STARS

University of Central Florida
STARS

Retrospective Theses and Dissertations

1978

Realization of a Fast Automatic Correlation Algorithm for Registration of Satellite Images

John E. Kassak
University of Central Florida, jkassak@cfl.rr.com

 Part of the Engineering Commons
Find similar works at: https://stars.library.ucf.edu/rtd
University of Central Florida Libraries http://library.ucf.edu

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation
Kassak, John E., "Realization of a Fast Automatic Correlation Algorithm for Registration of Satellite Images" (1978). *Retrospective Theses and Dissertations*. 293.
https://stars.library.ucf.edu/rtd/293

REALIZATION OF A FAST AUTOMATIC CORRELATION ALGORITHM FOR REGISTRATION OF SATELLITE IMAGES

BY

JOHN E. KASSAK
VANDERBILT UNIVERSITY, 1962

RESEARCH REPORT

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Engineering
in the Graduate Studies Program of the College of Engineering
of Florida Technological University

Orlando, Florida
1978

REALIZATION OF A FAST AUTOMATIC CORRELATION ALGORITHM FOR REGISTRATION OF SATELLITE IMAGES

BY

JOHN E. KASSAK

ABSTRACT

The requirement for a fast automated correlation algorithm for registration of satellite images is discussed. An overview of current registration techniques is presented indicating a correlator, matching binary maps compressed from the original imagery, may provide the required throughput when implemented with a dedicated hardware/processor. An actual registration problem utilizing GOES digitally processed imagery is chosen and defined. The realization of a fast correlator, matching image input data with sampled data base reference image data in real time is considered.

Director of Research Report

TABLE OF CONTENTS

ILLUSTRATIONS

LIST OF TABLES

LIST OF PHOTOGRAPHS

Picture

CHAPTER I

REGISTRATION CONCEPT

A. Need For Satellite Image Registration

Satellites provide researchers and investigators with a data source unmatched at comparable spatial and temporal coverage by any existing or practical alternate source. The data has been particularily advantageous in the study of Synoptic Meteorology, Atmospheric Profiling, and Measurement of Surface Features.[1]

Imagery from meteorological satellites and Landsat Earth Resources Satellite is ideally suited for both subjective enhancement via-time-lapse display and for objective measurements with time. This imagery is particularily useful for temporal studies because sequential surface views of a certain geographic area can be registered with minimum geometric error. The images are transmitted on succeeding orbital passes by scanning radiometer sensors aboard earth resources or meteorological satellites. The images are recorded at earth receiving stations. Very few investigations, to date, have reported on theme enhancement through temporal processing. Temporal analysis is often neglected because the facilities and techniques required to create well-registered multi-date sequences are somewhat specialized and not generally available.[2]

Preparation of even a simple movie-loop is annoyingly time consuming and then one is still left with the problem of extracting quantitative data. Each image or frame must be registered or aligned before photographed or stored so that there is no relative movement of fixed subjects when the images are viewed sequentially. Each image may occupy one or more frame periods (1/25 seconds) this movie-loop technique is preferred since spatially related trends are difficult to follow when operating from digital tapes.

The initial studies have shown that just as still photos are virtually indispensable when working with individual images, there is a basic need for some form of animated imagery as the workhorse tool for guidance in temporal studies. Nothing else seems to communicate the message to a human quite so quickly or clearly as time-lapse sequences, flicker comparisons, and the like.[3]

Currently, satellite pictures are either registered by hand or automatically by use of large computer systems. What is meant by the term registration? If we have two images A and B, let X_A, Y_A define a pixel element in image A and X_B, Y_B define a pixel in image B. Both images represent the same basic information but are displaced or captured from different camera stations. Assume that X_A, Y_A and X_B, Y_B represent the same fixed reference point on the surface of the earth, but are not the same horizontal element or vertical line with reference to the upper left corner of each picture or image. If sequential images are to be viewed as a motion picture, then all points such X_A, Y_A and X_B, Y_B must always be

viewed at constant distance from the viewing field reference boundaries; otherwise stationary subject matter will appear to move or jitter. Then one image must be moved relative to the other until all stationary points in the viewing field are aligned. The resulting image field can be no larger than the matched area of the subject images. Unmatched data cannot be used for temporal analysis. If two images have only a small segment which overlaps, then very little data is available for time comparisons.

Most weather field stations and many television weather stations register the picture electronically by means of video-analog comparative techniques. Only at Goddard Space Flight Center are satellite images completely registered by use of computers. The computational time is, however, considerable.[4]

B. Review of Applicable Registration Methods

Registration is basic to any image processing system. When it is desired to detect changes or perform a mapping of two similar images, it is necessary for meaningful results to have the images registered. If the pictures do not differ in magnification and rotation, then the best translational fit will yield the required registration. The problem is then to find a translation registration algorithm which can be implemented with a special purpose processor to provide automatic registration of images which are digitally processed and stored.

A search was made of existing automatic registration techniques in order to find a candidate algorithm which could be implemented

with a dedicated processing system. The goal being registration at the input frame rate, thus providing the required movie sequence. A full Geostationary Operating Environmental Satellite (GOES) image or frame is received every thirty minutes and requires fifteen minutes for transmission. The goal here is to correlate the incoming data points with a reference image at the actual data input rate. This means that a full image shall be registered in thirty minutes or less. Time must also be allotted for image enhancement and/or other data transformation before processing the next image.

A new image is thus processed and recorded every thirty minutes. This image is added to the sequence of previously transmitted satellite images providing a current movie loop of no more than thirty minutes delay. The image can be stored sequentially on an analog video disc and played back through a television monitor to produce the animated loop.

Three basic correlation methods exist: (1) Correlation, (2) Fast Fourier Transform Correlation, and (3) Binary Correlation (see Fig. 1). Let two images S (search area) and W (window), a subimage of S, be defined as shown in Figure 1. S is taken as an LxL array of digital picture elements which may assume one of K gray levels; i.e.[5]

$$0 \leq S(U+i-1, V+j-1) \leq K-1$$

$$1 \leq U,V \leq L-M+1$$

W is considered to be an MXM, M smaller than L array of digital picture elements having the same gray scale range:

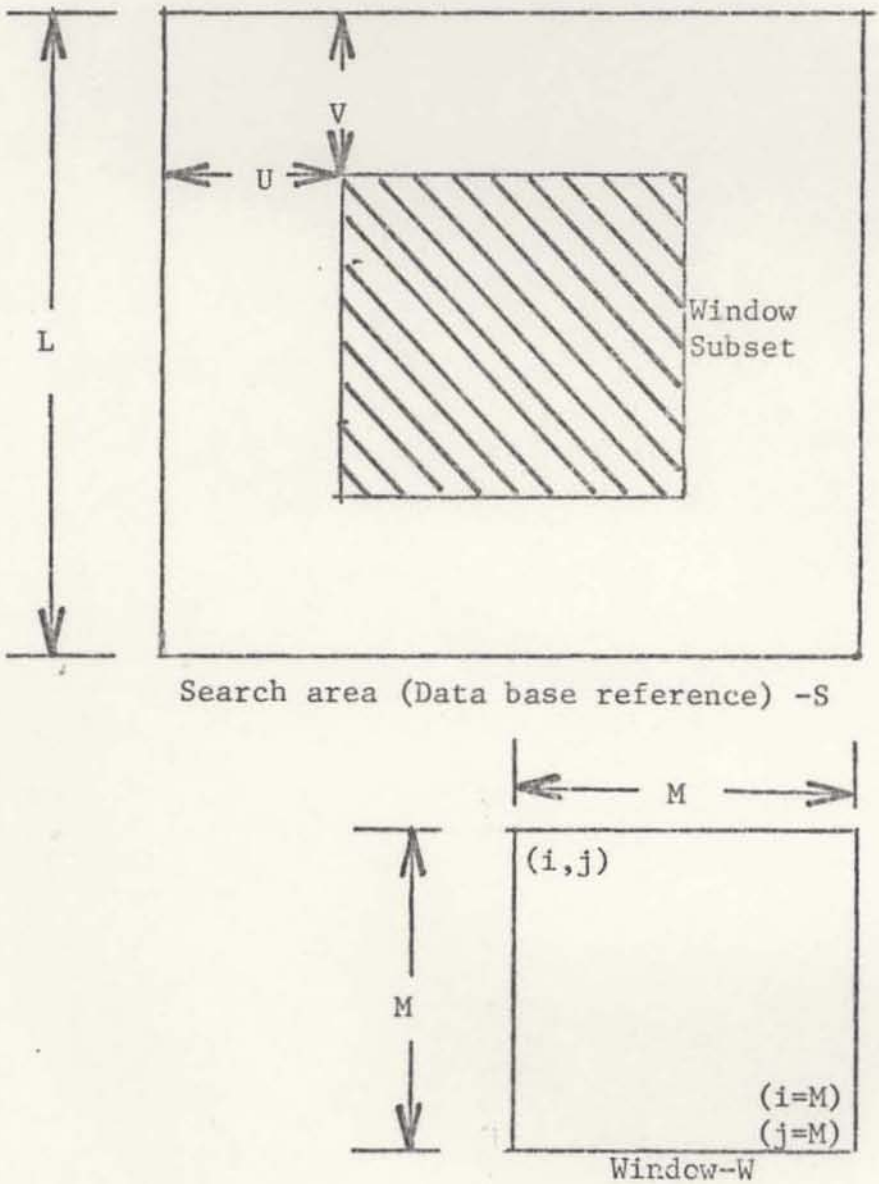


Fig. 1 Window and search area relationship[6]

$$0 \leq W(i,j) \leq K-1$$

$$1 \leq i,j \leq M$$

It will be assumed that enough a priori information is known about the dislocation between the window and the search area so that the parameters L and M may be selected with the virtual guarantee that at registration a complete subimage is contained in the search area (LxL) as shown in figure 1. Translational registration, therefore, is a search over some subset of the allowed range of reference points to find a point (U*,V*) which indicates a subimage in S that is most similar to the given window.

The measure of similarity is based on the Euclidean distance between two vectors (refer to Appendix A). As a result, one can use the normalized cross correlation as a measure of match. Correlation between two random variables of difference processes is by definition the expected value of their product[7] or

$$E(W,S) = \int_{-\infty}^{\infty} dW \int_{-\infty}^{\infty} WS\ P(W,S)dS \qquad (1)$$

where P(W,S) is equal to the joint probability of random variables W and S. Normalized cross correlation or normalized cross-covariance function can be defined as:

$$\rho_{WS}(U,V) = \frac{E\{(W-M_W)\ (S-M_S)\}}{E\{(W-M_W)^2\}\ E\{(S-M_S)^2\}} \qquad (2)$$

$$\rho_{ws}(U,V) = \frac{R_{ws}(U,V)-M_s M_w}{\sigma_s \sigma_w} \qquad (3)$$

where M = the mean value of the process and σ = standard deviation of a Gaussian distributed random process.

Note that σ_w is constant in the correlation process since it is not a function of U and V, but σ_s is dependent on U and V in the general case. Also, if both processes are Gaussian distributed with zero mean ($M_s=M_w=0$) and both processes are Ergodic, then [$R_{ws}(U,V) = \tau_{ws}(U,V)$] where τ_{ws} is equal to the spatial average of variables W and S. Then

$$E(WS) = \iint_{\ell} W(x,y)\ S(x+U,y+V)dxdy \qquad (4)$$

The resulting normalized cross correlation equation for two analog images is:

$$\rho_{ws}(U,V) = \frac{R_{ws}(U,V)}{\sigma_s(U,V)} = \frac{\iint W(x,y)\ S(x+U,y+V)dxdy}{\sqrt{\iint S^2(x+U,y+V)dxdy}} \qquad (5)$$

where σ_s = standard deviation of the S process.

The integrals are replaced with the summation signs for the discrete case. Current methods used for calculating the normalized cross correlation surface on most large machines are direct integration and shifting (discrete convolution) and Fast Fourier Transform (FFT) in Appendix B. Discrete convolution methods require considerable computational time. The FFT is a highly efficient procedure for computing the Direct Fourier Transform (DFT) of a time series. It can cut computer time by a factor $(\log_2 N)/N$, where N is the number of terms in each fourier series. Conventional techniques

require complex multiply and add operations proportional to N^2. The FFT makes application of an efficient computational technique called decimination in time which has been analyzed by Cochran.[8]

The choice of a similarity detecting algorithm should be justified by its probability of error and its computational complexity rather than by tradition. The reason generally given for using the correlation method is that correlation appears to be a natural solution for the mean-square error criteria. Therefore, correlation algorithms with lower computational complexity appear to be a more fitting choice.[9]

TABLE I

RUNNING TIME VERSUS CORRELATION ARRAY SIZE[10]

	Running Time in 60^{-1} Seconds/Times Slower Than Binary Correlation		
Array Size	Binary Correlation	FFT Correlation	Correlation
25 by 25	58	585/10.09	2450/42.24
29 by 29	124	1267/10.22	4465/36.01
33 by 33	230	983/ 4.27	7532/32.75
37 by 37	381	1778/ 4.67	11962/31.4
41 by 41	559		18097/32.37
45 by 45	850		
49 by 49	1167		
53 by 53	1710		
57 by 57	2377		
81 by 81	8371		
101 by 101			

C. Binary Cross Correlation Introduction

A method which generates a correlation function by the con-

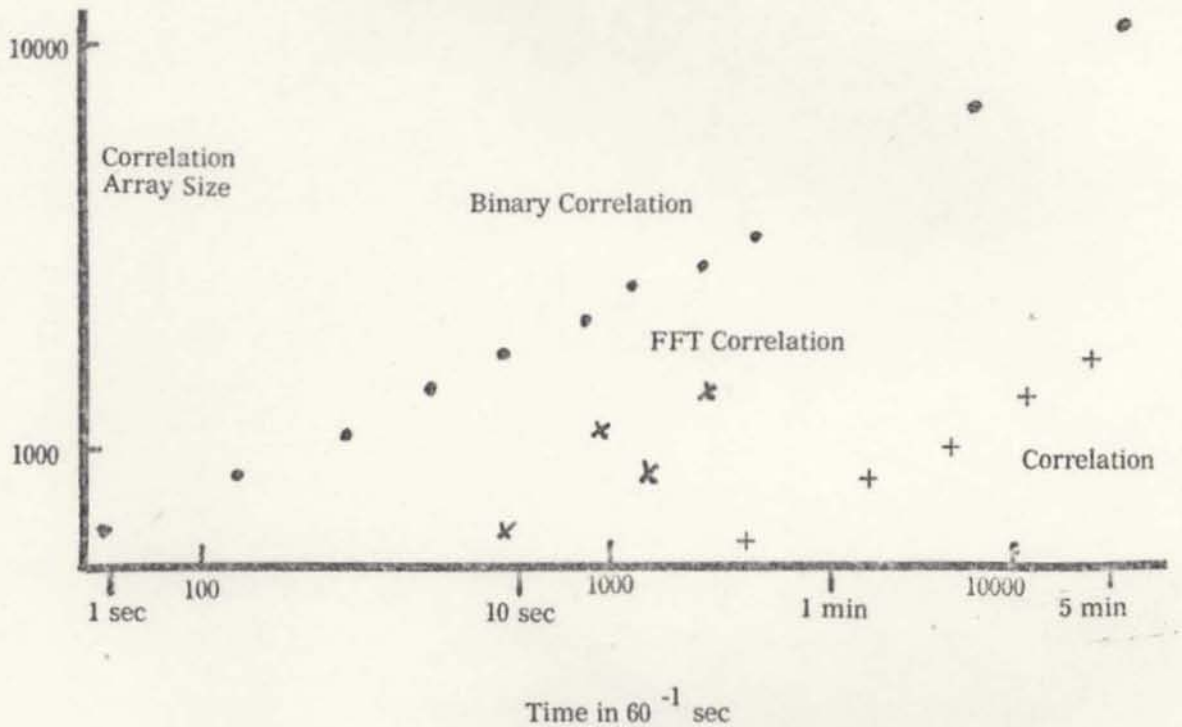


Fig. 2 Correlation array size versus running time [13]

volution of binary templates which have been compressed from the original digitized K gray scale images is known as binary cross correlation. This technique still guarantees a unique solution but is much simpler in terms of computational complexity. Binary images are formed by preprocessing both the data base reference map and the incoming satellite image to form binary (1 or 0) valued images. The assignment of image pixels is based on a predetermined signature, i.e., edges, contours, spatial frequency, texture, etc. The preprocessing algorithm used, depends on the statistics of the data contained in the satellite image and is further discussed in chapter two. The general equation (5) still applies, but S and W are now only binary valued.

Figure 2 is a graph of correlation array size (L=2M per Fig. 1) versus running time for three basic correlation methods (see Table I). The comparison was made by running times from an internal running clock on an IBM-7044 computer. It was demonstrated by Andrus[11] that the binary correlation scheme is superior to FFT #7 and direct correlation with respect to actual running time. Also, the binary data base map reduces memory storage requirements compared to the required by a K level gray scale image.

The following is a summary of the advantages of the binary correlation scheme:[12]

1. Binary maps can represent the time independent shapes of ground features, whereas, the image qualities of tone or gray level and texture can change.

dramatically with time.

2. Binary maps compress the data used as input for statistical correlation of temporal pairs resulting in

   a. Minimizing confusion caused by extraneous tone and texture data which would give equal weight in the decision making of a gray level registration system.

   b. Minimizing computer storage and processing time.

The correlation probability for the discrete case is then described by[14]

$$\rho_{ws_n}(U,V) = \frac{\sum_{j=i}^{1800} \sum_{i=1}^{2048} W(i,j)\, S(U-1+i,V-1+j)}{\sum_{j=i}^{1800} \sum_{i=1}^{2048} [S^2(U-1+i,V-1+j)]^{1/2}} \qquad (6)$$

which is of the form of a discrete normalized cross correlation of equation (5).

Since the data is to be registered at the input frame rate, each incoming processed pixel, W(i,j), will be matched against the elements in the expected search range and each correlation point thus updated sequentially. This point will appear clearer after further discussion in chapter two.

D. Registration by Subframe Partitioning

As the dislocation U*V* is derived empirically, it is proposed that the incoming image W(i,j) be matched in subframes thus giving n U*V* estimates for consideration. This yields the opportunity of applying a least squares solution to the n subframe estimates U*V* giving a rotation estimate for the entire test image (see fig. 3). This method will also reduce the correlation function memory size required and will also increase throughput by allowing correlation and search algorithms to be executed simultaneously.

Equation (6) becomes:

$$\rho_{ws_n}(U,V) = \frac{\sum_{j=100(n-1)+1}^{100n} \sum_{i=1}^{2048} W(i,j)\, S(U-1+i,V-1+j)}{\sum_{j=100(n-1)+1}^{100n} \sum_{i=1}^{2048} [S^2(U-1+1,V-1+j)]^{\frac{1}{2}}} \qquad (7)$$

where each frame consists of 100 lines (see Chapter II) and hence 18 $\rho_{ws_n}(U,V)$ are calculated for the full 1800 lines.

To accomplish the digital processing efficiently requires careful attention to the hardware and software design. Because of the large amount of computer time involved in image processing, consideration should be given to dedicated hardware. The next chapter will evaluate an actual registration problem with the intent of producing a dedicated correlator to accomplish the task.

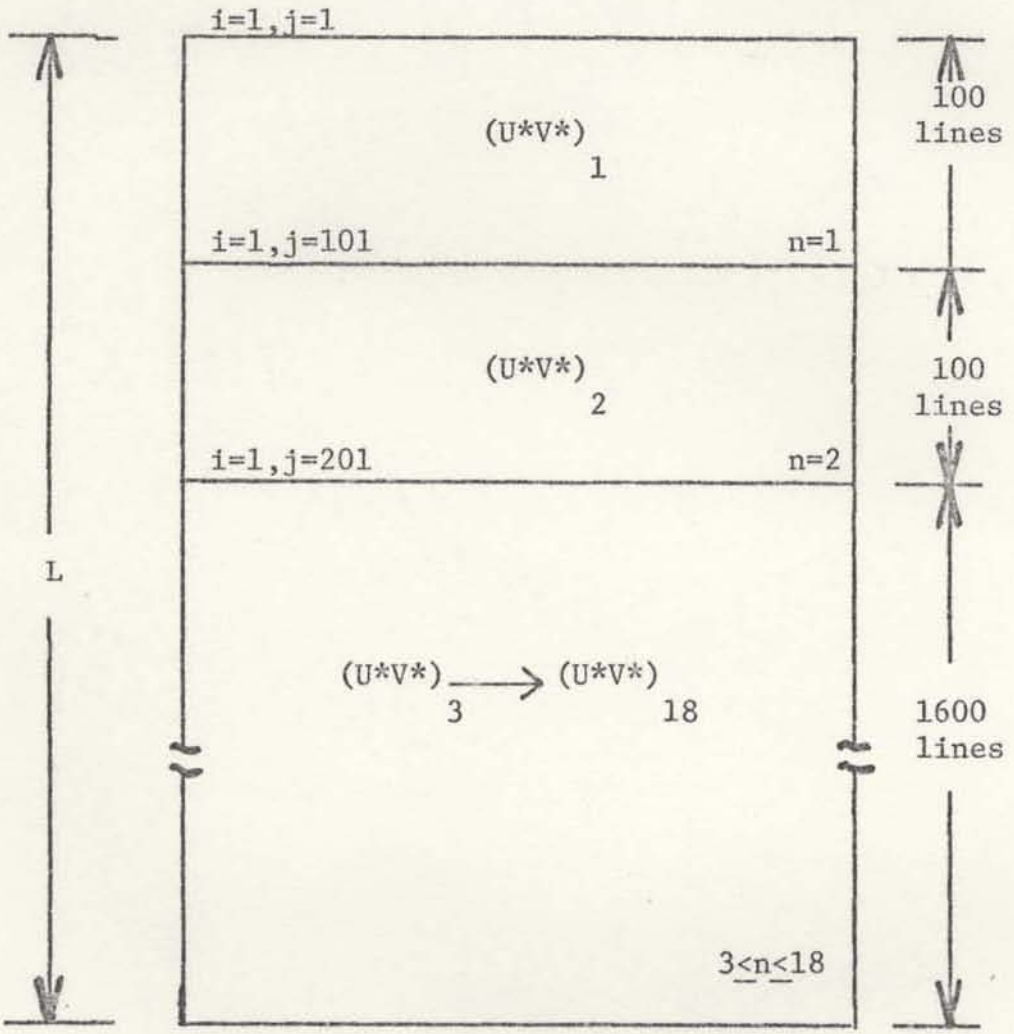


Fig. 3 Image frame sub-partitioning

CHAPTER II

PROBLEM SPECIFICATION

A. Preprocessing and Format of Input Data

A brief discussion of the classification and preprocessing concepts is necessary so that the nature of the data available for determination of the correlation surface is understood. As discussed in chapter one, a binary correlation scheme will be chosen for implementation. It is first necessary to segment the picture into specific regions or parts. In order to extract an object from a picture explicitly, we must somehow mark the points that belong to the object in a special way. This marking process can be thought of as creating a mask or overlay, congruent with the picture, in which there are marks at positions corresponding to object points. We can regard this overlay as a two-valued picture (e.g., 1's at object points, 0's elsewhere); the overlay thus has the value 1 for points belonging to the object and value 0 elsewhere. (There are many other ways of representing objects in regions, e.g., using the boundaries or skeletons).

Two basic methods:[15] (1) thresholding, (2) edge detection exist to segment an image map. Thresholding techniques are primarily designed to extract objects that have characteristic gray level ranges or textures; in other words, they yield objects

that have some type of uniformity. Another important approach to picture segmentation is based on the detection of discontinuity, i.e., of places where there is more or less abrupt change in gray level or in texture, indicating the end of one region and the beginning of another. For low altitude high resolution satellite images (Landsat), binary boundary maps are sucessfully used.[16] For low resolution high altitude weather satellites (the type considered in this report), the thresholding classification scheme where the surface terrain can be classified as either land or water is more appropriate. Fine structured high spatial frequency data is usually not present in GOES data, as can be seen from examination of a GOES weather satellite image in Appendix C. If an image of this type is classified by an intensity amplitude thresholded preprocessing scheme, then large uniform features may be extracted for matching. As shown in Appendix C, a t_1 is first selected which detects clouds, then a t_2 is selected which extracts or detects land mass. A resultant binary image may then be classified into large uniform objects with land = 1 and water = 0. Furthermore, matching two GOES weather maps which have been preprocessed by the thresholding technique will generate a low frequency correlation function. This will allow a sampled correlation algorithm to be used which will greatly reduce computation time. For the above rationale, the selection of thresholded GOES weather satellite images was chosen as input data for the binary correlation algorithm discussed in this report.

The output from the preprocessor will consist of a 2 bit word which will contain the following information where W(i,j) has a gray level range (z_1, z_k) and t is any number between z_1 and z_k:

$$\underline{\text{bit } 0} = W_{t1(i,j)} = \begin{matrix} 0 \text{ if } W(i,j) \\ 1 \text{ if } W(i,j) \end{matrix} = \begin{matrix} \text{t1---cloud or noise} \\ \text{t1---data noise free} \end{matrix}$$

$$\underline{\text{bit } 1} = W_{t2(i,j)} = \begin{matrix} 1 \text{ if } W(i,j) \\ 0 \text{ if } W(i,j) \end{matrix} = \begin{matrix} \text{t2---land} \\ \text{t2---water} \end{matrix}$$

Bit 0 indicates the presence of a noise pixel (cloud). As seen in Appendix C, cloud pixels are usually of the highest intensity and can be easily detected by picking some threshold t1 above t2. If cloud pixels are detected, the input data, W(i,j), will not be considered in calculation of the correlation function array. The new data, W(i,j), will be available at the input data rate, but delayed due to preprocessing time. The preprocessed reference image S will have been thresholded by t2 (the same as for W) and will be available "a priori" from the data base memory. Also, the reference image S will contain no noise or cloud pixels.

The following weighting scheme will be used to calculate the values to be incremented to the correlation estimates $R_{ws_n}(U,V)$.

1. If cloud present--no operation (add "0" to $R_{ws_n}(U,V)$ data).
2. If S(U+i-1,V+j-1) and W(i,j) match--increment $R_{ws_n}(U,V)$ by + 1.

3. If S(U+i-1,V+j-1) and W(i,j) do not match--decrement $R_{ws_n}(U,V)$ by 1.

The concept of awarding matches as well as penalizing for mismatches delineates the need to normalize every correlation point since every data base pixel is used for every shift in the search range. The normalizing factor thus becomes equal to the window array size minus the number of noise pixels in the window, and is a constant for each subframe. See example shown in Appendix A2.

B. Timing Requirements

The timing constraint is probably the one most challenging feature of this research report--correlation and subsequent registration at incoming data rate. The satellite image size will hereby be defined as 2048 pixels by 1800 scan lines. The expected dislocation and resultant correlation surface array size can be seen in figure 4. The mean time between sample points is equal to approximately 250 microseconds. The correlation surface array size of $R_{ws_n}(U,V)$ is seen to be 1024 by 1024. Thus, approximately 10^6 correlation data points must be updated every 250 microseconds based on the value comparison between incoming data point W(i,j) and the value of the reference search range S = S(U-1+i,V-1+j) where $1 \leq U,V, \leq 1024$ (the expected search range shifts with each new data pixel W(i,j). The represents a considerable computational task. Let us define the computational interval as T_a/r where T_a = the total time available (250 microseconds) and r = number of computation points (10^6). In this case T_a = 250 microseconds, and

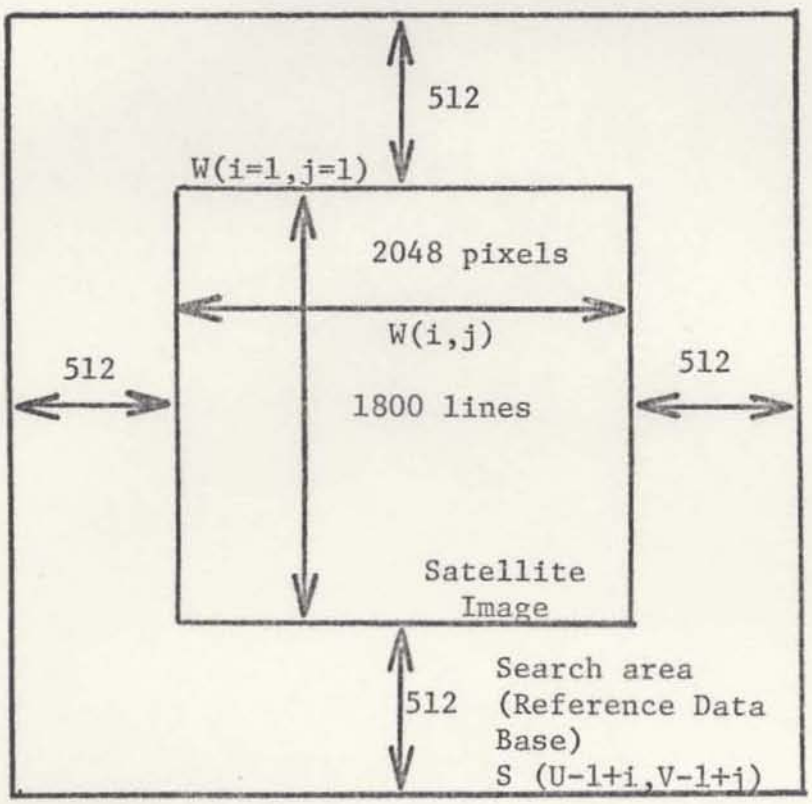


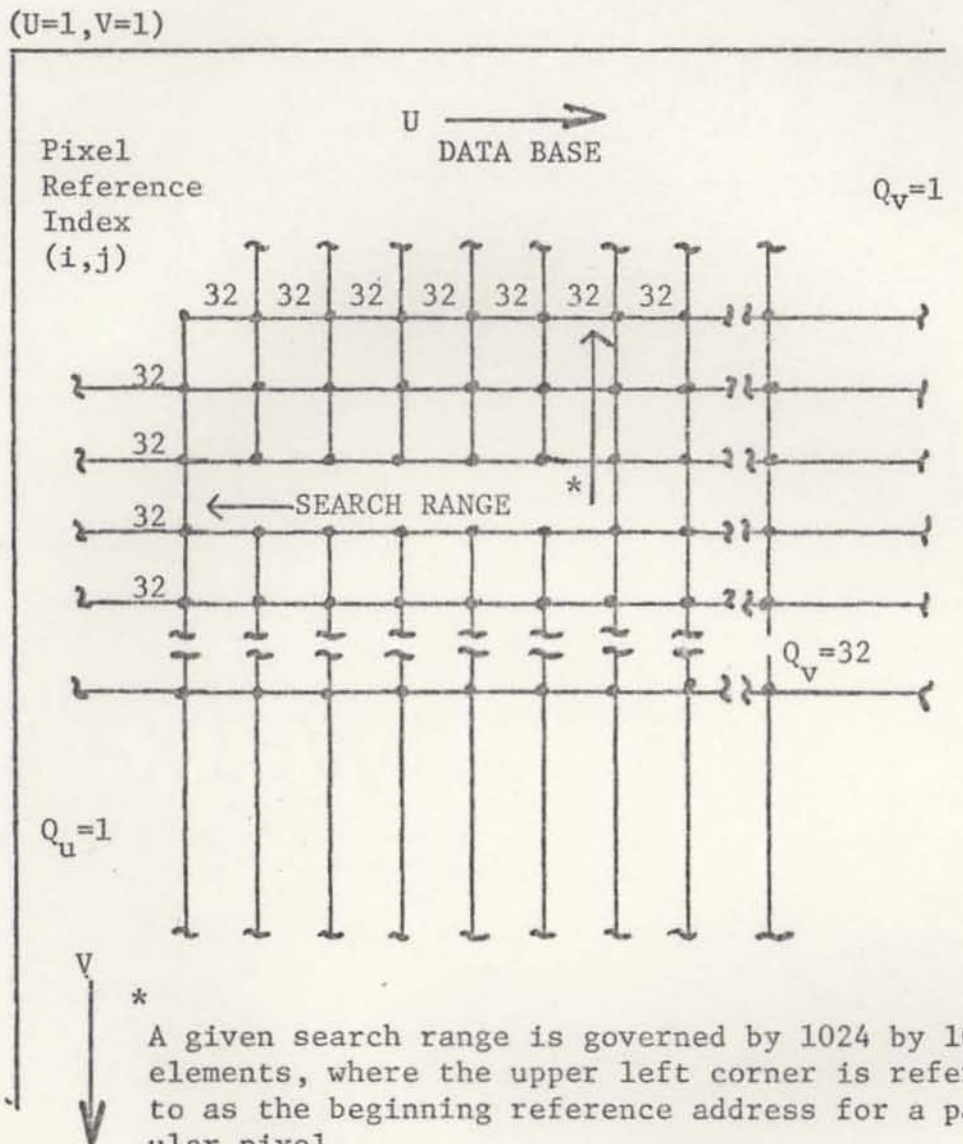
Fig. 4 Satellite image uncertainty

$r = 10^6$ representing a required computational interval of 0.250 nanoseconds. This data rate is unreasonable for any computational method known by this writer (fastest cycle time available being 60 to 100 nsecs). At this point, it becomes necessary to make application of the sampled correlation concept (see Fig. 5). Let us sample the correlation function at every thirty-second data base point. Equation (7) is now replaced by:

$$\text{weighted } \rho'_{ws_n}(U',V') = \frac{\text{weighted } R'_{ws_n}(U',V')}{2048 \times 100\text{-}P_n} = \tag{8}$$

$$\left[\sum_{j=100(n-1)+1}^{100n} \sum_{i=1}^{2048} [W_{t2}(i,j)-\overline{W}_{t2}(i,j)]\,[S(U'-1+i,V'-1+j)-\overline{S}(U'-1+i,V'-1+j)]\,[W_{t1}(i,j)] \right] \Big/ (2048 \times 100\text{-}P_n)$$

where the numerator is the mathematical expression representing 1 of N, parallel computed, course sampled, weighted correlation estimates; Q=the sample cycle; $1 \leq U', V' \leq 1024$; $1 \leq Q_u \leq 32/N$; $1 \leq Q_v \leq 32$; $U' = 32(Q_u-1)N$; $V' = 32(Q_v-1)$; P_n=constant and is the number of noise pixels in W per subframe. The R'_{ws_n} referred to in the rest of this chapter is the weighted function. Equation (8) is employed 18 times as $1 \leq n \leq 18$ as per figure 3. The correlation surface array size is reduced to 1024 (32x32) leaving a 240 nanosecond computational interval. It appears that even when employing a course



Note: The data base memory architecture will be assumed to be configured so that the left most address will unload N horizontal samples to the right inclusively. Each node represents a data base sample.

Fig. 5 Sampled correlation function grid pattern

sampled correlation function concept, a multi-processing scheme must also be considered to realize a solution.

The time allowed per computational interval can be increased by using a multi-processing scheme where the computational interval now becomes $T_a = N/r$ where N = the number of synchronous multi-processors. If N = 16, for example, then the computational interval is equal to $250 \times 10^{-6} \times 16/1024 = 3900$ nanoseconds. If one assumes a 100 nanosecond clock cycle or 100 nanoseconds per instruction (state of current technology), the 3900ns/100ns = 39 instructions that could be performed per computational interval. The data throughput is thus defined as r/T_a.

C. Memory Requirements and Dynamic Range

The memory requirements for storage of the data base reference map requires both large capacity and high performance. The data base map is binary valued (3072 x 2824 bits) and hence, approximately nine megabits must be stored. Also, the data computational interval discussed in the last section must also be divided between memory access time and control of the system resources. The access time allowed per computational interval can be improved by increasing the number of parallel processors. The tradeoff is either using fast memory or increasing the number of parallel computational modules. Due to the extent of the memory required, it is believed by this writer that utilization of fast memory (current RAM memories offer access times of 60ns or less) would invoke the greatest cost to the system. Bubble memories offer bulk storage but do not

present the high performance (access time) requirement needed here. Access time for Bubble memory is 4 microseconds or greater. The configuration of the data base memory and control is not being discussed in this report.

The other significant memory requirement is that for the storage of the correlation array $R'_{ws_n}(U,V)$. This memory requires only 1024 addressable locations. The dynamic range is determined by the condition where the binary images W and S are aligned and 204,800 possible increments are possible. The dynamic range of the correlation function determines that the data bus and memory word length by 18 bits (there are $2^{17} \leq 2048 \times 100 \leq 2^{18}$ possible comparisons per subframe). Note that this memory word size would be increased 18 times if not for the use of the subframe matching concept.

D. Organization of the Required Operations

The flow diagrams for the correlation and max search algorithms are shown in figures 6 and 7. These outline the basic tasks to be done. An analysis of the problem requires organizing the flow chart into two phases: Phase I and Phase II. This is primarily brought about by the relative frequency of operations per subframe. Phase I computes the numerator of the algorithm equation (8), while Phase II searches for the maximum correlation estimate $R'_{ws_n}(U^*,V^*)$. This allows Phase I and Phase II to operate in parallel.

The Phase I of the algorithm waits for a pixel to be received which initiates the correlation array update cycle. N data base

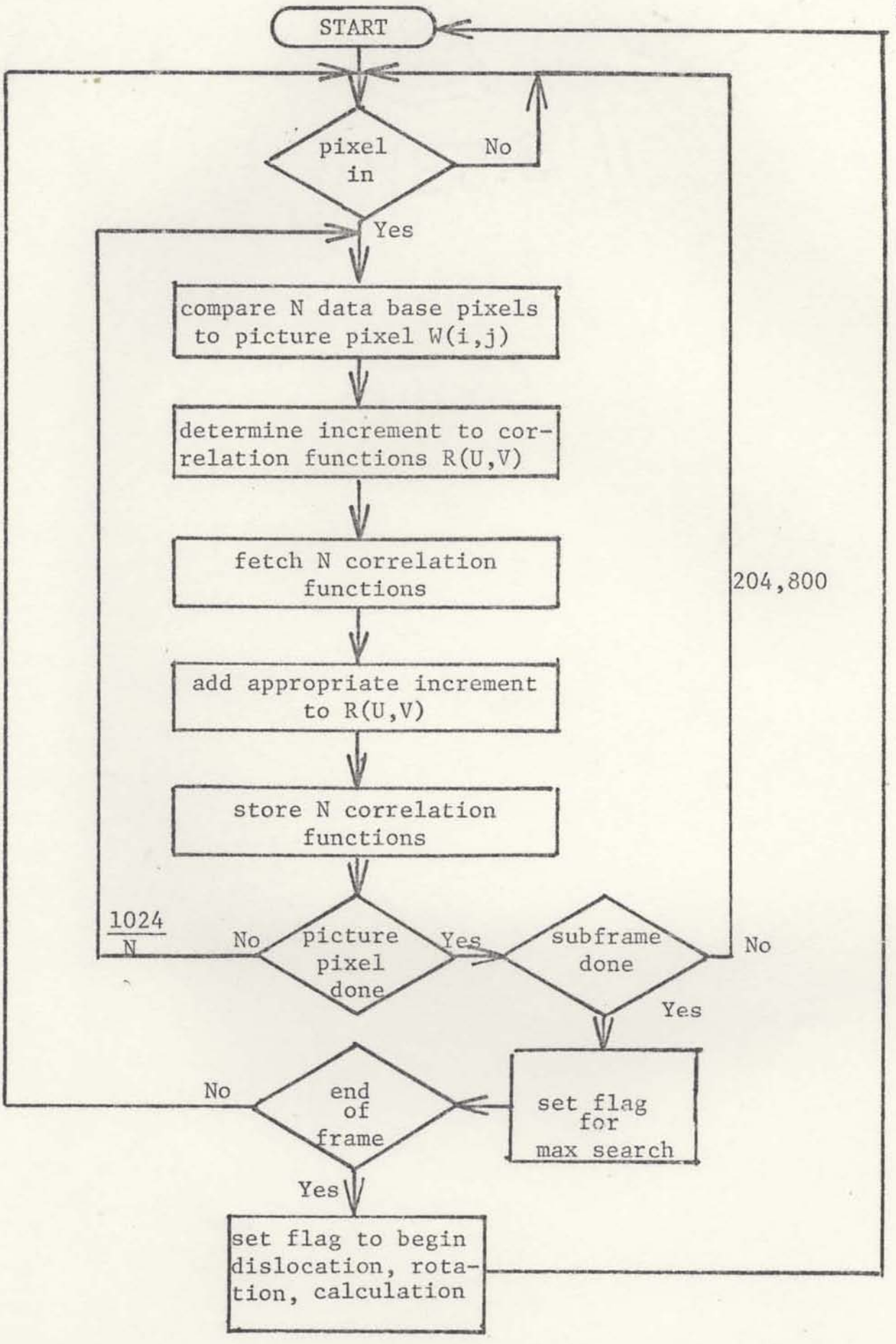


Fig. 6 Flow chart of Phase I, registration scheme

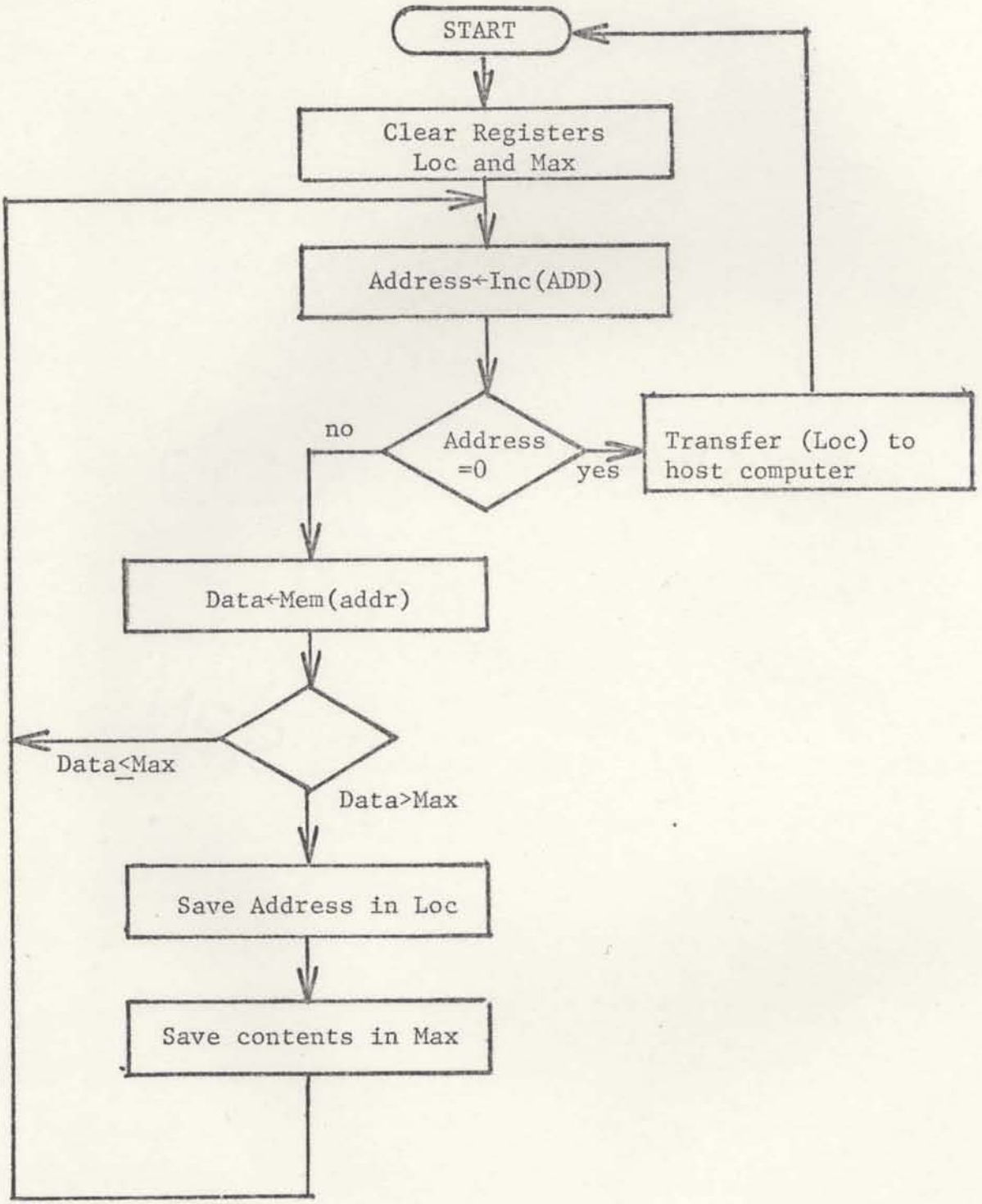


Fig. 7 Flow chart of Phase II, registration scheme

reference pixels are then simultaneously compared with the binary value of the classified input pixel W(i,j). On the basis of this comparison, N correlation estimates $R'_{WS_n}(U,V)$ are fetched, incremented/decremented and stored. This computational cycle is repeated for 2400 x 100 pixels (one subframe) or to the end of each 100 line subframe. At this time Phase II is implemented while the second subframe is in Phase I processing mode. Each Phase II cycle is entered 18 times since the image is subdivided into 18 subframes. The choice of 18 subframes is arbitrary and one could choose some other sub-partitioning scheme, if desired.

Phase II of the algorithm consists of fetching the 1024 correlation estimates from the correlation array memory. Each estimate in turn is compared to the previous estimate. Only the current max estimate is retained in a specified temporary register (max). The memory array address is also retained in a specified temporary register (loc). When the search is complete (loc) contains the address for $R'_{WS_n}(U^*,V^*)$. $R'_{WS_n}(U^*,V^*)$ represents only the approximate determination of the correlation peak (only every thirty-second reference data point is sampled) and resultant sub image translation. The contents of (loc) are transferred to a host computer which calculates the translation estimate $(U^*,V^*)_n$. At the end of the frame, the 18 $(U^*,V^*)_n$ are used by the host computer to calculate the rotation of W with respect to S by fitting the n translation estimates to a straight line and determining slope M.

TABLE II

BASIC TIMING
RELATIVE FREQUENCY OF PHASE I VS PHASE II

Time	C Sec	50 Sec	50 Sec	750 Sec	50 Sec	50 Sec	R Sec
	Delay C Sec	1st 100 Lines Sub Frames	2nd 100 Lines Sub Frames	3rd - 17th 100 Lines Sub Frames 1500 Lines	18th 100 Lines Sub Frames	100 Lines	Post Corr. Cal.
Pre-Processing	All Incoming Data Pts.	0	0	0	0	0	0
Phase I	0	2048 X 100	2048 X 100	2048 X 1500	2048 X 100	0	0
Phase II	0	0	1	1	1	1	0
Post Processing	0	0	0	0	0	0	TBD
Initiate Picture X Mission							

The overall timing and application of the two phase algorithm is shown in Table II. Eighteen, 50 second subframes exist (100 lines x 50 Msec. per line). Phase I is used 204,800 times per subframe. Phase II is really only entered once and performs 1024 computational intervals per subframe.

The fetching, incrementing and storage of correlation estimates suggest a microporcessor based architectural scheme be investigated.

Consideration of the timing and word size requirements of the GOES registration problem suggests a microprogrammable microprocessor and/or special purpose hardware be considered for implementation of the correlation algorithm. The microprogrammable CPU is best applied where high speed computation and control are required. Typical instruction cycle times of 100 ns and 20 bit word lengths are realizable with current Bipolar Schottky LSI Technology.[17] Also, the concept of N parallel computational modules, calculating N correlation estimates is facilitated.

CHAPTER III

SYNOPSIS

A. Summary and Conclusions

The object of this research report was to investigate and select a successful registration concept and consider the possibility of electronically realizing a fast, dedicated correlator for use in registering weather satellite images. In chapter one a comparison of current registration methods was made which indicated that binary correlation would provide the greatest computational savings with respect to speed and still yield the same registration result as correlating two K gray scale images. It was concluded that a binary correlation algorithm should be chosen for implementation. Computing the correlation of n image subframes will yield n translation estimates. This data can be used to determine image frame rotation by least squares solution.

The second chapter examined the input data and quantified the basic parameters of the registration problem which was to be implemented with the binary correlation algorithm. Weather satellite images, which contain low resolution and low spatial frequency data, lend themselves to classification or preprocessing by thresholding techniques. Large uniform features may be extracted and classified. Corrleation surfaces generated from matching two

images which have been preclassified by this technique will yield a low frequency correlation function. Data of this form was chosen as the input for the correlator, enabling the implementation of a fast course (sampled) correlation algorithm. It was decided to generate a course sampled correlation function since the amount of expected image dislocation and the size of the image array were large. These parameters are realistic and representative of many satellite imaging problems today. The correlation array was generated by sampling every thirty-second data base point. This sampling will yield an approximate translation $(U^*,V^*)_n$. In many cases, this may be accurate enough. If registration within one pixel is required, the exact location of the correlation peak may be estimated by assuming a known surface function. Examination of the correlation surface suggests that a two dimensional conical approximation could be used.

The correlation algorithm was organized into two phases. Phase I calculates the correlation estimates for a subframe while Phase II searches for the maximum correlation estimate of the previous subframe and transfers translation data to a host computer. The final translation/rotation and resultant registration would be performed by the host computer that is inputing, processing, sequencing and displaying the images. It is recommended that microprogrammable microprocessor hardware be investigated for implementation of Phases I and II.

B. Extensions

The correlation algorithm discussed in this paper has general application. Although the high speed algorithm produced a coarse or sampled correlation array, it could easily be modified to directly produce a fine or non-sampled correlation function. This would be required if one were sampling, for example, binary edge images processed from Landsat image data which contains high spatial frequency data. In the general case, it appears that many registration problems must be divided into two correlation operations. If one wanted to register Landsat images with considerable dislocation, one might threshold certain large objects to obtain a course registration, and then use binary edge images of the immediate area to obtain the fine registration. Maps where spatial frequency or texture were classified to obtain either a "1" for a certain texture, and "0" for everything else, might be used to provide a low frequency correlation function so that the fast sampling correlator can be used. Different image sizes, dislocation and transmission data rates would only require modifications to memory and I/O sizes, number of computational modules used, memory access time, etc.

The process is also adaptive to determining rotation since the registration is accomplished in 18 subframes. It would also be possible to weight the estimate of each subframe translation with respect to the presence of noise, which determines the likelihood of success. If one test subframe contained excessive noise pixels,

it may be completely disregarded. One subframe may be sufficient to register the image if rotation is not present. In some particular applications, it may be possible to introduce "a priori" information about the image and reduce the amount of memory storage required. For example, one knows the State of Florida is a curved surface and a continuous edge is being sought out. This concept falls under the realm of pattern recognition and is beyond the scope of this paper.

APPENDIX A

MATCHING BY CROSS-CORRELATION[18]

1. Theory

"Matching of patterns against pictures is of importance in many types of applications. The following are a few examples:

(a) The pattern to be matched may be a simple pattern such as a step or ramp, i.e., an edge. Other examples of such simple patterns are lines, spots, etc.

(b) The pattern may be a "template" representing a known object. For example, we can match templates of characters against pictures of printed pages; templates of targets against pictures obtained by reconnaissance systems; or templates of landmarks against pictures obtained by navigation systems, e.g., templates of star-patterns against pictures of the sky.

(c) The pattern may itself be a piece of picture, i.e., we may want to match a piece of one picture against another picture. For example, if we have two pictures of the same scene taken from different viewpoints, and we can identify the parts of the

two pictures that show the same piece of the scene, we can then measure stereoscopic parallax and determine the heights or depths of objects in the scene. Similarily, if we have two pictures taken at different times, we can use this information to measure the relative motions of objects in the scene (cars on a road, clouds in the sky, etc.). If we have two differently distorted pictures of the scene, we can use such piecewise matching to determine the relative distortion, so that it can be corrected if desired."

"There are many possible ways of measuring the degree of match or mismatch between two functions, f and g, over a region ℓ. For example, one can use as mismatch measures such expressions as:

$$\max_{\ell}|f-g| \text{ or } \iint_{\ell}|f-g| \text{ or } \iint_{\ell}(f-g)^2$$

and so on (where the integrals become sums in the digital case). It is easily verified that these expressions are all "distance measures" or metrics.

If we use $\iint(f-g)^2$ as a measure of mismatch, we can derive a useful measure of match from it. Note, in fact, that

$$\iint(f-g) = \iint f^2 + \iint g^2 - 2\iint fg$$

Thus if $\iint f^2 + \iint g^2$ are fixed, the mismatch measure $\iint(f-g)^2$ is large if, and only if, $\iint fg$ is small. In other words, for given $\iint f^2$ and $\iint g^2$ we can use $\iint fg$ as a measure of match.

The same conclusion can be reached by making use of the well-known Cauchy-Schwarz inequality, which states that for f and g non-negative, we always have

$$\iint fg \leq \sqrt{\iint f^2 \iint g^2}$$

with equality holding if and only if g = cf for some constant. (The analogous result in the digital case is

$$\sum_i \sum_j f(i,j)\, g(i,j) \leq \sqrt{\sum_i \sum_j f(i,j)^2 \sum_i \sum_j g(i,j)^2}$$

with equality holding if and only if g(i,j) = cf(i,j) for all i,j). Thus when $\iint f^2$ and $\iint g^2$ are given, the size of $\iint fg$ is a measure of the degree of match between f and g (up to a constant factor).

Suppose now that if f is a template, g a picture, and we want to find pieces of g that match f (We are tacitly assuming that f is small compared to g, i.e., f is zero outside a small region ℓ, and we are interested only in matching the nonzero part of f against g.), we can do this by shifting f into all possible positions relative to g, and computing $\iint fg$ for each such shift (U,V). By the Cauchy-Schwarz inequality, we have

$$\iint_\ell f(x,y)\, g(x+U,y+V)\,dxdy \leq \sqrt{\iint_\ell f^2(x,y)\,dxdy \iint_\ell g^2(x+U,y+V)\,dxdy}$$

Since f is zero outside ℓ the left-hand side is equal to

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, g(x+U,y+V)\,dxdy$$

which is just the cross-correlation C_{fg} of f and g.

Note that on the right-hand side, while $\iint f^2$ is a constant $\iint g^2$ is not, since it depends on U and V. Thus we cannot simply use C_{fg} as a measure of match; but we can use instead the normalized cross correlation

$$C_{fg} / \sqrt{\iint_{\ell} g^2(x+U,y+V)\,dxdy}$$

This quotient takes on its maximum possible value (namely, $\sqrt{\iint_{\ell} f^2}$) for displacements (U,V) at which g = cf."

2. Consider the following search area S and Window W (f=W, g=S per A.1).

(U+i-1) =	1	2	3	4	5	6	7	8
(V+j-1)=1	0	0	0	0	0	0	0	0
2	0	0	0	1	1	0	0	0
3	0	0	0	0	0	0	0	0
4	0	1	0	1	1	0	1	0
5	0	1	0	1	1	0	1	0
6	0	0	0	0	0	0	0	0
7	0	0	0	1	1	0	0	0
9	0	0	0	0	0	0	0	0

S

i =	1	2	3	4
j=1	0	0	0	0
2	0	1	1	0
3	0	1	1	0
4	0	0	0	0

W

The following arrays result for the normalized cross-correlation described in A.1:

(a)

	U=1	2	3	4	5
V=1	0	1	2	1	0
2	1	1	2	1	1
3	2	2	4	2	2
4	1	1	2	1	1
5	0	1	2	1	0

$R(U,V)=W*S$

(b)

U=1	2	3	4	5
3	5	4	5	3
5	8	6	8	5
4	6	4	6	4
5	8	6	8	5
3	5	4	5	3

$\Sigma\Sigma S^2$

(c)

	U=1	2	3	4	5
V=1	$\sqrt{3}$	$\sqrt{5}$	$\sqrt{4}$	$\sqrt{5}$	$\sqrt{3}$
2	$\sqrt{5}$	$\sqrt{8}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{5}$
3	$\sqrt{4}$	$\sqrt{6}$	$\sqrt{4}$	$\sqrt{6}$	$\sqrt{4}$
4	$\sqrt{5}$	$\sqrt{8}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{5}$
5	$\sqrt{3}$	$\sqrt{5}$	$\sqrt{4}$	$\sqrt{5}$	$\sqrt{3}$

$\sqrt{\Sigma\Sigma S^2}$

(d)

	U=1	2	3	4	5
V=1	0.00	0.45	1.00	0.45	0.00
2	0.45	0.35	0.82	0.35	0.45
3	1.00	0.82	2.00	0.82	1.00
4	0.45	0.35	0.82	0.35	0.45
5	0.00	0.45	1.00	0.45	0.00

$2\rho(U,V)$

(e)

U =1	2	3	4	5
0.00	0.22	0.50	0.22	0.00
0.22	0.17	0.41	0.17	0.22
0.50	0.17	1.00	0.41	0.50
0.22	0.17	0.41	0.17	0.50
0.00	0.22	0.50	0.22	0.00

$\rho(U,V)$

Note $\sqrt{\Sigma\Sigma W^2} = 2$

The following functions or arrays for the case where R(U,V) is incremented + 1 for a match and decremented by 1 for no match result:

(f)

	U =1	2	3	4	5
V=1	2	2	8	2	2
2	2	-4	4	-4	2
3	8	4	16	4	8
4	2	-4	4	-4	2
5	2	2	8	2	2

(g)

	U =1	2	3	4	5
V=1	1/8	1/8	1/2	1/8	1/8
2	1/8	-1/4	1/4	-1/4	1/8
3	1/2	1/4	1	1/4	1/2
4	1/8	-1/4	1/4	-1/4	1/8
5	1/8	1/8	1/2	1/8	1/8

weighted $R(U,V)=W*S+\overline{W}*\overline{S}-\overline{W}*S-W*\overline{S}=(W-\overline{W})*(S-\overline{S})$

$\rho(U,V)=R(U,V)/16$

Locations of W with respect to S where little or no correlation exists, but random matches of W and S occur, would be averaged to 0 by the subtraction of the W and S mismatch occurrences. The need to normalize every point in the R(U,V) array is hence eliminated. The normalization is thus performed to provide an upper bound of 1. The R(U,V) calculated in (f) is referred to as the weighted correlation function in chapter two. Note, $\overline{W}$, $\overline{S}$ refer to the compliment of the respective image W, S; i.e., 0's are replaced by 1's and 1's are replaced by 0's.

APPENDIX B

The following will illustrate that the Fast Fourier Transform can save considerable calculations when performing a correlation operation.

1. If A_k is a sequence of terms such as a digitally sampled function of N terms, then by definition the jth Fourier coefficient x(j) of the sequence is:

$$x(j) = 1/N \sum_{k=0}^{N-1} A(k)\, W^{jk} \quad \text{for } j = 0,1,\ldots N-1.$$

where $W = e^{2\pi i/N}$, the A(k)'s are the complex coefficients of the sequence A_k, and N is the number of terms in the sequence. If the x(j)s are expanded in the order indicated by the right side of the above equation, then N^2 operations are required. The above descirbes the DFT or Direct Fourier Transform of a time series. If the expansion is regrouped as described by Cochran[19], then the DFT calculation leads to a number of complex additions and multiplications which is proportional to $N\log_2 N$. This is considerably fewer operations than would be required by expanding the above equation directly. However, $N=2^m$, and N = large must be satisfied to realize computational savings. The following examples estimate the relative number of operations required to convolve two sequences

of length N by direct convolution, DFT and FFT. The results are approximate and serve to scope the operations required for comparative purposes. The one dimensional analysis can be extended to two dimensional applications and is discussed by Cochran.[20]

2. If one wished to convolve two sampled functions of N=16., The following operations result:

Direct Convolution

Multiply and shift 2 $\sum_{n=1}^{N} n = \frac{2N(N+1)}{2}$ 272 multiply-adds

DFT

$f_1 \rightarrow f_1, f_2 \rightarrow F_2$	$2xN^2$	=	512
$F_1 x F_2$		+	16
$f_1 * f_2 \leftarrow F_1 x F_2$	N^2	=	+ 256
			784 multiply-adds

FFT

$f_1 \rightarrow F_1, f_2 \rightarrow F_2$	$2x2N\log_2 N$	=	256
$F_1 x F_2$		+	16
$f_1 * f_2 \leftarrow F_1 x F_2$	$2N\log_2 N$	=	+128
			400 multiply-adds

3. The following example illustrates the condition when N=256.

Direct Convolution

Multiply and shift 2 $\sum_{n=1}^{N} n = \frac{2N(N+1)}{2}$ = 65,792 multiply-adds

DFT

$f_1 \rightarrow F_1, f_2 \rightarrow F_2$	$2xN^2$	131,072
$F_1 x F_2$		+ 256
$f_1 * f_2 \leftarrow F_1 x F_2$	N^2	+ 65,536
		196,864 multiply-adds

FFT

$f_1 \rightarrow F_1, f_2 \rightarrow F_2$	$2x2N\log_2 N$	8,192
$F_1 x F_2$		256
$f_1 * f_2 \leftarrow F_1 x F_2$	$2N\log_2 N$	4,096
		12,544 multiply-adds

4. Summary

It becomes apparent from the above example that the DFT would never be practical to use, and that FFT offers significant computational savings for large sequences. Also, when convolving two sequences by use of the DFT or FFT computational technique, the vector/array of f_1 and f_2 must be equal.

APPENDIX C

Photo 1 is a typical image frame received from the Geostationary Operating Environmental Satellite (GOES). The scene was provided by a scanning radiometer aboard the satellite which is a transducer sensitive to reflected radiation in the visible band. White represents the brightest optical intensity, and black the lowest. As the scene is scanned, it is seen that the land mass is characterized by a difference in texture and optical intensity than the ocean. The land features can then be classified or identified by the relative difference of optical intensity between land and water. Likewise, clouds are represented by the highest optical intensity in the scene and are distinguishable from both land or water.

The original images was input using a TV camera. Statistical data for determining threshold t_1 and t_2 were measured using the GE IMAGE 100 Processing System. Figures 8 and 9 are image intensity histograms with t_1 and t_2 selected. Classification or selection of the thresholds was made by the visual identification of modes (peaks) in the image intensity histograms and the correspondence with the original image was verified by displaying pixels associated with the particular mode on the system CRT (Photo 1 and 2).

Photo 1. Typical imagery from Geostationary Operating Satellite (½ mile resolution).

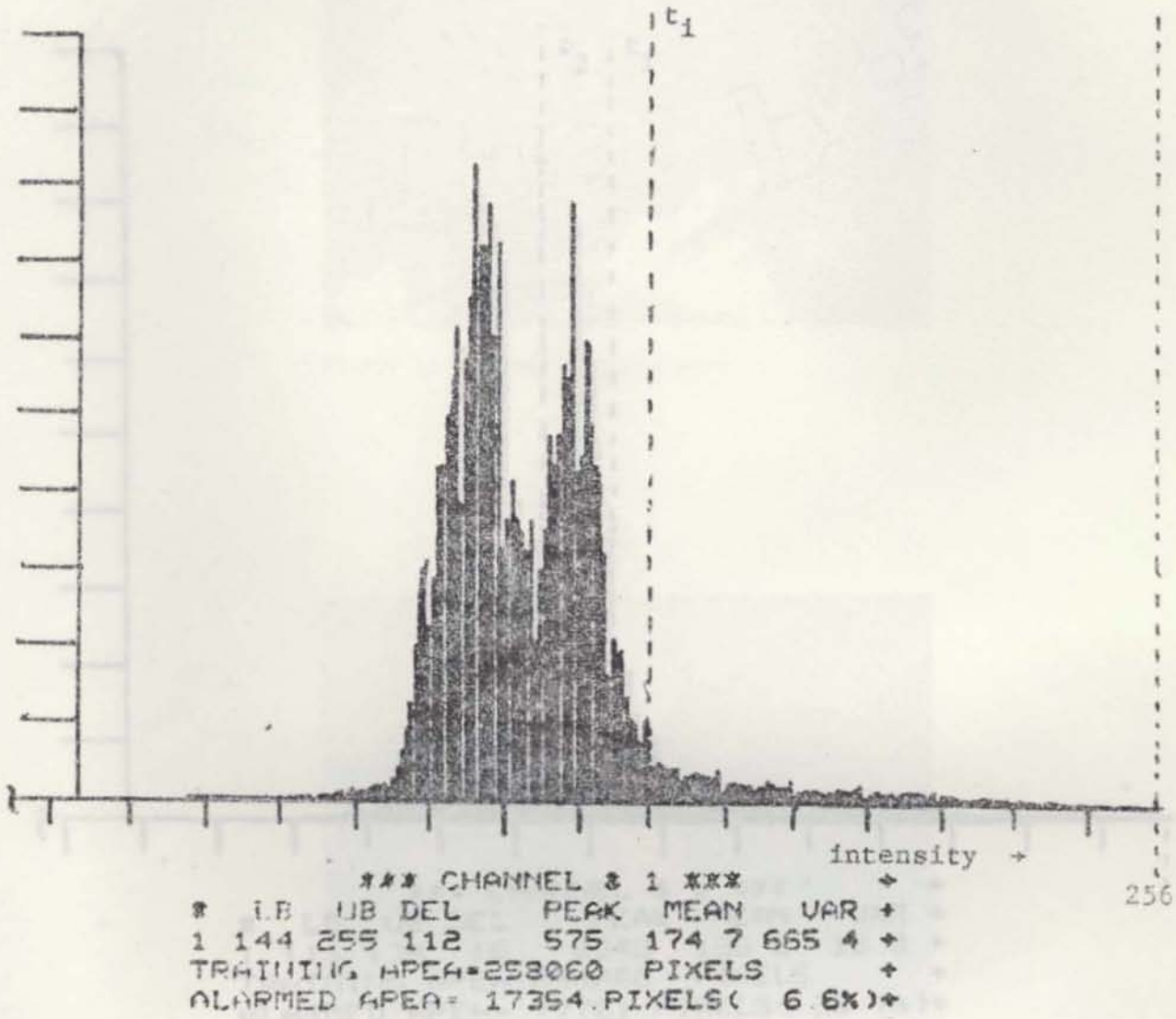


Fig. 8 GOES histogram of photo 1 with threshold 1 selection

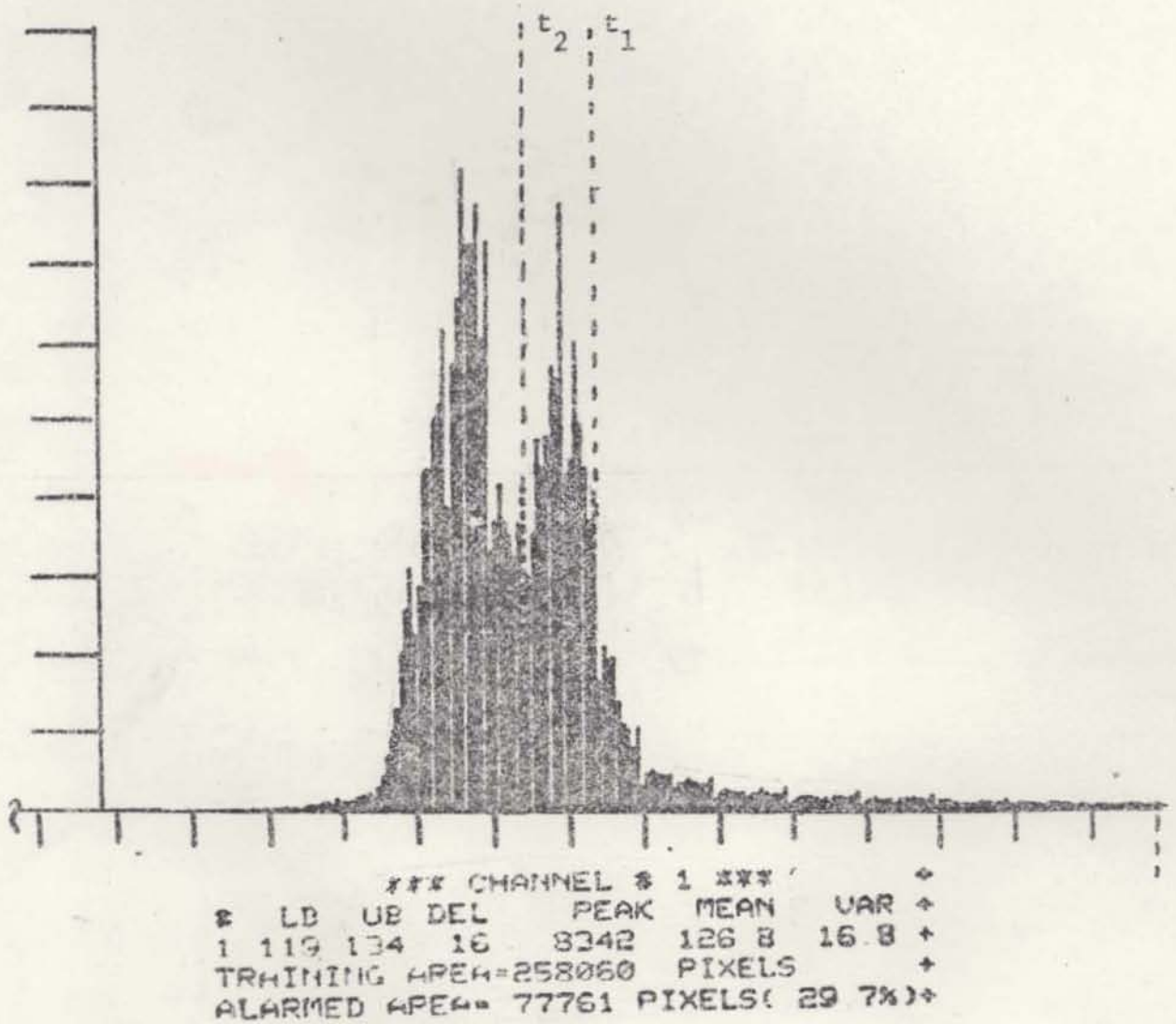


Fig. 9 GOES histogram of photo 1 with threshold $t_2 < t_1$

Photo 2. Cloud detection with t_1

Photo 2. Land detection with t_2

SYMBOL LIST

N: number of parallel processors.

R: cross correlation.

R_{ws}: cross correlation generated by the convolutions W and S.

ρ: normalized cross correlation.

ρ_{ws}: normalized cross correlation of R_{ws}.

S: search area or reference picture.

W: a subimage of S or piece of picture which matches S at (U*,V*).

(x,y): coordinates of analog subimage W.

(i,j): array of coordinates of digital picture elements of W.

(U,V): coordinates of W shifted relative to S.

(U*,V*) : coordinates of the maximum value of a correlation array

n: subframe number of W.

R_{ws_n}: cross correlation array generated by the convolution nth subframe of W and the search range of S.

ρ_{ws_n}: normalized R_{ws_n} correlation array.

R'_{ws}: course sampled cross correlation function R_{ws} generated by the convolution of W and every 32nd pixel of S in the search range.

ρ'_{ws}: normalized R'_{ws} correlation array.

R'_{ws_n}: course sampled cross correlation function of R_{ws_n} generated by the convolution of the nth subframe of W and every 32nd data pixel of S in the search range.

ρ'_{ws_n}: normalized R'_{ws_n} correlation array.

$(U^*,V^*)_n$: coordinates of the maximum value of the nth subframe correlation array

T_a: time available between input data samples.

r: number of data points or pixels in the search range.

Q_u: horizontal sample cycle.

Q_v: vertical sample cycle.

weighted R'_{ws_n}: the weighted cross correlation is a function which increments $R'_{ws_n}(U,V)$ for a match of W and S, and decrements $R'_{ws_n}(U,V)$ for a mismatch of W and S. The mathematical expression which represents this function is $(W-\bar{W})$ convolved with $(S-\bar{S})$. The concept can be thought of as virtually convolving two bi-valued images where land=1 and water=-1. The resulting correlation array does not require normalization of every point in order to determine the location of the peak. Locations of W with respect to S where little or no correlation exists would produce a function output equal to 0 since the sum of the matches between W and S which occur would be random.

weighted ρ'_{ws_n}: R'_{ws_n} divided by the expected dynamic range of the correlation array.

REFERENCES

[1] Harold W. Yates and William R. Bandeen, "Meteorological Applications of Remote Sensing From Satellites," Proceedings of IEEE 63 (January 1975): 148-163.

[2] Laboratory for Applications of Remote Sensing, Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, (Purdue University, West Lafayette, IN: 1975), p. 3A-12.

[3] Ibid., pp. 3A-12.

[4] Earth Resources Technology Satellite Association, Proceedings of the Third Symposium, (Menlo Park, CA: 1978), pp. 1725-1741.

[5] D. I. Barnea and H. F. Silverman, "A Class of Algorithms for Fast Digital Image Registration," IEEE Transactions on Computers C-21 (February 1972): 179-180.

[6] Ibid., p. 180.

[7] George R. Cooper and Clare D. McGillen, Probabilistic Methods of Signal and System Analysis (New York: Holt, Rinehart and Winston, 1971), pp. 82-83.

[8] William T. Cochran, et al, "What is the Fast Fourier Transform?" IEEE Transactions on Audio and Electroacoustics AU-15, No. 2 (June 1967): 45.

[9] D. I. Barnea and H. F. Silverman, "A Class of Algorithms for Fast Digital Image Registration," IEEE Transactions on Computers C-21 (February 1972): 179-180.

[10] R. R. Jayroe, J. F. Andrus, and L. W. Campbell, Digital Image Registration Method Based Upon Binary Boundary Maps, NASA TND-7607 (Marshall Space Flight Center, AL: Government Printing Office, 1974), p. 18.

[11] Ibid.

[12] Ibid., p. 15.

13 American Institute of Astronautics and Aeronautics, Proceedings of the Astrodynamics Specialist Conference, (Nassau, Bahamas: 1975), pp. 5-6.

14 William K. Pratt, Digital Image Processing (New York: Wiley-Interscience, 1978), p. 562.

15 A. Rosenfeld and A. C. Kak, Digital Picture Processing (New York: Academic Press, 1976), pp. 256-296.

16 American Institute of Astronautics and Aeronautics, Proceedings of the Astrodynamics Specialist Conference, (Nassau, Bahamas: 1975), pp. 1-6.

17 George Reyling, Jr., "Considerations in Choosing a Microprogrammable Bit-Sliced Architecture," Computer (July 1974): 81-84.

18 A. Rosenfeld and A. C. Kak, Digital Picture Processing (New York: Academic Press, 1976), pp. 296-302.

19 William T. Cochran, et al, "What is the Fast Fourier Transform?" IEEE Transaction on Audio and Electroacoustics, AU-15, No. 2, (June 1967): 45-55.

20 Ibid.

BIBLIOGRAPHY

American Institute of Astronautics and Aeronatuics. Proceedings of the Astrodynamics Specialist Conference. Nassau, Bahamas, 1975.

Barnea, D. I. and Silverman, H. F. "A Class of Algorithms for Fast Digital Image Registration." IEEE Transactions on Computers C-21 (February 1972): 179-180.

Cochran, William T., et al, "What is the Fast Fourier Transform?" IEEE Transactions on Audio and Electroacoustics AU-15, No. 2 (June 1967): 45-55.

Cooley, James W.; Lewis, Peter A. W.; and Welch, Peter D. "Application of the Fast Fourier Transform to Computation of Fourier Integrals, Fourier Series, and Convolution Integrals." IEEE Transactions on Audio and Electroacoustics AU-15, No. 2 (June 1967): 79-84.

Cooper, George R., and McGillen, Clare D. Probabilistic Methods of Signal and System Analysis. New York: Holt, Rinehart, and Winston, 1971.

Dietmeyer, Donald L. Logic Design of Digital Systems. Boston: Allyn and Bacon, 1972.

Earth Resources Technology Satellite Association. Proceedings of the Third Symposium. Menlo Park, CA: 1972.

Jayroe, R. R., Andrus, J. F., and Campbell, L. W. Digital Image Registration Method Based Upon Binary Boundary Maps. NASA TND-7607. Marshall Space Flight Center, AL: Government Printing Office, 1974.

Laboratory for Applications of Remote Sensing. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data. Purdue University, West Lafayette, IN: 1975.

Moik, J. G. Introduction to Computer Image Processing. NASA TMX-70689 Greenbelt, MD: Goddard Space Flight Center, 1973.

Nack, M. C. Final Report on Image Registration. NAS 5-11999, Task No. 699, Silver Spring, MD: Computer Science Corp., 1976.

Patz, Ben W. Class Discussion and Notes. Orlando, FL: Florida Technological University, May 1977.

Pratt, William K. Digital Image Processing. New York: Wiley-Interscience, 1978.

Rattner, Justin. Building Block Microprocessors. Santa Clara, CA: Intel Corporation Micro Computer Systems, 1974.

Reyling, Jr., George. "Considerations in Choosing a Micro-programmable Bit-Slice Architecture." Computer (July 1974): 81-84.

Rosenfeld, A. and Kak, A. C. Digital Picture Processing. New York: Academic Press, 1976.

Simpson, Richard S. and Houts, R. C. Fundamentals of Analog and Digital Communications Systems. Boston: Allyn and Bacon, 1971.

Soucek, B. Microprocessors and Microcomputers. New York: John Wiley and Sons, 1976.

Yates, Harold W. and Bandeen, William R. "Meteorological Applications of Remote Sensing from Satellites." Proceedings of the IEEE 63, No. 1 (January 1975): 148-163.