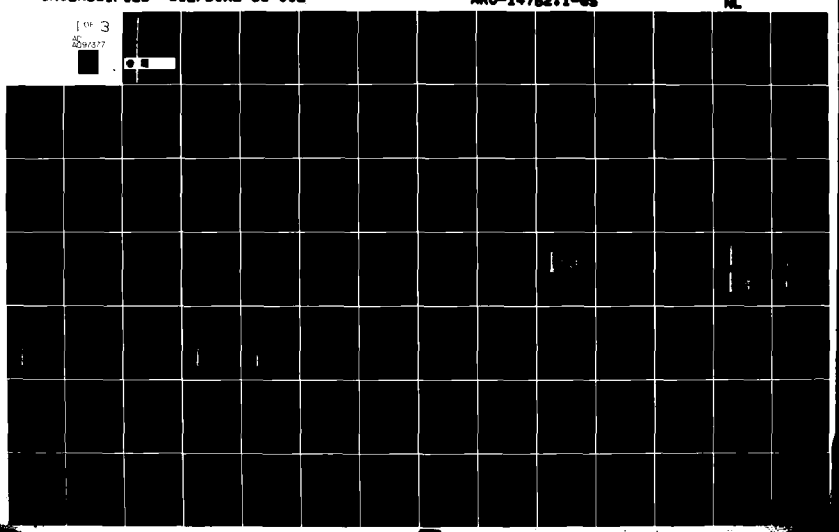
AD-A097 377 ARIZONA UNIV TUCSON DIGITAL IMAGE ANALYSIS LAB F/G 8/2
AUTOMATIC METHODS IN IMAGE PROCESSING AND THEIR RELEVANCE TO MA--ETC(U)
FEB 81 B R HUNT DAAG29-77-G-0175
UNCLASSIFIED SIE/DIAL-81-002 ARO-14782.1-MS NL



ARO 14782.1-GS

SIE/DIAL-81-002

AUTOMATIC METHODS IN IMAGE PROCESSING AND THEIR RELEVANCE TO MAP-MAKING

Final Report

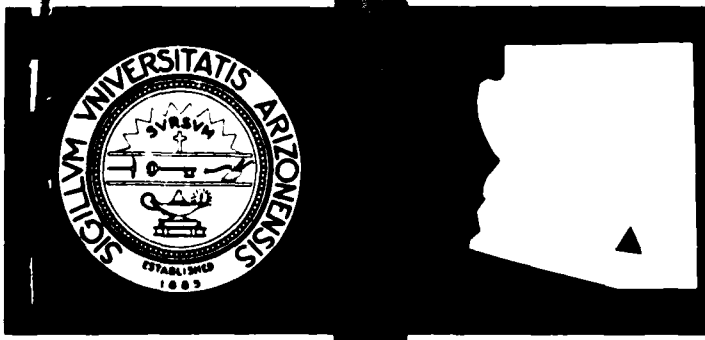
B. R. Hunt

February 11, 1981

U. S. ARMY RESEARCH OFFICE

DAAG29-77-G-0175

Digital Image Analysis Laboratory
Systems and Industrial Engineering Department
University of Arizona
Tucson, Arizona 85721



ENGINEERING EXPERIMENT STATION
COLLEGE OF ENGINEERING
THE UNIVERSITY OF ARIZONA
TUCSON, ARIZONA

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: C	2. GOVT ACCESSION NO: AD-A097 477
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): AUTOMATIC METHODS IN IMAGE PROCESSING AND THEIR RELEVANCE TO MAP-MAKING.	5. TYPE OF REPORT & PERIOD COVERED: Final Report.
	6. PERFORMING ORG. REPORT NUMBER: 15
7. AUTHOR(s): B. R. Hunt, Professor of Systems Engineering and Optical Sciences, and Director of the Digital Image Analysis Laboratory	8. CONTRACT OR GRANT NUMBER(s): DAAG29-77-G-0175
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Digital Image Analysis Laboratory, University of Arizona, Tucson, Arizona 85721	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, Post Office Box 12211, Research Triangle Park, NC 27709	12. REPORT DATE: February 11, 1981
	13. NUMBER OF PAGES: 220
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

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

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

The behavior of digital cross-correlation algorithms as applied to image matching problems is examined in terms of the relationship between measureable image properties and algorithm characteristics. Statistical image quality measures are developed which could be employed in a preprocessor to predict the performance of automatic stereo-compilation equipment. The measures include a quantity derived from the Cramer-Rao lower bound on the variance of any unbiased parameter estimator, various contrast

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

measures such as variance, contrast modulation, and median absolute deviation, and a stationarity detector related to the variance gradient. These measures are based on image and correlator models which describe the behavior of correlation processors under conditions of low image contrast or signal-to-noise ratio, geometric distortion, and image non-stationarity. Computer simulations using synthetic imagery were performed to verify the various models, and indicate the potential for the use of image quality measures in the the predicting of correlation behavior. Implications of the models in terms of correlation processor design and implementation are discussed.

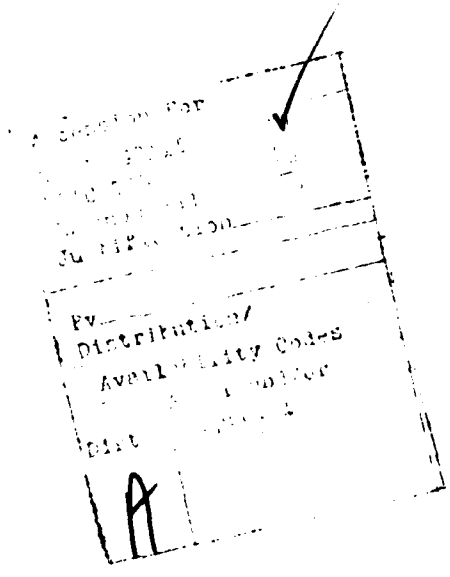


TABLE OF CONTENTS

TABLE OF CONTENTS--CONTINUED

TABLE OF CONTENTS--CONTINUED

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS--CONTINUED

LIST OF ILLUSTRATIONS--CONTINUED

LIST OF TABLES

ABSTRACT

The behavior of digital cross-correlation algorithms as applied to image matching problems is examined in terms of the relationship between measureable image properties and algorithm characteristics. Statistical image quality measures are developed which could be employed in a preprocessor to predict the performance of automatic stereo-compilation equipment. The measures include a quantity derived from the Cramer-Rao lower bound on the variance of any unbiased parameter estimator, various contrast measures such as variance, contrast modulation, and median absolute deviation, and a stationarity detector related to the variance gradient. These measures are based on image and correlator models which describe the behavior of correlation processors under conditions of low image contrast or signal-to-noise ratio, geometric distortion, and image non-stationarity. Computer simulations using synthetic imagery were performed to verify the various models, and indicate the potential for the use of image quality measures in the predicting of correlation behavior. Implications of the models in terms of correlation processor design and implementation are discussed.

CHAPTER 1

INTRODUCTION

In many image processing applications, it is necessary to spatially register two or more images of an object field obtained from different sensors or taken from the same sensor at different times or locations. The registration of two images consists of translating them relative to one another so that they exactly overlay each other. The translation (and possibly rotation) required to achieve registration provides information concerning the relative positions of the sensors and points in the object field. Applications of image registration are numerous and include such diverse areas as guidance of unmanned vehicles, change detection, target acquisition and tracking, and the production of topographic maps. It is primarily the latter topic that is addressed herein although many of the concepts and problems to be discussed are also found in other applications.

A problem of considerable interest within the cartographic community is that of mechanized "stereo-viewing", that is, automatically locating identical (conjugate) points on two or more aerial photographs. The result of this interest has been the development of several automated stereocompilation systems including the Universal Automatic Map

Compilation Equipment (UNAMACE) [Thompson 1966], the AS-11B-X Stereo Mapper [Brumm, et al. 1976], the Gestalt Photo Mapper GPMII [Crawley 1974], and the Heterodyne Optical Correlator (HOC) [Balasubramanian 1976]. The principal advantage of automated stereo-systems are their speed, economy and precision (although not necessarily accuracy). Different human operators using a stereo-plotter will not achieve the same degree of repeatability as will the automated system. The human operator, however, is far more reliable. The limitation in machine reliability is a function of the variability of terrain conditions over which it must operate. The basic function of the automatic stereo-compiler is to compare two nearly identical images (around corresponding conjugate points) and determine the position resulting in maximum similarity. The machine, however, is unable to recognize objects or shapes. When the machine is confronted with images that differ considerably in appearance, it does not function properly. As we will see, under conditions of steep terrain, the two images can differ significantly. If there are points in one image that are not visible in the other image, the results of the matching are, in general, unpredictable. The human operator, on the other hand, can recognize objects and shapes even if they differ in scale or aspect. Most present day automatic compilers incorporate terrain slope correction features which allow for on-line geometric correction of the distortion produced

by terrain relief. This added capability, however, reduces the overall system speed.

The automatic compiler is also less reliable in regions of imagery devoid of contrast, particularly in the presence of noise which tends to increase the difference between conjugate regions. Regardless of the sources of image differences, the machine occasionally "gets lost". That is, it will come to a point where it is unable to locate identical corresponding imagery. When this occurs, the machine will search until it is far off track. Finally an alarm is sounded to notify an operator of its condition. Because of the occurence of these conditions, the systems are only semi-automatic at best.

Stereo-compilation is accomplished by cross-correlation. The pictorial information is first converted into electrical signals via a scanning operation. The resulting signals (either analog or digital) are then compared in a correlator circuit or computer. The translation, or parallax, between the images is determined by the location of the peak of the cross-correlation function indicating maximum similarity between image segments. Performance for image registration is measured by the accuracy with which the offset of the peak of the correlation function matches the true parallax created by topographic relief and sensor orientation.

The term "correlation" is a generic term that includes a number of different algorithms and techniques which will be described in Chapter 2. Each of the techniques, however, generates a correlation function. An example of such a function, $c(\xi)$, is shown in Figure 1-1a. Without loss of generality, it is assumed that the correct peak position is at the origin of the ξ-axis. Due to a variety of noise sources to be described in Chapter 4, there are statistical fluctuations in the measured correlation function, $\hat{c}(\xi)$, which are indicated by the dashed lines in Figure 1-1. These fluctuations can produce two distinctly different types of correlation errors. An error occurring even though the correct lobe of the correlation function is chosen is called a "local registration error" as shown in Figure 1-1b. An error resulting from selecting an incorrect lobe is commonly termed a "false acquisition error" as shown in Figure 1-1c. The conditions under which these errors occur depend in part on the particular correlation technique. It is generally the case, however, that local registration errors occur in regions producing broad correlation functions while false acquisition errors occur in regions producing relatively narrow correlation functions.

No matter what the actual physics of the correlation process, the correlation calculations are subject to a certain amount of error, and the errors result in inaccurate determination of parallax values. Frequently these errors

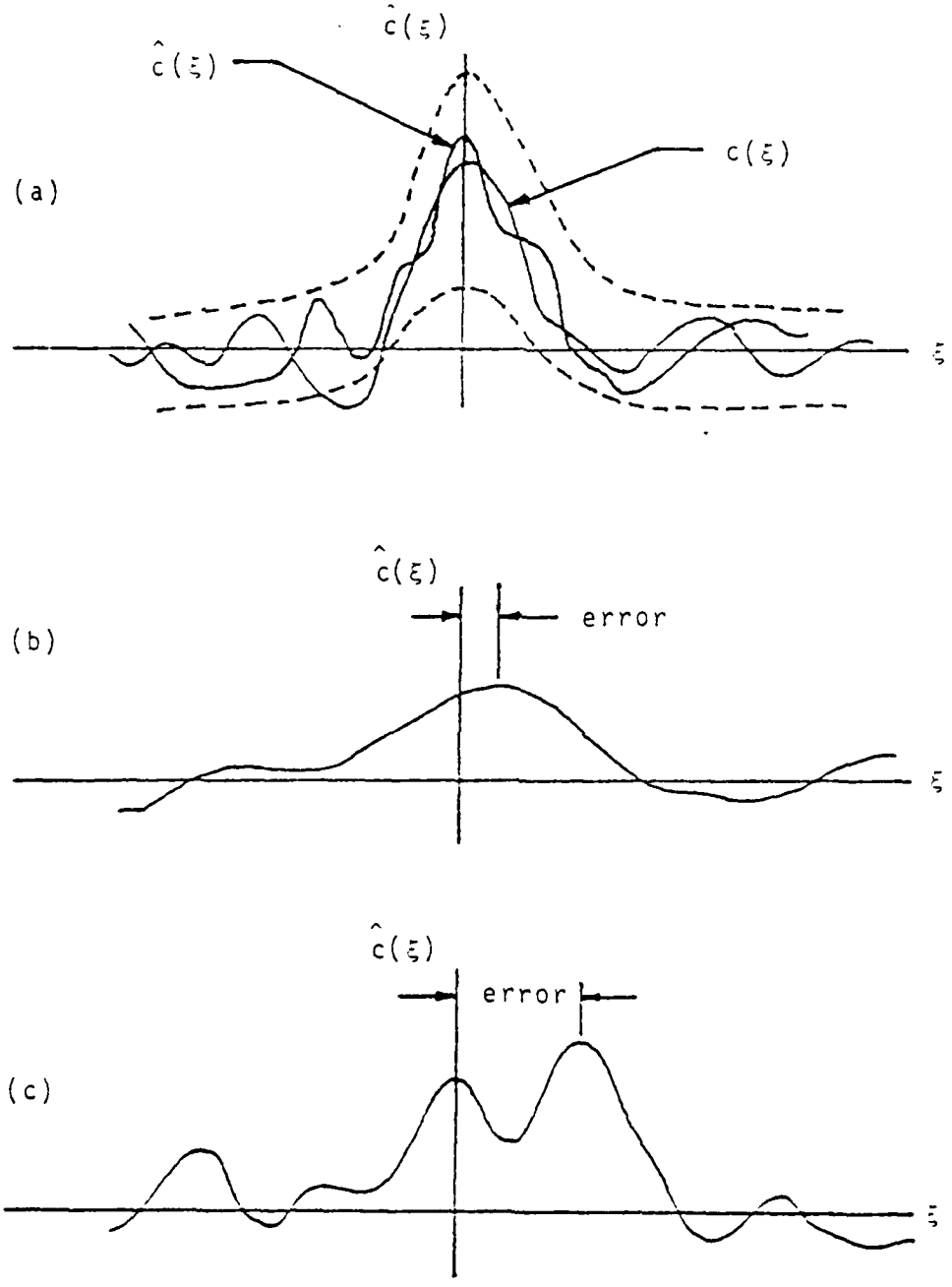


Figure 1-1. Correlation Errors
(a) no error
(b) local registration error
(c) false acquisition error

are catastrophic and lead to terrain elevation values that are impossible or inconsistent. Even when the errors are not catastrophic, it is of the greatest practical utility to know the magnitude of errors that are associated with a particular region of an image. The degree to which a particular region is "error prone" depends on the physical properties of the terrain, the correlation technique, the implementation of the technique, and the properties of the image function as well as the various contaminating noise sources. Knowledge of these relationships would allow the development of stereo-compilation pre-processors to detect image regions that are defective (error prone) in the sense that the probability of correct conjugate acquisition is inadequate. With such a "defect map", the correlation processor could operate adaptively within the following options:

(a) Reject the region and indicate the need for new data acquisition,

(b) Reject the region and indicate that the region can be successfully processed by human operators,

(c) Switch to a different correlation technique known to be less sensitive to the particular defect,

(d) Enhance the region in some way to improve the probability of correct conjugate acquisition,

(e) Accept the region as having sufficient quality to process as usual.

The ability to pre-process stereo images in this manner would allow improved resource allocation (human vs. machine) and thus improve the overall throughput rate and economy of the operation.

It is the primary purpose of this research effort to examine the behavior of digital image matching techniques in terms of the relationship between measurable image properties and known algorithm characteristics and to develop image quality measures which can be employed to predict the performance of automatic stereo-compilation equipment. In the remainder of this chapter we briefly review the basic concepts involved in stereo-compilation.

Stereo-Compilation Concepts

Figure 1-2 shows the geometry assumed in the analysis. To simplify notation, the functional relationships are defined in one dimension only. Extension to two dimensions is direct and does not alter any of the analysis. On the ground there are two quantities of interest: the elevation, $e(X)$, and a corresponding intensity pattern, $I(X)$, which is the result of light reflected from the features of the land. The camera is located at an altitude of H above reference elevation and the focal plane of the camera is assumed to be parallel to the reference plane. The camera has focal length

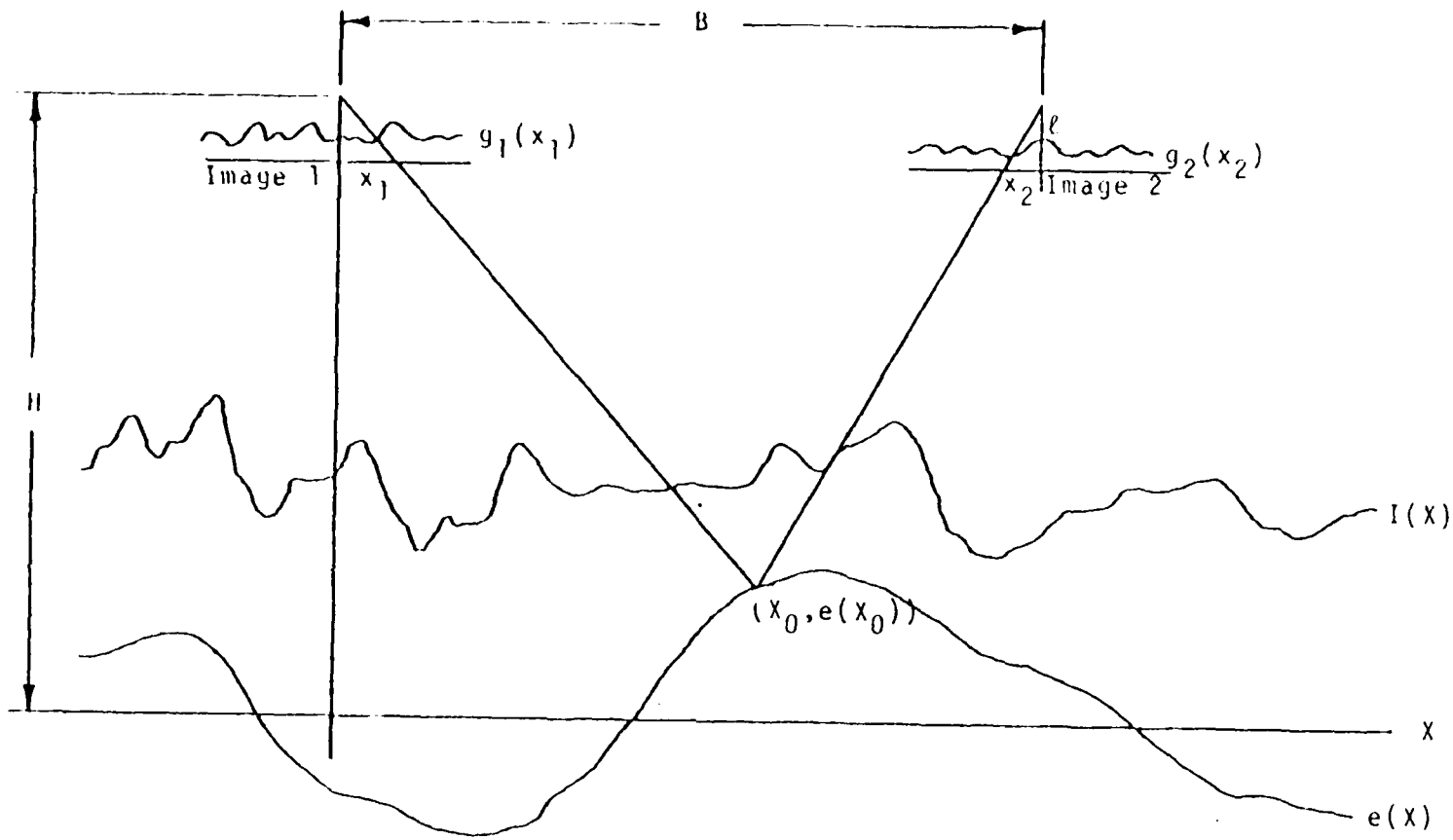


Figure 1-2. Image Formation Geometry

ℓ and the translation between images is the air base B. Two coordinate variables, x_1 and x_2, are ascribed to the two image planes. The geometry in Figure 1-2 is a so-called "geometric positive"; the true image lies behind the focal point of the camera and an inverted (geometric negative) image is formed. By reflecting the image through the focal point a geometric positive lying in front of the camera is constructed and has the advantage of having the same sense of coordinate algebraic signs as the ground coordinate system.

In practice, the acquisition of stereo-photographs rarely results in images which are precisely vertical. The rectification processes which correct for differences in altitude and orientation of the sensor stations are the domain of photogrammetry [Thompson 1966] and no further consideration is given them here.

It is assumed, then, that the following requirements are fulfilled in order to obtain "perfect" photographs.

(a) The optical axis of the camera lens is exactly vertical at the instant of exposure.

(b) There is no forward movement of the aircraft relative to the ground during the exposure time.

(c) The camera is free of distortion.

(d) Atmospheric conditions are ideal.

It is obvious that there is no practical possibility of regularly meeting these requirements. Fortunately, photogrammetrists have developed corrective procedures which relegate these effects as minor when compared to the contaminating effects discussed in Chapter 4.

Given the geometry of Figure 1-2, the coordinates x_1 and x_2 of the two images of a single ground point located at a distance X from the nadir point of image 1 can be determined by applying the rules of similar triangles. From Figure 1-2 it is clear that

$$x_1 = \frac{X\ell}{H - e(X)} \; ; \; x_2 = \frac{(X - B)\ell}{H - e(X)} \quad . \tag{1-1}$$

The accumulation of image points forms the intensity patterns observed in the image planes. The two intensity (or density) patterns, $g_1(x_1)$ and $g_2(x_2)$, correspond to the projection of the intensity pattern $I(X)$ into the two image planes. Assuming that the reflectance of the object field is independent of the camera station, the procedure for constructing an intensity pattern in the image from an intensity pattern on the ground can be characterized by the expressions

$$g_1(x_1) = I\left[\frac{x_1(H - e(X))}{\ell}\right] \tag{1-2a}$$

$$g_2(x_2) = I\left[B + \frac{x_2(H - e(X))}{\ell}\right] \tag{1-2b}$$

where x_1 and x_2 are the projections of point X.

Equation (1-1) can also be solved for e(X):

$$e(X) = H - \frac{B\ell}{x_1 - x_2} \qquad (1\text{-}3)$$

which shows that the elevation information is encoded in Equation (1-1) in the form of the difference in coordinates (parallax) of the conjugate points x_1 and x_2. The goal of stereo-compilation systems, then, is to accurately identify conjugate points.

The effects of a parallax error on the computed ground coordinates is shown in Figure 1-3. In a digital system, the images are sampled to produce discrete resolution elements (pixels). Assuming such a system, we substitute for the parallax in (1-3), $x_1 - x_2$, the parallax in terms of the number of pixels

$$x_1 - x_2 = i\Delta x - j\Delta x = (i - j)\Delta x \qquad (1\text{-}4)$$

where Δx is the sampling distance. From (1-3),

$$e(X) = H - \frac{B\ell}{(i - j)\Delta x} \quad . \qquad (1\text{-}5)$$

If the parallax is in error by ε pixels, the elevation error is given by

$$\delta e = \frac{- B\ell\varepsilon}{[(i - j) - \varepsilon](i - j)\Delta x} \quad . \qquad (1\text{-}6)$$

It is usually the case that the parallax errors are small with respect to the actual magnitude of the parallax, i.e., $(i - j) >> \varepsilon$. Thus (1-5) and (1-6) can be combined to give

$$\delta e = \frac{-\Delta x[H - e(X)]^2\varepsilon}{B\ell} \quad . \qquad (1\text{-}7)$$

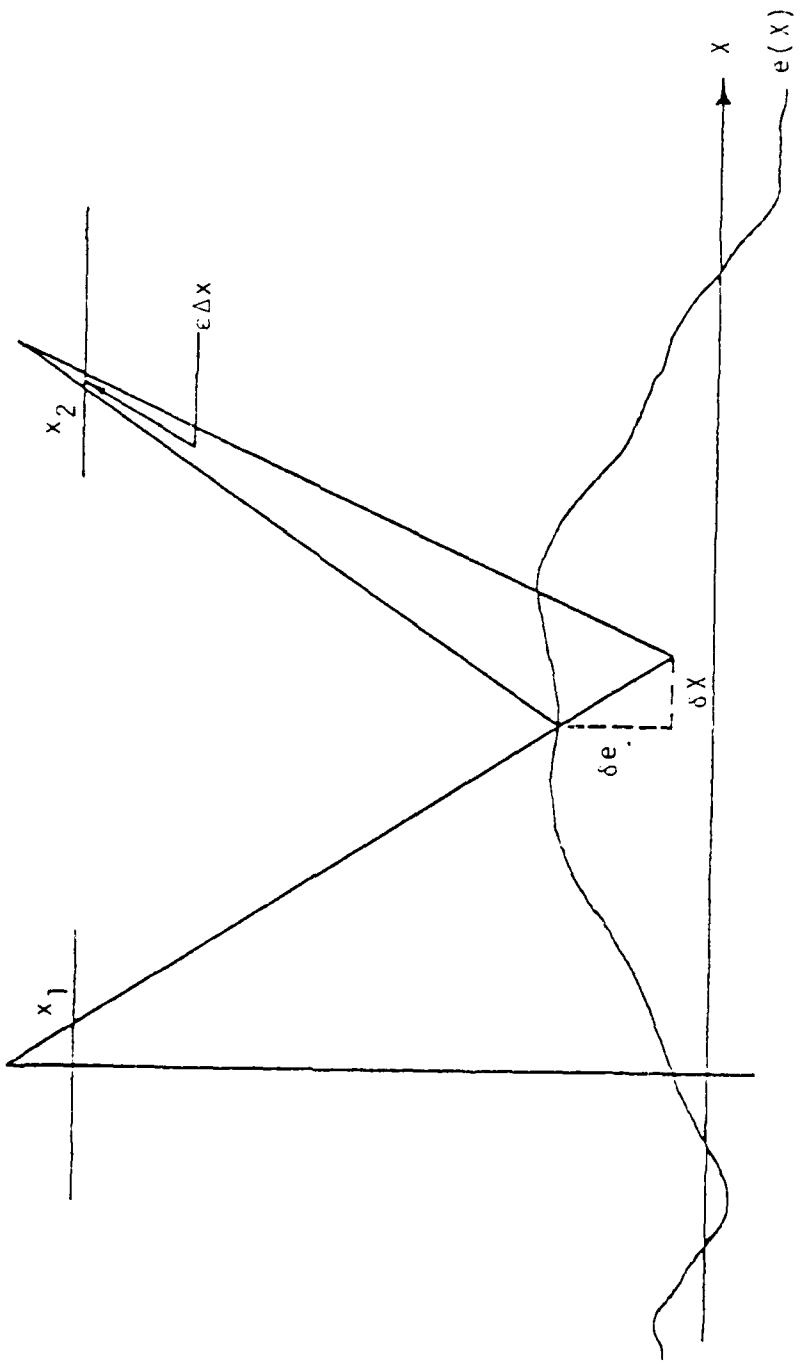


Figure 1-3. Errors in Terrain Coordinates Due to Parallax Computation Error

Equation (1-7) specifies the elevation resolution limit of the digital stereo-compiler. That is, the minimum detectable change in elevation is given by

$$\Delta e_{min} \triangleq \left|\frac{\delta e}{\epsilon}\right| = \frac{\Delta x[H - e(X)]^2}{B\ell} \quad . \qquad (1\text{-}8)$$

While the parameters H, B, and ℓ can be chosen during the data acquisition stage, the only parameter available to the processor is Δx, the sampling interval. Some image registration algorithms, however, attempt to improve on this resolution limit by interpolating the sampled correlation function in the vicinity of the peak [Pearson, et al. 1977]. This procedure is generally avoided in digital stereo-compilation systems due to the increased computational load. Correlation errors also result in ground position errors denoted by δX in Figure 1-3. Applying similar triangles to the geometry of Figure 1-3 reveals that

$$\frac{\delta X}{\delta e} = \frac{-i\Delta x}{\ell} \quad . \qquad (1\text{-}9)$$

Because of the two dimensional terrain error, the performance of stereo-compilation systems is expressed in terms of the parallax error.

Most of the stereo-compilation systems currently in use were designed primarily for producing contour and profile outputs. Recently, however, there has been an increasing need for topographic data in the form of digital grid-point elevations. Although many of the existing systems can be used to obtain such data, the AS-11B-X stereo-mapper was

developed specifically for this purpose [Brumm, et al. 1976]. As a result of improved image scanning techniques and highly parallel high speed digital processes, the AS-11B-X is capable of compilation speeds on the order of 10 to 50 times faster than for conventional systems with comparable accuracy. One of the features of the AS-11B-X which contributes to the increase in throughput rate is that image-intensity (or density) is scanned and measured on each photograph along corresponding epipolar lines. Epipolar lines are the lines of intersection between the photographs and the set of epipolar planes defined by the focal points of the two camera stations as shown in Figure 1-4. With such a scanning technique, the correlation function need only be one-dimensional since conjugate points lie on corresponding epipolar lines. Any residual parallax in the direction perpendicular to the epipolar lines is considered to be negligible [Brumm, et al. 1976].

Thesis Outline

In Chapter 2, we present a number of correlation techniques that are commonly employed in image matching applications and indicate the theoretical justification for their use. It is shown that the maximum likelihood similarity measure takes the form of covariance (matched filter), normalized covariance, or least squares, depending on the a-priori assumptions about the image formation model.

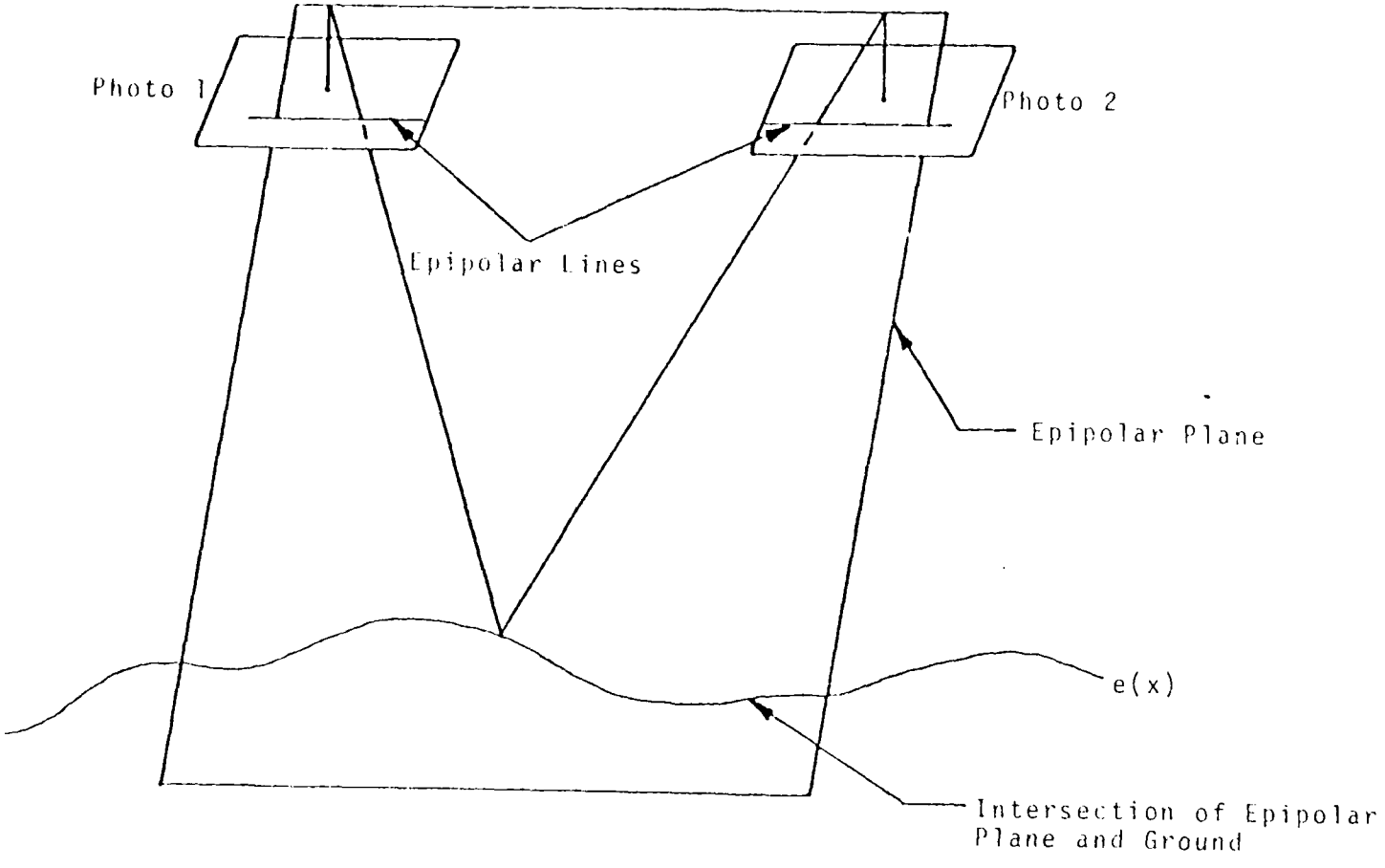


Figure 1-4. Epipolar Scan Lines

The Cramer-Rao lower bound on the accuracy of any unbiased estimator has been derived [Van Trees 1968] in the context of radar target range determination by estimation of signal round-trip time delay. A similar bound has been developed [Knapp and Carter 1976] in the context of estimating time delay between received signals at two spatially separated sensors used in passive sonar systems. In Chapter 3, these results are presented and compared and their application to the prediction of errors in stereo-compilation is discussed. The form of the bound derived by Knapp and Carter [1976] is shown to apply to the image formation model assumed and is thus considered as an image quality measure.

It is impossible, except under severely restrictive conditions, to derive a general expression for the probability of a false acquisition. Attempting such a derivation, however, leads to an image quality measure which is similar to the "figure of merit" ascribed to correlator performance by Webber and Delashmit [1974a]. This figure of merit is shown to be a function of both image area and local signal-to-noise ratio.

In order to employ the figure of merit as an image quality measure, it is necessary to determine the correlator output signal-to-noise ratio. In Chapter 4, the covariance function is analyzed in detail to determine the degree to which sensor noise, quantization noise, "self" noise, and machine noise corrupt correlator performance. Models for

computer round-off errors are developed for both fixed point and floating point correlation processors. The expressions developed for the purpose of predicting correlator performance all require the assumption of input data (image) stationarity. Chapter 4 concludes with a discussion of the effects of input data non-stationarity on correlator performance. Two schemes for detecting such regions are presented and the conditions under which such regions are considered "defective" are discussed. It is shown that the covariance processor is particularly susceptible to data non-stationarity.

In Chapter 5 we consider the effects of terrain variability on correlator performance and develop a model for "relief distortion" which is described by the function required to map an image region onto its conjugate. Under simplifying assumptions, this mapping is shown to be a spatial contraction or expansion in the direction parallel to epipolar lines. The distortion modelled by this affine transformation has been analyzed in detail by Mostafavi and Smith [1978a, 1978b]. We review these results and discuss the conditions under which they apply to the stereo-compilation problem. Of particular interest here, however, is the degree to which the effects of distortion on correlation accuracy can be predicted. To this end, an "image overlay quality" measure is developed in order to predict the statistics of a parameter used to describe distortion within a region of imagery. Due to the assumption required

to make the problem tractable and the computational complexity of the resulting procedure, it is concluded that pre-processing alone is unlikely to provide all the needed information for accurate error prediction but that a combination of pre-processing and on-line (i.e., during stereo-compilation) algorithms has the potential for detecting problem regions and increasing the overall system throughput rate.

Chapter 6 is concerned with the design and implementation of computer algorithms to assess the utility of the image quality measures developed in previous sections in terms of the degree to which image "quality" is indicative of the correlation behavior. Here we look at the image quality measures based on the Cramer-Rao bound and false acquisition probabilities as well as a number of other image features based on contrast measures such as variance and contrast modulation which appear to have potential for predicting error prone imagery. Results indicate that a measure based on the false acquisition probability appears to hold the greatest promise for predicting correlation behavior althougn, with few exceptions, the quality measures are remarkably similar in detection performance.

One dimensional correlation simulations in the presence of relief distortion indicate that extreme relief can, indeed, be a source of correlation error. Although some stereo-compilation systems correct for geometric

distortions described by affine transformations, we show that computational savings may be gained by performing these corrections only under certain conditions which can be partially specified in a pre-processing mode.

Finally, in Chapter 7, the implementation, both optical and digital, of promising pre-processing algorithms is considered. Also considered are image enhancement procedures which can be used to reduce the effects of noise on correlation behavior. Initial experiments indicate that these noise suppression operations can significantly improve correlation accuracy.

CHAPTER 2

CORRELATION TECHNIQUES

Numerous correlation techniques, both analog and digital, have been proposed for the purpose of image registration. None of these is equivalent to stereo-viewing as accomplished by human observers. Owing to his adaptability, the human observer is far superior. The interest in developing automatic systems, however, has led to a variety of computational procedures. Some of these procedures have a sound theoretical basis whereas others are at least partly ad hoc. In this chapter, we present the theoretical framework for some of the more common digital procedures and briefly discuss the rational behind the development of the ad hoc techniques.

Correlation Geometry and Notation

Let two images, S the search region and R the match window, be defined as in Figure 2-1. S is an L x K array of digital picture elements (pixels) and R is an M x N array such that $M \leq L$ and $N \leq K$. It is assumed that enough a-priori information is known about the search and match windows so that L, K, M, and N may be selected with the guarantee that, at registration, a complete subimage is

contained within the search area. In stereo-compilation systems, there are two mechanisms for establishing such a guarantee. In a "non-tracking" correlator, the search limits can be set by a-priori knowledge of the maximum and minimum elevation over the region of interest as indicated in Figure 2-2. In a tracking correlator, however, the search can be reduced by predicting the location of the correlation peak based on previously computed parallax values and computing the correlation function at several positions on either side of the predicted location. Because of the increased throughput rate achievable by reduction in search extent, the latter mode is commonly employed in present stereo-compilers. As mentioned previously, however, this mode has the disadvantage of being unable, in general, to recover from a "lost" condition without aid from an operator.

The various correlation procedures search over the allowed range of reference points to find a point which indicates a subimage that is most similar to the given window. In some applications (e.g., stereo-compilation with epipolar scanning), the search is one-dimensional (M = L in Figure 2-1). To simplify notation we will assume that the correlation functions are one-dimensional. The extension to two-dimensional searches is straighforward. Furthermore, we will denote the match window R by defining $\underset{\sim}{r}$ to be a vector consisting of the raster scanned elements of R, that is

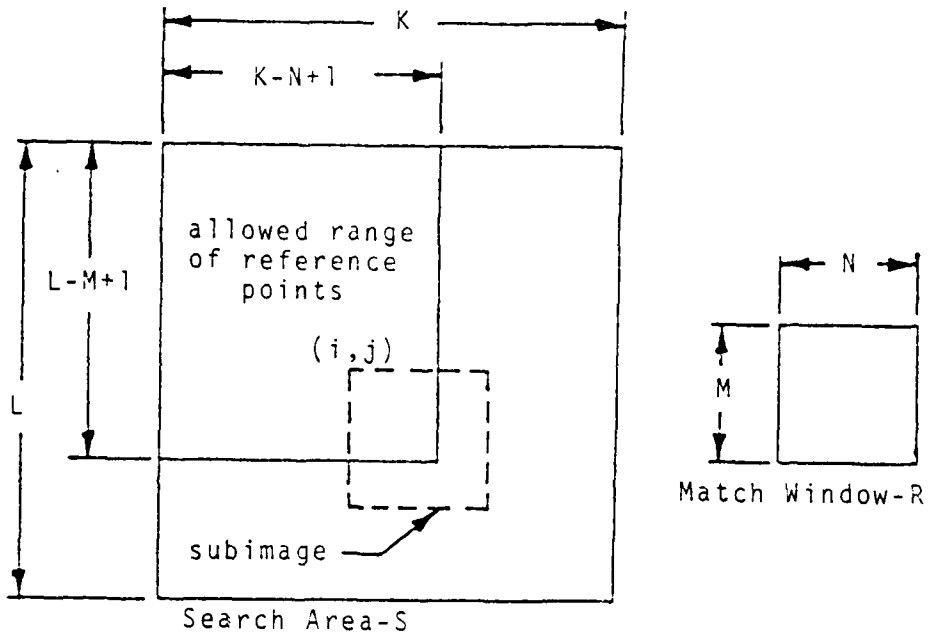


Figure 2-1. Correlation Search Space

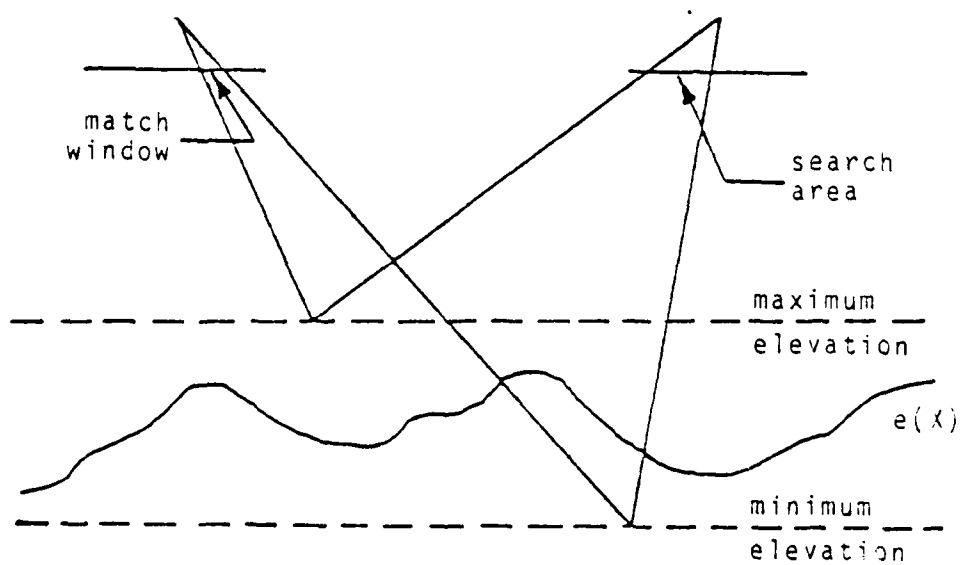


Figure 2-2. Establishing Correlation Limits in a Non-Tracking Correlator

$$\underset{\sim}{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{NM} \end{bmatrix} = \begin{bmatrix} R_{1,1} \\ R_{1,2} \\ \vdots \\ R_{M,N} \end{bmatrix} .$$

The various subimages of the search window will be denoted in a similar manner by $\underset{\sim}{s}^k, k = 1,\ldots,T$ where the superscript indicates the location of the subimage and T is the number of allowed reference points (taps[1]).

In general, one cannot expect any of the subimages $\underset{\sim}{s}^k$ to be identical to the match window $\underset{\sim}{r}$. Assuming the ideal image acquisition scenario described in Chapter 1, the primary sources of the difference between $\underset{\sim}{r}$ and the "correct" subimage, $\underset{\sim}{s}^c$, are film-grain noise (or shot noise/photon noise in a photoelectric sensor), scanning noise, and distortion due to terrain relief. Film-grain noise is a direct consequence of image acquisition, sensing and recording; it is a random quantity injected into both images. Although this type of noise is weakly signal dependent [Andrews and Hunt 1977], it is common procedure to characterize its behavior by the zero mean Gaussian probability density function (pdf):

[1] There is a correlator output associated with each reference point or relative location k. The term "tap" historically refers to the wire tap on a delay-line correlator.

$$p(d) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left[-\frac{1}{2}\frac{d^2}{\sigma_n^2}\right] \qquad (2\text{-}1)$$

where σ_n^2 is the noise variance. It is also assumed that the noise process is spatially uncorrelated, with a uniform ("white") power spectrum. Scanning noise will also be modelled as in (2-1).

The effects of terrain relief are not so easily characterized. For the present, we assume that the effects of relief are negligible. In Chapter 5, we discuss the conditions under which this assumption is valid.

If we let $\underset{\sim}{g}$ be the underlying image intensity or density sequence obtained via the projection relationship in (1-2), then a model for the effects of film-grain noise (and any other additive noise) is given by

$$\begin{aligned} \underset{\sim}{r} &= \underset{\sim}{g} + \underset{\sim}{n}_r \\ \underset{\sim}{s}^c &= \underset{\sim}{g} + \underset{\sim}{n}_s \end{aligned} \qquad (2\text{-}2)$$

where $\underset{\sim}{n}_r$ is the noise sequence contaminating the match window and $\underset{\sim}{n}_s$ is the sequence contaminating the <u>correct</u> subimage. From (2-2) it is easily seen that

$$\underset{\sim}{r} = \underset{\sim}{s}^c + \underset{\sim}{n}_r - \underset{\sim}{n}_s \quad . \qquad (2\text{-}3)$$

If we lump the noise terms together, i.e., replace $\underset{\sim}{n}_r - \underset{\sim}{n}_s$ by $\underset{\sim}{n}^c$, then (2-3) becomes

$$\underset{\sim}{r} = \underset{\sim}{s}^c + \underset{\sim}{n}^c \quad . \qquad (2\text{-}4)$$

Before proceeding, we note that $\underset{\sim}{n}^c$ and $\underset{\sim}{s}^c$ are correlated through Equation (2-2). Since we do not know which of the T subimages is $\underset{\sim}{s}^c$, each subimage $\underset{\sim}{s}^k$ must be tested. In the following section, it is shown that some of the more common correlation procedures are maximum likelihood parallax estimates.

Maximum Likelihood Correlation Procedures

If we consider $\underset{\sim}{r}$ to be a signal obtained at the receiver end of a transmission channel, then the set of subimages $\{\underset{\sim}{s}^k\}$ can be viewed as the set of possible transmitted messages. Given $\underset{\sim}{r}$, we wish to determine which transmitted signal $\underset{\sim}{s}^k$, k = 1,...,T most closely resembles $\underset{\sim}{r}$. The maximum likelihood (ML) processor selects the signal which maximizes the conditional pdf $p(\underset{\sim}{r}|\underset{\sim}{s}^k)$. As we will see, the form of the processor depends on the assumptions made about the set of signals $\{\underset{\sim}{s}^k\}$ as well as the properties of the "transmission channel". In order to generalize (2-2), we introduce the parameters a_c and $\underset{\sim}{m}_c$ in such a way that

$$\begin{aligned} \underset{\sim}{r} &= a_c \underset{\sim}{g} + \underset{\sim}{n}_r \\ \underset{\sim}{s}^c &= \underset{\sim}{g} + \underset{\sim}{m}_c + \underset{\sim}{n}_c \end{aligned} \qquad (2\text{-}5)$$

The parameter a_c can be viewed as a scale factor introduced to account for differences in image contrast and $\underset{\sim}{m}_c$ is a bias vector modelling possible differences in mean density level. We will assume m_c consists of identical

elements of value m. Rewriting (2-5) as in (2-3), we have

$$\underset{\sim}{r} = a_c(\underset{\sim}{s}^c - \underset{\sim}{m}_c - \underset{\sim}{n}_c) + \underset{\sim}{n}_r \quad . \tag{2-6}$$

The maximum likelihood problem can now be formulated in terms of the T hypotheses:

$$\begin{aligned} H_1 &: \underset{\sim}{r} = a_1(\underset{\sim}{s}^1 - \underset{\sim}{m}_1 - \underset{\sim}{n}_1) + \underset{\sim}{n}_r \\ H_2 &: \underset{\sim}{r} = a_2(\underset{\sim}{s}^2 - \underset{\sim}{m}_2 - \underset{\sim}{n}_2) + \underset{\sim}{n}_r \\ &\vdots \\ H_T &: \underset{\sim}{r} = a_T(\underset{\sim}{s}^T - \underset{\sim}{m}_T - \underset{\sim}{n}_T) + \underset{\sim}{n}_r \quad . \end{aligned} \tag{2-7}$$

Before proceeding with the determination of the ML processor, we need to obtain appropriate expressions for a_k and $\underset{\sim}{m}_k$. To determine $\underset{\sim}{m}_k$, we denote by r_i and s_i^k the i^{th} elements of the vectors $\underset{\sim}{r}$ and $\underset{\sim}{s}^k$ respectively. Taking expectations with respect to these elements gives

$$E\{r_i\} = a_k(E\{s_i^k\} - m) \quad . \tag{2-8}$$

If we estimate the ensemble means by computing the sample average over the appropriate region, then

$$m \cong \bar{s}^k - \frac{1}{a_k}\bar{r} \tag{2-9}$$

and (2-7) becomes

$$H_k : \underset{\sim}{r} - \bar{\underset{\sim}{r}} = a_k(\underset{\sim}{s}^k - \bar{\underset{\sim}{s}}^k - \underset{\sim}{n}_k) + \underset{\sim}{n}_r \quad . \tag{2-10}$$

To obtain a_k we compute

$$E\{(r-\bar{r})'(r-\bar{r})\} = a_k^2 E\{(s^k-\bar{s}^k-n_k)'(s^k-\bar{s}^k-n_k)\} + E\{n_r' n_r\} \tag{2-11}$$

where orthogonality eliminates the cross terms on the right hand side. If the vectors, $\underset{\sim}{r}$, $\underset{\sim}{s}^k$, etcetera consist of N elements, then the expectations in (2-11) can be estimated by computing sample variances over appropriate regions, i.e.,

$$N\hat{\sigma}_r^2 \cong a_k^2[N\hat{\sigma}_s^2 - 2E\{(\underset{\sim}{s}^k - \bar{\underset{\sim}{s}}^k)'\underset{\sim}{n}_k\} + N\hat{\sigma}_n^2] + N\hat{\sigma}_n^2 \qquad (2\text{-}12)$$

where $\hat{\sigma}_r^2$, $\hat{\sigma}_s^2$, and $\hat{\sigma}_n^2$ are sample variances and N is the total number of elements in the match window. The remaining expectations in (2-12) can be evaluated with the help of (2-5), i.e.,

$$E\{(\underset{\sim}{s}^k)'\underset{\sim}{n}_k\} = \frac{1}{a_k}E\{\underset{\sim}{g}'\underset{\sim}{n}_k\} + E\{\underset{\sim}{m}_k'\underset{\sim}{n}_k\} + E\{\underset{\sim}{n}_k'\underset{\sim}{n}_k\} = N\hat{\sigma}_n^2 \quad . \qquad (2\text{-}13)$$

Thus (2-12) reduces to

$$\hat{\sigma}_r^2 \cong a_k^2\,\hat{\sigma}_s^2 - \hat{\sigma}_n^2 + \hat{\sigma}_n^2 \qquad (2\text{-}14)$$

from which it follows that

$$a_k^2 \cong \frac{\hat{\sigma}_r^2 - \hat{\sigma}_n^2}{\hat{\sigma}_s^2 - \hat{\sigma}_n^2} \quad . \qquad (2\text{-}15)$$

For high quality stereo-pair imagery, it is usually the case that $\hat{\sigma}_r^2 >> \hat{\sigma}_n^2$ and $\hat{\sigma}_s^2 >> \hat{\sigma}_n^2$ (signal-to-noise ratios of 30dB are typical). Thus,

$$a_k^2 \cong \frac{\hat{\sigma}_r^2}{\hat{\sigma}_s^2} = \frac{(\underset{\sim}{r} - \bar{\underset{\sim}{r}})'(\underset{\sim}{r} - \bar{\underset{\sim}{r}})}{(\underset{\sim}{s}^k - \bar{\underset{\sim}{s}}^k)'(\underset{\sim}{s}^k - \bar{\underset{\sim}{s}}^k)} \quad . \qquad (2\text{-}16)$$

The problem of determining $p(\underset{\sim}{r}|\underset{\sim}{s}^k)$ is complicated by the fact that the noise in (2-10) is signal dependent. Useful results are obtained, however, by assuming that $\underset{\sim}{s}^k$ and $\underset{\sim}{n}^k$ are statistically independent so we proceed along this path for the present. Under this assumption, (2-10) is rewritten as

$$H_k : \underset{\sim}{r} - \underset{\sim}{\bar{r}} = a_k(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k) + \underset{\sim}{n}^k \qquad (2\text{-}17)$$

where

$$\underset{\sim}{n}^k = \underset{\sim}{n}_r - a_k\underset{\sim}{n}_k \qquad (2\text{-}18)$$

Since the only random quantity in (2-17) is n^k ($\underset{\sim}{r}$ and $\underset{\sim}{s}^k$ are known samples of a random process), then it follows that

$$p(\underset{\sim}{r}|\underset{\sim}{s}^k) = \frac{1}{(2\pi)^{\frac{N}{2}}\sigma_{nk}^N}\exp[-\frac{1}{2}(\underset{\sim}{r}-\underset{\sim}{\bar{r}}-a_k(\underset{\sim}{s}^k-\underset{\sim}{\bar{s}}^k))'C_n^{-1}(\underset{\sim}{r}-\underset{\sim}{\bar{r}}-a_k(\underset{\sim}{s}^k-\underset{\sim}{\bar{s}}^k))] \qquad (2\text{-}19)$$

where $C_n = \sigma_{nk}^2 I$ is the noise covariance matrix (I = identity matrix) and from (2-18), $\sigma_{nk}^2 = (1 + a_k^2)\,\sigma_n^2$. The selection of the hypothesis, H_i, to maximize (2-19) is known [Whalen 1971], [Van Trees 1968] to be the hypothesis which maximizes the quantity

$$c_k = 2a_k(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k) - a_k^2(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k) \qquad (2\text{-}20)$$

Here $(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)$ is simply the cross-covariance of $\underset{\sim}{r}$ with $\underset{\sim}{s}^k$ and $(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)$ is the energy of subimage $\underset{\sim}{s}^k$. If we make the a-priori assumption that $a_k = 1$ for all k, then the ML processor computes the quantity

$$c_k = 2(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k) - (\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k) \qquad (2\text{-}21)$$

and selects the hypothesis which maximizes c_k. Since $(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{r} - \underset{\sim}{\bar{r}})$ is independent of the hypothesis we could just as well compute

$$c_k = (\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{r} - \underset{\sim}{\bar{r}}) - 2(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k) + (\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)$$

$$= [(\underset{\sim}{r} - \underset{\sim}{\bar{r}}) - (\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)]'[(\underset{\sim}{r} - \underset{\sim}{\bar{r}}) - (\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)] \qquad (2\text{-}22)$$

and select the hypothesis H_k to minimize c_k. Thus, the assumption that $a_k = 1$ leads to the "smallest sum of squared differences" processor which selects the hypothesis which provides the best least-squares fit to $\underset{\sim}{r}$.

If the assumption is also made that the subimage energy is independent of position (or varies slowly with position) then the ML processor reduces to the simple covariance function

$$c_k = (\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k) \quad . \qquad (2\text{-}23)$$

For the most general case, $a_k \neq 1 \ \forall \ k$. Substituting (2-16) into (2-20) gives

$$c_k = \frac{2(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)\,[(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{r} - \underset{\sim}{\bar{r}})]^{\frac{1}{2}}}{[(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)]^{\frac{1}{2}}} - (\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{r} - \underset{\sim}{\bar{r}}) \quad . \qquad (2\text{-}24)$$

Eliminating quantities that are independent of the hypothesis leads to

$$c_k = \frac{(\underset{\sim}{r} - \underset{\sim}{\bar{r}})'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)}{[(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)'(\underset{\sim}{s}^k - \underset{\sim}{\bar{s}}^k)]^{\frac{1}{2}}} \quad . \qquad (2\text{-}25)$$

As before, we could just as well compute

$$c_k = \frac{(\underset{\sim}{r} - \bar{\underset{\sim}{r}})'(\underset{\sim}{s}^k - \bar{\underset{\sim}{s}}^k)}{\{[(\underset{\sim}{r} - \bar{\underset{\sim}{r}})'(\underset{\sim}{r} - \bar{\underset{\sim}{r}})][(\underset{\sim}{s}^k - \bar{\underset{\sim}{s}}^k)'(\underset{\sim}{s}^k - \bar{\underset{\sim}{s}}^k)]\}^{\frac{1}{2}}} \quad (2\text{-}26)$$

which is the well known normalized covariance function which selects the hypothesis which maximizes the correlation coefficient between $\underset{\sim}{s}^k$ and $\underset{\sim}{r}$. .

Because of the computational simplicity of (2-23), simple covariance has become the workhorse of present day stereo-compilation equipment. As we will see later, however, violations of the assumptions leading to (2-23) are cause for concern. Normalized covariance is commonly employed in applications having less demanding requirements on throughput rate although it is used by the recently developed heterodyne optical correlator [Balasubramanian 1976]. In the following section, the image matching problem is formulated in a slightly different manner and the corresponding ML processor is subsequently different.

Generalized Correlation

Since $\underset{\sim}{s}^k$ can be obtained from $\underset{\sim}{s}^{k-1}$ simply by a one pixel shift in subimage position, the matching problem can be reposed as

$$R(i,j) = S(i + D_x, j+D_y)\ \mathrm{rect}(\tfrac{i}{M}, \tfrac{j}{N}) + N(i,j) \quad (2\text{-}27)$$

where D_x and D_y are the relative shifts between the match windows R and the correct subimage of S, N(i,j) is the signal

dependent noise as before, and

$$\text{rect}\left(\frac{i}{M},\frac{j}{N}\right) \triangleq \begin{cases} 1 & |i| \leq \frac{M}{2},\ |j| \leq \frac{N}{2} \\ 0 & \text{otherwise} \end{cases} . \tag{2-28}$$

With this formulation, the ML processor selects the parameters $\hat{D}_x$ and $\hat{D}_y$ which maximize $p(R|D_x,D_y)$. As we shall see, the ML estimator can be realized as a pair of image prefilters followed by a cross-correlation as shown in Figure 2-3. A variety of "generalized correlation" techniques can be implemented in this manner with the choice of H_1 and H_2 based on the optimization of certain performance criteria. If $H_1(u,v) = H_2(u,v) = 1\ \forall\ u,v$ where u and v are the discrete horizontal and vertical spatial frequency variables, then the estimate $\hat{D}_x,\hat{D}_y$ is determined by the cross-covariance function

$$\hat{C}_{SR}(k,\ell) = \frac{1}{NM}\sum_{u=0}^{N-1}\sum_{v=0}^{M-1} \hat{G}_{SR}(u,v)W_N^{ku}W_M^{\ell v} \tag{2-29}$$

where

$$W_N \triangleq e^{j\frac{2\pi}{N}} . \tag{2-30}$$

Equation (2-29) is simply the inverse discrete Fourier transform of the estimated cross-power spectrum $\hat{G}_{SR}$ and the processor is the two-dimensional analog of (2-23). When S and R are filtered, however, the cross-power spectrum between the filtered outputs is given by

$$\hat{G}_{AB}(u,v) = H_1(u,v)H_2^*(u,v)\hat{G}_{SR}(u,v) \tag{2-31}$$

and the generalized correlation between S and R is

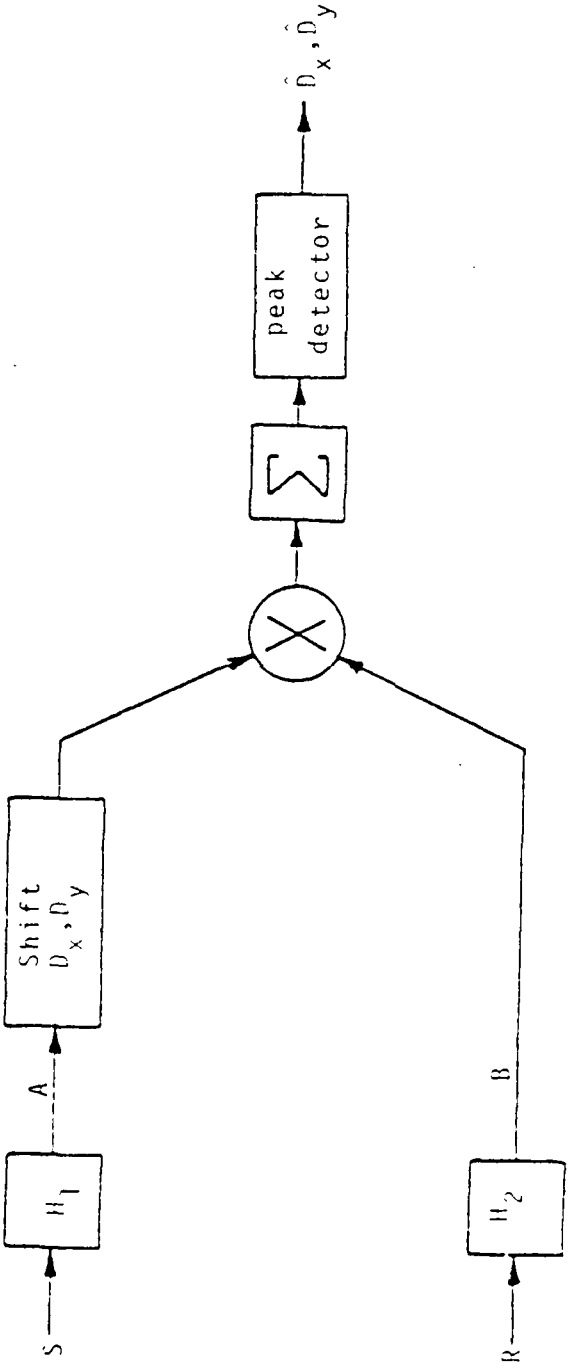


Figure 2-3. Generalized Correlator

$$\hat{C}_{AB}(k,\ell) = \frac{1}{NM} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} \psi(u,v)\hat{G}_{SR}(u,v)W^{km}W_M^{\ell v} \qquad (2\text{-}32)$$

where

$$\psi(u,v) = H_1(u,v)H_2^*(u,v) \quad . \qquad (2\text{-}33)$$

The ML estimator for D_x and D_y has been derived by Knapp and Carter [1976] in the context of determining time delay between signals received at two spatially separated sensors in the presence of uncorrelated noise. To avoid the problem of signal dependent noise encountered previously, Knapp and Carter compute the pdf $p(\underline{X};Q)$ where

$$\underline{X}(u,v) = \begin{bmatrix} R(u,v) \\ \\ S(u,v) \end{bmatrix} \triangleq \begin{bmatrix} F\{R(i,j)\} \\ \\ F\{S(i,j)\} \end{bmatrix} \qquad (2\text{-}34)$$

and Q is the power spectral density matrix

$$Q(k,\ell) \triangleq \begin{bmatrix} G_{RR}(k,\ell) & G_{RS}(k,\ell) \\ \\ G_{RS}^*(k,\ell) & G_{SS}(k,\ell) \end{bmatrix} \quad . \qquad (2\text{-}35)$$

The assumption is made that the elements of $\underline{X}(k,\ell)$ are uncorrelated Gaussian (thus statistically independent) random variables. This assumption requires that M and N in Figure 2-1 are large with respect D_x and D_y (this is not generally the case in stereo-compilation systems). The result, extended to two dimension, is given by

$$C_{AB}^{ML}(k,\ell) = \frac{1}{NM} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} \psi_{ML}(u,v)\hat{G}_{SR}(u,v)W_N^{ku}W_M^{\ell v} \qquad (2\text{-}36)$$

where

$$\Psi_{ML}(u,v) \triangleq \frac{1}{|G_{SR}(u,v)|} \frac{|\gamma_{SR}(u,v)|^2}{[1-|\gamma_{SR}(u,v)|^2]} \quad , \qquad (2\text{-}37)$$

and γ_{SR} is the coherence function

$$\gamma_{SR}(u,v) \triangleq \frac{\hat{G}_{SR}(u,v)}{\sqrt{\hat{G}_{SS}(u,v)\ \hat{G}_{RR}(u,v)}} \quad . \qquad (2\text{-}38)$$

Equation (2-36) can be interpreted as a phase estimator given by

$$\exp[j\hat{\theta}(u,v)] = \frac{\hat{G}_{SR}(u,v)}{|\hat{G}_{SR}(u,v)|} \qquad (2\text{-}39)$$

multiplied by a weighting function which weights the phase according to the strength of the coherence. To implement such a processor, Equations (2-38), (2-37), and (2-36) need to be computed for each match window location. This procedure is thus computationally prohibitive for the application at hand.

The correlation methods characterized by Figure 2-3 are all generalizations of (2-23). Normalized covariance can also be generalized in a similar manner, although the analytic development of the optimal processors is, in general, impossible. Pratt [1974] describes an image pre-processing procedure which simply "whitens" the images before performing the normalized covariance operation. A major consideration with any of these generalized techniques is the increase in computational complexity. To alleviate the computational load, a number of ad hoc correlation techniques have been proposed.

Ad Hoc Correlation Techniques

The techniques described below are included here primarily for completeness. They are not analyzed further.

Polarity Correlation. The input waveforms are quantized to two levels, then correlated as usual. The method is less reliable than covariance, particularly for small data sets, although it is far less demanding computationally.

Minimization of the Sum of Absolute Differences. With this method, averages are removed as in covariance, then the absolute values of the differences between the match window and subimage are summed. The position at which the sum is a minimum is considered to be the position of maximum correlation. This method is also inferior to covariance [Helava 1978] but is less demanding computationally and can produce useful results as in the following procedure.

Sequential Similarity Detection. Sequential similarity detection algorithms (SSDA's) are based on the fact that of the $(L - M + 1)(K - N + 1)$ reference positions shown in Figure 2-1, relatively few are near the correlation peak (this depends on the a-priori knowledge concerning the predicted location) and therefore only relatively few require high accuracy calculations. Furthermore, for a grossly mismatched windowing pair it may not be necessary to test all MN data pairs before rejecting the subimage as a possible match [Barnea & Silverman 1972]. Threshold algorithms are employed

in which an error sequence is generated using the previously described technique. If the accumulated error exceeds a threshold (fixed or variable) then that particular search window is rejected. Those positions which pass the threshold test are subjected to closer scrutiny, in some cases employing more reliable correlation algorithms to make final decisions.

Hierarchical Similarity Detection. A hierarchical approach which employs the SSDA technique just described has been developed by Wong and Hall [1978] in which a sequence of images is created, each of which is a filtered, subsampled version of the previous one. An SSDA is then employed on the filtered images to locate probable correlation peaks. Higher resolution images are employed at each stage to improve correlation reliability in the vicinity of the probable correlation peaks and the process continues until only one peak remains. This method is similar in concept to the multiband analog correlation technique which separates the image signal into several (usually two) bands. The smoothed imagery improves the "pull in" range (due to the broad correlation function) and the high frequency imagery provides the required correlation resolution [Thompson 1966].

The variety of correlation techniques makes it difficult to develop image quality measures which will reliably predict correlator performance for all the techniques. Since some methods are considerably more accurate than others, we

can only endeavor to tailor the image quality features to the particular correlation procedure chosen. In order to make the problem tractable, we will restrict our attention to a small number of techniques, in particular, covariance (2-23), normalized covariance (2-25) or (2-26), and least-squares (2-21). In the following section, the accuracy of these correlation schemes is assessed by computer simulation. The goal of the remaining chapters, then, will be to develop image quality measures which will predict the behavior of these models.

Correlation Processor Comparisons

In order to obtain valid comparisons of the image correlation techniques, it is necessary to have access to stereo-pair imagery for which the parallax values associated with each image point are known precisely, or equivalently, the corresponding elevation data is known precisely. It is also desirable to be able to control the image noise content as well as its spatial spectral characteristics (both intensity and elevation). In order to obtain the latter capability, it is necessary to generate purely simulated imagery. While we have taken this approach in part of the experimental results to be presented in Chapter 6, the imagery so generated lacks realism. The disadvantages of real imagery, however, are that the required parallax precision is not readily available

and secondly, real imagery is noise contaminated. Recent developments in digital sensor simulation have made available computer software able to produce images from a digital data base with the characteristics of photographs taken by a frame camera (Miknail, Unrun, and Alspaugh 1977). The data base used in the following simulations was created from actual stereo aerial photographs. The terrain model was developed using the Bendix AS-11B-1 analytical stereoplotter and consists of a large array of ground intensity and elevation data which correspond to samples taken at 15.75 ft. intervals. Using the software package SIM one can specify numerous image creation parameters (e.g., air base, altitude, focal length, tilt, pixel spacing, etc.) and introduce vertical scaling (Miknail, Unrun, and Alspaugh 1977). The resulting images can be manipulated (filtered, corrupted by noise, etc.) as desired. A stereo-pair generated using SIM is shown in Figure 2-4 along with the coded elevation data corresponding to the right image in Figure 2-4b. Pertinent image formation parameters are shown in Table 2-1.

The correlation procedures were performed by selecting a match window from the right image and a search area from the left image as shown in Figures 2-4a,b. Since maximum and minimum elevations are known in this particular model, the search window could be selected as shown in Figure 2-2. This guaranteed that the correlator will not "get lost." The image

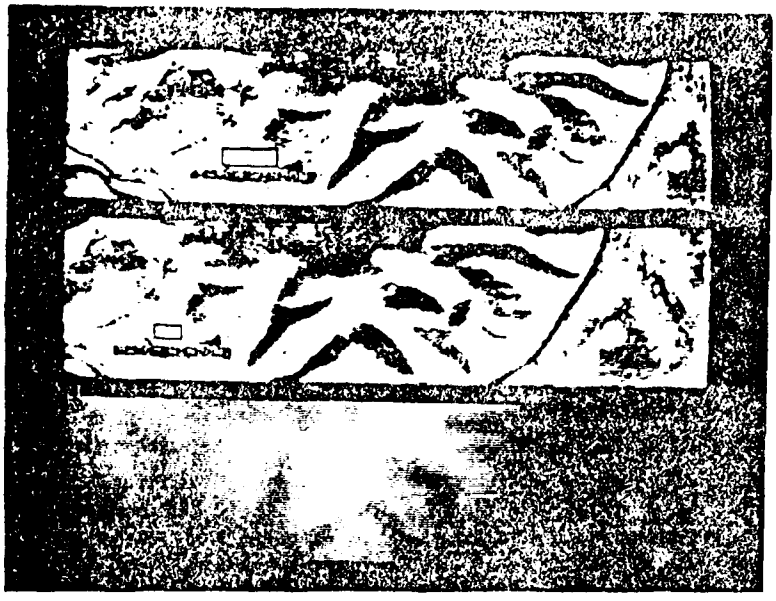


Figure 2-4. Synthetic Stereo-Pair and Elevation Data
(a) left image
(b) right image
(c) coded elevation

TABLE 2-1. Syntnetic Image Creation Parameters

Air Base	= 4724 ft.
Altitude (above sea level)	= 15984 ft.
Camera Focal Length	= 6 in.
Image Dimension	= 133 lines x 512 pixels
Pixel Spacing	= .15 mm
Elevation Range	= 1187 to 1481 ft.
Average Base/Height Ratio	= .33

lines in Figure 2-4 are epipolar lines of the stereo-pair so the correlation functions need only be one-dimensional.

Shown in Figures 2-5, 2-6, and 2-7 are the computed parallax matrices corresponding to covariance, least-squares, and normalized covariance respectively for match window sizes 3 x 7 (3 lines by 7 pixels/line), 3 x 15, and 3 x 23. The gray level (or intensity) displayed in these figures is linearly related to the parallax value. The source images for these results were noise free. Figure 2-8 shows the correct parallax matrix which appears as contours of the coded elevation data. The contours are the result of the resolution limit specified in Equation (1-8). For the image creation data specified in Table 2-1, it is approximately 44 ft. of elevation. Due to the large pixel spacing in the synthetic images, this resolution is far below the capabilities of present compilation systems. If, however, the performance of the correlation technique is measured in units of pixels, then the resolution of the correlation in terms of elevation may be ignored. The parallax error matrices can be determined by subtracting Figure 2-8 from the computed parallax matrices in Figure 2-5 to 2-7. The results are shown in Figure 2-9, 2-10, and 2-11 for each of the correlation methods.

A number of observations are in order. First, it is immediately clear from the correlation results that covariance processing is inferior to the other methods. Secondly, note

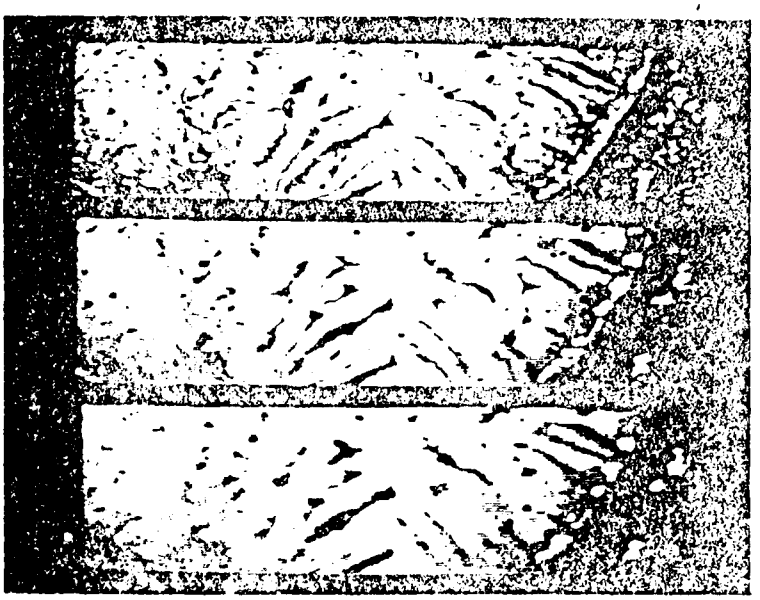


Figure 2-5. Computed Parallax-Covariance. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

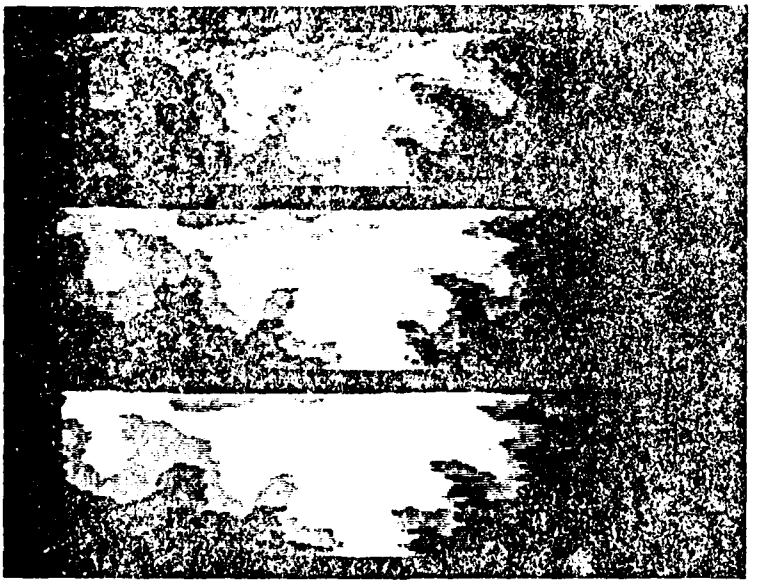


Figure 2-6. Computed Parallax-Least Squares. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

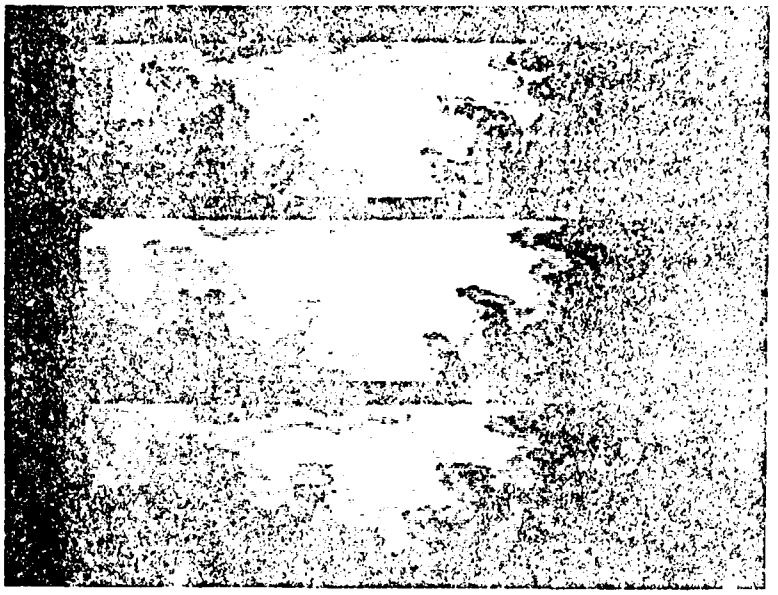


Figure 2-7. Computed Parallax-Normalized Covariance. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

Figure 2-8. Correct Parallax

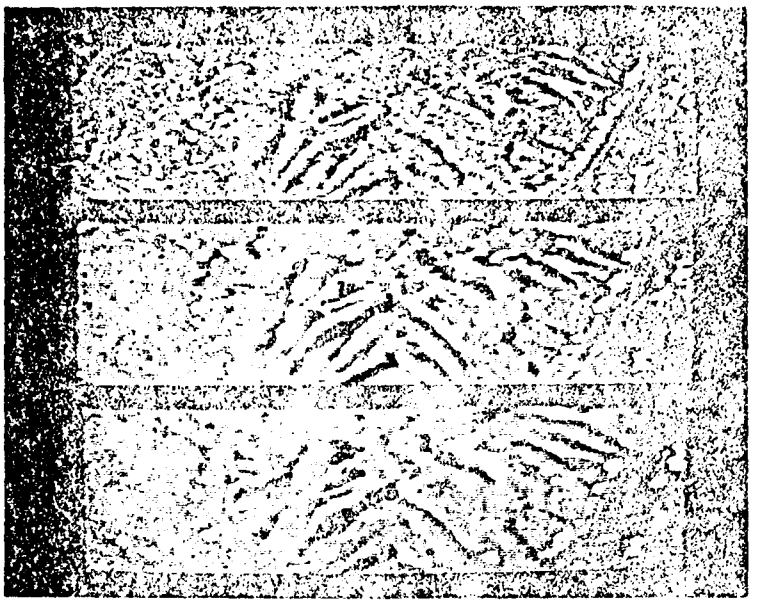


Figure 2-9. Error Maps-Covariance. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

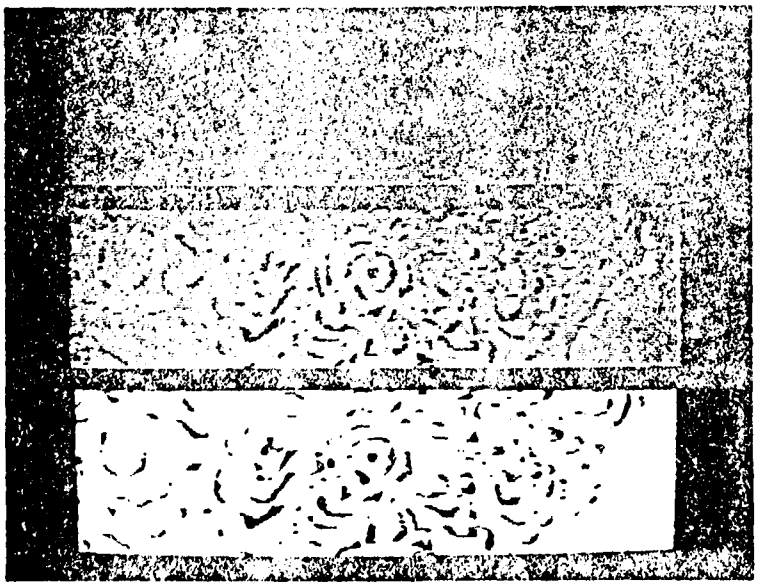


Figure 2-10. Error Maps-Least Squares. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

that many of the errors (most all with least squares and normalized covariance) lie along the contour lines of the parallax matrices. There are two possible reasons for this behavior. Since the correlation functions are sampled, it is possible that the true peak of the correlation function lies between two adjacent samples. This is precisely the case when the match window location corresponds to a region where the correct parallax is changing from one resolution element to another. It is not unreasonable to expect parallax errors of one pixel to occur in such a situation. The majority of errors lying along the contours are one-pixel errors. Also note, however, that the few larger errors that occur using least-squares and normalized covariance also tend to occur at these locations, since the peak of the correlation function is, in effect, suppressed. Another possible source of these errors could be the interpolation on the data base required to create the synthetic imagery. Because these errors are relatively small, they would not, in general, increase the probability of a "lost" condition in a tracking correlator. For this reason, these regions are treated in subsequent analysis as being 'no-error' regions.

Finally, we note the loss in correlator resolution with increased window length as seen by the irregular contours in the parallax matrices for the 3 x 23 window. This effect results from the fact that the computed parallax corresponds

to the mean elevation viewed through the match window. This effect is most noticeable in regions with steep terrain slopes.

The correlation experiment was repeated with noise contaminated imagery. Signal independent additive Gaussian white noise was introduced into each image of the stereo-pair resulting in an overall signal-to-noise ratio of 20dB. The resulting correlation parallax results are shown in Figures 2-12, 2-13, and 2-14. While the covariance parallax matrices appear to be almost unchanged (perhaps because of the inferior initial behavior), the degradation of performance for the normalized covariance and least squares processors is quite clearly observable. This chapter is concluded with a comparison of the computation complexity of each of these algorithms.

Comparison of Computational Efficiency

In order to properly compare the correlation algorithms, it is necessary to compare the computation complexity as well as accuracy since a trade-off generally results between these characteristics. Since many of the computations can be done in parallel with special purpose hardware it is difficult to compare algorithms in terms of total computation time. However, a common measure of computational complexity is the total number of real adds and multiplies required for a particular operation. This is the approach taken here. Since the least-squares procedure in (2-22) requires a two-pass procedure, one to compute the

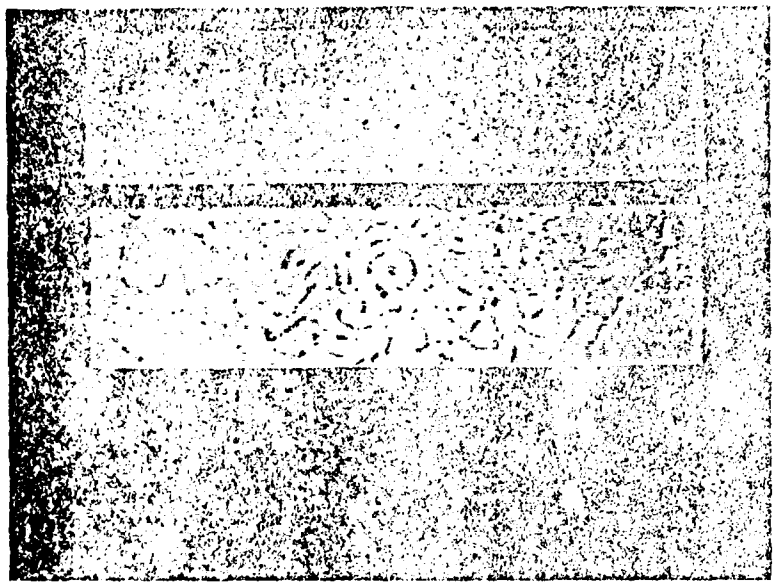


Figure 2-11. Error Maps-Normalized Covariance. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

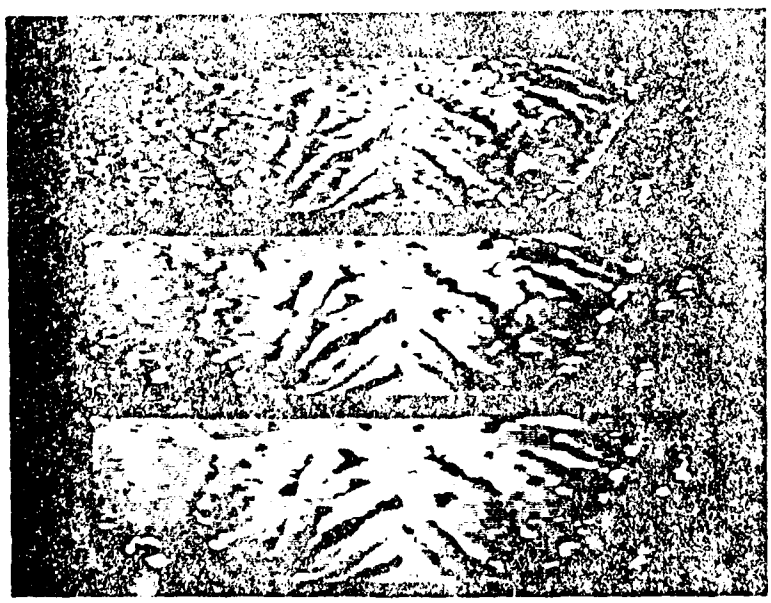


Figure 2-12. Computed Parallax-Covariance, SNR = 20 dB. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

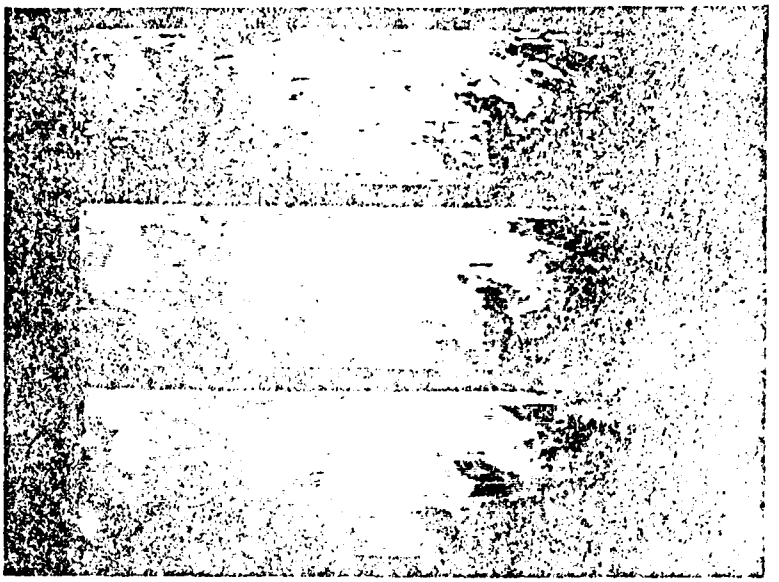


Figure 2-13. Computed Parallax-Least Squares SNR = 20 dB. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

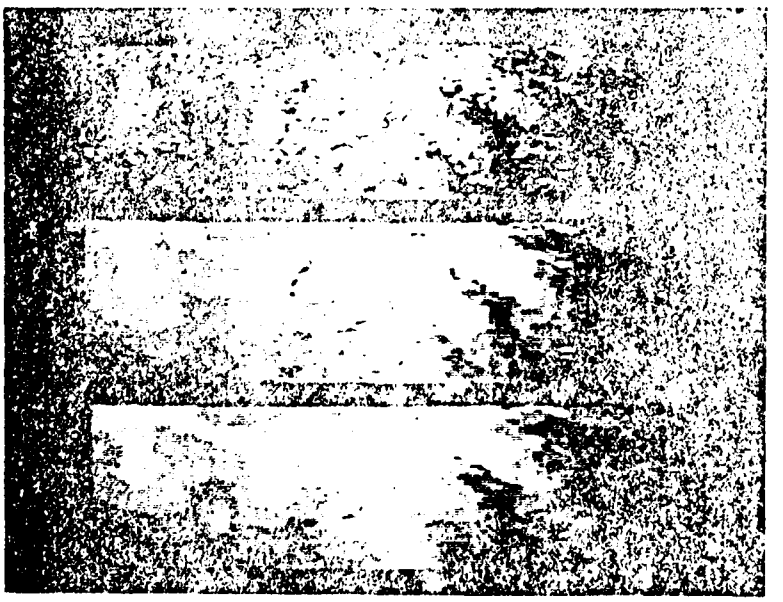


Figure 2-14. Computed Parallax-Normalized Covariance, SNR = 20 dB. Match Window Size: (a) 7 x 3, (b) 15 x 3, (c) 23 x 3

Adams and Goertzel [2-10] to the case of vector differences, we will use the result given in [2-11]. First, we rewrite each of the procedure algorithms in vector notation to explicitly show the number of terms of the process:

Covariance: $c_k = \underline{r} - \underline{\bar{r}}) \cdot (\underline{s}^k - \underline{\bar{s}}^k)$

$$= \sum_{j=1}^{N} r_j s_j^k - \frac{1}{N} \sum_{j=1}^{N} r_j \sum_{j=1}^{N} s_j^k \tag{2-40}$$

Least Squares: $c_k^{LS} = \sum (\underline{r} - \underline{\bar{r}}) \cdot (\underline{s}^k - \underline{\bar{s}}^k) - (\underline{s}^k - \underline{\bar{s}}^k) \cdot (\underline{s}^k - \underline{\bar{s}}^k)$

$$= 2c_k - \sum_{j=1}^{N} (s_j^k)^2 - \frac{1}{N} \left(\sum_{j=1}^{N} s_j^k \right)^2 \tag{2-41}$$

Normalized Covariance: $c_k^{NC} = \dfrac{(\underline{r} - \underline{\bar{r}}) \cdot (\underline{s}^k - \underline{\bar{s}}^k)}{[(\underline{s}^k - \underline{\bar{s}}^k) \cdot (\underline{s}^k - \underline{\bar{s}}^k)]^{\frac{1}{2}}}$

$$= \frac{c_k}{\left[\sum_{j=1}^{N} (s_j^k)^2 - \frac{1}{N} \left(\sum_{j=1}^{N} s_j^k \right)^2 \right]^{\frac{1}{2}}} \tag{2-42}$$

where N is the total number of elements in the match window, s_j^k denotes the j^{th} element in the k^{th} subimage and the superscripts LS and NC denote the correlation operation. Since $\sum_{j=1}^{N} r_j$ need be computed only once for each match window position, we will ignore its contribution to the total number of computations. From (2-41) and (2-42) it follows that c_k^{LS} and c_k^{NC} require essentially the same computational expense.

Table 2-2 shows the total number of adds and multiplies for each search area subimage location.

In image correlation, considerable savings can be obtained by storing the computational results from one subimage position to be used in the adjacent position. Figure 2-15 shows two adjacent subimages s^k and s^{k+1} with elements denoted by a_{ij} and dimension as indicated with LP = N. Note that:

$$(s^k)'s^k = \sum_{i=1}^{L} a_{i1}^2 + Q \qquad (2\text{-}43)$$

where

$$Q = \sum_{i=1}^{L} \sum_{j=2}^{P} a_{ij}^2 \qquad (2\text{-}44)$$

and

$$(s^{k+1})'(s^{k+1}) = \sum_{i=1}^{L} a_{i,p+1}^2 + Q \qquad (2\text{-}45)$$

so that

$$(s^{k+1})'s^{k+1} = (s^k)'s^k - \sum_{i=1}^{L} a_{i1}^2 + \sum_{i=1}^{L} a_{i,p+1}^2 \quad . \qquad (2\text{-}46)$$

The reduction in computation is particularly significant if L = P. For example, if L = 1 then the computation of $(s^{k+1})'s^{k+1}$ requires only 2 multiplies and 2 adds; or in general 2L multiplies and 2L adds. Assuming an epipolar search with T subimage locations, the total number of computations is listed in Table 2-3.

If the assumption is made that multiplication complexity (either time or hardware) is, say, three times the complexity

Table 2-2. Computational Complexity of Correlation Processors per Subimage

Covariance	N + 2 multiplies $2N$ - 1 adds
Least Squares and Normalized Covariance	$2N$ + 5 multiplies $4N$ - 1 adds

Table 2-3. Computational Complexity of Correlation Processor per Match Window Location

Covariance	$TN + 2(T - 1)$ + 2 multiplies $(T + 1)N + 2(T - 1)L - 1$ adds
Least Squares and Normalized Covariance	$(T + 1)N + (T - 1)(2L + 5) + 5$ multiplies $2(T + 1)N + (T - 1)(4L + 1) - 1$ adds

of an add operation then the results in Table 2-3 can be expressed in terms of equivalent adds as indicated in Table 2-4. Computational complexity is plotted as a function of N with the parameter L in Figure 2-16 for T = 9 (a typical value). Thus for T = 9, the computational complexity of least-squares and normalized covariance is about 74% higher than for covariance at N = 10 and varying nearly linearly, to 54% for N = 100. Clearly, these quantities depend on the implementation and are meant to be illustrative. One of the major differences between the general class of image registration problems and those that apply specifically to stereo-compilation is the size of the match window; in stereo systems a small match window is required to achieve the desired resolution and the computational savings of covariance (percentage wise) increases with decreasing window size. If cost is no object, highly parallel processors can be employed which nearly eliminate this differential in terms of computation time.

Table 2-4. Computational Complexity (Equivalent Adds) of Correlation Processor per Match Window Location

Covariance:	$(4T + 1)N + 2(T - 1)(L + 3) + 5$
Least Squares and Normalized Covariance	$5(T + 1)N + 2(T - L)(5L + 3) + 14$

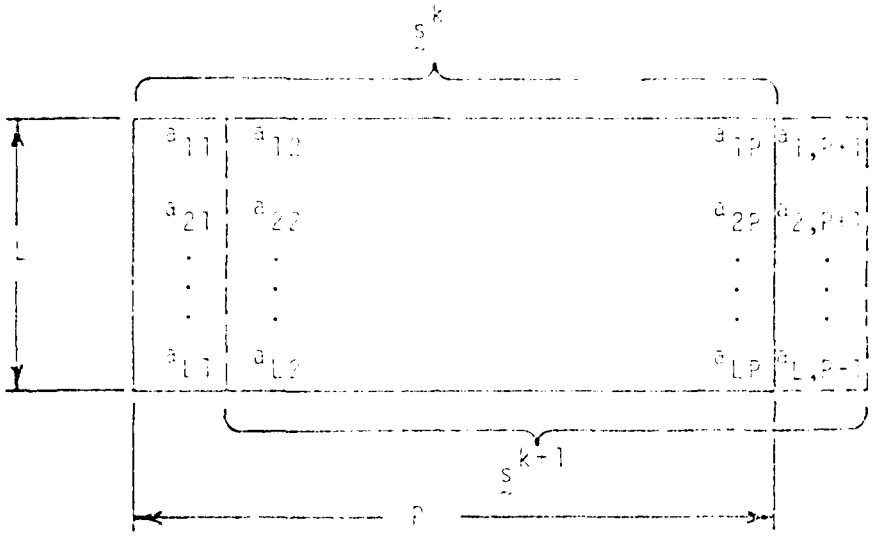


Figure 2-15. Adjacent Subimages

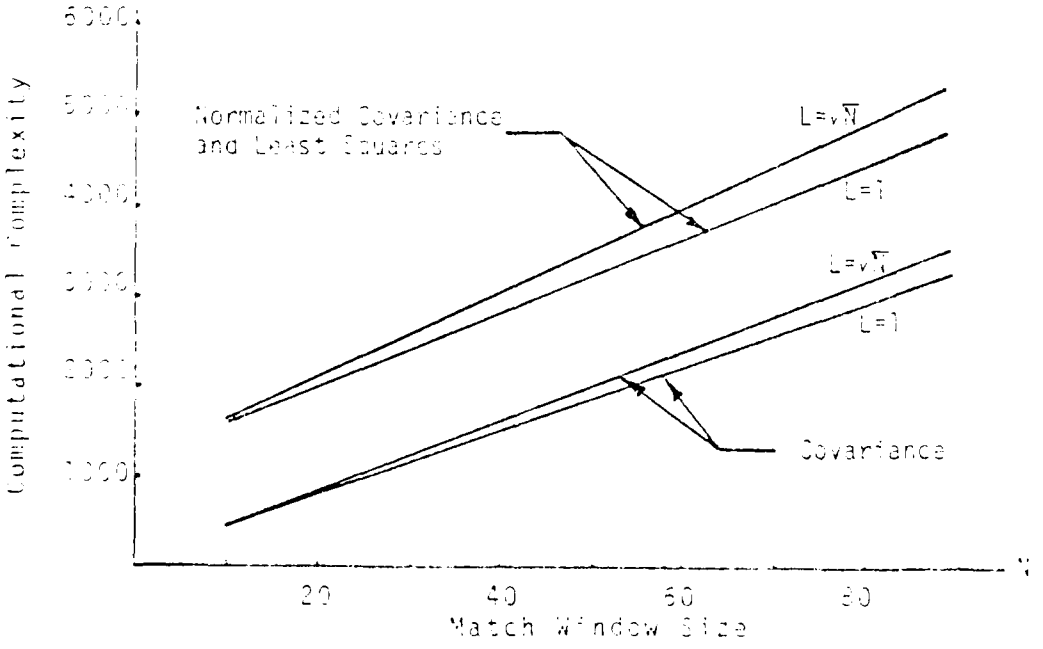


Figure 2-16. Computational Complexity of Correlation Processors

CHAPTER 3

CORRELATION AND ERROR

As shown in Chapter 2, the conjugate image intensity patterns are, in general, different, and the effect of this difference is the introduction of errors in the computation of the correlation value. The magnitude of these errors depends, in part, on the image density patterns themselves, the degree of dissimilarity between conjugate image regions, and, as we have seen, the correlation processor. The accuracy of correlation processors has been extensively studied in the context of radar and sonar target range determination and it has been shown by the methods of Cramer-Rao (Van Trees 1971) that in the presence of noise there is a lower bound on the accuracy of *any* correlation processor. The difference between the radar application and the image registration application is a rather subtle difference in the models. In the radar case, if the transmitted signal is $s(t)$, then the received signal is usually modelled by:

$$r(t) = \alpha s(t + \tau) + n(t) \qquad (3\text{-}1$$

where α is a scale factor accounting for the loss of signal power according to the radar equation (Skolnik 1970, 1980), $n(t)$ is white Gaussian noise that is uncorrelated with $s(t)$, and τ is the round trip time delay. For image registration

applications, the model is given by equation (3-6) which for the continuous case is written

$$r(x) = s(x + \tau_0) + n_r(x)$$
$$s(x + \tau_0) = s(x) + n_s(x) \qquad (3\text{-}2)$$

where τ_0 is the parallax between the adjacent regions.

As before, we can rewrite (3-2), as

$$r(x) = s(x + \tau_0) - n_s(x) + n_r(x)$$
$$= s(x + \tau_0) + n_r(x) - n_s(x) \quad . \qquad (3\text{-}3)$$

The difference between (3-1) and (3-3) is simply that in the latter, the noise ($n_s(x)$) is signal independent.

The Cramer-Rao bound is a lower bound on the variance of any unbiased estimate of the parallax, τ (or time delay τ), and its form depends on the model selected. The general Cramer-Rao bound is given by [Van Trees 1968, p.72]

$$E[(\tau - \hat{\tau})^2] \geq \frac{1}{E\left[\left(\frac{\partial \ln p(r,s,\tau)}{\partial \tau}\right)^2\right]} \qquad (3\text{-}4)$$

where $p(r,s,\tau)$ is the joint pdf of the observed signals r and s and the unknown parameter τ and $\hat{\tau}$ is an estimate of the true parallax τ_0. In the following section, two bounds are presented. One is based on (3-1), and the other on (3-3). It is then possible to show that the Cramer-Rao bound for (3-3) converges to the bound for (3-1), with increasing signal-to-noise ratio.

The Cramer-Rao Bound

The Cramer-Rao bound for the model of (3-3) has been derived by Knapp and Carter [1976] using the approach discussed in Chapter 2 (see (2-34) and (2-35)) to obtain an appropriate probability density function for use in (2-4). The result is given by

$$E[(\tau - \hat{\tau})^2] \geq \left\{ \int_{-1}^{1} (2\pi f)^2 \frac{|\gamma_{rs}(f)|^2}{1 - |\gamma_{rs}(f)|^2} \, df \right\}^{-1} \qquad (3\text{-}5)$$

where $|\gamma_{rs}(f)|^2$ is the continuous, one-dimensional form of (3-3), i.e.,

$$|\gamma_{rs}(f)|^2 = \frac{|\hat{G}_{rs}(f)|^2}{\hat{G}_{rr}(f)\hat{G}_{ss}(f)} \, , \qquad (3\text{-}6)$$

and D is the Hamming window length.

For the model of (3-2), it follows that

$$\begin{aligned} G_{rr}(f) &= |B|^2 G_{gg}(f) + G_n(f) \\ G_{ss}(f) &= G_{gg}(f) + G_n(f) \qquad (3\text{-}7) \\ G_{rs}(f) &= B\, G_{gg}(f) \, . \end{aligned}$$

Substituting (3-7) into (3-6) and assuming $G_n(f) = \frac{N_0}{2}$, where $\frac{N_0}{2}$ is the noise spectral height, yields

$$|\gamma_{rs}(f)|^2 = \frac{|B|^2 G_{gg}^2(f)}{\left(|B|^2 G_{gg}(f) + \frac{N_0}{2}\right)\left(G_{gg}(f) + \frac{N_0}{2}\right)} \, . \qquad (3\text{-}8)$$

If

$$G_{gg}(f) = \frac{N_0}{2} \qquad (3\text{-}9)$$

for all frequencies within the passband of the system, then the first term of the denominator of (3-8) reduces to the second term and

$$E[\epsilon - \hat{\epsilon}]^2 \geq \left\{ \frac{2}{1+\alpha^2} \int_{-\frac{1}{T}}^{\frac{1}{T}} (2\pi f)^2 G_{gg}(f)\, df \right\}^{-1} \qquad (3\text{-}10)$$

or

$$E[\epsilon - \hat{\epsilon}]^2 \geq \frac{\frac{N_0}{2}\left(\frac{1+\alpha^2}{2}\right)}{\int_{-\infty}^{\infty} (2\pi f)^2 G_{gg}(f)\, df} \,. \qquad (3\text{-}11)$$

Noting that

$$\sigma_g^2 = \int_{-\infty}^{\infty} G_{gg}(f)\, df \qquad (3\text{-}12)$$

(3-11) can be expressed in the form

$$E[\epsilon - \hat{\epsilon}]^2 \geq \frac{\left(\frac{1+\alpha^2}{2}\right)}{(2\pi)^2 \left(\frac{N_0}{2}\right)^2 \sigma_B^2} \qquad (3\text{-}13)$$

where

$$B = \left\{ \frac{\int_{-\infty}^{\infty} f^2 G_{ff}(f)\,df}{\int_{-\infty}^{\infty} G_{ff}(f)\,df} \right\}^{\frac{1}{2}} \tag{3-14}$$

is a measure of the bandwidth of the underlying image pattern $g(x)$. From (3-13) we see that correlation accuracy should improve with increased SNR ($\sigma_g^2/\frac{N_0}{2}$) and increased space bandwidth product (TB^2). Also, if the value of α is small, say due to an incorrect exposure of one of the images, then the correlation performance should degrade. These qualitative interpretations of (3-13) are certainly intuitively reasonable although, as we will see in Chapter 5, terrain relief will force a slightly different interpretation since the relationships in (3-7) will no longer be valid.

A bound can also be derived by evaluating (3-4) for the probability density function given in (2-19) which was based on the model of (3-3) but assumed independence of the noise. This model is essentially equivalent to the radar model for which the Cramer-Rao bound is given by (Van Trees 1968)

$$E[(\hat{\tau} - \tau)^2] \geq \frac{\frac{N_0}{2}(1 + \alpha^2)}{\alpha^2 \int_{x_l}^{x_u} \left[\frac{\partial s'(x,\tau)}{\partial \tau}\right]^2 dx} \tag{3-15}$$

where the limits x_ℓ and x_u are determined from the extremes of the match window. Application of Fourier derivative theorems and Parseval's theorem yields [Hunt and Ryan 1976]

$$E\{(\xi - \hat{\xi})^2\} \geq \frac{\frac{N_0}{2}\left(\frac{1 + \rho^2}{\rho^2}\right)}{(2\pi)^2 \sigma_s^2 D \beta^2} \tag{3-16}$$

where

$$\beta^2 = \frac{\int f^2 G_{ss}(f)\,df}{\int G_{ss}(f)\,df} \tag{3-17}$$

and $D = |x_u - x_\ell|$.

For the model of (3-2),

$$\beta^2 = \frac{\int f^2 (G_{gg} + \frac{N_0}{2})\,df}{\int (G_{gg}(f) + \frac{N_0}{2})\,df} \,. \tag{3-18}$$

Repeating the assumption of large SNR at each pass-band frequency, it follows that $\beta^2 \to B^2$ and $\sigma_s^2 \to \sigma_c^2$ with increasing SNR. Thus the lower bound in (3-11) and (3-13) are limiting forms of both (3-6) and (3-16) for high SNR's. The results in (3-16) and (3-17), however, point out the deficiency of the assumption that the noise and signal in (3-3) are uncorrelated. Note that if the SNR is low (say < 1) then the computation of β^2 in (3-17) is dominated by the noise. Thus we would expect (3-16) to be inconsistent in regions of low signal energy.

The term B^2 in (3-13) is the normalized second moment of the spectral density of the noise-free match window. From the moment theorem [Papoulis 1965] it follows that

$$(2\pi)^2 \int_{-\infty}^{\infty} f^2 G_{gg}(f)\,df = -\left.\frac{d^2 C_g(\tau)}{d\tau^2}\right|_{\tau=0} \qquad (3\text{-}19)$$

where $C_g(\tau)$ is the autocovariance function of the process g. Thus the lower bound in (3-11) is inversely proportional to the curvature of autocovariance function measured at the origin.

Equations (3-8) and (3-13) can be put in a more useful form for computation on sampled imagery. Assuming that the sampling rate is sufficiently high, the resulting sampled image is bandlimited in .5 cycles/sample spacing from which it follows that the noise variance, σ_n^2, is given by

$$\sigma_n^2 = \int_{-.5}^{.5} \frac{N_0}{2}\,df = \frac{N_0}{2} \qquad (3\text{-}20)$$

where, since the units are different, the equality holds in magnitude only. Thus (3-8) becomes

$$E\{(\xi - \hat{\xi})^2\} \geq \left\{ D \int_{-\infty}^{\infty} (2\pi f)^2 \frac{\alpha^2 G_{gg}(f)\,df}{(1+\alpha^2) G_{gg}(f)\sigma_n^2 + \sigma_n^2} \right\}^{-1} \qquad (3\text{-}21)$$

and the limiting form becomes

$$E\{(\xi - \hat{\xi})^2\} \geq \frac{\sigma_n^2 (1+\alpha^2)/\alpha^2}{(2\pi)^2 \sigma_g^2 D B^2} \quad . \qquad (3\text{-}22)$$

The lower bound in (3-21) is attainable only by a so-called "efficient" or "minimum variance" estimator (Whalen 1971). One of the assumptions leading to (3-21) is negligible topographic relief, which, as stated earlier, leads to the relationships in (3-7). For the more general case of

varying elevation, we must return to (3-5) to obtain an appropriate bound since $\hat{G}_{rs}(f)$ in (3-6) is dependent on the relief distortion.

The Cramer-Rao inequality specifies a lower bound on the parallax error variance but it does not guarantee that this bound can actually be reached or even closely approximated. For this reason, it becomes important to study physically realizable correlation processors and to compare their performance to the lower bound.

Generalized Correlation Local Error Variance

The generalized correlation processors described in Chapter 2 allow an analytical approach to error variance determination. Let

$$c(\xi) \triangleq \int_{-\infty}^{\infty} w(f)\hat{G}_{rs}(f)\, e^{j2\pi\xi f} df \qquad (3\text{-}23)$$

be the one-dimensional generalized correlation function analogous to (2-29). The average output $\bar{c}(\xi)$ of the generalized processor is shown in Figure 3-1 along with a few sample cross-correlation functions. Without loss of generality, we assume $\xi_0 = 0$. The actual output, $c(\xi)$, is only an approximation to $\bar{c}(\xi)$ due to the effects of noise and the fact that in a physically realizable system, the integration time is finite. Since at the peak,

$$\bar{z}(0) \triangleq \left.\frac{d\bar{c}(\xi)}{d\xi}\right|_{\xi=0} = 0 \quad , \tag{3-24}$$

the value of ξ for which $z(\xi) = \frac{dc(\xi)}{d\xi} = 0$ provides an estimate, $\hat{\xi}$, of the true value of ξ. If excursions of $z(\xi)$ are entirely confined to the linear segment near $\xi = 0$, as shown in Figure 3-2, then the parallax error variance is given by

$$\sigma_e^2 = \frac{\sigma_z^2}{\left|\left.\frac{dz(\xi)}{d\xi}\right|_{\xi=0}\right|^2} \quad . \tag{3-25}$$

Knapp and Carter [1976], extending a result from MacDonald and Schultheiss [1969] state that

$$\sigma_e^2 = E\{(\xi - \hat{\xi})^2\} = \frac{\int_{-\infty}^{\infty} |\psi(f)|^2 (2\pi f)^2 G_{rr}(f) G_{ss}(f) [1 - |\gamma(f)|^2] df}{D \left[\int_{-\infty}^{\infty} (2\pi f)^2 G_{rs}(f) \psi(f) df\right]^2} \tag{3-26}$$

where $\psi(f)$ is the pre-filter transfer function of the generalized correlator defined in (2-33).

For the maximum likelihood processor, $\psi(f)$ is given by (2-37) and repeated here for the continuous case:

$$\psi_{ML}(f) = \frac{1}{G_{sr}(f)} \cdot \frac{|\gamma(f)|^2}{[1 - |\gamma(f)|^2]} \quad . \tag{3-27}$$

Substituting (2-38) and (3-27) into (3-26), we have

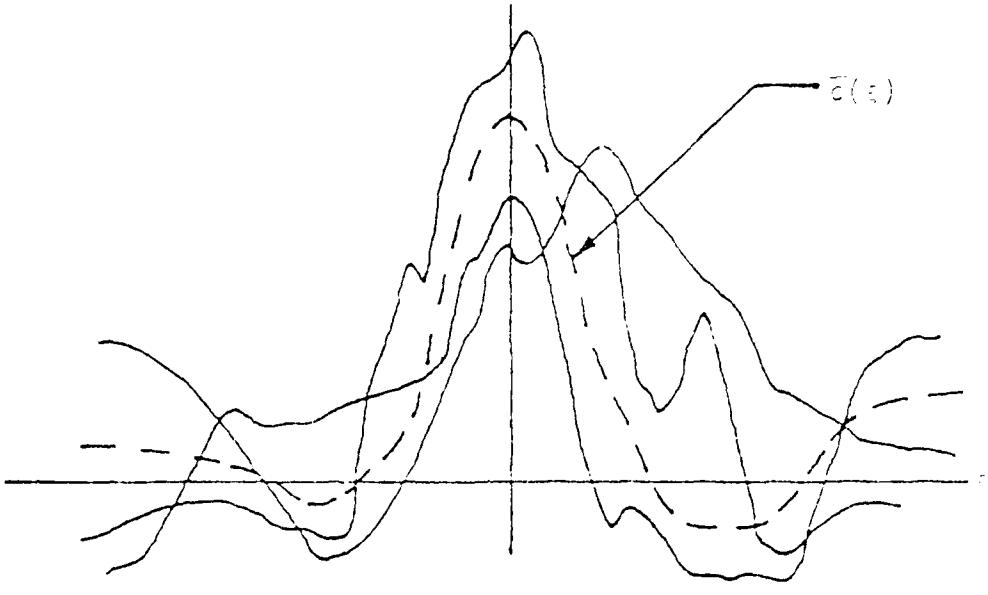


Figure 3-1. Correlator Output Waveforms

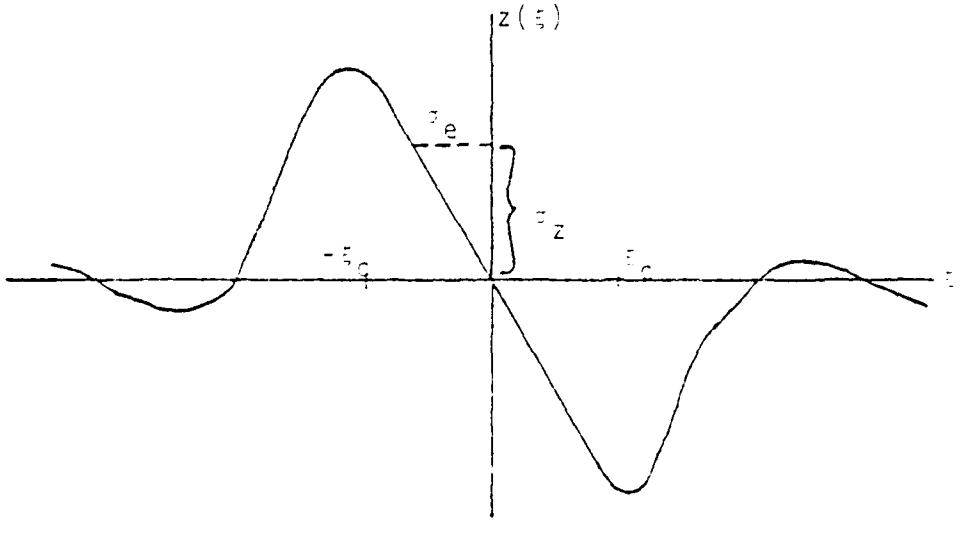


Figure 3-2. Derivative of Correlator Output

$$\sigma^2_{ML} = E\{(\varepsilon - \hat{\varepsilon})^2\} = \frac{1}{D\int_{-\infty}^{\infty} (2\pi f)^2 \frac{|\gamma(f)|^2}{1 - |\gamma(f)|^2} df} \quad (3\text{-}28)$$

We note that this is precisely the Cramer-Rao lower bound given in (3-5). There is a temptation at this point to conclude that the ML processor defined by (3-27) is a minimum variance processor. Before reaching such a conclusion, however, we must access the validity of the assumption leading to (3-5) and (3-27). Three assumptions are of particular concern. First, the assumption that D is large with respect to the correlation width is generally not valid in stereo-compilation systems. The effect of a small match window is to introduce random fluctuations into the correlation process which are not accounted for in (2-36) [Helava 1978]. Secondly, the analysis which leads to (3-26) assumes no false acquisition errors. For the present, suffice it to say that false acquisitions cause an increase in σ^2_e. Finally, all of the preceding analysis assumes wide-sense stationarity of both image and noise processes. Thus while (3-28) may provide an accurate measure of system performance in image registration problems involving large sections of imagery, we cannot expect the ML processor to achieve the accuracy specified by the lower bound in (3-5).

For the simple covariance estimator, $\lambda(f) = 1$, and from (3-26), the covariance error variance, σ^2_e, is given by

$$\sigma_e^2 = \frac{\int_{-\infty}^{\infty} (2\pi f)^2 [G_{rr}(f)G_{ss}(f) - |G_{rs}(f)|^2]df}{D\left[\int (2\pi f)^2 |G_{rs}(f)|df\right]^2} \quad . \qquad (3\text{-}29)$$

With substitutions from (3-7), (3-29) becomes

$$\sigma_e^2 = \frac{\int_{-\infty}^{\infty} (2\pi f)^2 [(1+\alpha^2)G_{gg}(f)G_n(f) + G_n^2(f)]}{D\left[\alpha \int (2\pi f)^2 G_{gg}(f)df\right]^2} \quad . \qquad (3\text{-}30)$$

For the conditions of (3-9), this again reduces to (3-11) although, in general $\sigma_e^2 \geq \sigma_{ML}^2$ since σ_{ML}^2 is the variance of the "minimum variance" estimator. Since for high signal-to-noise ratios, the theoretical accuracy of the covariance processor converges to the Cramer-Rao bound, the Cramer-Rao bound itself is a reasonable candidate for an image quality measure.

Specifically, we will consider (3-8) as an image feature. Since $G_{gg}(f)$ is generally unavailable, we must compute $G_{ss}(f)$ (or $G_{rr}(f)$) and estimate $G_n(f)$ so that (3-7) can be used to obtain

$$G_{gg}(f) = G_{ss}(f) - G_n(f) \quad . \qquad (3\text{-}31)$$

The quantity α can be estimated using (2-16). Experimental results (see Chapter 6) show that the Cramer-Rao measure is a potentially useful pre-processing algorithm. As we shall see, however, other features are also useful. In the following section, we consider the false acquisition probability as a possible quality measure. Before proceeding, however,

we note that for high SNR, there should be little, if any, difference between the ML processor given by (2-36) and (2-37) and simple covariance. Referring to Figures 2-9, 2-10, and 2-11, we see that both least-squares and normalized covariance result in a lower-mean-square error. Thus, covariance is definitely not a minimum variance estimator. As discussed in Chapter 4, the violation of local image stationarity is the primary reason for this discrepancy.

Correlation False Acquisition

As described in Chapter 1, a false acquisition occurs when the peak of the correlation function is associated with an incorrect lobe (Figure 1-1c). For an arbitrary correlation processor $c(\tau)$, the probability that such a peak occurs at location $\tau \neq 0$ is expressed as

$$P_{fa}(\tau) = \text{Prob}\,[c(\tau) > c(0)] \qquad (3\text{-}32)$$

The total false alarm probability, P_T, depends on the joint pdf of the correlator output locations and the statistical dependence between correlator output locations makes it impossible to obtain P_T in closed form except under the conditions of simplifying assumptions which will be discussed in Chapter 4. Although P_{fa} depends on the particular correlation processor, the following observations are appropriate:

1. Covariance (Equation (2-23)), least-squares (Equations (2-21) or (2-22)) and normalized covariance (Equations (2-25) or (2-26)) all require computation of the covariance.

2. Under the assumption of stationarity the energy terms are constant or change slowly with subimage location. Thus each of the processors is approximately proportional to covariance.
3. The mean value of the covariance function at the correct peak is directly related to the total energy of the match window.
4. The fluctuation of the correlation functions far from the peak depend on the noise content in the particular region as well as the spatial frequency content of the image patterns themselves.

Applying these observations to the qualitative behavior of $P_{fa}(\xi)$, it follows that we should expect an increase in false acquisition frequency in regions of low signal-to-noise ratio. If we assume that the functions $r(x)$ and $s(x)$ are zero-mean, then the covariance processor is given by

$$c(\xi) = \int_{-D/2}^{D/2} r(x)s(x + \xi)dx \quad . \tag{3-33}$$

Assuming further that $c(\xi)$ and $c(0)$ are Gaussian distributed random variables with

$$E[c(\xi)] = \int_{-D/2}^{D/2} E[r(x)s(x + \xi)]dx = D\,C_{rs}(\xi) \tag{3-34}$$

and

$$\mathrm{Var}\{c(\xi)\} = \sigma_c^2(\xi) \tag{3-35}$$

with $\sigma_c^2(\xi)$ as yet undetermined, then

$$\mathrm{Prob}\{c(\xi) > c(0)\} = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\frac{\sigma_c(\xi)v - D(C_{rs}(0) - C_{rs}(\xi))}{\sigma_c(0)}} \mathrm{Gaus}(u)du \right] \mathrm{Gaus}(v)dv \tag{3-36}$$

where

$$\mathrm{Gaus}(x) \triangleq \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] . \tag{3-37}$$

The derivation of (3-36) is given in Appendix A.

If $\sigma_c(0) = \sigma_c(\xi)$ and $C_{rs}(\xi) = 0$ $(\xi > \xi_c)$ then since $C_{rs}(0) = D\sigma_g^2$ we have

$$\mathrm{Prob}\{c(\xi) > c(0)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{v - D\sigma_g^2}{\sigma_c(0)}} \mathrm{Gaus}(u)du\, \mathrm{Gaus}(v)dv . \tag{3-38}$$

Thus the false acquisition probability depends on the ratio

$$\zeta = D\sigma_c^2/\sigma_c(0) = \frac{\text{signal energy}}{\text{correlator output standard deviation}} . \tag{3-39}$$

It has been shown [Andrews 1974] that the probability density functions of the correlator tap outputs are approximately Gaussian although this approximation becomes less valid away from the correlator peak. The quantity ζ in (3-39) can be shown to be equivalent to the "figure of merit" used by

Webber and Delashmit [1974a] to describe correlator performance for Gaussian random scenes.

In Chapter 4, we will see that $\tau_c(0)$ depends on additive noise as well as "self noise" and "machine noise" which are error sources to be discussed in the following chapter.

CHAPTER 4

THE COVARIANCE FUNCTION

The covariance processor defined by Equation (2-23) is the most common technique for locating conjugate imagery in digital stereo-pair images. It is also commonly employed by correlation processors in radar and sonar systems to estimate target range by measuring signal round trip travel time. In the context of these applications, the correlation processor is also known to be the "matched filter" which maximizes the processor output signal-to-noise ratio at the peak of the correlation function. For the analysis to follow, it is more convenient to express the covariance function in summation notation rather than the vector inner product notation. We thus consider the discrete covariance function estimate

$$c(k) = \sum_{i=1}^{N} r_i s_{i+k} - \frac{1}{N} \sum_{i=1}^{N} r_i \sum_{i=1}^{N} s_{i+k} \qquad (4-1)$$

where N is the number of elements in the match window and we assume one-dimensional sequences. The extension to two dimensions is straightforward and representative results will be included in a later section.

Assuming negligible topographic relief, there are four factors which influence correlator performance:

(1) Sensor noise attributable to the data acquisition process.

(2) Self noise resulting from the fact that each correlation function is a sample from a random process with associated statistical fluctuations.

(3) Machine noise resulting from computer truncation or round-off due to limited machine register lengths.

(4) Input data non-stationarity which violates the assumption leading to covariance being equivalent to the ML processor.

In a digital system, the correlation functions are sampled at 2L + 1 evenly spaced locations. It will be assumed that an error in the cross-correlation computation occurs if a value of k in (4-1) other than zero corresponds with the maximum value of $c(k)$. Denoting the probability of error by P_e and the probability that the peak occurs at tap m by P_m we have that

$$P_e = 1 - P_0 \tag{4-2}$$

and

$$MSE = \sum_{\ell=-L}^{L} \ell^2 P_\ell \tag{4-3}$$

where MSE is the estimated mean-square-error. A tacit assumption in (4-3) is that the correct peak location is at the center of the search region. If this is not the case the limits in (4-3) are not so simply related and a registration bias can result in a non-zero mean correlation error

[Mostafavi and Smith 1978b]. In general, the computation of P_m involves a 2L + 1-fold integration over the joint pdf of the correlator tap outputs. For even the smallest values of L, this is computationally prohibitive. Because of the correlation between tap outputs, the only situations that can be handled analytically occur when the input data consists of either very smooth lowpass signals or wide-bandwidth signals approaching white noise. In the former case, we can assume that the tap outputs are so highly correlated that the self noise is negligible and in the latter, we assume the tap outputs are statistically independent. Wernecke [1978] presents an analysis of multitap correlator self-noise for analog signals in which the tap covariances are accounted for by performing a least-squares parabolic fit to the average correlator output in the vicinity of the correlation peak. The results obtained will be presented in the sections to follow. They do not differ substantially, however, from the results associated with the Cramer-Rao lower bound which, we recall, is related to local registration error and neglects false acquisition errors.

Sensor Noise

Let each image be corrupted by additive zero-mean Gaussian uncorrelated noise. The covariance calculation in the presence of such noise becomes

$$\hat{c}(k) = \sum_{i=1}^{N} (g_i + \xi_i)(g_{i+k} + n_i) - \frac{1}{N} \sum_{i=1}^{N} (g_i + \xi_i) \sum_{i=1}^{N} (g_{i+k} + n_i) \tag{4-4}$$

where $\hat{c}$ is an estimate of the true discrete covariance function

$$c(k) = \sum_{i=1}^{N} g_i g_{i+k} - \frac{1}{N} \sum_{i=1}^{N} g_i \sum g_{i+k} \quad , \tag{4-5}$$

and $\{\xi\}$ and $\{n\}$ are statistically independent white noise sequences. The computational error due to the additive noise is given by

$$e_c(k) = \hat{c}(k) - c(k) = \sum_{i=1}^{N} g_i n_i - \frac{1}{N} \sum_{i=1}^{N} g_i \sum_{i=1}^{N} n_i + \sum_{i=1}^{N} \xi_i g_{i+k} - \frac{1}{N} \sum_{i=1}^{N} \xi_i \sum_{i=1}^{N} g_{i+k} + \sum_{i=1}^{N} \xi_i n_i - \frac{1}{N} \sum_{i=1}^{N} \xi_i \sum_{i=1}^{N} n_i \quad . \tag{4-6}$$

Since the noise sequences are uncorrelated and $E\{e_c(k)\} = 0$, it follows that

$$\sigma_{e_c}^2(k) = E\{e_c^2(k)\} = E\{(\Sigma g_i n_i - \frac{1}{N}\Sigma g_i \Sigma n_i)^2\}$$
$$+ E\{(\Sigma s_i g_{i+k} - \frac{1}{N}\Sigma s_i \Sigma g_{i+k})^2\}$$
$$+ E\{(\Sigma s_i n_i - \frac{1}{N}\Sigma s_i \Sigma n_i)^2\} \qquad (4\text{-}7)$$

Taking expectations with respect to the noise terms it is straightforward to show that

$$\sigma_{e_c}^2(k) \cong \sigma_n^2[(\Sigma g_i^2 - \frac{1}{N}(\Sigma g_i)^2)$$
$$+ (\Sigma g_{i+k}^2 - \frac{1}{N}(\Sigma g_{i+k})^2) + N\sigma_n^2] \quad . \qquad (4\text{-}8)$$

If we also assume stationarity of the input process as well, expectations in (4-7) can also be taken with respect to the process g resulting in

$$\sigma_{e_c}^2(k) \cong N\sigma_n^2[2\sigma_g^2 + \sigma_n^2] \qquad (4\text{-}9)$$

where we let

$$E(\Sigma g_i^2 - \frac{1}{N}(\Sigma g_i)^2) \cong N\sigma_g^2 \qquad (4\text{-}10)$$

and

$$E(\Sigma n_i^2 - \frac{1}{N}(\Sigma n_i)^2) = N\sigma_n^2 \qquad (4\text{-}11)$$

Since the error in (4-6) usually involves a large number of terms, it is reasonable to assume that the distribution of e_c is Gaussian.

Self Noise

If the image input signals consisted solely of the desired signals (i.e., g_i and g_{i+k}) there would still be statistical fluctuations at the correlator output due to the random nature of the desired signals. These fluctuations are known as correlator "self-noise".

Taking expectations in (4-1),

$$E(\hat{c}(k)) = N R_{rs}(k) - \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} R_{rs}(j + k - i) \qquad (4\text{-}12)$$

where R_{rs} is the cross-correlation function

$$R_{rs}(k) = E\{r_i s_{i+k}\} \qquad (4\text{-}13)$$

and stationarity is assumed. The cross-correlation between correlator taps is given by

$$\begin{aligned} E(\hat{c}(k)\hat{c}(\ell)) &= E\{(\Sigma r_i s_{i+k} - \tfrac{1}{N} \Sigma r_i \Sigma s_{i+k})(\Sigma r_i s_{i+\ell} - \tfrac{1}{N} \Sigma r_i \Sigma s_{i+\ell})\} \\ &= \sum_{i=1}^{N} \sum_{j=1}^{N} E(r_i s_{i+k} r_j s_{j+\ell}) \\ &+ \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{m=1}^{N} \sum_{n=1}^{N} E(r_i s_{m+k} r_j s_{n+\ell}) \qquad (4\text{-}14) \\ &- \frac{1}{N} \sum_{i=1} \sum_{j=1} \sum_{m=1} \{E(r_i s_{i+k} r_j s_{m+\ell}) \\ &+ E(r_i s_{m+k} r_j s_{m+\ell})\} \end{aligned}$$

In general, the fourth order moments in (4-14) cannot be simplified. However, if we assume that the error incurred by estimating the means in (4-1) is negligible then the covariance function can be written as

$$\hat{c}(k) = \sum_{i=1}^{N} r_i s_{i+k} \tag{4-15}$$

where r and s are zero-mean sequences. With this assumption, only the first term in (4-14) remains.

If we then employ the image model

$$\begin{aligned} r_i &= g_i + \xi_i \\ s_i &= g_i + \eta_i \end{aligned} \tag{4-16}$$

then

$$\begin{aligned} E(\hat{c}(k)\hat{c}(\ell)) = &\sum_{i=1}^{N}\sum_{j=1}^{N} E(g_i g_{i+k} g_j g_{j+\ell}) \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{N} [E(\xi_i \xi_j g_{i+k} g_{j+\ell}) \\ &+ E(\xi_i \xi_j \eta_{i+k} \eta_{j+\ell}) \\ &+ E(\eta_{i+k} \eta_{j+\ell} g_i g_j)] \end{aligned} \tag{4-17}$$

$$\begin{aligned} = &\sum_{i=1}^{N}\sum_{j=1}^{N} E(g_i g_{i+k} g_j g_{j+\ell}) \\ &+ N\sigma_n^2 C_g(k-\ell) + N\sigma_n^4 \delta(k-\ell) \\ &+ N\sigma_n^2 \sigma_f^2 (k-\ell) \\ &+ 2\sigma_n^2 (N - |k-\ell|) C_f(k-\ell)(1 - \delta(k-\ell)) \end{aligned} \tag{4-18}$$

where $\delta(k)$ is the Kronecker delta function

$$\delta(k) = \begin{cases} 1, k = 0 \\ \\ 0, \text{otherwise} \end{cases} . \tag{4-19}$$

Thus the cross-covariance between correlator taps is given by

$$K_c(k,\ell) = E(\hat{c}(k)\hat{c}(\ell)) - E(\hat{c}(k)E(\hat{c}(\ell))$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} E(g_i g_{i+k} g_j g_{j+k})$$

$$+ \sigma_n^2(3N - |k-\ell|)C_g(k-\ell) \qquad (4\text{-}20)$$

$$+ N\sigma_n^2(\sigma_n^2 - \sigma_g^2)\delta(k-\ell) - N^2 C_g(k)C_g(\ell) \quad .$$

Since g is zero-mean, the fourth order moment in (4-20) can be simplified if we assume the process generating g is Gaussian. In this case [Whalen 1971, p.97],

$$E(g_i g_{i+k} g_j g_{j+k}) = C_g(k)C_g(\ell) + C_g(i-j)C_g(i+k-j-\ell)$$

$$+ C_g(j+\ell-i)C_g(i+k-j) \qquad (4\text{-}21)$$

and (4-20) becomes

$$K_c(k,\ell) = \sum_{i=1}^{N}\sum_{j=1}^{N} (C_g(i-j)C_g(i-j+k-\ell) + C_g(j+\ell-i)C_g(i+k-j)$$

$$+ \sigma_n^2[3N-|k-\ell|]C_g(k-\ell) + N\sigma_n^2(\sigma_n^2-\sigma_g^2)\delta(k-\ell) \quad . \qquad (4\text{-}22)$$

If $k = \ell$ we obtain the variance of the ℓ^{th} tap:

$$\sigma_c^2(\ell) = K_c(\ell,\ell) = \sum_{i=1}^{N}\sum_{j=1}^{N} (C_g^2(i-j) + C_g(j+\ell-i)C_g(i+\ell-j))$$

$$+ N\sigma_n^2(2\sigma_g^2 + \sigma_n^2) \quad . \qquad (4\text{-}23)$$

Note that the last term in (4-23) is the contribution to the tap variance due to the additive sensor noise. Thus the additive noise and the self-noise of the desired signal are uncorrelated. If g is a white process then (4-22) becomes

$$K_c(k,\ell) = N\sigma_g^4(\delta(k-\ell) + \delta(k+\ell)) + N\sigma_n^2(2\sigma_g^2 + \sigma_n^2) \quad (4\text{-}24)$$

and the individual tap variances are given by

$$\sigma_c^2(\ell) = N(\sigma_g^2 + \sigma_n^2)^2 + N\sigma_g^2 \quad . \quad (4\text{-}25)$$

The contribution of self-noise to the degradation of correlation accuracy is, in general, difficult to determine due to the correlation between correlator tap outputs given by (4-22). As mentioned earlier, Wernecke [1978] presents an analysis in which the tap covariances are accounted for by performing a least-squares parabolic fit to the average correlator output in the vicinity of the correlation peak. With the assumptions that:

(1) The correlator taps are centered on the correct peak location,

(2) The variance of the correlation function curvature is negligibly small, and

(3) The correct correlation lobe is acquired,

Wernecke concludes that the parallax error variance is given by

$$\sigma_e^2 = \frac{3\sum_{i=-L}^{L}\sum_{j=-L}^{L} ij\, K_c(i,j)}{(2\tau)^4\sigma_g^4(NB^2)^2\left[(L+1)(2L+1)L\right]^2} \quad (4\text{-}26)$$

where B is defined in (3-14) and the total number of correlator taps is 2L+1. Note that assumption (3) above ignores false acquisition errors. If the correlator tap outputs are statistically independent then (4-26) becomes

$$\sigma_e^2 = \frac{6 \sum_{i=1}^{L} i^2 \sigma_c^2(i)}{(2\pi)^4 \sigma_g^4 (NB^2)^2 [(L+1)(2L+1)L]^2} \quad . \qquad (4\text{-}27)$$

Comparing this result with (4-3) we see that under the above assumptions,

$$P_\ell \cong \frac{3 \sigma_c^2(\ell)}{(2\pi)^4 \sigma_g^4 (NB^2)^2 [(L+1)(2L+1)L]^2} \quad . \qquad (4\text{-}28)$$

The fact that σ_e^2 in (4-27) is inversely proportional to L^6 is somewhat misleading since the addition of correlator taps reduces σ_e^2 only as long as the additional correlator taps improve the parabolic curve fitting technique. For severely bandlimited signals resulting in broad correlation functions, such a result may be appropriate. However, in the case of severely bandlimited signals, it is highly unlikely that the correlator taps are statistically independent. To avoid these problems, one must use (4-26) and constrain L to a reasonably small value. The computation of (4-26) on match window size subsections is cumbersome due to the calculation of the covariance matrix K_c. Equation (4-26) is thus not well suited to be used as an image quality measure or pre-processor algorithm.

Machine Noise

Figure 4-1 shows a block diagram of the computation of the covariance function as performed by the AS-118-X stereo-compiler. The computation of c_k is subject to errors due to truncation or round-off in the limited precision computer. The magnitude of the errors incurred depend on the number of binary digits (bits) used to represent the input data, the number of bits used in the arithmetic operations and the type of arithmetic employed. Most present stereo-compilers employ fixed-point arithmetic with 8 bit precision. Accuracy in fixed point correlators suffers from the need to scale the input data in such a way as to avoid register overflow in the accumulators. With continuing improvements in microprocessor and minicomputer technology, we will see the advent of 16 bit stereo-compilers in the very near future. Floating point processors eliminate the need for prescaling and improve overall system precision at the expense of increased system complexity and cost and decreased throughput rates. Developments in the digital processing of radar and sonar signals, however, have led to implementations of hybrid floating point processors which are based on the idea that the exponent tends to grow only in the positive direction, particularly near the peak of the covariance function. This data format is easily implemented in hardware and has performed satisfactorily [Oppenheim 1978, p.272] in the radar setting.

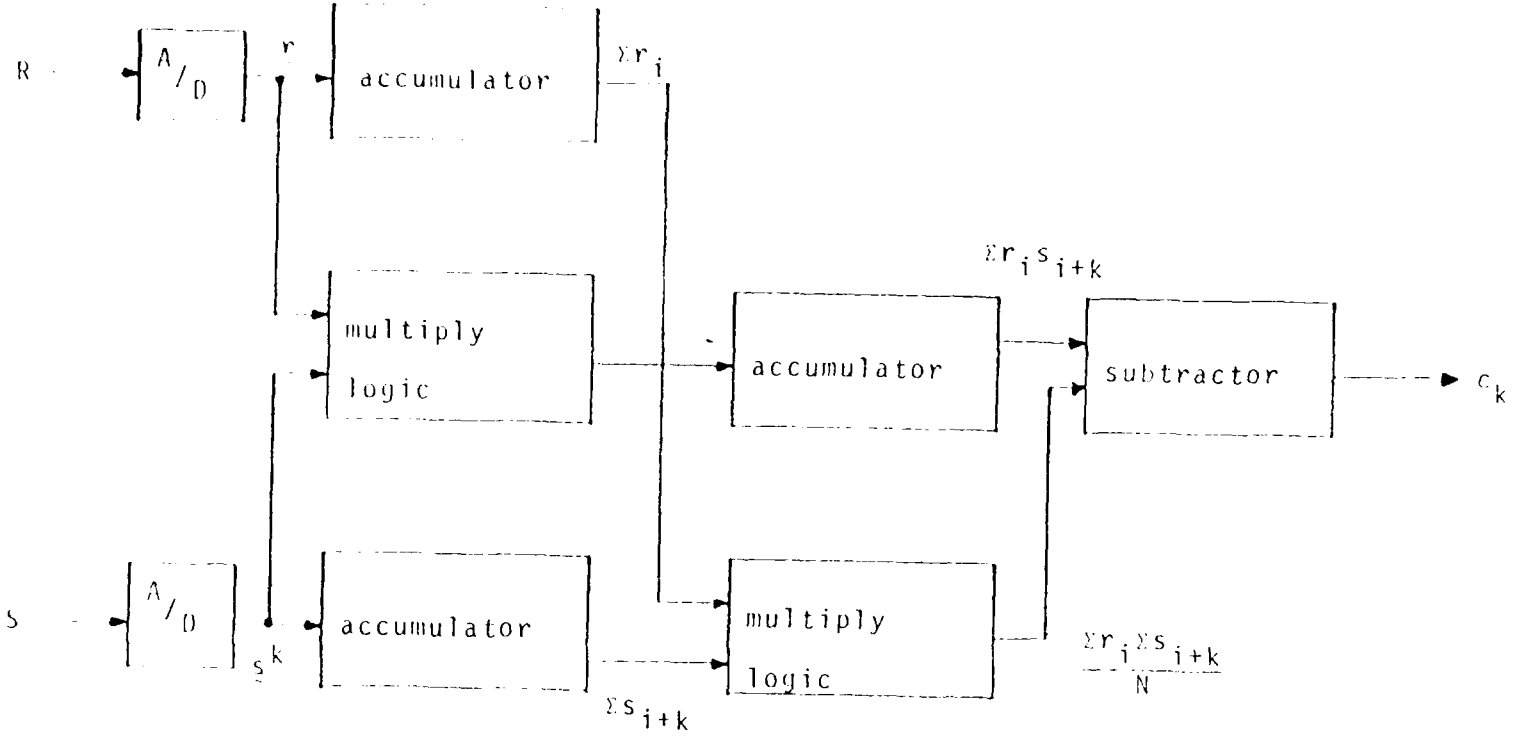


Figure 4-1. Digital Computation of Covariance

In this section, an analysis of both fixed-point and floating-point covariance processors is presented and expressions for the contribution to correlator output noise due to finite register lengths are derived. We will consider the "correct" computation to be

$$c_k = \sum_{i=1}^{N} r_i s_{i+k} - \frac{1}{N} \sum_{i=1}^{N} r_i \sum_{c=1}^{N} s_{i+1} \qquad (4\text{-}29)$$

with the actual computed function given by

$$\hat{c}_k = (\sum r_i s_{i+k} + e_1) - \frac{1}{N} (\sum r_i + e_2)(\sum s_{i+k} + e_3) + e_4 \qquad (4\text{-}30)$$

where e_1, e_2, and e_3 are the resulting accumulator output errors and e_4 is the result of division by N and subtraction. It will be assumed that the distributions of e_2 and e_3 are identical.

Fixed Point Processing

For fixed-point arithmetic, one usually considers a register to be a fixed-point fraction. In this way the product of two numbers remains a fraction which must be truncated or rounded off to remain within the available register length of b bits. The sum of two numbers, however, does not require truncation although it can produce an overflow condition, i.e., the sum is no longer a fraction. This overflow condition imposes a constraint on the dynamic range of the correlator output which, in turn, requires scaling of the input data. More specifically, if we assume that

$$0 \leq r_i \leq 1$$
$$0 \leq s_{i+k} \leq 1 \qquad (4\text{-}31)$$

for all values of i and k, then it is necessary that

$$\sum_{i=1}^{N} r'_i s'_{i+k} \leq 1 \quad ,$$

$$\sum_{i=1}^{N} r'_i \leq 1 \quad ,$$

$$\sum_{i=1}^{N} s'_{i+k} \leq 1 \qquad \forall k$$

where the prime notation denotes pre-scaled quantities. This condition can be achieved by computing

$$\frac{c_k}{N} = \sum_{i=1}^{N} \frac{r_i s_{i+k}}{N} - \sum_{i=1}^{N} \frac{r_i}{N} \sum_{i=1}^{N} \frac{s_{i+k}}{N} \quad . \qquad (4\text{-}33)$$

The computed covariance function is thus given by

$$\frac{c_k}{N} = \sum_{i=1}^{N} \left(\frac{r_i s_{i+k} + \varepsilon_i}{N} + \varepsilon_i\right)$$

$$- \sum_{i=1}^{N} \left(\frac{r_i}{N} + \eta_i\right) \sum_{i=1}^{N} \left(\frac{s_{i+k}}{N} + \xi_i\right) + e_4$$

where ε represents the multiplication round-off errors and η_i, ξ_i, and e_4 represent the errors induced by the reduction in precision caused by division by N. The value of N is commonly chosen to be a power of 2, say $N = 2^m$, so that division by N simply involves shifting the data by m bits at the input to the accumulator. If N is large the loss in precision can be significant. In fact, if $N \geq 2^b$ then

division by N completely eliminates the input data! This problem can be avoided to some degree by performing the scaling on partial sums (Brum, et al. 1976). To see this let

$$c_k(1) = \sum_{i=1}^{N_1} \frac{r_i s_{i+k}}{N_1} - \frac{1}{N_2} \sum_{i=1}^{N_1} \frac{r_i}{N_1} \sum_{i=1}^{N_1} \frac{s_{i+k}}{N_1}$$

$$c_k(2) = \sum_{i=N_1+1}^{2N_1} \frac{r_i s_{i+k}}{N_1} - \frac{1}{N_2} \sum_{i=N_1+1}^{2N_1} \frac{r_i}{N_1} \sum_{i=N_1+1}^{2N_1} \frac{s_{i+k}}{N_1} \qquad (4\text{-}35)$$

$$\vdots$$

$$c_k(N_2) = \sum_{i=N-N_1+1}^{N} \frac{r_i s_{i+k}}{N_1} - \frac{1}{N_2} \sum_{i=N-N_1+1}^{N} \frac{r_i}{N_1} \sum_{i=N-N_1+1}^{N} \frac{s_{i+k}}{N_1}$$

where $N_1 N_2 = N$. It is easily seen that

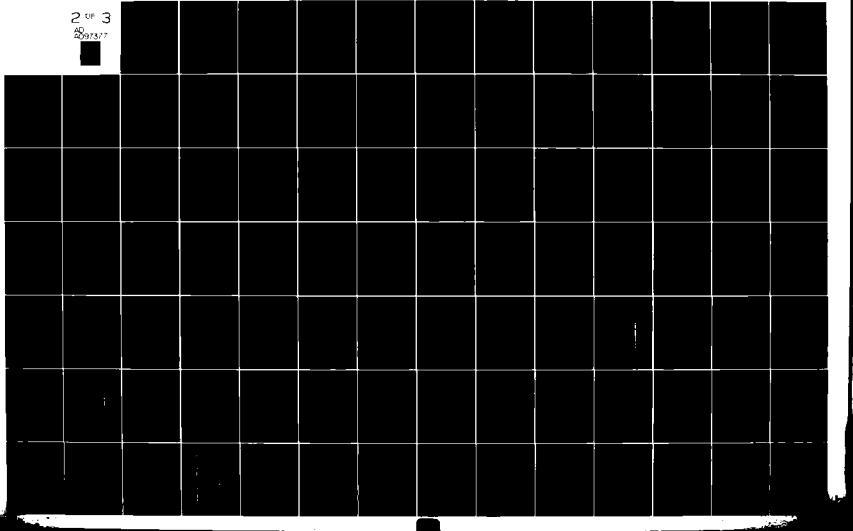
$$\frac{c_k}{N} = \sum_{i=1}^{N_2} \frac{c_k(i)}{N_2} \quad . \qquad (4\text{-}36)$$

Since $N_i < N$, $i = 1,2$ the loss in precision at each division is reduced. If the resulting error is still unacceptable, (4-36) can also be broken into partial sums and scaled accordingly. To analyze the accumulation of error in such a procedure we first determine the distribution of errors in the partial sum of (4-35) and then the accumulation of these errors is analyzed using (4-36). At this point the subscript k will be dropped since it serves no useful purpose in the analysis.

In addition to the round-off error, there is quantization error due to the analog-to-digital conversion. Since

AD-A097 377 ARIZONA UNIV TUCSON DIGITAL IMAGE ANALYSIS LAB F/G 8/2
AUTOMATIC METHODS IN IMAGE PROCESSING AND THEIR RELEVANCE TO MA--ETC(U)
FEB 81 B R HUNT DAAG29-77-G-0175
UNCLASSIFIED SIE/DIAL-81-002 ARO-14782.1-MS NL

2 OF 3
AD A097377



the quantization error is additive, the output error due to quantization has precisely the same form as the output error due to sensor noise. It follows from (4-9) that

$$\sigma^2_{c_q} \cong \frac{1}{N_1} \sigma_q^2 [2\sigma_g^2 + \sigma_q^2] \tag{4-37}$$

where $\sigma^2_{c_q}$ is the output variance due to input quantization noise having variance σ_q^2. For quantization to b bits, it can be easily shown that [Oppenheim and Schafer 1975]

$$\sigma_q^2 = \frac{2^{-2b}}{12} \quad . \tag{4-38}$$

The factor of N_1 in (4-37) (when compared to (4-9)) is the result of division by N_1 in (4-35). From (4-33), (4-34), and (4-35) it follows that the error in the partial sum injected by round-off and scaling is given by

$$\begin{aligned} e &= \frac{\hat{c}}{N_1} - \frac{c}{N_1} \\ &= \sum_{i=1}^{N_1} \left(\frac{\xi_i}{N_1} + \sigma_i\right) - \left\{\frac{1}{N_2} \sum_{i=1}^{N_1} \frac{r_i}{N_1} \sum_{i=1}^{N_1} \delta_i + \sum_{i=1}^{N_1} \frac{s_i}{N_1} \sum_{i=1}^{N_1} \eta_i \right. \\ &\quad \left. + \sum_{i=1}^{N} \eta_i \sum_{i=1}^{N} \delta_i + \beta \right\} \end{aligned} \tag{4-39}$$

where β is the error incurred with division by N_2. For round-off errors, $E\{e\} = 0$. For truncation errors, there is a bias term since truncation always rounds down when the numbers to be truncated are positive. We will ignore the distinction here and consider only round-off errors. Since $\{\xi\}$, $\{\sigma\}$, $\{\eta\}$, and $\{\delta\}$ are statistically independent and, we assume, uncorrelated with $\{s\}$ and $\{r\}$, it follows that

$$E(e^2) = \sigma_e^2$$

$$= E\left(\sum_{i=1}^{N_1} \frac{\delta_i}{N_1}\right)^2 + E\left(\sum_{i=1}^{N_1} \sigma_i\right)^2 + \frac{1}{N_2^2}\left[E\left(\sum_{i=1}^{N_1} \frac{r_i}{N_1} \sum_{i=1}^{N_1} \delta_i\right)^2\right.$$

$$\left. + E\left(\sum_{i=1}^{N_1} \frac{s_i}{N_1} \sum_{i=1}^{N_1} n_i\right)^2 + E\left(\sum_{i=1}^{N_1} n_i \sum_{i=1}^{N_1} \delta_i \sum_{i=1}^{N_1} \delta_i\right)^2\right] + E(\beta^2)$$

$$= \frac{1}{N_1}\sigma_q^2 + N_1\sigma_q^2 + \frac{2}{N_2^2}\sigma_q^2 \sum_{i=1}^{N_1}\sum_{j=1}^{N_1} R(i-j) + \frac{1}{N_2^2}\sigma_q^4 + \sigma_q^2 \qquad (4\text{-}40)$$

where $R(n) = R_s(n) = R_r(n) = R_g(n) + \sigma_n^2\delta(n)$. Since $\sigma_q^2 >> \sigma_q^4$, (4-40) is approximated by

$$\sigma_e^2 \cong \sigma_q^2\left[N_1 + \frac{1}{N_1} + 1 + \frac{2}{N_2^2}\left(N_1\sigma_n^2 + \sum_{i=1}^{N_1}\sum_{j-1}^{N_1} R_g(i-j)\right)\right]. \qquad (4\text{-}41)$$

Furthermore, $\frac{1}{N_1} << N_1$ and since $\frac{2}{N_2^2} N_1\sigma_n^2 < 1$ for reasonable choices of N_1 and N_2 it follows that

$$\sigma_e^2 \cong \sigma_q^2\left[N_1 + 1 + \frac{2}{N_2^2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_1} R_g(i-j)\right] \quad . \qquad (4\text{-}42)$$

Since $R_g(n) \leq R_g(0)$ it is easily shown that

$$\frac{2}{N_2^2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_1} R_g(i-j) \leq 2\frac{N_1^2}{N_2^2}(\mu_g^2 + \sigma_r^2) \qquad (4\text{-}43)$$

where $\mu_g = E\{g\}$. It is usually the case that $N_2 > N_1$ and since $\mu_g^2 < 1$ and $\sigma_g^2 < 1$ it follows that the error variance

associated with partial sums is approximately given by

$$\sigma_e^2 \cong \sigma_q^2[N_1 + 1] \quad . \tag{4-44}$$

Replacing $\hat{c}_k(i)$ in (4-36) by $c(i) + e_i$ and accounting for the division round-off error, e_i', we have

$$\frac{\hat{c}}{N} = \sum_{i=1}^{N_2} \left(\frac{c(i) + e_i}{N_2} + e_i'\right) \quad . \tag{4-45}$$

The total accumulated error is thus given by

$$e_t = \frac{1}{N_2} \sum_{i=1}^{N_2} e_i + \sum_{i=1}^{N_2} e_i' \quad . \tag{4-46}$$

Therefore, the round-off error variance, σ_{ro}^2, is

$$\sigma_{ro}^2 = \frac{1}{N_2} \sigma_e^2 + N_2\sigma_q^2 \tag{4-47}$$

$$\cong \frac{1}{N_2} \sigma_q^2(N_1 + 1) + N_2\sigma_q^2$$

$$\cong \sigma_q^2\left(\frac{N_1 + 1 + N_2^2}{N_2}\right) \quad . \tag{4-48}$$

But $N_2 = N/N_1$ so

$$\sigma_{ro}^2 \cong \sigma_q^2\left(\frac{N_1^3 + N_1^2 + N^2}{NN_1}\right) \quad . \tag{4-49}$$

Clearly, σ_{ro}^2 can be minimized by proper selection of N_1 although N_1 and N_2 are usually chosen for convenience in implementation. The AS-11B-X correlator uses $N_1 = 8$ and N_2 is adjusted to alter the window size (recall that $N = N_1N_2$). For large values of N (say $N > 64$),

$$\sigma_{ro}^2 \cong \frac{N}{8} \sigma_q^2 \tag{4-50}$$

Note that if $N_1 = N$ and $N_2 = 1$, then from (4-49)

$$\sigma_{ro}^2 \cong (N + 2)\ \sigma_q^2 \qquad (4\text{-}51)$$

so the scaling of partial sums has reduced the round-off error by slightly more than a factor of 8 in this case. This translates to a gain of approximately 1.5 bits of effective precision.

Floating Point Processing

Floating-point processing eliminates the need for data scaling since register overflow conditions can be detected and handled by the exponent of the floating-point number. For this reason, we need only consider round-off errors. To simplify the analysis, each of the two terms in the covariance function is analyzed separately. First consider the sum of products term

$$c_1 = \sum_{i=1}^{N} r_i s_i \quad . \qquad (4\text{-}52)$$

The accumulation of error in the computation of a sum of products has been analyzed [Oppenheim and Schafer 1975, p.439] in the context of quantization errors due to finite register lengths in floating point digital filters. The approach here is similar although the results developed here apply specifically to correlation.

With floating-point computations the errors are signal dependent and may be written as

$$e_i = r_i s_i \xi \qquad (4\text{-}53)$$

where we assume that ξ is a sample from a distribution uniform on $(-\frac{2^{-b}}{2}, \frac{2^{-b}}{2})$ where, again, b is the available register length. As before, the round-off error variance is given by (4-38). If we let $\{S_i\}$ be the sequence of partial sums:

$$S_\ell = \sum_{i=1}^{\ell} r_i s_i \tag{4-54}$$

then

$$\begin{aligned} S_1 &= r_1 s_1 (1 + \xi_1) \quad , \\ S_2 &= [r_2 s_2 (1 + \xi_2) + S_1](1 + \eta_2) \\ &= [r_2 s_2 (1 + \xi_2) + r_1 s_1 (1 + \xi_1)](1 + \eta_2) \quad , \\ S_3 &= [r_3 s_3 (1 + \xi_3) + S_2](1 + \eta_3) \\ &= [r_3 s_3 (1 + \xi_3) + r_1 s_1 (1 + \xi_2)(1 + \eta_2) \\ &\quad + r_1 s_1 (1 + \xi_1)(1 + \eta_2)](1 + \eta_3) \end{aligned} \tag{4-55}$$

where $\{\eta\}$ is the sequence of addition round-off errors which are also assumed to be zero-mean Gaussian with variance given by (4-38).

Continuing in the above manner, it can be shown that

$$S_N = \sum_{j=1}^{N} s_j r_j (1 + \xi_j) \prod_{i=j}^{N} (1 + \eta_i) \quad ; \quad \eta_1 = 0 \tag{4-56}$$

From the assumptions on $\{\xi\}$ and $\{\eta\}$ we can write

$$S_N = \sum_{j=1}^{N} A_j r_j s_j \tag{4-57}$$

where

$$A_j = \prod_{i=j}^{N} (1 + \xi_i) \tag{4-58}$$

and we have dismissed with the distinction between $\{\xi\}$ and $\{\eta\}$. The cumulative round-off error, e_N, is then given by

$$\begin{aligned} e_N &= S_N - c_1 \\ &= \sum_{i=1}^{N} A_i r_i s_i - \sum_{i=1}^{N} r_i s_i \\ &= \sum_{i-1}^{N} (A_i - 1)\, r_i s_i \quad . \end{aligned} \tag{4-59}$$

Taking expectations in (4-59) with respect to round-off noise, we have

$$E\{e_N\} = \sum_{i=1}^{N} r_i s_i \, E(A_i - 1) \quad . \tag{4-60}$$

From (4-58), it is seen that

$$E\{A_j\} = \prod_{i=j}^{N} (1 + E\{\xi\}) = 1 \tag{4-61}$$

so that

$$E\{e_N\} = 0 \quad . \tag{4-62}$$

The correlation output variance is thus given by

$$\begin{aligned} \sigma_{e_N}^2 &= E(e_N^2) \\ &= E\Big(\sum_{i=1}^{N} r_i s_i (A_i - 1) \sum_{j=1}^{N} r_j s_j (A_j - 1) \Big) \\ &= \sum_{i=1}^{N} \sum_{j=1}^{N} r_i r_j s_i s_j \, E\{(A_i - 1)(A_i - 1)\} \quad . \end{aligned} \tag{4-63}$$

Now

$$E((A_i - 1)(A_j - 1)) = E(A_i A_j) - 1 \quad ,$$

and

$$E(A_i A_j) = E\left(\prod_{=i}^{N} (1 + \xi) \prod_{k=j}^{N} (1 + \xi_k) \right)$$

$$= E\left((1 + \xi_N)^2 (1 + \xi_{N-1})^2 \ldots (1 + \xi_L)^2 (1 + \xi_{L-1}) \right.$$

$$\left. \ldots (1 + \xi_M) \right) \tag{4-64}$$

where

$$\begin{aligned} L &= \max(i,j) \\ M &= \min(i,j) \end{aligned} \quad . \tag{4-65}$$

Therefore

$$\begin{aligned} E(A_i A_j) &= \left[E(1 + \xi)^2 \right]^{N-L+1} \\ &= (1 + \sigma_q^2)^{N-\max(i,j)+1} \end{aligned} \quad . \tag{4-66}$$

It can easily be shown that if $\varepsilon << 1$, then

$$(1 + \varepsilon)^M \cong M\varepsilon + 1 \quad . \tag{4-67}$$

Combining (4-63) to (4-67), we have

$$\sigma_{e_N}^2 = \sigma_q^2 \sum_{i=1}^{N} \sum_{j=1}^{N} r_i r_j s_i s_j (N - \max(i,j) + 1) \quad . \tag{4-68}$$

The value of $\sigma_{e_N}^2$ clearly depends upon the input data sequences. Reintroducing the shift parameter k and taking expectations over the input processes yields

$$\sigma_{e_N}^2(k) = \sigma_q^2 \sum_{i=1}^{N} \sum_{j=1}^{N} E(r_i r_j s_{i+k} s_{j+k})(N - \max(i,j) + 1) \quad . \tag{4-69}$$

As before, the fourth order moment can be simplified for zero-mean Gaussian processes for which (4-21) applies. For non-zero-mean processes, consider

$$E((r_i' + \mu)(r_j' + \mu)(s_m' + \mu)(s_n' + \mu)) = E(r_i r_j s_m s_n)$$

where we assume the processes have identical means and the primed quantities are zero-mean. Since all odd order moments of zero-mean Gaussian processes are zero [Whalen 1971, p.97] it follows that

$$E(r_i r_j s_m s_n) = E(r_i' r_j' s_m' s_n') + \mu^2 C_r(i-j) + C_s(m-n) + C_{rs}(i-m) + C_{rs}(i-n) + C_{rs}(j-m) + C_{rs}(j-n) + \mu^4 \quad . \quad (4\text{-}70)$$

Combining (4-21), (4-69), and (4-70), we have

$$\sigma^2_{e_N}(k) = \sigma_q^2 \sum_{i=1}^{N} \sum_{j=1}^{N} [\, C^2(i-j) + C_{rs}^2(k) + C_{rs}(i+k-j)C_{rs}(j+k-i) + \mu^2(2C(i-j) + 2C_{rs}(i-j) + C_{rs}(j+k-i) + C_{rs}(i+k-j)) + \mu^4] \max(i,j) \quad (4\text{-}71)$$

where $C(n) = C_r(n) = C_s(n)$ and $\max(i,j)$ replaces $N - \max(i-j) + 1$ due to the symmetry of the autocovariance functions and the fact that all terms depend only on the difference $i - j$ and not on these individual values.

To complete the analysis of round-off errors in floating-point correlators we need to consider the term

$$C_2 = \frac{1}{N} \Sigma\, r_i \; \Sigma\, s_i \quad . \quad (4\text{-}72)$$

If the images are quantized to $N = 2^b$ levels, then errors in each of the summations in (4-72) occur only as a result of overflow. Thus to develop an expression for the cumulative error in (4-72) we should consider the probability that such an overflow will occur with each addition, or equivalently, the expected total number of overflows that occurs in the

computation of (4-72). Even if the distributions of r and s were accurately known, however, the computation of overflow probability would be cumbersome since the likelihood of an overflow at a given addition operation depends on whether or not an overflow occurred during previous additions. To avoid these complications, we assume that round-off errors occur with each addition and realize that the result will overestimate the error variance. We thus consider

$$c_2(k) = \frac{1}{N} \sum_{i=1}^{N} r_i(1+\xi_i) \sum_{i=1}^{N} s_{i+k}(1+\eta_i)\,(1+e_1)\,(1+e_2) \qquad (4\text{-}73)$$

where $\xi_1 = \eta_1 = 0$ and e_1 and e_2 result from the computation of the product of sums and division by N respectively. The accumulated error in (4-73) is given by

$$e_s(k) = c_2 - \frac{1}{N} \sum_{i=1}^{N} r_i \sum_{i=1}^{N} s_{i+k} \qquad (4\text{-}74)$$

$$= \frac{1}{N}(1+e_1)(1+e_2) \sum_{i=1}^{N} r_i \sum_{i=1}^{N} s_{i+k}\eta_i + \sum_{i=1}^{N} s_{i+k} \sum_{i=1}^{N} r_i\xi_i + \sum_{i=1}^{N} r_i\xi_i \sum_{i=1}^{N} s_{i+k}\eta_i + \frac{1}{N}(e_1+e_2+e_1e_2) \sum_{i=1}^{N} r_i \sum_{i=1}^{N} s_{i+k} \qquad (4\text{-}75)$$

The statistical independence of error terms implies that

$$\sigma_{e_s}^2(k) = E(e_s^2(k)) = \frac{1}{N^2}(1+\sigma_q^2)^2 E\Big[\Big(\sum_{i=1}^{N}\sum_{j=1}^{N} r_i s_{j+k}\eta_j\Big)^2 + \Big(\sum_{i=1}^{N}\sum_{j=1}^{N} r_i s_{j+k}\xi_i\Big)^2 + \Big(\sum_{i=1}^{N}\sum_{j=1}^{N} r_i\xi_i s_{j+k}\eta_j\Big)^2\Big] + \frac{1}{N^2} E(e_1+e_2+e_1e_2)^2\, E\Big(\sum_{i=1}^{N}\sum_{j=1}^{N} r_i\, s_{j+k}\Big)^2 \quad . \qquad (4\text{-}76)$$

Eliminating all terms involving powers of σ_q greater than 2(since $\sigma_q^4 << \sigma_q^2$) we have

$$\sigma_{e_s}^2(k) \cong \frac{2}{N^2}\sigma_q^2 \sum_i^N \sum_j^N \sum_m^N E(r_i r_j s_{m+k}^2) + \sum_i^N \sum_j^N \sum_m^N \sum_n^N E(r_i r_j s_{m+k} s_{m+k}) \quad .(4\text{-}77)$$

Application of (4-70) results in

$$\begin{aligned}\sigma_{e_s}(k) \cong \frac{2}{N^2}\sigma_q^2 \sum_i \sum_j \sum_m & (C(i-j)C(0) + 2C(m+k-j)C(m+k-i) \\ & + \mu^2(c(i-j) + C(0) + rC_{rs}(m+k-i)) + \mu^2) \\ & + \sum_i \sum_j \sum_m \sum_n (C(i-j)C(n-m) + C_{rs}(m+k-i)C_{rs}(n+k-j) \\ & + C_{rs}(n+k-i)C_{rs}(m+k-j + \mu^2(c(i-j)+C(m-n) + C_{rs}(m+k-i) \\ & + C_{rs}(n+1-i) + C_{rs}(m+k-j) + C_{rs}(n+k-j)) + \mu^4) \quad . \qquad (4\text{-}78)\end{aligned}$$

Summary of Correlation Noise Quantities

At this point, it is worthwhile to collect the various error terms for purposes of comparison. The expression for sensor noise, self-noise, and machine noise are listed below.

Sensor Noise: From (4-9),

$$\sigma^2_{e_c}(k) \cong N\ \sigma_n^2\ (2\sigma_g^2 + \sigma_n^2) \quad . \tag{4-79}$$

Self-Noise: From (4-23), removing the sensor noise,

$$\sigma_c^2(k) \cong \sum_{i=1}^{N}\sum_{j=1}^{N} [C_g^2(i-j) + C_g(j+k-i)C_g(i+k-j)] \quad . \tag{4-80}$$

A/D Quantization Noise: From (4-37),

$$\sigma^2_{c_q}(k) \cong \frac{1}{N}\ \sigma_q^2\ (2\sigma_g^2 + \sigma_q^2) \quad . \tag{4-81}$$

Fixed-Point Scaling and Round-off Noise: From (4-49),

$$\sigma^2_{sro}(k) \cong \frac{1}{N}\ \sigma_q^2\ \left(\frac{N_1^3 + N_1^2 + N^2}{N_1}\right) \quad . \tag{4-82}$$

Floating-Point Round-off Noise: From (4-71) and (4-78),

$$\begin{aligned}\sigma^2_{ro}(k) &= \sigma^2_{e_N}(k) + \sigma^2_{e_S}(k) \\ &\cong \sigma_q^2 \sum_i^N \sum_j^N [C^2(i-j) + C_{rs}^2(k) + C_{rs}(i+k-j)C_{rs}(j+k-i) \\ &\qquad + \mu^2(2C(i-j) + 2C_{rs}(i-j) + C_{rs}(j+k-i) \\ &\qquad + C_{rs}(i+k-j)) + \mu^4]\max(i,j)\end{aligned} \tag{4-83}$$

(contd)

$$+ \frac{6\sigma_q^2}{N^2} \sum_i \sum_j \sum_m C(i-j)C(0) + 2C(m+k-j)C(m+k-i)$$

$$+ \mu^2(c(i-j) + C(0) + rC_{rs}(m+k-i)) + \mu^4$$

$$+ \sum_i \sum_j \sum_m \sum_n C(i-j)C(m-n) + C_{rs}(m+k-i)C_{rs}(n+k-j)$$

$$+ C_{rs}(n+k-i)C_{rs}(m+k-j)$$

$$+ \mu^2(C(i-j) + C(m-n) + C_{rs}(m+k-i) + C_{rs}(n+k-i)$$

$$+ C_{rs}(m+k-j) + C_{rs}(n+k-j) + \mu^4 \qquad (4\text{-}83)$$

with the assumptions that

$$C(n) = C_g(n) + \sigma_n^2 \xi(n)$$

$$C_{rs}(n) = C_g(n) \qquad (4\text{-}84)$$

and $$\mu = E(r) = E(s) = E(g) \quad .$$

Example

The signal dependence of $\sigma_c^2(k)$ and $\sigma_{ro}^2(k)$ and the complexity of (4-83) make it difficult to ascertain the relative magnitudes of the noise terms for arbitrary source autocovariance functions. We must also note that the fixed point errors resulted from the computation of c_k/N rather than c_k so any comparison must account for this factor by multiplying $\sigma_{cq}^2(k)$ and $\sigma_{sro}^2(k)$ by N^2. Equivalently we could compute the correlator output SNR given by

$$\Lambda(k) = \frac{E^2(c(k))}{\sigma^2(k)} \qquad (4\text{-}85)$$

where $\sigma^2(k)$ represents any of the noise terms. There is some question as to whether Λ or $\Lambda^{\frac{1}{2}}$ should be called the output SNR. The argument in favor of $\Lambda^{\frac{1}{2}}$ is that the correlator output is actually a "power signal". We have selected Λ, however, since it represents the output signal-to-noise power ratio, regardless of the physical units.

For our example, consider

$$C_g(n) = \sigma_g^2 \exp\left[-\frac{1}{2}\left(\frac{n}{\Delta}\right)^2\right] \tag{4-86}$$

where Δ represents the correlation width. Since the noise and signal processes have been assumed to be Gaussian on the interval (0,1) (see (4-31)), we will let $\sigma_g^2 = .05$, $\Delta_g = .5$, and $\sigma_n^2 = 5 \times 10^{-4}$. This value of σ_g^2 implies that about 3% of the total area of the Gaussian distribution lies outside the unit interval. The input signal-to-noise ratio is 20dB. We also select $N = 64$ and $N_1 = 8$ since they are typical values employed by the AS-11B-X compiler. The correct peak output SNR's (at $k = 0$) associated with each noise source are plotted in Figure 4-2 for the selected parameters. For this example, it appears that self-noise is the dominating error source regardless of the correlation width. This result is misleading, however, because the nature of self-noise differs significantly from the other noise sources. The self-noise curve in Figure 4-2 represents the expected amount of fluctuation in the magnitude of the peak of cross-covariance function due to the random nature

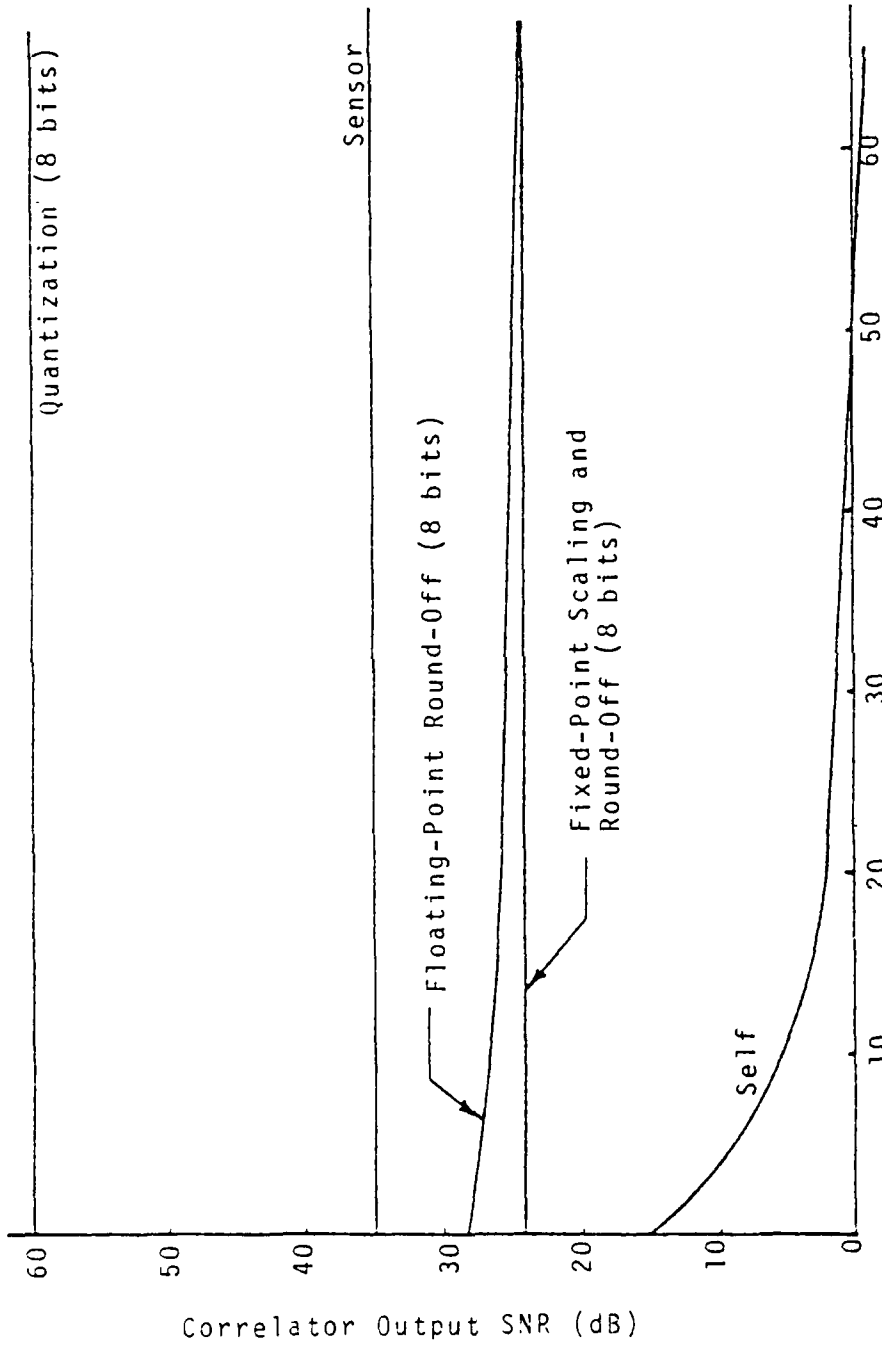


Figure 4-2. Correlator Output SNR. Gaussian Auto-Correlation, N = 64

of the underlying image formation process within a region of imagery that is assumed to be characterized, in the statistical sense, by the autocovariance function C_g. The remaining SNR curves, on the other hand, represent the random fluctuation in peak correlation magnitude at a given location due to the random nature of the noise source. A second fundamental difference between self-noise and the other noise sources is that the self-noise cross-covariance between taps increases with Δ. Figure 4-3 shows the relationship between the normalized autocovariance function and the normalized covariances between the center tap and all other taps for $N = 16$ and $\Delta = 1., 5., 10.,$ and $20.$, assuming the autocovariance function of (4-86). For this particular example, the tap covariances are slightly greater than the input signal covariances for identical function sample separations, i.e.,

$$E(C_0 C_\ell) > E(g_k g_{k+\ell}) \quad \forall \ \ell > 0 \quad . \tag{4-87}$$

These properties of self-noise make it difficult to determine the contribution of self-noise to correlation error production. Since our goal is the development of usable image quality measures, an experimental approach based on the probability of false acquisitions is developed in the following section. Two image quality features, one of which ignores the contribution of self-noise, are developed using (3-39). The performance of the measures in a detection scenario will be discussed in Chapter 6.

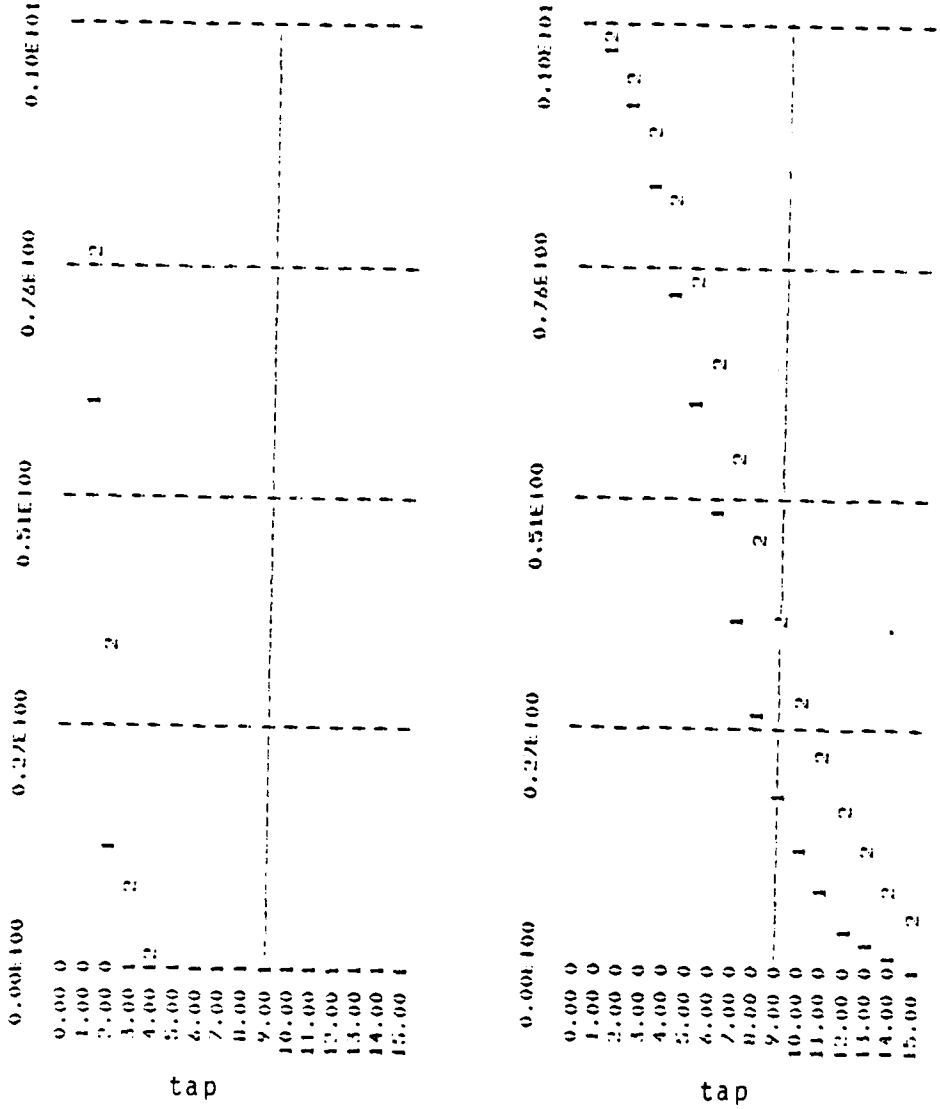


Figure 4-3. Center/Side Tap Covariance

N = 16
1 = data autocovariance function
2 = center/side tap covariance

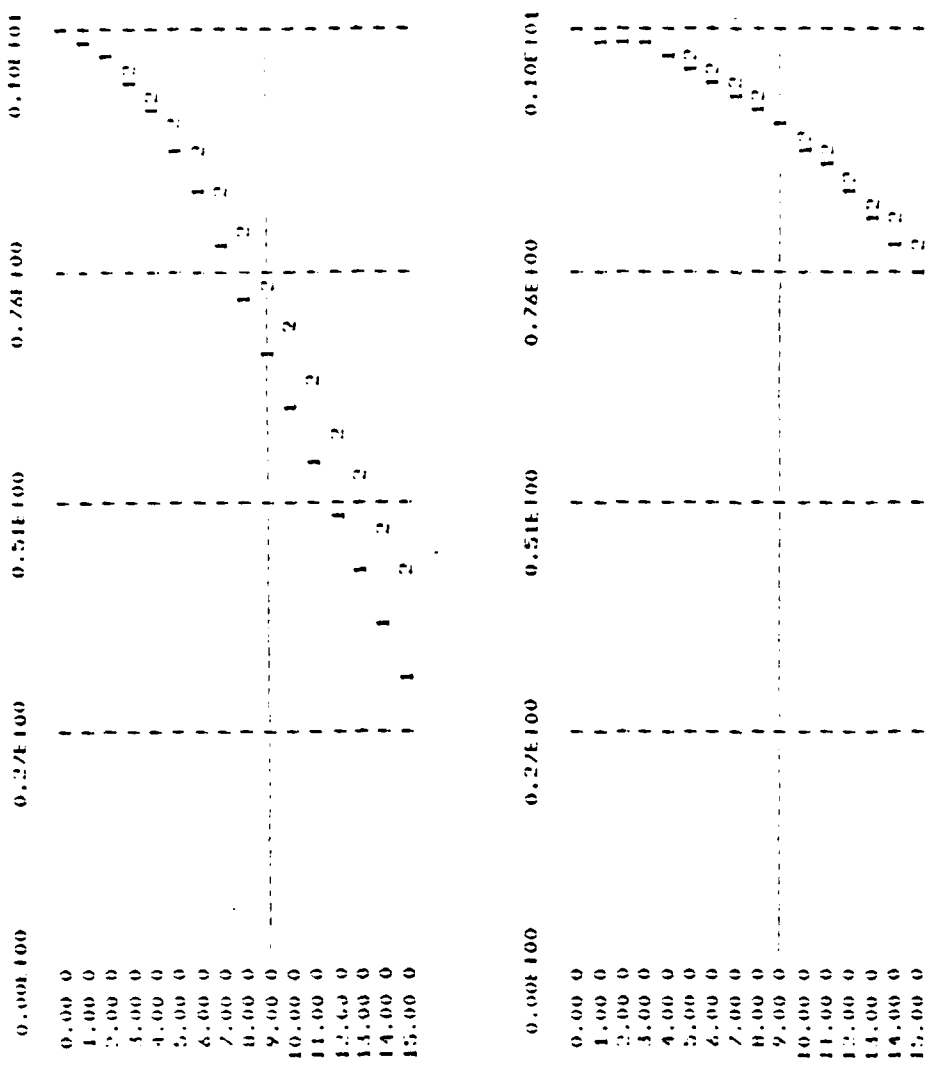


Figure 4-3 (continued). Center/Side Tap Covariance

False Acquisition Measures

Equation (3-39) defines an image quality measure

$$Q = \frac{N\sigma_g^2}{\sigma_c(k)} \tag{4-88}$$

where $\sigma_c(k)$ is the correlator output standard deviation at tap k due to noise injection of the various types described. For the present, we will ignore quantization noise and machine noise (simulations have been performed using floating point arithmetic with 24 bit precision which, provides an SNR of over 100 dB) and consider two expressions for Q based on sensor noise only for which

$$\sigma_c(k) = \sqrt{N}\,\sigma_n\,(2\sigma_g^2 + \sigma_n^2)^{\frac{1}{2}} \tag{4-89}$$

and for sensor noise plus self-noise for which

$$\sigma_c(k) = [N\,\sigma_n^2(2\sigma_g^2 + \sigma_n^2) + \sum\sum C_g^2(i-j)$$

$$+ C_g(i+k-j)C_g(j+k-i)]^{\frac{1}{2}} \tag{4-90}$$

which for large k becomes

$$\sigma_c(k) \approx [N\,\sigma_n^2(2\sigma_g^2 + \sigma_n^2) + \sum\sum C_g^2(i-j)]^{\frac{1}{2}} \tag{4-91}$$

For white noise input, (4-91) becomes

$$\sigma_c(k) = \sqrt{N}\,(\sigma_g^2 + \sigma_n^2) = \sqrt{N}\,\sigma_s^2 = \sqrt{N}\sigma_r^2 \tag{4-92}$$

where the model of (3-7) with $\alpha = 1$ is assumed. Since the contribution of self-noise to error production is unclear, and to avoid the computation of the autocovariance function (which may not be a good statistical representation of correlation behavior when estimated over small data sets), we introduce the factor λ and let

$$\sigma_c(k) = \sqrt{N}\left[\lambda N\sigma_g^4 + 2\sigma_n^2\sigma_g^2 + \sigma_n^4\right]^{\frac{1}{2}} \tag{4-93}$$

where

$$0 \leq \lambda \leq 1$$

since

$$0 \leq N\sigma_g^4 \leq \sum_{i=1}^{N}\sum_{j=1}^{N} C_g^2(i-j) \leq N^2\sigma_g^4 \quad . \tag{4-94}$$

If $\lambda = 0$, (4-93) reduces to (4-89) and if $\lambda = \frac{1}{N}$, (4-93) reduces to (4-92). While there are a variety of ways to make λ adaptive (such as computing (4-90) directly) we consider the expression

$$\lambda = 1 + \lambda_0\left(\frac{1}{N} - 1\right) \tag{4-95}$$

where

$$\lambda_0 \triangleq \begin{cases} \dfrac{\sigma_r^2 - \sigma_n^2}{\sigma_r^2} & \text{if } \sigma_r^2 > \sigma_n^2 \\ 0 & \text{otherwise} \end{cases} \tag{4-96}$$

This choice of λ is motivated by the reasonable assumption that low variance imagery typically displays a broad

autocorrelation function whereas high variance imagery tends to generate narrow autocorrelation functions. Substituting (4-93) into (4-88) yields

$$Q = \frac{\sqrt{N}\ SNR}{[\lambda N(SNR)^2 + 2SNR + 1]^{1/2}} \tag{4-97}$$

where $SNR = \sigma_g^2/\sigma_n^2$ is the input signal-to-noise ratio. Since σ_g^2 must be estimated from $\underline{r}$ (or $\underline{s}$) we compute

$$\hat{SNR} = \begin{cases} \dfrac{\sigma_r^2 - \sigma_n^2}{\sigma_n^2} & ; \ \sigma_r^2 > \sigma_n^2 \\ 0 & ; \text{otherwise} \end{cases} \tag{4-98}$$

Figure 4-4 displays the single tap false acquisition probability, P_{FA}, as a function of Q. The behavior of Q as a function of SNR and λ is shown in Figure 4-5 for N = 21 and 64. Note that the larger fixed values of λ (say $\lambda > .5$) do not result in reasonable (or observed) values of P_{FA}. For example, for $\lambda = .5$, the single tap false acquisition probability cannot be improved beyond $P_{FA} = .1$ regardless of the input SNR and window size. While P_{FA} is only an indication of the true total false acquisition probability P_T (over all taps), it is necessary that $P_T > P_{FA}$.

Regardless of how Q is computed, there is a one-to-one relationship between SNR and P_{FA}. Thus it is reasonable to consider the alternative quality measure Q_1 given by

$$Q_1 = SNR \tag{4-99}$$

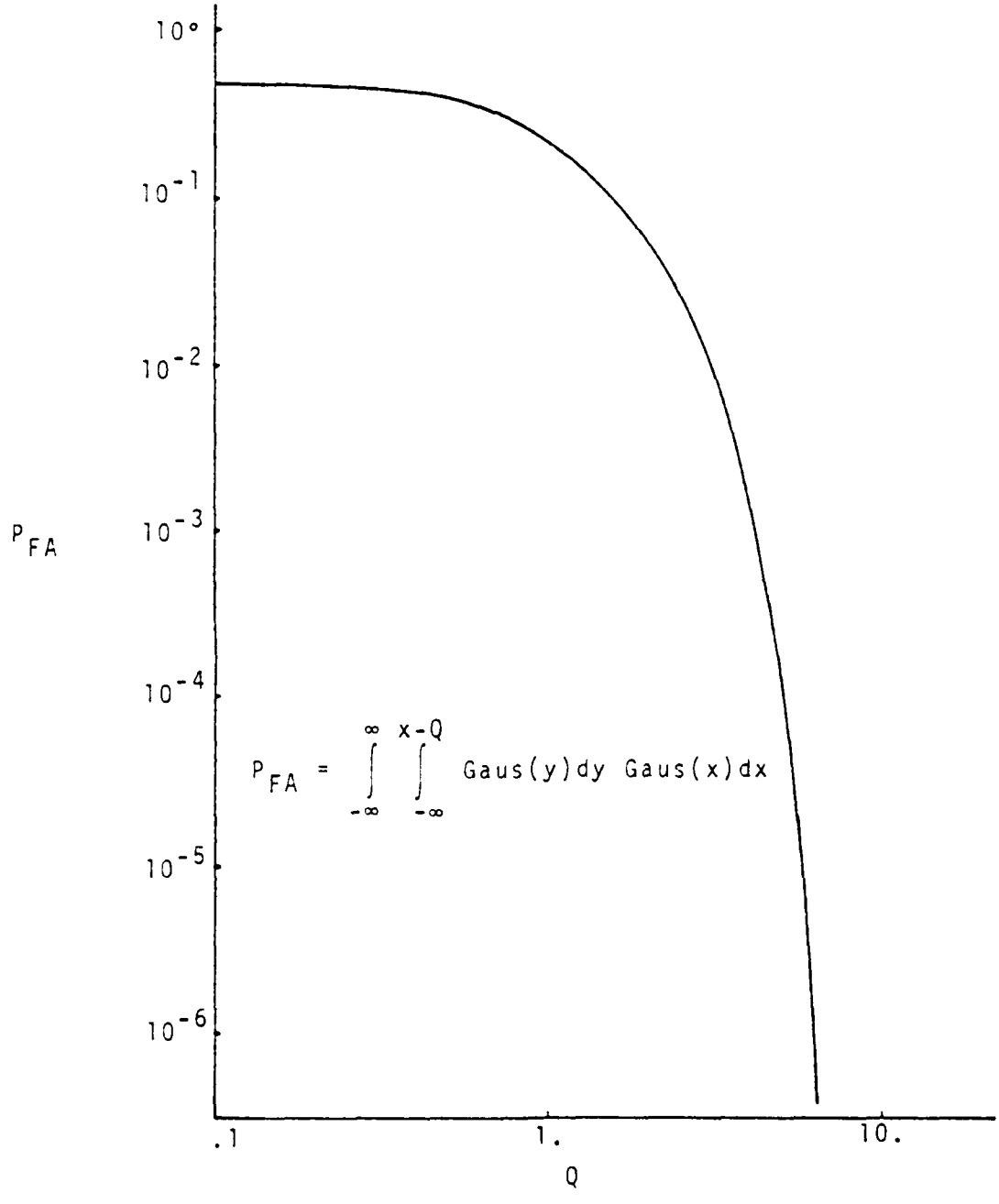


Figure 4-4. Single Tap Probability of False Acquisition vs. Q

which is also shown in Figure 4.5. If σ_n^2 is a known constant then Q_1 varies directly with the local image variance. The use of the quality measures Q and Q_1 in a threshold detection algorithm will be discussed in Chapter 6. It is clear, however, that the choice of Q and λ alters the pdf of the quality measure which, in turn, affects the selection of appropriate threshold levels.

Estimated Mean Square Error

A closed form expression for the estimated mean square error defined by (4-3) can be derived by extending the result in (3-36) if we assume that the correlator tap outputs are statistically independent. If such is the case, then

$$P_\ell = \int_{-\infty}^{\infty} \prod_{\substack{n=-L \\ n\neq\ell}}^{L} \int_{\infty}^{x} p_{c_n}(y)dy \; p_{c_\ell}(x)dx \qquad (4\text{-}100)$$

where $p_{c_k}(x)$ is the correlator output pdf at tap k. Under the assumption of Gaussian pdf's, (4-100) becomes

$$P_\ell = \int_{-\infty}^{\infty} \prod_{\substack{n=-L \\ n\neq\ell}}^{L} \int_{-\infty}^{u(x,n,\ell)} \text{Gaus}(y)dy \; \text{Gaus}(x)dx \qquad (4\text{-}101)$$

where

$$u(x,n,\ell) = \frac{\sigma_c(\ell)x + N(C_{rs}(\ell) - C_{rs}(n))}{\sigma_c(n)} \quad . \qquad (4\text{-}102)$$

Computation of (4-101) generally leads to overestimates of $\hat{MSE}$ because the tap covariances are neglected. This can be

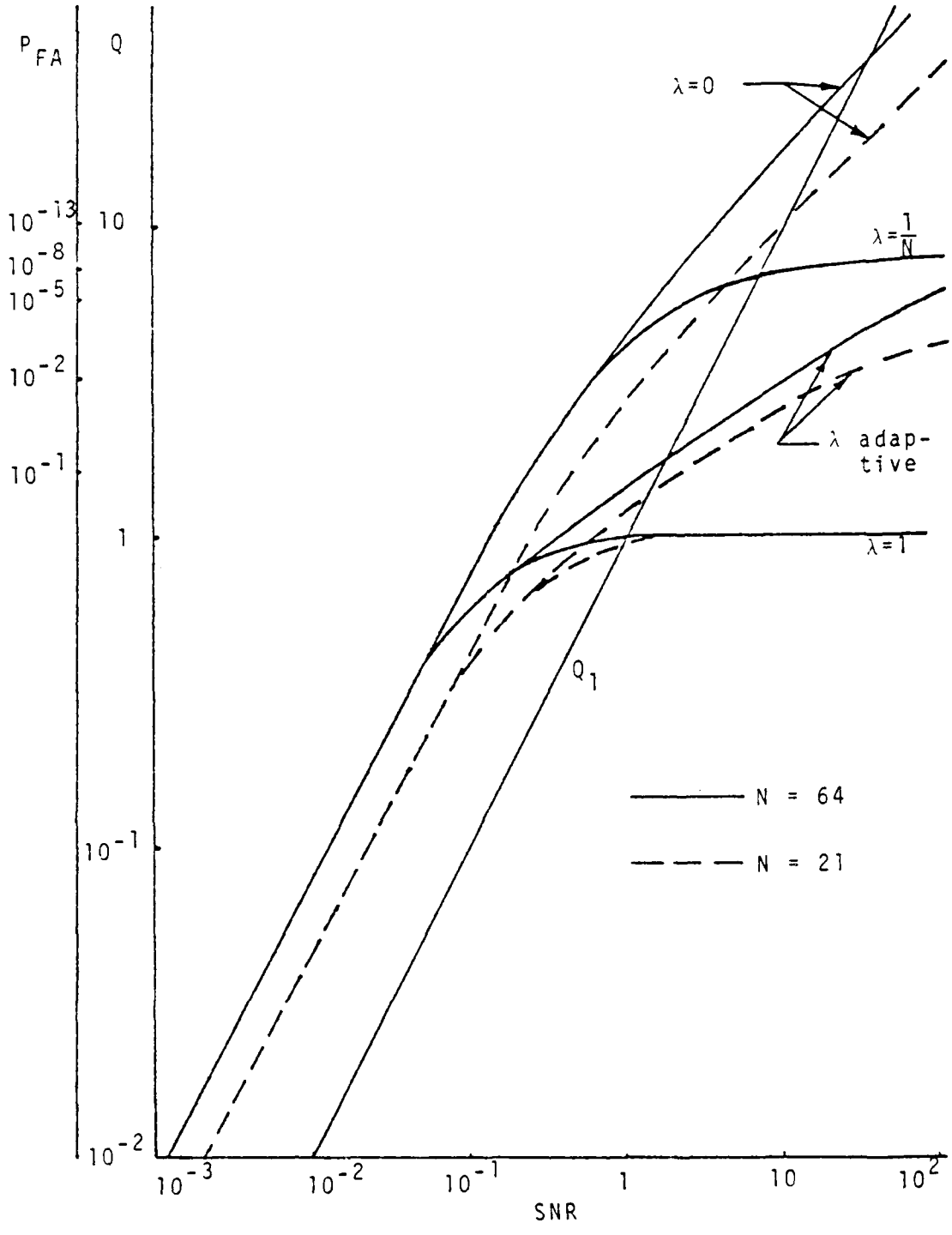


Figure 4-5. P_{FA} and Q vs. SNR

counteracted to some degree by imposing the constraint that [Ryan and Hunt 1980]

$$\sum_{\ell=-L}^{L} P_\ell = 1 \quad . \tag{4-103}$$

The computational complexity of (4-101), however, makes it cumbersome to use as an image quality feature.

Input Data Non-Stationarity

Virtually all of the literature on correlation behavior ignores the effects of non-stationarity of input data. Each of the expressions derived in the preceding sections require the assumption of stationarity to express the expected behavior in terms of ensemble statistical averages (means, variance, autocorrelations, etc.). To make use of these results, however, requires estimating the statistical averages from finite data records which, in turn, requires that the input data be both stationary and ergodic. Images, however, are notoriously non-stationary. The local statistics within a windowed image can change drastically as the window moves from one location to another particularly if the window passes an edge (a local discontinuity in image luminance or amplitude level) between reasonably smooth regions. Many image processing algorithms (for restoration, enhancement, etc.) are improved by adapting algorithm parameters to local image statistics to account for the non-stationarity of the image. Improvements in correlation accuracy can be achieved in a similar manner as the following discussion implies.

The assumptions leading to covariance as the ML processor include the assumption that the variance (or energy) of the search area subimage does not change appreciably over the search area. If this assumption is violated due to the presence of an edge within the search area, the resulting behavior of the covariance function is, in many cases, predictable as the following example implies.

Consider the sequence of picture element values shown in Figure 4-6. We select from this sequence a match window of length 5 centered on position 2. This match window is thus the sequence {4,6,4,5,8}. This window is then correlated with the entire sequence using the covariance function as a measure of similarity. The resulting correlation function is shown in Figure 4-7a. The location of the correct correlation peak is denoted by location 0. The peak actually occurred at location 4 so there is a registration error of 4 pixels. Figures 4-7b to 4-7j display similar results for all other match window positions selected from Figure 4-6. All match window positions chosen have maximum cross-covariance corresponding to center positions 6, 7, or 8 in Figure 4-6. These points form the transition region (edge) between two regions in which the statistics change much more slowly. The only correct correlation results in this experiment occurred for match window center positions 7 and 8. Since covariance removes the means from the sequences, the covariance processor can be thought of as imposing stationarity of the mean upon the process. However,

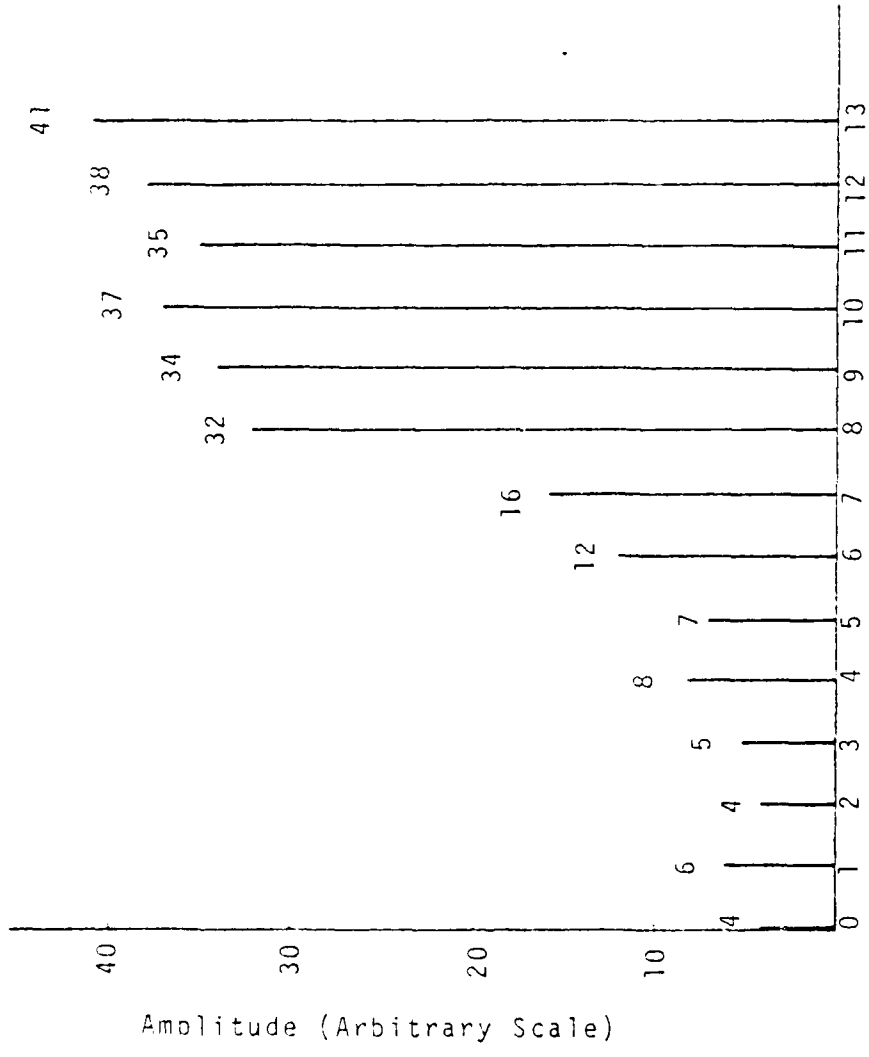


Figure 4-6. Sample Sequence of Pixels Forming on Edge

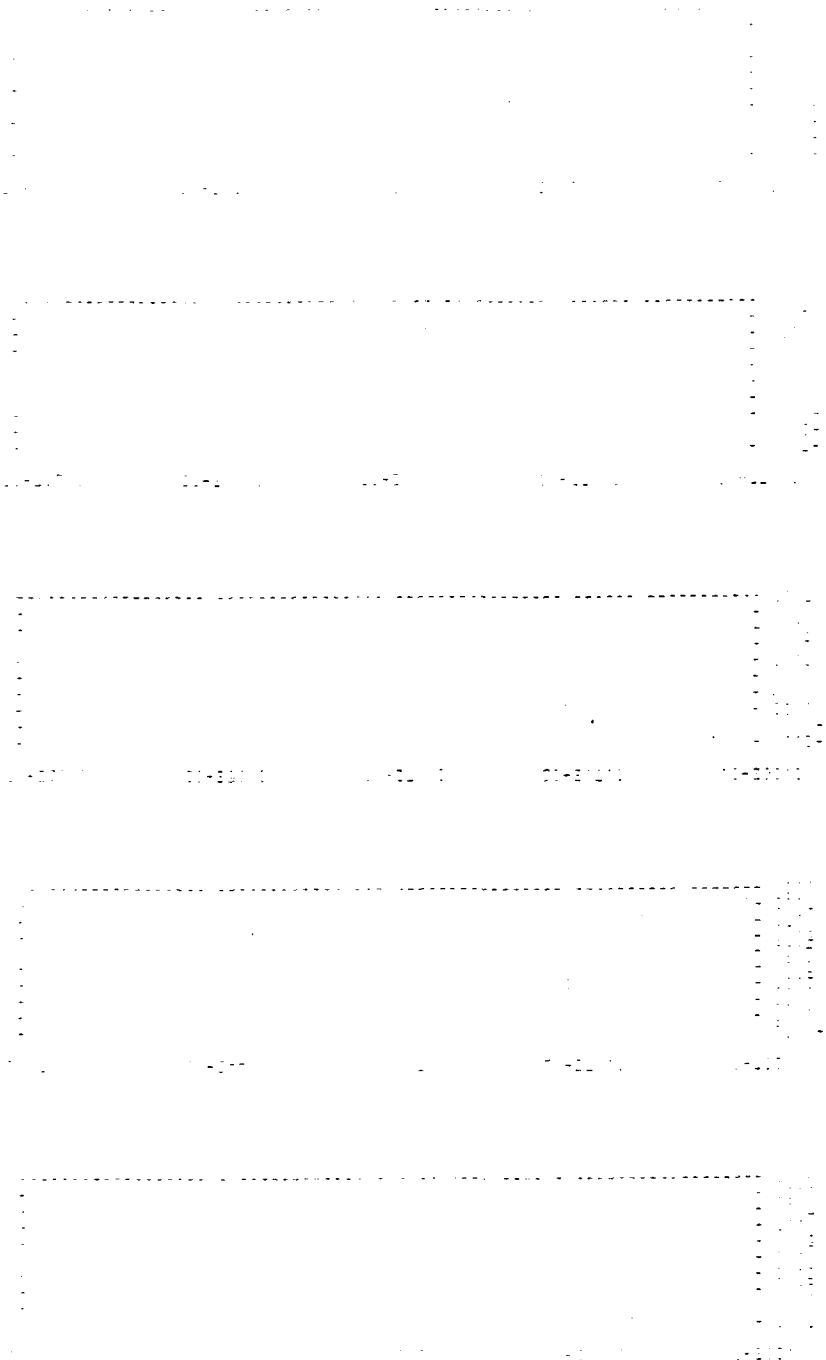


Figure 4-7. Covariance Functions for the Sequence in Figure 4-6

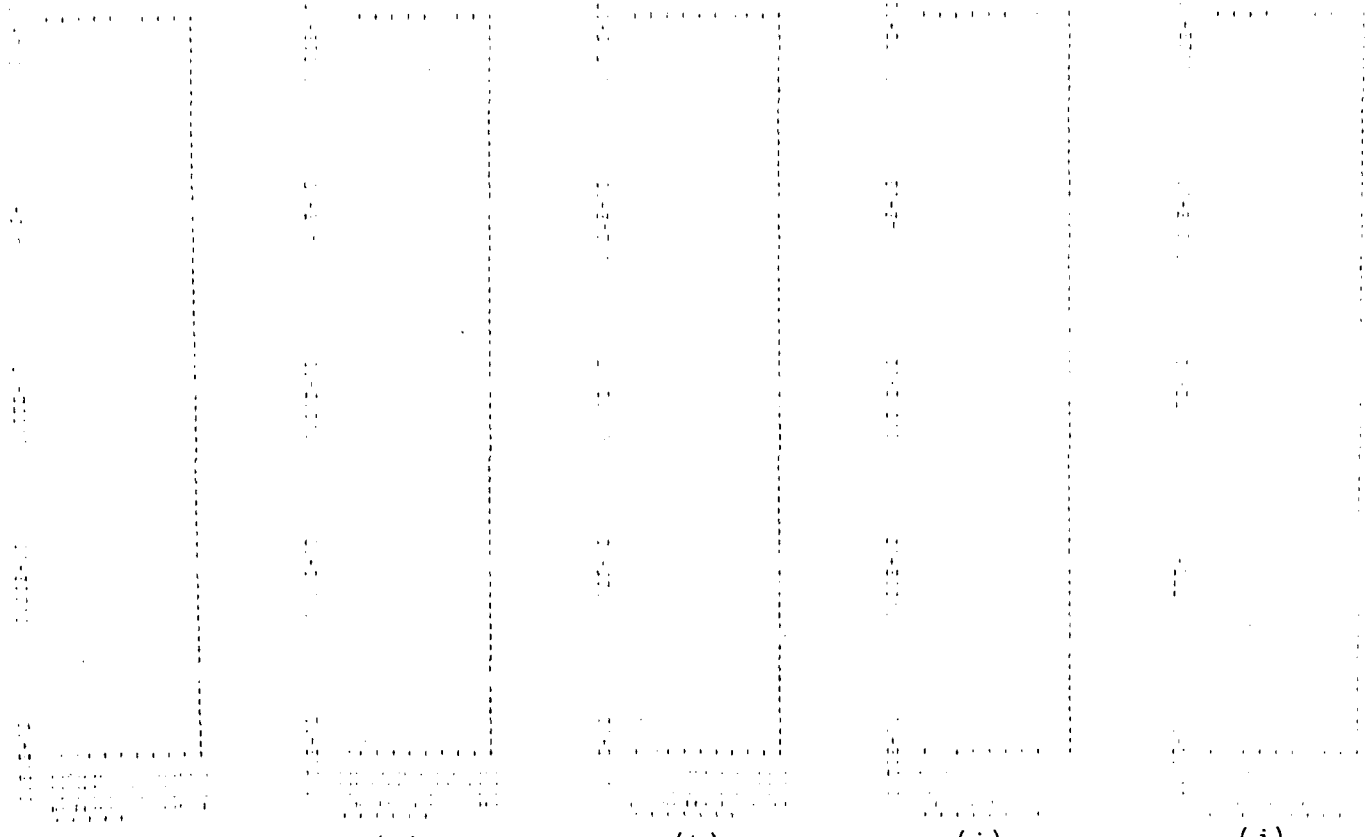


Figure 4-7 (continued). Covariance Functions for the Sequence in Figure 4-6

no correction is made for non-stationarity of the variance. Normalized covariance, however, imposes stationarity of the variance upon the process by normalizing the signal energy at each position. As seen in Figure (2-7), the error behavior appearing in Figure (2-5) does not occur. Normalized covariance is a special case of generalized least-squares given in (2-20) so we expect least squares to behave in a manner similar to normalized covariance in the vicinity of an edge. Figures 2-10 and 2-11 indicate that this is, indeed, the case.

A close comparison of Figures 2-9 and 2-4b (by using an overlay) shows that there are regions of considerable error located along the edges of Figure 2-4b. It is also clear that positive errors generally lie on one side of an edge, while negative errors lie on the opposite side. This behavior is also noted in Figure 4-6. It is thus reasonable to conclude that at least some of the errors occuring in Figure 2-9 as a result of covariance processing are due to the presence of edges which violate the assumption of image stationarity.

We present here two approaches to the detection of such regions. For stereo-pair images with negligible relief distortion, the value of the covariance function at the correct peak location is an estimate of the energy within the image region under observation. As before, we assume the distribution of image energy is Gaussian, i.e.,

$$p_e(x) = \frac{1}{\sqrt{2\pi}\sigma_c(0)} \exp\left[\frac{-(x - C(0))^2}{2\sigma_c^2(0)}\right] \quad . \qquad (4\text{-}104)$$

The presence of an edge-produced region of non-stationarity results in a large peak covariance value due to the high local energy (or variance) across the edge. Thus given the distribution of (4-104), we can select a threshold t such that

$$\text{Prob}[C_0 > t] = \int_t^\infty p_e(x)dx \qquad (4\text{-}105)$$

is some desired value. Such a detection scheme could be implemented either as a preprocessor or as a reliability measure used during the stereo compilation procedure itself. Detection of these regions would enable an adaptive processor to convert to normalized covariance or least-squares correlation in the region of expected difficulty. The processor returns to covariance processing when the numerator of the normalized covariance function or one-half of the first term in least-squares correlation falls below the selected threshold.

It is also interesting to consider the conditions under which covariance produces errors and least-squares (and presumably normalized covariance) produces correct peak locations. Suppose the correct hypothesis (using the notation of Chapter 2) is H_i but that covariance selects H_j. This condition implies that

$$\text{(covariance):}\quad r's^{(j)} - r's^{(i)} > 0 \tag{4-106}$$

$$\text{(least squares):}\quad 2r's^{(i)} - s^{(i)}'s^{(i)} < 2rs's(j) - s^{(j)}'s(j) \tag{4-107}$$

where we assume the means have been previously removed. Equations (4-106) and (4-107) can be combined to yield

$$s^{(j)}'s^{(j)} - s^{(i)}'s^{(i)} > 2(r's^{(j)} - r's^{(i)}) > 0 \tag{4-108}$$

This condition simply states that the covariance processor can select H_j only if the energy of subimage $s^{(j)}$ is greater than the energy of subimage $s^{(i)}$. It is at least intuitively reasonable that the quantities

$$\varepsilon_{ij} \triangleq s^{(j)}'s^{(j)} - s^{(i)}'s^{(i)} \tag{4-109}$$

and

$$\delta_{ij} \triangleq 2r'(s^{(j)} - s^{(i)}) \tag{4-110}$$

are correlated so that as ε_{ij} increases, the probability that $\delta_{ij} > 0$ also increases. If the distribution of δ_{ij} conditioned on ε_{ij}, $p(\delta_{ij}|\varepsilon_{ij})$ were known, it would be possible to establish thresholds on ε_{ij} which would allow the detection of image regions for which Prob $[\varepsilon_{ij} > 0]$ exceeds a maximum acceptable error rate. Such a distribution could be determined experimentally with sufficient computing power and a set of "representative" images. A processor which might be employed in such a detection scheme is shown in Figure 4-8.

Since the computation of $p(\varepsilon|\varepsilon)$ may be computationally prohibitive, a more realistic approach can be developed by noting that ε is related to the gradient of the image energy. A simplified pre-processor based on this realization is shown in Figure 4-9. In Chapter 6, we present experimental results to indicate the feasibility of such a procedure. This disadvantage of this approach is that the resulting error map indicates only the locations of points corresponding to high error probabilities and information concerning the error magnitude is lost. Such a binary error map could be used to switch the compilation processor from covariance to a more reliable technique, and this may be all that is required.

The Extension to Two Dimensions

The various correlator output noise variances given in (4-79) to (4-83) were derived under the assumption of one-dimensional source sequences and a one-dimensional search. Of these quantities, only self-noise, (4-80), and floating-point round-off noise, (4-83), are dependent on the signal statistics and thus, on the shape of the match window. Since both of these expressions result from the expansion of the fourth order moment given in (4-21), the extension to two-dimensional simply requires the corresponding expression for the two-dimensional case. Beginning with the 2-D expression for the covariance function given in (4-5) and ignoring the sensor noise, we have

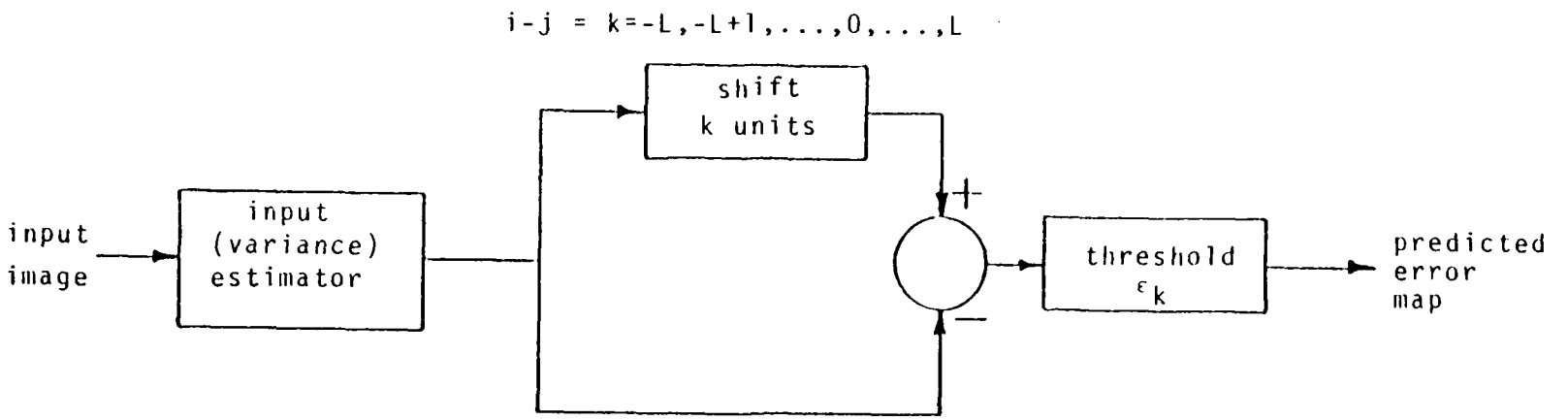


Figure 4-8. Detector for Edge-related Errors

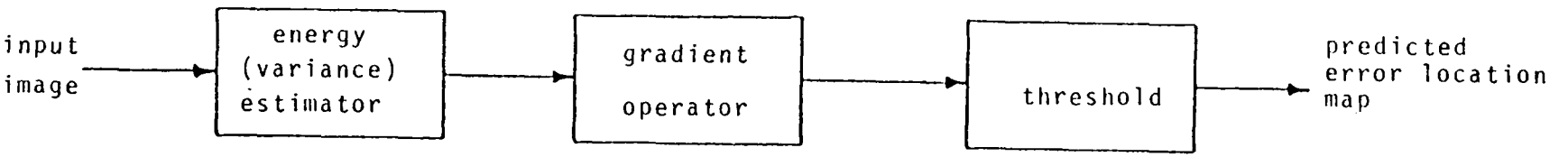


Figure 4-9. Simplified Error Location Detector

$$c(k,\ell) = \sum_{i=1}^{M}\sum_{j=1}^{N} g_{i,j}g_{i+k,j+\ell} - \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} g_{i,j}\sum_{i=1}^{M}\sum_{j=1}^{N} g_{i+k,j+\ell} \quad . \tag{4-111}$$

Assuming, as before, that the error in estimating the mean is negligible, we let g be zero-mean. Thus

$$E(c(k,\ell)c(m,n)) = \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{a=1}^{M}\sum_{b=1}^{N} g_{i,j}g_{i+k,j+\ell}g_{a,b}g_{a+m,b+n} \quad . \tag{4-112}$$

But the two-dimensional counterpart of (4-21) results in

$$\begin{aligned} E(c(k,\ell)c(m,n)) \cong & \sum_{i}\sum_{j}\sum_{a}\sum_{b} C_g(k,\ell)C_g(m,n) \\ & + C_g(a-i,b-j)C_g(a+m-i-k,b+n-j-\ell) \\ & + C_g(a+m-i,b+n-j)C_g(i+k-a,j+\ell-b) \quad . \end{aligned} \tag{4-113}$$

The individual tap variances result from (4-113) and (4-111) when $k = m$ and $\ell = n$ for which

$$\begin{aligned} \sigma_c^2(k,\ell) &= E(c^2(k,\ell)) - E^2(c(k,\ell)) \\ &\cong \sum_{i}\sum_{j}\sum_{a}\sum_{b} C_g^2(a-i,b-j) \\ &\quad + C_g(a+k-i,b+\ell-j)C_g(i+k-a,j+\ell-b) \quad . \end{aligned} \tag{4-114}$$

CHAPTER 5

RELIEF DISTORTION

In all of the preceding discussion, we have made the assumption that topographic relief is negligible so that the two image functions differ only by additive noise terms; that is, the underlying image functions are identical. The presence of topographic relief, however, introduces distortion between the two image functions. The primary effect of this distortion is a reduction in the peak value of the cross-correlation function which, in turn, increases the probability of both local registration and false acquisition errors. In this chapter, a model of relief distortion is developed and used to determine the effects of relief distortion on the covariance function and on the image quality measures previously discussed.

Relief Distortion Model

Consider the image formation geometry shown in Figure 1-2. The image generation equations were given in Equations (1-2) and (1-3) and repeated here:

$$g_1(x_1) = I\left(\frac{x_1(H - e(X))}{\ell}\right) \qquad (5\text{-}1a)$$

$$g_2(x_2) = I\left(B + \frac{x_2(H - e(X))}{\ell}\right) \quad . \qquad (5\text{-}1b)$$

If

$$x_2 = x_1 - \frac{B\ell}{H - e(X_0)} \tag{5-2}$$

where X_0 is as shown in Figure 1-2, then x_1 and x_2 are conjugate points and $g_1(x_1) = g_2(x_2)$. If there is no terrain relief, then $g_1(x_1 + \Delta x) = g_2(x_2 + \Delta x)$. In order to determine the effect of distortion on the image functions, consider

$$g_1(x_1 + \Delta x) = I\left[\frac{(x_1 + \Delta x)(H - e(X_1)}{\ell}\right] = I(X_1) \tag{5-3a}$$

and

$$g_2(x_2 + \Delta x_2) = I\left[B + \frac{(x_2 + \Delta x_2)(H - e(X_2)}{\ell}\right] = I(X_2) \tag{5-3b}$$

where $x_1 + \Delta x_1$ and $x_2 + \Delta x_2$ are the images of terrain points X_1 and X_2 respectively. We wish to determine the value of Δx_2 so that $X_1 = X_2 = X$. After some algebraic manipulation involving (1-1), (5-2), and (5-3), it follows that

$$\Delta x_2 = \Delta x_1 \left[1 - B \frac{e(X) - e(X_0)}{XH_0 - X_o(H - e(X))}\right] \tag{5-4}$$

where $H_0 = H - e(X_0)$. Letting

$$D(X) = \frac{e(X) - e(X_0)}{X - X_0} \quad , \tag{5-5}$$

we can rewrite (5-4) as

$$\Delta x_2 = \Delta x_1 \left[1 - B \frac{D(X)}{H_0 + X_0 D(X)}\right] \quad . \tag{5-6}$$

Finally, if we let

$$f_1(x) = g_1(x_1 + x)$$

and (5-7)

$$f_2(x) = g_2(x_2 + x) \quad ,$$

then the model for distortion can be written as

$$f_1(x) = f_2(h(x)) \tag{5-8}$$

where

$$h(x) = x\left[1 - B\,\frac{D(X)}{H_0 + X_0 D(X)}\right] \tag{5-9}$$

and x is related X by

$$X = \frac{x(H - e(X))}{\ell} \quad . \tag{5-10}$$

Equation (5-8) expresses the effects of relief distortion in terms of the functional composition [Apostal 1974] of an image function f_2 with a distortion function $h(x)$. Now suppose that $D(x) = m$ where m is a constant specifying the terrain slope. Then

$$h(x) = x\left[1 - \frac{mb}{H_0 + X_0 m}\right] \quad . \tag{5-11}$$

It is usually the case that $H_0 >> mX$ so that

$$h(x) \cong x\left[1 - \frac{mB}{H_0}\right] \triangleq \alpha x \quad . \tag{5-12}$$

This result implies that constant terrain slope produces spatial contraction or expansion of one image function with respect to the other. The value of α depends on the base-height ratio, B/H_0, and the magnitude of the terrain slope m, with distortion increasing with increases in either of these parameters.

The Effect of Distortion on Correlation Accuracy

The effect of signal distortion on the cross-correlation accuracy is most easily analyzed in the continuous domain. Consider the cross-covariance function

$$r(\xi) = \int_{-\frac{D}{2}}^{\frac{D}{2}} f(x + \xi)g(x)dx \qquad (5\text{-}13)$$

where D is the length of the match window and f and g are sample functions from stationary, zero-mean Gaussian processes. If $g(x) = Kf(x - x_0)$,

then

$$E(r(\xi)) = K \int_{-\frac{D}{2}}^{\frac{D}{2}} C_f(\xi + x_0)dx = KDC_f(\xi + x_0) \qquad (5\text{-}14)$$

and the expected peak location is the translation between the functions f and g as desired. If $g(x) \neq Kf(x - x_0)$ for fixed but arbitrary K and x_0, then we refer to g as a distorted version of f, assuming, of course, that f and g are similar and that true points of conjugacy exist.

We now assume the distortion model of (5-8) and let

$$g(x) = f(h(x)) \quad . \qquad (5\text{-}15)$$

Without loss of generality, let K = 1 and $x_0 = 0$ so that the presence of distortion implies $h(x) \neq x$. Substituting (5-15) into (5-13) and taking expectations gives

$$E(r(\xi)) = \int_{-\frac{D}{2}}^{\frac{D}{2}} C_f(\xi + x - h(x))dx \quad . \tag{5-16}$$

If $u(x) \triangleq h(x) - x$ is a monotonic function of x, then there exists a function z with $z(u) = x$ such that (5-16) can be written as

$$E(r(\xi)) = \int_{-\infty}^{\infty} z'(u)\,\mathrm{rect}\left(\frac{u - \bar{u}}{\tau}\right) C_f(\xi - u)du \tag{5-17}$$

where $z'(u) = dz/du$,

$$\mathrm{rect}(x) = \begin{cases} 1 \text{ for } |x| < \frac{1}{2} \\ \\ 0 \quad \text{otherwise} \end{cases} ,$$

$$\bar{u} = \frac{1}{2}\left[h\left(\frac{D}{2}\right) + h\left(-\frac{D}{2}\right)\right] \quad , \tag{5-18}$$

and

$$\tau = \left|D + h\left(-\frac{T}{2}\right) - h\left(\frac{T}{2}\right)\right| \quad . \tag{5-19}$$

The derivation of (5-17) is presented in Appendix B. The monotonic restriction on $u(x)$ implies that the distortion consists of spatial compression or spatial expansion but not both. For aerial imagery taken over reasonably smooth terrain, the restrictions are minor. If $u(x)$ is not monotonic, then the resulting ensemble correlation function can be determined directly from (5-16). Equation (5-17) is the convolution

$$E(r(\xi)) = \int z'(u) \frac{1}{D} \text{rect}\left(\frac{\xi-\bar{u}}{\tau}\right) * D\, C_f(\xi) \qquad (5\text{-}20)$$

which shows that the effect of relief distortion is to smooth the correlation function and reduce its peak value. Equation (5-20) also implies that the peak of the correlation function may be biased ($\bar{u} \neq 0$) so that the choice of the true correlation peak results in incorrect matching of imagery. The fluctuation of $\bar{u}$ provides a limit on the attainable accuracy even if the imagery is noise-free. It is straightforward to show that under the assumptions leading to (5-12),

$$E(\bar{u}) = 0 \qquad (5\text{-}21a)$$

$$E(\bar{u}^2) = \sigma_{\bar{u}}^2 = \frac{D^2}{8}\, \sigma_\alpha^2 \qquad (5\text{-}21b)$$

where

$$\sigma_\alpha^2 = E(\alpha^2) - 1 \quad . \qquad (5\text{-}22)$$

Equation (5-21b) implies that the registration bias fluctuation is proportional to the variation in α (i.e., variability of terrain slope) over the terrain region viewed through the match window. As D increases, the match window "sees" more terrain and it is likely that σ_α^2 will also increase. Thus while a larger match window provides better noise suppression, it increases the susceptibility to registration bias errors.

For arbitrary terrain, the function $z(\cdot)$ in (5-17) may be impossible to obtain in closed form (it may not be unique). It is obtainable, however, for the special but useful case in which h(x) is given by (5-12). Substituting

(5-12) into (5-20), it follows that

$$u(x) = (\alpha - 1)x \quad , \tag{5-23a}$$

$$z'(u) = \frac{1}{\alpha - 1} \quad , \tag{5-23b}$$

and

$$E(r(\xi)) \triangleq DC_{fg}(\xi) = \frac{1}{D|\alpha - 1|} \text{ rect}\left(\frac{\xi}{D(\alpha - 1)}\right) * DC_f(\xi) \quad . \tag{5-24}$$

As an example, let the image function f be white noise so that $C_f(\xi) = \delta(\xi)$, the Dirac impulse. Then (5-24) becomes

$$E(r(\xi)) = \frac{1}{|\alpha - 1|} \text{ rect } \left(\frac{\xi}{D(\alpha - 1)}\right) \tag{5-25}$$

in which the location of the peak is certainly ambiguous, but most importantly, the peak height is reduced. The reduction is most significant for large values of the product $D|\alpha - 1|$ and for imagery containing high spatial frequency energy resulting in a narrow autocorrelation function. The distortion characterized by $h(x) = \alpha x$ is analogous to the doppler shift which occurs in radar systems when there is relative motion between the target and signal source-receiver [Skolnik 1962].

The effect of the affine transformation $h(x) = \alpha x$ on correlator performance has been analyzed by Mostafavi and Smith [1978a, 1978b]. They also consider a more general affine transformation which includes a possible rotation between conjugate regions.

In radar systems the effect of a doppler shift on the matched filter mismatch is usually considered to be negligible unless the lengthening or shortening of the signal is

or the order of the inverse of the signal bandwidth. In other words, the shift (distortion) can be ignored if

$$|\Delta D| << \frac{1}{B_g}$$

where ΔD is the change in signal extent and B_g is the signal bandwidth. In the present instance, $\Delta D = (1 - \alpha)D$. We will thus define the distortion measure as

$$d \triangleq B_g D|1 - \alpha| \quad . \tag{5-27}$$

Continuing the analogy between the effects of doppler distortion and relief distortion, we require that

$$d << 1 \quad . \tag{5-28}$$

In Chapter 6, simulations of the distortion modelled by $h(x) = \alpha x$ show that (5-28) also applies in image cross-correlation applications. The use of (5-27) and (5-28) as a predetection feature depends on the ability to estimate α from the stereo-pair without actually performing the stereo compilation procedure. Although such an estimation procedure will be considered shortly, it is more practicable to obtain a space-bandwidth map of the images (i.e., a matrix containing $B_g D$ values corresponding to each image point), then compute the distortion measure d during the compilation process where α (or more generally $h(x)$) can be obtained from successive parallax values. Values of d exceeding a pre-established threshold indicate regions of probable correlator performance degradation. Improved correlation can be obtained by warping the appropriate image functions according to (5-15) (effectively forcing α to be equal to 1). An algorithm for the on-line

spatial warping of image patches has been described by Panton [1978]. Such an on-line procedure either requires interpolation on previously stored data, or rescanning the image with updated sampling rates based on the expected terrain slope at the next match location. Increased throughput rate could thus be achieved by performing the warping only when required as specified by the distortion measure d.

Webber and Delashmit [1974b] have considered the filtering of imagery to desensitize the correlation process to linear scale-factor distortion described by $h(x) = \alpha x$. They conclude that for low SNR and Gaussian autocorrelation functions given by (4-86) that Δ should be adjusted by filtering so that $\Delta = D|\alpha|/\sqrt{6}$. This desensitization process, however, applies only to local registration accuracy. Furthermore, as we have previously stated, low SNR imagery typically results in broad autocorrelation functions which are already insensitive to distortion.

As an example of the effect distortion can have on the expected correlation function, consider the autocovariance function used in previous examples:

$$C_f(\xi) = \exp\left[-\frac{1}{2}\left(\frac{\xi}{\Delta}\right)^2\right] \quad . \tag{5-29}$$

Using (3-14), the bandwidth associated with this autocovariance function is given by

$$B_g = \frac{1}{2\pi\Delta} \quad . \tag{5-30}$$

Evaluating (5-20) for $h(x) = \alpha x$ leads to

$$E(r(\xi)) = \frac{\sqrt{2\pi}\Delta}{|\alpha-1|} \int_{-\frac{D}{2\Delta}|\alpha-1|-\frac{\xi}{\Delta}}^{\frac{D}{2\Delta}|\alpha-1|-\frac{\xi}{\Delta}} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv \quad . \tag{5-31}$$

Expressing this result in terms of the distortion measure d, we have

$$\frac{E(r(\xi))}{D} = \frac{1}{\sqrt{2\pi}d} \int_{-\pi d-\frac{\xi}{\Delta}}^{d-\frac{\xi}{\Delta}} \frac{1}{\sqrt{2\pi}} e^{-\frac{v}{2}^2} dv \tag{5-32}$$

Figure 5-1 displays some representative correlation curves for several values of the parameter d. The effects of prime importance are the reduction in peak height which effects the false acquisition rate and the alteration in the curvature of the function at the origin which affects the local registration accuracy (see 3-19). The reduction in normalized peak magnitude and the reduction in normalized curvature magnitude are plotted in Figure 5-2 as a function of d. These effects have been observed experimentally by Casasent and Psaltis [1976].

The Effect of Distortion on Error Prediction

In Chapter 3, we saw that the local registration accuracy depends on the second moment of the image power spectrum

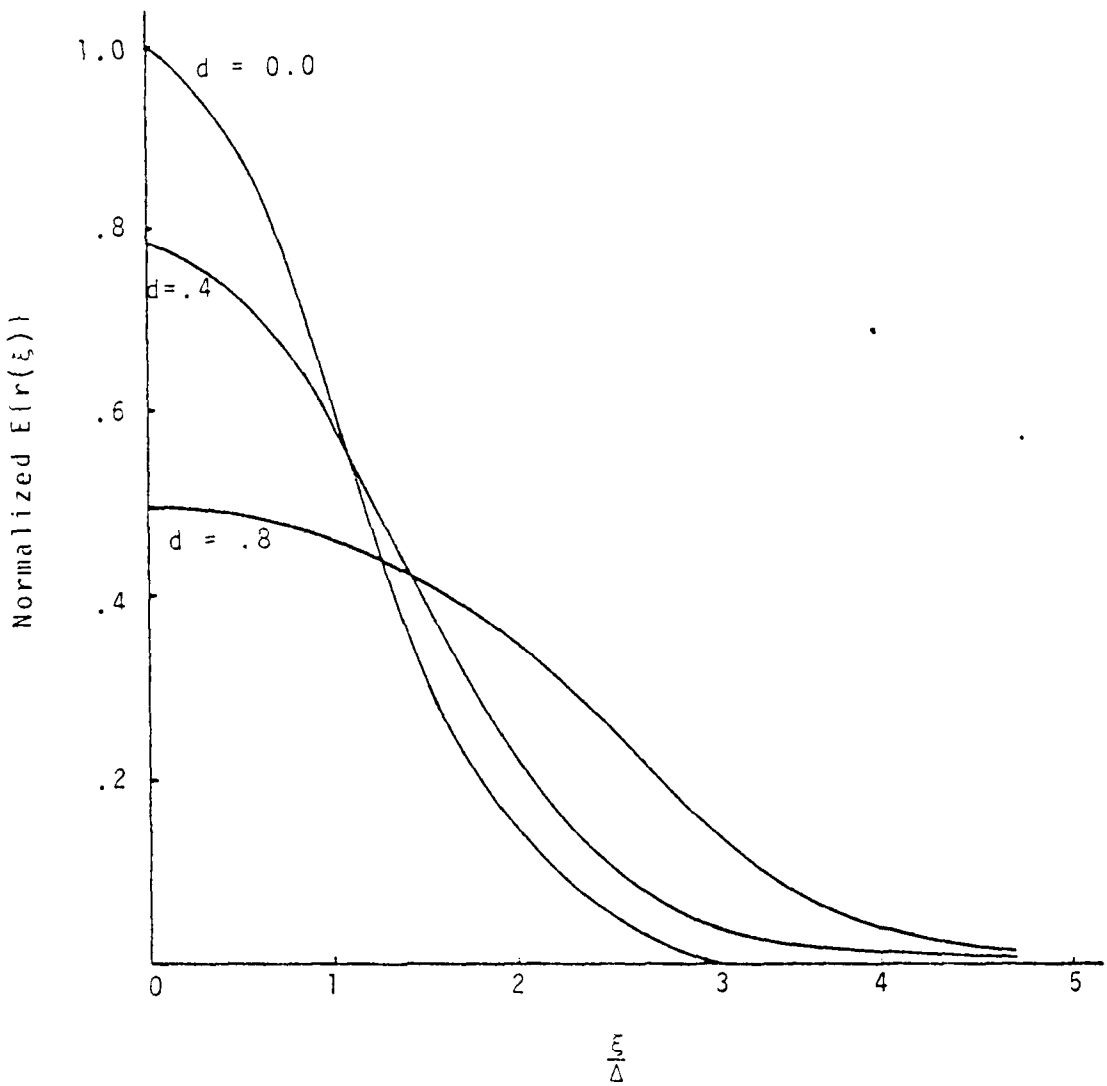


Figure 5-1. The Effect of Distortion on Cross-Covariance Functions

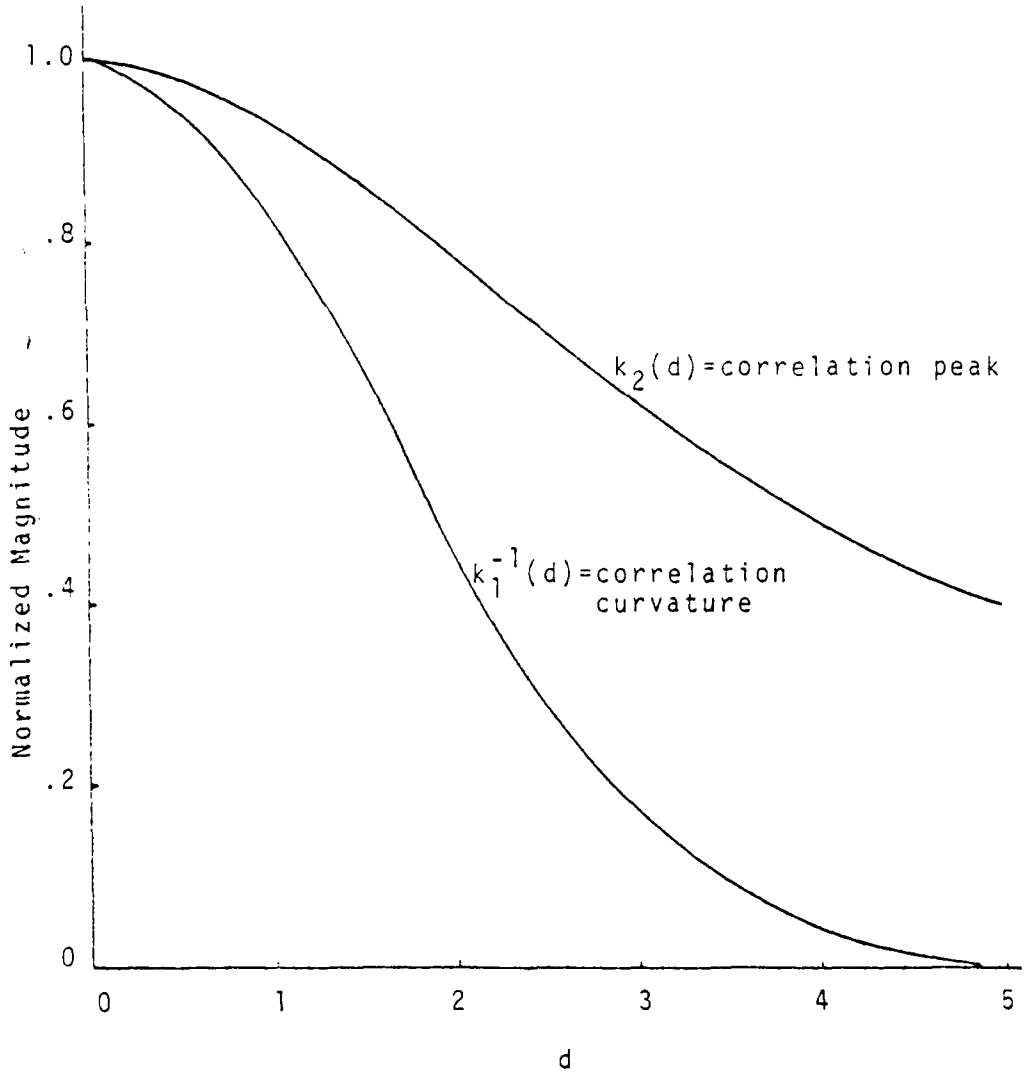


Figure 5-2. Normalized Correlation Peak Height and Curvature vs. Distortion Measure for Gaussian Auto-Correlation Functions

in the manner described by (3-11). From (3-19) and (3-11) we see that

$$E(\xi - \hat{\xi})^2 \geq \frac{N_0/2}{D\frac{1}{2\pi}\left|\frac{d^2C_g(\xi)}{d\xi^2}\right|_{\xi = 0}} \quad . \tag{5-33}$$

Since the curvature of the cross-covariance function decreases with increased distortion, a correction factor, k_1, can be included in (5-33) to obtain

$$E(\xi - \hat{\xi})^2 \geq \frac{N_0/2\ k_1}{D\frac{1}{2\pi}\left|\frac{d^2C_g(\xi)}{d\xi^2}\right|_{\xi = 0}} \tag{5-34}$$

where k_1 depends on the form of the autocorrelation function and the distortion measure d. For the Gaussian autocorrelation function, k_1 is obtained from Figure 5-2 and

$$k_1 = e^{\frac{1}{2}(\pi d)^2} \quad . \tag{5-35}$$

Equation (5-35) describes the approximate increase in the Cramer-Rao lower bound due to the presence of distortion. The increase in local registration MSE is shown in Figure 5-3.

The false acquisition probability as a function of the quality measure Q is seen in Figure 4-4. The reduction in correlation peak value effectively reduces the signal-to-noise ratio at the correlator output. For Gaussian autocorrelation functions we can obtain a correction factor k_2 so that the adjusted quality measure Q' is given by

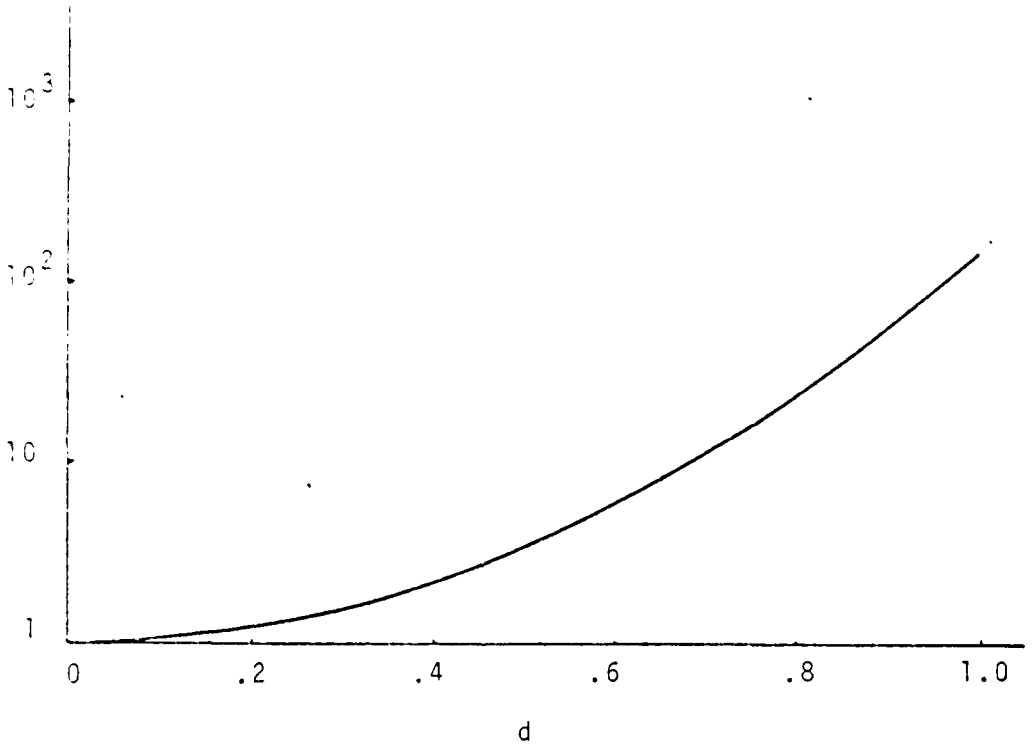


Figure 5-3. Increase in Local MSE Due to Distortion

$$Q' = Qk_2 \qquad (5\text{-}36)$$

where k_2 is shown in Figure 5-2 and given by

$$k_2 = \frac{2}{\sqrt{2\pi}d} \int_0^{-d} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad . \qquad (5\text{-}37)$$

The resulting increase in false acquisition probability depends on the magnitude of Q as seen in Figure 5-4.

The Effect of Match Window Size on Correlation Accuracy

In the absence of relief distortion, the correlation accuracy is expected to improve with increasing window size. This behavior is predicted by the Cramer-Rao bound as well as P_{FA} since Q increases with match window extent. In the presence of terrain relief, however, the distortion measure varies directly with match window size. The result is that for $d > 0$, there is an optimum window size, N_{opt}, which will minimize the correlation error in some sense. The optimum value of N, however, depends on the criterion to be minimized and there is no general procedure for selecting the "best" criterion. Obvious candidate criteria include (a) the minimization of P_{FA} or (b) the minimization of MSE assuming correct acquisition. These criteria have been examined by Mostafavi and Smith (1978a, 1978b) assuming a Gaussian autocorrelation function. They show that the optimum value of

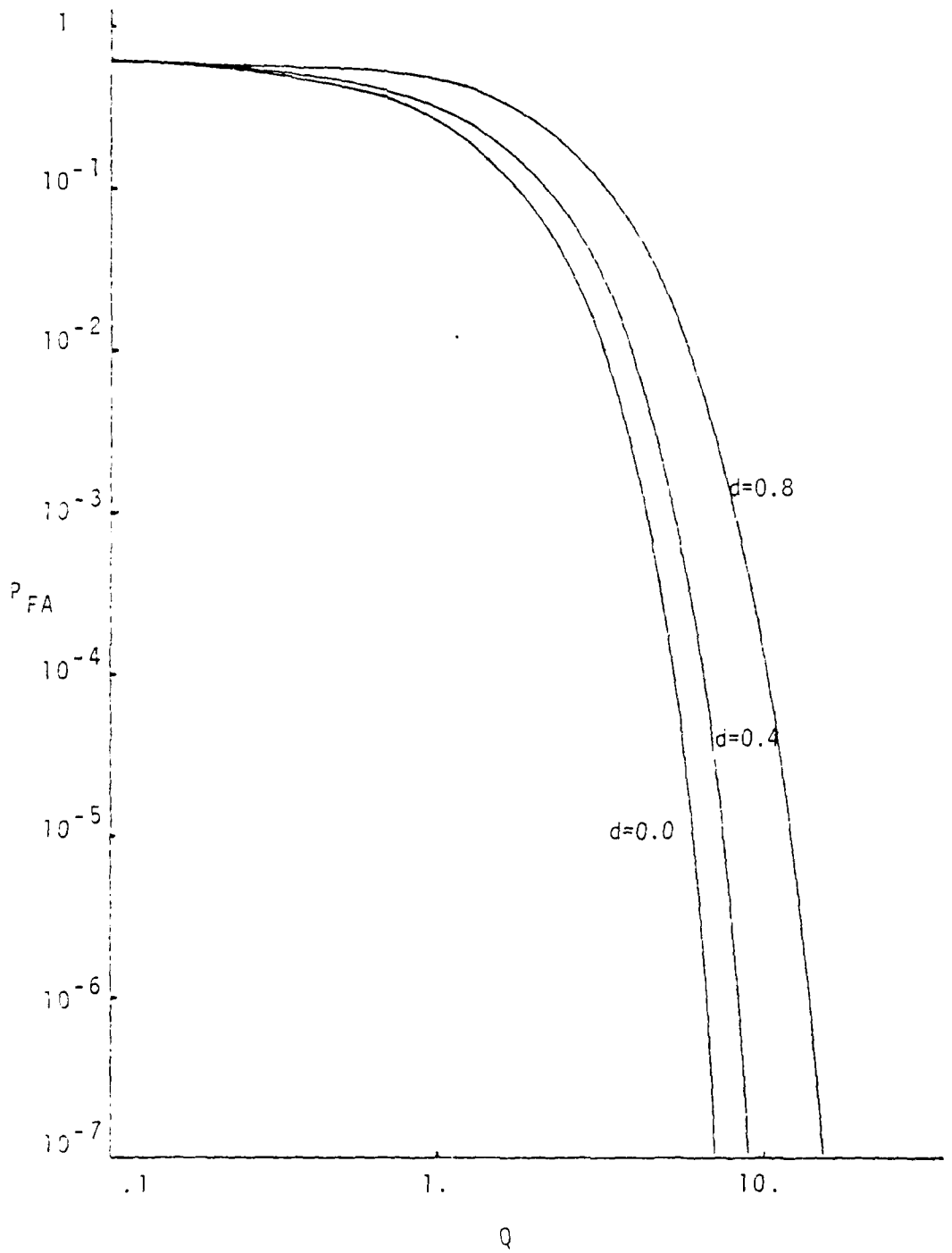


Figure 5-4. Effect of Distortion on Single Tap False Acquisition Probability

N is a function of the correlation width L and the distortion d and that the optimum value of N under criterion (a) is slightly larger (~30%) than under criterion (b). In either case N_{opt} decreases with increased distortion as one would expect. Mostafavi [1979] has also considered the optimization of window functions for image correlation. These results were derived under the assumption that the affine transformation describes the distortion globally, i.e., the geometric distortion is independent of window size and shape. In stereo-compilation systems, however, the relief distortion is location dependent and as the window size increases the distortion becomes more complex; for large match windows, the distortion function $h(x) = \alpha x$ may not be valid. Most present stereo-compilers operate with a fixed match window size and perform spatial warping as described earlier to adjust for local terrain variations. As stated previously, it would be useful to be able to obtain distortion information prior to the compilation process. In the following section, an image overlay quality measure is developed by considering the effect of relief distortion on the arithmetic difference between two registered images.

Estimation of Relief Distortion

We will assume that within sufficiently small image regions, that the relief distortion is modelled by (5-12). Repeating the earlier result,

$$\alpha \cong 1 - \frac{mB}{H_0} \tag{5-38}$$

where $H_0 = H - e(X)$. The quantities m and H_0 can be viewed as random variables which depend on the match window location. Since $H >> e(X)$, we will let

$$H_0 = H - E(e(X)) \tag{5-39}$$

and ignore the effect of variations in elevation on the value of α. From (5-38), it follows that

$$E(\alpha - 1)^2 \cong (\frac{B}{H_0})^2 E(m^2) \quad . \tag{5-40}$$

But $m(X) = \frac{de(X)}{d\,X}$ so (5-40) becomes

$$E(\alpha - 1)^2 \cong (\frac{B}{H_0})^2 \left(- \frac{R_e^2(\tau)}{d\tau^2}\bigg|_{\tau = 0}\right) \tag{5-41}$$

where R_e is the elevation profile autocorrelation function.

As we have seen, the difference between the stereo-pair images is the result of topographic relief and noise sources. Letting $g_1(x_1)$ and $g_2(x_2)$ represent the underlying image intensity profiles, we have seen that if x_1 and x_2 are conjugates then, in general,

$$g_1(x_1) = g_2(x_2) \tag{5-42a}$$

but

$$g_1(x_1 + \Delta) \neq g_2(x_2 + \Delta) \quad . \tag{5-42b}$$

Now suppose we select a region of imagery, S_1 (considerably larger than match window size), and register S_1 with the other member of the stereo-pair. The value of parallax at registration, Δ, provides us with the mean elevation,

$\bar{e} = E(e(X))$, over the selected region and defines a region S_2 which is most similar to S_1. If the elevation in the region is constant, then assuming the registration is correct, the difference between elements of S_1 and S_2 is given by

$$S_2(x_2) - S_1(x_1) = g_2(x_2) + n_2(x_2) - g_1(x_1) - n_1(x_1)$$
$$= n_2(x_2) - n_1(x_1) \qquad (5\text{-}43)$$

Now define

$$\sigma_d^2 = E(S_2(x_2) - S_1(x_1))^2 \qquad (5\text{-}44)$$

where σ_d^2 is the difference variance. From (5-43),

$$\sigma_d^2 = E(n_2(x_2) - n_1(x_1))^2 = 2\sigma_n^2 \quad . \qquad (5\text{-}45)$$

In the presence of topographic relief, however, the overlay is less accurate and $\sigma_d^2 > 2\sigma_n^2$. For the mean elevation $\bar{e}$, the parallax is given by

$$\delta = \frac{B\delta}{H - \bar{e}} \qquad (5\text{-}46)$$

and the coordinate transformation that results in correct registration is given by

$$x_2 = x_1 - \frac{B\delta}{H - \bar{e}} = x_1 - \delta \quad . \qquad (5\text{-}47)$$

If we now define

$$n(x_1) = S_2(x_2)\Big|_{x_2 = x_1 - \delta} - S_1(x_1) \qquad (5\text{-}48)$$

then $\sigma_d^2 = E(n^2(x_1))$. An expression for σ_d^2 that includes the relief distortion can be derived by considering the geometry of Figure 5-5. If at point X in the object space, $e(X) = \bar{e}$, then $g_1(x_1) = g_2(x_2)$. If $e(X) \neq \bar{e}$, then as a first order

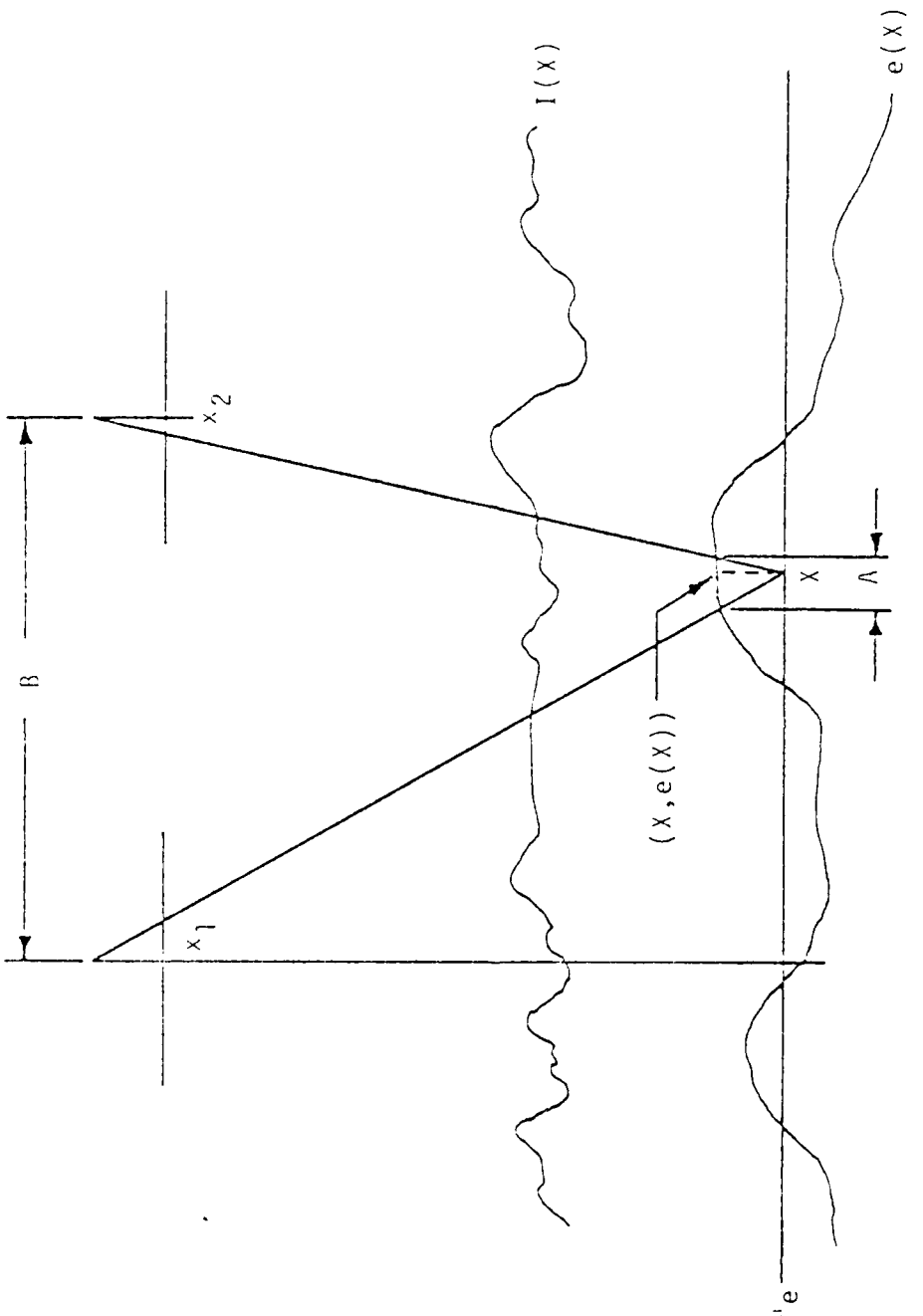


Figure 5-5. Relief Distortion Geometry

approximation

$$g_2(x_2) = g_1(x_1) + AI'(X) \qquad (5\text{-}49)$$

where A is the resulting difference in ground position of the pre-images of $g_1(x_1)$ and $g_2(x_2)$ due to the change in e(X). If we assume that e(X) is a constant over the interval A, then by similar triangles,

$$A = \frac{B}{H_0}\,[e(X) - \mu_e] \quad . \qquad (5\text{-}50)$$

Substituting (5-50) into (5-48) and (5-49) results in

$$n(x_1) = \frac{B}{H_0} I'(X)[e(X) - \mu_e] + n_2(x_2) - n_1(x_1) \quad . \qquad (5\text{-}51)$$

Thus

$$\sigma_d^2 = \left(\frac{B}{H_0}\right)^2 E(I'(X))^2 E(e(X)) - \mu_e)^2 + 2\sigma_n^2 \qquad (5\text{-}52)$$

where the assumption is made that the terrain elevation and terrain reflectance processes are independent. Taking the expectations in (5-52), we have

$$\sigma_d^2 = \left(\frac{B}{H_0}\right)^2 \left(- \left.\frac{d^2R_I(\tau)}{d\tau^2}\right|_{\tau = 0}\right) \sigma_e^2 + 2\sigma_n^2 \qquad (5\text{-}53)$$

where σ_e^2 is the variance of the elevation data within the region enclosed by S_1. Since the photograph intensity pattern is the image of the ground pattern, it follows that

$$R_I''(\tau) = \frac{d^2R_I(\tau)}{d\tau^2} = K\frac{d^2R_g(\tau)}{d\tau^2} = KR_g''(\tau) \qquad (5\text{-}54)$$

where K is a constant which accounts for the change in scale and units from object to image space. Equations (5-53) and

(5-54) can be combined to yield

$$\sigma_e^2 = (\sigma_d^2 - 2\sigma_n^2)\left(\frac{H_0}{S}\right)^2 \frac{1}{K\, R_g''(0)} \quad . \qquad (5\text{-}55)$$

The quantity σ_d^2 can be estimated by computing

$$\hat{\sigma}_d^2 = \frac{1}{M} \sum_{S_1, S_2} (S_2(x_2) - S_1(x_1)) \qquad (5\text{-}56)$$

where M is the total number of elements in S_1. If we now assume that the elevation bandwidth, β_e, is known and assumed to be constant over the ensemble of images, then from the moment theorem

$$- \left.\frac{d^2 R_e(\tau)}{d\tau^2}\right|_{\tau = 0} = (2\pi)^2 \sigma_e^2 \beta_e^2 \quad . \qquad (5\text{-}57)$$

Substituting (5-57) and (5-55) into (5-41) results in

$$E(\alpha - 1)^2 \cong (2\pi)^2 \beta_e^2 (\sigma_d^2 - 2\sigma_n^2) \frac{1}{K\, R_g''(0)} \quad . \qquad (5\text{-}58)$$

Equation (5-58) represents an estimate of the fluctuation of the distortion parameter α measured over the terrain region viewed through section S_1.

To summarize, the procedure for estimating $E(\alpha - 1)^2$ is as follows:

(1) Select region S_1

(2) Register S_1 with the other image to define S_2 and μ_e

(3) Compute the sample difference variance, $\hat{\sigma}_d^2$

(4) Using either S_1 or S_2, compute

$$R_g''(0) = -\left.\frac{d^2R_g(\tau)}{d\tau^2}\right|_{\tau = 0} = (2\pi)^2 \int_{-\infty}^{\infty} f^2 G_g(f)df \quad .$$

(5) Compute $E(\alpha - 1)^2$ as in (5-58).

The assumption that ε_e is known is certainly an oversimplification since ε_e is also a random variable that is location dependent. ε_e cannot be determined without performing the stereo-compilation process. In its place, then, the best we can do is to use $\hat{\varepsilon}_e = E(\varepsilon_e)$ in (5-58) where the ensemble mean is estimated over an ensemble of known terrain conditions. Note also that if there are other differences between the images such as a reflection from an object appearing bright in an image (due to aspect angle) and dark in the other image, the computation in (5-58) will overestimate the variance of the distribution of α. In these situations, Equation (5-58) may, however, find utility as an overlay quality measure since any factors that reduce the overlay quality are also expected to degrade the correlation performance. We thus refer to $\sigma_\alpha^2 = E(\alpha - 1)^2$ computed as in (5-58) as the "effective" relief distortion.

The Effect of Distortion on Estimated Mean Square Correlation Error

Assuming that the distribution of α, $p_\alpha(\alpha)$, were known, it is possible to adjust (at least theoretically) the estimate of MSE from equations (4-3) and (4-100) to (4-102)

by realizing that the cross-covariance functions C_{rs} in (4-102) are functions of the parameter α. We can thus write

$$u(x,n,\ell,\alpha) = \frac{\sigma_c(\ell)x + N(C_{rs}(\ell,\alpha) - C_{rs}(n,\alpha))}{\sigma_c(n)} \quad (5\text{-}59)$$

and compute

$$\hat{MSE} = \int_{-\infty}^{\infty} MSE(\alpha)p_{\alpha}(\alpha)d\alpha \quad (5\text{-}60)$$

where $MSE(\alpha)$ is computed from (4-3) with

$$P_{\ell}(\alpha) = \int_{-\infty}^{\infty} \sum_{n=-L}^{L} \int_{-\infty}^{u(x,n,\ell,\alpha)} Gaus(y)dy\ Gaus(x)dx \quad . \quad (5\text{-}61)$$

In practice, however, it is extemely unlikely that image data will meet all the requirements leading to (5-61) (namely Gaussian, stationary data with uncorrelated correlator top outputs). However, (5-59) provides a means of adjusting the quality measure Q should the distortion be deemed a potential source of difficulty. This problem is discussed further in the following chapter after observing the behavior of correlator performance in a simulation of relief distortion.

CHAPTER 6

CORRELATION ERROR PREDICTION - EXPERIMENTAL RESULTS

In the previous chapters, an attempt has been made to develop the theoretical framework on which to base practical image pre-processing schemes. In this chapter, we consider the problem of implementing image processing algorithms within such a framework. Although the resulting complexity and throughput rate of the pre-processor are of some concern, we will postpone discussion of this topic until the behavior of a variety of processing schemes has been observed. To simplify the processing, however, we will restrict our attention to procedures which create "quality maps" based on only one image of the stereo-pair. This eliminates any scheme that requires the computation of cross spectra or, equivalently, cross correlation as, for example, in (3-29). Instead, we will assume the model of (3-7) with $\alpha = 1$. The quality measures developed in previous sections fall into three categories: (a) those based on local registration accuracy, (b) those based on false acquisition probability, and (c) those based on the detection of input data non-stationarity. In category (a) we will look specifically at the Cramer-Rao bound expression (Equation (3-8)), since it is independent of the particular correlation algorithm. In category (b), we will be concerned with the quality measure

Q in its various forms which depend on the parameter α in (4-97). Since the measure Q_1 in (4-99) depends on the signal variance which is a measure of local image contrast or texture, we will also consider some nonparametric local statistics such as contrast modulation and median absolute deviation. These quantities will be defined and discussed in detail in a later section. In category (c) we will confine the discussion to the variance gradient as discussed in Chapter 4.

The objective of the experiments in the following sections is to observe the behavior of the quality measures and determine the degree to which the quality measures can be used to detect error-prone regions of imagery. Before proceeding, however, it is necessary to formulate a definition of "error-prone" and to establish appropriate detection criteria. To this end, consider the stereo-images $\underset{\sim}{A}$ and $\underset{\sim}{B}$ and the two ensembles of "noise-images" $\{\underset{\sim}{\eta}\}$ and $\{\underset{\sim}{\sigma}\}$. We assume the noise-images are statistically independent. We form a noisy stereo-pair $\underset{\sim}{A} + \underset{\sim}{\eta}_1$ and $\underset{\sim}{B} + \underset{\sim}{\sigma}_1$, perform the stereo-compilation process and generate an error map, $\underset{\sim}{e}_1$, as in Figures 2-9 and 2-11. If this procedure is then repeated for $\underset{\sim}{\eta}_i$ and $\underset{\sim}{\sigma}_i$, $i = 1,\ldots,N$ we can generate an average squared error map,

$$\underset{\sim}{\bar{e}^2} = \frac{1}{N} \sum_{i=1}^{N} \underset{\sim}{e}_i^2 \quad . \tag{6-1}$$

A point a_{ij} in $\underset{\sim}{A}$ can be defined as "error-prone" if the associated error point, e_{ij}, is such that $e_{ij} > t_c$ where t_c is some error threshold. This definition is appropriate for the quality measures which fall into categories (a) and (b) above. For category (c), however, we note from Figure 2-9 that the covariance processor produces significant errors in the absence of additive noise, presumably due to the presence of edges as discussed in Chapter 4. For this particular case, an error-prone point will simply be one that corresponds to an error value that is greater than t_c.

There are a variety of methods for ascertaining the degree to which a particular quality measure "follows" the error behavior. Once an error threshold t_e has been established, the detection criterion reduces to a simple binary hypothesis test with hypotheses H_0: the point is not error prone and H_1: the point is error prone. For a particular measure Q, we can thus select a quality threshold t_Q and compute the usual detection statistics. Although the decision regions will depend on the particular Q chosen we will assume for purposes of illustration that the larger Q is, the less susceptible the image point should be to correlation error (this is reversed for the Cramer-Rao bound). With this convention we can compute the fraction of hits (H: $e_{ij} > t_c$ and $Q < t_Q$), misses (M: $e_{ij} > t_e$ and $Q > t_Q$), false alarms (FA: $e_{ij} < t_e$ and $Q < t_Q$), and correct rejections (CR: $e_{ij} < t_e$ and $Q > t_Q$). Under this convention,

H + M = 1 and FA + CR = 1. Generation of this data allows the construction of ROC (receiver operating characteristic) curves which are plots of the detection frequency (P_D), versus the false alarm frequency (P_{FA}).

It is also instructive to compute the histograms of the quality measures conditioned on each of the two hypotheses. Since it is desirable that these distributions be separated, a measure of the degree of separation is given by

$$\Delta_Q = \frac{E\{Q|H_0\} - E\{Q|H_1\}}{\sigma_{Q|H_0} + \sigma_{Q|H_1}} \quad . \tag{6-2}$$

where $\sigma^2_{Q,H_i} = E(Q^2|H_i) - E^2(Q|H_i)$. The larger Δ_Q is, the less sensitive the detector will be to changes in the threshold levels t_Q in the vicinity of the optimal (say in the Bayes sense) threshold. This is desirable since the choice of threshold level in a real operating pre-processor must be based on simulation results (and experience) and for a particular image, is unlikely to be optimal.

We shall proceed by first considering the covariance processor errors produced by edges. We will then restrict our attention to the normalized covariance processor when we consider the quality measures in categories (a) and (b). Although the false acquisition quality measures were derived for the covariance processor we recall that normalized covariance simply forces stationarity on the correlation process. There is no reason to believe that the error behavior of covariance and normalized covariance are significantly

different under conditions of stationarity. In fact, in Chapter 2, it was shown that they are equivalent processors under conditions of stationarity.

Finally, we will present some simulations designed to assess the correlation behavior in the presence of distortion modelled by $h(x) = ax$ and to determine the utility of the distortion measure d discussed in Chapter 5.

Input-Data Non-Stationarity

The synthetic image shown in Figure 2-4b was passed through the processor shown in Figure 4-8. The variance estimator was implemented by computing the sample variance within a window having the same dimensions as the match window (7 pixels/line x 3 lines, etc.). Since the similarity search is along epipolar lines, the gradient operator computes the horizontal gradient only using the operator shown in Figure 6-1. The number of elements in the gradient operator is chosen to be the same as the number of pixels/line in the search window since, for each match window location, we wish to detect the presence of any portions of the search window displaying rapidly changing energy. We assume that the match window is centered in the search window, or more precisely, that the correct subimage is centered in the search window. This is a reasonable approximation for a tracking correlator since the search is carried out in a region that is symmetric about the predicted correct

subimage location. For a non-tracking correlator, the search window is usually determined by the expected extremes in elevation. Thus a match window corresponding to an elevation that is near one of the extremes will result in a correct subimage which is located near an extreme of the search window.

The covariance error maps from Chapter 2 and the corresponding variance horizontal gradient maps are shown in Figure 6-2 for match window sizes 7 x 3 and 15 x 3. The distributions of errors for these models is nearly symmetical about zero. For this reason, the ROC curves in Figure 6-3 were formed by taking the absolute value of the data in Figure 6-2 and altering the gradient threshold t_g for a given error threshold $t_e = 1.5$.[1] The distribution separation as a function of t_e is shown in Figure 6-4. The staircase appearance results from the fact that the errors in Figure 6-2a are integer-valued so any threshold between n and n + 1 will result in the same separation.

The regions detected by three different thresholds are shown in Figure 6-5 where the statistics are color coded. The point at which the false alarm rate becomes unacceptable depends on the spatial relationship between the hit (or detection) locations and the false alarm locations; false alarms appearing in close proximity to hits can contribute, in this

[1] The separate ROC curves for positive and negative valued errors are not significantly different from Figure 6-3.

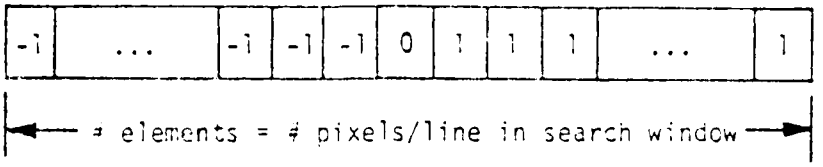


Figure 6-1. Horizontal Gradient Operator

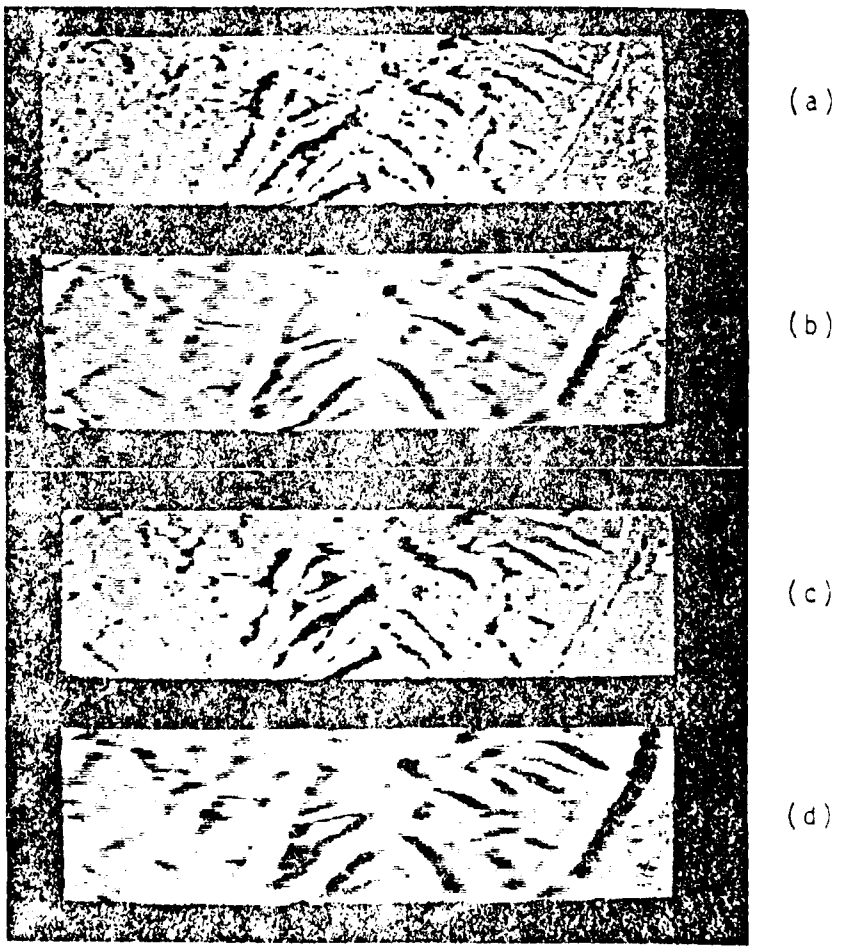


Figure 6-2. Comparison of Covariance Error Maps with Horizontal Gradient of Image Variance (HGV) (a) Error Map - 7 x 3 window, (b) HGV - 7 x 3 window, (c) Error Map - 15 x 3 window, (d) HGV - 15 x 3 window

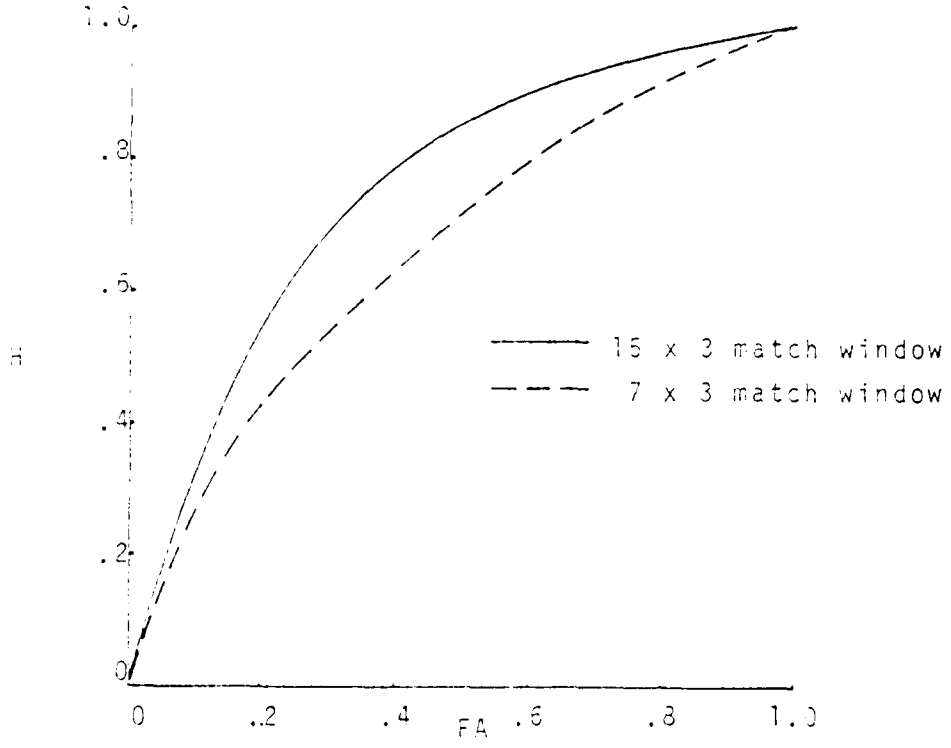


Figure 6-3. ROC Curves for Variance Horizontal Gradient Error Threshold = 1.5

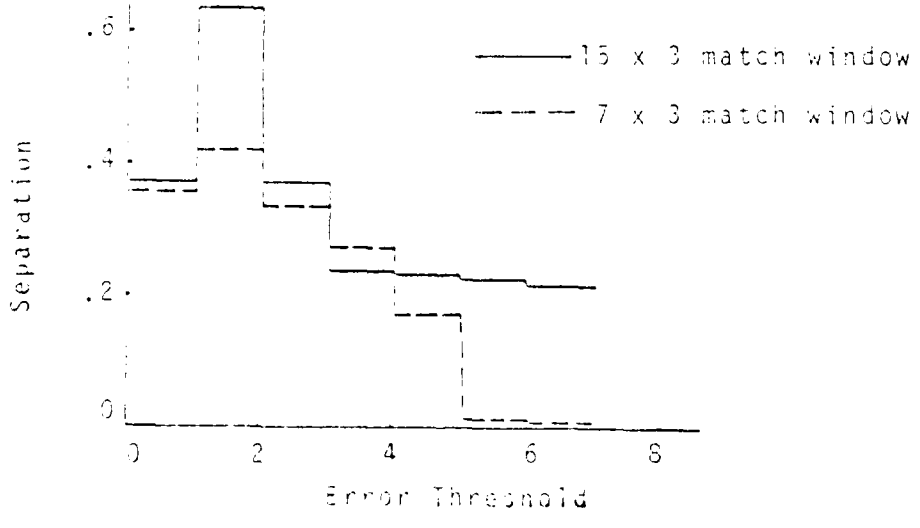


Figure 6-4. Separation Factor for Variance Horizontal Gradient

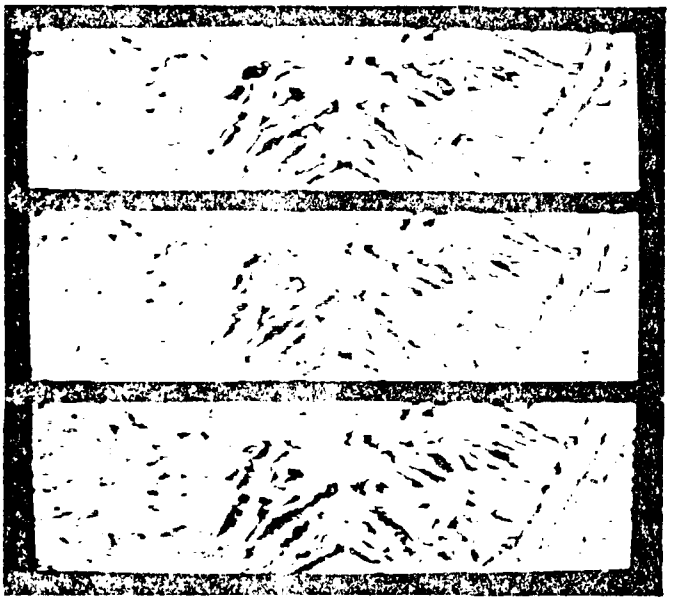


Figure 6-3. Detection of Image Non-Stationarity (Red = HIT, Green = FA, Blue = MISS)

author's view, to acceptable detection. Unfortunately, the point at which false alarms become excessive is somewhat subjective.

Although the overall structure of the variance gradient shows similarities to the covariance error map, it is clear that there are errors produced by other mechanisms which, in effect, degrade the detection statistics. Also, since the simulation was performed using a non-tracking correlator, some of the predicted error locations are not seen on the actual error map (for example, the white and black parallel features in Figure 6-2(b) corresponding to the road at the right-hand side of Figure 2-4(b). This is due to the fact that the corresponding elevation is near the lower extreme and the non-tracking search window does not overlap the feature (road) resulting in the large variance gradient). Furthermore, the presence of an edge does not guarantee that correlation errors will occur; they are only more likely at these locations.

Prediction of Correlation Errors for Stationary Data

The normalized covariance processor imposes stationarity by normalizing the image data sequences. By employing the normalized covariance processor, the presence of edges can be ignored as a source of correlation error and we need be concerned only with local registration and false acquisition errors. Without observing each correlation function,

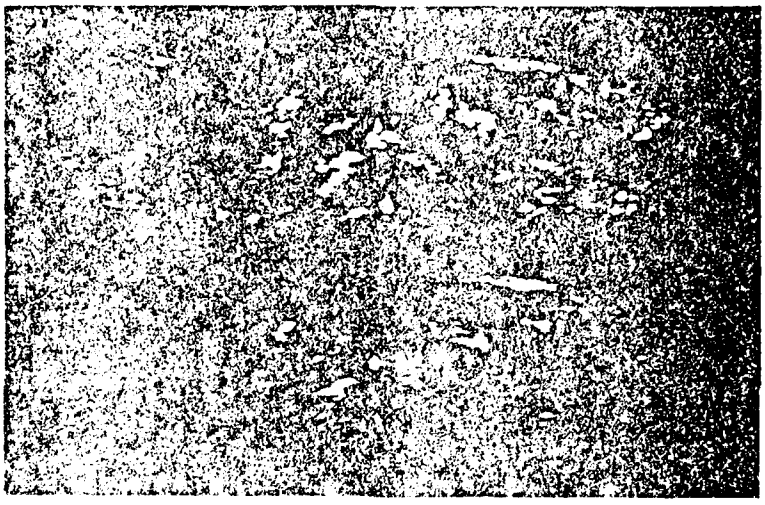
however, it is difficult to classify an error as due to local registration or false acquisition. For this reason, we will be concerned only with the degree to which the image quality measures predict the error behavior.

Figure 6-6 shows the error maps (MSE) generated in the Monte Carlo simulation as described earlier for the match window sizes 7 x 3 and 15 x 3 with an overall input signal-to-noise ratio of 20dB. As in Chapter 2, the image intensity associated with a particular point in Figure 6-6 is linearly related to the mean-square-error associated with that point. In order to observe the sensitivity of the various quality measures to noisy source data, the quality measures were computed over both clean and noise contaminated imagery.

Computation of the Quality Measures

A list of the measures employed and their definitions is given in Table 6-1. Computation of the false acquisition and contrast measures are straightforward. The vector $\underline{r}$ represents the image data, r_{min} and r_{max} are the maximum and minimum pixel values within the test match window, and MED represents the median operator. These quantities are computed without regard to the location of data within the test window.

The estimate of the signal-to-noise ratio in the computation of Q is given in (4-98). When computing this quantity on clean imagery, the numerator, $\sigma_r^2 - \sigma_n^2$ is replaced by σ_g^2.



(a)

(b)

Figure 6-6. MSE Maps-Normalized Covariance, SNR = 20 dB. Match Window Size: (a) 7 x 3, (b) 15 x 3

Table 6-1. Image Quality Features

Cramer-Rao Bound:

$$CRB = \left\{ N \int_{-\infty}^{\infty} (2\pi f)^2 \frac{G_g^2(f)}{2G_g(f)\sigma_n^2 + \sigma_n^4} df \right\}^{-1}$$

False Acquisition:

$$Q = \frac{\sqrt{N}\ SNR}{[N(SNR)^2 + 2SNR + 1]^{\frac{1}{2}}}$$

Contrast Measures:

$$\text{variance} = \sigma_r^2 \triangleq \frac{1}{N}(\underset{\sim}{r} - \bar{\underset{\sim}{r}})'(\underset{\sim}{r} - \bar{\underset{\sim}{r}})$$

$$\text{contrast modulation} = C_m \triangleq (r_{max} - r_{min})/(r_{max} + r_{min})$$

$$\text{contrast difference} = C_d \triangleq r_{max} - r_{min}$$

$$\text{contrast ratio} = C_r \triangleq r_{max}/r_{min}$$

$$\text{median absolute deviation} = MAD \triangleq MED|\underset{\sim}{r} - r_{MED}|$$

The computation of the Cramer-Rao bound expression requires the estimation of the test window power spectrum. Since the correlation search is one-dimensional (epipolar lines), CRB is computed by averaging these quantities over the lines of the test window. Since the number of data samples per line (7 or 15) is small, the data is first multiplied by an approximation of the Kaiser window to smooth the spectral estimate [Harris 1978]. The Kaiser window approximation is given by

$$w(n) = a_0 - a_1 \cos(\frac{2\pi n}{N_p}) + a_2 \cos(\frac{2\pi}{N_p}2n)$$

$$- a_3 \cos(\frac{2\pi}{N_p}3n) \; ; \; n = 0,1,\ldots,N_p-1 \qquad (6\text{-}3)$$

where

$$a_0 = 0.40243$$

$$a_1 = 0.49804$$

$$a_2 = 0.09831$$

$$a_3 = 0.00122$$

and N_p is the number of pixels per test window line. Although other windows could certainly be employed, it was found that computation of CRB without any windowing led to extremely erratic results.

After windowing the test line, an estimate of the power spectrum is obtained by computing ther periodogram [Oppenheim and Schafer 1975]. Since this procedure provides

an estimate of $G_r(f)$ instead of $G_g(f)$, we form the estimate

$$G_g(f) \cong \begin{cases} G_r(f) - \sigma_n^2 & \text{if } G_r(f) > \sigma_n^2 \\ 0 & \text{otherwise} \end{cases} \tag{6-4}$$

If it should happen that $G_g(f) = 0$ for all samples over the periodogram, then CRB is set to an arbitrary large value. When computing the Cramer-Rao bound on clean imagery, $G_g(f)$ can be measured directly and σ_n^2 is based on knowledge of the associated noisy image statistics.

Error Prediction Performance

Figure 6-7 shows ROC curves for the measurements made on clean imagery. The corresponding separation measure is given in Figure 6-8. The ROC curves and separation measure for the noise contaminated case are shown in Figure 6-9 and 6-10 respectively. Comparisons of measurements made on clean vs. noisy data for selected features are shown in Figure 6-11 which displays the percent reduction in detection frequency due to the presence of noise. Although these results are based only on the synthetic images of Figure 2-4, the following observations are noted:

(1) The Cramer-Rao measure provides reasonably good detection rates only for measurements based on the 15 x 3 window over imagery that is free of noise contamination. The poor performance for the 7 x 3 window case

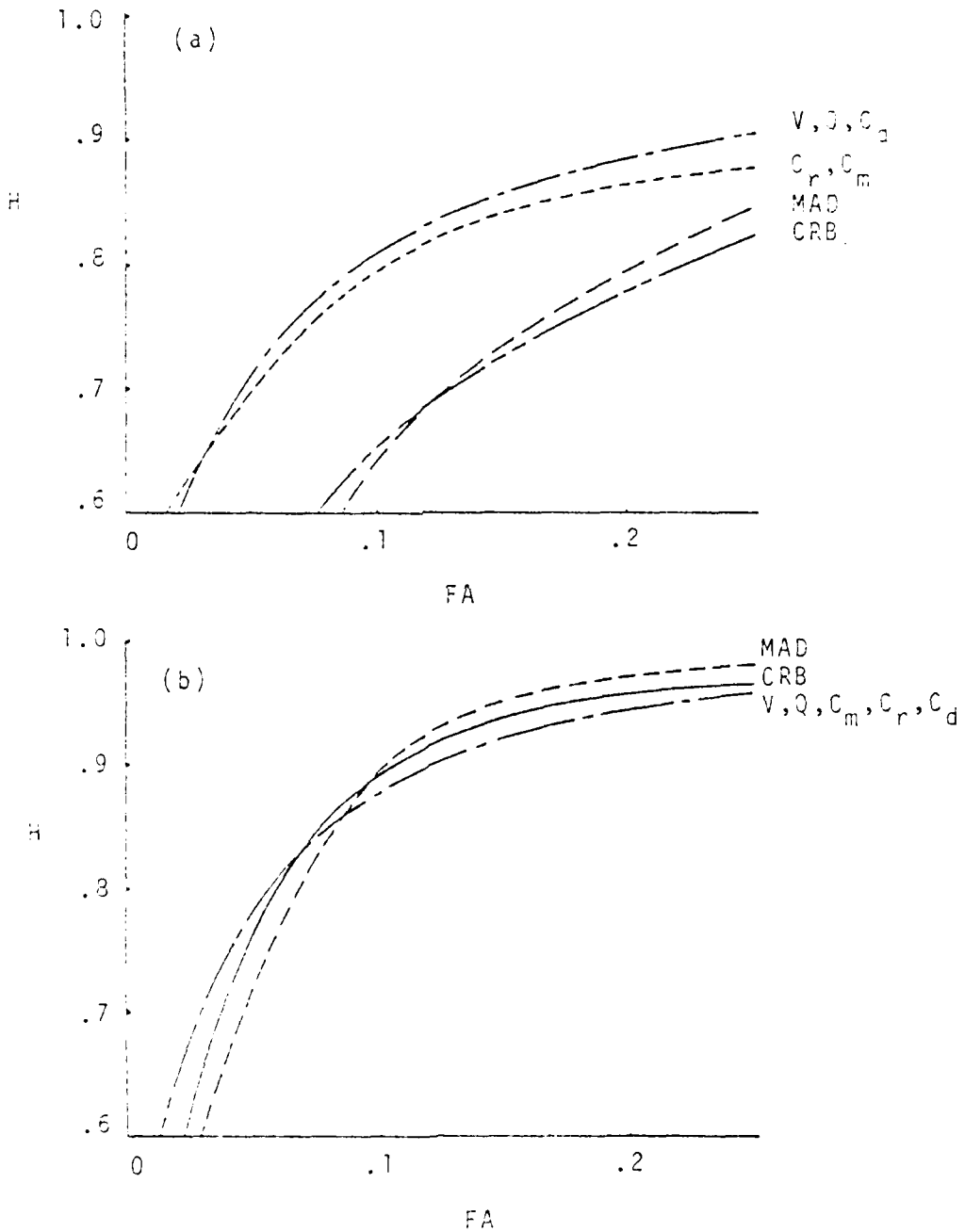
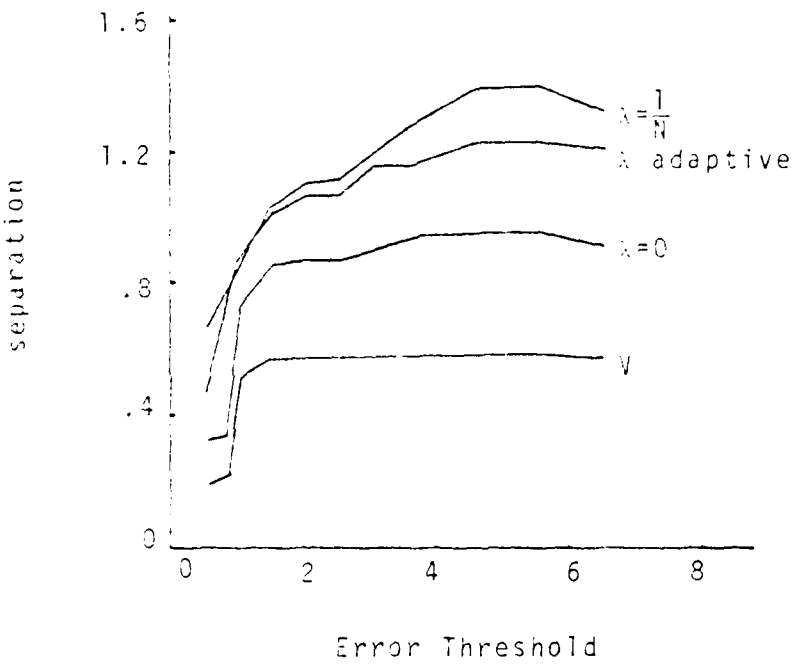


Figure 6-7. ROC Curves for Quality Measures on Clean Imagery (a) 7 x 3 match window (b) 15 x 3 match window



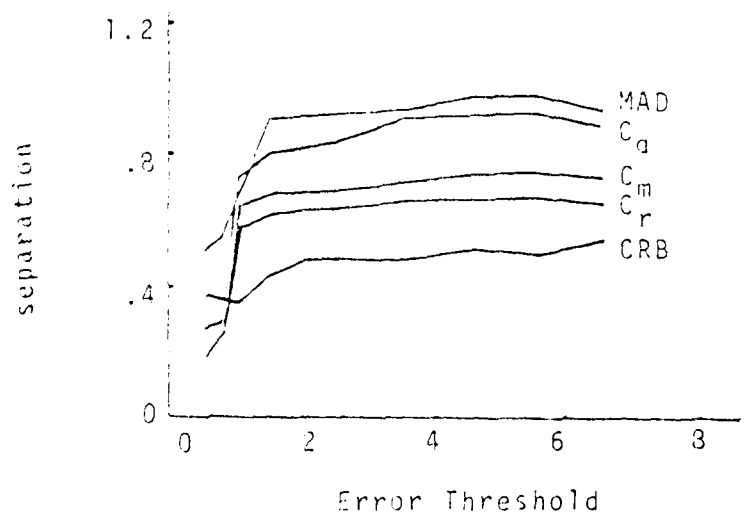
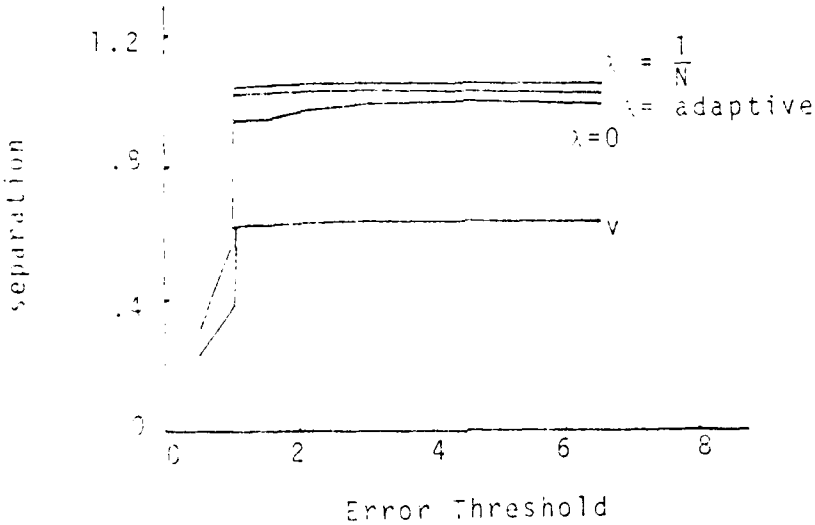


Figure 6-8a. Separation Measure--Clean Images
7 x 3 match window



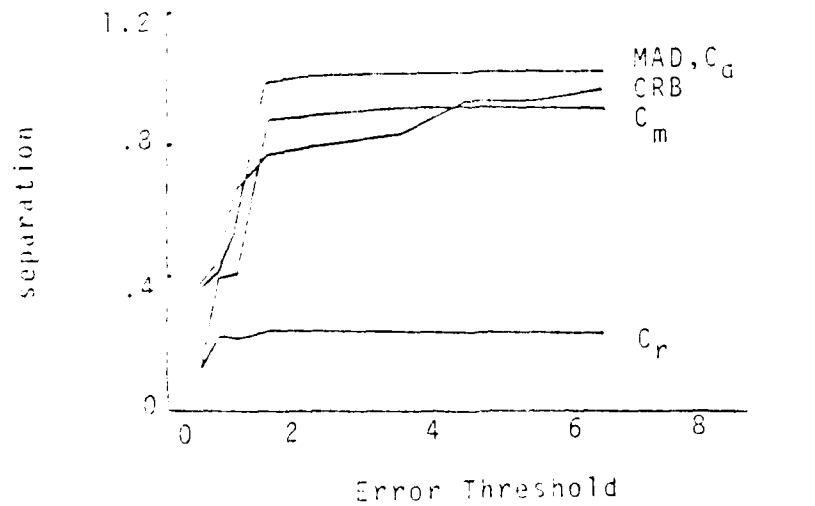


Figure 6-8b. Separation Measure--Clean Images
15 x 3 match window

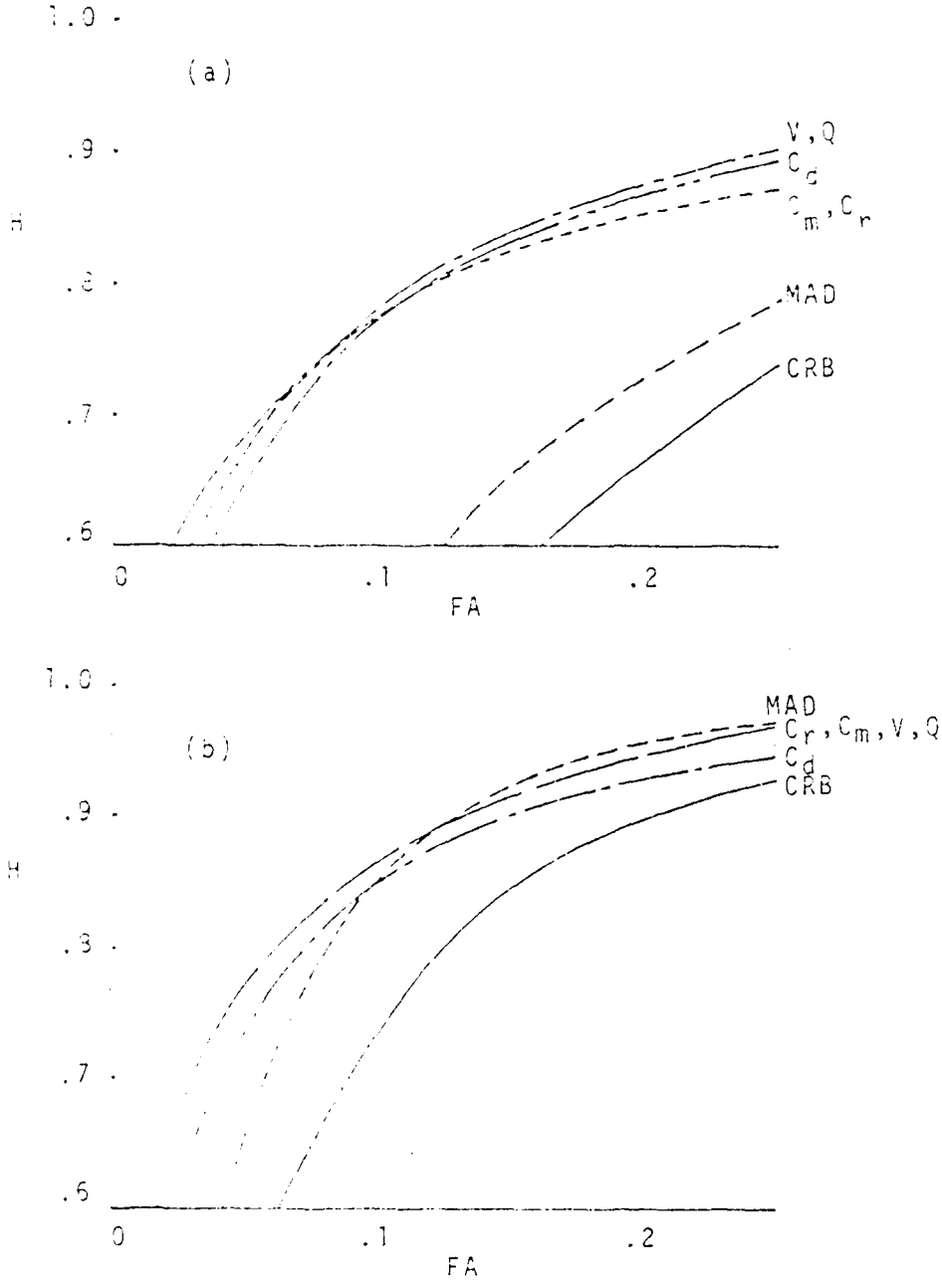
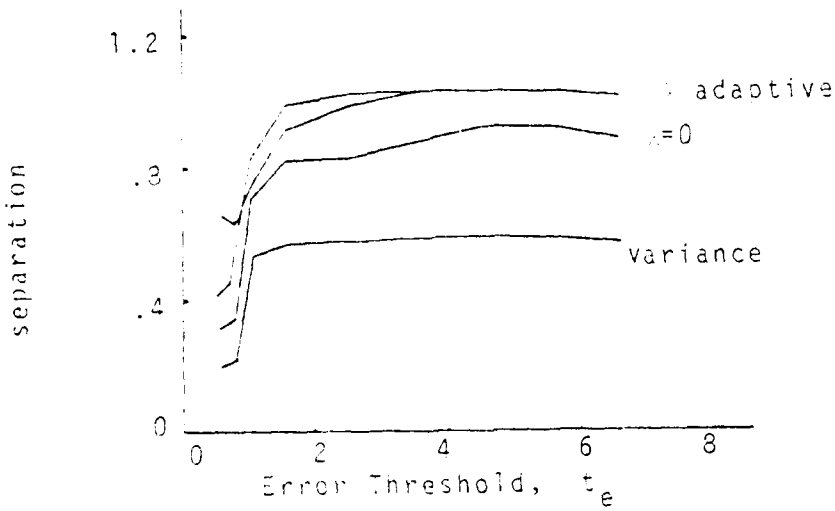


Figure 6-9. ROC Curves for Quality Measures on Noisy Images (SNR = 20 dB)
(a) 7 x 3 match window,
(b) 15 x 3 match window



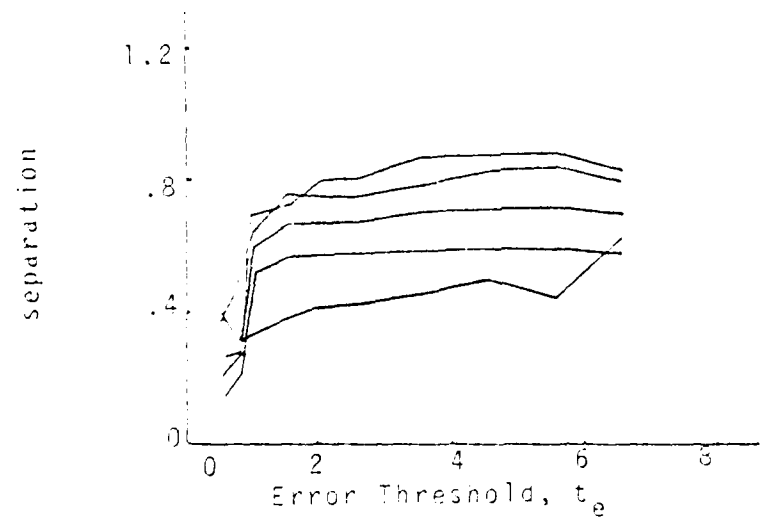
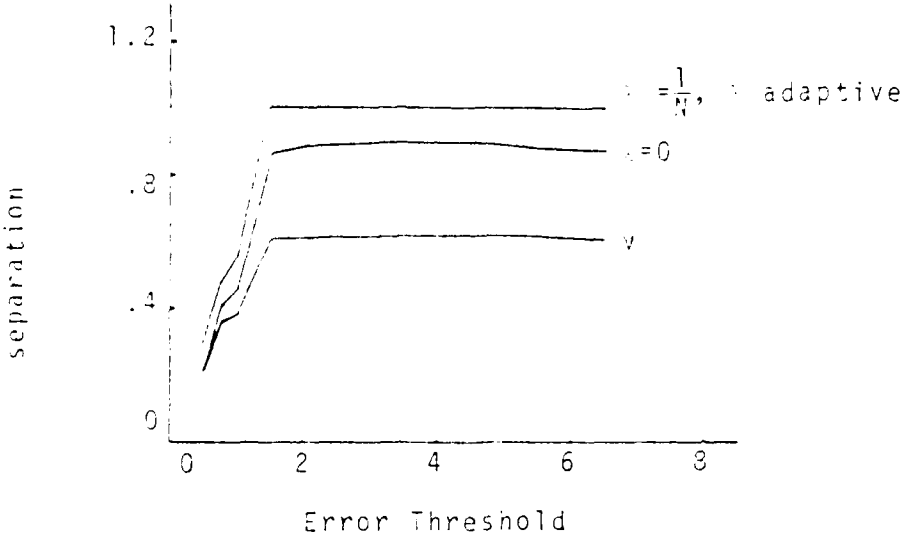


Figure 6-10a. Separation Measure--Noisy Images
7 x 3 match window



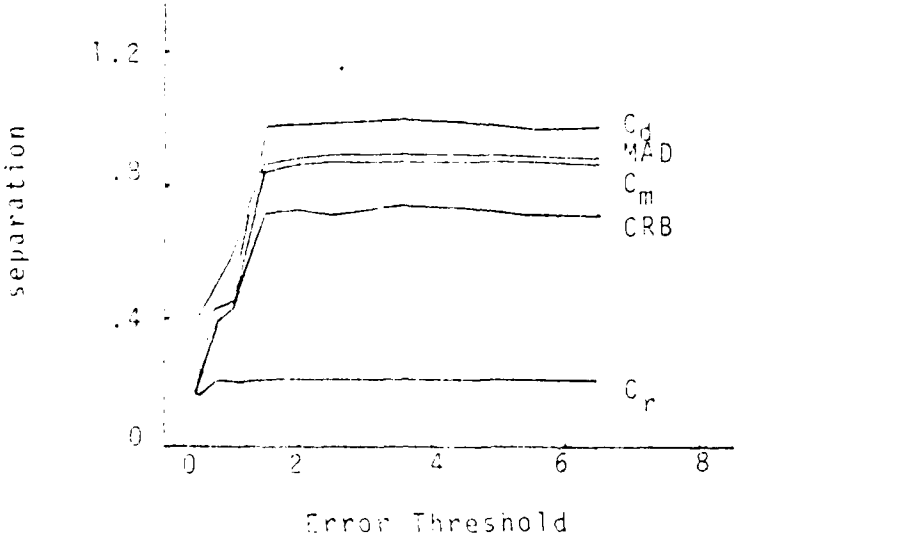
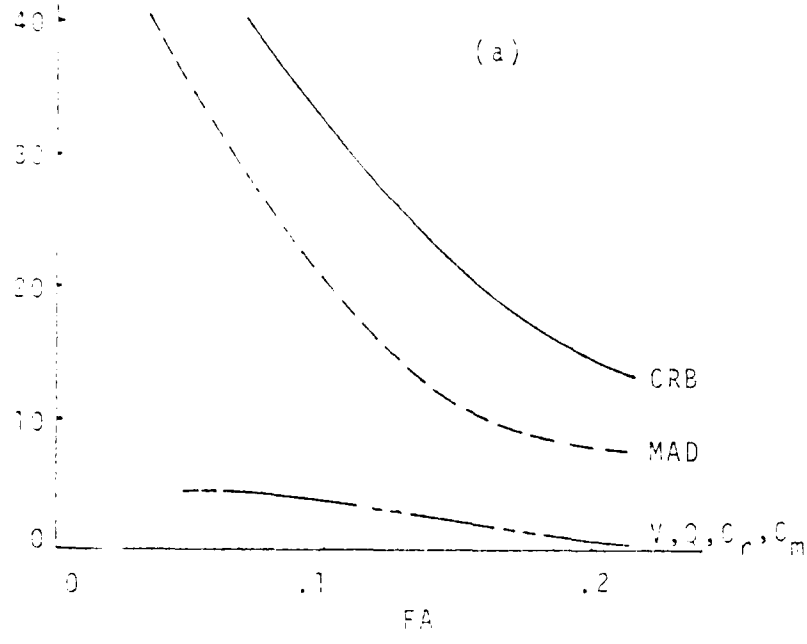


Figure 6-12b. Separation Measure--Noisy Images
15 x 3 match window



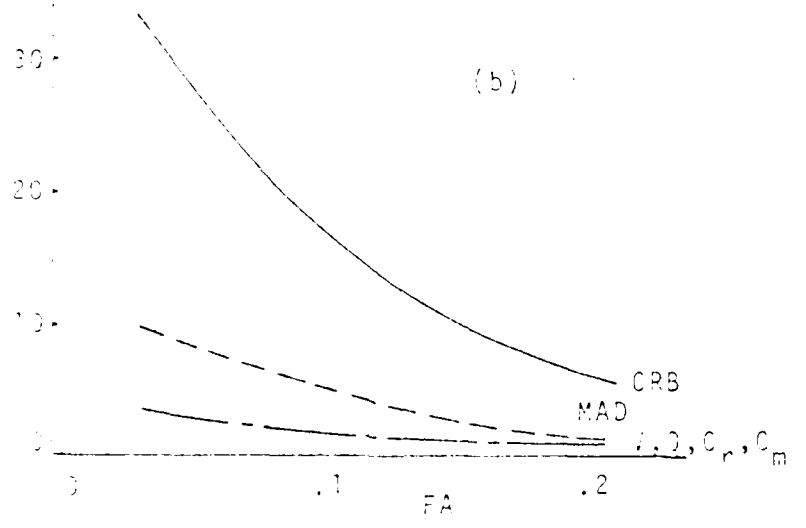


Figure 6-11. Percent Reduction in Detection Frequency Due to Noise Contamination
(a) 7 x 3 match window,
(b) 15 x 3 match window

is attributed to the poor spectral estimates obtainable on such small data records. Since the Cramer-Rao measure weights the image power spectrum by the square of the frequency, it is not surprising that it is sensitive to the presence of white noise.

(2) The ROC curves for the variance and the false acquisition quality measures, Q, are identical. This results from the assumption that the noise power is constant with sub-image location and the one-to-one relationship between SNR and Q specified by Equation (4-97). These quantities differ greatly, however, in separation.

(3) The low separation for low error thresholds seen in Figure 6-8 and 6-10 are due primarily to the one-pixel errors which are independent of the local image statistics when the error threshold exceeds 1.0, these are, in effect, considered as non-errors and the separation measure responds accordingly.

(4) For a given quality measure, an increase in separation implies improved detection (i.e., the hit/false alarm ratio increases). With few exceptions, the separation is essentially independent of error threshold for thresholds greater than 1.5. When comparing two features

however, the fact that one feature has a greater separation than another feature does not necessarily imply improved detection behavior (e.g., MAD with 7 x 3 window).

(5) The variance, Q, and the various contrast measures provide remarkably similarly ROC behavior and all are relatively insensitive to noise.

(6) The ROC curves of Figure 6-10 result from measurements made on a single noise contaminated image. A Monte-Carlo approach similar to the procedure used to define "error-prone points" could be employed but is computationally prohibitive. Further experimentation has shown, however, that the curves of Figure 6-9 are representative although some variation (and inversion of order) occurs between trials. One can conclude only that the behavior of V, Q, C_r, C_m, and C_d are nearly equivalent (in the ROC sense) and that MAD seems to improve with window size more than the other features do.

(7) The separation measure can be better understood by considering Figure 6-12. Figures 6-12a, b, and c are representations of the

quality measure Q for three different values of the parameter α. Comparing these figures to the separation measures in Figure 6-8a, it can be seen that increased separation results in a data distribution that is increasingly bimodal; that is, the feature data is separated into two classes which become more distinct as the separation measure increases. The decrease in separation resulting from measurements on noisy imagery can be observed in Figures 6-12d, e, and f. We also note that the decrease in separation resulting from measurements on noisy images increases with separation. For example, the separation of the variance map changes little, if any, but the separation of Q for $\alpha = \frac{1}{N}$ is considerably reduced.

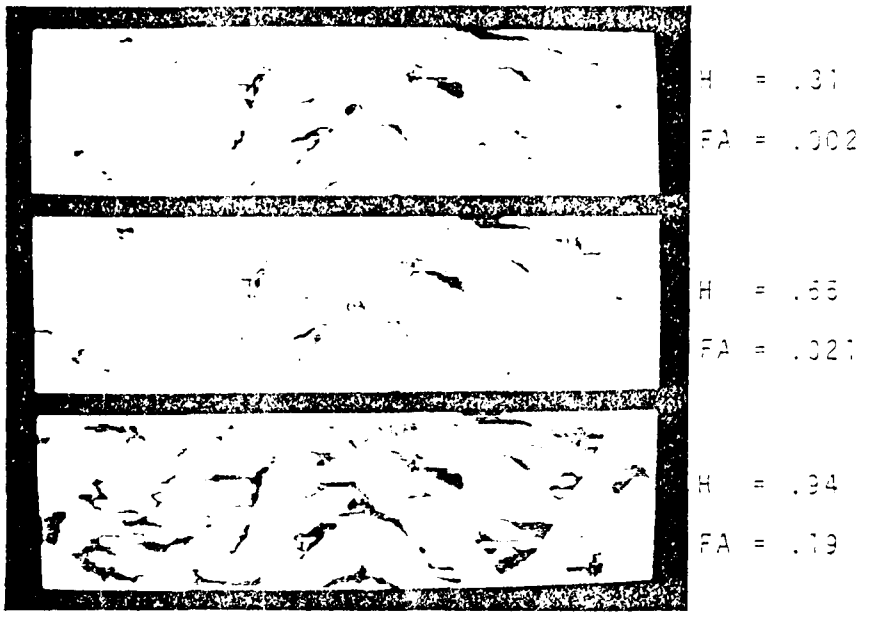
(8) The detection capability for the 15 x 3 window is superior in all respects to that for the 7 x 3 window. Since parameter estimation (e.g., variance, spectrum) generally improves with the quantity of data, it is reasonable to expect further improvement for larger window sizes. The parameters, however, must be representative of the match window at a given location and thus the test window cannot increase in size indefinitely.

Figure 6-12. Image Quality Measures. (a) $\lambda = 0$, no noise; (b) $\lambda = 1/21$, no noise; (c) λ adaptive, no noise; (d)-(f) same as (a)-(c) with measurements made on noisy (SNR = 20 dB) images

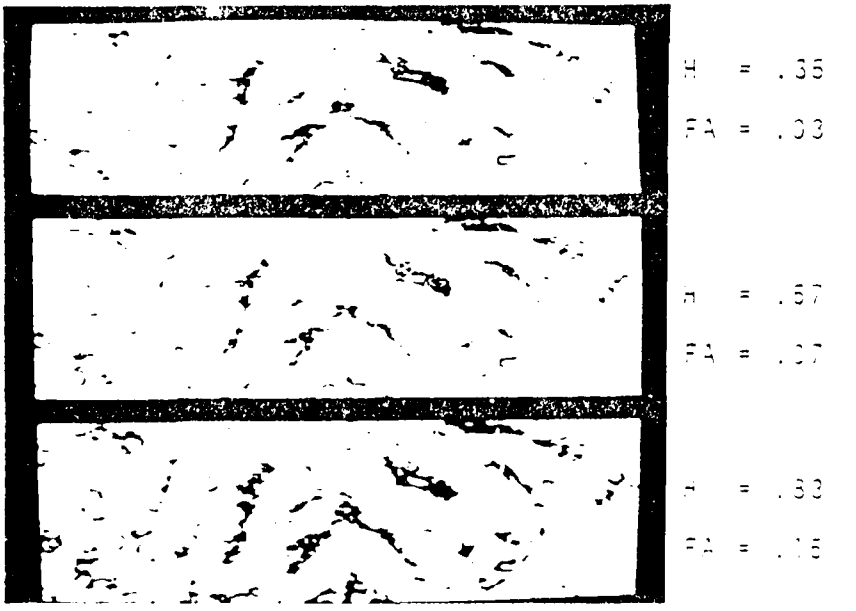
The ROC curves and separation measure do not provide a complete description of the performance of a particular image feature for the simple reason that a given false alarm rate provides no information concerning the spatial relationship between the locations of hits and false alarms; a false alarm that is in close proximity to a hit location is more indicative of the true error behavior than a false alarm in a region where few or no errors occur. Because of this spatial relationship, the point at which the false alarm frequency becomes excessive is difficult to define. As an example, the color-coded detection statistics for the variance feature are shown in Figure 6-13a for several different thresholds and associated hit and false alarm rates. Note that for a false alarm frequency of .021, the false alarm locations (green) cluster near the hit locations (red). However, for FA = .19, numerous false alarms appear at locations distant from hit locations. This behavior is also feature dependent; in Figure 6-13b, we see that isolated false alarms appear at very low hit rates (.35) for the Cramer-Rao feature measured over noisy imagery.

Threshold Sensitivity

Figure 6-11 implies that the variance, Q, and the various contrast measures are relatively insensitive to the introduction of additive noise. These curves are generated by altering the feature decision regions by adjusting the feature threshold level and it is important to determine the



(a) Variance Map (No Noise)



(b) Cramer-Rao Bound Map (SNR=20 dB)

Figure 6-10. Sample Detection Statistics
Red = Hit, Green = FA, Blue = Miss

sensitivity of the threshold setting to the presence of noise. If a given feature provides a 10% false alarm rate with a threshold setting of t_f when measured over clean imagery, it is desirable that the change in threshold required to provide a 10% false alarm rate in the presence of noise be minimal, or equivalently, we would like the change in false alarm rate (or hit rate) for a given threshold to be insignificant. For the features which are based on parametric statistics (variance, Q, CRB) the appropriate threshold depends on the accuracy of the noise power estimate, σ_n^2. Since the estimate of the SNR incorporates the noise power estimate, the false acquisition based measures (Q) adjust accordingly. The variance measure could be likewise adjusted. As an example, Table 6-2 presents some of the raw data used to generate the ROC curves of Figures 6-7b and 6-9b. We note that for $\sigma_n = 0$, Q provides nearly equivalent hit rates (H) for the given feature thresholds (T). This is not the case, however, for the adaptive σ or the variance. The variance, however, can be adjusted by subtracting the noise power (σ^2) from the threshold for the noise case to provide nearly equivalent detection rates.

The features based on non-parametric statistics (MAD, C_m, C_r, C_d) cannot be adjusted so simply since the increase in feature value due to the injection of known noise power is difficult to ascertain and is best determined empirically.

Table 6-2. ROC Curve Raw Data

Q, λ = 0 Noise Estimate: σ_n^2 = 25.

	T	4.0	5.0	6.0	8.0	10.0	12.0
Clean	H	.49	.61	.70	.83	.90	.94
	FA	.007	.016	.029	.067	.124	.200
Noise	H	.48	.62	.70	.83	.90	.94
	FA	.007	.018	.033	.072	.129	.202

Q, λ = adaptive Noise Estimate: σ_n^2 = 25.

	T	1.0	1.25	1.3	1.5	2.0
Clean	H	.72	.83	.85	.90	.97
	FA	.032	.070	.079	.128	.300
Noise	H	.53	.76	.79	.87	.96
	FA	.010	.044	.052	.097	.272

Variance

	T	50	100	125	150	200
Clean	H	.65	.83	.87	.90	.94
	FA	.021	.069	.097	.130	.190
Noise	H		.75	.82	.87	.92
	FA		.042	.068	.096	.16

Relief Distortion Simulations

A series of experiments was performed to observe the behavior of the correlation processors in the presence of geometric distortion described by the distortion function $h(x) = \alpha x$ where α is defined in (5-12). Table 6-3 shows the approximate value of $|\alpha-1|$ as a function of the effective base-height ratio B/H_0 and the terrain slope m. The simulations were performed for values of α satisfying $|\alpha-1| < .2$ as follows

(1) A "white" noise sequence is generated using a random number generator.

(2) A "signal" is produced by filtering the noise sequence with the low-pass transfer function $e^{-f/s}$ where f is the normalized spatial frequency (Nyquist folding frequency = .5) and s is the "shaping factor" which controls the spatial frequency content of the signal; the signal bandwidth increases with s.

(3) A sample sequence, $\underset{\sim}{s}$, consisting of N_s samples is then selected and stored.

(4) A central portion of $\underset{\sim}{s}$ is distorted and resampled using Lagrange interpolation (Stark 1970) to produce a sequence $\underset{\sim}{r}$.

(5) White Gaussian uncorrelated noise of known power is added to each sequence independently.

Table 6-3. $|\alpha - 1|$ as a Function of Base-Height and Terrain Slope

$\frac{B}{H_0}$	m = .2	.4	.6	.8	1.0
.5	.10	.20	.30	.40	.50
.4	.08	.16	.24	.32	.40
.3	.06	.12	.18	.24	.30
.2	.04	.08	.12	.16	.20
.1	.02	.04	.06	.08	.10
0					

(6) The noisy match sequence, $r + n_r$, is then correlated with the noisy search sequence, $s + n_s$, to generate a correlation function.

(7) The computed peak of the correlation function is compared to the correct peak location and the error is retained.

(8) Steps 3-7 are repeated M times and the results averaged to yield MSE values for each value of the distortion parameter.

(9) Steps 2-8 are repeated for several shaping factors and signal-to-noise ratios.

One product of this procedure is a two-dimensional function which is similar in nature to the so-called "ambiguity function" encountered in radar signal analysis. The ambiguity function, $c(\tau,\alpha)$, describes the behavior of the correlation function as the parameter α is varied. Figure 6-14 shows typical mean ambiguity functions (i.e., statistical averages of M = 20 sample functions) for $s = 10.0$ and $s = 0.01$ for an eleven-pixel match window. Note that for $s = 10.0$, the primary effect of the distortion parameter is the reduction of the magnitude of the correlation function at the correct peak location. This effect is negligible, however, for broad correlation functions ($s = .01$). The functions in Figure 6-14 were generated using a normalized covariance processor; the results for unnormalized covariance are similar in appearance.

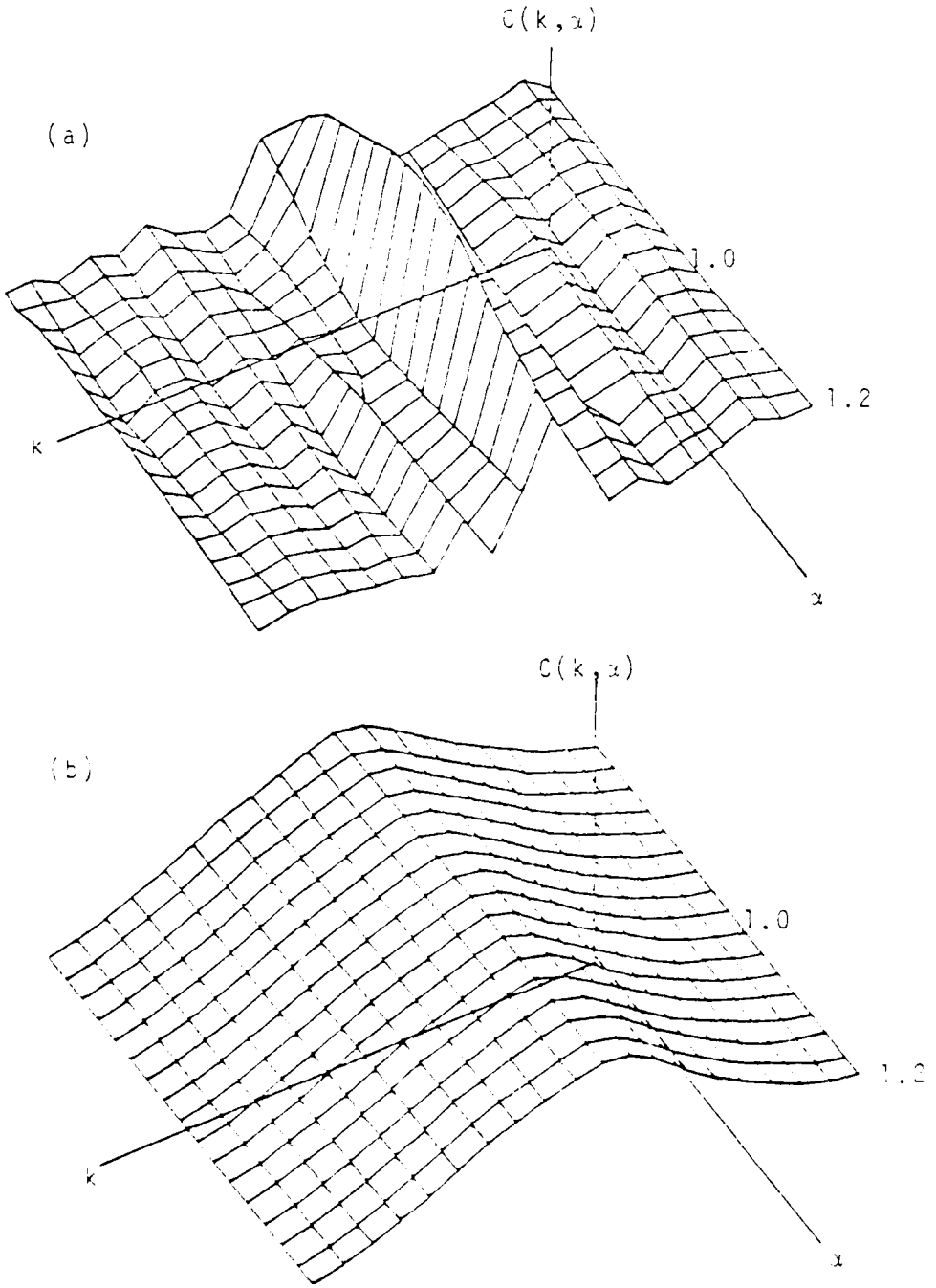


Figure 6-14. The Effect of Relief Distortion on the Normalized Covariance Function (a) S = 10.0, (b) S = .01

The mean-square correlation errors associated with the previously described procedure are shown in Tables 6-4 and 6-5 for simple covariance and normalized covariance respectively. Also given are the associated distortion measures defined by Equation (5-27). Each entry in the table corresponds to 20 repetitions (M = 20) at an input SNR of 30dB. The SNR is defined in terms of the total energy of the signal generated in step 2 and not in terms of the individual match window energies. For this reason, the true input SNR tends to increase with shaping factor s. For s > .1, the errors are primarily due to distortion; for s < .1 the low SNR results in increased error magnitude which is essentially independent of the distortion. A comparison of tables 6-4 and 6-5 implies that simple covariance is slightly more sensitive to distortion and degrades more rapidly with decreasing SNR than does normalized covariance. This data also indicates that one can expect distortion produced errors for values of the distortion measure d exceeding, say, 0.1. Thus the inequality in (5-28) should be satisfied by at least a factor of 10 for this application if distortion is to be ignored.

Some, but certainly not all, correlators in use at this time are capable of correcting for geometric distortion up to a terrain slope of about 50° (m = 1.2). From (6-12), this implies that

$$s - 1 \geq 1.2 \frac{B}{f_0} \qquad (6-5)$$

Table 6-4. The Effect of Distortion on Covariance Accuracy

(a) Correlation Mean Square Error

	Shaping Factor							
	0.010	0.017	0.030	0.100	0.300	1.000	3.000	10.000
.800	24.45	19.95	7.20	7.50	9.35	21.55	9.05	25.00
.825	26.45	24.75	7.70	7.50	7.65	15.70	7.55	19.60
.850	27.35	25.75	7.95	7.40	9.50	7.55	.70	13.10
.875	26.45	26.05	8.80	7.85	2.05	8.10	3.60	8.20
.900	25.90	21.20	9.40	7.30	.80	5.65	3.40	5.45
.925	30.15	25.30	9.65	7.15	.80	2.45	.20	5.00
.950	24.35	25.55	9.60	7.15	.80	0.00	0.00	0.00
.975	26.80	21.00	8.85	7.20	4.00	0.00	0.00	0.00
* 1.000	25.20	25.15	9.75	7.15	4.00	1.25	0.00	0.00
1.025	25.25	22.20	9.70	9.00	4.00	1.25	0.00	0.00
1.050	30.95	20.70	9.35	8.05	4.45	1.25	0.00	0.00
1.075	25.00	25.80	9.40	8.60	4.45	7.75	1.25	7.25
1.100	28.45	21.95	9.15	8.00	6.90	9.00	3.90	7.25
1.125	25.30	26.05	14.35	8.75	6.70	17.20	8.70	8.15
1.150	30.50	22.25	14.90	8.25	8.55	21.25	9.15	11.35
1.175	29.30	22.60	14.75	8.65	8.55	22.45	10.35	16.35
1.200	27.45	22.00	14.75	8.65	12.15	21.20	10.35	16.80

(b) Distortion Measure

	Shaping Factor							
	0.010	0.017	0.030	0.100	0.300	1.000	3.000	10.000
.800	.05	.04	.05	.06	.26	.45	.45	.57
.825	.05	.04	.04	.05	.23	.40	.40	.50
.850	.04	.03	.03	.05	.20	.34	.34	.42
.875	.03	.03	.03	.04	.17	.28	.29	.35
.900	.03	.02	.02	.03	.13	.23	.23	.28
.925	.02	.02	.02	.02	.10	.17	.17	.21
.950	.01	.01	.01	.02	.07	.11	.11	.14
.975	.01	.01	.01	.01	.03	.06	.06	.07
* 1.000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.025	.01	.01	.01	.01	.03	.06	.06	.07
1.050	.01	.01	.01	.02	.07	.11	.11	.14
1.075	.02	.02	.02	.02	.10	.17	.17	.21
1.100	.03	.02	.02	.03	.13	.23	.23	.28
1.125	.03	.03	.03	.04	.17	.28	.29	.35
1.150	.04	.03	.03	.05	.20	.34	.34	.42
1.175	.05	.04	.04	.05	.23	.40	.40	.50
1.200	.05	.04	.05	.06	.26	.45	.46	.57

Table 6-5. The Effect of Distortion on Normalized Covariance Accuracy

Shaping Factor

	0.010	0.017	0.030	0.100	0.300	1.000	3.000	10.000
.800	14.45	2.90	0.25	0.10	9.05	17.55	10.00	26.35
.825	5.85	1.15	0.15	0.05	6.05	14.25	7.35	18.00
.850	11.50	0.85	0.20	0.00	0.00	3.85	0.45	12.30
.875	10.70	0.90	0.05	0.00	0.80	3.20	3.20	5.05
.900	4.40	1.90	0.00	0.00	0.80	3.20	3.20	2.90
.925	7.60	1.65	0.05	0.00	0.00	0.00	0.00	0.00
.950	7.75	1.55	0.05	0.00	0.00	0.00	0.00	0.00
.975	8.85	1.10	0.00	0.00	0.00	0.00	0.00	0.00
1.000	6.10	3.55	0.05	0.00	0.00	0.00	0.00	0.00
1.025	6.35	1.40	0.05	0.00	0.00	0.00	0.00	0.00
1.050	10.00	2.75	0.05	0.00	0.00	0.00	0.00	0.00
1.075	11.70	0.70	0.05	0.00	1.25	2.45	0.00	7.25
1.100	10.15	3.70	0.05	0.00	1.25	10.55	1.45	7.25
1.125	14.70	1.30	0.10	0.00	3.70	12.15	9.15	7.30
1.150	12.35	0.55	0.15	3.20	3.65	16.20	10.35	7.75
1.175	7.50	2.05	0.15	3.20	3.70	16.20	10.35	7.75
1.200	13.35	0.80	0.25	3.25	3.75	16.75	10.40	7.75

(a) Correlation Mean Square Error

Shaping Factor

	0.010	0.017	0.030	0.100	0.300	1.000	3.000	10.000
.800	.05	.04	.05	.06	.26	.45	.46	.61
.825	.05	.04	.04	.05	.23	.40	.40	.49
.850	.04	.03	.03	.05	.20	.34	.34	.43
.875	.03	.03	.03	.04	.17	.29	.29	.35
.900	.03	.02	.02	.03	.13	.23	.23	.28
.925	.02	.02	.02	.02	.10	.17	.17	.21
.950	.01	.01	.01	.02	.07	.11	.11	.14
.975	.01	.01	.01	.01	.03	.06	.06	.07
1.000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.025	.01	.01	.01	.01	.03	.06	.06	.07
1.050	.01	.01	.01	.02	.07	.11	.11	.14
1.075	.02	.02	.02	.02	.10	.17	.17	.21
1.100	.03	.02	.02	.03	.13	.23	.23	.28
1.125	.03	.03	.03	.04	.17	.29	.29	.35
1.150	.04	.03	.03	.05	.20	.34	.34	.43
1.175	.05	.04	.04	.05	.23	.40	.40	.49
1.200	.05	.04	.05	.06	.26	.45	.46	.61

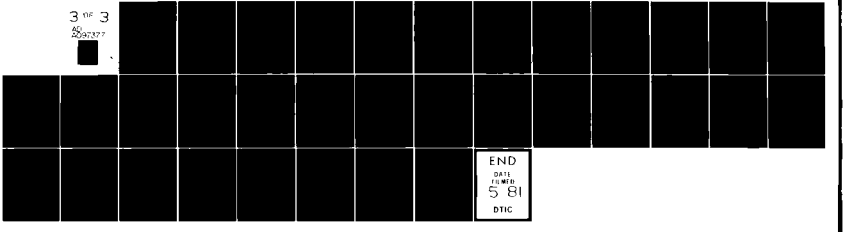
(b) Distortion Measure

If we can conclude that distortion can be ignored for $d < .1$, then we need not make geometric corrections if the signal bandwidth satifies

$$f_g < \frac{1}{\sqrt{12}\, N(B/H_0)} \quad . \tag{6-6}$$

Equation (6-6) suggests that one could pre-screen the stereo imagery to obtain a map of the locations that violate these conditions in order to expedite the processing. Note that if (6-6) is violated, it may be possible to reduce the window size N at these locations to reduce the distortion sensitivity although this must be done without significantly altering the false acquisition probability which depends on the output signal-to-noise ratio which, in turn, varies directly with $\sqrt{N}$.

AD-A097 377 ARIZONA UNIV TUCSON DIGITAL IMAGE ANALYSIS LAB F/G 9/2
AUTOMATIC METHODS IN IMAGE PROCESSING AND THEIR RELEVANCE TO MA--ETC(U)
FEB 81 B R HUNT DAAG29-77-G-0175
UNCLASSIFIED SIE/DIAL-81-002 ARO-14782.1-MS NL



CHAPTER 7

SUMMARY AND CONCLUSIONS

The major contribution of this dissertation is the development and experimental comparison of a number of image quality measures which, based on the results of Chapter 6, have exhibited potential for predicting the locations of "error prone" image data. In the process of developing these features, it was possible to establish a theoretical framework for the comparison of covariance, least squares, and normalized covariance as maximum likelihood processors (Chapter 2). Observations regarding the accuracy of these processors revealed that the small match window size and image non-stationarity violate the assumptions leading to expressions for theoretical correlator accuracy (Chapter 3) with the result that the performance, in terms of MSE, of the covariance processor is far worse than is predicted theoretically. Methods for detecting covariance errors associated with image non-stationarity have been considered and have met with limited success. The covariance function has been analyzed in detail to determine the relative correlation output signal-to-noise ratios associated with several noise sources (Chapter 4). This information was required in order to specify the parameters associated with the false acquisition measure Q. Finally, the effects of image distortion have been considered (Chapter 5) and we have

concluded that the prediction of distortion related errors is best accomplished "on-line", i.e., as the correlation process is underway, although needed signal bandwidth information could be provided in a pre-processing mode.

In this chapter, we provide a brief summary of the more important results and indicate, where applicable, the implications regarding the physical implementation of the various processing schemes. In Chapter 1, we alluded to the possibility of enhancing imagery to make it less susceptible to error producing mechanisms. We will briefly consider this topic here and provide some experimental results. Finally, a few comments concerning the possible directions of future work in this area will be presented.

Correlation Algorithms

In Chapter 2, we showed that if we assume the simple image formation model,

$$\underset{\sim}{r} = a\underset{\sim}{g} + n_r \quad , \tag{7-1}$$

then maximum likelihood correlation processors could be derived, the form of which depends on the a-priori assumptions regarding the parameter a. Covariance, least-squares, and normalized covariance are the products of this approach. If the stationarity assumption which leads to covariance as the ML processor is violated, disastrous error conditions can occur. If we assume that the conditions (original stereopair, geometry, etc.) under which this behavior was observed (see Figure 2-5) are not pathological, then the capability

of switching to a more reliable algorithm in the vicinity of regions violating stationarity is certainly desirable. The implementation of such a switching capability was discussed earlier (Chapter 4). The question arises, however, as to which alternative algorithm should be selected. Both least-squares and normalized covariance do not require the stationarity assumption. Least-squares requires the assumption that a = 1 while normalized covariance requires only the assumption of the image model which, as discussed earlier, is not entirely valid due to the signal dependence of the noise. Since these procedures are nearly equivalent in terms of computational load, it would seem natural to select normalized covariance. The correlation experiments in Chapter 2, however, seem to indicate that it may be advantageous to employ least-squares. In order to assess the adviseability of using least-squares as opposed to normalized covariance, the Monte Carlo correlation experiment was repeated for least-squares. The resulting overall MSE values for the 15 x 3 window and 20dB additive noise were .182 for least-squares and .615 for normalized covariance (measure in units of pixel spacing). Thus for the images and parameters of this simulation, least-squares is considerably more accurate than normalized covariance, particularly since the MSE values include the unavoidable contour-related errors. We note further that the maximum MSE associated with any image point was 20.0 for least-squares and 60.0 for normalized covariance. A close look at these error maps reveals that the contour-

related errors are slightly more pronounced for the least-squares processor which results in the contours being more "smeared out". The least-squares processor, however, is less susceptible to errors resulting from low SNR, at least for the source images employed in this simulation.

Since the ROC curves in Figures 6-7 and 6-9 are based only on the normalized covariance, the experiment was repeated for the least-squares error map. Figure 7-1 compares the ROC curves for least-squares and normalized covariance for the variance measure. The improvement in detection capability for least-squares, as well as the improvement in MSE, provides a strong argument in favor of least-squares as an alternative algorithm. If, however, the assumption that $a = 1$ is not valid, then it would be necessary to know the point at which normalized covariance surpasses least-squares in performance (if, indeed, such a point exists). We have not pursued this topic further.

The implementation of least-squares or normalized covariance requires only a few simple additions to Figure 4-1 in order to compute the energy of the search window subimage. The resulting processor is shown in Figure 7-2 in which covariance, least-squares, and normalized covariance are all shown as processor outputs. With such an implementation there is a small decrease in throughput rate and only a slight increase in complexity. Because least-squares requires only a shift register (multiplication by 2) and a

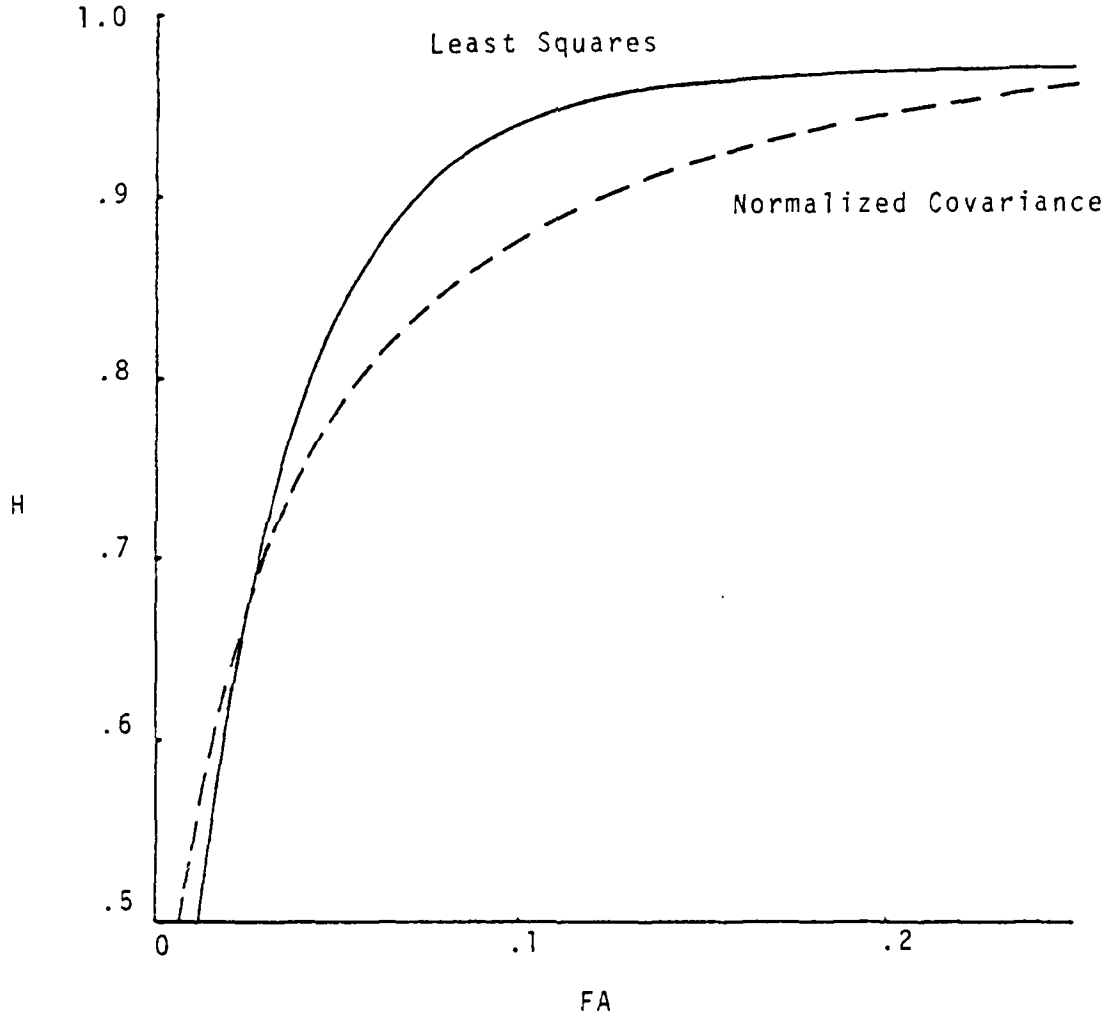


Figure 7-1. ROC curves for the Variance Measure

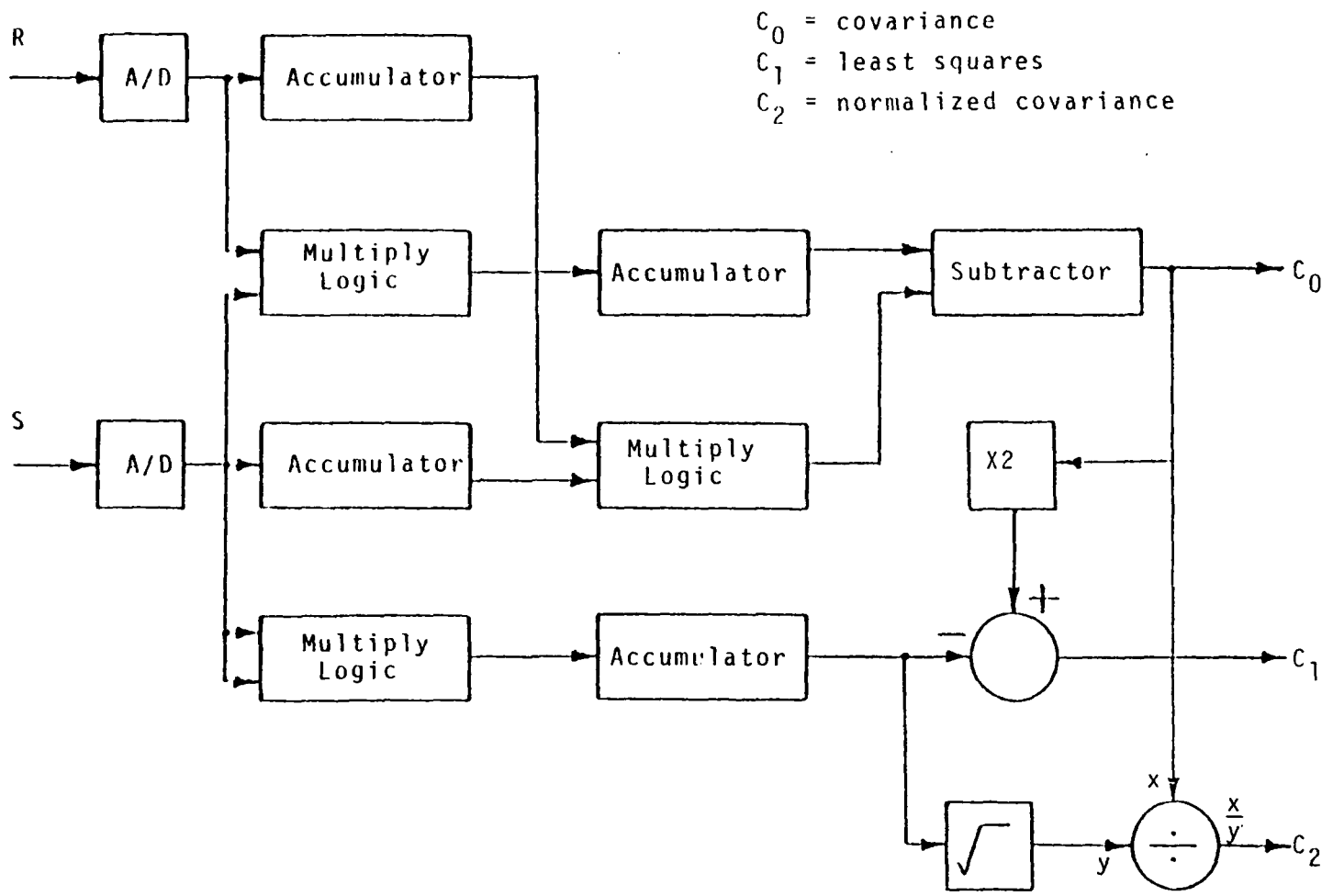


Figure 7-2. Correlation Processors

subtraction circuit, it is the more desireable from a hardware viewpoint. From our observation regarding relative processor accuracy, it would seem that such a minor increase in complexity is well worth the investment.

Error Variance

In the analysis of local registration errors, we assumed that the correct correlation lobe is selected, and furthermore that the image subsections are large enough so that the Fourier components are uncorrelated. This approach led to the development of the Cramer-Rao bound on the accuracy of any unbiased correlation processor. An analysis of the generalized correlator resulted in a generalized maximum likelihood correlator (given by (2-36) and (2-37)) which achieves the lower bound under aforementioned assumptions plus stationarity. Under conditions of high signal-to-noise ratio, the covariance processor is nearly equivalent to this "minimum variance" processor. This equivalence, however, implies that the generalized ML processor will also be susceptible to edge-produced errors since these edges typically are associated with regions of high SNR. Thus, while the theory establishes relationships which aid in the understanding of the correlation process, we have found that the violation of assumptions necessitates an empirical approach regarding specific processor accuracies, at least for the application at hand where the small size of the match window is a significant factor. On the other hand,

when correlating large regions of imagery, the statistics of the match area are usually representative of the statistics of the search area and the presence of small regions with atypical statistics can be ignored.

The analysis of false acquisition errors led to the development of the image quality measure Q which depends on the local image signal-to-noise ratio. The value of Q is related to the single-tap probability of false acquisition. The overall probability of false-acquisition was found to be impossible, in general, to obtain analytically due to the non-zero correlator tap cross-covariances. Thus, unlike the Cramer-Rao bound, the magnitude of Q cannot be directly related to mean-square-error without extensive simulation. The form of Q depends on the assumptions made concerning the degree to which self-noise contributes to error production. The results of Chapter 6 have shown, however, that the ROC curve associated with Q is independent of this assumption (i.e., choice of λ). Furthermore, the low apparent sensitivity of the threshold of the $\lambda = 0$ version of Q to noise contamination makes this feature an attractive candidate. A comparison of the range of magnitudes of the features associated with the three versions of Q ($\lambda = 0$, $\lambda = \frac{1}{N}$, λ adaptive) are shown in Figure 7-3. The numbers marked on the range designation bars indicate the hit rate associated with the particular value of Q. Since we cannot translate the single-tap false acquisition probability to overall false acquisition probability it is difficult to determine

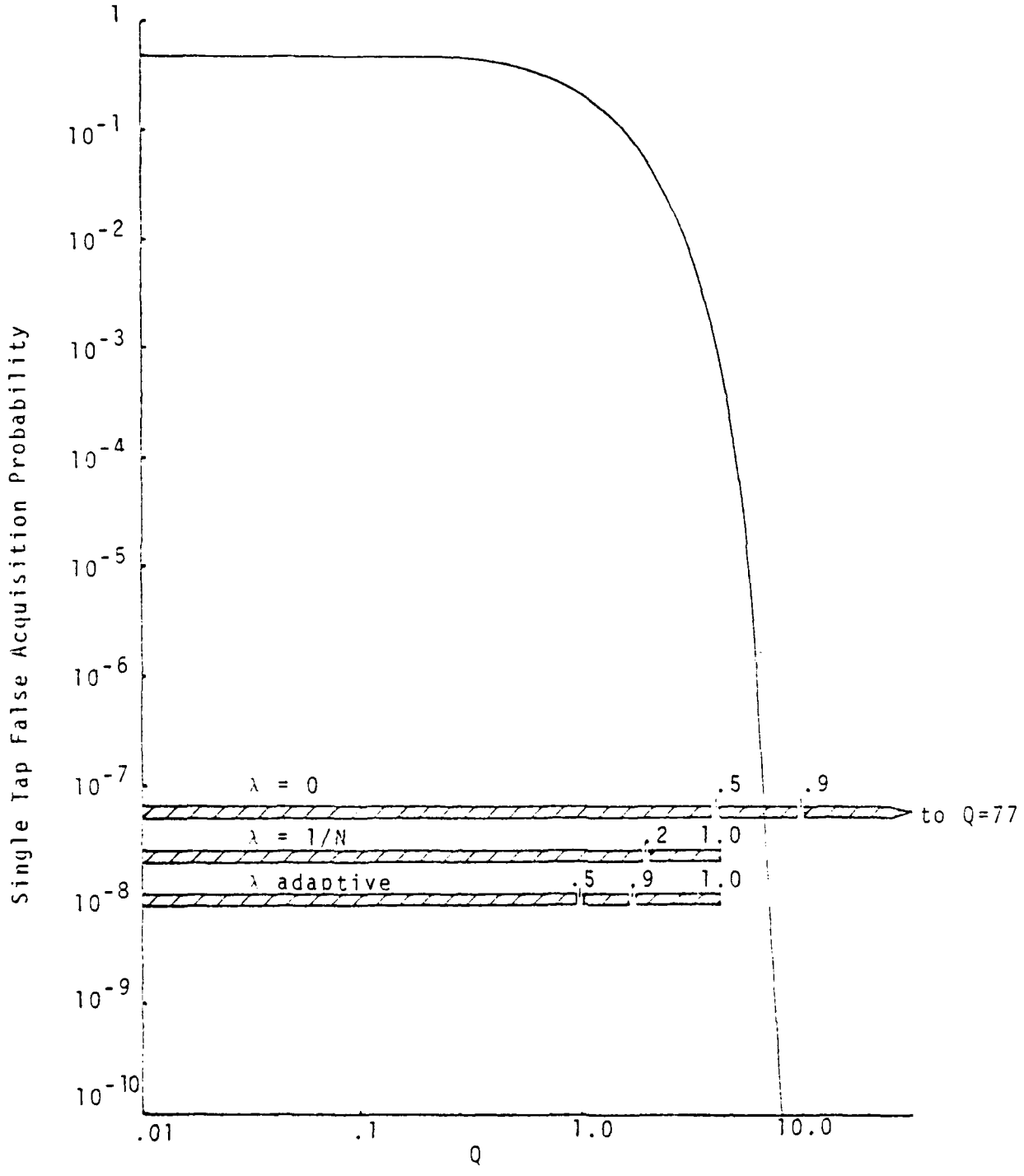


Figure 7-3. Comparison of the Computed Magnitudes of Q

which model most accurately represents the process. One advantage of the $\lambda = 0$ model is that the change in detection rate, say from .5 to .9, requires a larger percentage change in threshold (Q) than for the other models. Thus this measure is least sensitive to errors in the choice of Q.

The detection capabilities of the Cramer-Rao bound and Q are remarkably similar for the 15 x 3 match window, no-noise measurements. This similarity is not so surprising however, when the nature of these measures is considered. As stated previously, the Cramer-Rao bound measures the curvature of the autocorrelation function at the origin (or peak) whereas Q is directly related to the peak magnitude of the autocorrelation function. It is not unreasonable to expect these quantities to be highly correlated with the resulting coupled behavior depicted in Figure 7-4; that is, as the peak of the function decreases ($c_2(0) < c_1(0)$), the autocorrelation width increases. Although there are exceptions to this behavior, they do not appear to occur with sufficient frequency to alter the detection capabilities. Thus the erratic behavior of the Cramer-Rao measure in the presence of noise makes the use of Q all the more attractive. Furthermore, since the magnitude of the false acquisition errors tend to be larger by their very nature (see Figure 1-1) than local registration errors, they are more likely to produce a "lost" condition in a tracking correlator and, therefore, a measure which is sensitive to image data resulting in false acquisitions is certainly preferable.

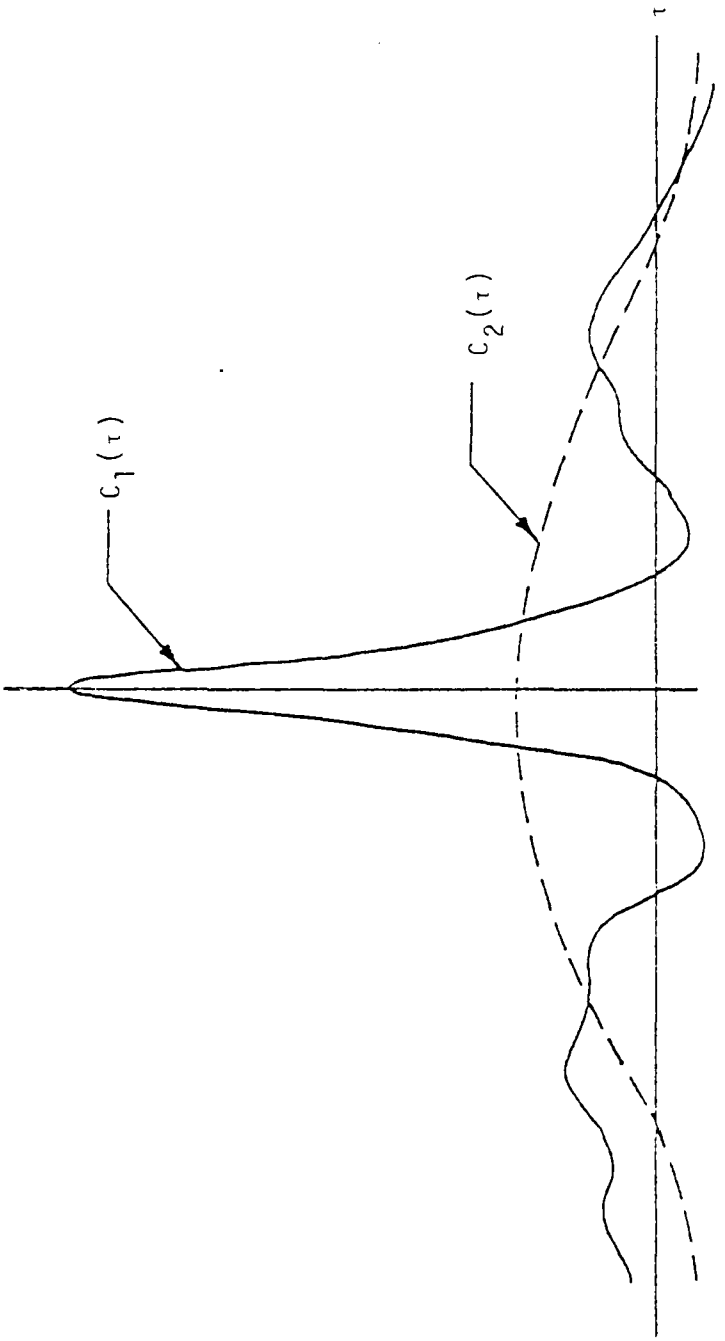


Figure 7-4. Peak Magnitude-Curvature Dependence

Implementation of Image Quality Measures

The computation of image features can be performed digitally by scanning the film in a manner similar to the scanning required by the stereo-compiler. Once the subimage of interest has been digitized, the features listed in Table 6-1 can be computed in a straightforward manner. The contrast measures require the least computation since a series of compare operations will yield the necessary maximum and minimum subimage density values. The Cramer-Rao bound requires the most computation since we must compute either the pericdogram or the autocorrelation function.

The limiting form of the Cramer-Rao measure given by (3-22), however, can be computed optically without digitizing the source data. This is accomplished by using coherent illumination and appropriate lenses to create the optical Fourier transform. A filter which weights each frequency domain component by the radial component of the spatial frequency is placed in the frequency plane and the resulting irradiance distribution is sensed, integrated, and recorded. This energy represents the second moment of the subimage power spectrum. An aperture which provides a pre-transform windowing function can also be provided. A quick calculation involving the noise power estimate then yields the Cramer-Rao measure. It should be mentioned that a limiting factor in any conerent optical computation is the so-called

"speckle effect" [Considine 1966] which is a consequence of the fact that the impulse response function of the optical system is complex which allows for destructive interference which would not occur in an incoherent system. It is partly because of this effect that incoherent optical spatial frequency analysis techniques are being developed [Cole 1980], [Rhodes 1980].

The variance of a subsection can also be sensed optically by coherently illuminating the desired subsection, blocking "dc" in the frequency plane and integrating over the Fourier irradiance distribution. Given the variance and the estimated noise power, the quality measure Q can be easily computed.

The optical measurement of the contrast measures is not so straightforward. Recent advances in non-linear optical image processing have resulted in feedback systems using Fabry-Perot interferometers containing phase recording media [Bartheolemew and Lee 1980]. At the time of this writing, these systems are capable of performing optical image thresholding and analog to digital conversion with a resolution of about 8 gray-levels [Atkins et al. 1980] which does not appear to be sufficient for the application at hand.

Relief Distortion

In Chapter 5, a model was developed which describes the geometric distortion between stereo-pair images as a function of the terrain topography. The effect of this distortion on the correlation process was shown to be a reduction in the peak height of the correlation function with resulting increases in local registration and false-acquisition errors. This behavior, however, was shown to be independent of the statistics of the local image subsection in a manner described by the distortion measure given in (5-27). Simulations described in Chapter 6 verified the nature of the distortion measure. The use of such a measure as an error predictor is not straightforward however, since the distribution of the linear distortion parameter α is difficult to obtain without performing the stereo-compilation process. An attempt at estimating the statistics of α, however, led to an "image overlay quality" measure given by Equation (5-58). The utility of such a measure in the prediction of correlation behavior is, at this point, uncertain.

The most promising approach for dealing with relief distortion in a pre-processing mode is to compute a map of the image signal bandwidth to be combined with terrain slope information acquired during the correlation processing to yield an estimate of the distortion measure in locations where the bandwidth exceeds the constraint specified in Equation (6-6). Geometric correction could be applied at these points with well-known procedures.

Image Enhancement for Correlation

In this section, we present three procedures for suppressing additive noise which may find an application in improving correlation accuracy. Since the statistics of an image are usually non-stationary, both spatially variant and spatially invariant techniques are considered.

The first approach involves simple low-pass filtering using a spatial averaging filter. The motivation for this procedure derives from considering the effect of low-pass filtering on the correlator input SNR. Let $\{x_1, x_2, \ldots, x_N\}$ represent a sequence of samples from a stochastic process and let

$$y = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{7-2}$$

represent the averaging process. If we assume that the $\{x_i\}$ each have mean $\bar{x}$ and variance σ_x^2 then it follows that

$$\bar{y} = E\{y\} = E\{x\} = \bar{x} \tag{7-3}$$

and

$$\sigma_y^2 = E\{(y-\bar{y})^2\} = \frac{1}{N}\sigma_x^2 + \frac{1}{N^2} \sum_{\substack{i=1 \\ i\neq j}}^{N} \sum_{j=1}^{N} C_x(i,j) \tag{7-4}$$

where $C_x(i,j) = E\,(x_i-\bar{x})(x_j-\bar{x})$ is the $(i,j)^{th}$ element of the covariance matrix of $\{x_i\}$. If $\{x_i\}$ represents samples of uncorrelated noise, then

$$\sigma_y^2 = \frac{\sigma_x^2}{N} \quad . \tag{7-5}$$

If we now let $\underline{g}$ represent the underlying image data, then the signal-to-noise ratio of the filtered image, SNR_f, is given by

$$SNR_f = \frac{\frac{1}{N}\sigma_g^2 + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N} C_g(i,j)}{\frac{1}{N}\sigma_\eta^2} \tag{7-6}$$

$$= SNR + \frac{1}{N\sigma_\eta^2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N} C_g(i,j) \tag{7-7}$$

Thus if the image data is correlated, there can be a net improvement in SNR (this can also be easily seen by considering the effect of low-pass filtering in the frequency domain). The fractional improvement in SNR is given by

$$\frac{SNR_f - SNR}{SNR} = \frac{1}{N\sigma_g^2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N} C_g(i,j) \tag{7-8}$$

Equation (7-8) implies that there may exist an optimum value of N which depends only on the form of C_g. Thus if

$$\sum_{i=1}^{N} [C_g(i,N+1) + C_g(N+1,i)] > \sigma_g^2 \tag{7-9}$$

we can obtain further improvement in SNR by increasing the size of the averaging window. It is also possible that (7-9) can be continued to be satisfied for all N as N increases without bound. Although it is not shown

explicity, (7-9) can be used to determine the optimum shape and orientation (if one exists) of the averaging filter as well as the total number of points since the spatial proximity of points in the averaging window will have a bearing on the values of the elements in C_g. The geometry of the filter which maximizes the SNR, however, does not necessarily minimize the correlation MSE, particularly if the correlation process is carried out along epipolar lines. For example, a 24-point filter can have a variety of shapes and orientations (2x12, 12x2, 4x6, etc.). Smoothing in the direction parallel to epipolar lines will reduce the false acquisition probability (due to improved SNR) at the expense of an increase in local registration errors due to the smoothing of the resulting correlation functions. Since local registration errors tend to be smaller than false acquisition errors, this effect may be desirable. Furthermore the reduction in signal bandwidth produced by such filtering results in reduced sensitivity to relief distortion. On the other hand, smoothing in the direction perpendicular to the epipolar lines will result in improved SNR without drastically smoothing the correlation functions.

Because of the spatial warping due to relief, however, points in one image lying along a line perpendicular to an epipolar line are not necessarily colinear in the other image. Thus averaging data in the cross-epipolar

direction can reduce the similarity (correlation coefficient) between the filtered signals. With these considerations in mind, it is clearly more difficult to specify the optimum window geometry.

A further limitation is imposed on the filtering process by the quantization required for subsequent digital processing. Low-pass filtering reduces the dynamic range of the signal which decreases the correlator output signal-to-quantization noise ratio as well as the signal-to-machine noise ratio (see Equations (4-81) and (4-82). Thus excessive filtering can result in inferior correlator behavior.

An experiment was performed to observe the effect of low-pass spatially invariant filtering (using averaging) on the accuracy of the normalized covariance stereo-compilation. Table (7-1) displays the MSE resulting from tne correlation process as a function of the geometry of the averaging window. From these results, it is clear that significant improvement in MSE (on the order of 70%) can be achieved by simple low-pass filtering. These results also imply tnat orientation is, indeed, a significant factor. Note that better results are obtained by smoothing primarily in the direction perpendicular to the epipolar search.

Finally, we note that (7-2) can be generalized to provide for more arbitrary filtering by allowing a

Table 7-1. Residual MSE as a Functioning of Averaging Window Geometry, Unfiltered MSE = .2659

Number of Lines \ Number of Pixels	1	3	5	7	9
17	.309				
15	.320	.223			
13	.329	.222	.240		
11	.358	.240	.238	.336	
9	.383	.256	.261	.279	.291
7		.325	.294	.283	.287
5		.420	.360	.350	
3		.481	.426		
1	.866				.707

weighting sequence $\{w_i\}$ so that

$$y = \frac{1}{N} \sum w_i X_i \quad . \tag{7-10}$$

The second approach to noise suppression is based on the common enhancement procedure known as unsharp masking [Pratt 1978]. This procedure consists of subtracting a blurred version (the "mask") of the original from the original, and then adding the weighted difference image back to the original. The process can be described by the relation [Gray, Hunt, and McCaughey 1979]

$$Y = G[X-M(X)] + \alpha M(X) \tag{7-11}$$

where X is the original image, M(X) is the mask of the original, G is an arbitrary gain function which may be spatially variant, and α is a fixed parameter which weights the mask image with respect to the difference image. For the present, we let $\alpha = 1$ and define G by

$$G = \begin{cases} \sqrt{\dfrac{\sigma_x^2 - k\sigma_n^2}{\sigma_x^2}} & \text{if } \sigma_x^2 > k\sigma_n^2 \\ 0 & \text{otherwise} \end{cases} \tag{7-12}$$

where σ_x^2 is the measured local variance and k is an 'adjustment factor' which adds some flexibility to the process. In regions dominated by noise, the output image approaches the mask, whereas in high SNR regions, the output approach the original. A preliminary experiment was performed in which the masking function was simply an averaging filter which replaces the center pixel of an MxM window with the mean

value of the window. Results are shown in Table 7-2 as a function of the parameter k for M = 3 and M = 5. As $k \rightarrow \infty$, this method approaches simple averaging. The results from the invariant averaging, however imply that improved results may be achieved by a judicious choice of window geometry. Comparison of identical geometries between the methods implies that there is no advantage to unsharp masking.

Finally we consider a procedure known as short space spectral subtraction which has been recently developed [Lim 1980] and has shown considerable promise as a noise cleaning scheme which is performed prior to the application of image restoration procedures (e.g., deblurring). This approach consists of sectioning the original image into overlapping blocks (16x16 pixels), windowing the individual subsections and removing the estimated noise by power spectral subtraction, a procedure wherein only the amplitude of the Fourier components are altered according to the estimated noise spectrum which is weighted by a factor k similar to Equation (7-12). The filtered blocks are then reassembled to form the cleaned image. Results are tabulated in Table 7-2 for several values of k.

Although all three noise suppression procedures provide reduction in correlation MSE, the most encouraging approach, from both computational complexity and resulting MSE viewpoint, is the simple low-pass filter. Spectral subtraction is certainly the most complex in terms of implementation and does not provide the flexibility available

with the other methods. As stated previously, further reductions in MSE can probably be obtained by optimizing the window geometry and filter weights for the low-pass filtering technique. By observing the histograms of the correlation errors before and after the filtering, it can be seen that, regardless of the technique, there is a trade-off between false acquisition and local registration errors. In other words, the filtering reduces the frequency of the relatively large errors at the expense of a small (< 5%) reduction in correct acquisitions and an increase in small error (particularly 1 pixel-errors) frequency.

Table 7-2. Residual MSE-Unsharp Masking and Short Space Spectral Subtraction Noise Suppression, Unfiltered MSE = .8659

Method	k	Residual MSE
Unsharp Masking (5x5 mask)	1.0	.8909
	2.0	.6421
	3.0	.6133
Unsharp Masking (3x3 mask)	3.0	.5928
	4.0	.5537
	6.0	.5261
Short Space Spectial Subtraction	1.92	.7410
	2.56	.7123
	3.20	.6997

Extensions

The experimental results presented herein are based on the properties of the synthetic images shown in Figure 2-5. While there is no reason to believe that the results are not representative of the nature of the correlation process and its predictability, there are obvious shortcomings which would require extensive simulation work to alleviate. To begin, the accuracy of the correlation processors is dependent on the data acquisition geometry and while we have included this factor in the form of the base-height ratio $(\frac{B}{H_0})$, there are trade-offs involved. As an example, we note from Equation (1-8) that an increase in base-height ratio results in improved elevation resolution but from Equation (5-12) we see that the increase in elevation resolution comes at the expense of increased distortion which degrades the correlator performance. It would be interesting to consider this trade-off in more detail to determine if optimum conditions can be specified.

The correlator models developed in Chapter 4 provide information regarding the necessary degree of machine precision in order to relegate machine noise as a second order error source. Further simulation work is required to verify these models for both fixed point and floating point processing.

The detection of sources of edge-related errors was based on the gradient of the variance map. It is clear from the results that improvement in the detection capability will require a look at the form of the individual image data sequences with respect to the form of the edge itself. However, it is preferable that such a processor not be required to use correlation as a similarity measure (due to the computational load). In Chapter 4, we presented two approaches to the detection of such regions (variance gradient and and variance outlyer) and it is conceivable that a combination of these approaches in a pattern recognition mode would yield an improvement in detection capability.

The use of classical pattern recognition schemes [Fukunaga 1972], [Kasdan 1971] in which a variety of features are combined in the detection processes has been considered in the context of correlation prediction [Ryan, Gray, and Hunt 1980]. In such a procedure, feature values associated with known correlation behavior (MSE) are input to a pattern recognition package in a "learning" mode. The pattern recognition algorithms rank and weight each feature to maximize the detection statistics over the known behavior. The weightings can then be used in the detection mode on data for which the correlation behavior is to be assessed. Preliminary investigations in this area using the features developed herein, have indicated that little, if any, improvement in detection capability is gained by such a combination, presumably because

all the features are, in effect, measuring the same image qualities (texture or contrast). Further work in this area, however, is certainly worthwhile although the acquisition of multiple features as well as the required software overhead for such a recognition capability must be considered.

APPENDIX A

DERIVATION OF EQUATION (3-36)

We assume that $c(0)$ and $c(\xi)$ are both Gaussian distributed with pdf's

$$p_{c(0)}(x) = \frac{1}{\sqrt{2\pi}\sigma(0)} e^{-\frac{(x-DC(0))^2}{2\sigma^2(0)}} \qquad (A-1)$$

and

$$p_{c(\xi)}(x) = \frac{1}{\sqrt{2\pi}\sigma(\xi)} e^{-\frac{(x-DC(\xi))^2}{2\sigma^2(\xi)}} .$$

If we assume that $c(0)$ and $c(\xi)$ are statistically independent, then

$$\text{Prob}[c(\xi)>c(0)] = \int_{-\infty}^{\infty} \int_{-\infty}^{x} p_{c(0)}(y)dy \; p_{c(\xi)}(x)dx \quad . \quad (A-2)$$

With substitution from (A-1) and a change of variables given by

$$V = \frac{x-DC(\xi)}{\sigma(\xi)} \quad ,$$

(A-2) becomes

$$\text{Prob}[c(\xi)>c(0)] = \int_{-\infty}^{\infty} \int_{-\infty}^{v\sigma(\xi)+DC(\xi)} \frac{1}{\sqrt{2\pi}\sigma(0)} e^{-\frac{(y-DC(0))^2}{2\sigma^2(0)}} \text{Gaus}(v)dv \quad . \qquad (A-3)$$

Similarly, the substitution

$$u = \frac{y - DC(0)}{\sigma(0)}$$

leads directly to (3-36).

APPENDIX B

DERIVATION OF EQUATION (5-17)

We begin with Equation (5-16):

$$E(r(\xi)) = \int_{-\frac{D}{2}}^{\frac{D}{2}} C_f(\xi + x - h(x))dx \quad . \qquad (B\text{-}1)$$

If $u(x) = h(x) - x$ is a monotonic function of x, then there exists a function z with $z(u) = x$. With this change of variables, (B-1) becomes

$$E(r(\xi)) = \int_{h(-\frac{D}{2})+\frac{D}{2}}^{h(\frac{D}{2})-\frac{D}{2}} z'(u)C_f(\xi - u)du \quad . \qquad (B\text{-}2)$$

Since $u(x)$ is monotonic then either $du/dx > 0$ or $du/dx < 0$ for all $x\varepsilon[-\frac{D}{2}, \frac{D}{2}]$. But

$$\frac{du}{dx} = h'(x) - 1 = \frac{1}{z'(u)} \quad . \qquad (B\text{-}3)$$

Thus if $du/dx > 0$, then $z'(u) > 0$ and

$$h(\tfrac{D}{2}) - \tfrac{D}{2} > h(-\tfrac{D}{2}) + \tfrac{D}{2} \qquad (B\text{-}4)$$

so that

$$E(r(\xi)) = \int_{h(-\frac{D}{2})+\frac{D}{2}}^{h(\frac{D}{2})-\frac{D}{2}} |z'(u)|C_f(\xi - u)du \quad , \qquad (B\text{-}5)$$

and if $du/dx < 0$, then $z'(u) < 0$ and

$$h(\tfrac{D}{2}) - \tfrac{D}{2} < h(-\tfrac{D}{2}) + \tfrac{D}{2} \qquad \text{(B-6)}$$

so that, again,

$$E(r(\xi)) = \int_{h(\frac{D}{2})-\frac{D}{2}}^{h(-\frac{D}{2})+\frac{D}{2}} |z'(u)| C_f(\xi - u)du \quad . \qquad \text{(B-7)}$$

Combining (B-5) and (B-7), we have

$$E(r(\xi)) = \int_{-\infty}^{\infty} |z'(u)| \text{rect}\left[\frac{u - \frac{1}{2}(h(\frac{D}{2})+h(-\frac{D}{2}))}{|D+h(-\frac{D}{2})-h(\frac{D}{2})|}\right] C_f(\xi-u)du \qquad \text{(B-8)}$$

from which (5-17) follows.

LIST OF REFERENCES

1. Akins, R. P., Athale, R. A., and Lee, S. H., "Feedback in Analog and Digital Optical Image Processing", Optical Engineering, Vol. 19, No. 3, May/June, 1980.

2. Andrews, H. C. and Hunt, B. R., Digital Image Restoration, Prentice Hall, New Jersey, 1977, pp. 20-23.

3. Andrews, L. C., "Output Probability Density Functions for Cross Correlators Utilizing Sampling Techniques", IEEE Trans. on Aero. and Elec. Systems, Vol. 10, No. 1, January, 1974, pp. 78-80.

4. Apostol, T. M., Mathematical Analysis, Addison Wesley, Reading, Mass., 1974, p. 37.

5. Balsubramanian, N., "Photogrammetric Aspects of the Heterodyne Optical Correlator", U. S. Army Engineer Topographic Laboratories Report ETL-0071, Fort Belvoir, Virginia, 1976.

6. Barnea, D. I. and Silverman, H. F., "A Class of Algorithms for Fast Digital Image Registration", IEEE Trans. on Computers, Vol. C-21, February, 1972, pp. 179-186.

7. Bartholomew, B. J. and Lee, S. H., "Nonlinear Optical Processing with Fabry-Perot Interferometers Containing Phase Recording Media", Applied Optics, Vol. 19, No. 2, January, 1980, pp. 201-106.

8. Brumm, G. A., Whiteside, A. E., Bedross, G. M., Forrest, L. A., and Hutchenreuther, A. J., "AS-11B-X Automated Stereo-Mapper", Bendix Research Lab., Final Technical Report RADC-TR-76-100, April, 1976.

9. Casasent, D. and Psaltis, D., "Position, Rotation, and Scale Invariant Optical Correlation", Applied Optics, Vol. 15, No. 7, July, 1976, pp. 1795-1799.

10. Cole, T. W., "Spatial Frequency Analysis with a New Incoherent Optical Approach", Applied Optics, Vol. 19, No. 10, May, 1980, pp. 1665-1669.

11. Considine, P. S., "Effects of Coherence on Imaging Systems", J. Opt. Soc. Am., Vol. 56, No. 8, 1966,

12. Crawley, B. G., "Gestalt Contours", The Canadian Surveyor, Vol. 28, No. 3, 1974,

13. Fukunaga, K., Introduction to Statistical Pattern Recognition, Academic Press, New York, 1972.

14. Gray, R. T., Hunt, B. R., and McCaughey, D. G., "A Median Masking Technique for the Enhancement of Digital Images", Proc. of SPIE, Vol. 27, San Diego, Calif., August, 1979, pp. 142-144.

15. Harris, F. J., "On the Use of Windows for Harmonic Analysis with the DFT", Proc. IEEE, Vol. 66, No. 1, January, 1978, pp. 51-83.

16. Helava, U. V., "Digital Correlation in Photogrammetric Instruments", Photogrammetria, 34:10-41, 1978.

17. Hunt, B. R. and Ryan, T. W., "Prediction of Correlation Errors in Parallax Computations from Digital Stereo Images", Proc. SPIE, Vol. 149, San Diego, Calif., August, 1978, pp. 222-231.

18. Kasdan, H. L., Non-Parametric Pattern Recognition, Ph.D Dissertation, UCLA, 1971.

19. Knapp, C. H. and Carter, G. C., "The Generalized Correlation Method for Estimation of Time Delay", IEEE Trans. ASSP, Vol. ASSP-24, No. 4, August, 1976, pp. 320-327.

20. Lim, J. S., "Image Restoration by Short Space Spectral Subtraction", IEEE Trans. ASSP, Vol. ASSP-28, No. 2, April, 1980, pp. 191-197.

21. MacDonald, V. H. and Schultheiss, P. M., "Optimum Passive Bearing Estimation", J. Acoust. Soc. Amer., Vol. 46, No. 1, January, 1969, pp. 37-43.

22. Mikhail, E. M., "Photogrammetric Aspects of the Heterodyne Optical Correlator", U. S. Army Engineer Topographic Laboratories Report ETL-0095, Fort Belvoir, Virginia, 1976.

23. Mikhail, E. M., Unruh, J. E., and Alspaugh, D. H., "Sensor Simulation from Spectral and Digital Terrain Data", Purdue Research Foundation, Final Technical Report, April, 1977.

24. Mostafavi, H., "Optimum Window Functions for Image Correlation in the Presence of Geometric Distortion", IEEE Trans. on ASSP, Vol. ASSP-27, No. 2, April, 1979, pp. 163-169.

25. Mostafavi, H. and Smith, F. W., "Image Correlation with Geometric Distortion, Part I: Acquisition Performance", IEEE Trans. on Aero. and Elect. Systems, Vol. AES-14, No. 3, May 1978a, pp. 487-493.

26. ibid, "Image Correlation with Geometric Distortion, Part II: Effect on Local Accuracy", Vol. AES-14, No. 3, May, 1978b, pp. 494-500.

27. Oppenheim, A. V., ed., Applications of Digital Signal Processing, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1978.

28. Oppenheim, A. V. and Schafer, R. W., Digital Signal Processing, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1975.

29. Panton, D. J., "A Flexible Approach to Digital Stereo Mapping", Photogrammetric Engineering and Remote Sensing, Vol. 44, No. 12, December, 1978, pp. 1499-1512.

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

31. Pearson, J. J., Hines, D. C., Golosman, S., and Kuglin, C. D., "Video-Rate Image Correlation Processor", Proc. of SPIE, Vol. 119, 1977, pp. 197-205.

32. Pratt, W. K., "Correlation Techniques of Image Registration", IEEE Trans. on Aero. and Elect. Systems, Vol. AES-10, No. 3, May, 1974, pp. 353-358.

33. ibid, Digital Image Processing, John Wiley and Sons, New York, 1978.

34. Rhodes, W. T., "Incoherent Spatial Filtering", Optical Engineering, Vol. 19, No. 3, May/June, 1980, pp. 323-330.

35. Ryan, T. W. and Hunt, B. R., "The Prediction of Accuracy in Digital Cross-Correlation of Stereo-Pairs Images", Proc. of SPIE, Vol. 219, Hollywood, Calif., February, 1980a, pp. 121-128.

36. Ryan, T. W., Gray, R. T., and Hunt, B. R., "Prediction of Correlation Errors in Stereo-Pair Images", Optical Engineering, Vol. 19, No. 3, May/June, 1980, pp. 312-322.

37. Skolnik, M. I., Introduction to Radar Systems, McGraw Hill, New York, 1962.

38. ibid, ed., Radar Handbook, McGraw Hill, New York, 1970.

39. Stark, P. A., Introduction to Numerical Methods, MacMillan, New York, 1970.

40. Thompson, M. M., ed., Manual of Photogrammetry, American Society of Photogrammetry, Falls Church, Virginia, 1966.

41. Van Trees, H. L., Detection Estimation and Modulation Theory, Part I, John Wiley and Sons, New York, 1968.

42. Webber, R. F. and Delashmit, W. H., "Product Correlator Performance for Gaussian Random Scenes", IEEE Trans. on Aero, and Elect. Systems, Vol. AES-10, No. 4, July, 1974a, pp. 516-520.

43. ibid, "Linear Scale-Factor Error and Optimum Filter Bandwidth for Correlation Accuracy", Electronics LeHers, Vol. 10, October, 1974b, pp. 414-415.

44. Wernecke, S. J., "Statistical Analysis of Multitap Correlator Output Errors and TDOA Measurement Precision", Technical Memorandum 892, ESL, Inc., Sunnyvale, Calif., 1978.

45. Whalen, A. D., Detection of Signals in Noise, Academic Press, New York, 1971.

46. Wong, R. Y. and Hall, E. L., "Sequential Hierarchical Scene Matching", IEEE Trans. on Computers, Vol. C-27, No. 4, April, 1978, pp. 359-367.

END

DATE
FILMED

5 81

DTIC