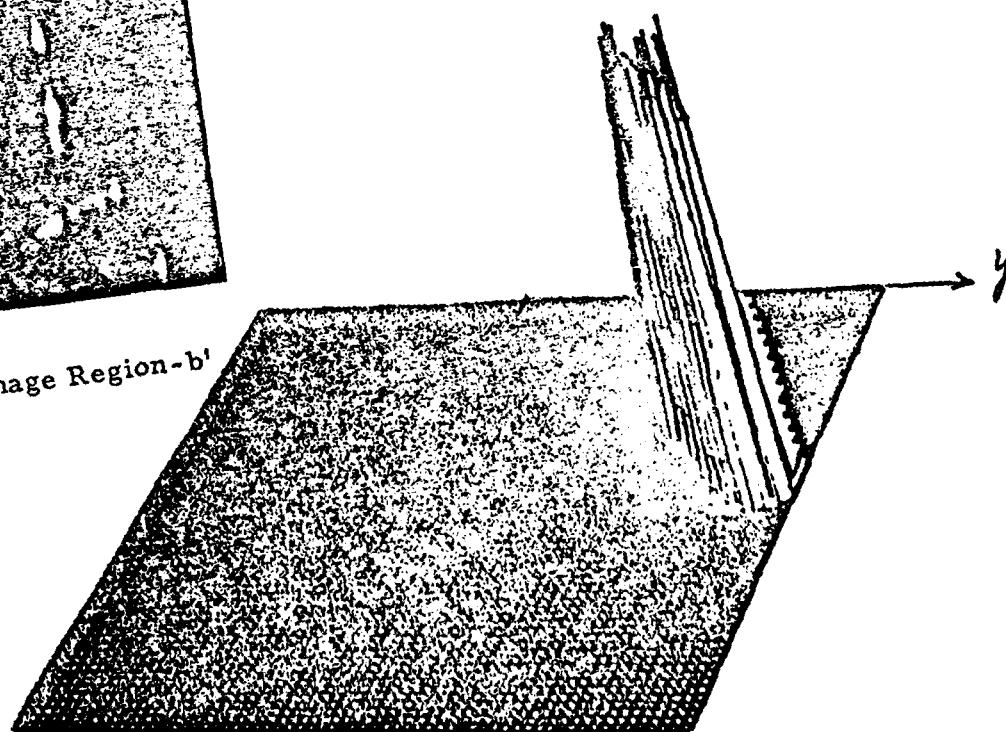




Region Selection



Selected Image Region-b'



Correlation Function

Figure 55.

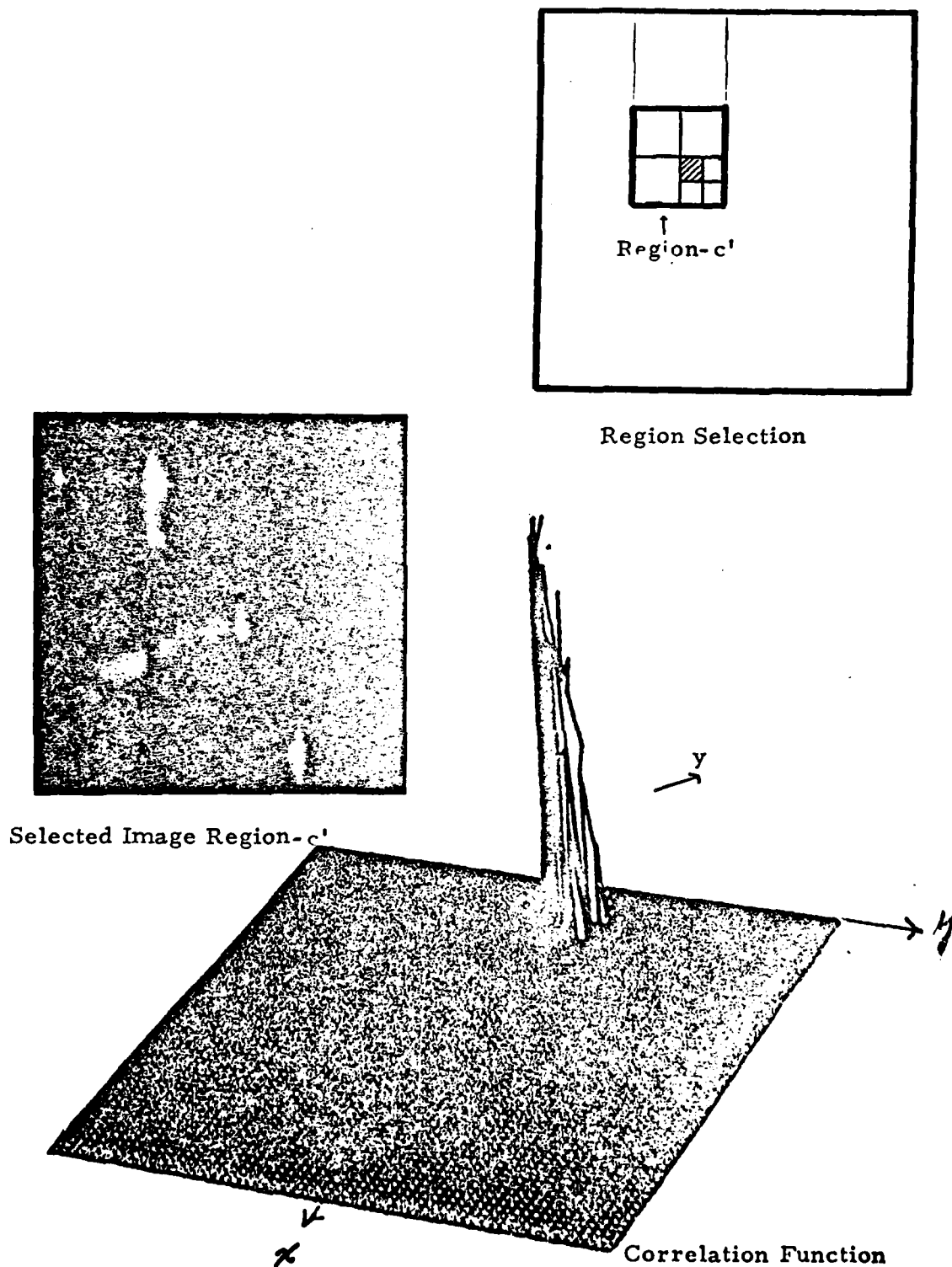


Figure 56.

REFERENCES

1. A. Klinger and C.P. Dyer, "Experiments on Picture Representation Using Regular Decomposition," Computer Graphics and Image Processing 5, pp. 68-105, 1976.
2. S. Tanimoto and T. Pavlidis, "A Hierarchical Data Structure for Picture Porcessing," Computer Graphics and Image Processing 4, pp. 104-119, 1975.
3. H. K. Ramapriyan, "A Multilevel Approach to Sequential Detection of Pictorial Features," IEEE Transaction on Computers, pp. 66-78, January 1976.

SECTION VI. SCENE CONTENT MEASUREMENTS AND RECOMMENDATIONS FOR FUTURE WORK

The general problem of map matching of images of the same scene taken under different conditions with different sensors was considered in detail during this study. Several innovative approaches and concepts were investigated and demonstrated by experimental computations on images from optical and radar sensors. Among these were:

- Techniques of geometric and sensor intensity corrections
- Hierarchical search for image registration
- Extraction of invariant features and invariant feature correlations
- Edge extraction and edge correlations

Since most correlation techniques work well on locating a region from an image, one approach to map matching of two different images is to develop inverse transformations to make the images appear as a region from one image. This general transformation depends upon both sensor and imaging characteristics and is difficult if not impossible to determine. However, it was possible to develop simple approximations which rectify not only the geometric but also the intensity of two arbitrary images. A "map warping" approach which involves a two dimensional polynomial mapping from corresponding points in the two images was used for geometric corrections. A new technique based upon the Karhunen-Loeve transformation was used to determine a principal component of the image with

a corrected intensity distribution. This approach is an example of a class of intensity matching transformations which because of their effectiveness and simplicity may become important techniques for map matching. These techniques only require that intensity statistics be assumed or experimentally determined. Other information that could be obtained about the scene, such as object composition or height, would permit the develop of more sophisticated methods.

Hierarchical search techniques for image registration were introduced in this report and promise to be a highly efficient and effective method for map matching. An impressive feature of these methods is logarithmic computational efficiency. A special case of the hierarchical techniques permits a "zoom" to be used to locate an object of interest. That is, an image can be searched at low resolutions and a region is located approximately. Then the approximation can be refined to any desired level by investigating selected regions in higher resolution images. The hierarchical technique was demonstrated for locating regions from an optical image in an optical image, regions in a radar image in a radar and regions of a radar image in an optical image. The matching of radar and optical image requires the extraction of measurements which are invariant to sensor and geometric distortions. This requirement led to the study of invariant measurements.

Certain measurements have a high degree of invariance to the types of geometric and sensor transformations produced in optical and side looking radar imaging. Moment measurements of the image intensity can be transformed into a set of invariants which are invariant to translation, rotation, scale and other transformations. These invariants were computed and used as measurements in correlation experiments. It was noted that due to the digital nature of the computation the measurements were only approximately invariant but did work quite well for map matching optical and radar images.

Edge structure also remains invariant to many sensor transformations and thus edge correlation for map matching was also considered in detail. A new decomposition method for 3 by 3 windows was developed and used for edge detection. Although edge structure is very invariant to sensor transformations, the edge correlator may be sensitive to geometric transformations depending upon the edge detection technique. Experimental optical-radar correlations were performed to demonstrate the efficiency of the edge correlation approach.

A common problem encountered in all correlation methods is a sensitivity of performance to the content of the scene. This phenomena was demonstrated in several examples and requires further investigation.

Each of these techniques and developments have been successfully applied to a limited set of images. The use of a limited number of images was necessary at this time in order to provide a well-defined environment for the development

of the basic techniques.

To demonstrate the true usefulness of these approaches, they must be applied to many scenes of interest. Therefore, it is recommended that the following work be performed:

1. Establish the performance of these approaches to sequential map matching using hierarchical search, edge features, and structured invariant moments with respect to terrain and type.
2. Identify the types of scenes and features which are best suited for these approaches.
3. Perform theoretical and experimental statistical analysis on these approaches and determine suitability of these approaches to the perturbed scenes.
4. Verify the validity of experimental performances through statistical analysis.

