FINAL TECHNICAL REPORT

ANALYSIS OF PRENORMALIZATION TECHNIQUES

FOR

IMAGERY SCREENING

[redacted] 25X1

[redacted] 25X1

29 APRIL 1966

Prepared by:

[redacted] 25X1

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

FINAL TECHNICAL REPORT

ANALYSIS OF PRENORMALIZATION TECHNIQUES

FOR

IMAGERY SCREENING

[redacted] 25X1

[redacted] 25X1

29 APRIL 1966

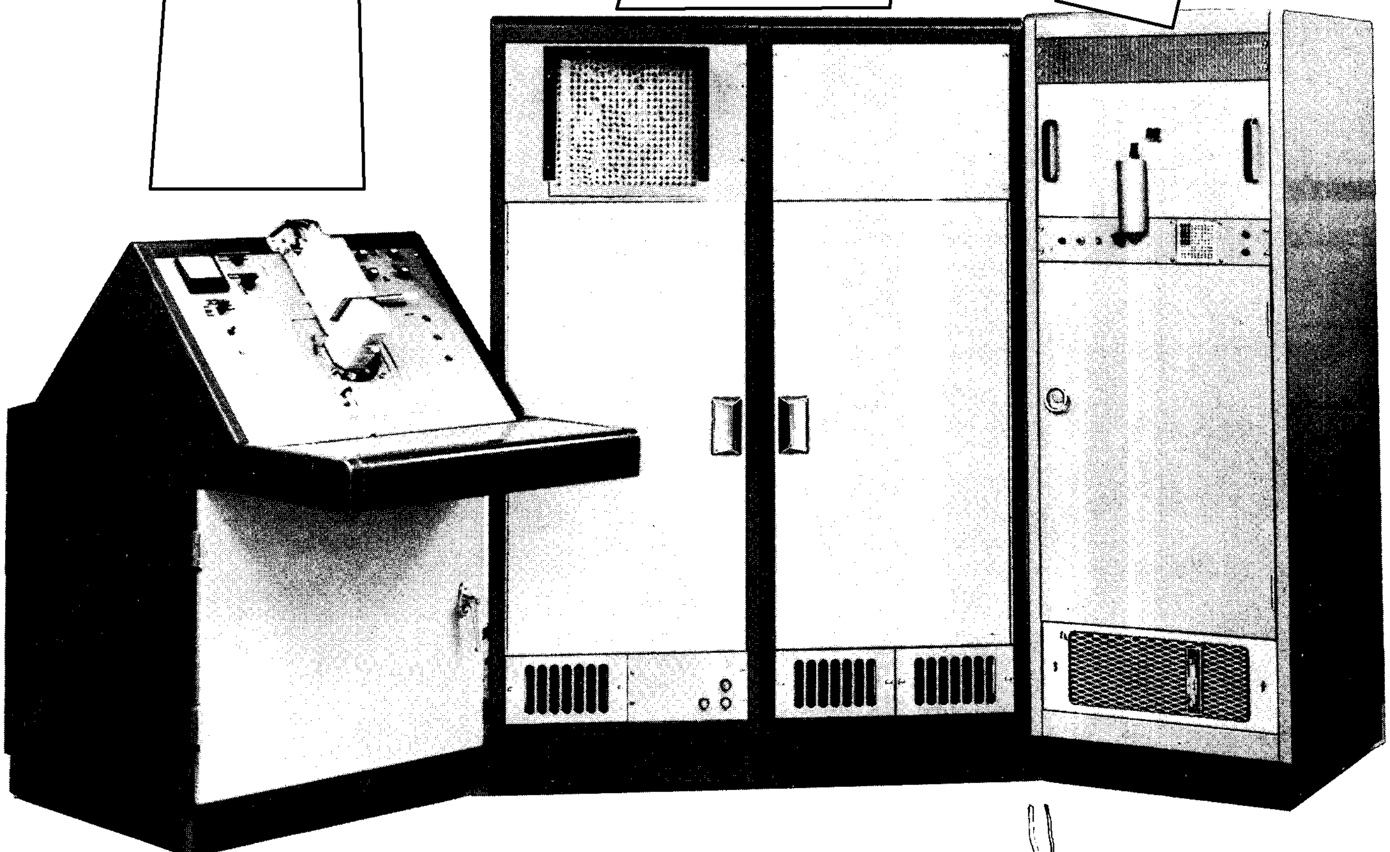
Prepared by:

25X1 [redacted]

ILLEGIB
ILLEGIB

ILLEGIB

[redacted]



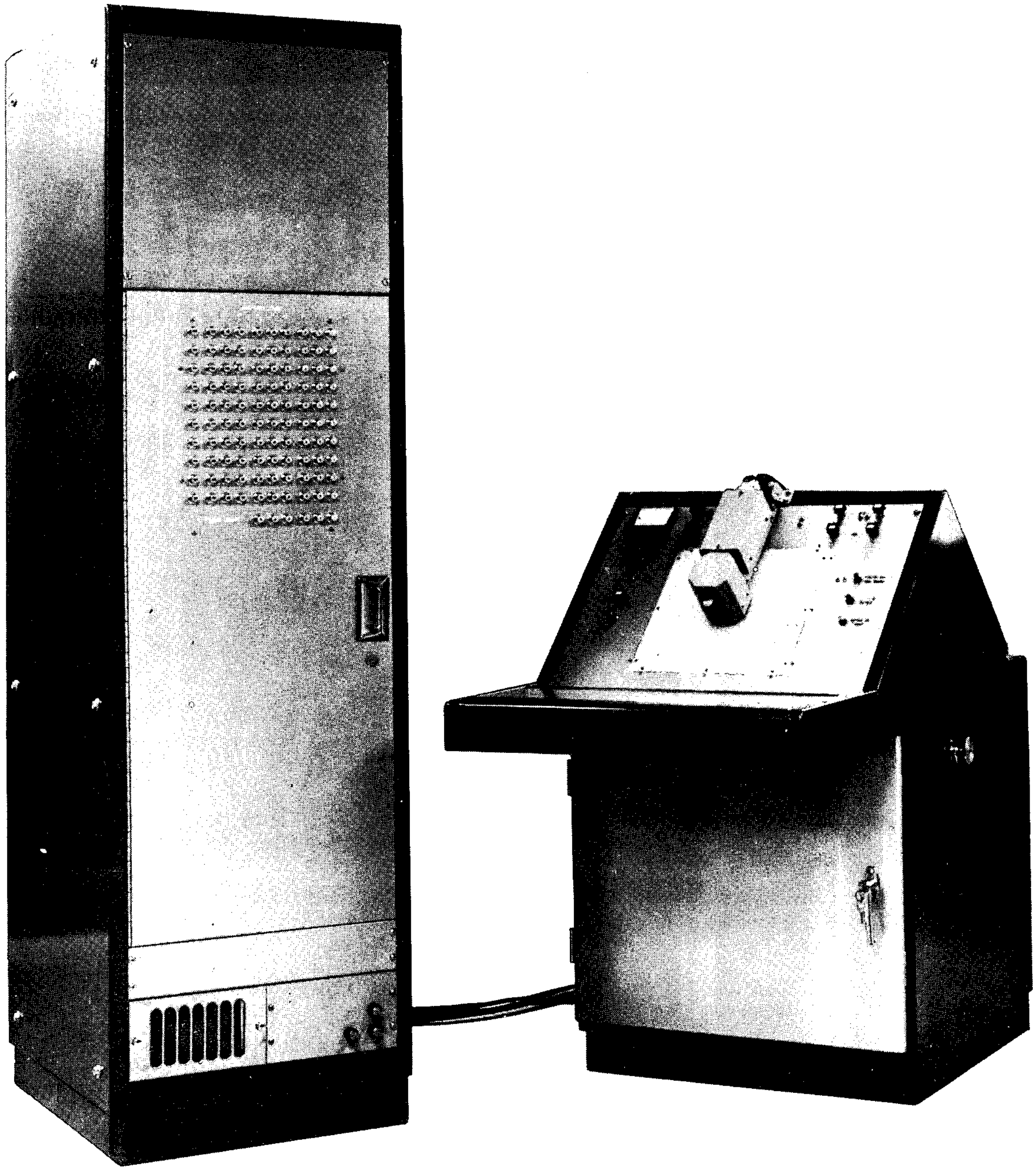
25X1

on contract [redacted]

on contract [redacted]

25X1

FREQUENCY-DOMAIN SYSTEM



RADC
Sponsored

TIME-DOMAIN SYSTEM

FOREWORD

This final technical report describes the results obtained by [redacted] during the course of the program conducted under [redacted] The objective of the program was to evaluate the time- and frequency-domain Integral-Scan Parameter-Generator.

25X1 25X1

[redacted]

25X1

ABSTRACT

This is the final technical report on a study of the time- and frequency-domain Integral-Scan Parameter-Generator. The report includes the details of the experiments conducted to evaluate the performance of the systems with aerial reconnaissance imagery. The test results indicate that the measurements made by the prenormalizing system are descriptive of the imagery under view and that the measurements have a high degree of invariance to translation and rotation.

TABLE OF CONTENTS

ILLUSTRATIONS

FRONTISPIECES: FREQUENCY-DOMAIN SYSTEM

TIME-DOMAIN SYSTEM

TABLES

I. INTRODUCTION

Research at [redacted] including a survey of the overall developments in automatic recognition equipment, shows that the development of procedures for removing variations in pattern-recognition error-rates caused by translation and rotation should be emphasized over the development of complex sensing or decision logic. However, so long as research programs use only cut film, the need for such prenormalization is not evident, and the processing equipment can consist of a scanner and a decision logic. (See figure 1.) ?

However, once roll film is introduced into the research program or operational procedures are simulated, rotation and translation of imagery occur immediately, and one of three techniques for handling this variability is required. First, the decision logic can be trained for all possible orientations of the target imagery. Second, some form of invariant parameter-generation can be used. Third, the film stage can be designed to rotate and translate the imagery. It is easier to add video-processing techniques to eliminate rotation and translation variances (figure 2) than to develop a film stage to rotate or translate the film.

The program covered by this report concentrated on demonstrating both the capabilities of the two prenormalizing techniques developed by [redacted] and the invariance, 25X1 to translation and rotation, of the prenormalization equipment implementing these techniques. Photographs of the hardware

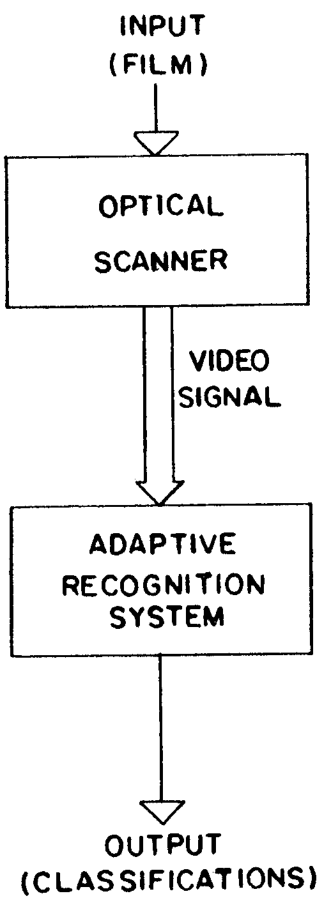


Figure 1. Automatic Target-Recognition Model
(Without Video Processing)

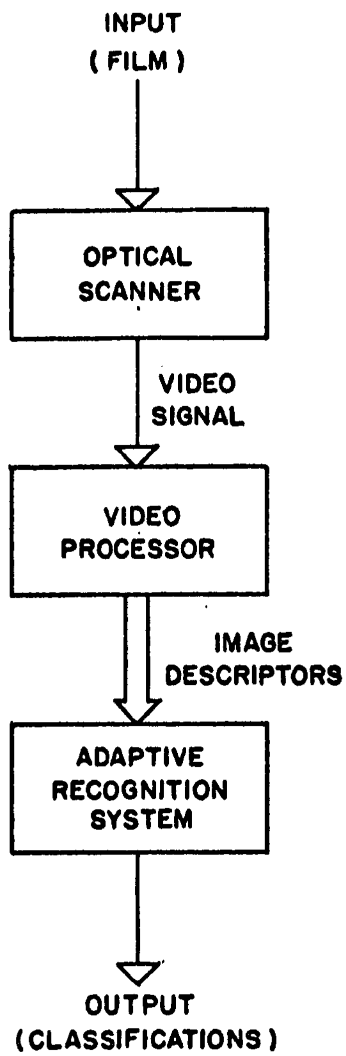


Figure 2. Automatic Target-Recognition Model
(With Video Processing)

are presented as Frontispieces to this report. Available decision logics were used, and no effort was made to improve existing logics or to develop further the scanning or processing techniques originally generated in the program.

The principal objective of the program was to demonstrate experimentally that target images can be recognized even when information comprising several orders of magnitude has been eliminated. It is not possible to determine exactly which measures of an image completely classify or describe it. However, it has been found that it is possible to judiciously choose a small number of meaningful measures and still accomplish recognition of the imagery.

The two generalized preprocessing techniques employed in this program are shown in conceptual form as figures 3 and 4. Frequency-domain processing (figure 3) is based on the generation of D-cells from two-dimensional power spectrum data, and time-domain processing (figure 4) is based on the correlation of image derived video signals with sets of orthogonal polynomial functions. Both techniques will be discussed in more detail in a later section. In particular

what data?

target search problems, if some preliminary data about the imagery is supplied, it is possible to develop preprocessing equipment which is devised to extract parameters best suited to describing the targets being sought. This would improve the error rate and eliminate the necessity for having all of the band-pass filters and general orthogonal functions available.

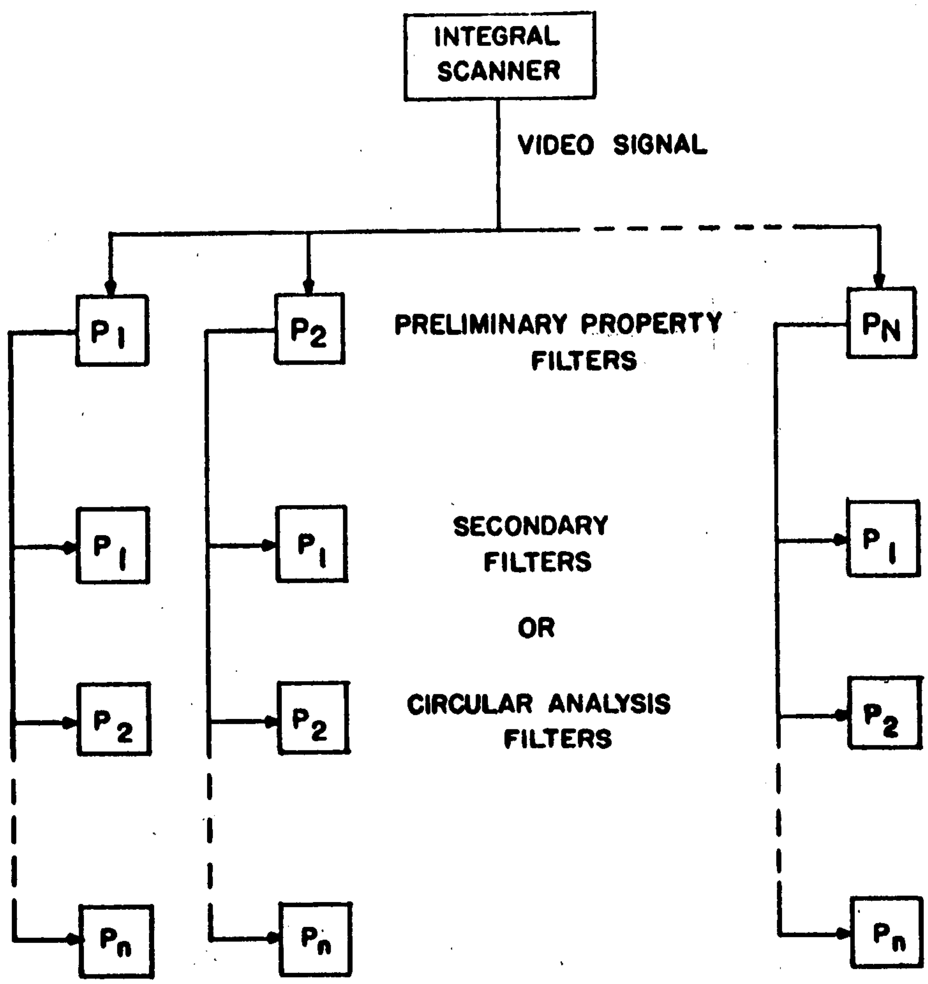


Figure 3. Frequency-Domain Processor

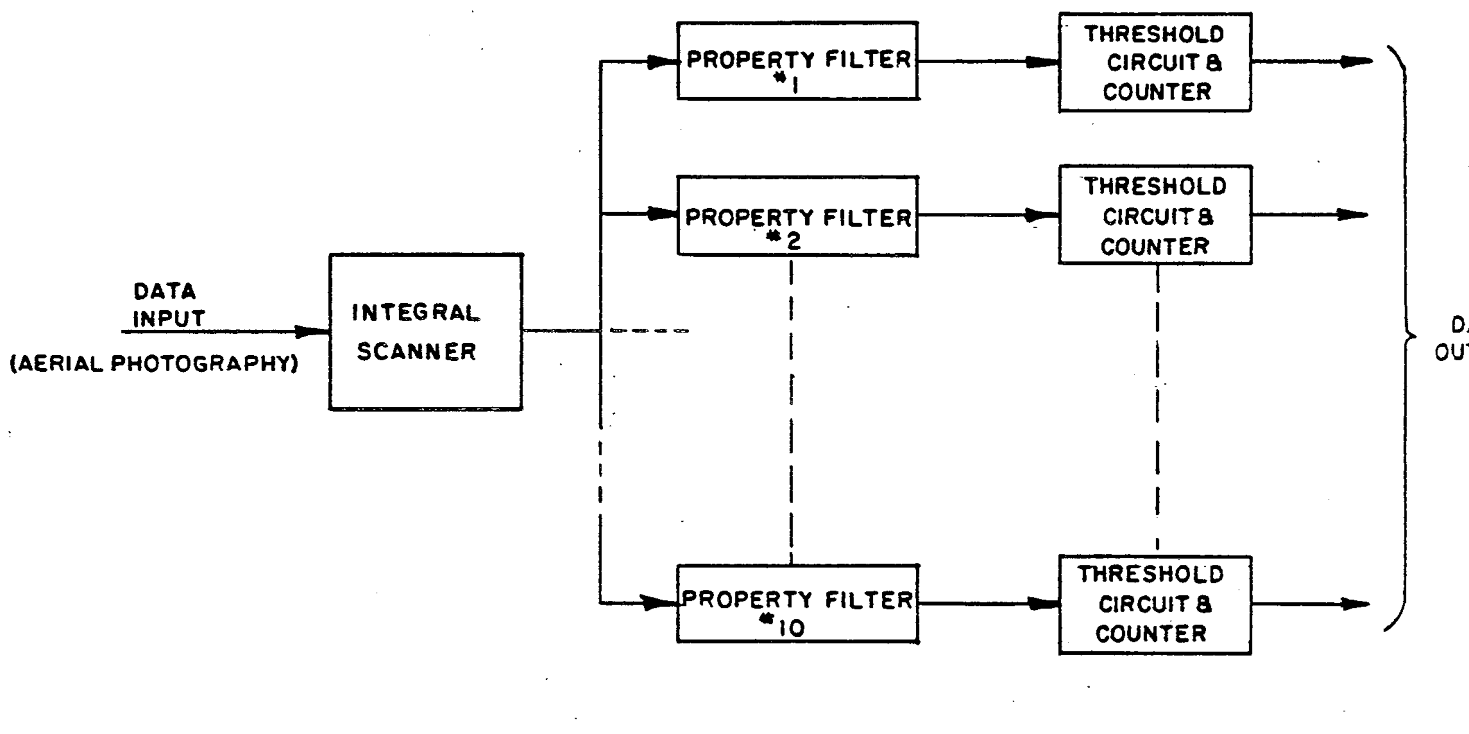


Figure 4. Time-Domain Processor

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

During this program, the stress was not on defining the absolute recognition-rates but on proving the invariance of these rates to rotation and translation. Recognition rates can always be improved by modifying existing equipment; however, the invariance of those rates must be a fundamental part of the approach. Also, simple error-calculations are not sufficient for analysis of the experimental results; the calculations must be analyzed for invariance during rotation and translation, invariance between target classes, and consistency of recognition rate.

The Chi-square test has been chosen for analysis of the data. Other tests, such as the analysis of variance, or factorial analysis, would have yielded additional information. However, the imagery data required for these tests could not be obtained during the program.

The integral scanner used during this program was the optomechanical device outlined in Appendix I and detailed in the final hardware report submitted under separate cover. The resolution and flexibility of the equipment were limited; we now feel that integral scanning equipment should be based on an easily programmed electronically-generated flying-line. Because of the unique sensing capability of the flying spot and the flying-line, a combination of the two would be most desirable. However, the full potential of the existing optomechanical flying-line scanner can be more fully exploited.

A. PROGRAM DESCRIPTION

The original objectives of this program were to define the operational goals for automatic screening of photographic intelligence data; to study, test, and evaluate the techniques applicable to the problem; and to generate a design for an operational prototype system. At the sponsor's request, work on the first and third objectives was discontinued so that the second could be emphasized.

A problem arose when the GFE CONFLEX I had a failure in its disc memory, necessitating the programming of a CONFLEX simulator on the CDC G15 and SDS 925 computers. Because of limits of the memory size of the computers, only two-class experiments could be run.

The CONFLEX I malfunction also entailed making the extraction equipment independent of the CONFLEX. The original plan was to go directly from the integral scan prenormalization equipment into the CONFLEX I. Since the CONFLEX became unavailable it became necessary to revert to a punched-paper-tape output and to process all of the experimental data in three phases. The first phase was data collection and punched-paper-tape output; the second phase was conversion of the punched-paper-tape output to SDS 925 format input; and the third phase was the use of the CONFLEX simulator on the SDS 925 computer.

An extensive testing program requires a considerable amount of visual imagery. Several sources of this imagery were contacted, and imagery was received from the Air Force and the Army. Although ideally some measure of quality should have been assigned to the imagery, this was far beyond the scope of the program; therefore, imagery subjectively established as being of differing complexity was used in the test data-base.

Unfortunately, most of the imagery was of cultural rather than military targets, a factor that in no way affected the validity of the program results but did limit the results since the inclusion of more military targets would have eliminated potential questions concerning the military significance of the data. Because it was possible to conduct only two-class experiments and there were but limited examples for each class, most of the tests were simulated screening operations. That is, the task of the decision logic was to distinguish man-made or other significant visual features from imagery having non-significant content.

Work on defining the operational objectives for automatic screening of photographic intelligence-data and the generation of a design for an operational prototype system was discontinued part way through the program at the sponsor's request. That effort has not been included in this report, but it is worthy of note that the completed research indicated the possiblity that an automatic screener could be designed to separate significant targets from background.

A review of the test results will show that many more experiments are possible. A full evaluation of the integral scanning technique has just begun, although several thousand images have been processed. The urgent need for a standardized data base of several hundred examples for each of several hundred target classes has been demonstrated. A large portion of each automatic target-recognition program is always concerned with the collection or generation of suitable input materials.

In summary, although many outside factors have caused a change in the original direction of the program, many useful results have been obtained. The integral scanner and its video processing equipment are now completely independent of the CONFLEX I. The processor output is punch-paper-tape that can be utilized in any decision logic with suitable modifications. Thus, this equipment can be used with any suitable input imagery and with any suitable decision logic.

The capability now exists for the processing of several thousand images a month and by simple change of output to a higher-speed peripheral unit, or by going on line, this test capability can be increased manyfold. The system has become a research tool capable of generating a standardized input for differing decision logics; the integral scanner alone can be used to generate standardized linear signals for differing preprocessors. Finally, extensive testing has demonstrated that the integral scanning equipment with time-domain and frequency-domain processing does yield statistically significant invariant error rates.

II. EXPERIMENTAL PROGRAM

A. TIME-DOMAIN ANALYSIS

During the intial phase of this contract two parallel research efforts were being conducted. One dealt with a prenormalization approach based on the time-domain analysis of signals obtained by the integral-scan sequence, the other considered frequency-domain analysis. When the breakdown of CONFLEX I dictated the instrumentation of a paper-punch-tape output of the frequency-domain analysis scanner developed under this contract, the time-domain processor and scanner were used in conducting the tests. The purpose of this procedure was to establish some basis of comparison between the two processing approaches (i.e., frequency-domain versus time-domain) by using identical imagery in similar tests on both systems.

A brief description of the time-domain analysis approach, the system developed, and the computer processing used precedes a delineation of the tests conducted on this system.

1. System Description

A detailed description of the integral-scan technique itself and the motivation behind it is presented in the final report submitted under [redacted] and synopsized in Appendix I. There it was established that changes in rotation and translation of an object in the scanned field result in changes in the time of occurrence of the video waveshapes of the objects.

Thus, when a method independent of time of occurrence was used in processing this waveform, a set of measures (or properties) invariant to translation and rotation could be established. Time-domain analysis, diagrammed in figure 4, is matching a set of ten orthogonal property filters to the video waveform passed through delay lines. Essentially, the approach consists of looking, as it were, through a weighted window at the waveshape and establishing a level of correlation between this waveform and the various property filters. The correlation level is then thresholded and can be counted. A tailored form of the standard correlation measure used in statistics is employed, and a generalized circuit implementation of this measure is shown in figure 5.

Of the 31 operational amplifiers that make up the component circuits, 11 are used in the normalizing calculation of the video information. The remaining 20 perform the final summing operation of the ten orthogonal property filters to which the video is matched. The output pulses from the threshold circuits of these ten filters are counted and displayed on ten-stage binary counters, one of which is set for each of the ten property filters.

2. <u>Computer Program</u>

Since CONFLEX I was not available for testing, a computer
25X1 program was written for [redacted] CDC G-15 computer. The program

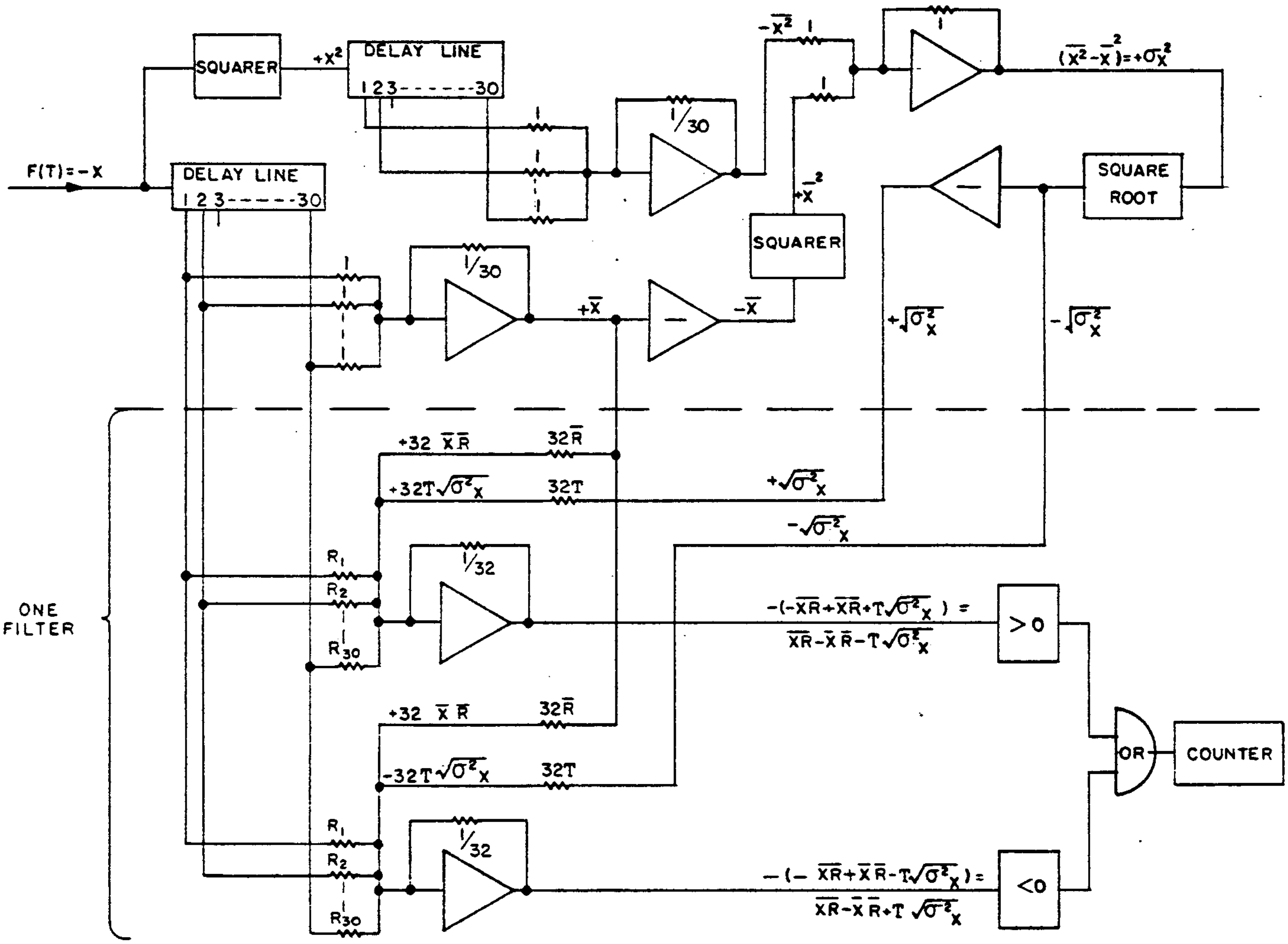


Figure 5. Circuit Implementation of the Correlation Function

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

simulated an adaptive structure,[1] using a linearly independent coding[2] on the values obtained from the property counters for each target sample. These values were recorded on paper-punch-tape and computer processed.

The computer program is presented in the function diagram of figure 6. The output count of each filter is used to generate 100 D-cells, in either a +1 or -1 state. This state is dependent on a threshold level that varies linearly from the first to the 100th D-cell. Hence, a property count (filter output) that is exactly half of the maximum count acceptable would produce 50 + D-cells and the same number of - D-cells.

The 100 D-cells of each of the ten filters are in turn multiplied by the disposition of 1000 weights in each of two fields. (For a two-class experiment, one field would be sufficient, but this program is designed to handle a four-class problem; hence, the two fields.) At the beginning of any learning cycle, the weights are set to zero; hence, the results after multiplication are all zero. All of the

---

1 Widrow and Hoff, "Adaptive Switching Circuits," Stanford Electronic Laboratory, Stanford, California, 1960.

2 Smith, "Contractor Control by Adaptive Pattern Recognition Techniques," Stanford Electronic Laboratory, Stanford, California, 1964.

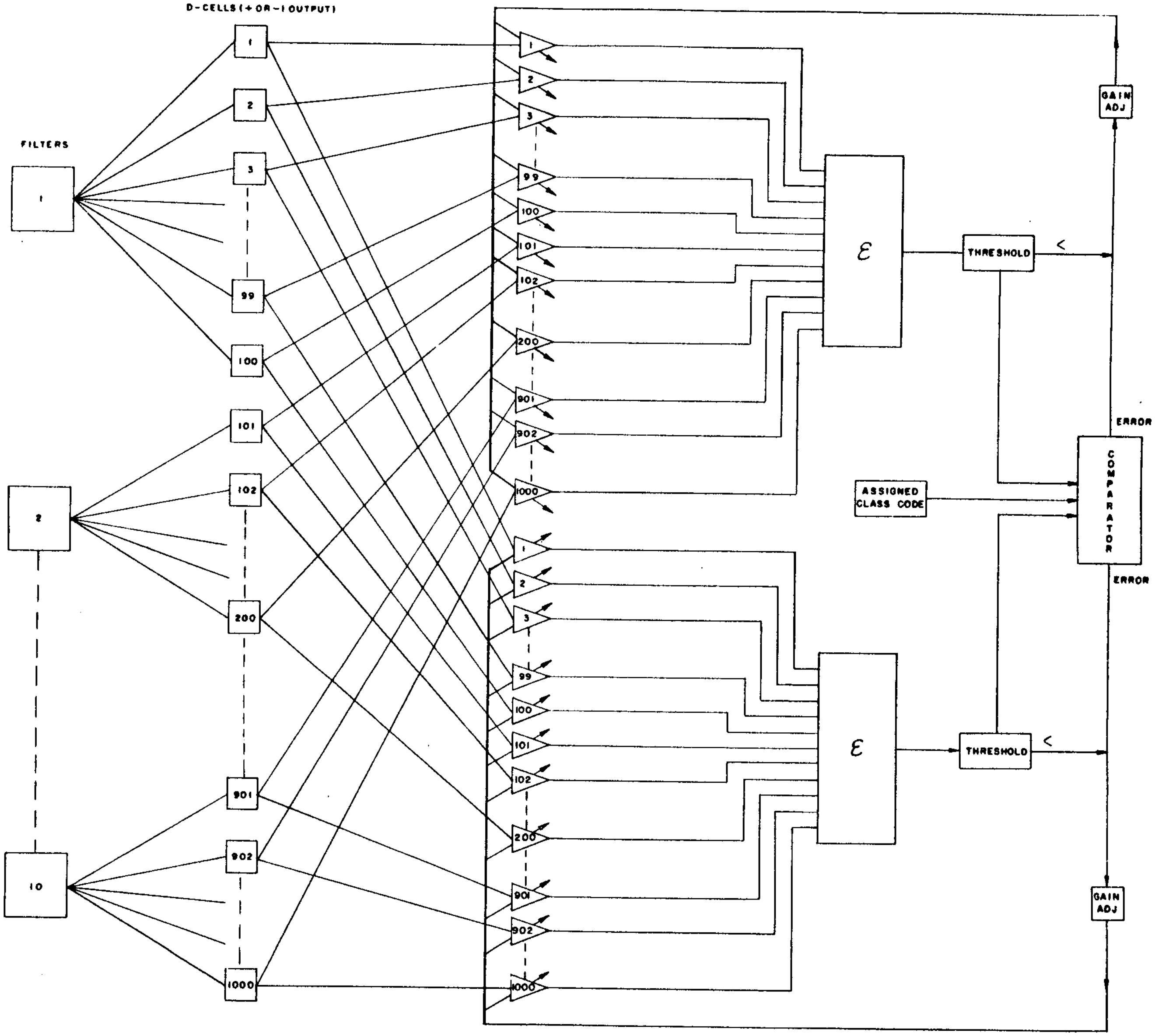


Figure 6. Function Diagram: Adaptive Learning Program (CDS G-15 Computer)

weights in each field are summed, and the signs of these sums are compared to an assigned class-code that is recorded on the paper-punch-tape with the filter count of a given sample.

For this four-class program, the code can be either ++, +-, --, or -+, where each pair of signs is compared to the signs of the sums of the two fields. If either or both do not compare (as would be the case here where all weights are initially zero) the weights in either or both fields are adjusted ("Gain Adjust") by incrementing or decrementing the weight counters according to the D-cell inputs; i.e., a + D-cell increments the counter, and a - D-cell decrements it. Once again, weights are summed and the signs compared with the assigned code. This process is repeated and the memory (i.e., weights) is thus updated until the signs of the sums of both fields agree with the correct class code for the given sample.

In a like manner, each sum is compared to a threshold set to define the degree of certitude with which a sample is learned. If either or both sums do not exceed this threshold, the weights in either or both fields are adjusted in the manner explained above until the threshold level is exceeded. These weights or counters can now be referred to as a memory adapted to the sample represented by the input-filter count.

In a typical experiment in a learn mode, numerous samples in up to four different classes are fed into this computer program through tapes, and recycled through until the memory is completely adapted to all samples in each class. For each iteration, the sample number, an ERROR or NO ERROR indication, and the field sums are typed out. Once the memory convergence is reached, any number of unknown samples can be tested against this memory for classification into one of the four classes.

The recognize mode is similar in its operation to the learn mode, except that no memory adjustment is involved. The D-cell representation of the tested sample is multiplied by the 2000 adapted weights; the results are summed and tested against the class code assigned to the test sample, and against any recognition threshold that may be programmed in. If any operation involves an error, this information is simply typed out, and the next unknown sample is tested.

3. Experiments

Tests were conducted using the time-domain system and computer processing described above, and the limited amount of imagery on hand. The classes of targets were aircraft, bridges, road intersections, tanks and various backgrounds. Varied numbers of learn samples were taught in one orientation, and the resulting adapted memories were tested against appropriate classes of test samples.

In some cases, the test samples were the same learn samples in similar orientations, to test scanner/processor repeatability. In others, the test samples were the same samples in different orientations to check invariance to translation and rotation. Further, unknown (i.e., unlearned) samples were tested for correct classification as well as invariance to position. The results of these and subsequent tests on the frequency-domain processor are detailed in another section of this report.

B. FREQUENCY-DOMAIN ANALYSIS

1. System Description

A detailed theoretical and hardware description of the frequency-domain system -- the integral scanner, processor, and punch-output -- is presented in the final report written under [redacted] The present description is concerned with the output format of the processor data, the computer program involved in the conversion of this data to a compatible form for processing on a second, high-speed computer system, and finally, the actual learn/recognize, CONFLEX-simulation processing program. The functional block diagram in figure 7 outlines the information flow of this frequency-domain system approach.

25X1

It should be pointed out that the hardware implementation of an ideal automatic recognition system would combine all of the components included in figure 7 with a more sophisticated, mission-oriented form of input and output,

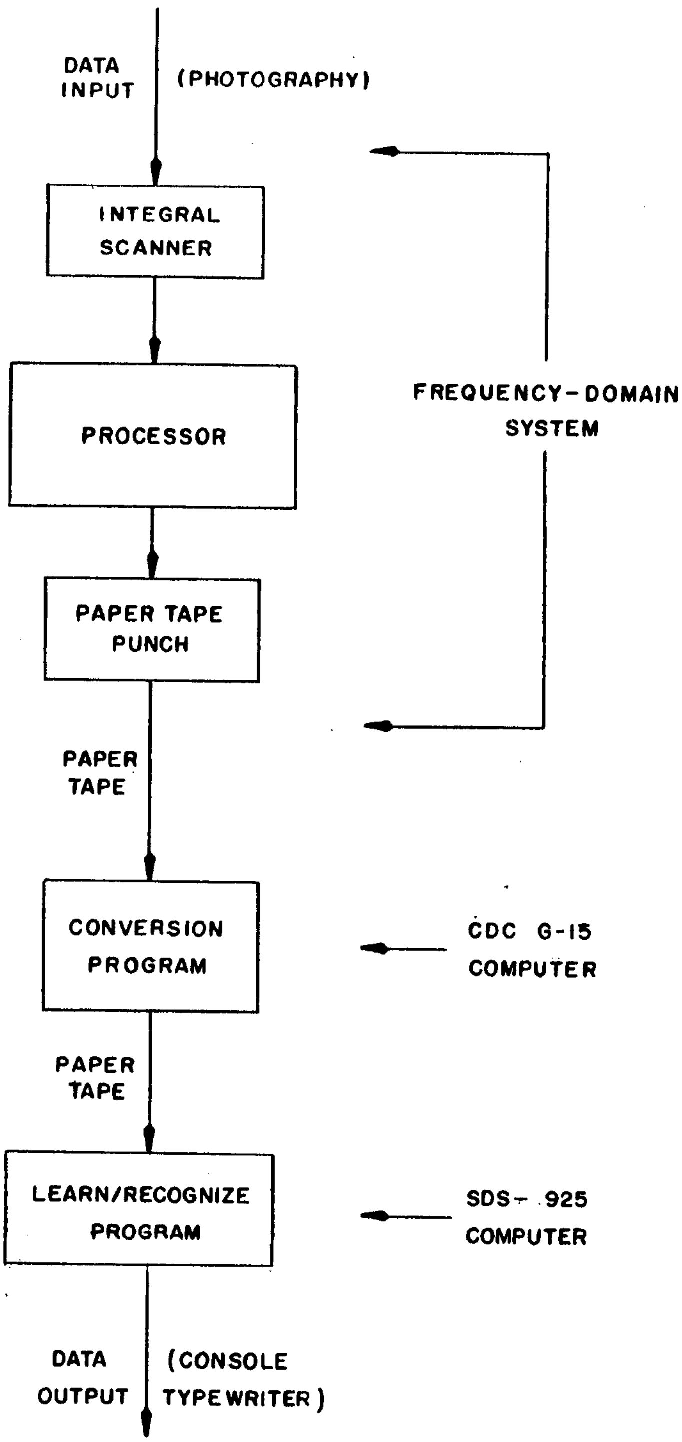


Figure 7. Block Diagram: Frequency-Domain System and Computer Analysis

and without, of course, the paper-tape medium. The function-breakdown approach was used in this program to provide greater flexibility in testing and defining system-component parameters.

Experience with the CDC G-15 program used in the time-domain approach indicated the advisability of using the SDS-925 computer for the learn/recognize processing. Because tests involving a large number of samples showed a need for more rapid processing than could be obtained from the G-15, an adaptive learning program was written for the speedier SDS-925, and a conversion of the paper-tape-punch output format originally designed for processing on the G-15 became necessary. Computation time per sample was thus reduced by an estimated factor of 40, and the resulting more versatile program allowed closer examination of the processes. The details of this development are discussed in a later section of this report.

2. Computer Program

Conversion Program: The purpose of the conversion program is: first, as mentioned above, to translate the punch output into a tape input-format acceptable to the high-speed SDS-925 computer used in the actual data processing; second, to give each data sample an assigned class-code by which a computed classification is checked with the actual identification of the sample.

It would be helpful here to review the output data-format of the tape perforator, described in the Punch Output and Logic Section of the Final Report. Briefly, 25X1 a data sample is represented by a punch output of 400 D-cells in either of three states, +1, 0, or -1, on five-level paper tape. Each punch character contains two D-cells in the form pictured below:

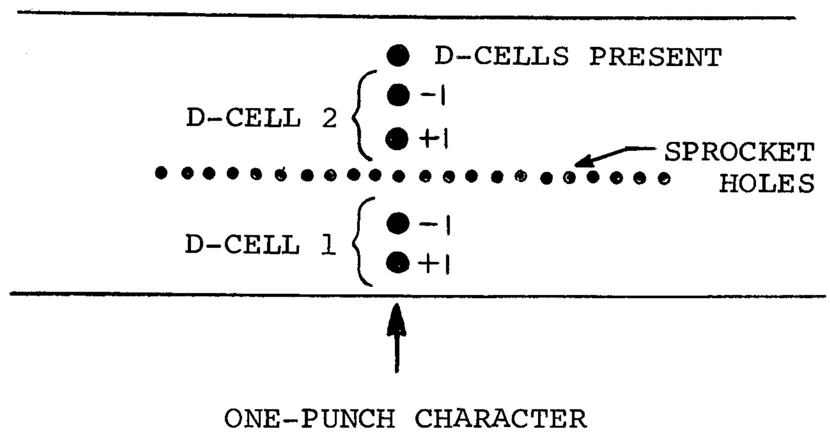


Figure 8. Example of Paper-Tape-Punch Output Format

A punch in the fifth level indicates the presence of two D-cells in that punch character. If no other punch exists, both D-cells are in a 0-state, but any other punch in the first or second level represents the disposition of one D-cell, and in the third or fourth level, the disposition of the second D-cell, according to the above figure. Taking the hardware aspects of the CDC G-15 computer into consideration, the paper-punch logic is designed to output a tape block made up of seven punch characters plus a "tab" code four times, followed by a "stop" code. The tape block, then, contains 56 D-cells. Four hundred D-cells of a given sample require a total of eight tape blocks, the last of which contains only four punch characters of information (i.e., eight D-cells, since 56 D-cells/tape block x 7 tape blocks = 392 D-cells + 8 D-cells (of 8th tape block) = 400 D-cells/ data sample).

Figure 9 shows a generalized flow diagram of the conversion program. This program accepts the 400 D-cells and, using a table look-up approach, converts each D-cell into an appropriate + 1, 0, or -1 representation acceptable to the SDS-925 computer. The G-15 then punches these out on five-level tape in ten blocks of 40 characters each, each block separated by an "end of block" (carriage return) code. The eleventh or last block is made up of one character, the assigned class-code. The class-code is inserted, after the conversion of a data sample, under program control according to an instruction typed in at the beginning of a conversion run, which designates how many samples are in which class and the order of occurrence they follow.

<u>Adaptive Learning Program</u>: The approach used in the adaptive learning program written for the SDS-925 computer is basically similar to that used in the G-15 program for the time-domain processing described above. The main difference in the two approaches is that, because of memory-size limitations, the adaptive learning program can handle only two-class, not four-class, experiments. Instead of the 1000 internally coded, two-state D-cells used with the

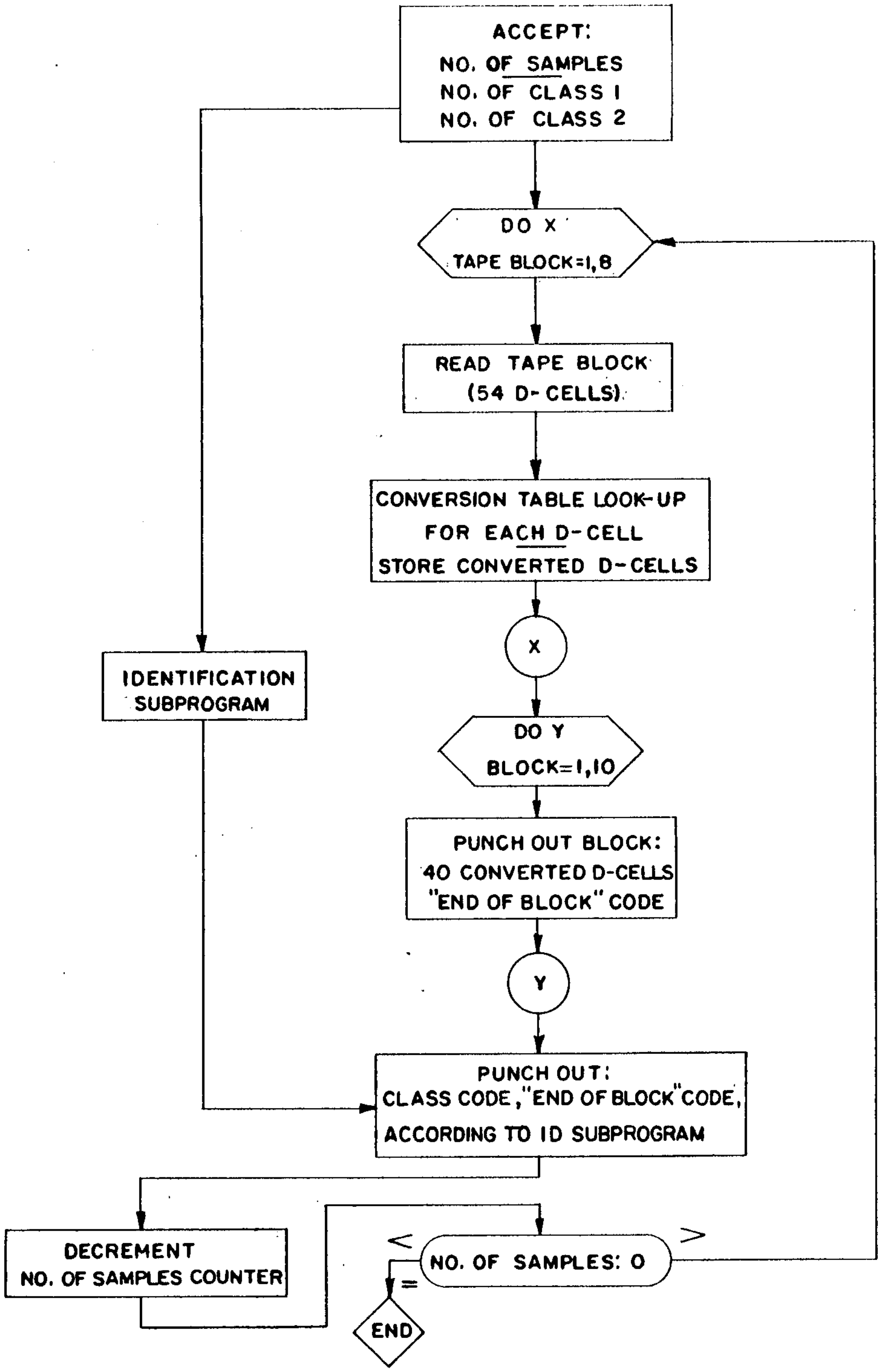


Figure 9. Conversion Program Flow Chart

time-domain approach, here 400 three-state D-cell representations of the samples, generated externally by the frequency-domain processor and fed in via a format conversion in the CDC G-15 computer, are employed. To facilitate more rapid convergence upon an adapted memory in a learning mode, the memory adjust ("Gain Adjust" referred to above) routine was revised. Previously, an error resulting from misclassification and/or threshold requirements caused repeated adjustment of the memory weights until a non-error condition existed, but in this program any error causes only one memory adjustment. In this manner, progressive learn samples in various classes cause a more gradual adaptation of the memory weights, hence a faster convergence.

The function diagram in figure 10 represents the computer program. Four hundred D-cells and an assigned class code on five-level paper tape are read in. Each D-cell is equivalently multiplied by the disposition of 400 weights initially set to zero at the beginning of any learn cycle. The weights are then summed, and the magnitude of this sum is compared to a threshold set to define the degree of certitude with which a sample is learned. If the threshold is not exceeded, as would be the case at the beginning of any learn cycle where all weights are zero, the memory is adjusted by incrementing, maintaining, or decrementing the weight counters according to the +1, 0, or -1 disposition of the D-cell input to that counter. If the threshold requirement is met,

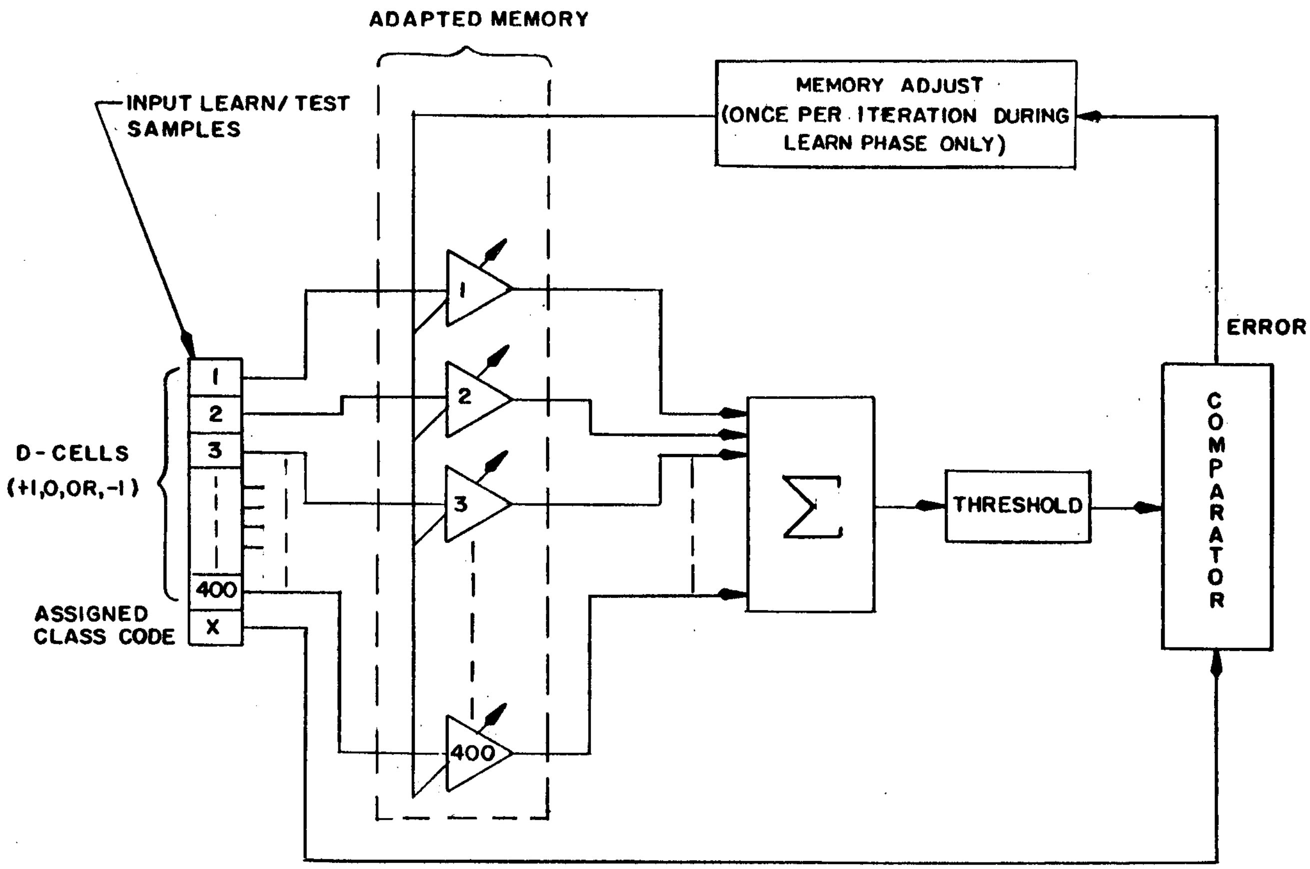


Figure 10. Function Diagram: Adaptive Learning Program (SDS-925 Computer)

the sign of the sum (+ or -) is compared to the assigned class code recorded on the punched tape with the D-cells of the given sample. If this comparison does not hold, the memory weights are adjusted in the manner described above. In any case, whether the error involves the threshold and/or the classification, the memory is adjusted only once each pass for a given sample. Thereupon, the next learn-sample tape is read in and the entire process repeated until all learn samples are processed without a memory adjustment (i.e., a pass involving no errors).

The generalized flow chart in figure 11 describes details of the actual mechanics of the computer program. With each sample input, the sample number and weight sum over the threshold is typed out before any memory adjustment takes place. To make errors more easily identifiable, a routine was incorporated whereby any error appears as a minus sum in the print-out. The routine consists of equivalently multiplying the sum of the negatively identified class (Class 1) by a -1 (i.e., the SUM = -SUM instruction in the flow chart) and always subtracting the threshold. In keeping with this convention in the memory adjust routine, each weight is updated according to its input D-cell and class code. If the sample is coded as a Class 1, each + D-cell decrements its appropriate weight counter, and each - D-cell increments its counter. For a Class 2 sample, the + D-cell increments its counter, and the - D-cell causes an appropriate decrement in its weight counter. In each class, a 0-state D-cell maintains the existing state of its weight.

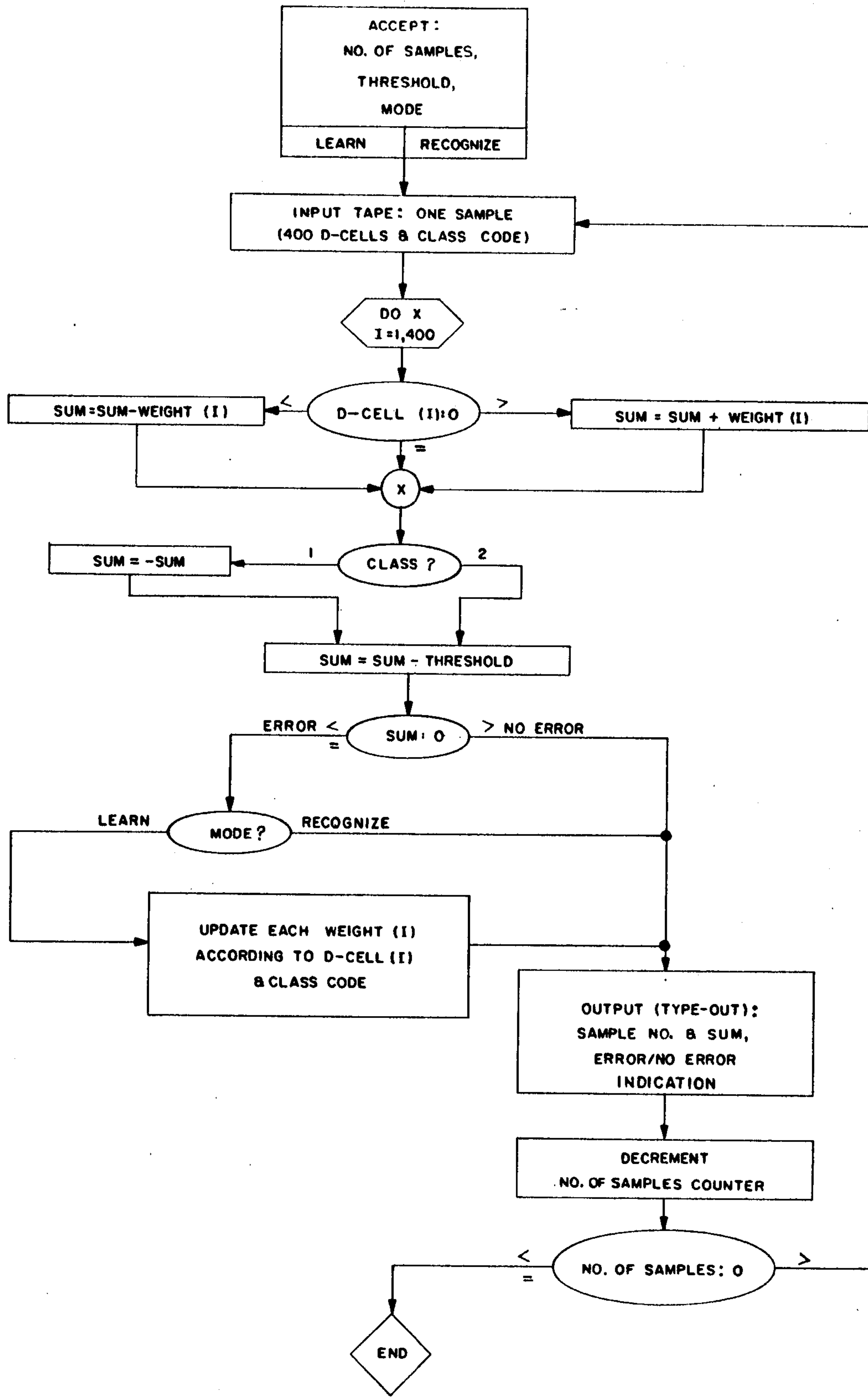


Figure 11. Adaptive Learning Program Flow Chart

Several sense-switches are programmed in to provide a variety of output. It is desirable to have the complete print-out (for each sample, its sample number, its sum of memory weights and error history, plus the pass number and total errors per pass) when samples are being tested, but limited print-outs can be used during extended learn phases to conserve time. Provisions have been made also to dump memory at any point in the learn cycle in order to test against the unadapted memory and compare the results with those from tests with a completely adapted memory. The dump-memory provision also makes it possible to preserve memories and later cross-test with samples of different tests.

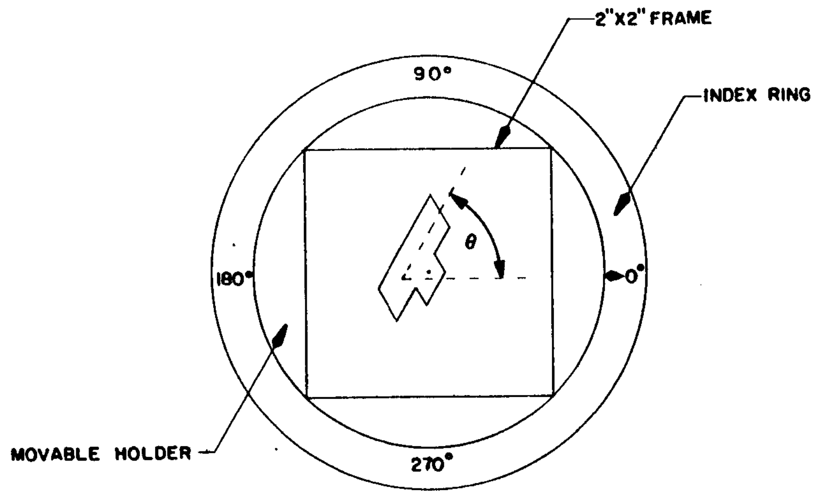
3. Experiments

Two distinctively different types of experiments were run. The first, called generalization, consisted of training the decision logic with images from two classes and testing this decision logic with different images from the same two classes. The second, called memorization, consisted of testing the decision logic with the identical imagery used for training the logic. In both types of experiments, rotation and translation took place during generation of the testing samples. Thus, for the results derived from the generalization experiments to be meaningful, it was necessary to have a standardized format for the imagery.

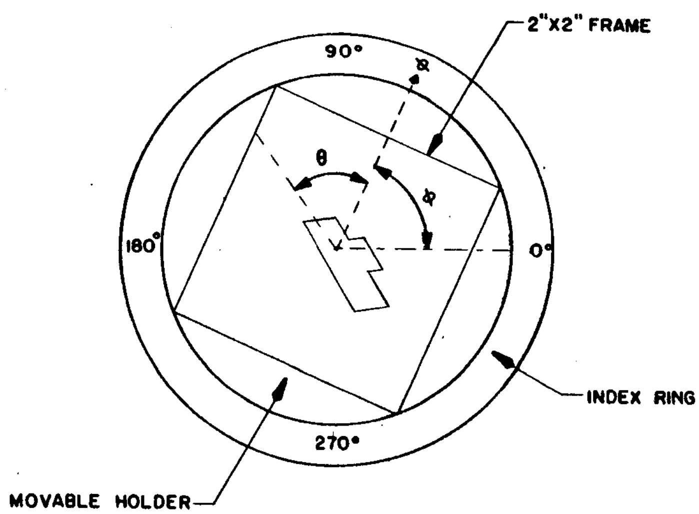
Figure 12(A) shows the random orientation of the imagery within a frame. Figure 12(B) shows the actual orientation of this imagery after a specific Rotation φ. It can be seen that after this rotation has occurred the actual orientation of the imagery is unknown because of the first incremental angle θ.

In memorization experiments, image orientation is not a problem because the target has been seen in its original orientation. However, with a separate group of images such as those used in generalization experiments, some form of standard orientation must be maintained. Figure 13(A) shows a standardized image frame. Figure 13(B) shows that a known rotation has rotated this image through a known angle. Thus, if the performance in this experiment is related to rotation, it will be quite easy to determine error rate changes with rotation angle. However, if the imagery had not been standardized, the rotation of the testing imagery in relation to the rotation of the training imagery would be completely unknown. Unfortunately, many targets do not lend themselves to establishment of a standard orientation. During the program, the major axis of the imagery was aligned horizontally.

It is obvious that conducting a generalization experiment is much more complex than conducting a memorization experiment; therefore, most of the experiments were limited to the latter type. However, it should be kept in mind

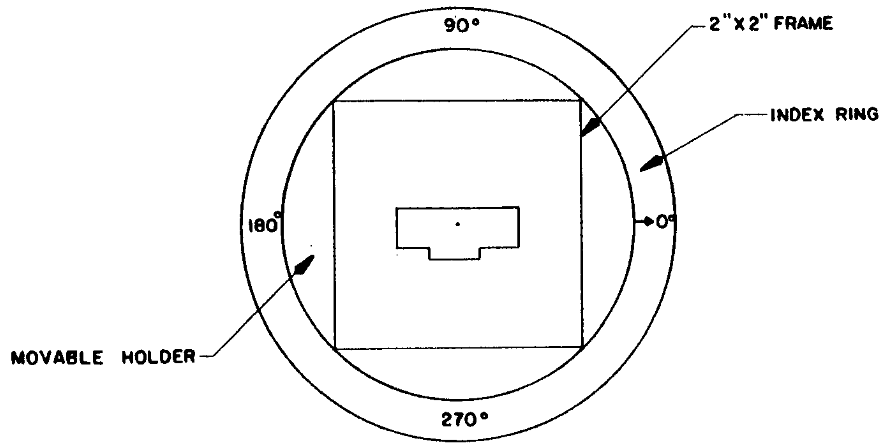


(A)

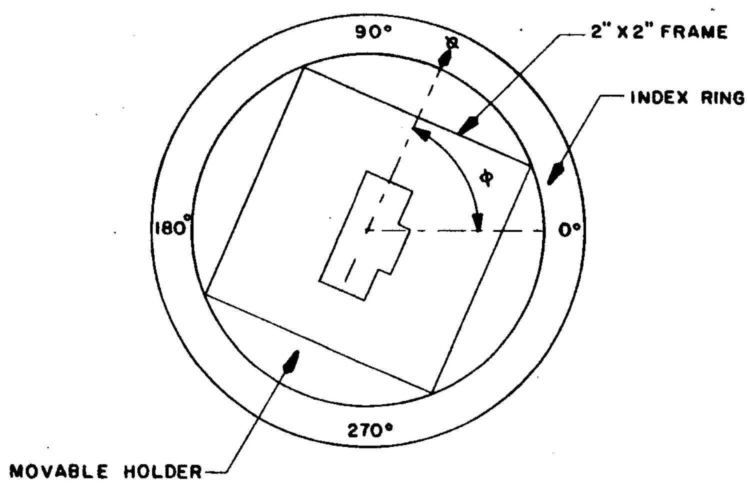


(B)

Figure 12. Random Orientation of Imagery



(A)



(B)

Figure 13. Standardized Orientation of Imagery

that this program was not established as a target-recognition program, but as a recognition-invariance program. Therefore, the demonstration of invariance can take place regardless of the type of experiment used to demonstrate it.

Another facet of the experimental effort was directed towards investigating the parameters of the existing integral scanner. Tests were conducted to examine the information content of single-sweep video signals. Positive and negative transparencies of primitive shapes in varying widths and field sizes such as straight lines and line intersections, and other simple geometric shapes (e.g., triangles, squares, hexagons, circles, etc.) of a constant area of transmission were produced for examinations of linear and angular resolution. Experiments with scanner noise and scanner repeatability were also included.

C. IMAGERY

An extensive pattern-recognition program requires an extremely large, well-controlled, and representative data base. During the present program, [redacted] 25X1
developed a limited data base consisting of imagery in 2" x 2" super slide-holders. The imagery was assigned to the data base on the basis of its content. Performance of the target-recognition equipment during the program was related only to the subjective content of the imagery.

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

Other possible variations such as detail (characteristic of individual image), context (the surrounding imagery), density (number of targets), and image variables will all have to be verified. Because it is possible to introduce variations in performance throughout any experiment, an extensive image-evaluation program was not warranted. Such an analysis, either subcontracted or conducted using GFE data, should be considered with any future program.

The characteristics required of the imagery selected included:

1. The sensing signal would be transmitted by the imagery.
2. The size of the target on the imagery was approximately 1/10th the total frame size.
3. The imagery was of individual targets or groups of identical targets.
4. The imagery was of targets with clearly delineated outlines. Non-target imagery, such as opaque edges and auxiliary information, was also included with some of the targets. Table 1 lists the imagery used in the experiments.

TABLE 1. IMAGERY USED IN EXPERIMENTAL PROGRAM

<u>TARGETS</u>: (Positive transparencies)

Tanks	Dams
Runways	Tactical vehicles (e.g., trucks,etc.)
Cities	Railroads
Roads	Buildings
Bridges	Aircraft

BACKGROUND: (Positive transparencies)
Natural terrain, varying from featureless terrain to that having different levels of vegetation, shrub, and tree cover as well as snow-covered and snow-patched terrain. Bodies of water and shorelines.

PRIMITIVE SHAPES: (Positive and negative transparencies)
Geometric: Triangle, semicircle, rectangle, square, parallelogram, trapezoid, pentagon, hexagon, octagon, and circle.
Linear: Spoked figures varying from two (i.e., a straight line centered on the field) to eight spokes, in symmetric and unsymmetric configurations. (Each spoked figure has line widths of .025, .015, .005 and .002 inches.)

D. STATISTICAL ANALYSIS

The integral scanning and video processing techniques used on this program were designed to yield recognition results invariant to image translation and rotation. The main purpose of the program was to determine whether or not this invariance was being realized. Two hypotheses were tested. The first was that system recognition accuracy is independent of image translation. The second was that recognition accuracy is independent of image rotation. Each hypothesis was tested for various imagery-classification problems.

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

The hypotheses were tested by means of the well-known Chi-square test, a test discussed in most elementary texts on statistics. This discussion, which is based on Cramer's development,[3] will point out certain properties of the test that apply to this problem.

The Chi-square test is generally used to test the hypothesis that a random variable X has a given distribution function F(x). The expected values of the random variable determined from the hypothesized distribution are compared with values of the variable observed in an experiment. This comparison is effected by dividing the range of the variable into some finite number of mutually exclusive categories, computing the probability of obtaining a value in each category, and comparing the computed (expected) values with experimentally obtained values. Actually, the hypothesis that the unknown distribution function of the variable is some distribution function giving the same probabilities as the hypothesized distribution function, is essentially what is tested.

Let the sample size be n and the number of categories be k. If the (calculated) probability of obtaining a variable value in group i is p_i, then the expected number of variable values falling in group i is np_i. Let the

---

3 Harold Cramer, The Element of Probability Theory and Some of Its Applications, John Wiley & Sons, New York, 1955.

observed number of values in group i be denoted by f_i. The statistic χ^2 is defined as

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i} \tag{1}$$

Obviously, a significant deviation from the expected values tends to increase the value of χ^2.

The number of categories, k, determine the number of degrees of freedom in the resulting χ^2 distribution. Since we are subject to the linear constraint $f = \sum_{i=1}^{k} f_i$, there are k-1 degrees of freedom.

A level of significance for our test must be chosen. For instance, we might say that we will <u>reject</u> the hypothesis that our random variable has a certain distribution function if less than p% of the χ^2 values have a value as large as the calculated value. Thus, we have a p% probability of rejecting the hypothesis when it is, in fact, true. We can minimize this probability, of course, but only at the expense of increasing the probability of accepting the hypothesis when it is false. Normally, a level of significance between one and five percent is chosen.

For a p% level of significance, we <u>accept</u> the hypothesis if $\chi^2 \leq \chi_p^2$ and reject the hypothesis if $\chi^2 > \chi_p^2$. Here χ_p^2 denotes the p% value of χ^2; i.e., the probability that $\chi^2 > \chi_p^2$ is $\frac{p}{100}$.

Now, suppose that we wish to test the hypothesis that the recognition accuracy of our system is invariant to image rotation; i.e., we hypothesize that the recognition rate p is independent of the rotation of the image. Hence, the error rate $q = 1-p$ is also independent of the rotation of the image. We can use the χ^2 test to test the hypothesis that p and q are constant for imagery scanned at different rotations. Suppose that we have made n_i classifications at rotation r_i, $i = 1, 2, \ldots, k$, and that f_i is the observed number of correct classifications at rotation r_i. At each rotation we have two categories: f_i correct classifications and n_i-f_i errors. The expected number of correct classifications is $n_i p$ and the expected number of errors is $n_i q$. Since there are two categories, we have one degree of freedom, and we obtain a term

$$\frac{(f_i-n_ip)^2}{n_ip} + \frac{\left[(n_i-f_i)-n_iq\right]^2}{n_iq} = \frac{(f_i-n_ip)^2}{n_ipq} \qquad (2)$$

However, there are k such terms contributing to χ^2. Summing over the k rotations, we obtain

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i-n_ip)^2}{n_ipq} \qquad (3)$$

with k degrees of freedom.

If we do not a priori know p, we must estimate p from the sample. The total number of observations is $n = \sum_{i=1}^{k} n_i$

and the total number of correct classifications is $f = \sum_{i=1}^{k} f_i$.

Under the hypothesis that the probability of correct recognition is constant, we obtain the maximum likelihood estimate $p* = f/n$. Then $q* = 1-p*$. However, since we have estimated a parameter from the sample, we must reduce the degrees of freedom by one. Hence, we now have k-1 degrees of freedom.

This test will be applied in the next section in order to determine whether the probability of correct recognition is invariant with respect to rotation and translation of the imagery.

The Chi-Square test described above is the most powerful one that could be conducted on the existing data. The test is used to determine whether there is a constant error rate throughout the whole experiment. It is possible to determine or establish a probability of rejecting a claim when that claim is true. The probability of such an error is called the level of significance. We are establishing the amount of deviation that we will allow from a known mean. However, in such a test, there is a danger inherent in the acceptance of that hypothesis, because the probability of its false acceptance cannot be obtained. Thus, we are establishing a null hypothesis such that we are willing to "reserve judgement" about its validity unless there is clear evidence to lead to its rejection.

We can establish the following criteria in a screening test. The null hypothesis would be the following: Target is present. Thus, we will establish the limits such that this hypothesis will be accepted unless data is available to disprove it. This level of significance could be called the missed-target rate. Thus, in a screening test, we have control over the missed-target rate. However, because it is a significance test, we have no control over the false-alarm rate. As the missed-target rate or Type I rate is lowered more and more, imagery will be indicated as possessing targets. If the missed-target rate is made small enough, then every piece of imagery will be inspected.

In most military pattern-recognition decisions, it is necessary to specify both the acceptable false-alarm rate and the acceptable missed-target rate. Because the missed-target rate or the probability of Type I error must also be controlled, a loss matrix must be established. Costs are associated with both the false-alarm rate and the missed-target rate, and a balance is struck between the two. Once we have established both error rates, we no longer have a significance test. Thus, we accept the null hypothesis of the presence of targets in two directions.

It is sometimes useful to fit distributions that are either univariate or multivariate to the data. Once such distributions have been fitted, it is possible to calculate any of the potentially useful moments. In our analysis of the time-domain filtering equipment, it was possible to form histograms of the output data and to determine both which filters were most significant and which memory weights contributed to the decision. The frequency-domain filtering did not lend itself to such an analysis because the only output from the equipment was D-cell values. The original number of D-cell values were cut from 5000 to 400. Because of the slow paper-tape output from the frequency-domain filtering equipment, it would have taken excessive amounts of time to punch out 5000 D-cells. However, the equipment was still designed to inspect the output of each filter.

III. EXPERIMENTAL RESULTS

The results of the experimental program described above, which involved more than 4500 samples, are presented according to three general categories. First are the General Results, which indicate the screening capability of the system approaches used by presenting the percentage of samples correctly classified. Because of the vast amount of test-result data, the General Results combine all of the individual experiments, memorization as well as generalization, and include all samples, regardless of the time-domain tests, where two-class experiments were not used exclusively. Next, the results of any diagnostic tests, designed to analyze the validity or strength of various aspects of the overall approach, are presented. Finally, the invariance of the integral scanner approach to translation and rotation is described in a statistical analysis of the results of all of the varied experiments.

A. TIME DOMAIN

1. General Results

Initial three-class experiments resulted in an overall 81.3 percent correct classification. These were generalization experiments; i.e., the samples tested were not included in the learn phase of the tests. In addition, a random translation and rotation of the samples were tested. In detail, the results were: aircraft, 96.8% correct; crossroads, 67.9% correct; bridges, 60.0% correct.

The second set of generalization tests was limited to two classes to permit comparison with subsequent identical experiments using the frequency-domain system. Controlled rotations of the test samples were also included in these experiments. Basically, a target-versus-background discrimination was tested, using tanks as targets and natural terrain and other man-made objects as background. Correct identifications were obtained for 88.9% of the tank samples and 61.4% of the background for a weighted, overall correct classification of 84.3% for these tests.

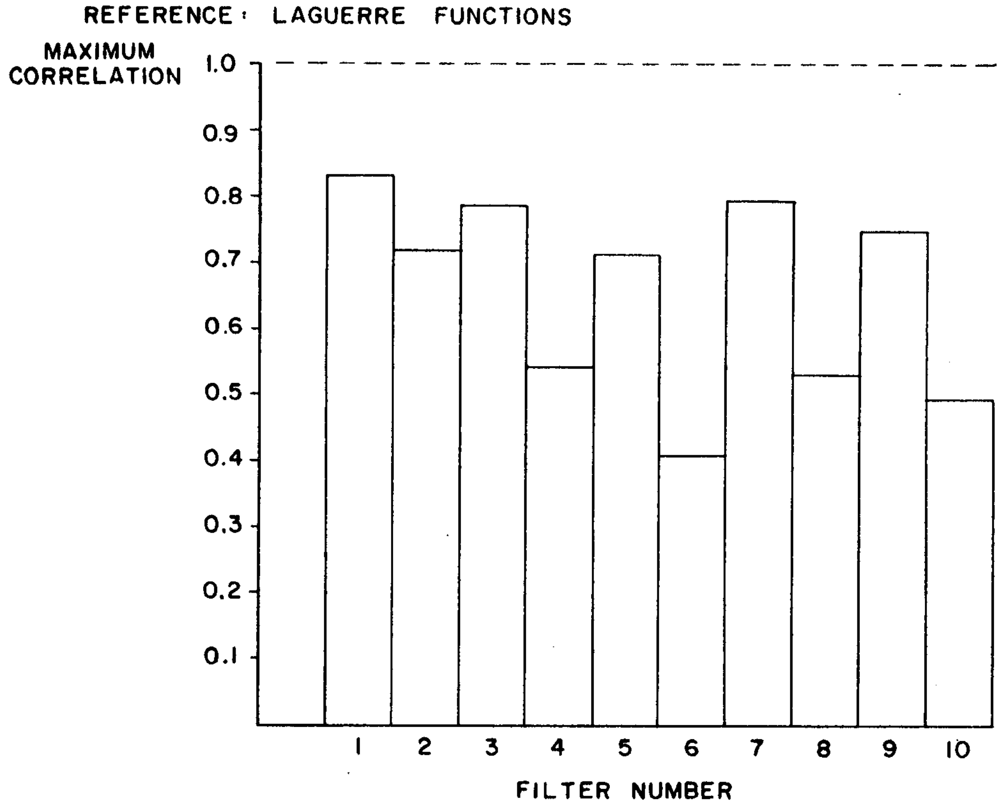
2. Diagnostic Tests

As mentioned in the description of the time-domain analysis, the video waveform is matched to a set of ten orthogonal-function property-filters, each implementing a particular weighting function. Interchangeable resistor boards were used in defining the weighting functions. For the majority of experiments, the reference-function resistor-boards used defined a series of ten orthogonal sine functions. A second set of boards defining ten orders of Laguerre functions were also constructed to permit examination and comparison of the outputs of each filter for the different classes of targets tested. This was, in effect, testing the strength of each filter in distinguishing various classes.

First, for each of the ten filters and each of the two target classes, histograms were used to record the number of filter-output counts falling in various counting intervals (e.g., 0-9, 10-19, 20-29, $\cdots$, to the maximum counter output).

Each increasing interval represents a higher level of correlation between the video input of the sample and the weighting function of the filter represented by the output count. For all of the intervals, the count total in a given counting interval in one class is compared to the total in the same interval of the second class. In this manner, an overall correlation between one class and the other for each filter is computed. For good class discrimination, this correlation should be low. Hence, an examination of the correlation values of each filter indicates how effective that filter is in distinguishing between the two classes tested. Figure 14 shows the results of the computations when, in one case, the filters implement ten Laguerre functions and in the other, they implement ten sine functions. The data in this figure shows that the sixth Laguerre function and the second and third sine functions are the most effective in distinguishing between tanks and background and that the sixth through tenth sine functions have lesser effects. Since the average of the ten Laguerre-function filters is 0.66 as compared to 0.76 for the ten sine-function filters, one can conclude that the former functions would be more effective in correctly classifying tanks and backgrounds.

From the foregoing, it is easy to see that an optimum set of filters can be selected to distinguish between any given classes of targets. The choice of weighting functions would be determined by the selection of targets to be classified by a given recognition system, and the number of these functions would define the degree of certitude with which a



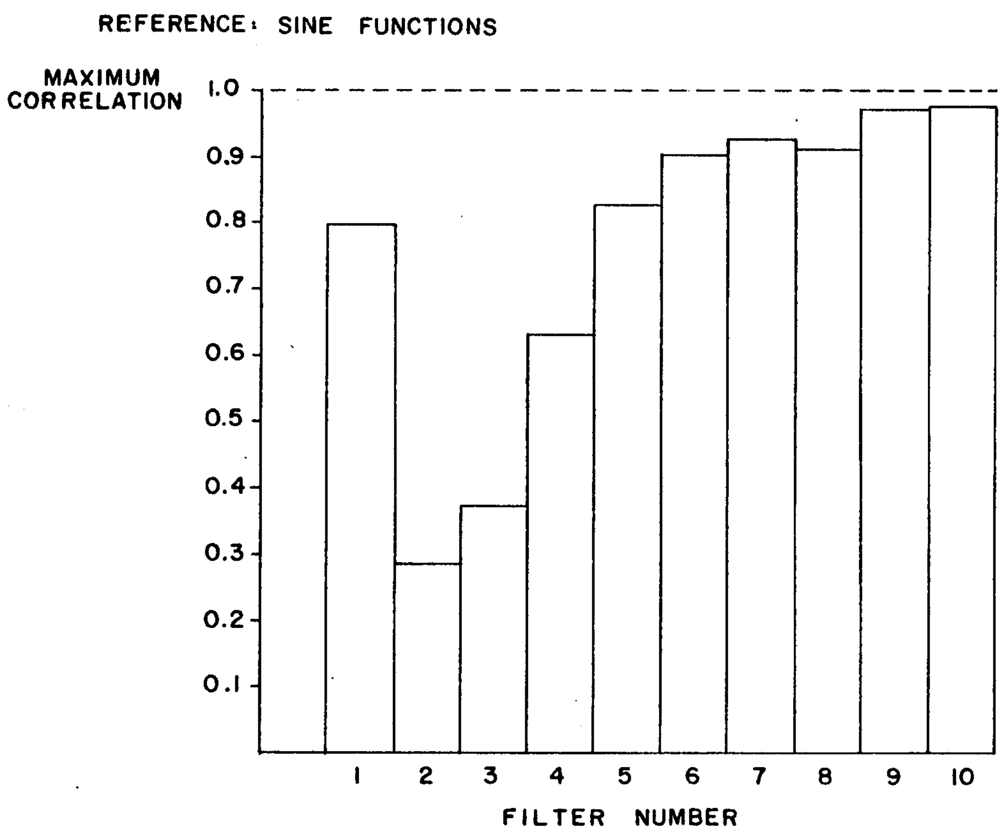


Figure 14. Correlation of Filter Outputs for Tanks-Versus-Background Experiments

classification is made. The development of an operational system using a time-domain and matched-filter approach would include such an assessment in the filter design.

3. Statistical Analysis

The primary emphasis of this program was on frequency-domain processing, but sufficient data was collected to perform the following test for rotational invariance under time-domain processing. The system was trained to recognize tanks against a background of nontargets, then a second set of tank imagery was processed through the system with the imagery oriented at controlled rotations from a fixed standard orientation. The following results were obtained:

TABLE 2. TIME-DOMAIN STATISTICAL DATA

ROTATION	0°	90°	180°	270°	TOTAL
CORRECT	45	45	45	43	178
ERROR	5	7	7	9	28
TOTAL	50	52	52	52	206

The maximum likelihood estimate p^*, for the probability of correct recognition is $p^* = \frac{178}{206} = 0.86$. Let us hypothesize that the probability of correct recognition, p^*, is invariant to image rotation. This hypothesis will be tested by the χ^2 test at a 5% level of significance. $\chi_p^{\ 2}$, for three degrees of freedom and p = 5, is 7.815. Calculating the χ^2 statistic as defined in equation (3) of section II-D on the basis of

Table 1, we obtain $\chi^2 = 1.29$. Since $\chi^2 < \chi_p^2$, we accept the hypothesis that the probability of correct recognition is invariant with respect to image rotation.

B. FREQUENCY DOMAIN

1. General Results

To avoid the confusion of many pages of figures and more clearly present the screening ability of the equipment, effort has been made to combine the large amount of data of the frequency-domain tests into two general classes: targets-versus-background tests, and targets-versus-targets tests. The former represents the basic level of discrimination; i.e., between man-made objects (e.g., tanks, roads, buildings, etc.) and natural terrain. The latter is a higher level of screening between target types, such as roads-versus-bridges, clearings-versus-woods, etc.

Targets-versus-Background:

The detections in each category are represented in Table 3(A); the percentages are in Table 3(B). The Type I error percentage is 23.1%; the Type II error percentage is 36.8%. The Type I error (missed targets) can be controlled during the experiments, but reduction of it increases the Type II error (false alarm). The costs associated with varying these errors determine their final values. Usually, missing a target incurs a larger penalty than causing a false

TABLE 3. TARGETS VS. BACKGROUND: TEST RESULTS

ACTUAL

DECISION		TARGET	NON-TARGET	TOTAL
	TARGET	362	150	512
	NON-TARGET	109	258	367
	TOTAL	471	408	879

(A)

ACTUAL

DECISION		TARGET	NON-TARGET
	TARGET	76.9%	36.8%
	NON-TARGET	23.1%	63.2%

(B)

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

alarm. For example, crying "wolf" over missiles in Cuba would have been embarrassing; missing them could have been disastrous!

Targets-versus-Targets:

The results of these tests are presented in Table 4, where the percentage of correct classification is listed for each of the two target types tested in each experiment. In evaluating these results, it should be remembered that, since these are two-class experiments, a misclassified Class I was identified as a Class II and vice versa.

TABLE 4. TARGETS VERSUS TARGETS: TEST RESULTS

TARGET CLASS I	CORRECT	TARGET CLASS II	CORRECT
Runway #1	84.2%	Runway #4	95.0%
Runway #2	94.7%	Runway #3	100. %
Roads	53.9%	Bridges & Dams	78.9%
Clearings	63.2%	Woods	58.0%
Contiguous* I	84.0%	Continuous #I**	60.0%
Contiguous II	72.3%	Continuous #II	71.1%
Point-targets	65.4%	Extended-targets	77.8%
Runways	70.0%	Urban areas	75.0%
Runways	16.7%	Orchards	96.7%
Runways	88.0%	Roads	80.0%

* Contiguous targets are man-made objects, mainly buildings and vehicles, the area of which occupies 25% to 50% of the scanner's field-of-view.

** Continuous targets are railroads, highways, bridges, and dams, all of which extend into and out of the field-of-view.

2. Diagnostic Tests

As the basis for investigating the parameters of the existing optomechanical scanner, sets of imagery representing various primitive shapes were produced. One set is of geometric shapes, all with equal areas and with bounds falling within the scanner field-of-view. The background of each test slide is opaque in respect to the transparent area of the geometric figure on that slide. These geometric shapes were designed for examination of the angular resolution, so that as many other parameters as possible were kept constant (e.g., equal areas of light transmission for each shape, limiting shapes to within the scanner field-of-view, etc.).

Essentially, the angular resolution of the integral scanner is a measure of the ability of the scanner to distinguish an n-sided polygon from an (n-1)-sided polygon. The two-class experiments were set up to distinguish between the shapes of most similar characteristics. To avoid any variation caused by system output, three samples of each figure were used as learn samples for that figure. Once the memory had adapted to the six learn samples (i.e., three per class), various rotations and translations of the figures were tested. The classes of figures tested and the percentage of correct classifications of the rotated and translated test samples are presented in Table 5.

TABLE 5. GEOMETRIC SHAPES: TEST RESULTS

CLASS I	CORRECT	CLASS II	CORRECT
Triangle	88.0%	Semi-Circle	92.0%
Rectangle	84.0%	Square	54.2%
Parallelogram	88.0%	Trapezoid	100%
Pentagon	100%	Hexagon	80.0%
Octagon	100%	Circle	100%

More important in this experiment, however, was an examination of the learning curve (the rate at which the memory was able to adapt to the shapes taught). Most laudable was the rapid convergence of the adaptive memory in distinguishing an octagon from a circle; only three learn passes were required. More than twice that number of passes were required in tests comparing shapes that to the human observer appear to be more easily distinguishable (e.g., triangle versus semicircle, etc.). A close investigation of the error rates involved in the learn cycle showed that scanner repeatability played some part in the varying convergence rates. As evidenced by the overall results, all of the geometric shapes were well within the angular-resolution capabilities of the scanner.

A more rigorous test of both the angular and the linear resolution of the integral scanner was implemented using linear spoked figures. The figures varied from two (i.e., a straight line centered on the field) to eight spokes in symmetric and unsymmetric configurations. The unsymmetric configuration involves a discrete angular displacement of one spoke of the symmetric figure. Each spoked figure was produced in line widths of 0.25, 0.15, .005, and .002 inches, and tests were run to discriminate between symmetric and unsymmetric configurations of the same figure. In addition, various translations and rotations were tested. Here, also, three samples of each figure were used as learn samples.

The rate of convergence to an adapted memory clearly showed, in all experiments, the ability of the scanner to discriminate between the symmetric and unsymmetric configurations. However, the overall results with the test samples produced a 52.1% correct classification level. This, although better than chance, seemed to indicate that the information content of the video signal, particularly for the figures of the smaller line widths, closely bordered the video noise-level caused by non-uniform slit-widths in the scanner; i.e., slit-width dimensions approximating those of the spoked figures. This area is definitely one that could be improved by using an electronic flying-line scanner.

3. Statistical Analysis

The following test was conducted to determine the effects of imagery rotation on frequency-domain analysis. The system was trained to recognize buildings against a background of cultural features, then the imagery was rotated through a number of angles. The recognition rate at the end of the rotation was noted, and the results are presented in Table 6.

TABLE 6. FREQUENCY-DOMAIN STATISTICAL DATA (ROTATION)

Rotation	30°	60°	90°	180°	210°	240°	270°	300°	330°	Total
Correct	12	10	7	9	12	11	11	7	13	94
Error	3	5	8	6	3	4	4	6	2	41
Total	15	15	15	15	15	15	15	15	15	135

The maximum likelihood estimate, p^*, for the probability of correct recognition is $p^* = \frac{94}{135} = 0.70$. Again, our hypothesis is that the probability of correct recognition, p^*, is invariant to image rotation. Using a five percent level of significance, we find the χ_p^2 is equal to 15.507 for eight degrees of freedom. Calculating the χ^2 statistic on the basis of Table 6, we obtain $\chi^2 = 9.85$. Since $\chi^2 < \chi_p^2$, we accept the invariance hypothesis.

Finally, a test was run in which a class of several types of targets was compared with mixed backgrounds. After the system was trained on imagery presented at a standard orientation, recognition tests were run with the imagery both translated and rotated. The results are presented in Table 7.

TABLE 7. FREQUENCY-DOMAIN STATISTICAL DATA (ROTATION-TRANSLATION)

Orientation	45°	270°	0.2" down	0.2" left	Total
Correct	29	30	30	31	120
Error	10	8	9	8	35
Total	37	38	39	39	155

The maximum likelihood estimate for the probability of correct recognition is $p^* = \frac{120}{155} = 0.77$. The hypothesis in this case is that the recognition rate is invariant to both rotation and translation of the imagery. The 5 percent value of χ^2 for three degrees of freedom is 7.815. In accordance with Table 7, χ^2 is 0.44. Hence $\chi^2 < \chi_p^2$, and our hypothesis is acceptable.

The above results support the hypothesis that the recognition capabilities of the system are invariant to rotation and translation of the imagery.

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

IV. CONCLUSIONS AND RECOMMENDATIONS

As has been explained in the introduction of this report, several changes in final objectives took place during the program as a result of sponsor request. These changes affected primarily the physical products of this program. Although it had originally been planned to supply diagrams and fabrication data for an automatic target recognition system, such highly detailed data is not now included due to redirection of the effort by the sponsor; however, the basic information necessary to build such equipment is included in this report. Based on our research our conclusions are as follows:

1. The equipment can extract parameters from visual imagery that are representative of the imagery and are invariant to translation and rotation.

2. A scientific basis for the fabrication of a screening device has been established.

3. Discrimination of man- and non-man-made or significant and nonsignificant imagery can be made by automatic equipment with a high degree of accuracy.

4. Automatic equipment, combined with the cognitive processes of the human interpreter, would be a potentially very significant combination for image processing.

5. Improper training or "over-training" may have occurred because the decision logic was trained until no errors were present. This implies that the examples shown to the training

routine included not only good examples of the class but very poor examples of the class. Thus, when the actual testing program was run, a rotated and translated image was being compared with good and poor examples of a class. The degradation in the class model set up by this method may have caused some of the error.

Based on our research, our recommendations are as follows:

1. Fabrication of a prototype screening device that can be used with humans to test all kinds of militarily significant imagery.

2. Continuation of research in the fields of scanning, preprocessing, and decision logic.

3. Existing equipment should be fully exploited and continued extensive testing be conducted on it.

4. An electronic analog of the opto-mechanical scanner should be developed. The capability of such a scanner would exceed that of a serial flying-spot scanner and only the optical techniques of spatial filtering and hologram generation could come close to matching the versatility of the integral scanning approach to image processing. In addition, the electronic or opto-mechanical integral scanner is not affected by those problems which affect the optical systems used in the other systems.

5. The two different kinds of video processors should be compared. Thus, the effects of varying image quality, image content, and target density, on the performance of the time-domain filtering and the frequency-domain filtering could be examined.

Each one of these prenormalizers may have a group of targets for which it will have optimum performance. The time-domain analyzer also is flexible in that the weighting functions can be changed easily. Thus, the current sine wave orthogonal functions can be replaced with other orthogonal functions. Some comparisons have been made using Laguerre functions, and further investigation would be beneficial.

In summary, it is evident that many useful ideas have been generated by this program. These ideas will permit the fabrication of a prototype of an operational imagery screener. However, what is more important is the fact that a capability has been established that has only begun to be exploited. With a suitable, well-controlled data base of 10 or 20,000 examples of imagery, meaningful results can be made immediately available to the experimenters. It is felt that it would be a waste of the sponsor's investment not to continue to use this equipment. While further tests are conducted using this existing hardware, the fabrication of an all-electronic flying-line scanner could commence for later incorporation in the system.

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

APPENDIX I

INTEGRAL-SCAN PARAMETER-GENERATOR

INTEGRAL-SCAN PARAMETER- GENERATOR

INTRODUCTION

"Prenormalization" designates the removal of unessential variability (i. e., variability which has no effect on the class membership of a pattern) from patterns prior to pattern-learning and pattern-recognition operations in order to achieve invariance to translation and rotation. For example, in target-detection and target-classification applications of prenormalization, knowledge of the position and orientation of the target with respect to some limited viewing is unessential information.

THE INTEGRAL SCAN

A means for extracting image data and retaining the essential characteristic of image continuity has been developed at [redacted] during a study of complex image analysis. The study, beginning with analyses of various optical transforms, resulted in a novel scanning technique utilizing an integral rather than a point scan of the image field. This prenormalization approach is based on the time-domain analysis of signals obtained by the integral-scan sequence developed at [redacted]

25X1

25X1

The ensuing brief description of the integral-scan technique and of the motivation behind it is presented with reference to figures 1 and 2. Figure 1 shows an idealized airfield being scanned by an extremely narrow slit. The slit is being moved across the viewing field bounded by Y_0, Y_1, and X_0, X_1 at a certain velocity. The output due to the airfield is also shown in figure 1. It should be noted that the integral scan maintains image continuity and connectivity information that is normally lost with the conventional spot-scanning techniques. An important feature of this method of scanning is that the waveshape produced by an object is invariant to translation. If the airfield is moved up or down, the entire waveshape remains unchanged; if the airfield is moved right or left, the waveshape due to the airfield remains unchanged and only its time of occurrence varies.

In figure 2, another shape of airfield is being scanned by four slits in the order indicated. The waveform resulting from the scanning operation is also shown. If the airfield is rotated within the resolution capability of the scanning arrangement (i.e., in 45° steps), the only change that results is a change in the phase (i. e., time-order) of the signal. If a large number of slits, each at a slightly different angle, are used, then the rotational resolution of the scanning process will be quite good and the only effective change due to any rotation would be a phase change (i. e., change in time-order).

Thus, for the integral scan with many slits having small angular displacements, the only changes caused in the output waveform are changes in the time of occurrence of the waveshapes due to the objects in the field of view. Therefore, if a method of processing this waveform independent of the time of occurrence is used, a set of measures (or properties) invariant to translation and rotation can be obtained.

SYSTEMS UTILIZING THE INTEGRAL SCAN

The integral-scan approach to pattern recognition has been incorporated in several research programs at [redacted] primarily in the fields of character recognition and prenormalization of aerial reconnaissance photography. Under these programs, several scanners have been built and are presently in operation in conjunction with electronics using time-domain and frequency-domain analysis, and subsequent computer processing to examine and distinguish various classes of primitive shapes, character fonts, and aerial imagery. At the same time, examination of the parameters and limitations of these existing scanners has led to the design of new and improved methods of implementing the integral scan for use in future systems.

25X1

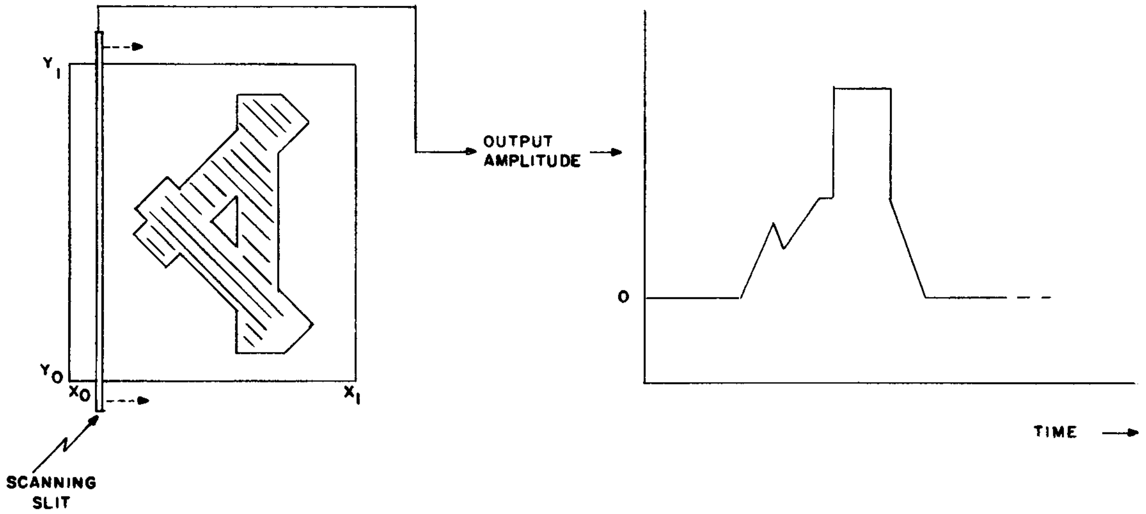


FIGURE 1

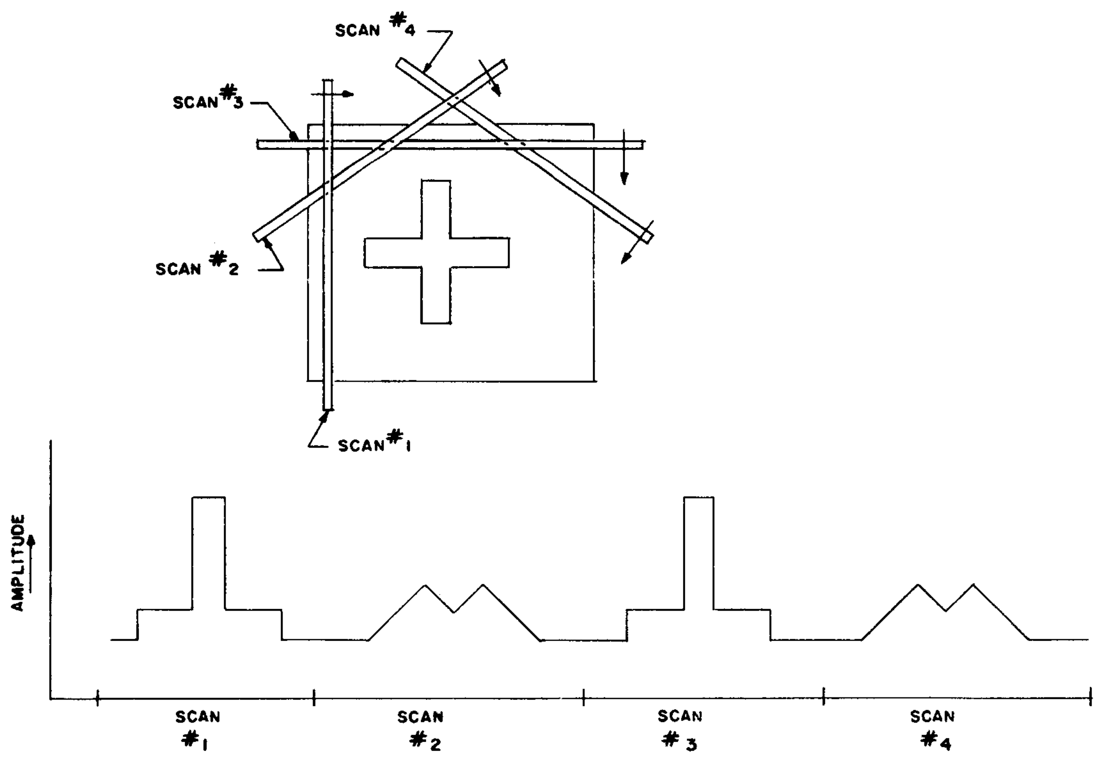


FIGURE 2

APPENDIX II

GENERAL COMMENTS
ON
AUTOMATIC TARGET-RECOGNITION RESEARCH

GENERAL COMMENTS
ON
AUTOMATIC TARGET-RECOGNITION RESEARCH

The time for actually fabricating the equipment is here! Technology has increased the generation of reconnaissance or intelligence information, but the photointerpreter is still limited to traditional PI aids. Thus, very little has been done to alleviate the "dumbbell effect" (figure 1). The ability to utilize the large amount of imagery now being acquired is severely restricted by current data handling capabilities. Consequently, the data necessary for military decisions cannot flow through to the users in a short enough time to be wholly useful. Recent experiences have shown that emergencies can quickly strain imagery-interpretation capabilities and thus further restrict this flow, but this appendix will make no effort to discuss these problems because they fall outside the scope of the research.

The existing equipment will be quite valuable in the field of automatic research. Because the parameters used in this program to describe the imagery are being generated independent of a decision logic, it is possible to use them in any available decision logic. Thus, the integral scanner can be used as a standardized parameter for testing differing decision logics. This application would not necessarily be based upon its invariance characteristics but only upon the uniqueness of the equipment and the parameters it generates.

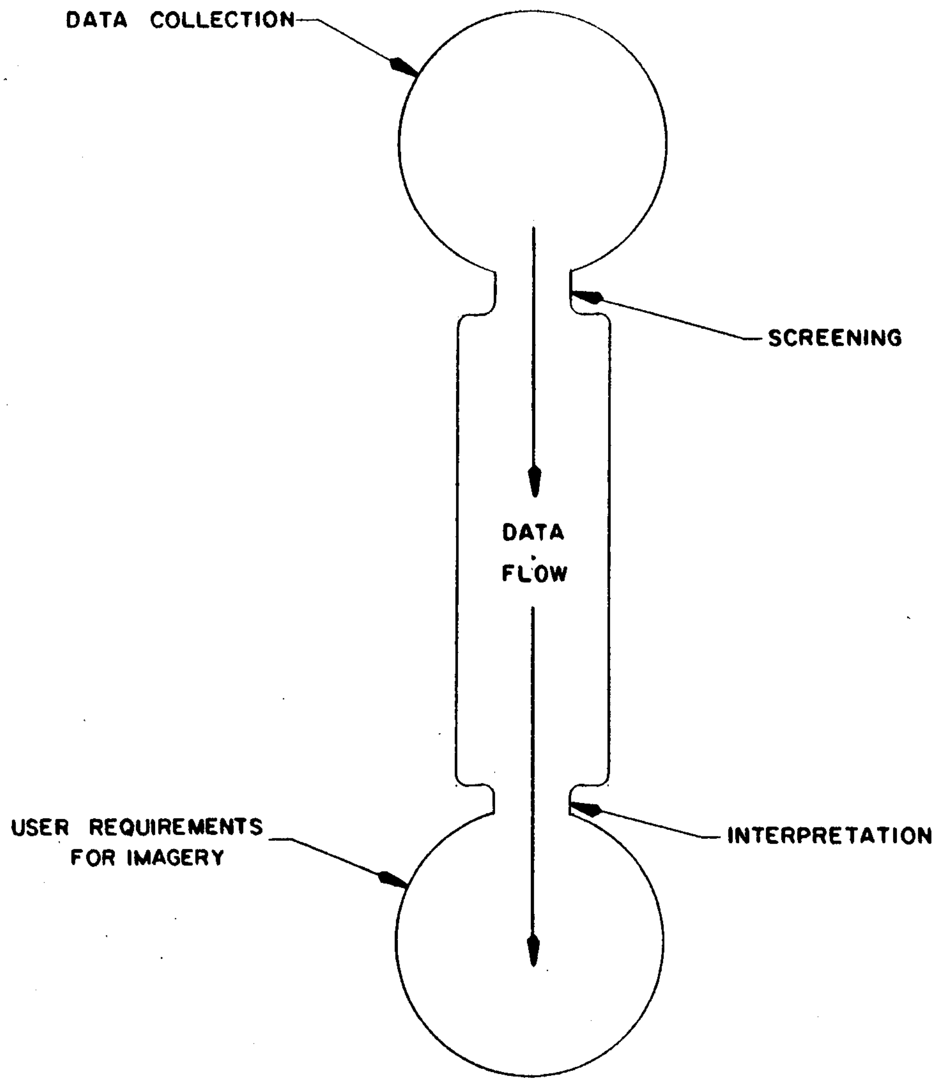


Figure 1. Dumbbell Effect

The original criteria for selecting test imagery should be supplemented with an approach having a sound theoretical basis. First, sufficiently high-quality imagery should be used in order that system performance can be evaluated under conditions of uniform degradation. Target density and target detail should be varied. Low target-density could be defined as one target per photograph; high target-density, as two or more targets. Low target-detail could be defined as seven or less significant features per target, which could be clearly seen on the negative; high target-detail could be regarded as eight or more such features. These photographs should also be processed to various levels of contrast, resolution, and scale.

Of equal importance, looking at the broader picture of the research effort in automatic target-recognition, is the establishment of a common data base. Several organizations are working on varied approaches to this problem. Ultimately, the effort should converge at a common point, the recognition of a target. But what is a typical target? Because of the nature of aerial photography, this can be only generally defined. But despite this factor, a common data base can be established for use in comparing the efforts of individual companies. Without such a data base to use as a standard, it would be possible for one device to classify a photograph as an airfield and for another device to classify the same photograph as a bridge. The inbred idiosyncrasies of each company's target-recognition device seem to have been the determining factor in what photography has been used for testing, since good results promote any program. For this reason, it is certainly difficult for a contractor to compare the research efforts of various companies, or for those companies themselves to know where they stand in the overall picture.

Effort should be made, in cooperation with the agencies that are the sources of the aerial photography needed, to establish a common data base to provide this basis for comparing all target-recognition equipment. Since there is no representative aerial photograph of any target, this data base could take the form of a library of indexed film rolls containing multiple examples of many classes of targets under varied conditions (e.g., rotation, scale change, environment, seasonal change, film quality, etc.). Were such a data base available to each company involved in target-recognition work, it would benefit not only the sponsor but the other individual companies. Duplication of effort could be avoided, and system merits could be more equally weighed.

Another area of importance is a review of psychological test programs. These programs have performance data for human beings with various kinds of stimulus materials. With such materials, it is possible to compare the performance of the automatic equipment with that of the unassisted human. The best approach would be to have direct man-versus-machine comparison. However, until such experimentation is possible the use of previously scaled imagery will be sufficient. There are two screening experiments for which human performance has been established. One, conducted by [redacted] 25X1
25X1 [redacted] under Air Force sponsorship, has

already been reported in RADC-TDR-63-313. The second, conducted [redacted] under Army sponsorship, is in the approval stage. It appears that the IBM report will contain much human performance data that would be useful for these purposes.

Several programs have been conducted to demonstrate a relationship between image parameters and photointerpreter performance. Unfortunately, most of the programs have been inconclusive, partially because the kinds of image parameter measurements that can be made are almost endless. Because it is desired to demonstrate that automatic equipment is competitive with human performance , selections of image parameters should be restricted to parameters already used in experiments on human-factors.

There are two main sources for information on such experiments. The United States Army Personnel Research Office has started an extensive program in screening research that includes contracts to outside sources as well as in-house research. These programs can provide guidelines for the experiments that should be conducted on automatic equipment. It has always been felt that the image quality of reconnaissance photography would affect the performance of the photointerpreter. Resolution, contrast, and scale also affect this performance. The Air Force at Wright-Patterson is supporting a program to develop other measures of photo quality. About all that has been discovered from these various experiments is that it is difficult to relate photointerpreter performance directly to any single or group of image parameters. However, it would be wise to make some quality measurements on the imagery as a data base is developed. Until some kind of agreement can be reached by the experimenters, an extensive image mensuration program would be a waste of time. The best approach would be to limit measurements to target contrast and to image density, detail, scale, content, and context.

Four statistical values which can be useful as diagnostic tools for determining the characteristics of the distributions of imagery parameters are measures of the following:

1. Central tendency.
2. Dispersion, including range, average or mean deviation, or standard deviation.
3. Skewness.
4. Kurtosis.

These measures will be valuable if there is a univariate distribution. If there are multivariate distributions, the calculations of the covariance matrix can be used as a useful diagnostic tool.

Many statistical approaches have been developed that can be used in pattern-recognition experiments. However, the application of many of them depends on a greater volume of more controlled data than is now on hand for testing.

Many considerations for future research have been included in this appendix, some generated as the result of this program and some reflecting the general attitude of [redacted] 25X1

Approved For Release 2005/11/21 : CIA-RDP78B04770A000200020017-1

towards this area. It appears that the successful conduct of research now depends almost solely on an ample supply of imagery. Every effort should be made to assure that such a supply is available to the researchers.