To the Graduate Council:

I am submitting herewith a thesis written by Jyh-Chyuan Su entitled "A Comparison of Registration Metrics for Correlating Non-Compatible Images." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Electrical Engineering.

L. Pinson

Lewis J. Pinson, Major Professor

We have read this thesis
and recommend its acceptance:

Roy D Joseph

[illegible]

Accepted for the Council:

L Evans Roth

Vice Chancellor
Graduate Studies and Research

A COMPARISON OF REGISTRATION METRICS FOR CORRELATING NON-COMPATIBLE IMAGES

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Jyh-Chyuan Su

March 1982

ACKNOWLEDGMENTS

The work reported herein was supported in part by U. S. Air Force Contract AFOSR-80-0275.

The author wishes to express his appreciation to his advisor and committee chairman, Dr. Lewis J. Pinson, for his guidance and professional assistance throughout this thesis. The author would also like to thank the other committee members, Dr. Roy D. Joseph and Dr. Bharat K. Soni, for their suggestions.

Special thanks are due to Dr. C. L. Rao and Mr. Jeffrey Lankford for their assistance in developing computer programs. Thanks are likewise expressed to The University of Tennessee Space Institute library staff for cheerfully providing all the needed reference material, and to Mrs. June F. Jarrell for typing the manuscript.

The author would like to thank his parents, his wife and son, for their support and encouragement. This work is dedicated to them.

ABSTRACT

In the field of pattern recognition it is often necessary to locate objects in a given image by referring to their position in another image of the same scene. Thus, in image processing it is important to accurately and efficiently register (match) two images of the same scene even when they may be recorded by separate sensors under different conditions.

In a search for an improved registration measure, in this work various metrics were used to register a set of compatible and non-compatible images. Specifically, the standard metrics, including the direct cross-correlation (DCC), fast Fourier transform (FFT) and mean absolute different (MAD) metrics, were compared to the new thresholded difference (TD) metric. In the case of compatible images, the TD metric produced a sharper matching peak than the other pixel-by-pixel methods investigated. For non-compatible images representing different spectral bands, times of day and scale, the TD metric, with a suitable threshold, gave results which were generally as good as or better than the other metrics used.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The field of digital image processing has grown considerably during the past decade through increased utilization of imagery in many applications coupled with improvements in size, speed, and cost effectiveness of digital computers and related digital signal processing hardware. A digital image is a spatial function f(x,y) which has been digitized both in spatial coordinates and in amplitude (intensity or gray-level). Digitization of the spatial coordinates (x,y) is accomplished by image sampling, while amplitude digitization is the result of gray-level quantization. One may consider a digital image as a matrix whose row and column indices identify a point in the image while the corresponding matrix element value specifies the gray-level at that point. Each element of the digital array is called a pixel (picture element) [1].

In many image processing applications it is necessary to form a pixel-by-pixel comparison of two images of the same object field obtained from different sensors, or taken from the same sensor at different times. Image registration (or scene matching) refers to the process of locating or matching a region of an image with a corresponding region of another view of the scene, perhaps taken with a different sensor. Applications of image registration include change detection in aerial or satellite pictures [2], recognition of objects [3], and vehicle guidance [4,5]. For example, a number of missile systems being developed by the military utilize precision tracking systems based on correlating live imagery from an on-board sensor with reference imagery stored in a computer [6].

Registration of two images requires a quantitative measure of their similarity in order to determine their match location. The method most widely used for similarity measurement is correlation. In fact, the similarity detection problem itself is generally called "correlation." In this report, however, the distinction between the mathematical correlation and the more general class of similarity detection problems will be maintained. To include methods such as the mean absolute difference (MAD) algorithm and the thresholded difference (TD) algorithm, it is more accurate to refer to the methods for image registration as registration metrics.

A classical technique for registering a pair of functions is to form a correlation measure between the functions and determine the location of the maximum correlation. If $f(x,y)$ and $g(x,y)$ are functions of continuous variables, their cross-correlation is defined as

$$C(\alpha,\beta) = \iint_{-\infty}^{\infty} f(x,y)g(x+\alpha,y+\beta)dxdy \quad . \tag{1-1}$$

If both functions are sampled and held, then Equation (1-1) can be approximated by the following discrete equation:

$$C(p,q) = \Delta x \Delta y \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m,n)g(m+p,n+q) \quad , \tag{1-2}$$

where $f(m,n) = f(m\Delta x,n\Delta y)$,

$g(m+p,n+q) = g[(m+p)\Delta x,(n+q)\Delta y]$,

$\Delta x,\Delta y$ are sampling intervals.

Because practical images are of finite extent, the double summation over m and n in Equation (1-2) takes on finite limits. Since the objective is to find the maximum of C(p,q), multiplication by $\Delta x\Delta y$ is a constant scaling factor and can be eliminated. Thus Equation (1-2) becomes

$$C(p,q) = \sum_{m=1}^{K} \sum_{n=1}^{L} f(m,n)g(m+p,n+q) \quad ; \tag{1-3}$$

$p = 0,1,2,\ldots,M-K$; $q = 0,1,2,\ldots,N-L$.

Equation (1-3) is the correlation estimate for the reference image array size K×L and the live sensor image array size M×N . Since the cross-correlation can exhibit false peaks (that is, the maximum value may occur at a point other than the registration point), it is desirable to use algorithms which involve a normalization to reduce the probability of finding false peaks. Hence the normalized equation can be written as

$$C(p,q) = \frac{\sum_{m=1}^{K} \sum_{n=1}^{L} f(m,n)g(m+p,n+q)}{\left[\sum_{m=1}^{K} \sum_{n=1}^{L} f^2(m,n)\right]^{1/2} \left[\sum_{m=1}^{K} \sum_{n=1}^{L} g^2(m+p,n+q)\right]^{1/2}} . \qquad (1\text{-}4)$$

Algorithms which normalize the correlation function in some manner typically reduce the occurrence of false peaks at the expense of increased processing complexity and computation time. In order to reduce the computational burden for the correlation process, it may be desirable to quantize the amplitude into fewer gray-levels—for example, only two or three gray-levels.

Our purpose is to seek registration metrics which minimize error and uncertainty in the estimate and which are also efficient computationally.

A number of registration metrics have been described in the literature [6,7,8,9,10], including the direct cross-correlation method, fast Fourier transform method, phase correlation method, and sequential similarity detection algorithm. These methods have similarities and differences which affect computational burden and expected performance. A comparison of different methods must take into consideration the degree of false peak rejection and the processing complexity required for a normalization scheme.

It can be verified that the height and sharpness of the correlation peak obtained when two images are correctly registered is dependent on the contents of the two images. Thus we develop a new registration metric called the thresholded difference metric which is derived from an attempt to utilize the joint probability density function for the two images in the registration process. The thresholded difference metric has the potential for improving the matching accuracy with less sensitivity to scene content.

In practice, the reference and live sensor images of the same scene obtained by non-compatible sensors will generally exhibit substantial differences. These differences can be classified as [11]:

1. Intensity difference—for example, the intensity reversals exhibited by thermal images taken at different times of day.
2. Structural difference—the common object-scene may have changed in structural detail in the

time between the taking of the two images.

3. Geometrical difference—the two images may have been taken from different sensor positions and directions, giving rise to differences in rotation, aspect, and scale.

Even though the reference and live sensor images may both be of a common object-scene, these differences of intensity, structure, and geometry will often be sufficient to severely attenuate or eliminate the matching peak in the correlation function. In order for registration to be accomplished with non-compatible images, it is necessary to implement some preprocessing algorithms which bring the two images into compatibility before registration metrics can be applied. Potential preprocessing algorithms include edge detection [12,13], geometric transformation [14], and intensity transformation [14].

To test the preprocessing algorithms and registration metrics, various practical imagery was studied and compared. Real-world imagery taken from day-TV and 8-12 μ passive IR sensors was obtained in digitized form on magnetic tapes received from the U. S. Army Missile Command and Wright-Patterson Air Force Base. These images were used in simulations to evaluate quantization methods, preprocessing algorithms, and different registration metrics.

CHAPTER II

THEORY

In this chapter four registration metrics are defined, including the direct cross-correlation (DCC), fast Fourier transform (FFT), mean absolute difference (MAD), and thresholded difference (TD) metrics. These registration metrics are based on similar fundamentals. A search procedure is performed to find the relative translation between two images, whereby a suitably defined similarity measure is an extremum. The shift at which an extremum occurs is designated as the registration position. In the latter portions of this chapter, the quantization methods, the edge-detection algorithms, and registration statistical analysis are discussed.

Direct Cross-Correlation (DCC) Metric

The direct cross-correlation metric is defined as

$$C(p,q) = \frac{\sum_{m=1}^{K} \sum_{n=1}^{L} R(m,n)S(m+p,n+q)}{\left[\sum_{m=1}^{K} \sum_{n=1}^{L} R^2(m,n)\right]^{1/2} \left[\sum_{m=1}^{K} \sum_{n=1}^{L} S^2(m+p,n+q)\right]^{1/2}} \qquad (2\text{-}1)$$

for $p = 0,1,2,\ldots,M-K$; $q = 0,1,2,\ldots,N-L$.

Figure 1 illustrates the relationship between the reference image R of size K×L and the sensor image S of size M×N . The properties and performance of the DCC metric can be described as follows:

1. To determine the location of maximum correlation, C(p,q) must be computed at each shift (p,q). A decision about registration can only be made after the correlation array C(p,q) is computed for all (M-K+1)(N-L+1) translations of the reference image within the sensor image. A large number of computations are needed since the reference and sensor image areas are usually large.
2. The normalization factor in the denominator in Equation (2-1) is needed to reduce the occurrence of false peaks in C(p,q). For the normalized correlation estimate, false peaks can occur because of noise but not because of different amplitude statistics for the two image signals.
3. Computational complexity can be reduced considerably by using fewer quantization levels to represent the two images.

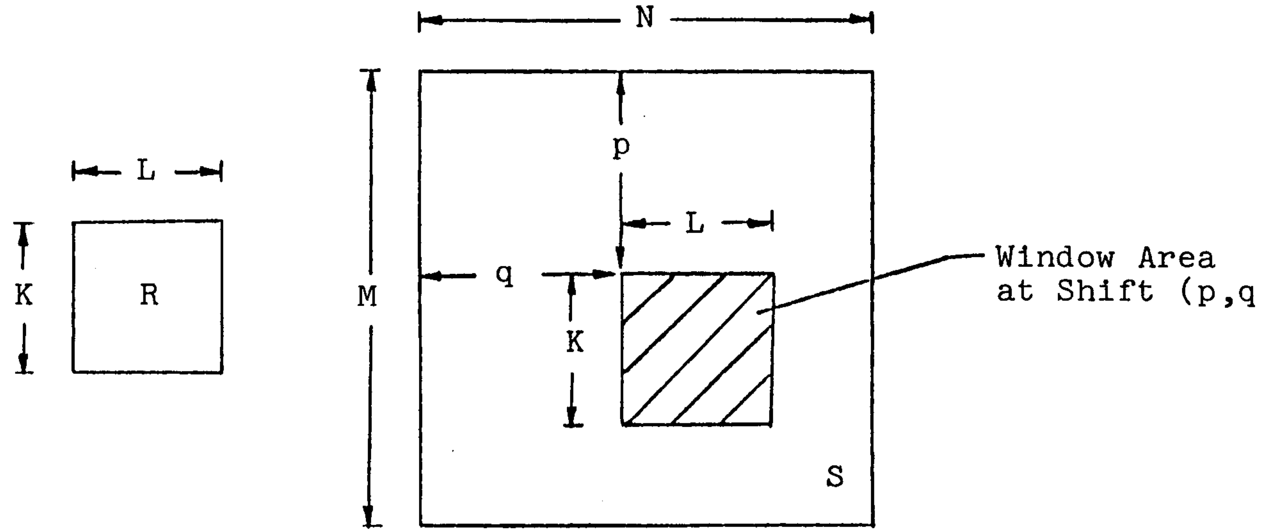


Figure 1. Relationship between the Reference Image R(K,L) and the Sensor Image S(M,N).

Fast Fourier Transform (FFT) Metric

Fast computation of the discrete correlation function is possible using FFT algorithms [15]. For the one-dimensional continuous case, correlation is implemented from Fourier transform by

$$\phi(k) = \int_{-\infty}^{\infty} x(t)y(t+k)dt = F^{-1}\{\Phi(\omega)\} = F^{-1}\{X(\omega)Y^{*}(\omega)\} \; . \quad (2\text{-}2)$$

For the two-dimensional discrete case, we have the analogous relationship

$$C(p,q) = \sum_{m=1}^{K} \sum_{n=1}^{L} R(m,n)S(m+p,n+q) = \mathrm{IFFT}\{\mathrm{FFT}[R]\mathrm{FFT}^{*}[S]\} \quad , \quad (2\text{-}3)$$

where IFFT denotes the inverse FFT and * denotes the complex conjugate. Zeros should be added to R so that both R and S have size M by N . The algorithm is most efficient if M and N are powers of 2 . The properties and performance of the FFT metric can be described as follows:

1. Relatively fast with FFT hardware [16].
2. This algorithm must also be normalized to reduce false peak occurrence. Therefore Equation (2-3) yields

$$C(p,q) = \frac{\mathrm{IFFT}\{\mathrm{FFT}[R]\mathrm{FFT}^*[S]\}}{\left[\sum_{m=1}^{K}\sum_{n=1}^{L} R^2(m,n)\right]^{1/2}\left[\sum_{m=1}^{K}\sum_{n=1}^{L} S^2(m+p,n+q)\right]^{1/2}} \quad . \tag{2-4}$$

3. A particularly interesting registration metric is obtained if the normalization is done in the spatial frequency domain. Then

$$C(p,q) = \mathrm{IFFT}\left\{\frac{\mathrm{FFT}[R]\mathrm{FFT}^*[S]}{|\mathrm{FFT}[R]\mathrm{FFT}^*[S]|}\right\} \quad . \tag{2-5}$$

The term inside the larger brackets is the phase component of the cross power spectral density between the two images. For this reason the metric in Equation (2-5) has been referred to as a phase correlation metric [8]. The phase correlation metric improves resolution and false peak rejection. Unfortunately, this improved performance is achieved at the cost of increased processing complexity.

Mean Absolute Difference (MAD) Metric

The mean absolute difference metric is derived from the sequential similarity detection algorithm (SSDA) [9] which increases processing speed by computing a function based on the difference of data point amplitudes instead of their product. Since subtraction is a faster process than multiplication, a time saving is realized. For the reference image array R(K,L) and the sensor image array S(M,N), the MAD metric is defined as

$$M(p,q) = \sum_{m=1}^{K} \sum_{n=1}^{L} |R(m,n) - S(m+p,n+q)| \qquad (2\text{-}6)$$

for $p = 0,1,2,\ldots,M-K$; $q = 0,1,2,\ldots,N-L$.

The performance of the MAD metric is given as follows:

1. Ideally, at registration M(p,q) equals zero. If R and S both contain a random additive noise, then at registration M(p,q) is a minimum.
2. A false minimum can occur for M(p,q) if the two images have different average intensities at registration. Thus, this algorithm needs to be normalized by subtracting the array average intensity value from the value at each pixel. Hence Equation (2-6) yields

$$M(p,q) = \sum_{m=1}^{K} \sum_{n=1}^{L} \left| R(m,n) - \hat{\mu}_R - S(m+p,n+q) + \hat{\mu}_s(p,q) \right| \quad , \tag{2-7}$$

$$\hat{\mu}_R = \frac{1}{KL} \sum_{m=1}^{K} \sum_{n=1}^{L} R(m,n) \quad , \tag{2-8}$$

$$\hat{\mu}_s(p,q) = \frac{1}{KL} \sum_{m=1}^{K} \sum_{n=1}^{L} S(m+p,n+q) \quad . \tag{2-9}$$

3. A potential advantage is derived from the fact that the MAD metric exhibits a minimum at registration. Thus the metric must be below a given threshold for a decision of registration and the computation can be halted any time the threshold is exceeded. Therefore a more computationally efficient algorithm is defined as

$$I(p,q) = \left\{ r \,\middle|\, r = \text{minimum}[(m-1)L+n], \text{ and} \right.$$
$$\left. \varepsilon(p,q) = \sum_m \sum_n \left| R(m,n) - \hat{\mu}_R - S(m+p,n+q) + \hat{\mu}_s(p,q) \right| \geq T \right\} \quad , \tag{2-10}$$

where T is a predetermined threshold. At registration $I(p,q)$ is a maximum. From Equation (2-10), if the accumulative error function $\varepsilon(p,q)$ exceeds threshold T before all $K \times L$ points are examined, then it is assumed that the test has failed for shift (p,q), and a new shift is checked. In this algorithm only a relatively few number of computations are necessary for shifts that correspond to a gross misregistration, so it saves much computation time.

Thresholded Difference (TD) Metric

The thresholded difference metric [17] is derived from an attempt to utilize the joint probability density function between the two images with shift in two dimensions (p,q) as parameters. The TD metric which is adaptive to scene content has the potential for improving correlation accuracy. It is defined as

$$TD(p,q) = \sum_{m=1}^{K} \sum_{n=1}^{L} g\left\{ \left| R(m,n) - S(m+p,n+q) \right| \right\} \quad (2\text{-}11)$$

for $p = 0,1,2,\ldots,M-K$; $q = 0,1,2,\ldots,N-L$,

where $g\{\cdot\}$ is a threshold function,

$$g\{\cdot\} = 1 \text{ if } |R(m,n) - S(m+p,n+q)| \leq T ,$$

$$= 0 \text{ otherwise.}$$

T is a selected threshold.

A model for the TD metric is shown in Figure 2.

The properties and performance of the TD metric are given as follows:

1. TD(p,q) ranges from 0 to KL . At registration TD(p,q) is a maximum. Ideally, a value of KL corresponds to a perfect match between R and S at registration and is achievable with T = 0 .
2. Registration properties of two images are completely characterized by the joint probability density function (pdf) between the two images with shift in two dimensions (p,q) as parameters. An estimate for the joint pdf between two images is given by their joint histogram. For 8-bit images the joint histogram has a domain of 256×256 values and presents a relatively complicated statistical description of the joint properties of the two images. Additionally, a complete histogram surface is generated for each relative spatial shift (p,q) between the two images. Thus it appears that characterizing two images by their joint histogram is a formidable task which produces too much information to be usable. Strictly speaking, this is true; however, there is useful

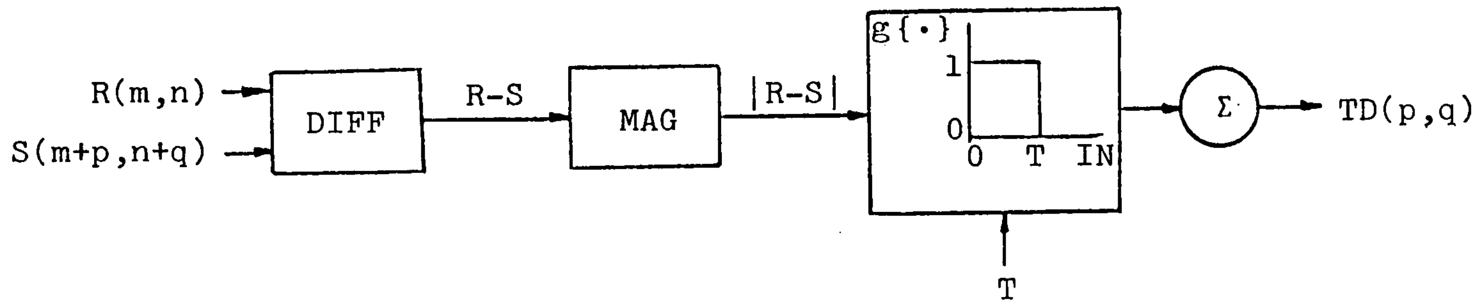


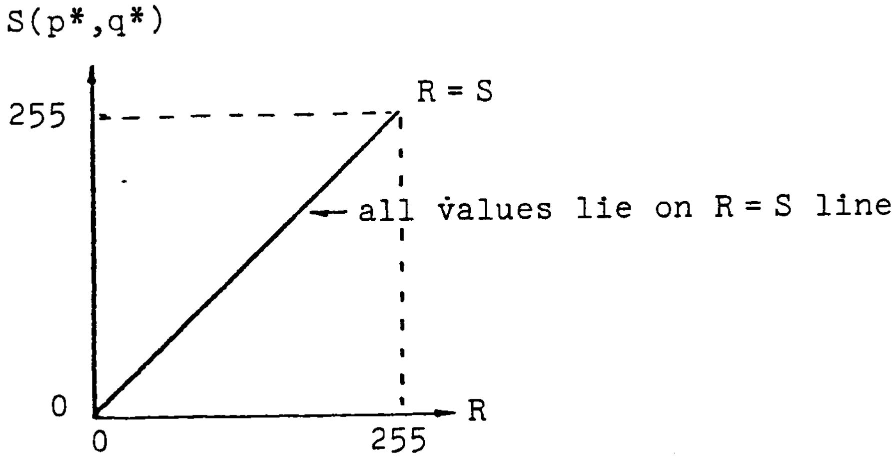
Figure 2. Implementation of a TD Metric Correlator.

information and enhanced understanding to be achieved from examining the joint histogram. For images R and S with eight-bit pixel values and size $K \times L$, the joint histogram for shift (p,q) is given by

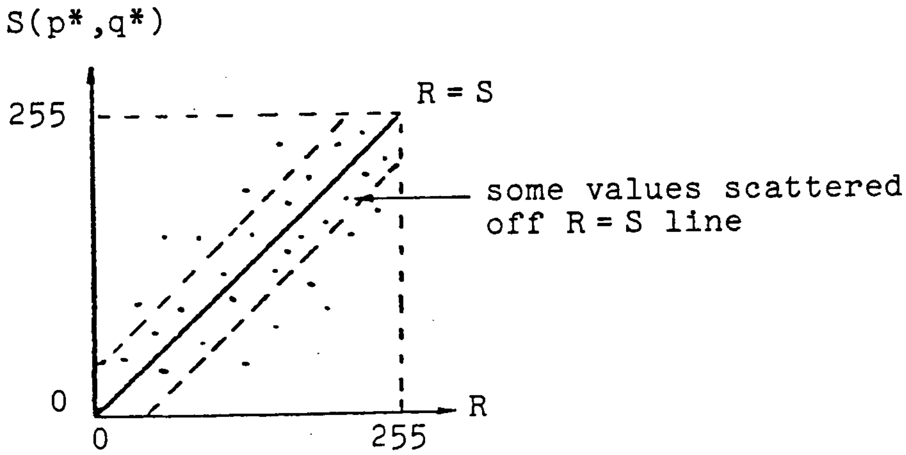
$$HIS(a,b,p,q) = \frac{n(R_a,S_b)}{KL} \quad , \tag{2-12}$$

$$a,b = 0,255 \quad ,$$

where $n(R_a,S_b)$ = number of pixel pairs wherein the amplitude in R was a , while the amplitude in S was b . At registration position (p^*,q^*) , ideally, all pixels match exactly and all points in the joint histogram will be along the R = S line as shown in Figure 3(a). Realistically, at registration some pixels will not match due to uncorrelated noise in the images or due to geometric distortions. These mismatched pixels are represented as points scattered off the R = S line in Figure 3(b). The scattering distance can be related to the signal-to-noise ratio or percent distortion.



(a) At Registration, Ideal Case



(b) At Registration, Practical Case

Figure 3. Joint Histogram Properties.

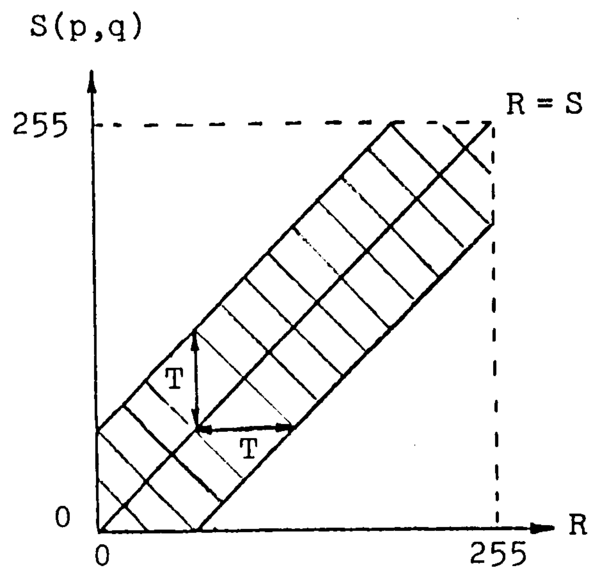
3. In terms of the joint histogram between two $K \times L$ images, the relationship between TD(p,q) and the joint histogram is shown in Figure 4. The relation can also be shown as

$$TD(p,q) = \sum_{m=1}^{K} \sum_{n=1}^{L} g(|R(m,n) - S(m+p,n+q)|)$$

$$= KL \times \{\text{Summation of values of joint histogram in the shaded region in Figure 4}\} \quad . \qquad (2\text{-}13)$$

The sharpest peak for the TD metric is achieved for a threshold $T = 0$; however, this peak is reduced in magnitude if the images are not perfect replicas at registration. T can be increased to allow for small mismatches between pixels. This restores peak height at registration but broadens the peak.

4. Tradeoff analyses are needed to show interrelationships between peak height, peak broadness and noise, threshold value, scale errors and scene content.



TD(p,q) = KL × {summation of all values in the shaded region}

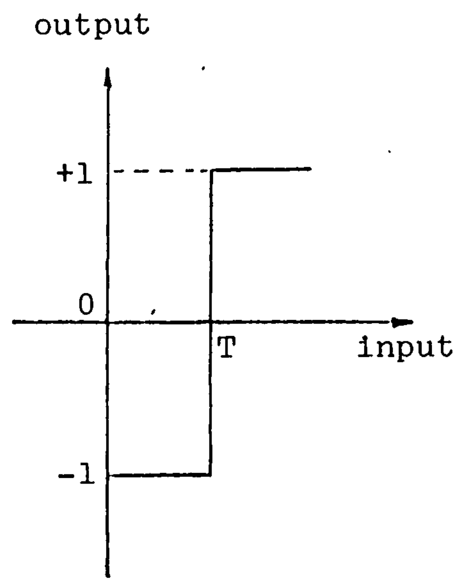
T is a selected threshold

Figure 4. Relationship between the Joint Histogram and TD(p,q).

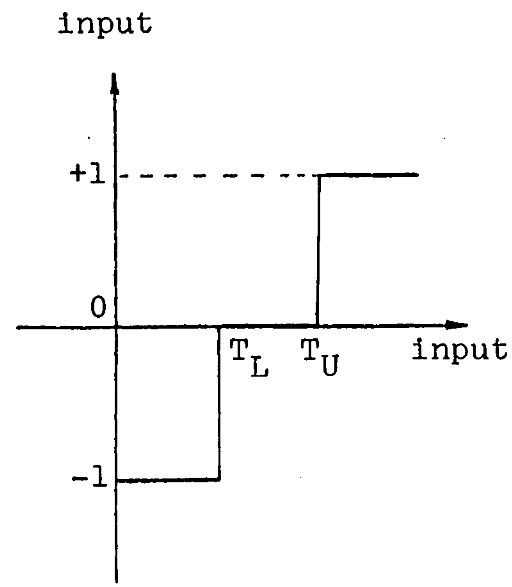
Quantization

Computation for all the registration metrics is time-consuming for eight-bit images. Algorithms which sacrifice gray-scale for increased processing speed are reported in the literature [18,19,20]. For two or three quantization levels, the DCC and MAD metrics can be performed by hardware implementations which exhibit high computational efficiency. In fact, if an extreme coarse quantization (two-level or one-bit) is used, then multiplication and addition are reduced to simple logic functions. The DCC and MAD metrics are then computationally equivalent, the only difference being in the specific logic operation required for implementation.

The input-output relationship for the binary and tri-level quantizers is shown in Figure 5. For the binary quantizer shown in Figure 5(a), the weight of the output is 1 if the input is above a given threshold T , and is -1 if the input is below that threshold. For the tri-level quantizer in Figure 5(b), the weight of the output is +1 if the input is above a given upper threshold T_U , is -1 if the input is below a specified lower threshold T_L , and is 0 for the input between the lower and upper thresholds. Thresholds are typically chosen in terms of such statistical parameters as the mean, median, and standard deviation of the image.



(a) Binary Quantizer



(b) Tri-Level Quantizer

Figure 5. Binary and Tri-Level Quantizers.

These values can be either global or adaptive in a local window.

The direct cross-correlation (DCC) metric was given in Equation (2-1). Now if R and S are subjected to a 1-bit (2-level) quantization so that the only possible values for R and S are ±1 , then the denominator of Equation (2-1) is simply KL . Thus for the binary correlator a normalized algorithm is given by

$$C(p,q) = \frac{1}{KL} \sum_{m=1}^{K} \sum_{n=1}^{L} g_1[R(m,n)]g_2[S(m+p,n+q)] \quad , \qquad (2\text{-}14)$$

$$\hat{C}(p,q) = \sum_{m=1}^{K} \sum_{n=1}^{L} g_1[R(m,n)]g_2[S(m+p,n+q)] \quad , \qquad (2\text{-}15)$$

where the functions g_1 and g_2 represent the input-output characteristics for the two-level quantization for R and S respectively.
A 1-bit correlator which quantizes R and S to "0" and "1" is equivalent statistically to one which quantizes to ±1 and is easily implemented using an "EXCLUSIVE OR" logic element as shown in Figure 6 [20].

The degradation in the correlation due to quantizing to only 2 levels shows up as a $\pi^2/4$ increase in the estimate variance as

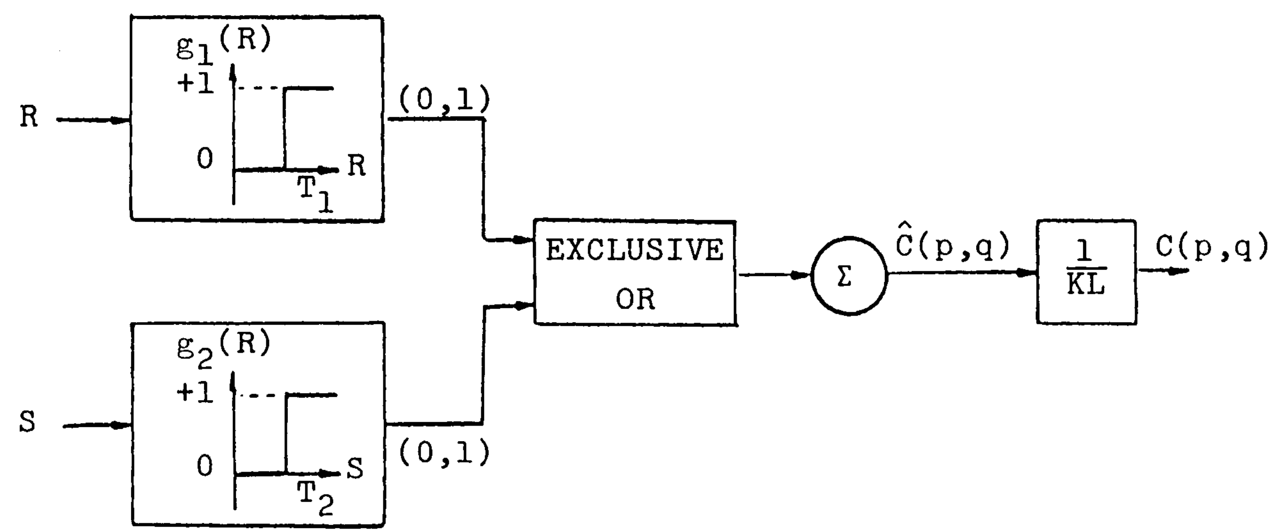


Figure 6. Implementation of a Binary Correlator.

compared with a multi-bit (infinite resolution) quantizer [21]. The statistical disadvantage incurred by coarse quantization of the images is offset by being able to include more pixels in the correlation computation for a real-time requirement assuming that geometric distortion is not a problem. Thus for an application which requires real-time correlation, quantization of the data to two levels is an attractive technique.

Edge-Detection Algorithms

In practice, the two images (reference and sensor) of the same scene obtained by non-compatible sensors will generally exhibit substantial differences. For this reason it is necessary to use preprocessing methods to bring the two images into compatibility before registration metric can be applied. A preprocessing method that should be considered is the edge-detection algorithm. It has the potential to pull out only the edge features of the images, thereby reducing the sensitivity of the correlator to the contrast reversals in the images. For example, an edge-extraction operation performed on thermal images will extract the invariant edges between a building and its surroundings, irrespective of the variations in magnitude and sign of the temperature difference between the two.

A considerable amount of research has been done in extracting edges from images using digital computers [13]. This section explains 2 edge detection algorithms that could be implemented as a preprocessing step in the registration system: the 2×2 edge detector and the 3×3 edge detectors [19].

1. 2×2 edge detector: The method of edge detection presented here is known as the Roberts absolute estimate of the gradient operator [1]. Figure 7 is the pixel representation of a 2×2 area in the digital picture. If the digital picture is represented by the two-dimensional function $g(x,y)$, then the magnitude of the gradient at pixel (i,j) can be given by

$$F(i,j) = \left\{[g(i,j) - g(i+1,j+1)]^2 + [g(i,j+1) - g(i+1,j)]^2\right\}^{1/2} . \tag{2-16}$$

From Equation (2-16) it can be seen that in picture areas of constant gray-level, $F(i,j)$ will be zero; and in picture areas of high gray-level change in either x or y or both directions,

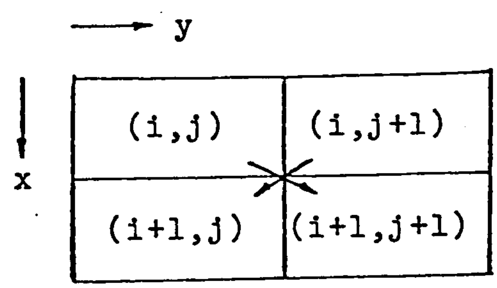


Figure 7. Pixel Representation for the Roberts Cross Operator.

F(i,j) will be large. A more computationally efficient algorithm is given by

$$F(i,j) = |g(i,j) - g(i+1,j+1)| + |g(i,j+1) - g(i+1,j)| \quad . \qquad (2\text{-}17)$$

To implement binary quantization using the gradient properties of Equation (2-17) requires the selection of a gradient threshold for values of F(i,j). The quantization process is given in Equation (2-18).

$$Fq(i,j) = \begin{cases} 1 & ; \quad F(i,j) \geq GTH \\ 0 & ; \quad F(i,j) < GTH \end{cases} \quad , \qquad (2\text{-}18)$$

where $Fq(\cdot,\cdot)$ is the quantized gradient array and GTH is the gradient threshold. The application of Equations (2-17) and (2-18) to the original picture gives an outline or edge drawing of objects in the picture.

2. 3×3 edge detectors: The estimation of the gradient magnitude at a particular pixel can be extended to include the effects of the eight nearest pixels surrounding the pixel of interest. Figure 8 illustrates a 3×3 pixel area

$\longrightarrow y$

$\downarrow x$

(i-1,j-1)	(i-1,j)	(i-1,j+1)
(i,j-1)	(i,j)	(i,j+1)
(i+1,j-1)	(i+1,j)	(i+1,j+1)

Figure 8. Pixel Representation for the 3×3 Edge-Operator.

used to estimate the gradient magnitude at pixel (i,j). The operation of calculating the gradient magnitude can be defined using an inner product formulation. Assume that the digital picture function is given by g(x,y) and an appropriate weighting function W(x,y) where W is a 3×3 matrix. Then the inner product of g(x,y) with W(x,y) is given by

$$(g,W)_{i,j} = \sum_{k=1}^{3} \sum_{\ell=1}^{3} g(i+k-2,j+\ell-2)W(k,\ell) \quad . \tag{2-19}$$

A number of weighting matrices can be found in the literature [13]. Figure 9 illustrates 3 sets of weighting matrices. By letting $S_x(i,j) = (g,W_1)$ and $S_y(i,j) = (g,W_2)$ the magnitude of the gradient at pixel (i,j) is

$$F(i,j) = \left[S_x^2(i,j) + S_y^2(i,j)\right]^{1/2} \quad . \tag{2-20}$$

A more computationally efficient algorithm is

$$F(i,j) = |S_x(i,j)| + |S_y(i,j)| \quad . \tag{2-21}$$

$$W_1 \qquad\qquad\qquad\qquad W_2$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

(a) Prewitt weighting functions.

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

(b) Sobel weighting functions.

$$\begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix}$$

(c) Isotropic weighting functions.

Figure 9. Weighting Matrices Used for Gradient Estimation.

Registration Statistical Analysis

For the practical case, optical path transmission fluctuations and electronic noise in the sensor combine to make the image random in nature. It is desirable to determine the statistical performance of the registration metrics. The primary measures of performance for any registration metric are mean, variance, probability of false fix, probability of correct registration, and registration accuracy. Calculation of these performance measures without knowledge of the joint probability density function, as dependent on various parameters, is possible for only a few special cases where simplifying assumptions can be made. Based on given statistical models for the signal and noise, these performance measures in terms of selected threshold, quantization, and signal-to-noise ratios are derived in the literature [6,20,22,23,24,25]. For analyzing the thresholded difference (TD) registration metric in statistical terms, we propose the following model. Each image is the sum of a shift-variant signal with an independent shift-invariant noise term. The noise term is further assumed to be zero-mean Gaussian. That is, the images are modeled as

$$\gamma_1(m,n) = \mu_1(m,n) + n_1(m,n) \quad ,$$

$$\gamma_2(m+p,n+q) = \mu_2(m+p,n+q) + n_2(m+p,n+q) \quad , \qquad (2\text{-}22)$$

where $\gamma_1(m,n)$ is the reference image, $\gamma_2(m+p;n+q)$ is the window area of live sensor image at shift (p,q). The probability density function for the noise term is

$$p_{n_i}(n_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i}\, e^{-\frac{n_i^2}{2\sigma_i^2}} \quad ; \quad i = 1,2 \quad . \tag{2-23}$$

The TD registration metric is defined as Equation (2-11). We analyze its statistical properties for two special cases:

Case (1): For compatible sensors at registration in the absence of geometric distortion. Then $\mu_1(m,n) = \mu_2(m+p^*,n+q^*)$ where (p^*,q^*) is the required shift for the 2 images being at registration. Now, we derive $TD(p^*,q^*)$ in terms of the probabilities associated with each pixel pair comparison.

Let

$$z = \gamma_1(m,n) - \gamma_2(m+p^*,n+q^*) = n_1(m,n) - n_2(m+p^*,n+q^*) \quad . \tag{2-24}$$

Then

$$TD(p^*,q^*) = \sum_{m=1}^{K} \sum_{n=1}^{L} g\left\{ |n_1(m,n) - n_2(m+p^*,n+q^*)| \right\}$$

$$= \sum_{m=1}^{K} \sum_{n=1}^{L} g\left\{ |z| \right\} \quad . \tag{2-25}$$

Since n_1 and n_2 are independent shift-invariant zero-mean Gaussian, then

$$E(z) = E[n_1(m,n) - n_2(m+p^*,n+q^*)] = 0 \quad , \tag{2-26}$$

$$\sigma_z{}^2 = \sigma^2_{n_1-n_2} = \sigma^2_{n_1} + \sigma^2_{n_2} \quad , \tag{2-27}$$

$$p_z = p_{n_1-n_2} = p_{n_1} * p_{n_2} \quad , \tag{2-28}$$

where "*" means convolution, p_z is also Gaussian [26].

$$p_z(z) = \frac{1}{\sqrt{2\pi}\ \sigma_z} e^{-\frac{z^2}{2\sigma_z^2}} \quad , \tag{2-29}$$

$$P[|z| \leq T] = \int_{-T}^{T} p_z(z)dz = 2\ \mathrm{erf}(\frac{T}{\sigma_z}) \quad , \tag{2-30}$$

where erf(β) is the error function,

$$\mathrm{erf}(\beta) = \int_0^\beta \frac{1}{\sqrt{2\pi}} e^{(-\lambda^2/2)} d\lambda \quad . \tag{2-31}$$

Thus

$$E\left[TD(p^*,q^*)\right] = \sum_{m=1}^{K} \sum_{n=1}^{L} E\left[g\{|z|\}\right] = 2KL\,\mathrm{erf}\left(\frac{T}{\sigma_z}\right) \quad . \tag{2-32}$$

Case (2): At relative shift (p,q) where μ_1 and μ_2 essentially are statistically independent.

Then

$$\gamma_1(m,n) = \mu_1(m,n) + n_1(m+p,n+q) \quad , \tag{2-33}$$

$$\gamma_2(m+p,n+q) = \mu_2(m+p,n+q) + n_2(m+p,n+q) \quad . \tag{2-34}$$

Let

$$\begin{aligned} z &= \gamma_1(m,n) - \gamma_2(m+p,n+q) \\ &= \mu_1(m,n) + n_1(m,n) - \mu_2(m+p,n+q) - n_2(m+p,n+q) \quad . \end{aligned} \tag{2-35}$$

Since μ_1 , μ_2 , n_1 , n_2 are statistically independent, then

$$E[z] = E[\mu_1(m,n)] - E[\mu_2(m+p,n+q)] \quad , \tag{2-36}$$

$$\sigma_z{}^2 = \sigma^2_{\mu_1} + \sigma^2_{\mu_2} + \sigma^2_{n_1} + \sigma^2_{n_2} \quad , \tag{2-37}$$

$$p_z = p_{\mu_1} * p_{\mu_2} * p_{n_1} * p_{n_2} \quad , \tag{2-38}$$

$$P\left[|z| \le T\right] = \int_{-T}^{T} p_z(z)dz \quad . \tag{2-39}$$

Clearly, for the non-registration case, $P[|z| \le T]$

decreases as σ_z^2 increases.

When the images can be considered as uncorrelated, then the output from the threshold function $g\{\cdot\}$ for a given threshold T is reduced if any of the signals or noise standard deviations are increased.

From the analyses of Case (1) and Case (2), the important factor for the TD metric is to select a suitable threshold T which can make $E[TD(p^*,q^*)] > E[TD(p,q)]$ for any other shift (p,q) . Selection of threshold T depends on signal-to-noise ratio, geometric distortion, different spectral response and scene content. Further analysis is needed to allow more quantitative evaluation of the thresholded difference registration metric.

CHAPTER III

SIMULATION RESULTS AND DISCUSSIONS

This chapter describes the simulation experiments using the available imagery received from the U. S. Army Missile Command and Wright-Patterson Air Force Base. The real-world imagery taken from day-TV and passive IR sensors is available in digitized form on magnetic tapes. These images were used in simulations to evaluate quantization methods, preprocessing algorithms, and registration metrics. Algorithms used in the simulations were described in Chapter II. They include the DCC metric in Equation (2-1), FFT metric in Equation (2-4), MAD metric in Equation (2-7), TD metric in Equation (2-11), and 3×3 Prewitt gradient operator [27] in Equations (2-19), (2-21). Amplitudes corresponding to 1/3 percentiles were chosen as the lower and upper thresholds T_L and T_U in the tri-level quantizer, and the global or locally adaptive mean and median values were chosen as threshold T in the binary quantizer. The simulation programs were coded in FORTRAN and structured using subroutines for ease of algorithm substitution or modification. A summary of the simulations done is given in Figure 10. The effects of noise, scene content, time of day variations, different spectral bands and different

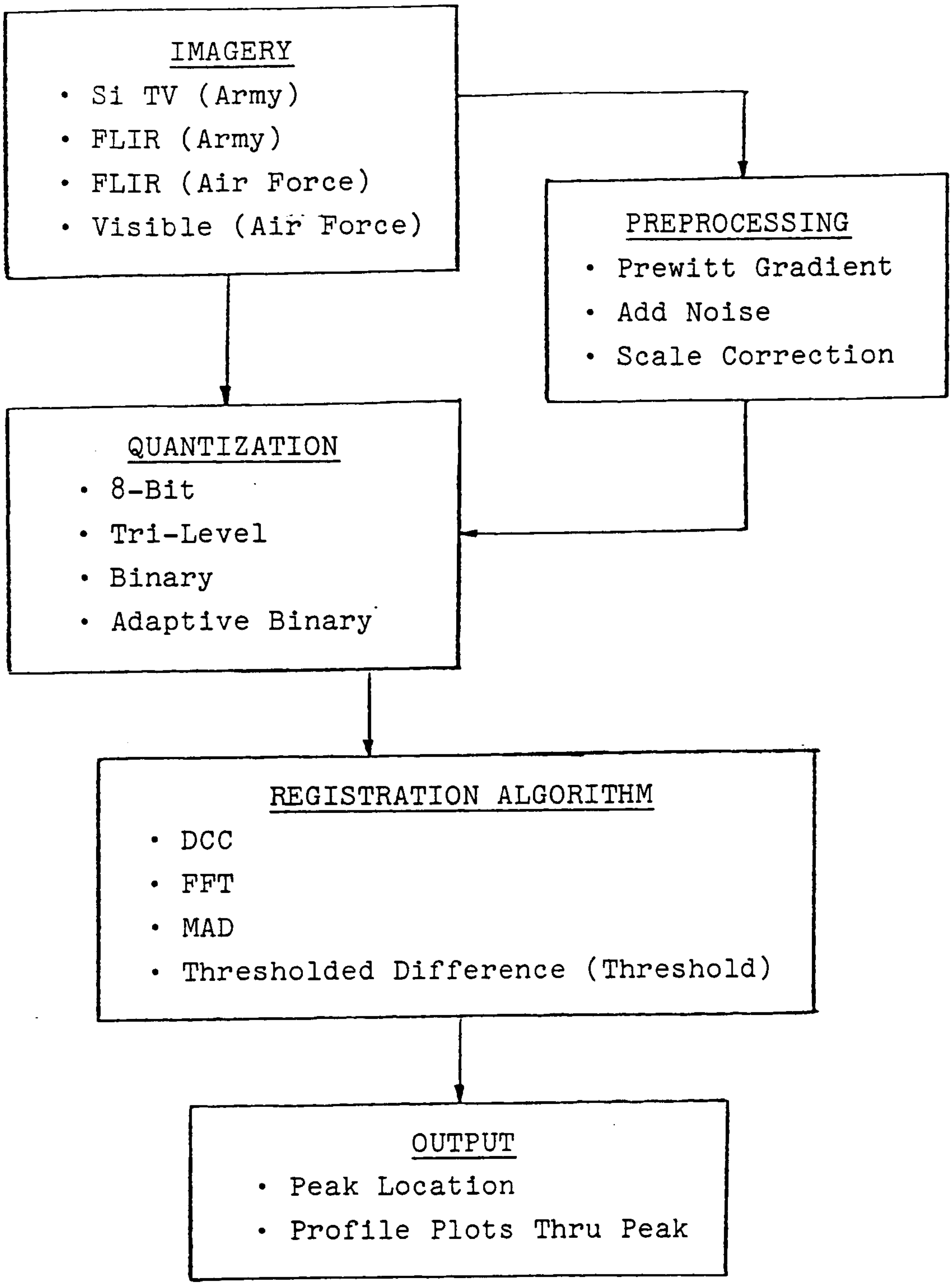


Figure 10. Simulation Options Summary.

scale are illustrated by the simulations. Further, the effects of binary, tri-level and eight-bit quantization as well as gradient preprocessing and an adaptive binary quantizer are shown by selected simulations. The simulation cases and results are discussed below.

Simulation Case 1: Compatible Images, Ideal Case

The imagery used in simulation Case 1 is shown in Figure 11. This image was taken from the wide field-of-view (WFOV) TV sensor and stored in digitized form as File 1 on the magnetic tape received from the U. S. Army Missile Command. It is a 256 gray-level (8-bit) image with array size 407×392. A subimage 127×127 array was segmented from the 407×392 array as the live sensor image S and a 64×64 array from the center of S was taken as the reference image R . This relationship is shown in Figure 12. In this case the reference image R(64,64) is a subset of the live sensor image S(127,127). So this simulation is the compatible ideal case; that is, the two images are taken from the same sensor, are noise-free and have no geometric distortion. In this ideal case, the DCC, FFT, MAD, and TD registration metrics were tested using 256 gray-level, tri-level, and binary images of the original and gradient images. Two different subimages (Scene 1 and Scene 2) were used. Scene 1 was a target with field-tree background and Scene 2 was a

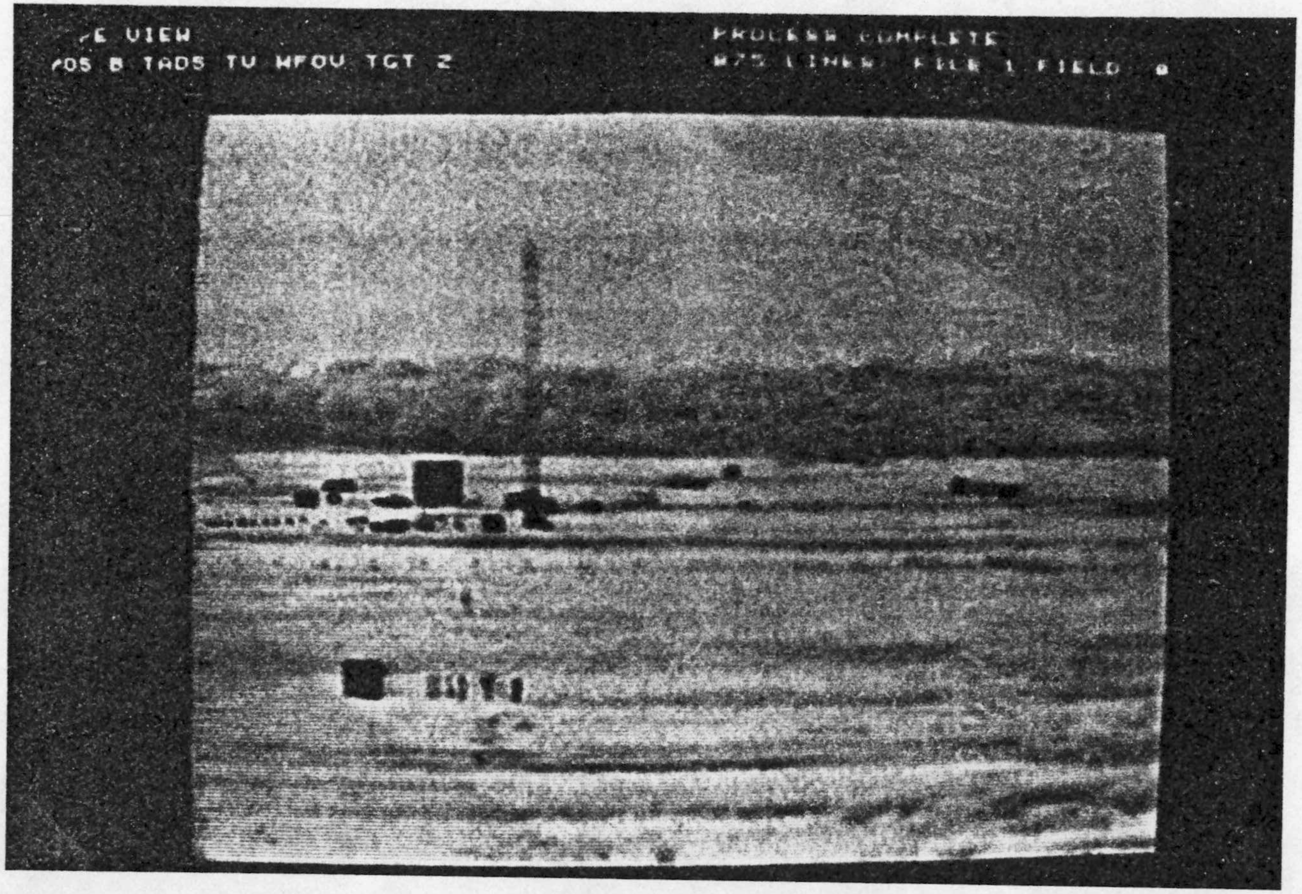


Figure 11. Army File 1 WFOV TV Image.

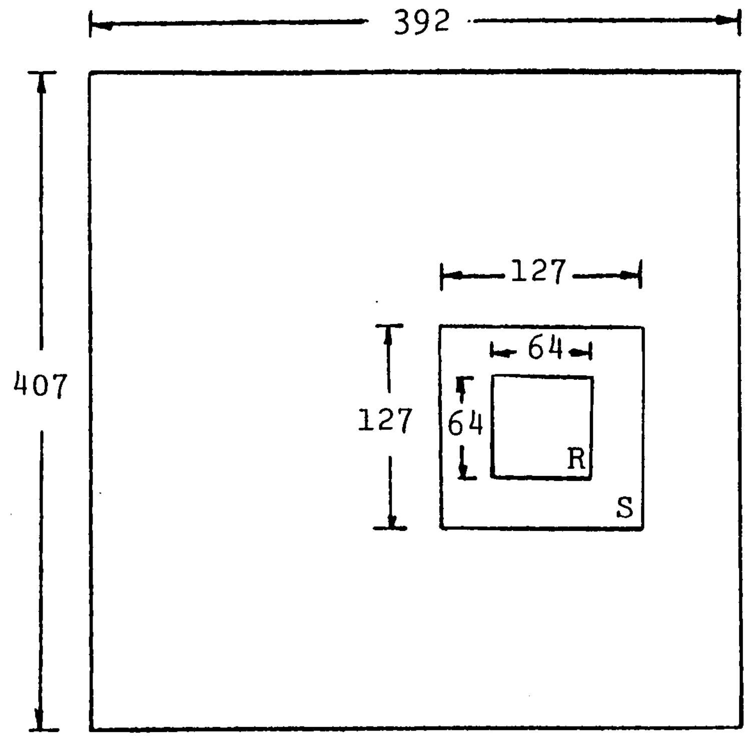
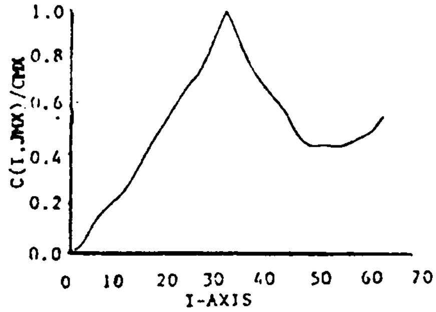


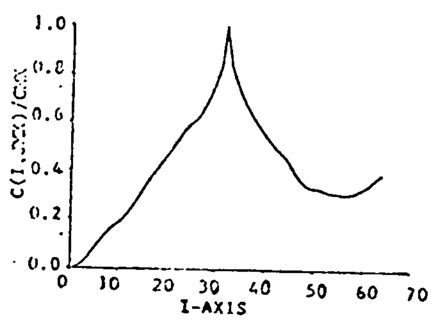
Figure 12. Relationship between R(64,64) and S(127,127) in Simulation Case 1.

target with field background. The simulation results showed that all methods gave the correct peak location (32,32) and peak amplitudes as expected from algorithm definitions for ideal registration. Thus the features for comparison are relative peak sharpness and computer CPU time. The cross-sectional plots through the matching peak in vertical (I-axis) and horizontal (J-axis) directions are shown in Figures 13 and 14. The CPU time required for different algorithms is given in Table 1. Based on simulation results, the following evaluations are made:

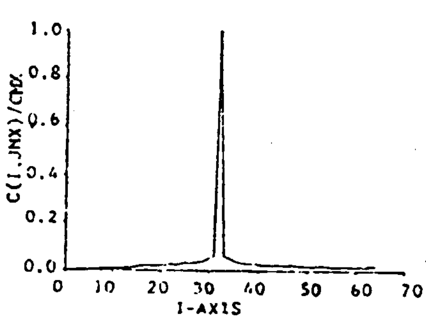
1. Performance of Registration Metrics for 256 Gray-Level Images. Figures 13 and 14 show cross-sectional plots through the matching peak of Scene 1 using 8-bit DCC, MAD, TD ($T = 0$) and FFT metrics. Comparisons of Figure 14 show that the correlation surface of the DCC metric is the same as that of the FFT metric. Comparisons of Figure 13 show that the TD ($T = 0$) metric provided the sharpest peak of any of the metrics. The thresholded difference metric is clearly superior in terms of matching peak sharpness if the threshold, T, can be kept small. It should be emphasized that a threshold of $T = 0$ gives almost perfect



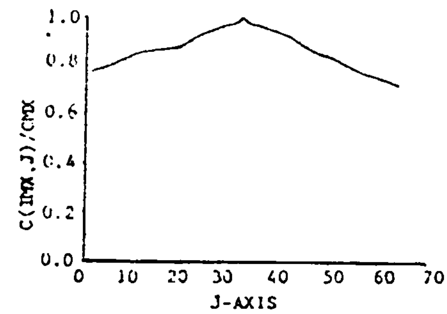
(a) DCC, I-Axis



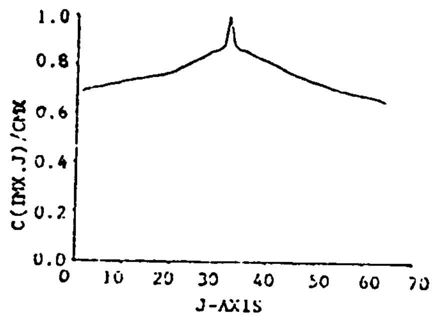
(b) MAD, I-Axis



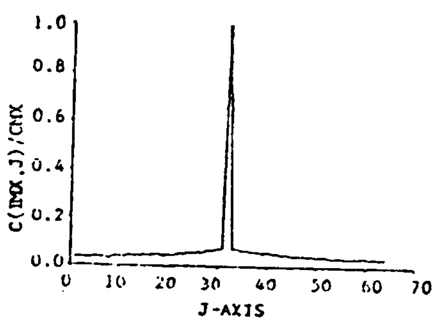
(c) TD(T = 0), I-Axis



(d) DCC, J-Axis

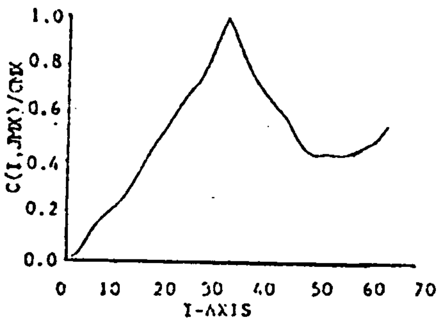


(e) MAD, J-Axis

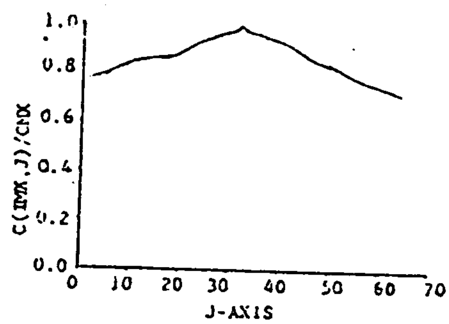


(f) TD(T = 0), J-Axis

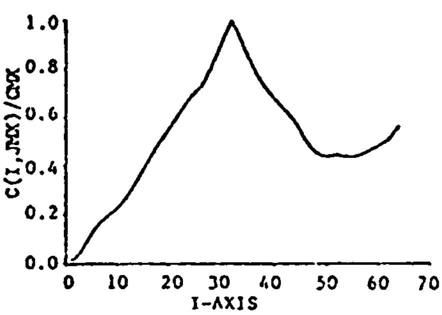
Figure 13. Cross-Sectional Plots of Scene 1 Registrations Using 8-Bit DCC, MAD, TD(T = 0) Metrics.



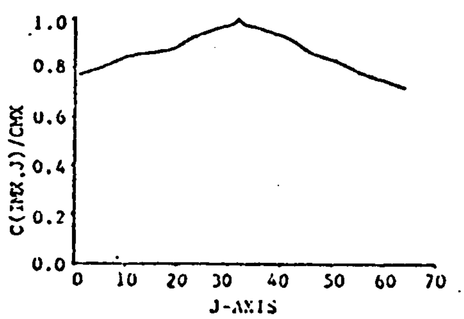
(a) DCC, I-Axis



(c) DCC, J-Axis



(b) FFT, I-Axis



(d) FFT, J-Axis

Figure 14. Cross-Sectional Plots of Scene 1 Registrations Using 8-Bit DCC, FFT Metrics.

TABLE 1

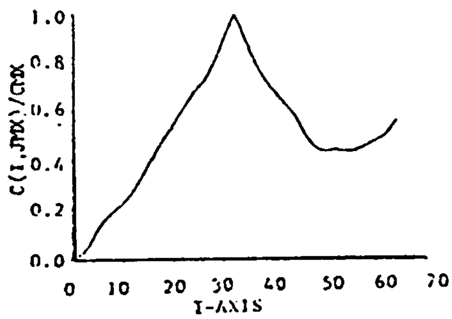
THE COMPUTER CPU TIME FOR SIMULATION CASE 1

ALGORITHM	CPU TIME (MIN : SEC) Scene 1 (Original)	Scene 1 (Gradient)	Scene 2 (Original)	Scene 2 (Gradient)
DCC (8-Bit)	11:05	10:48	11:06	7:11
DCC (Tri-Level)	6:56	7:03	6:59	4:33
DCC (Binary)	6:06	5:57	6:03	3:37
FFT (8-Bit)	4:37	3:27	4:38	3:30
FFT (Tri-Level)	4:29	3:14	4:30	3:16
FFT (Binary)	4:34	3:13	4:35	3:12
MAD (8-Bit)	11:04	12:03	11:57	8:32
MAD (Tri-Level)	11:32	11:19	11:39	7:41
MAD (Binary)	11:28	11:21	11:38	7:43
TD (T = 0)	7:22	7:09	7:18	5:01
TD (T = 2)	7:14	7:09	7:22	4:49
TD (T = 5)	7:12	7:09	4:57	4:50
TD (T = 10)	7:18	7:16	5:06	4:55

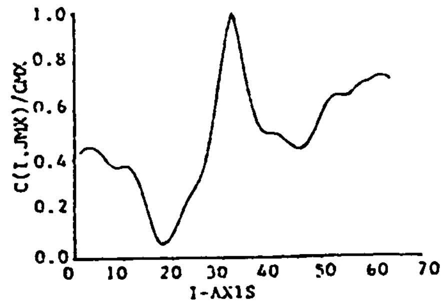
results for the ideal conditions under which the correlation was done (i.e., at registration the images were known to be a perfect match).

2. Scene Dependence. Figure 15 shows cross-sectional plots of Scene 1 and Scene 2 simulations. These plots show that the performance (the sharpness of the peak) of registration metrics was strongly influenced by different scene statistics except for the TD metric. The TD metric has the potential for improving the matching accuracy with less sensitivity to scene content.

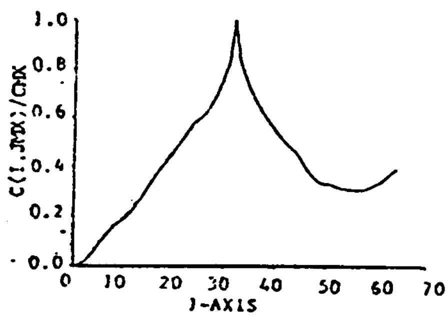
3. Effect of the Prewitt Gradient Algorithm. Gradient preprocessing by application of a 3×3 Prewitt operator to Scenes 1 and 2 was done prior to quantization and correlation. The resultant high-frequency-emphasized images should exhibit sharper correlation peaks. The results are shown as peak profiles in Figure 16 Scene 1, I-axis; Figure 17, J-axis, for various registration metrics. Results without and with gradient preprocessing are shown side-by-side for easy comparison. In every case except for the TD metric, gradient preprocessing did produce a sharper



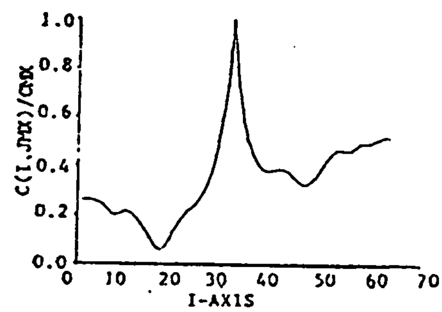
(a) DCC, Scene 1



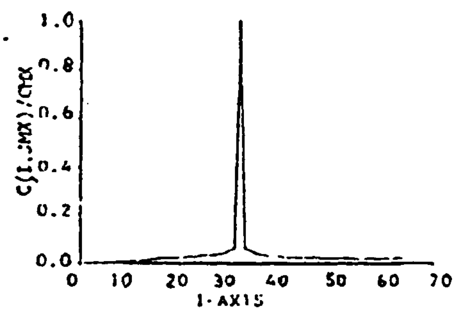
(d) DCC, Scene 2



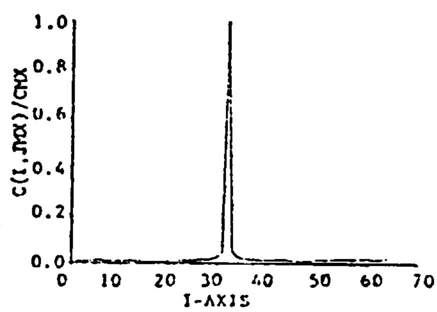
(b) MAD, Scene 1



(e) MAD, Scene 2

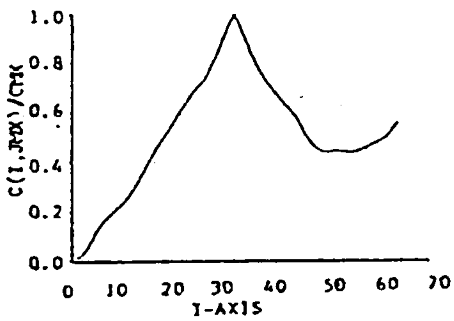


(c) TD(T = 0), Scene 1

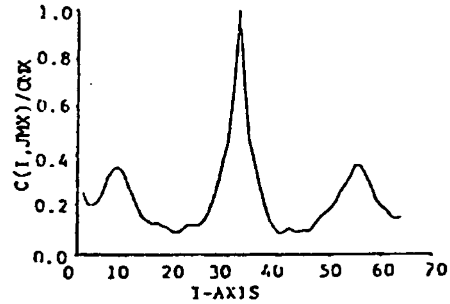


(f) TD(T = 0), Scene 2

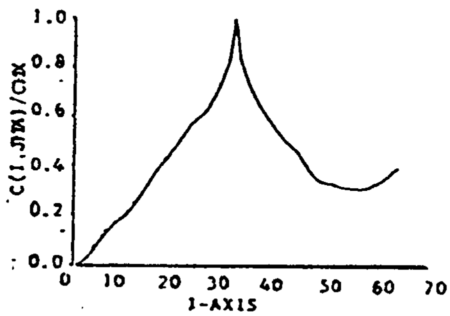
Figure 15. I-Axis Cross-Sectional Plots of Scene 1 and Scene 2 Registrations Using 8-Bit DCC, MAD, TD(T = 0) Metrics.



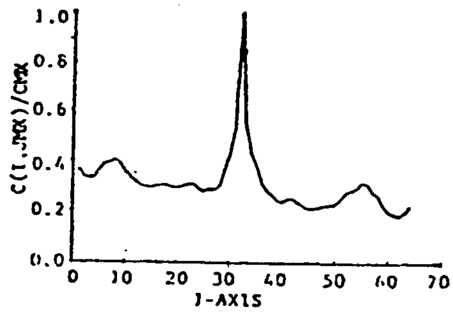
(a) DCC, Original



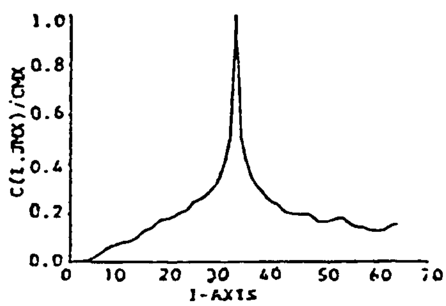
(d) DCC, Gradient



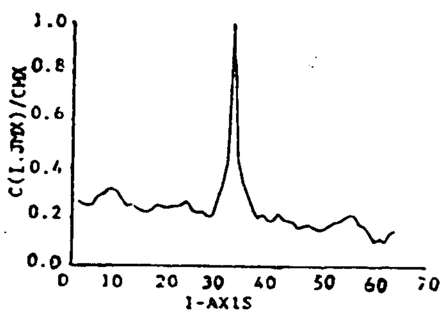
(b) MAD, Original



(e) MAD, Gradient

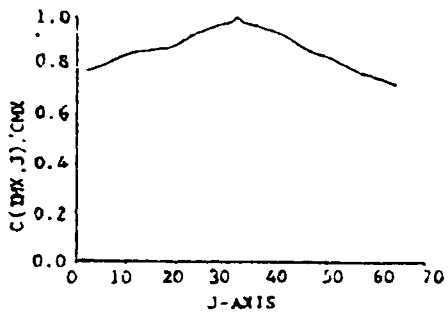


(c) TD(T = 5), Original

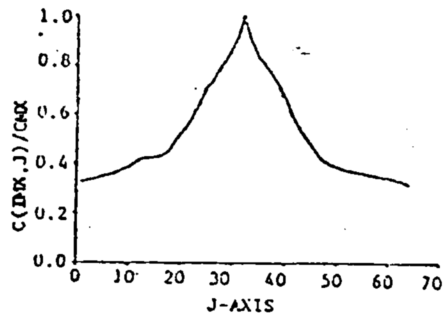


(f) TD(T = 5), Gradient

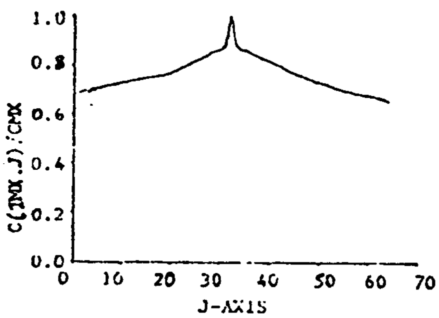
Figure 16. I-Axis Cross-Sectional Plots of Scene 1 Original and Gradient Image Registrations Using 8-Bit DCC, MAD, TD(T = 5) Metrics.



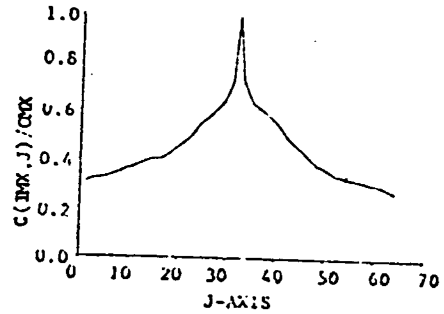
(a) DCC, Original



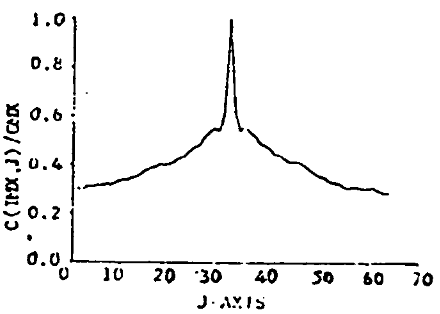
(d) DCC, Gradient



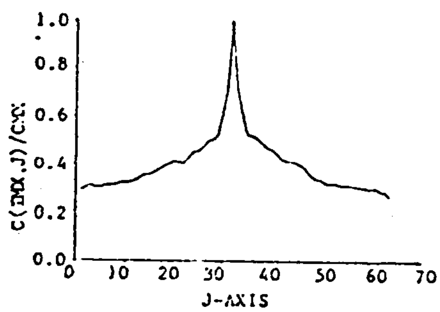
(b) MAD, Original



(e) MAD, Gradient



(c) TD(T = 5), Original



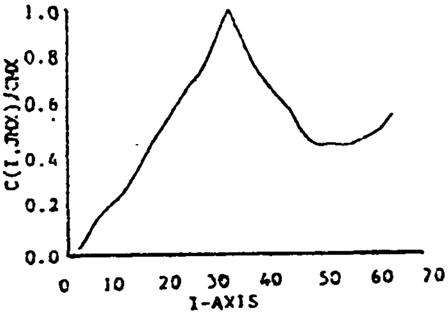
(f) TD(T = 5), Gradient

Figure 17. J-Axis Cross-Sectional Plots of Scene 1 Original and Gradient Image Registrations Using 8-Bit DCC, MAD, TD(T = 5) Metrics.

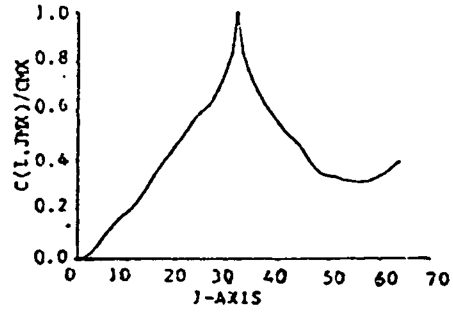
correlation peak. It is postulated that since the TD method exhibits high frequency emphasis characteristics itself, then high frequency emphasis filtering provided by the gradient processor offers no additional gain.

4. Effect of Quantization. Figure 18 shows cross-sectional plots of Scene 1 256 gray-level, tri-level and binary image registrations. Since these tri-level and binary quantizers used global thresholds which tend to reduce high spatial frequency in the imagery, the peaks were broadened in the tri-level and binary correlators.

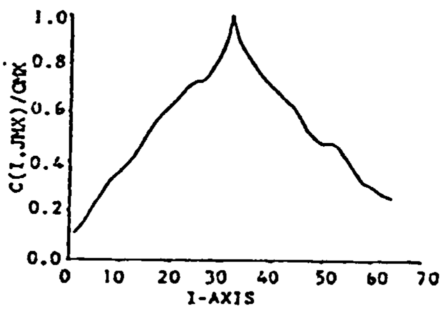
5. Effect of Different Threshold T on TD Metric. Figure 19 shows cross-sectional plots of Scene 1 registration using the TD metric with different threshold values. For the ideal case, results showed that the sharpest peak was achieved for a threshold $T = 0$. An increase in T broadened the peak. If the images are known to be of high quality (low noise, low distortion), it is possible to achieve high correlation accuracy using the TD method with a low value of T .



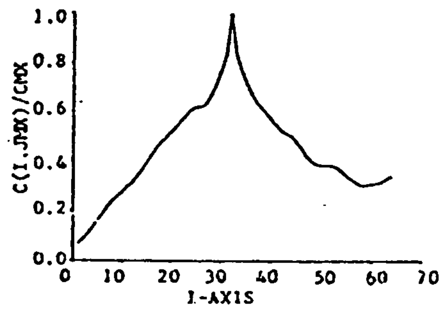
(a) DCC, 8-Bit



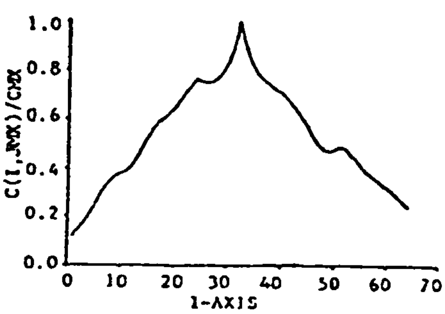
(d) MAD, 8-Bit



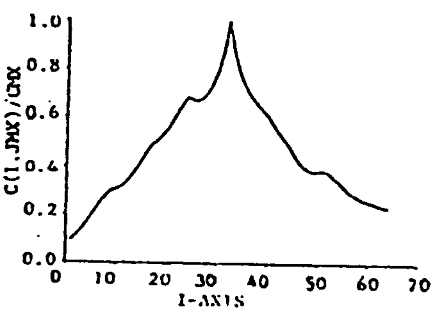
(b) DCC, Tri-Level



(e) MAD, Tri-Level

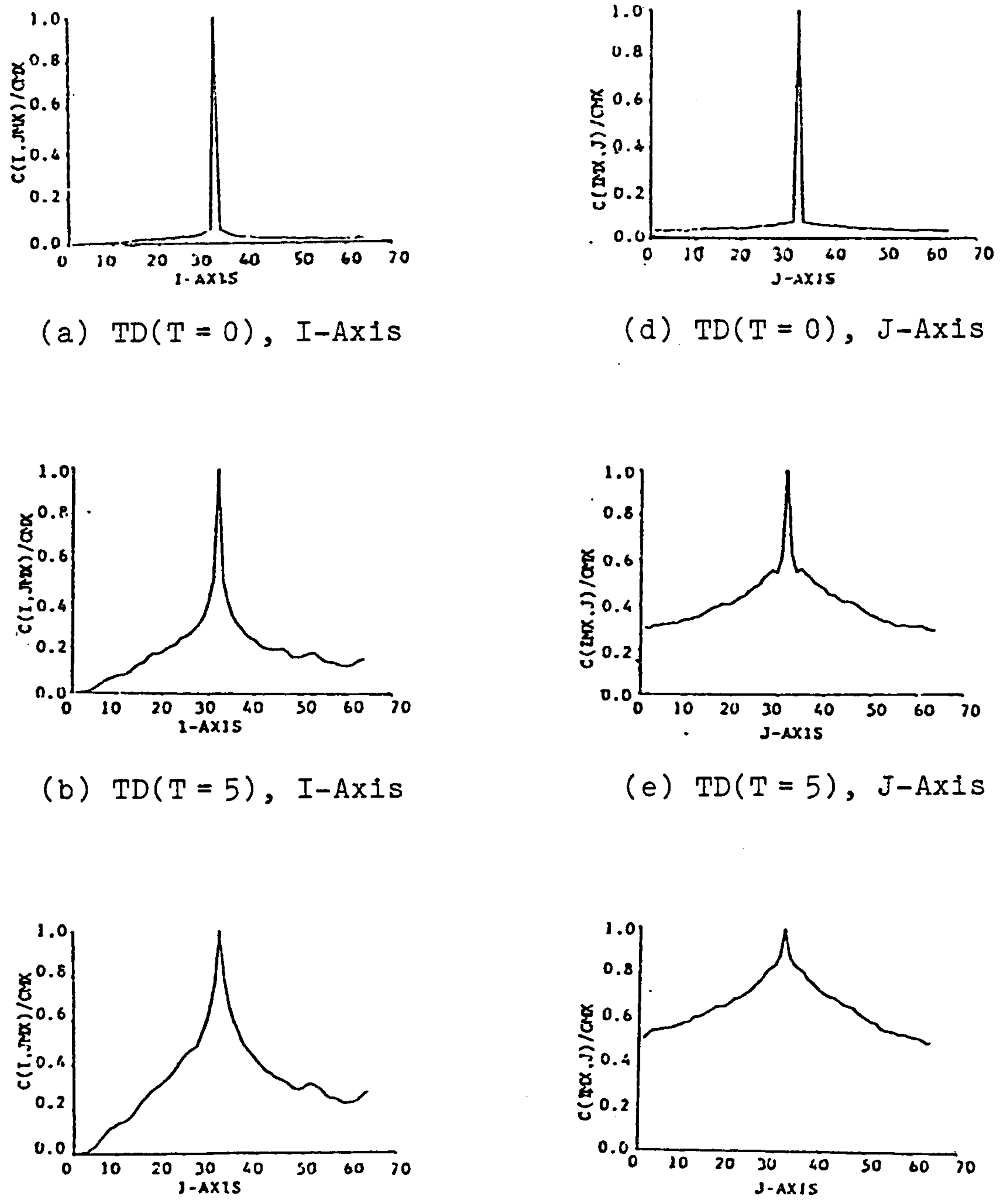


(c) DCC, Binary



(f) MAD, Binary

Figure 18. Cross-Sectional Plots of Scene 1 Registrations Using 8-Bit, Tri-Level, and Binary DCC, MAD Metrics.



(a) TD(T = 0), I-Axis

(d) TD(T = 0), J-Axis

(b) TD(T = 5), I-Axis

(e) TD(T = 5), J-Axis

(c) TD(T = 10), I-Axis

(f) TD(T = 10), J-Axis

Figure 19. Cross-Sectional Plots of Scene 1 Registrations Using TD(T = 0,5,10) Metrics.

6. The Computer CPU Time Consideration.

A comparison of the data in Table 1, page 46, shows that the FFT metric is the fastest algorithm. In the case of correlating 256 gray-level images using the FFT metric, if the normalization scheme was not included, the CPU time was only 30 seconds, but the result was a false fix peak at location (32,16). If the normalization scheme was included, the FFT metric took 4 minutes and 37 seconds CPU time and gave the correct matching peak at location (32,32). This result shows that the normalization scheme rejects the false peaks at the expense of increased computation time. The simulations were implemented on a VAX-11-780 computer whose basic memory unit is the BYTE (8 bits). Thus it was not possible to actually store and manipulate individual binary pixel values in one bit or tri-level pixel values in two bits. Therefore no CPU time advantage was achieved with the tri-level and binary images. If the algorithms were implemented using hardware logic gates, then registration of tri-level (two-bit) and binary (one-bit) images would be computed more efficiently.

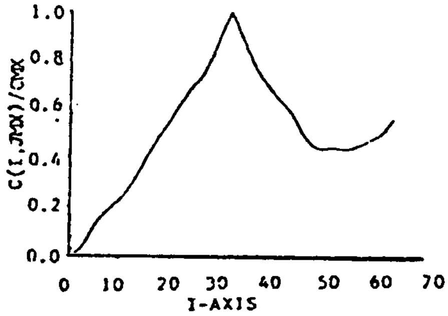
Simulation Case 2: Compatible Images, Noise Addition

To evaluate the effects of uncorrelated noise on the different registration metrics, random noise was added to R and S on a pixel-by-pixel basis and independently for each image. Noise-free R and S images were taken from Scene 1 in Case 1. Computer generated zero-mean Gaussian noise matrix [28] with specifiable standard deviation was added to each image. The pixel values of the noise-added reference and sensor array were then rescaled to 256 gray-levels prior to quantization and correlation.

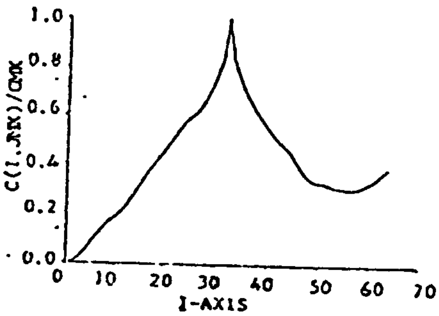
Two sets of simulations were done representing fairly high noise levels. Signal-to-noise ratio was defined as

$$SNR = \frac{\sigma}{\sigma_n} \quad , \qquad (3\text{-}1)$$

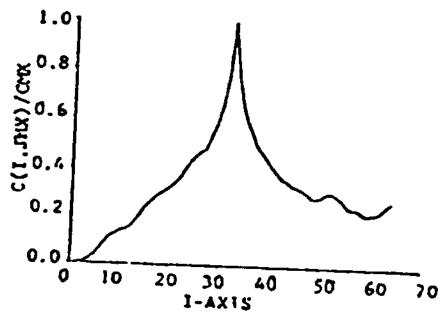
where σ is the image standard deviation and σ_n is the specifiable noise standard deviation. For 1 set of simulations a SNR of 0.43 was used, and for the other set a SNR of 1.42 was used. Figure 20 shows cross-sectional plots of registration for the ideal and the $SNR = 1.42$ noise added cases. Additional noise dependence is indicated in Table 2, which gives calculated peak locations for all the algorithms tested for Scene 1 with $SNR = 0.43$ and 1.42. Note that errors



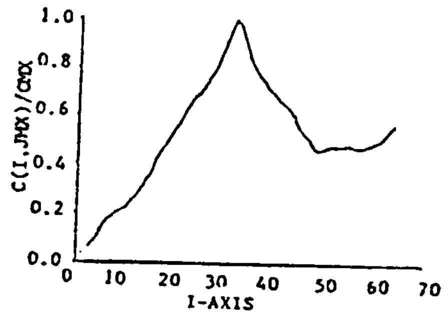
(a) DCC, Original



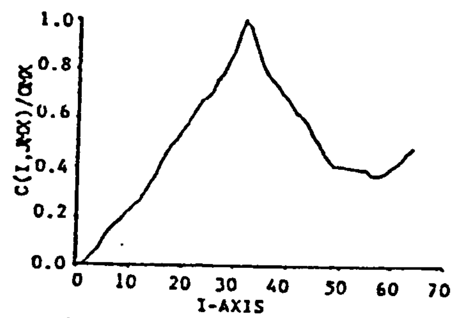
(b) MAD, Original



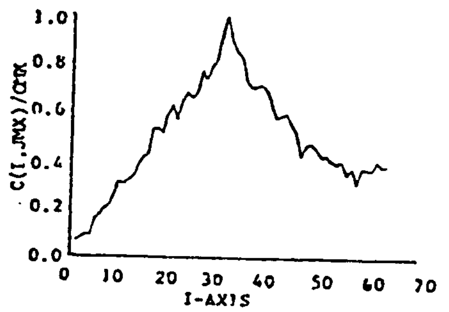
(c) TD(T = 10), Original



(d) DCC, SNR = 1.42



(e) MAD, SNR = 1.42



(f) TD(T = 10), SNR = 1.42

Figure 20. Cross-Sectional Plots of Scene 1 Original and SNR = 1.42 Image Registrations Using 8-Bit DCC, MAD, TD(T = 10) Metrics.

TABLE 2

RESULTS OF SIMULATION CASE 2

Correct Peak Location (32,32)		
	Computed Peak Location	
ALGORITHM	SNR = 0.43	SNR = 1.42
DCC (8-Bit)	(33,32)	(32,32)
DCC (Tri-Level)	(31,59)	(32,37)
DCC (Binary)	(31,59)	(32,32)
MAD (8-Bit)	(33,32)	(32,32)
MAD (Tri-Level)	(31,59)	(33,32)
MAD (Binary)	(31,59)	(32,30)
TD (T = 0)	(25,51)	(31,26)
TD (T = 2)	-	(33,38)
TD (T = 5)	-	(32,35)
TD (T = 10)	-	(32,32)
TD (T = 20)	(30,35)	(32,32)
TD (T = 30)	-	(32,32)
TD (T = 40)	(33,32)	-
TD (T = 60)	(31,40)	-
TD (T = 80)	(33,41)	-
TD (T = 100)	(33,32)	-

are larger for the horizontal (J-axis) peak location. This is expected since high frequency noise was added to an image with inherent low frequency in the horizontal direction.

The simulation results show that the TD metric (which is a high frequency emphasis metric) is more sensitive to the high frequency (pixel-by-pixel) noise added to the images. It even produces false peaks for very low values of T . Performance of the TD metric is similar to the DCC metric if T is "large enough." The value of T which is "large enough" depends on noise level. For the 2 simulations, the TD metric began to behave resonably well at a threshold of $T = 0.5\ \sigma_n$. This is illustrated in Table 2 where for SNR = 0.43, $T = 40 = 0.6\ \sigma_n$ ($\sigma_n = 65$) and for SNR = 1.42, $T = 10 = 0.51\ \sigma_n$ ($\sigma_n = 19.5$).

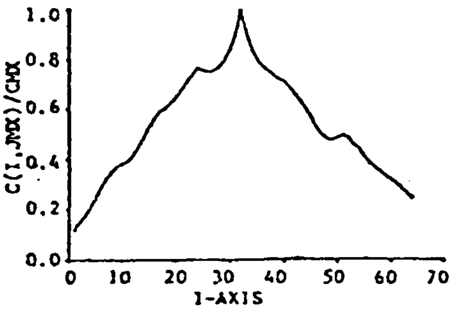
Simulation Case 3: Adaptive Binary Quantization

For binary quantization in Case 1 and Case 2, the global median was chosen as the threshold. This global thresholding tends to reduce high spatial frequencies in the images and broadens the matching peak. A locally adaptive threshold was implemented in Case 3 to quantize pixel values to 1 bit (binary quantization) before correlation. The threshold was computed as either the mean or median in a 3×3 moving window. At each location the center pixel was quantized about the threshold computed for that location.

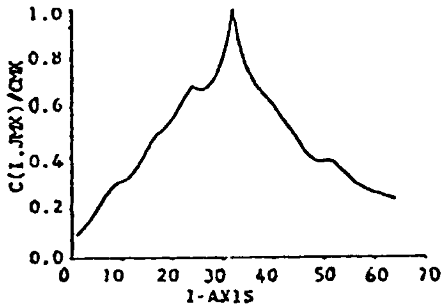
This adaptive threshold in effect emphasizes the image's high frequency content and should result in a much sharper correlation function. Figure 21 shows cross-sectional plots of registration for Scene 1 using global threshold, a 3×3 adaptive median threshold, and a 3×3 adaptive mean threshold. Table 3 summarizes the noise effects on the global and the adaptive threshold method in terms of calculated peak locations. The results show that a binary correlator, using locally adaptive thresholding in a 3×3 moving window, gave a sharper matching peak than the global thresholding method; however, it was more sensitive to the noise in the images. If the images to be correlated have low noise levels at higher spatial frequencies and if image distortions are strong only at low spatial frequencies, then the adaptive quantizer offers superior registration performance over the global quantizer.

Simulation Case 4: Non-Compatible Images, Different Spectral Bands

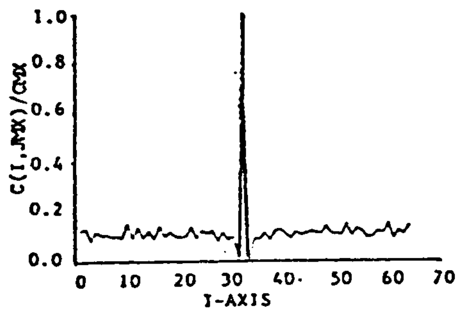
The two images used in this simulation were images of the same scene but in different spectral bands. These



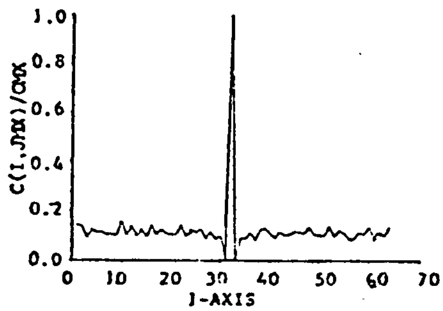
(a) Global-Median, DCC



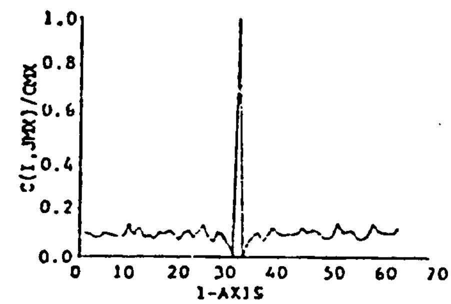
(d) Global-Median, MAD



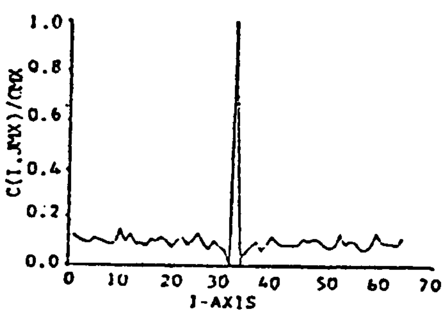
(b) Local-Median, DCC



(e) Local-Median, MAD



(c) Local-Mean, DCC



(f) Local-Mean, MAD

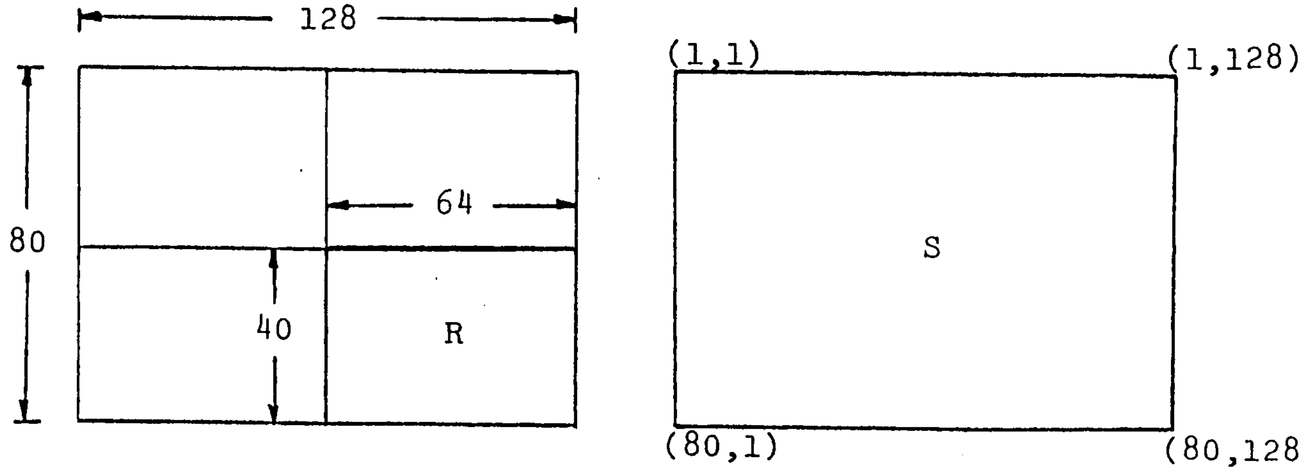
Figure 21. Cross-Sectional Plots of Scene 1 Global-Median, Local-Median, and Local-Mean Quantized Binary Image Registrations Using Binary DCC, MAD Metrics.

TABLE 3

RESULTS OF SIMULATION CASE 3

ALGORITHM (Binary)	Correct Peak Location (32,32) Global Median	Locally Adaptive 3 × 3 Mean	Locally Adaptive 3 × 3 Median
DCC (Ideal)	(32,32)	(32,32)	(32,32)
MAD (Ideal)	(32,32)	(32,32)	(32,32)
DCC (SNR = 1.42)	(32,32)	(32,1)	(11,1)
MAD (SNR = 1.42)	(32,30)	(32,1)	(11,1)
DCC (SNR = 0.43)	(31,59)	(12,15)	-
MAD (SNR = 0.43)	(31,59)	(59,64)	-

2 images in digitized form are stored as File 26 and File 27 on the magnetic tape received from Wright-Patterson Air Force Base. They are 128 gray-level (7-bit) images with array size 80 by 128. The File 26 image was taken with an 8-12 μ passive IR sensor, and the File 27 was the visible image taken with an optical sensor. The reference image R(40,64) was taken as the lower right quadrant of File 26. This segment was correlated with the sensor image S(80,128) which was the entire File 27 image. The relationship between R(40,64) and S(80,128) is shown in Figure 22. The correct peak location is (41,65). Simulations were done to compare the DCC, MAD and TD registration metrics with and without gradient preprocessing. A summary of the results is given in Table 4 in terms of calculated peak locations. Clearly, the closest match is (41,50) for the TD (T = 10, Gradient) method. Yet the match is not accurate in the horizontal (J-axis) direction. The results indicate that further preprocessing to extract similarities between the two images is required. Further information on the spectral response of the two sensors would be useful in improving non-compatible image registration results.



File 26 (8-12 μ Image) File 27 (Visible Image)

Figure 22. Relationship between R(40,64) and S(80,128) in Simulation Case 4.

TABLE 4

RESULTS OF SIMULATION CASE 4

Air Force File 26 (8-12 μ IR Image)
Air Force File 27 (Visible Image)

Correct Peak Location (41,65)

ALGORITHM	Computed Peak Location
DCC (8-Bit)	(37,2)
MAD (8-Bit)	(40,29)
TD (T = 10)	(41,29)
TD (T = 50)	(41,29)
DCC (8-Bit)(Gradient)	(2,25)
MAD (8-Bit)(Gradient)	(2,25)
TD (T = 10)(Gradient)	(41,50)
TD (T = 50)(Gradient)	(4,32)

Simulation Case 5: Non-Compatible Images, Different Times of Day

The two images used in this simulation were taken at different times of day by an 8-12 μ passive IR sensor. These 2 images in digitized form are stored as File 2 and File 10 on the magnetic tape received from Wright-Patterson Air Force Base. They are 128 gray-level (7-bit) images with array size 150 by 150. The reference image R(50,50) was segmented from the central region of the File 2 image which was taken at 12:00 o'clock noon. The entire image of the File 10 which was taken at 2:00 o'clock p.m. was chosen as the sensor image S(150,150). The relationship between R(50,50) and S(150,150) is shown in Figure 23. The correct matching peak location is (51,51). Since the sensor positions and directions were the same, there should be minimal geometric distortion. The only difference was intensity differences caused by the two images being taken at different times of day. Simulation results are summarized in Table 5. Figures 24 and 25 show cross-sectional plots of registration for the original and gradient images. As can be seen, all methods correctly identified the registration point, so the best comparison of methods is in terms of relative peak sharpness and sidelobe levels. The data in Table 5 show that all algorithms gave the

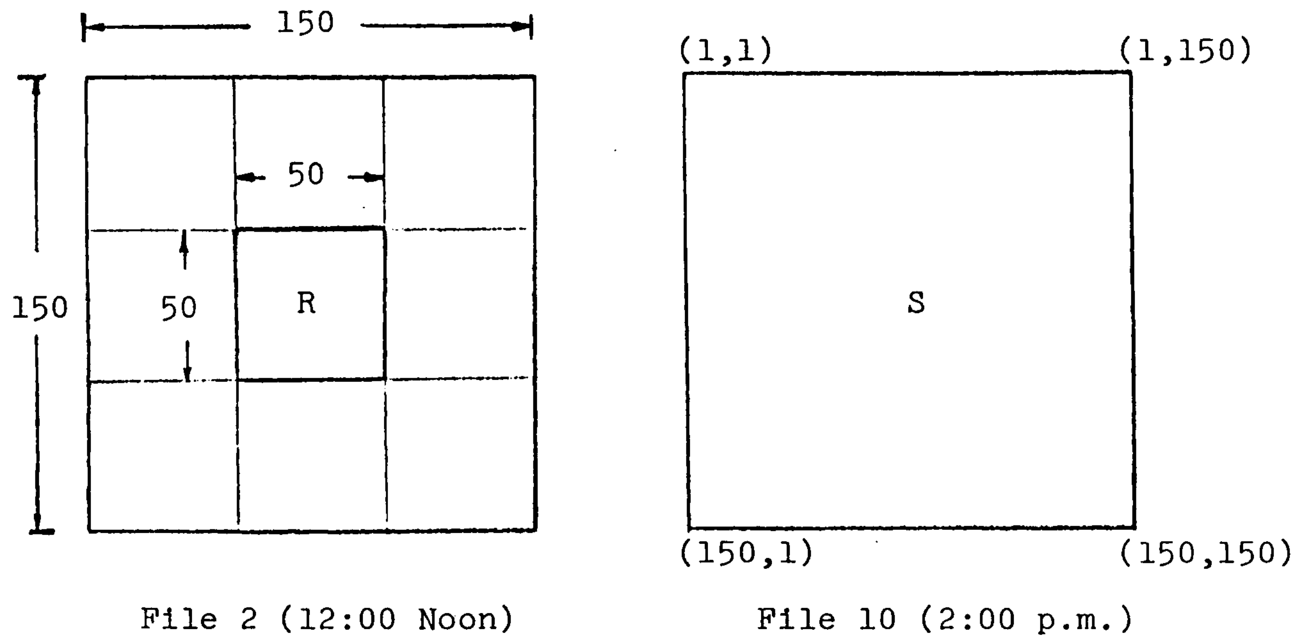
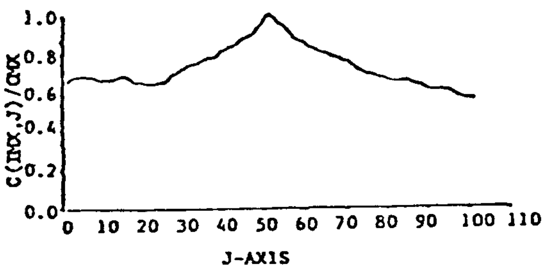


Figure 23. Relationship between R(50,50) and S(150,150) in Simulation Case 5.

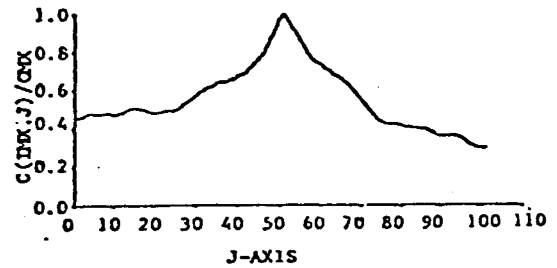
TABLE 5

RESULTS OF SIMULATION CASE 5

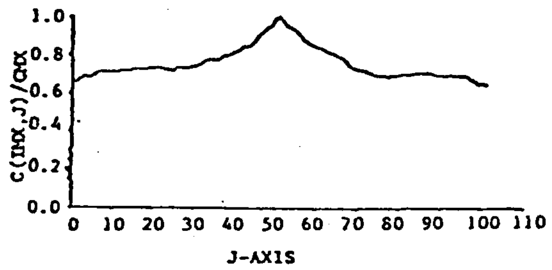
Air Force File 2 (Noon) Air Force File 10 (2:00 o'clock p.m.)	
Correct Peak Location (51,51)	
ALGORITHM	Computed Peak Location
DCC (8-Bit)	(51,51)
MAD (8-Bit)	(51,51)
TD (T = 5)	(51,51)
TD (T = 10)	(51,51)
TD (T = 15)	(51,51)
DCC (8-Bit)(Gradient)	(51,51)
MAD (8-Bit)(Gradient)	(51,51)
TD (T = 5)(Gradient)	(51,51)
TD (T = 10)(Gradient)	(51,51)
TD (T = 15)(Gradient)	(51,51)



(a) DCC

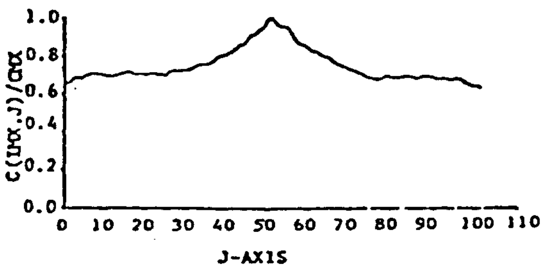


(b) MAD



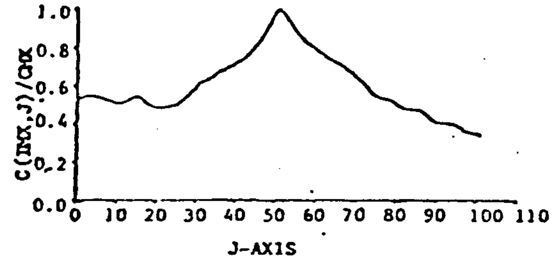
(c) TD(T = 5)

(d) TD(T = 10)

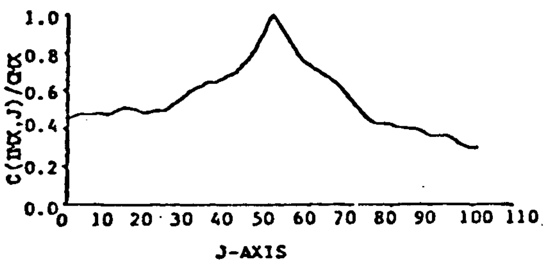


(e) TD(T = 15)

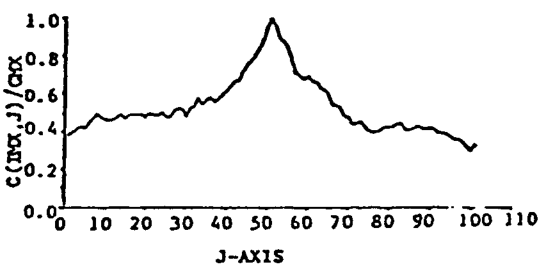
Figure 24. Cross-Sectional Plots of Case 5 Original Image Registrations Using 7-Bit DCC, MAD, and TD(T = 5,10,15) Metrics.



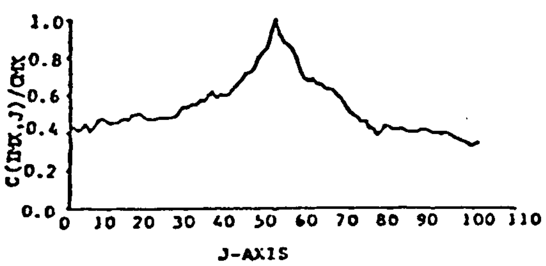
(a) DCC



(b) MAD



(c) TD(T = 5)



(d) TD(T = 10)

(e) TD(T = 15)

Figure 25. Cross-Sectional Plots of Case 5 Gradient Image Registrations Using 7-Bit DCC, MAD, and TD(T = 5,10,15) Metrics.

correct peak location (51,51). These results were expected to be good because there were relatively small intensity differences in images taken at 12:00 o'clock noon and at 2:00 o'clock p.m.

Simulation Case 6: Non-Compatible Images, Different Scale and Different Spectral Band

The two images used in this simulation were taken with different sensors and different fields-of-view. These 2 images in digitized form are stored as File 9 and File 10 on the magnetic tape received from the U. S. Army Missile Command. Figure 26 shows File 9 image and Figure 27 shows File 10 image. They are 256 gray-level (8-bit) images with array size 407 by 392. The File 9 image was taken with a wide field-of-view (WFOV) TV sensor and the File 10 image was taken with a narrow field-of-view (NFOV) forward-looking infrared (FLIR) sensor. A comparison of Figures 26 and 27 shows that these two images have different scale factors and contrast reversals. It is necessary to preprocess one of the images to match the scale before registration. Secondly, it is desirable to perform an intensity transformation to eliminate contrast reversals. The procedures were as follows:

1. Scale difference was computed from dimensions of a large rectangular target board in the two

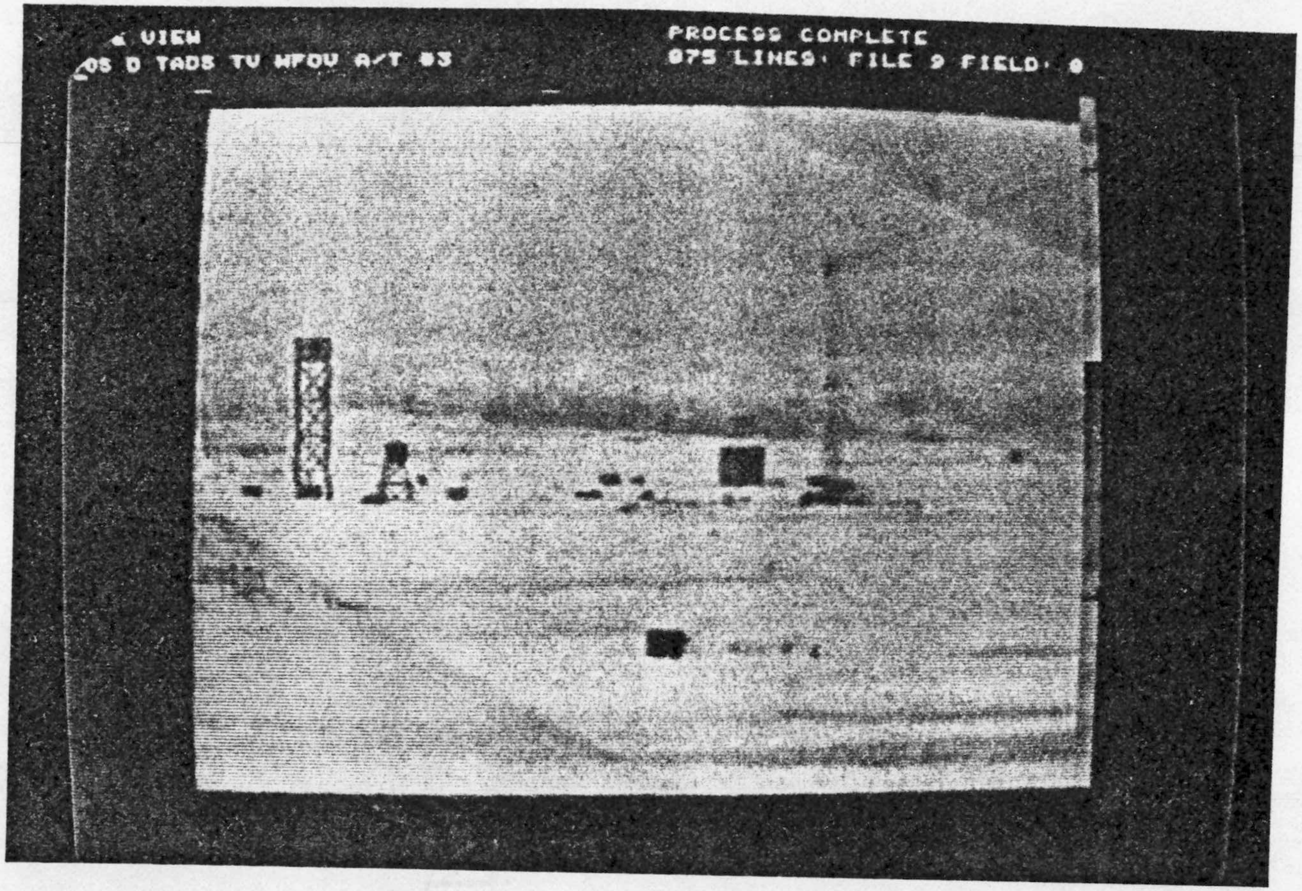


Figure 26. Army File 9 WFOV TV Image.

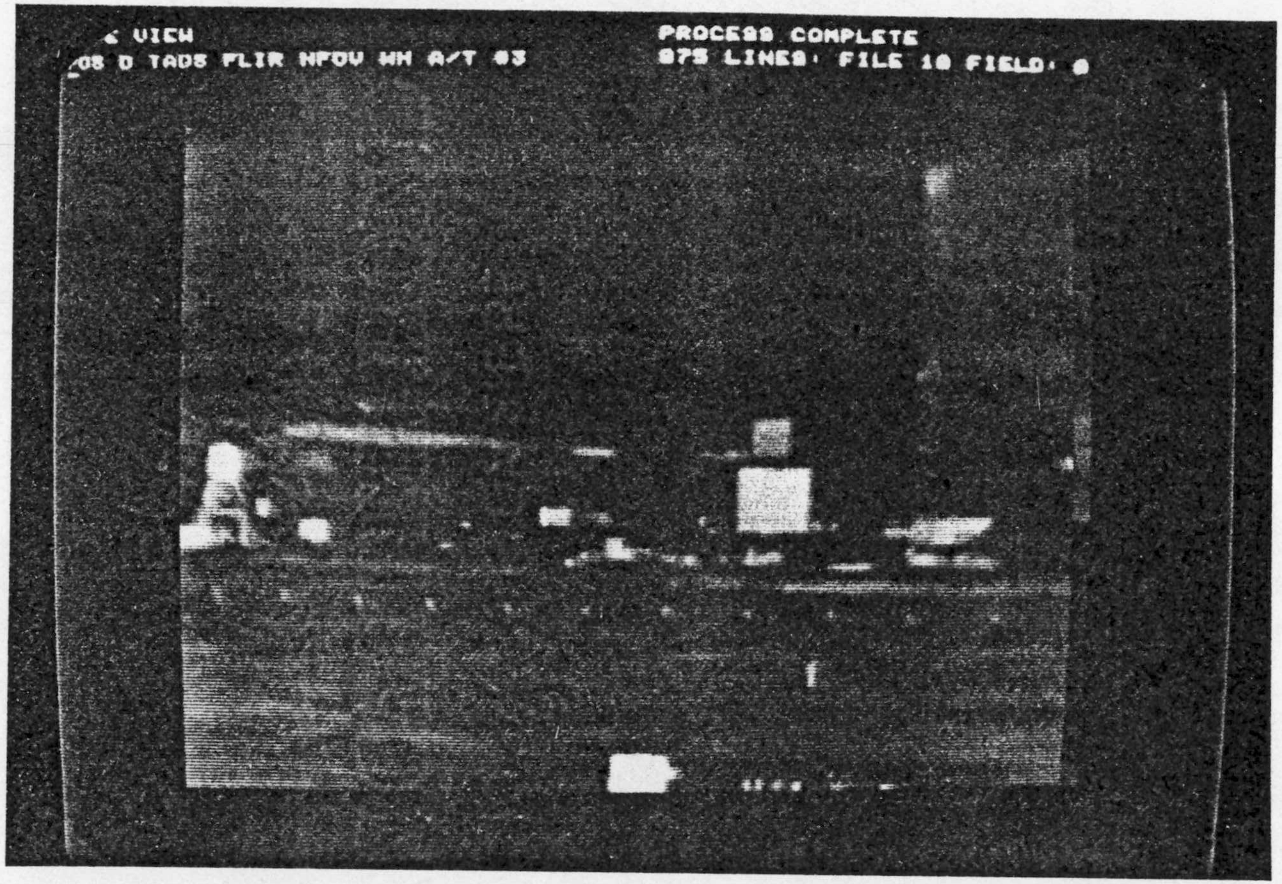
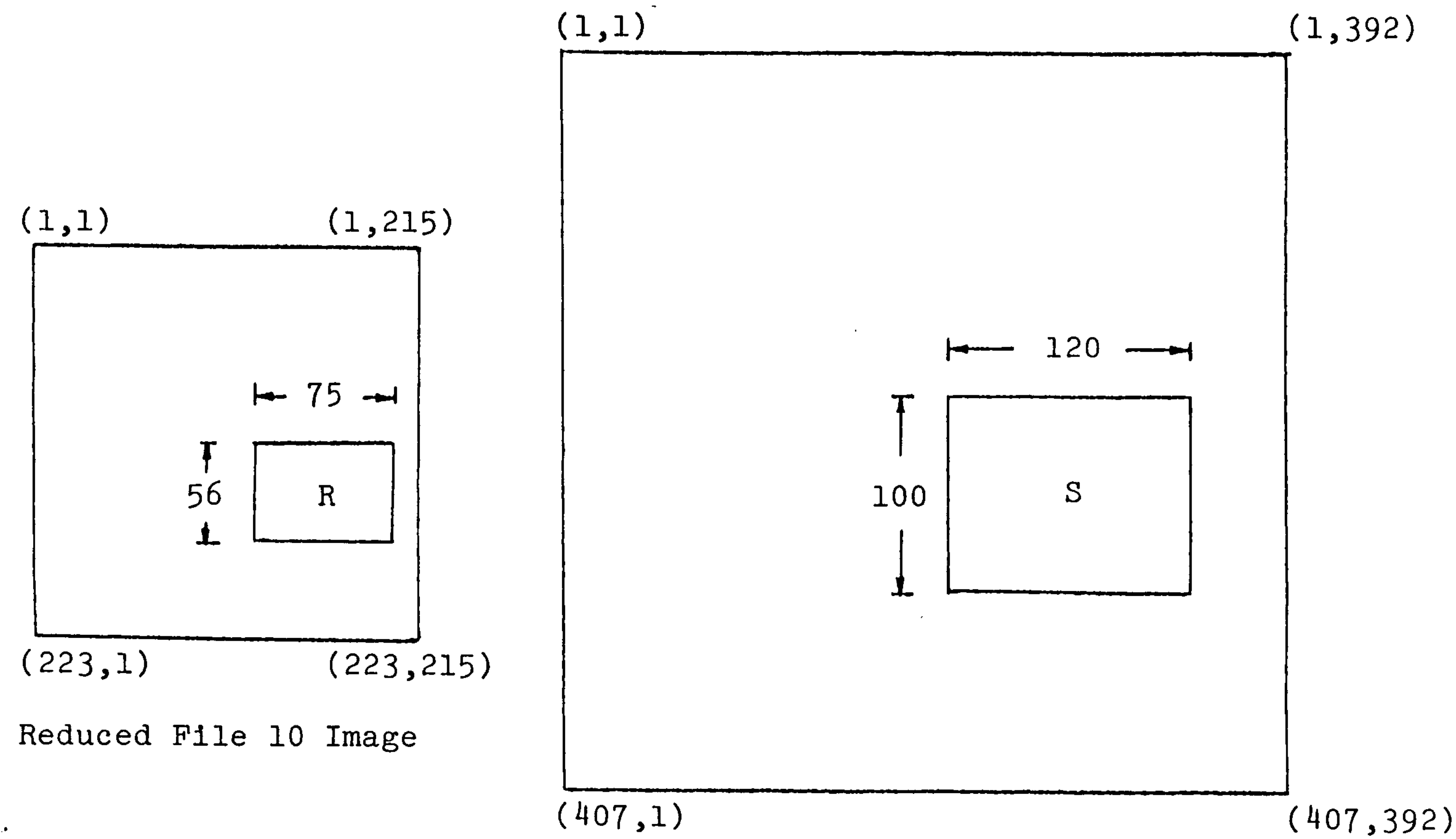


Figure 27. Army File 10 NFOV FLIR Image.

images. Since File 10 was a narrow field-of-view image, objects in File 10 appeared larger than in File 9 by a ratio determined to be 1.0/0.55. Scaling was applied to reduce the size of objects in File 10 to match the smaller size in File 9. After scaling, File 10 then had only 223 by 215 pixels.

2. A simple intensity transform $F(i,j) = 255 - F(i,j)$ was performed on the pixel values of the reduced File 10 to eliminate contrast reversals.

A subimage R(56,75) segmented from the reduced File 10 was taken as the reference image. The other subimage S(100,120) segmented from File 9 was taken as the sensor image. The relationship between R(56,75) and S(100,120) is shown in Figure 28. The correct matching peak location is (11,6). Simulation results are summarized in Table 6 in terms of calculated peak locations. Peak profiles are given in Figure 29 for the original and gradient images. The DCC eight-bit, MAD eight-bit and TD (various thresholds) metrics are compared. As can be seen from Table 6, false peaks were obtained for the TD metric with low threshold values. This was expected because of the significant dissimilarity between the two images (different scale and different spectral bands). From Table 6, it is clear that for the



Reduced File 10 Image

File 9 Image

Figure 28. Relationship between R(56,75) and S(100,120) in Simulation Case 6.

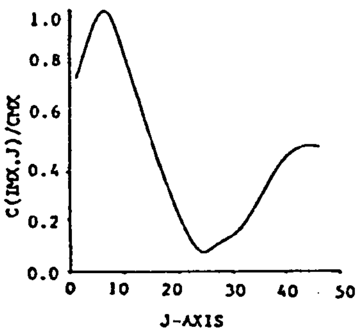
TABLE 6

RESULTS OF SIMULATION CASE 6

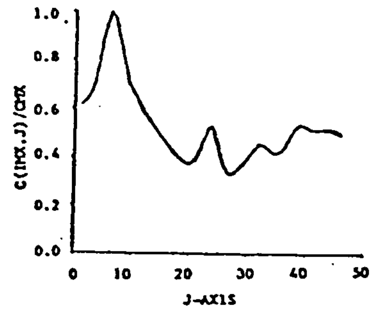
Army File 9 (Si TV, WFOV, Scale 0.55)
Army File 10 (FLIR, NFOV, Scale 1.00)

Correct Peak Location (11,6)

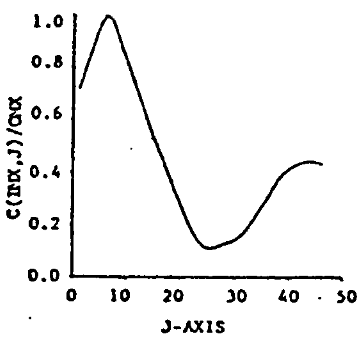
ALGORITHM	Computed Peak Location: Original	Gradient
DCC (8-Bit)	(11,6)	(11,6)
MAD (8-Bit)	(11,6)	(11,6)
TD (T = 5)	(41,28)	(44,44)
TD (T = 10)	(43,25)	(44,45)
TD (T = 15)	-	(45,44)
TD (T = 20)	(11,5)	(11,6)
TD (T = 25)	(11,6)	(11,7)
TD (T = 30)	(11,6)	(11,6)
TD (T = 35)	(11,6)	(11,6)
TD (T = 40)	(20,46)	(11,6)
TD (T = 45)	-	(11,6)
TD (T = 60)	(21,44)	(11,6)



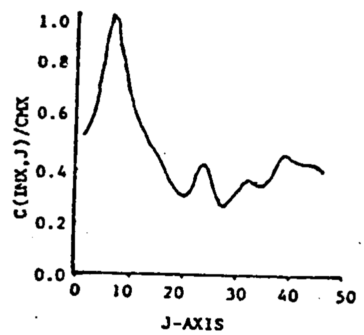
(a) DCC, Original



(d) DCC, Gradient

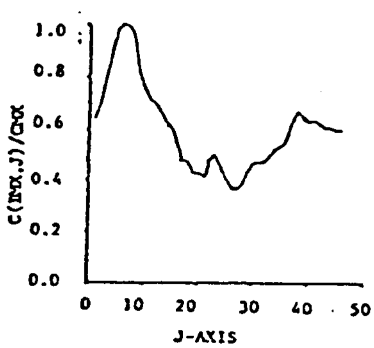


(b) MAD, Original



(e) MAD, Gradient

(c) TD(T = 30), Original



(f) TD(T = 30), Gradient.

Figure 29. Cross-Sectional Plots of Case 6 Original and Gradient Image Registrations Using 8-Bit DCC, MAD, TD(T = 30) Metrics.

TD metric, proper selection of the threshold T is important. The selection of the optimum threshold T depends on the signal-to-noise ratio, percent of geometric distortion, different spectral response and scene statistics. It is significant that the TD metric, with a suitable threshold T, performs as well as or better than the DCC and MAD metrics.

CHAPTER IV

CONCLUSIONS AND RECOMMENDATIONS

Based on the simulation results and discussions presented in Chapter III, a number of conclusions and recommendations may be made:

1. For ideal compatible images, all registration metrics gave the correct peak location for original and gradient images. The sharpness of the matching peak was improved in registering gradient images. The fastest metric is the fast Fourier transform (FFT) metric while the thresholded difference (TD) metric yields the sharpest matching peak. For noisy compatible images, the correlator performance is related to signal-to-noise ratio SNR in terms of a lower SNR giving higher incidence of false peaks.
2. An adaptive binary quantizer was implemented which sets the quantization threshold as the mean or median in a 3×3 moving window. It showed far superior peak sharpness when compared with a global thresholding method; however, it showed greater sensitivity to noisy images.

3. For non-compatible images which represent different spectral bands, different times of day and different scale, the preprocessing algorithms are important. Preprocessing algorithms including edge detection, geometric transformation and intensity transformation must bring the two images into compatibility (to the extent possible) before applying any registration metrics.
4. Scene Dependence. Relative sharpness of the registration function, probability of correct registration and the success of preprocessing algorithms have been shown through simulation studies to depend strongly on scene characteristics. Further analysis may lead to the use of this dependence to allow optimization of correlator performance.
5. The simulations reported in Chapter III demonstrate the potential of the new registration metric—the thresholded difference (TD) metric. By a proper choice of threshold T, it yields the sharpest matching peak. Computationally, it is faster than all 8-bit registration metrics except for the FFT metric; its algebraic simplicity (see Figure 2,

page 17) makes it easily implementable in microprocessor based systems. The only difficulty with the TD metric is that there is still uncertainty as to how the optimum threshold T should be chosen. An optimum threshold for the TD metric should be based on some measure of noise and/or distortion between the two images to be correlated. Further work is recommended to provide the theoretical basis and a practical measurement technique to establish an optimum value for T in a realistic application.

6. Simulation Case 6 showed a successful correlation of a day-TV image with imagery from a co-located forward-looking infrared (FLIR) system. Such an approach to correlating non-compatible images from different sensor types is important for the development of a stand-alone, all-weather guidance system without requiring that a reference image be stored for each sensor.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Gonzalez, R. C., and P. Wintz. Digital Image Processing. Mass.: Addison-Wesley Publishing Co., Inc., 1977.

2. Frei, W., T. Shibata, and G. C. Huth. "Environmental Change Detection in Digitally Registered Aerial Photographs," Proceedings of the SPIE on Application of Digital Image Processing III, 207:26-31, 1979.

3. Hall, E. L. Computer Image Processing and Recognition. New York: Academic Press, Inc., 1979.

4. Kuglin, C. D., A. F. Bulmenthal, and J. J. Pearson. "Map-Matching Techniques for Terminal Guidance Using Fourier Phase Information," Proceedings of the SPIE on Digital Processing of Aerial Images, 186:21-29, 1979.

5. Lo, T. K., G. Gerson. "Guidance System, Position Update by Multiple Subarea Correlation," Proceedings of the SPIE on Digital Processing of Aerial Images, 186:30-40, 1979.

6. Pinson, L. J. "Statistical Performance for a Digital Image Correlator—Scene Dependence," Final Technical Report Contract DAAG29-76-D-0100, 31 May 1979.

7. Anuta, P. E. "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques," IEEE Trans. on Geosci. Electron., GE-8:353-368, Oct. 1970.

8. Kuglin, C. D., and D. C. Hines, Jr. "The Phase Correlation Image Alignment Method," Proceedings of the IEEE 1975 Conference on Cybernetics and Society, pp. 163-165, Sept. 1975.

9. Barnea, D. I., and H. F. Silverman. "A Class of Algorithms for Fast Digital Image Registration," IEEE Trans. on Computer, C-21:179-186, Feb. 1972.

10. Svedlow, M., C. McGillem, and P. E. Anuta. "Image Registration: Similarity Measure and Preprocessing Method Comparisons," IEEE Trans. on Aerospace and Electron. Syst., AES-14 (No. 1): 141-150, Jan. 1978.

11. Merchant, J. "Full-Frame Image Registration by Address Modification," Proceedings of the SPIE on Digital Processing of Aerial Images, 186:49-52, 1979.

12. Pinson, L. J., C. L. Rao, and J. C. Su. "Adaptive Correlation Concepts for Non-Compatible Imagery," Interim Technical Report Contract AFOSR-80-0275, 15 Jan. 1981.

13. Frei, W., and C. C. Chen. "Fast Boundary Detection: A Generalization and a New Algorithm," IEEE Trans. on Computer, C-26 (No. 10):988-998, Oct. 1977.

14. Wong, R. Y. "Sequential Scene Matching Using Edge Features," IEEE Trans. on Aerospace and Electron. Syst., AES-14 (No. 1):128-140, Jan. 1978.

15. Cooley, J. W., and J. W. Tukey. "An Algorithm for Machine Calculation of Complex Fourier Series," Math. Comp., 19:297-301, Apr. 1965.

16. Buijs, H. L., A. Pomerleau, M. Fourier, and W. G. Tam. "Implementation of a Fast Fourier Transform (FFT) for Image Processing Applications," IEEE Trans. on Acoust., Speech, Signal Process., ASSP-266 (No. 6):420-424, June 1974.

17. Pinson, L. J. (Private Communication). University of Colorado, Colorado Springs.

18. Cooper, B.F.C. "Correlators with Two-Bit Quantization," Aust. J. Phys., 23:521-527, 1970.

19. Boland, J. S. III, L. J. Pinson, and E. G. Peters. "Automatic Target Hand-Off for Non-Compatible Imaging Systems." Final Technical Report Contract DAAK40-77-C-0156, 15 Sept. 1978.

20. Pinson, L. J., J. S. Boland, III, and W. W. Malcolm. "Statistical Analysis for a Binary Image Correlator in the Absence of Geometric Distortion," Optical Engineering, 17 (No. 6):635-639, Nov.-Dec. 1978.

21. Boland, J. S. III, L. J. Pinson, etc. "Automatic Target Hand-Off Using Correlation Techniques," Final Technical Report Contract DAAH01-76-C-0396, 31 Jan. 1977.

22. Webber, R. F., and W. H. Delashmit. "Product Correlator Performance for Gaussian Random Scenes," IEEE Trans. on Aerospace and Electron. Syst., AES-10 (No. 4):516-520, July 1974.

23. Wong, R. Y., and E. L. Hall. "Performance Comparison of Scene Matching Techniques," IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-1 (No. 3):325-330, July 1979.

24. Pratt, W. K. "Correlation Techniques of Image Registration," IEEE Trans. on Aerospace and Electron. Syst., AES-10 (No. 3):353-358, May 1974.

25. McGillem, C. D., and M. Svedlow. "Image Registration Error Variance as a Measure of Overlay Quality," IEEE Trans. on Geosci. Electron., GE-14 (44-49), Jan. 1976.

26. Peebles, P. Z. Jr. Probability, Random Variables, and Random Signal Principles. New York: McGraw-Hill Book Co., 1980.

27. Prewitt, J.M.S. "Object Enhancement and Extraction," in Picture Processing and Psychopictorics, (B. S. Lipkin and A. Rosenfeld, eds.), New York: Academic Press, Inc., 1970.

28. Forsythe, G. E., M. A. Malcolm, and C. B. Moler. Computer Methods for Mathematical Computations. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1977.

VITA

Jyh-Chyuan Su was born in Tainan, Taiwan, Republic of China, on April 10, 1954, the son of Fong-Mei and Ching-Mo Su. He attended elementary school in the Tainan area and graduated from The First Tainan High School in June 1972. The following October he entered National Tsing Hua University at Hsinchu, China, and in June 1976 he received a Bachelor of Science degree in Nuclear Engineering. He then served in the Chinese Navy Fourth Shipyard as an Engineering Officer under the ROTC Program for two years. In September 1978 he entered the Graduate School of National Tsing Hua University, and in June 1980 he earned his Master's degree in Nuclear Engineering. In September 1980 he accepted a graduate research assistantship from The University of Tennessee, Knoxville, and began study toward a Master's degree in Electrical Engineering at The University of Tennessee Space Institute (UTSI); this degree was awarded in March 1982.

He is married to the former Miss Chin-Chu Chang, and they have a one-year-old son, Jiunn-Chao.