R. M. Centner*
E. D. Hietanen
Bendix Research Laboratories
Southfield, Michigan 48075

Automatic Pattern Recognition

Applied to photogrammetric operations, it would provide significant benefits over present annual methods

Abstract: *Automatic pattern recognition is an important new tool for the processing of photographic image data. Significant success has been achieved with an adaptive pattern recognition technique wherein a decision logic network is "trained" to classify imagery. In the training process, the network accumulates statistical information from a variety of image samples of known classification. Once it has been trained, the network can classify additional image samples with a high degree of accuracy. This technique has been successfully applied to two significant problems in automatic photomapping: (1) pass-point selection and (2) planimetry extraction. This paper describes the logic technique and the statistical training method, summarizes experimental results to date, and discusses the overall potential of automatic pattern recognition.*

Introduction

A number of key tasks in photogrammetry require interpretation of complex image patterns. The interpretation may involve identification of specific objects or planimetric features, or it may involve judgment decisions such as determining the relative quality of pass-points. Up to now, interpretation tasks of this nature have been performed by skilled operators. Operators who have had considerable training and experience can perform such tasks with a high degree of accuracy and efficiency. However, even highly skilled operators do not always work at peak efficiency, and the tedious nature of the task can inhibit both accuracy and efficiency. Also, the supply of highly skilled personnel is limited, a problem which becomes increasingly severe as the quantity of photographic image data to be processed grows rapidly.

The technology of automatic pattern recognition offers a means for automating photogrammetric interpretation tasks, thereby alleviating the above-stated problems. Significant research and development effort is now being applied to this area. In general, the term *automatic pattern recognition system* refers to a computer-like system which accepts input image data and generates output signals which identify or classify the image data in some way. A general system configuration is shown in Figure 1. An electro-optical transducer (generally a scanner) converts the image data on a small area of photography to electrical form. A preprocessor extracts parametic properties of the image data and supplies these to a decision logic network which generates the output signals. The nature of the output signals is of course dependent on the specific application; they might provide one or more of the following indications:

- Presence or absence of specific objects (buildings, ships),
- Terrain classification,
- Presence or absence of specific planimetric features (roads, shorelines),
- Relative uniqueness of terrain (for pass-point selection).

An overall system controller operates to position the image transducer relative to the film so as to process sequentially an entire photographic frame. Depending on the ap-

* Presented at the Annual Convention of the American Society of Photogrammetry, Washington, D. C., March 1969.

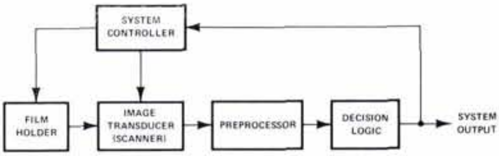

FIG. 1. Configuration of system for automatic pattern configuration.

plication, the controller may use feedback data from the system output to determine the positioning strategy. For example, it may be desired to simply track a road rather than process an entire frame and classify all areas as *road* or *no-road*.

The following sections of this paper describe a pattern-recognition approach developed by Bendix for photogrammetric applications. This approach emphasizes the use of *trainable* design techniques, wherein the system is trained (much like an unskilled operator) by repetitive application of sample imagery of known classification. Reward-punish signals are provided by a *teacher* in accordance with the decisions made by the system being trained. Gradually, the system builds up a memory of the correct decisions for various types of imagery, until it reaches a trained state and is ready to be separated from the teacher and used for actual operation. At this point, it is analogous to a skilled operator.

The next section describes the basic approach, and subsequent sections describe two specific applications: (1) automatic selection of pass-points and (2) automatic planimetry extraction.* The final section indicates other potential applications, summarizes advantages, and indicates present problem areas and research needs.

GENERAL DESIGN APPROACH

In designing pattern-recognition systems, the major emphasis is generally placed on the preprocessor and decision logic stages. These stages, as shown in Figure 1, generate a system output based on signals provided by the image transducer. The design of the preprocessor and decision logic is accomplished by the general design approach indicated in Figure 2. This approach is described in the following paragraphs.

* This work was sponsored by Rome Air Development Center, U. S. Air Force, Automatic pass-point selection was investigated under Contract AF 30(602)-3417, and automatic planimetry extraction was investigated under Contract AF 30(602)-4139.

The function of parameter evaluation is to determine an optimum preprocessor for a given pattern recognition problem. This is accomplished by analyzing various sets of parameters and determining a relative significance rating of the parameters in each set. Several reasons occur for performing parameter evaluation. From a data-processing standpoint, it reduces the number of parameters which must be considered in designing the decision logic stage. It also reduces the number of parameters which must be implemented in the ultimate fabrication of the pattern recognition system.

Parameter evaluation is generally based on two considerations. One consideration is the amount of information contributed independently by each parameter. This determines how well pattern recognition can be achieved using each parameter by itself. The second consideration is the degree of dependency between parameters. For example, if two parameters are rated highly by the independent consideration but are also highly dependent (or correlated), it may be desirable to use only one of the two parameters. Parameter evaluation techniques generally use tradeoffs between the two considerations to arrive at an overall optimum selection.

The decision logic design determines a means of recognition based on the selected parameters. In other words, a decision logic stage is designed which receives, as inputs, the values of the selected parameters and provides the classification (such as pass-point quality) which the given sample represents. The decision logic may be designed by either incremental or statistical approaches. In the incremental approach, which is essentially an adaptive or trainable process, adjustments are made to the decision logic for individual samples of parameter values. After adjustments for a large number of samples, the decision logic achieves a *trained* state. In the trained state, the decision logic is able to make correct classifications based on parameter values alone. In the statistical

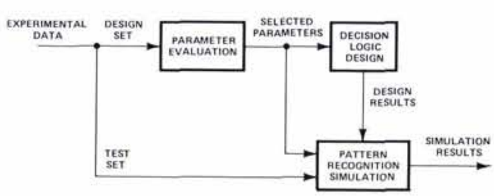

FIG. 2. Design approach for pattern recognition network.

approach, a large number of samples are examined simultaneously to determine means of classification. The major tradeoffs between the incremental and statistical approaches are data storage and speed. Incremental approaches generally require less data storage but require a longer training time than statistical approaches.

Following the design of the decision logic, the next step is to simulate the overall pattern recognition design. The simulation determines how well the design performs in correctly classifying new input patterns, and indicates whether additional refinement is needed. The simulation is usually tested with samples that were not used in parameter evaluation or decision logic design. The objective of this type of simulation, as described further below, is to determine how well pattern recognition can be accomplished for previously unseen samples. This provides an indication of the expected operational performance of the pattern recognition system.

All of the above steps can be implemented by computer programs. This enables detailed studies of the effects of various classification strategies or the exclusion of certain parameters. Performing these studies by computer simulation is considerably more efficient than other means.

The general procedure in the computer simulations is to use a portion of the experimental data to design the pattern-recognition system, and the remainder of the data for testing purposes. These two portions of data may be designated as the design set and test set, respectively. The performance of the pattern recognition system is generally somewhat better for samples included in the design set than for those in the test set. This difference in performance is a result of imperfect extrapolation to previously unseen test data. As the number of samples is increased, however, the performance for the test set generally approaches the performance for the design set. The difference in performance between the design and test sets can be used as a criterion to determine whether a large enough number of samples has been used for training.

The simulation performance may also be used to indicate any corrective measures that may be required. For example, if a given group of classes is frequently misclassified, parameter evaluation may be repeated in order to locate parameters which are particularly helpful in detecting the problem classes. In addition, the decision logic configuration may be modified to add distinguishing capability for the problem classes.

The use of this basic pattern recognition approach for automatic pass-point selection and for automatic planimetry extraction is described in the next two sections.

Automatic Pass-Point Selection

BASIC CONCEPT

The objective of this investigation was the development of automatic techniques for searching a designated area of a photograph in order to select and locate precisely one or more points suitable for use as pass-points. The selected pass-points are to be subsequently employed in an automatic point-transfer device with the objective of accurately determining the photo coordinates of corresponding points on other photographs. In this process, the imagery surrounding a selected pass-point is automatically matched by means of correlation techniques with the imagery surrounding the corresponding points on other photographs. Therefore, in the selection of pass-points, the basic objective is the location of points on the photograph where the subsequent automatic image matching process can be performed with high precision.

An automatic pass-point selection system accepts as input information a pair of overlapping photographs and produces as output information the locations of the selected pass-points. In addition to the photography, the pass-point selection system accepts specifications of the areas in which pass-points are to be selected and specification of the number of pass-points to be selected. A selected pass-point may be defined by a photographic image of the selected pass-point and its surrounding area together with a superimposed reference mark or marks, and/or the coordinate location of the pass-point on a reference photograph. A pass-point selection system might also accept other information to facilitate the selection of suitable pass-points and might provide other output information to facilitate the point-transfer of the selected pass-points.

This investigation used a breadboard measurement system which consisted of two flying-spot scanners, correlation and other special-purpose analog circuitry, and a punched-card recorder. Data collected by the *breadboard* measurement system was subsequently processed by a set of computer programs to design a pattern recognition system. The following paragraphs describe this investigation and the results achieved in further detail.

BREADBOARD MEASUREMENT SYSTEM

The basic requirement of the breadboard measurement system was to measure photographic information indicative of pass-point quality. Three basic considerations were desirable in selecting the measurements taken: first, the expected significance of the parameter; second, the complexity of the required measurement circuitry; and, third, the compatibility with the general characteristics of a point-transfer device. Both stereo- (two images) and single-image measurements were used.

Both the stereo- and single-image measurements are based on scanning the photo image(s) using a cathode-ray tube (CRT) and photomultiplier tube (PMT). Basically, the PMT measures the amount of light from the CRT which is transmitted by a small point in the photo image. By varying the position of the light source using CRT deflection signals, the resulting video signal from the PMT describes the variations in transmittance in the photo image. The simultaneous scanning of each of a conjugate pair of photo images yields two video signals. One of these signals is used to extract the single image measurements, and both are used to extract the stereo measurements.

The stereo measurements were extracted using electro-optical correlation techniques. Basically, correlation of the two video signals provides an indication of the degree of alignment of the conjugate imagery. A total of 52 stereo measurements were extracted for each of a large number of samples of conjugate. imagery.

Several properties of the correlation function itself were measured. One property, for example, was the peak value of the cross-correlation function, and another was a measure of the slope of the orthogonal correlation function in the X and Y directions. These properties are indicative of the quality of the image match and the sensitivity of signals which control the positioning of the photographs. Terrain shape properties were also extracted by the correlation technique. The average terrain slope for the area scanned was obtained from measurement of the scan shaping required to produce the highest correlation peak. Scan shaping modifies the X and Y deflection signals applied to the CRT's to compensate for the relative change in scale due to terrain slope. Also, measures of terrain roughness and curvature were obtained by making several relative elevation measurements around the center of the scan area.

An additional 37 measurements were extracted from the video signal obtained by scanning a single image. These measurements consisted of both amplitude and time information.

The amplitude measurements were the average DC and AC video signal level, and a special type of measurement which describes the basic structure of the signal amplitude distribution. The latter measurement was essentially the ratio of outputs of a linear rectifier and square-law device applied to the video signal. This ratio indicates whether the signal is basically random, or resembles a triangular, sine, or square wave.

The time measurements were basically the first and second moments of the zero-crossing interval of the video signal. The first moment indicates the average spatial frequency of the image scanned, and the second moment indicates the amount of variation in spatial frequency.

DESIGN OF PASS-POINT SELECTION SYSTEM

To design a pattern-recognition system for determining the pass-point quality of an image area, the above described measurements were extracted from a large number of randomly selected potential pass-points. In addition, each potential pass-point was examined visually to determine the probable permanence of the surrounding imagery. Permanence was determined by the class of imagery, which, for the photography used, was (1) cultural, (2) barren, (3) wooded, or (4) cultivated. These classes were assigned decreasing performance ratings, that is, cultural areas were estimated to have the highest permanence, followed by barren, wooded and cultivated in that order. A numerical weighting factor was then assigned to each permanence class. A pass-point figure-of-merit was then defined as the product of the permanence weighting factor and a quality estimate derived from the measured parameters. The pattern recognition problem was then defined as the problem of estimating the pass-point figure-of-merit from the measured parameters alone.

The first step performed after data collection was parameter evaluation. As mentioned above, parameter evaluation is the process of selecting the parameters to be provided by the preprocessor. This was accomplished by means of a technique known as *correlation analysis*. Correlation analysis entails the computation and analysis of a matrix of quantities known as *correlation coefficients*. These coefficients indicate the degree of linear relationship between all possible pairs of parameters and also between each single parameter and the figure-of-merit. A high correlation coefficient between a pair of parameters indicates redundancy; that is, they contribute essentially the same information. A high correlation coefficient between a given parameter and the figure-of-merit indicates a high degree of significance for the given parameter. The program developed for parameter evaluation used a sequential selection process which used both redundancy and significance information. Application of this program resulted in the selection of 24 parameters for use in the subsequent design steps. The 24 parameters consisted of 13 parameters from stereo measurements and 11 parameters from single image measurements.

The next step was the design of the decision logic used for classification. The decision logic consisted of two stages: a parameter quantizer and a threshold-logic network. The parameter quantizer is essentially a filter which extracts significant data from a set of analog-valued parameters. The quantizer encodes the values of the various parameters according to statistically determined quantization levels. These quantization levels are the values at which the bits representing the parameter will undergo transitions. For example, in the quantization of a measured parameter P_m, the value of a given bit f_n at the output of the quantizer is given by

$$f_n = +1 \qquad \text{if } P_m \geq \theta$$
$$f_n = -1 \qquad \text{if } P_m < \theta$$

where θ is a quantization level for the parameter P_m. The optimum set of quantization levels for a given set of measured parameters is one for which the output pattern contains a maximum of information indicative of the figure-of-merit of the potential pass-point.

A computer program was used for the selection of quantization levels. This program is flexible so that some parameters may be encoded into four or five bits, and others into only one bit. The main point is that each parameter is encoded in a manner that will transmit the most useful information to the threshold logic network.

The threshold logic network implements the pass-point figure-of-merit based on the output pattern from the parameter quantizer. The term *threshold logic* denotes a general class of logic networks which have the property of being trainable. A schematic diagram of a conventional, or linear, threshold-logic circuit is shown in Figure 3. The binary inputs to the circuit, $f_1, f_2, \cdots, f_n$,

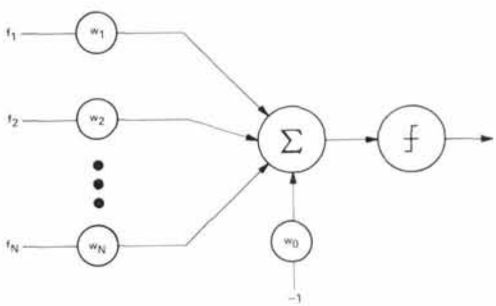

FIG. 3. Circuit for linear threshold logic.

are weighted, or multiplied, by quantities, $w_1, w_2, \cdots, w_n$. If the analog sum of the weighted inputs is greater than the threshold w_0, the circuit produces a $+1$ output, otherwise, a -1 output is produced. Training of the circuit is accomplished by adjustment of the weights and threshold whenever the circuit produces an incorrect output.

A single linear threshold logic circuit becomes inadequate as the size of the input pattern $f_1, f_2, \cdots, f_n$ increases. The common procedure is to use more than one linear threshold-logic circuit in such instances. However, where several circuits are used, the procedure for training becomes much more complex. Thus, the linear threshold-logic approach is severely handicapped in large-scale systems.

To overcome these disadvantages, a more powerful technique known as *nonlinear* threshold logic has been used. This approach offers a significant improvement over linear threshold logic in the number of functions that can be realized, and in the ease of training. A nonlinear threshold-logic circuit differs from the linear form in the presence of a variable cross-multiplier stage. This stage is placed between the input variables and the weights, as shown in Figure 4. The outputs of the variable cross-multiplier stage, $\beta_1, \beta_2, \cdots, \beta_Z$, are independent cross-products of certain combinations of the input variables. The training procedure for the nonlinear threshold logic circuit uses a procedure for

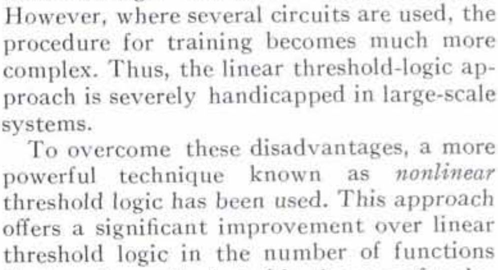

FIG. 4. Circuit for nonlinear threshold logic.

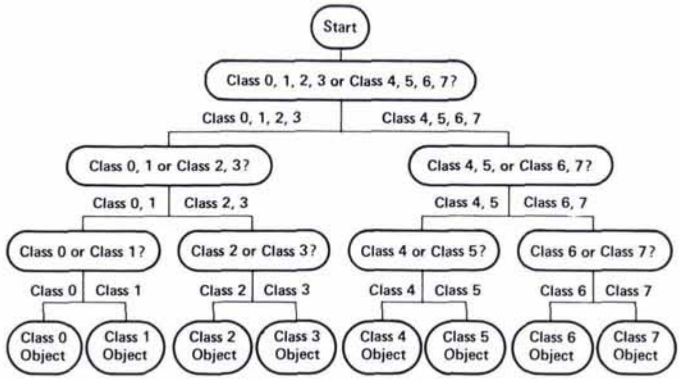

FIG. 5. Sample decision tree.

selecting an optimum set of cross-product terms and finding appropriate weights.

A single nonlinear (or linear) threshold logic circuit provides a single binary output of the *yes-no* type. In order to obtain more than two levels of figure-to-merit classification, several circuits must be used. One method of using multiple circuits is in a form known as a *decision-tree*. One type of simple decision tree for detecting eight levels of pass-point quality is shown in Figure 5. Each encircled statement represents a single circuit designed to realize a particular function. For example, the statement *Class 0, 1, 2, 3* or *Class 4, 5, 6, 7?* indicates a circuit designed to produce a $+1$ output for Classes 0, 1, 2, 3 and a -1 output for Classes 4, 5, 6, 7. Depending on the first output, either the statement to the left or right is next applied.

It should be noted that implementation of a decision tree does not require a large group of physically separate circuits. A single time-shared circuit can be used repetitively, with weights and cross-product terms for each part of the decision tree stored in an external digital memory.

EXPERIMENTAL RESULTS

Results obtained from application of the nonlinear threshold logic and decision tree

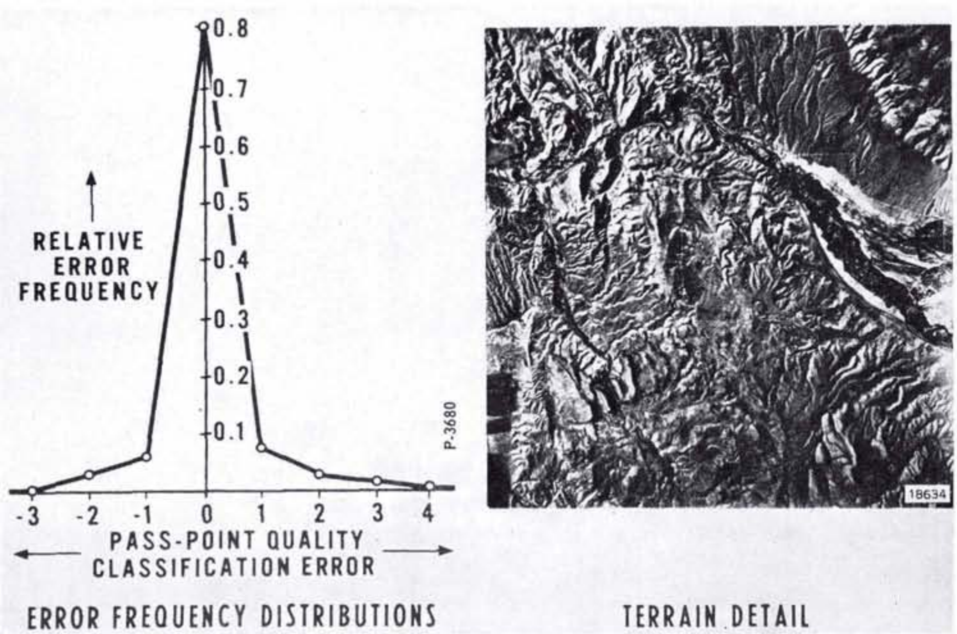

FIG. 6. Error histogram of pass-point results.

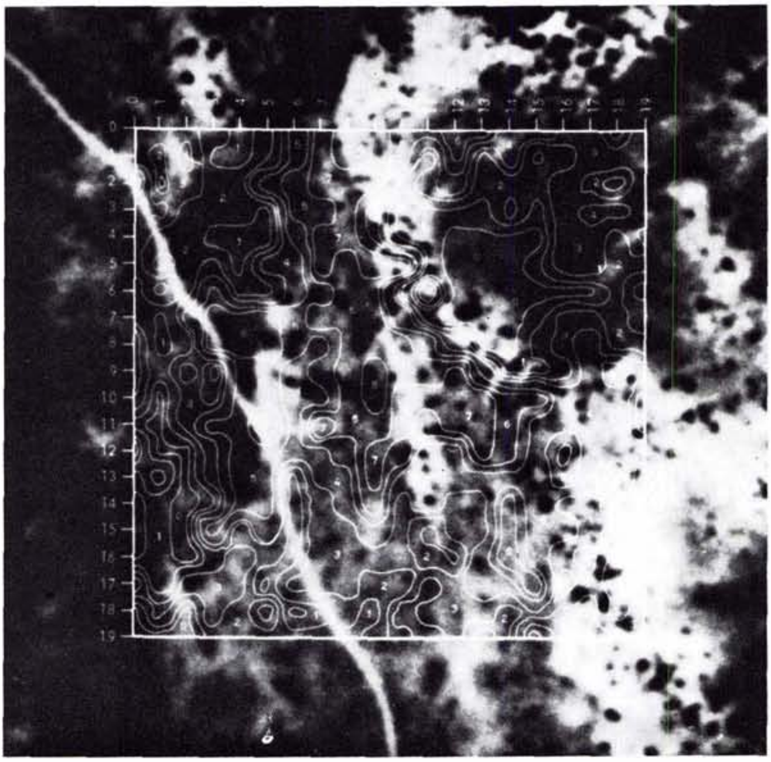

FIG. 7. Contour representation of pass-point results.

concepts have been quite encouraging. One illustration of these results is shown in the error histogram in Figure 6. These results were obtained for 160 test points, that is, samples not used to design the pattern recognition network. In this study, the pass-point figure-of-merit was divided into eight groups. Each group contained approximately the same number of design samples, and the group numbers were a linear indication of pass-point figure-of-merit. As shown in Figure 6, approximately 80 percent of the test points were classified correctly, that is, they produced the same figure-of-merit class as the visual evaluation. Also, more than 95 percent of the points were within one class of the correct value.

Another form of results can be shown by superimposing figure-of-merit values over the corresponding photograph location. This type of display, shown with contour lines connecting points of equal figure-of-merit, is shown in Figure 7. The information used to determine these contours was obtained by taking measurements at 400 uniformly spaced points and computing figure-of-merit values for each point. Because many peaks (or high figure-of-merit locations) occur in the contour representation, a two-mode strategy would be required to select optimum pass-points. A coarse search would first be used to location regions of high figure-to-merit, and then a finer search would be applied around these regions.

AUTOMATIC PLANIMETRY EXTRACTION

CLASSIFICATION TYPES

Another investigation was conducted of techniques for the automatic extraction of planimetric features from aerial photographs. Emphasis was placed on extraction of four gross area-type features, namely, urban areas, vegetated areas, cultivated areas, and hydrographic areas. Attention was focused on techniques applicable to the AS-11B type of analytical stereoplotter which uses electro-optical scanning and computer-controlled plotting.

This investigation considered the basic problem of classifying the type of planimetry present within a single small section of photograph. This type of information could ultimately be used to trace boundaries of planimetric area types, or to perform a point by-point analysis of an entire photograph.

BREADBOARD MEASUREMENT SYSTEM

The breadboard-measurement system used for this investigation was based on the system used in the pass-point investigation. Several modifications were made, however, in the particular measurements extracted.

A total of 37 stereo measurements were extracted for each sample of conjugate imagery. The measurements extracted were similar to those taken under the pass-point investigation, although changes were made in scan sizes and scan positioning techniques. For example, the area scanned was displaced about the central area in a different way. Previously, the size of the area scanned was varied to determine curvature and roughness. In this investigation, however, the size of the area scanned was held constant and its position was varied. This approach offers higher sensitivity and is simpler to implement.

More extensive changes were made in the circuitry for extracting image measurements. A total of 46 measurements were extracted for samples in this group. The major change in the single image measurements was the extraction of measurements using 11 reference levels. These reference levels, which were not used in the pass-point measurements, provide detailed amplitude and time properties of the video signal.

Both amplitude and time statistics were measured for each of the 11 reference levels. The 11 reference levels were chosen to cover the normal range of the video signal. The amplitude measurement extracted for each level indicated the percentage of time the video signal was approximately equal to the level. The group of 11 amplitude measurements thus derived provides a form of amplitude distribution for the video signal. This distribution can be used to indicate the basic form of video signal produced by the imagery, that is, sinusoidal, square wave, triangular wave, noise, and so forth.

The time-statistics extracted from the 11 reference levels indicate the predominant frequency, if any, present in the video signal. Both the first and second moments of the level-crossing statistics were measured. A high value of the first moment and low value of the second moment indicates a high degree of periodicity in the video signal.

DESIGN OF PLANIMETRY EXTRACTION SYSTEM

The data processing steps used to design the pattern recognition network for planimetry extraction were similar to those described above for pass-point selection. A total of 39 parameters were developed and evaluated. On the basis of this evaluation, a selection was made of 24 parameters, distributed equally between the stereo- and single-image types. To perform parameter evaluation, arbitrary numbers were assigned to the four classes of planimetry considered.

It was found desirable to investigate many forms of decision trees as it was not obvious which form would produce the best results. This was because the classes of planimetry had no natural ordering as is the case with figure-of-merit data. For example, a decision tree which first determines *wooded, cultivated, or urban, hydrography*? might perform markedly better than one which first determined *wooded, urban, or cultivated, hydrography*? It was also found desirable to apply types of error correction as illustrated in Figure 8. As shown in this decision tree, a final correction function *C or H*? is applied before the final output is determined. This function provides the ability to correct classification errors made in the first classification step.

EXPERIMENTAL RESULTS

Evaluation of various decision trees was accomplished by simulating the resulting designs. Each design was simulated using design and test sets of data with 272 points in each set.

The highest overall percentage accuracy achieved was 95.6 and 91.2 percent for the design and test sets, respectively. This level of performance was achieved using the decision tree shown in Figure 8. A more complete picture of the results can be shown by means of a confusion matrix. A confusion matrix, as shown in Figure 9, shows the numbers of points which fall into different categories. For example, row *C* in Figure 8(a) shows that 74 out of 78 cultivated samples were classified as cultivated.

In addition to the simulations of various classification networks using design and test sets, a number of simulations were run on data

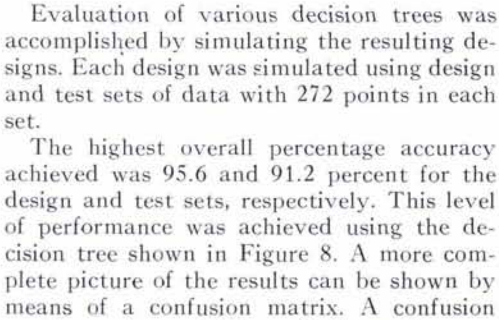

FIG. 8. Decision tree.

Classification Results →

Actual ↓	C	U	W	H
C	74	–	3	1
U	1	15	–	–
W	5	2	151	–
H	–	–	–	20

(a) Design Set

Classification Results →

Actual ↓	C	U	W	H
C	68	1	7	1
U	3	13	–	–
W	5	–	153	1
H	3	1	2	14

(b) Test Set

FIG. 9. Confusion matrix display of results of planimetry extraction.

collected in a grid-like array of points. This type of test was considered to provide a good indication of the potential feasibility of automatic planimetry extraction since a grid-like array of points would probably also be used in such a system.

Three small areas were processed in these tests; the areas are identified in Figure 10(a). The results achieved for the three small areas are shown in Figure 10(b), (c), (d). Data was collected at each point where a symbol is shown. Each symbol represents the classification for the particular data point as determined by computer simulation. The results can be seen to agree quite well with the actual terrain. A number of errors occurred in the small urban area; these points were generally called wooded or cultivated. Classification of these points as wooded is probably due to the high density of trees in the area. The hydrography appears to be quite well defined, and the boundaries between vegetated and cultivated areas in Area 2 are generally correct.

FUTURE POTENTIAL

The experimental results described in the preceding sections demonstrate that automatic pattern recognition can be used for image processing applications. In addition to pass-point and planimetry extraction, numerous other image processing problems can be handled by this technique. These include:

* *Specific Object Recognition.* Scan photographs to detect specific objects such as buildings and ships.
* *Identification of Specific Planimetric Features.* Scan photographs to detect specific planimetric features such as roads and shorelines. May be incorporated with automatic control equipment to track and plot rivers and roads.
* *Bandwidth Compression for Transmission of Image Data.* Process raw data from remote sensors such as satellite cameras and linear or planetary cameras, to extract significant features for transmission to Earth stations. For example, detection of storm patterns from weather satellites, detection of life-like objects from planetary explores, detection of geological features.

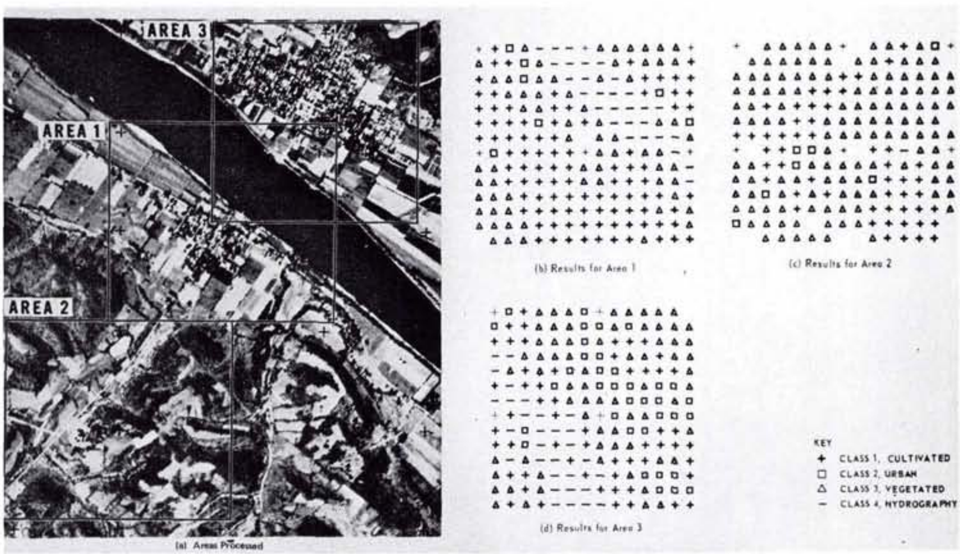

FIG. 10. Results from processing of grid-array points.

The use of automatic pattern recognition for photogrammetric applications would provide significant benefits over the manual methods presently being used. In particular, the potential advantages include:

- Relief of operators from tedious non-creative tasks.
- Provision of greater consistency of results, with objective classification standards.
- High-speed processing of large volumes of data with resultant decreased costs.
- Elimination of long training programs to develop skilled operators.
- Ability to provide output data directly to other data processing equipment for automatic mapping, compilation of statistical records, process control, etc.

To achieve these advantages, it will be necessary to develop the required pattern-recognition equipment to a level of capability equivalent to that of trained skilled operators. This is not a simple task for most practical applications. It requires the collection and analysis of large bodies of experimental data, and the development of complex data processing hardware with high reliability and reasonable cost. However, the payoff for such developments is considerable, and significant investments are justifiable. This is evidenced by the fact that many laboratories throughout the world are conducting major research and development programs related to automatic pattern recognition. These programs are investigating such areas as the development of improved training algorithms for decision network design, the development of efficient high-speed means to digitize image data for subsequent computer analysis, the use of optical image processing techniques to achieve low-cost high-speed implementation, and the development of techniques for using context data to improve recognition accuracy.

The continued development of ever more complex and less costly integrated circuitry will make it possible in the near future to produce highly sophisticated systems that are economically competitive with manual operations. Thus, it is highly probable that the automatic pattern recognition systems will ultimately become commonplace in photogrammetry and other image processing applications.