Automatic Target Recognition: State of the Art Survey

BIR BHANU, Member, IEEE
University of Utah

In this paper a review of the techniques used to solve the automatic target recognition (ATR) problem is given. Emphasis is placed on algorithmic and implementation approaches. ATR algorithms such as target detection, segmentation, feature computation, classification, etc. are evaluated and several new quantitative criteria are presented. Evaluation approaches are discussed and various problems encountered in the evaluation of algorithms are addressed. Strategies used in the data base design are outlined. New techniques such as the use of contextual cues, semantic and structural information, hierarchical reasoning in the classification and incorporation of multisensors in ATR systems are also presented.

Manuscript received January 31, 1985; revised November 20, 1985.

Author's address: Department of Computer Science, 3160 Merrill Engineering Bldg., University of Utah, Salt Lake City, UT 84112.

0018-9251/86/0700-0364 $1.00 © 1986 IEEE

I. INTRODUCTION

One of the key components of present and future defense weapon systems to be used on autonomous vehicle missions is the automatic target recognition (ATR) system. The ATR system effectively removes man from the process of target acquisition and recognition. This is desirable since the system with a man in the loop is generally slow, unreliable, vulnerable, and may limit the performance of the overall system or mission in real situations [1]. One important application of the ATR is in helping and guiding pilots of high-performance aircrafts flying close to the ground during bad weather or at night. Examples of systems incorporating ATR are the low altitude navigation and targeting infrared for night (LANTIRN) system, and cruise missile and remotely piloted vehicle (RPV) applications such as the Aquila RPV [2, 3].

Basically, the ATR system performs automatic target acquisition, identification, and tracking by processing a sequence of images. Fig. 1 shows an air-to-ground FLIR (forward-looking infrared) image with tank targets. In general, the target set could consist of tanks, trucks, armored personnel carriers, ships, etc. The algorithmic components of an ATR system can be decomposed into preprocessing, detection, segmentation, feature computation, selection and classification, prioritization,

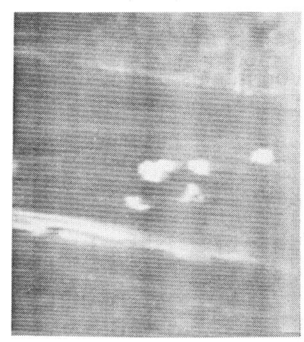


Fig. 1. An air-to-ground FLIR image.

tracking, and aimpoint selection (Fig. 2). The goal is to perform these functions in real time and to be able to adapt to dynamic tactical situations. The problem domain requires the tools of image processing (IP), image analysis (IA), pattern recognition (PR), and artificial intelligence (AI). Research work in this area has been going on for the last 25 years, but it has been only recently that sophisticated algorithms, microprocessors, and VLSI and VHSIC technology [3] have become available so that with the improvement in infrared, millimeter wave, and laser sensor technology it is now feasible to accomplish the ATR objectives. However, work in the area of evaluation of ATR algorithms is in its infancy [4].

ATR systems use a wide variety of algorithm suites. It is important to develop quantitative performance criteria for ATR systems for several reasons, 1) to compare various ATR systems and to predict their performance in a given scenario, 2) to study the behavior

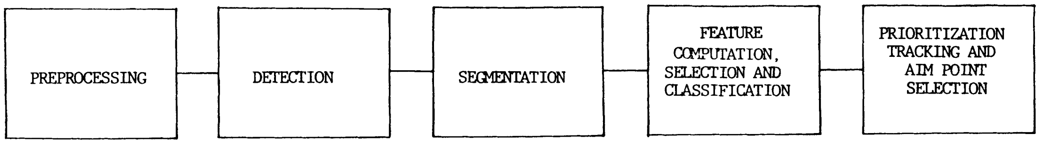


Fig. 2. Basic block diagram of typical ATR system.

of an ATR system and its components under different conditions and parameter settings, so as to be able to find the optimum performance achievable and the allowable tolerances of its components, 3) to understand the characteristics of the target and the background which affect the performance of algorithms, 4) to find common functional elements among the algorithms currently in use, 5) to help the algorithm developer to choose the appropriate algorithms for his application, and 6) to provide a fair and complete evaluation methodology for standardization purposes. So far, ATR algorithms having no provision for countermeasures have been tested on a very limited data set, and good classification performance has been reported. However, in practice these efforts have been only partially successful and have produced high false alarm rates. Some of the key reasons for this are the nonrepeatability of the target signature, competing clutter objects having the same shape as the actual targets, experience with a very limited data base, obscured targets, very little use of available information, related to and present in the image, such as context, structure, range, etc. If use were made of these diverse sources of information, then it is expected that the target signature characteristics would be extracted reliably and the effectiveness of current systems would be improved in both acquisition and classification.

As most of the work in the area of ATR systems has been done using FLIR sensors, in this presentation, FLIR imagery is assumed to be the input to an ATR system. FLIRs image the thermal radiations emitted by an object [5]. Generally the images exhibit high contrast and show few shadows. For a historical perspective of military applications of infrared techniques refer to [6]. Issues related to image processing and pattern recognition architectures are not discussed in detail here (for a survey refer to [7]). Sometimes in the literature the term "autocuer" is used for an ATR system.

There are a number of purposes to this paper. In Section II the techniques which have been used to solve the ATR problem are reviewed with emphasis on the commonality in the approaches. The recognition expert systems, which allow the use of contextual cues in an AI framework to improve the performance of ATR systems, are also discussed. Since the image data base is an invaluable resource for the development and testing of algorithms, issues related to data base are addressed in Section III, answering such theoretical questions as, "Is the data base representative, independent, and large enough?" As the current algorithm work is hampered by a lack of image characterization capability, the image information measures and characteristics are also discussed. In Section IV methods are proposed for evaluating ATR algorithms. Several new quantitative criteria are presented and major problems in the evaluation of algorithms and their possible solutions are discussed. As FLIR, laser detection and ranging (LADAR), and millimeter wave sensors complement each other in many respects (discussed in Section V), it is necessary to perform integrated information processing on multisensor data to achieve high system performance. The use of semantic, structural, and statistical information in a hierarchical classifier in a state of the art multisensor system is presented in Section V. Finally Section VI presents the conclusions and the future trends in this field.

II. ATR ALGORITHMS

Two approaches have been used to solve the ATR problem. First is the classical pattern recognition approach (Fig. 2), which uses statistical and structural techniques. Such techniques are based on the hypothesis that features of objects from different classes lie in easily separable regions of the multidimensional feature space, while features from the same class cluster together. Such an approach has limited knowledge and almost no intelligence and reasoning capability to learn from the dynamic environment and adapt to it. There could be substantial variation due to changing weather conditions, etc., even in a limited geographical area. It is desired that the ATR system be able to learn from the environment and be able to use contextual cues. The other AI-based approach provides these capabilities. It not only requires low-level processing, image analysis, and pattern recognition methods, but also requires high-level symbolic manipulation [8].

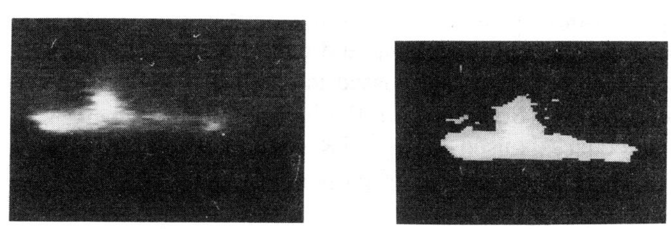
Pattern Recognition Approach for ATR

In this section various processing steps, which constitute the classical pattern recognition approach, as shown in Fig. 2, are described and then the knowledge-based approach is discussed.

Preprocessing: This step is designed to improve target contrast and reduce noise and clutter present in the image. It is usually accomplished by a local filter such as the median filter or the prototype automatic target screener (PATS) [9–11]. Other techniques shown to be effective use a high-pass filter for edge crispening and locally variable scaling for contrast stretching. The median filters have been commonly used as prefilters in edge detection because of their edge preservation property. They have also been used as postfilters after edge enhancement. As compared to a standard 2-dimensional median filter, Narendra [12] has presented a separable median filter which results in a simpler implementation in real-time hardware and whose performance is comparable to the 2-dimensional filter in image noise smoothing. Lo [13] has compared the performances of a variable threshold zonal filter, unsharp masking, PATS, histogram equalization, statistical difference operators, and a constant variance technique. He has found that the first three filters give the best enhancement results on FLIR images.

Target Detection: This is the process of localizing those areas in the image where a potential target is likely to be present. In some techniques such as "superslice" [14], localization and segmentation are inseparable. Most of the techniques can be adapted to detect either light or dark targets (targets which are hotter than the background generally appear as bright contrasting objects in FLIR images). Burton and Benning [15] and Schachter [16] present an evaluation of some target detection methods. Burton and Benning [15] and Politopoulos [17] use a double window filter. This filter is based on the contrast between the target and its immediate background. It is conventionally used in the detection of targets in radar data. It consists of two rectangular windows, in which the inner window surrounds the target, and the outer window contains background. Range is used to control the window sizes [15, 18]. (Note that although range is very useful in detection, segmentation, and classification of targets, it is not available directly from a passive FLIR sensor. The use of active sensors has the risk of possible detection and susceptibility to countermeasures. Techniques for generating range information using scene dynamics in a sequence of passive sensor images have been proposed [19].) The metrics used to determine the likelihood of a target being localized at a pixel are different in [15] and [17], which make no use of a statistical model in their experimental design. Schachter [16] uses a simple statistical model and estimates the probability density function of the random variables designating target and background windows. Minor and Sklansky [20] use a spoke filter (an eight-spoke digital mask) which is an extension of the Hough circle detector. It examines the local edge magnitude and direction. It needs preprocessing which includes intensity normalization, a dc notch filter, and an edge detector. Schachter [16] and Frey et al. [21] have also used the basic spoke filter idea. Bhanu et al. [18] use intensity, edge, and range information. Mitchell and Lutton [22] use intensity and texture measures. Rubin and Tseng [23] use a linear discriminant function of local features of the image to obtain interest points. The target is assumed to be in the neighborhood of these points. Tisdale [24] uses gradient operators. Soland and Narendra [11] extract the "object intervals" along each scan line in the edge image. These lines are concatenated to obtain the outline of each object. Object boundaries are finally smoothed. In the mode seeker technique [25] a pixel is iteratively replaced by the average of a selected set of its neighbors. The neighbors are chosen such that they belong to the same histogram peak as the given pixel. Upon termination a two-spiked histogram is obtained, one for the target and the other for the background. The global extension of this method called "superspike" gives good results. The border-following algorithms and pyramid approach did not perform well as target detectors [16].

Segmentation: Once a potential target is localized, it is extracted from the background as accurately as possible. The superslice algorithm [14] assumes that the objects are distinguished from their surroundings by the presence of edges and that different thresholds may be required to extract different objects in the same scene. First it computes a thinned edge image. It then uses several thresholds and finds the borders of the connected components in the image. Segmentation is achieved by selecting the threshold such that the coincidence between the thinned edge and the border of the connected components is maximized. Minor and Sklansky [20] use edge direction in addition to edge magnitude like superslice. These techniques have problems when there is a significant intensity distribution across the target. Brown and Frei [26] use a sequential region growing technique which uses size and rate of growth to reject nontargets and terminate. It does not use intensity thresholding or edge information. When it is preceded by smoothing it may be suitable for the segmentation of nonhomogeneous objects of known size in noisy and blurred images. Rosenfeld et al. [27, 28] and Bhanu et al. [18, 29, 30] use two-class (target and background) and three-class (target, background, and clutter) relaxation methods. These are iterative parallel schemes which make use of contextual information to remove any inconsistency and ambiguity in the labeling of pixels. For an example of segmentation of a ship target using a relaxation method see Fig. 3. Note that there is a significant intensity variation across the target. Part of the target image is whiter and part darker (the target is partly hot and partly cold). The pyramid spot detection approach [31] for extracting compact objects from a contrasting background is based on detecting spots in a succession of lower resolution images. Spots are detected by comparing each pixel with its 8 nearest neighbors. To obtain segmentation, thresholds can be calculated in the low resolution image or the original image. A variation of the above technique is the pyramid linking approach [32]



(a) (b)

Fig. 3. Segmentation of ship target using relaxation. (a) FLIR image of ship target. (b) Segmentation of (a).

which involves the creation of links between pixels in the successive levels of a pyramid. Hartley et al. [33] compare superslice and pyramid approaches, relaxation, and mode seekers on a data base of 51, 128×128 FLIR images. They observe that the superspike algorithm outperformed all the others. Chen and Yen [34] use Fisher's linear discriminant to segment by pixel classification. Their theoretical error probability computation for the percent of pixels misclassified compares favorably with the experimental results. Yu and Fu [35] consider a spatial stochastic model for recursive contextual classification which results in the elimination of misclassified samples in homogeneous regions and smoothing of noisy boundaries. Several researchers have used local statistics and edge information for the segmentation of the targets from the background [22–24].

Feature Computation, Selection, and Classification: After segmentation, a set of features is computed for each object. The reliability of these features is essential for target classification. Most of the features used by researchers are geometric, topological, and/or spectral. Hu's moments are most extensively used for the classification of aircraft, ships, ground targets, bridges, buildings, etc. [11, 18, 23, 36]. Cheatham et al. [37] use geometric moments computed optically in a hybrid system for ship classification. Various shape, gray scale, and projection features are commonly used. Semantic features such as geographical, temporal context, and environmental have been used only to a very limited extent. In addition to the specific restrictions imposed by the classification techniques (e.g., linear and quadratic classifiers), the desirable properties of the features are invariance with respect to geometry (rotation, scale, and translation), computational efficiency, and extractability under adverse conditions. The robustness of feature classification with respect to different segmentation results has not been thoroughly investigated.

The primary goal of feature selection is to obtain features which maximize the similarity of objects in the same class while maximizing the dissimilarity of objects in different classes. It also results in computational efficiency and reduces memory requirements of the classifier. Nonextensive sequential feature selection techniques are not optimal. It is necessary to do feature selection on uncorrelated features, since statistical dependence or even conditional independence among the features may cause the best subset to exclude the best feature. Feature selection in the context of the ATR problem has been performed informally by histogram examination [14, 36], Bhattacharyya measure [18], Kolmogorov-Smirnov test [36], F-statistic, exhaustive scheme [11], physical reasoning and linear regression techniques, etc. [38]. In feature selection the assumption of Gaussian distribution of the data is generally made even though it may not have a multivariate Gaussian distribution. Features and classifiers are optimized with respect to aspect. Classification has been mostly done by a K-nearest neighbor (K-NN) algorithm [11, 18, 23], using projections [22], linear and quadratic classifiers [14], structural classifiers [39], tree-based classifiers [11, 14], or using clustering techniques. In a tree classifier design it is required to partition the samples at each node into two classes and select the subset of features which is most effective in separating the two classes. As an example, a tree can be generated by maximizing the amount of average mutual information gain (AMIG) at each partitioning step [40]. Note that optimizing the decision at each node in a tree classifier may not necessarily yield the best results. It is important to understand the hierarchical ordering of features and class separability at each node together with the global cost minimization criteria. Statistical-structural classifiers have been constructed to obtain a classifier which can work at all practical ranges [41, 42] and which allows the incorporation of scenario specific knowledge and decision smoothing. Linear and quadratic classifiers have been combined synergistically with a K-NN classifier to reduce the amount of computation and increase recognition performance [41, 43]. Generally, in these classification studies there has been no statistical analysis of performance and the data base is limited. Interframe analysis has been used to improve classification performance [10, 44].

Prioritization, Tracking, and Aimpoint Selection: Prioritization is the process of assigning priorities to the targets in the field of view (FOV). This information, which is prestored, is normally based on the type of the target, and the probability of its correct classification. Once the targets are prioritized, they are handed off to a tracker. Some of the shortcomings of the earlier target trackers were inability to track more than one target at a time, frequent target breaklocks encountered during high clutter and low signal/noise ratio condition, and the great difficulty in reacquiring a target once breaklock occurred. As a result, the concept of the "intelligent" tracker emerged. It combined target cueing and tracking methodologies for near-zero breaklock performance. Reischer [45] presents a summary of target tracking methodologies. Noting that any detected scene change from frame to frame is potentially a target, both signal- and symbolic-based approaches have been used [46, 47]. Iyala et al. [46] use a symbolic approach for image registration and motion analysis whereas Holben [47] uses correlation tracking over a subregion in a FLIR image. Trackers based on correlation, feature, intensity, and contrast complement each other by allowing switching from one approach to the other when the confidence level of a given tracking approach is low. This results in tracking with a high confidence level and minimization of loss-of-lock. Dorrough et al. [48] consider a multimode tracker consisting of an intensity centroid tracker, an edge tracker and a correlation tracker (based on a sequential similarity detection algorithm). The tracker has two modes: tracking and coasting. In the tracking mode target motion analysis is done and checks are made to determine if the target continues to be tracked. If it is not, the tracker enters the coasting mode, and the target characteristics before the loss-of-lock are used to reacquire the target. Note that the method will work only if the velocity of the target does not change while it remains obscured. Some tracking algorithms have been implemented in hardware and others are programmable. Gilbert et al. [49, 50] present a real-time video tracking system, called RTV (real-time videotheodolite). They use adaptive statistical clustering techniques to classify pixels (separating the target from the background) and the projections of the target as features [22] to identify and track objects.

Aimpoint selection involves the determination of the critical aimpoint of a target. It may be an interior point in the silhouette of the target. A stored feature vector corresponding to the target class and aspect (obtained from the previous classification step) is used for aimpoint designation. In scenarios involving a missile approaching the target, the maintenance of the aimpoint is important so that the missile is locked on to the selected aimpoint. It is carried out by using correlation, feature matching, etc.

Knowledge-Based Approach for ATR

The limitations of the ATR systems, as mentioned before, have led to the realization that better performance can be achieved by suboptimal ways of handling context, rather than by optimal ways of handling the local structure as conventionally accomplished in the PR approach. The ATR problem is suited for building a knowledge-based (K-B) system for a specific operational environment and geographical area. Recently attempts have been made to use context (temporal, global, local, and ancillary information, such as map data, sensor data, seasonal, and intelligence information), semantics, and problem domain knowledge in an AI framework to improve the performance of ATR systems [41, 51–55]. Most of this work has started only in the last few years and the results are yet to be seen. Tseng et al. [51] implemented a simple K-B system in PASCAL as part of the intelligent bandwidth compression (IBC) program. Kim et al. [52] and Spiessbach and Gilmore [53] incorporate contextual cues using AI techniques. Drazovich et al. [54] describe a K-B approach for radar target classification capable of distinguishing sea targets. Using "if-then" rules, the system would match radar features and images with known sizes, shapes, and dimensions of vessels stored in the data base.

The K-B approach basically has three parts. The first part consists of the development of low-level image analysis techniques and PR methods, as in the classical pattern recognition approach. The second part includes the AI [8] techniques for symbolic representation, strategies to be used for the integration of knowledge, search and control methods, and the implementation of a knowledge base so that appropriate knowledge is available at the right place in the search and decision making process. The third part combines the first two parts so that the system as a whole can be implemented. Fig. 4(a) shows the basic structure of a K-B system. It consists of a knowledge base, a global data base, and a control structure. The global data base contains all the current information about the objects, their properties and classifications, and the system's view of the situation or interpretation of the features. The knowledge base consists of the knowledge about the objects to be recognized, the knowledge needed to recognize the objects and perform low-level image analysis and processing and the knowledge to control the interaction of various expert modules. Here semantic nets, production systems, and frame representations are used. The control structure matches the current world model in the global data base against the conditions in the set of rules in the knowledge base. When a match is found, the control structure evaluates the corresponding action specification by activating an expert and putting the results of processing into the global data base. This usually results in placing a new symbol, erasing an old symbol, or executing a process. Control strategies such as bottom up, top down, and a mixture of these two are used in a hierarchical setup and hypotheses are generated and tests are made about the target type.

As an example of a K-B system, Fig. 4(b) shows a block diagram of a target recognition expert system [41] which uses a blackboard model [7]. The blackboard model was first used by Matsuyama [56] in image understanding for the structural analysis of aerial photographs. In this model, all the information about the properties of objects is stored in the blackboard. The experts shown in Fig. 4(b) can be classified as feature experts, control and decision making experts, spatial and temporal experts. The system is initiated by start/stop expert. Each expert checks with the blackboard for the possible condition of its activation and it puts the result of processing on the blackboard. The individual experts derive their control from the central control of the blackboard. The experts can be sufficiently intelligent and autonomous by themselves for the activation of other experts. For example, a segmentation expert may call for a region, edge, or texture expert, say on the basis of the information received from the horizon detector. The blackboard has 1) rule-based computational mechanism, 2) image formation, mission oriented, and other relevant information, and 3) information about the objects and their relations. Model-based recognition in a hypothesis and test paradigm is used for finding object type and orientation.

In short the representation, modeling, and control strategy are the key factors which determine the capability of a K-B system. It is now realized that the extraction of high-level symbolic information from image data is more critical than the use of this information. Note that the realtime considerations influence the design of the whole system. It would be interesting to evaluate the performance of the K-B system at the boundaries of

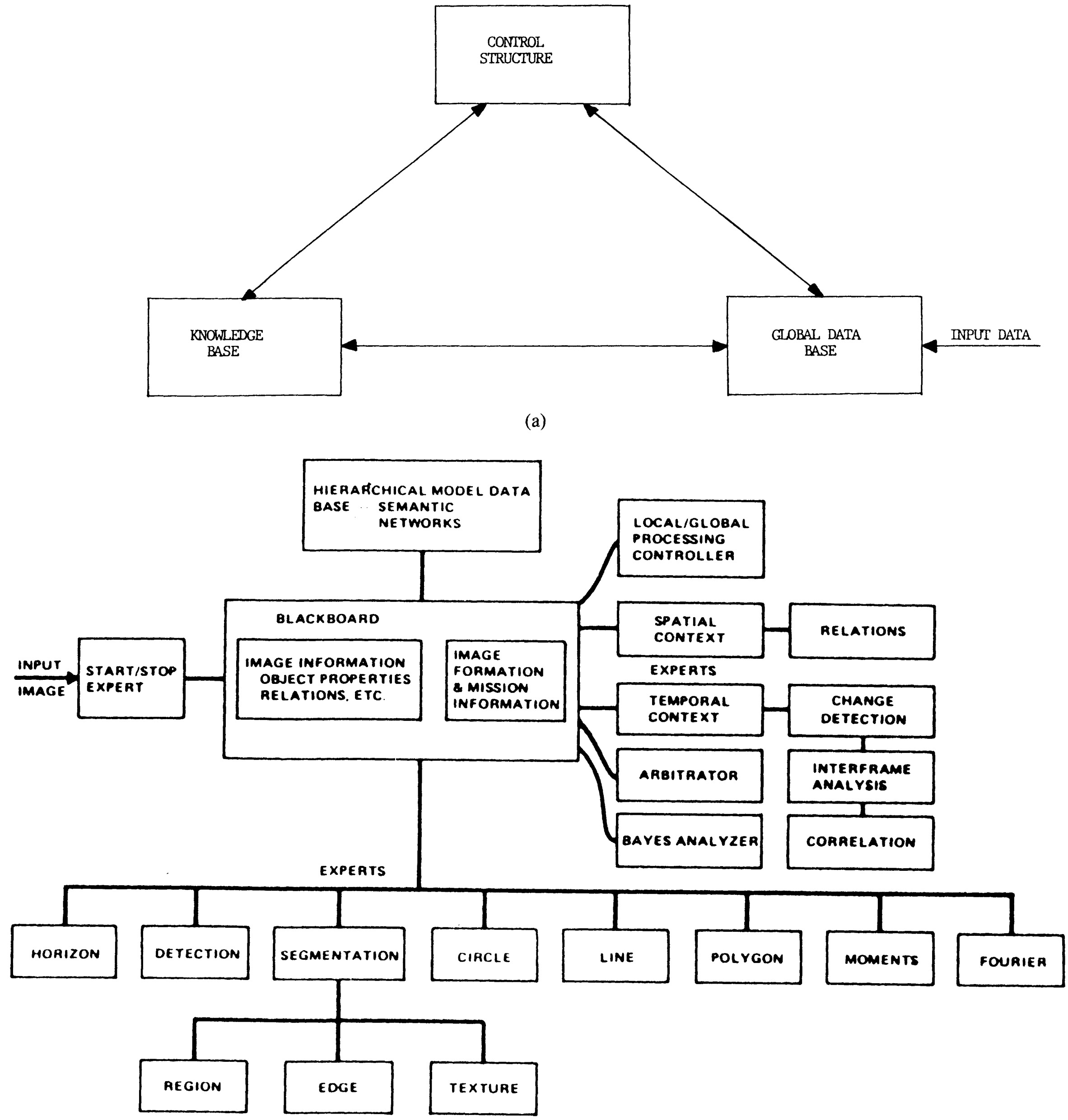


Fig. 4. K-B approach for ATR. (a) Basic structure of K-B approach. (b) Block diagram of target recognition expert system.

the knowledge it contains and to know when the specific image encountered is beyond its capacity.

III. DATA BASE CONSTRUCTION AND EVALUATION

One of the key requirements in the development of ATR algorithms is the availability of the image data base on which algorithms are to be evaluated. Since the reliability and accuracy of the results of different techniques used to evaluate the ATR algorithms depend upon the data base, its design is important for the systematic evaluation of algorithms. The assumptions that an image or data at the input of an ATR system should satisfy are often not stated explicitly and there has been a general absence of models. Preliminary work has been done to mathematically model the ATR system. However, such models are not realistic because of the difficulty of modeling the background and because the tasks of preprocessing, segmentation, feature selection, and classification are highly interrelated, application-dependent, and nonlinear. Assuming an Earth-viewing sensor, there is an enormous variety of backgrounds even in a restricted operational scenario in a given geographical area. At present there is a lack of adequate theoretical models for background characterization. In the absence of such models, a heuristic approach is used to design and verify ATR algorithms. Therefore, such algorithms must be validated experimentally, which requires a large data base. The interaction of algorithms and data base is shown in Fig. 5. Note that the models for algorithm evaluation drive the requirements for data base, its collection, and organization.

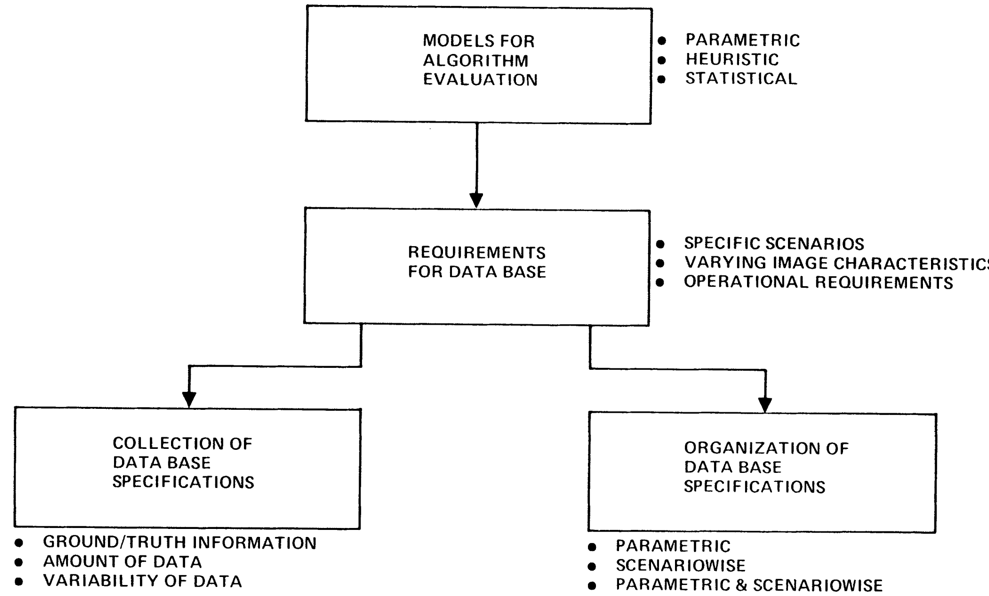


Fig. 5. Interaction of algorithms and data base.

In a PR approach, it is generally assumed that most of the information about the data environment of the system is contained in the labeled samples which are apportioned to a training and a test set in such a way that they both represent similar scenarios. The "leaving-one-out" approach can also be used. (For other approaches see [57].) As a result, the performance of the classifier depends upon the feature set used on the training and test sets. The classification performance is measured by training the classification algorithm on the training set, testing it on the test set, and noting the results. Therefore, the training and test samples should be fairly representative of the real-world scenario in which an ATR system is to be operated. Since it is impossible to obtain images exhibiting all scenarios and conditions which may exist in reality, the training data is generated using real images which exhibit the typical characteristics of FLIR scenes. Synthetic data are initially useful in the design and evaluation of the individual components of the system. The classifier performance is ultimately evaluated on the real data, not on the synthetically generated data. The number of samples in different classes in the training set is chosen in proportion to their most likely outcome in practice.

The *requirements* of the data base are tied to the specific scenario in which the cuer is to be operated. It includes the images exhibiting varying sensor characteristics (FOV, signal/noise ratio, range, etc.) and operational requirements such as terrain; climatic, atmospheric, and day and night conditions; false, cold and partially occluded targets; decoys; burning objects; targets in groups; targets on roads; different aspects of targets; various types of clutter, etc. *Organization* of the data is done parametrically, scenariowise, or as a combination of both. Parametric specification can be made as to the type, size, aspect, and velocity of the target; range; depression angle; altitude of the sensor; day/night conditions, etc. With respect to scenarios it may be interesting to include examples of closing in or popping up type of sequence of images against various targets in numerous background, clutter, and environmental conditions. A combination of these allows the selection of any desired set of images. The *size* of the data base is a function of specific cuer scenario requirements. It should have sufficient samples for each aspect, class, and scenario (10 to 200) to train a classifier so as to obtain statistically meaningful results. Theoretical bounds as a guide to the amount of data for a given performance of the ATR system in a particular scenario can be obtained by using inequalities from probability theory and testing to verify that the samples are uncorrelated [41].

In order to be able to quantify the data base according to the type of images that it contains, measures of the image information content are needed. In the following, several image information measures are first presented and then discussed. The specific FLIR image characteristics which form the basis for developing evaluation techniques help to define the interrelation between input image parameters and ATR performance, and identify representative sets of real test images for algorithm evaluation purposes.

Information Measures: A FLIR image can consist of tactical targets, background, noise, and clutter. Target refers to an object which an ATR system must detect and classify. Clutter is defined as an "object" which resembles a target, but is not a target. Noise is characterized as a random intensity distribution over the image caused by the sensor and the image formation process. Following are some quantitative information measures.

Edge points measure: The edge points are characterized by high contrast with their neighbors. The potential target-like objects are usually present in the neighborhood of these points. Information content of an image can thus be measured by finding the points in an image at which the magnitude of an edge operator is greater than a certain threshold. Several measures are defined in terms of these edge points such as 1) number of edges per unit area, 2) information theoretic measure I present in the image, $I = -\log_2 P$ bits, where P is the probability of possible pictures made up of edge points. For example, if in an image of 500×600 pixels with 8 bits/pixel, there are 10 percent edge pixels, then the information content of the image is

$$I = \mathrm{Log}_2\, 256^{(3\times 10^5)\times \frac{1}{10}} = 2.4 \times 10^5 \text{ bits}$$

and 3) the entropy H, the degree of uncertainty present in the image. Compute the probability p at an edge point pixel of being white $p_i(1)$ or black $(p_i(2) = 1-p_i)(1))$, then obtain the entropy given as:

$$H = -\sum_i \sum_{l=1}^{2} p_i(l) \log p_i(l),\ i \in \text{edge point}.$$

Uniformity and structural measures: These criteria measure the consistency of a pixel with respect to its neighbors and reveal the homogeneity of the regions. Also, they provide structural information about the image. One such measure is,

$$U = \sum_x \sum_y [f(x,y) - \bar{f}(x,y)]^2$$

where $f(x,y)$ is the gray level at a pixel in the image and $\bar{f}(x,y)$ is the average gray level in a 3×3 window centered about the pixel (x,y).

Other criteria which measure structural information are structural entropy [58], and the linear and cultural features present in the image. Measures such as cooccurrence matrices are useful as general scene information characterizing criteria. The number of target-like objects in an image which are not targets characterize the amount of clutter present in the image.

Characteristics of FLIR Images: In order to achieve optimum performance from an ATR system it is essential that its design make maximum possible use of the specific characteristics of FLIR images. For example, segmentation techniques should be based on models which take into account the real aspects of FLIR scenes so that the resulting information derived from these models can be useful in predicting realistic performance measures for the technique. It would be useful to distinguish the characteristics originating in the image formation process (sensor noise and signal transfer characteristics) and in the actual physical scene. Following are the characteristics and modeling approaches to FLIR images which are important to the general ATR problem.

1) The simplest model of a FLIR image is a two-class, black and white intensity definition. The target is brighter or darker than the immediate background.

2) In many situations the simple model of 1) may not be sufficient to extract the target if part of the target is hotter and part is colder than the background. This is a common occurrence in FLIR images. In such cases, the image model should make use of the context and some shape information about the targets (K-B system).

3) When multiple targets are present in the image they may occlude or be in close proximity of each other. Thus, separation of individual targets may be difficult. Scene modeling should include occlusion considerations [42].

4) Size of the target is an important parameter in the ATR system design. As the range increases, the target occupies a reduced number of pixels in the image. If the number of pixels on the target becomes very small, the target may dissolve into the background. Thus, range information may be of crucial importance to the scene model.

5) Targets may be obscured by or be partially hidden in smoke, dust, and shadows. Scene modeling should accordingly account for these effects.

6) Often it happens that target boundaries are poorly defined and buried in the background. In such cases, the use of textural, structural, and contextual scene information may be useful.

7) Frame to frame analysis is important for scene modeling. Such an analysis may reduce the amount of computation. It increases classification accuracy by requiring repeated consistency of the classification decision.

The information measures and characteristics listed above can be used to 1) develop practical and realistic algorithms and techniques to evaluate existing ATR systems, 2) analyze various components of an ATR system, 3) create a synthetic data base in the laboratory which incorporates the realism of FLIR scenes, 4) Obtain a subset of images from the data base which can be used as a representative set in testing ATR systems and finally, 5) Specify requirements for data bases of FLIR images which are increasingly representative of the real world.

IV. EVALUATION OF ATR ALGORITHMS

An ATR system should be evaluated on the basis of the task it is able to perform in a given environment considering such factors as sensor type, resolution of data, type of objects, and complexity and information content of the scene. The design of the system should be such that each of its components makes maximum use of the input data characteristics and its goals are in conformity with the end result of reducing classification errors and false alarms. To obtain optimum performance, it is essential to obtain the maximum attainable performance from each of the components of the ATR system. Thus it is necessary that each component have its own quantitative figures of merit against which it can be individually evaluated. However, since the ultimate goal of the ATR system is correct classification and the intermediate steps, such as preprocessing, segmentation, feature selection, classification, etc. are subservient to that goal and not an end in themselves, it is logical that each of the components must not only be evaluated with respect to its own figures of merit but also against its effect on the overall classification.

Thus classification performance of the system is evaluated using a black box approach, together with the evaluation of each of the components and its effect on the classification results. For example detection and segmentation techniques are evaluated on the basis of how well they are able to locate and preserve the shape of the target. The assumptions inherently involved in the development of these algorithms are checked against the input data and their effects on the output results. A segmentation technique based only on the concept of contrast will not be able to produce the target boundaries faithfully. This in turn provides valuable information to the algorithm developer guiding him in a reevaluation of his assumptions so that they are compatible with the realistic images and so that consistently good results can be obtained. The approach allows one to evaluate an ATR system efficiently, to identify better system components among various ATR systems, to understand them in depth, and to design better systems in the future.

The three parameters which are generally used to characterize the *overall performance* of an ATR system are the probability of target detection, the probability of correct classification, and the number of false alarms per frame. False dismissal is accounted for in the probability of target detection. By classification is meant recognition, which is the determination of the target type, e.g., tank versus truck. At present there is no standardization of these definitions. Normally they are defined as follows. The probability of target detection (classification) is defined as the ratio of the total number of targets correctly detected (classified) in the testing set to the total number of targets in the testing set. False alarms per frame is defined as the total number of false alarms divided by the total number of images in the testing set. Note that there is an inherent conflict between the probability of target detection and false alarms per frame.

Performance Criteria for ATR Components: In the following, quantitative figures of merit are presented for ATR system components. They are based on the system concept. The effect of the system on the input is measured by its output.

The *preprocessing* operations condition the incoming data stream to reduce the sensor and/or environment-dependent perturbations. They improve image quality so that the effectiveness of subsequent processing steps is enhanced. Examples of preprocessing functions are noise suppression, dc restoration, focus control, adaptive contrast enhancement, gain and bias adjustments, etc. The figure of merit depends upon the operation. As an example a median filter is supposed to maintain the sharpness of the edges, i.e., it does not blur edges as a linear low-pass filter would. To evaluate this filter several criteria can be used. 1) An edge operator is applied to the input image and to the output median filtered image. Now both images are thresholded at the same gray value. The ratio of the number of edges of the thresholded input images to those which are present in the thresholded output images gives an indication of the effectiveness of the median filter in preserving edges, but not its effectiveness in removing noise. 2) The median filter is commonly compared with an average filter by using the ratio of variance of the filtered images. (Refer to [9] for a thorough discussion.) The edge evaluation schemes such as [59] which combine two desirable properties of well-formed edges, viz. good continuity and thinness, can be used to measure the effectiveness of various types of preprocessing operations. A weighted combination of two components viz., discrepancy and the busyness, is used to evaluate the smoothing of an image [60]. Here discrepancy measures the difference between the original and the smoothed images by finding the sum of squared difference between gray values of the corresponding pixels. The busyness measure is computed on the images by finding the sum of absolute values of an operator, such as the gradient or Laplacian. As the preprocessing step increases the separability between the target and the background, Bhattacharyya distance is used to measure the separability [61].

The *detection* operation localizes those areas in the image where a potential target is likely to be present. If a target is missed in this process, it will be missed altogether. Several performance measures are 1) probability of target detection—a target is said to have been detected if its centroid lies within a small window (its dimensions are function of the true target size) centered at the centroid of the true target [18, 33]; 2) the ratio of the localized area to the image area—it measures the computational efficiency of the algorithm [18]; and 3) the number of localized areas to the actual number of targets present in the image—it affects the false alarm rate [18].

The *segmentation* operation extracts the target from the background after it has been detected. Segmentation techniques are evaluated on the basis of how well the implicit or explicit model in the technique is able to predict performance. Some quantitative measures are: 1) number of target pixels misclassified with respect to the true target [62], 2) correlation coefficient between the true and extracted target, 3) mean square error between the true and extracted target, 4) object-to-background contrast, intensity difference and Bhattacharyya distance between the true target and clutter objects are used with thresholding, where the threshold is fixed a priori, or determined in a global, local, or object adaptive manner [61], and 5) shape number to estimate the shape difference between the true and extracted targets [63]. Weszka and Rosenfeld [64] propose two methods for measuring the "goodness" of a thresholded image. They use a busyness criterion and a discrepancy criterion. The busyness measure is based on the assumption that for simple compact shapes, which are not strongly textured, it is desired that the thresholded images look smooth rather than busy. For a given threshold it is measured by summing those entries of the cooccurrence matrix which represent the percentage of object-background adjacencies. The discrepancy criterion in principle involves measuring the discrepancy of a thresholded image in terms of classification error, i.e., misclassifying an object point as a background point or vice versa. It is obtained by fitting a pair of Gaussian curves (object and background distributions are assumed normal) to the histogram of an image. Using the parameters of the fitted curve, a threshold is selected that minimizes the probability of misclassification. These two measures are found to be useful in facilitating threshold selection in FLIR images.

Evaluation of *features* and their clustering is important in the design of ATR algorithms. Computational efficiency and accuracy of the features and their power to distinguish different targets are also important, since the classification results depend upon the accuracy and reliability of the features. For example, Hu's moments which are invariant with respect to size, position, and orientation have been commonly used. However, they are not contrast invariant and the contrast change of an image introduces a nonlinear scaling effect. Resolution is affected by scale, so that the invariant moments are no longer strictly invariant under rotation and scale changes. To examine segmentation effects, the features of the segmented target and the true target are compared by measuring the distance between them in feature space and evaluating different features by carrying out a hypothesis test and a variance analysis [41]. To evaluate the clustering of features a clustering fidelity criterion such as $\beta = \mathrm{Tr}(S_B) \cdot \mathrm{Tr}(S_W)$, where S_B and S_W are between-cluster and within-cluster scattering matrices, is used. The behavior of β is such that it passes through a maximum at the intrinsic number of clusters, and at the maximum the ratio of $\mathrm{Tr}(S_B)/\mathrm{Tr}(S_W)$ is exactly 1. The maximum of β can be determined by incrementing the number of clusters until a decrease is detected. This allows one to obtain the number of clusters inherent in the data. Departure of the number of clusters thus obtained from the known number of clusters tells about the quality of the features.

It is also important to determine if the clustering is really present or it is a statistical artifact because a clustering method always finds clusters, whether or not they are real. Furthermore, it may impose a particular structure on the data rather than find the actual structure present. For example, for the patterns lying along two long parallel lines, a mean-squared algorithm will probably cut the lines rather than group the patterns on each line. Different techniques are likely to give different solutions unless the data are very clearly structured. Methods based on using several clustering techniques and partitioning the data randomly and testing for the randomness of the data determine the structure of stable clusters, if any.

The performance of a *classifier* is measured by finding the probability of classification and the false alarm rate. The probability of classification is given in the form of a confusion matrix and the confidence associated with each classification together with the size of the feature vector and data base. It is usually measured by the well known techniques of partitioning the data set into a training set and a testing set and/or using the leaving-one-out approach. Reliability of a classifier is a function of the clustering quality of the features. It is measured by finding the stability of the feature clustering. Stability is obtained by finding the average of the sum of the squares of the distances between the reference cluster centers and the cluster centers actually obtained in the leaving-one-out technique. The efficiency of the K-NN algorithm is also an important consideration in the evaluation of the classifier. Several approaches including ordered search procedures, branch, and bound, and $k-d$ tree have been used for finding k-NNs [65, 66].

In practice an *interframe analysis* is carried out to improve the performance of an ATR system. Here a tracker is normally used. The basic performance measures of a tracker are centroid drift, jitter, peak track rate, probability of loss-of-lock, and reacquisition. Other measures are the ability of the tracker to perform adequately in the presence of platform motion and target obscuration.

V. MULTISENSOR TARGET RECOGNITION SYSTEM

In this section we present an example of the state of the art ATR system [48]. The schematic diagram of this multisensor ATR system is shown in Fig. 6. It uses semantic, structural, and statistical information in a hierarchical manner. FLIR (8–12 micron), LADAR (10.6 micron), and MM wave (3.2 mm) sensors operate in a synergistic manner to obtain the best performance, even in the presence of adverse conditions. Co_2 laser radar (LADAR) images are based on reflectivity differences. They are similar to FLIR images. They suffer from speckle, but are less sensitive to target operating conditions and environmental changes. MM wave sensors provide better penetration of battlefield obscurants (smoke, dust) and weather, but the image quality is relatively poor. The multisensor approach has some counter-countermeasure advantages [67–69]. Many of the current algorithms designed for FLIR images for preprocessing, detection, segmentation, and classification could be applied to LADAR and MM wave images [48, 54].

The sensed images are preprocessed (Fig. 6) to enhance the quality of the images or restrict the areas of the images which are to be searched. Then the targets are localized and segmented from the background. Their features are computed and preliminary information regarding their classification is obtained. At the same time, interframe analysis and tracking are carried out which by directing the focus of attention not only help to reduce the amount of computation involved in the detection and segmentation of the targets in subsequent

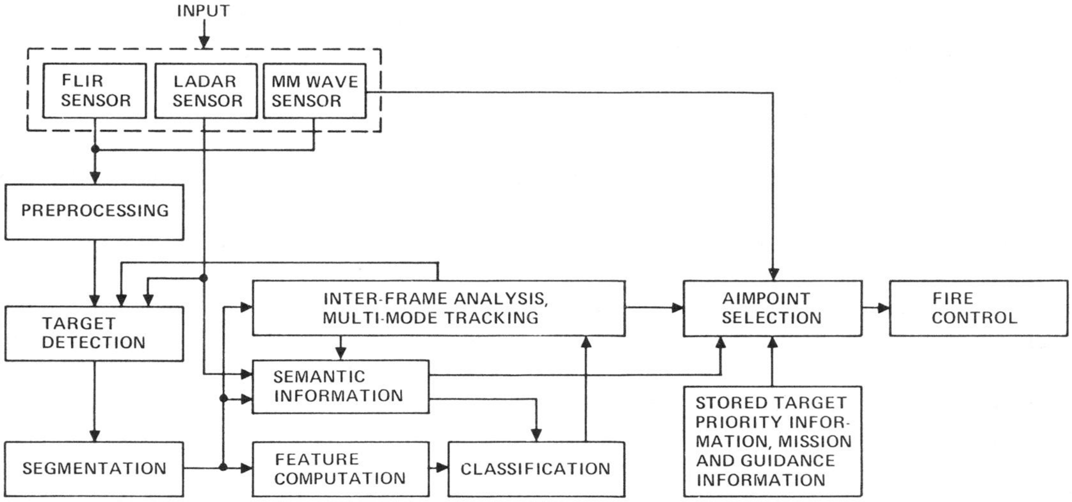
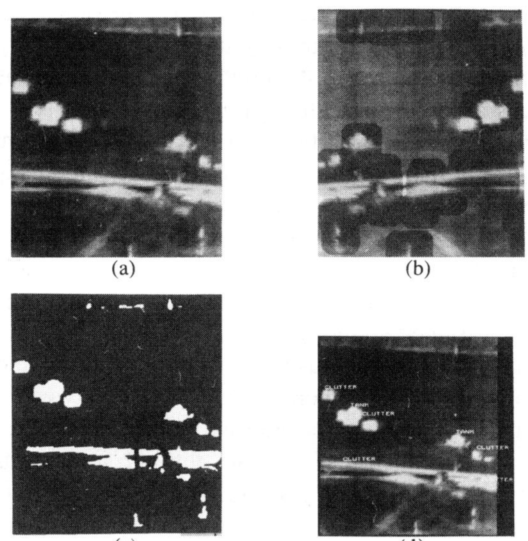


Fig. 6. Schematic diagram of multisensor ATR system.

frames, but also to improve classification results. The slant range (distance of the target from the center of the FOV) is used during target detection and in improving the classification results. This in turn aids the aimpoint selection process. When the FLIR and LADAR sensors are not operational, a MM wave sensor is used to determine the relevant information regarding the target. This information is then used for aimpoint selection. Using the target's intensity, texture, and slant range (note that with range not only the size of the object varies but also the contrast of the object/background varies, which is a function of the weather conditions) an intelligent preprocessing step is carried out. This step reduces the image area to be processed to segment the target by over 80 percent, and thus greatly reduces the amount of computation [18]. This percentage can be raised by use of the data from the other sensors. An efficient relaxation scheme is used to extract the target from the background, while preserving the gross boundary of the target. The technique allows control over segmentation through the use of three parameters. It has been thoroughly evaluated with respect to the size of the target, contrast, and signal/ratio [62]. As an example, Fig. 7 shows the segmentation of a tank target.

A number of shape, geometric, moment, and gray scale features [18, 41] are computed after the signature of the target has been extracted from the background. Features which are important in discriminating between different types of target are automatically selected. Feature selection is carried out by using the Bhattacharyya distance, K-means algorithm, and the discrete Karhunen-Loeve transform. The technique based on selected feature set have been found to be more robust and comprehensive than those based only on features such as moments. These features can be correlated with those obtained from other sensors. The classification scheme consists of a number of linear or quadratic classifiers with ancillary clustering techniques. These classifiers are designed to work on a reduced set of features after a feature selection process. They also allow



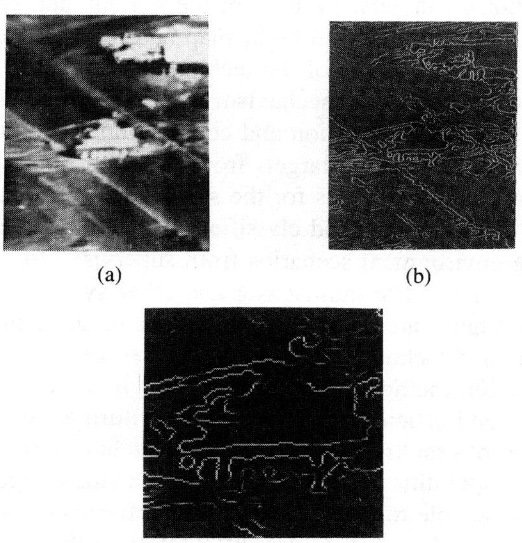
(a) (b) (c) (d)

Fig. 7. Preprocessing, segmentation, and classification of tank target. (a) Original FLIR image. (b) Preprocessed image. (The shaded portion of the image is the area where a potential target is likely to be present. (c) Segmented image using a relaxation method. (d) Classification results.

the design of a tree classifier such as [70], where at each nonterminal node a K-means algorithm and a Bhattacharyya measure are used. A classifier giving the smallest error rate is selected. The classifier controls the detection and segmentation of the target image. The K-NN algorithm is implemented using an efficient tree approach. Classification results on the image of Fig. 7(a) are shown in Fig. 7(d) [18, 41, 48].

The final assignment of a target to one of several classes is done using the classification results from two sources of information, 1) classification of a target based on several frames of the input data by an interframe analysis, or 2) classification of a target using the range image from the LADAR sensor. The target is finally classified by using simple logic with inputs 1) and 2). In order to improve significantly the performance of an ATR system, it is essential to utilize both multiframe information and a multimode tracker with look-ahead capability. The multimode tracker allows the use of centroid, correlation, and target feature tracking for maintaining target track. Input to the aimpoint selection block comes from the MM wave sensor, semantic information, and interframe analysis modules. These provide classification and aspect information. Target prioritization, mission, and guidance information is stored as a feature vector for each target class with several aspects per class. The aspect data obtained from the sensors is used to appropriately select a feature vector from the set of prestored information on that class of target. From this, an aimpoint is computed on the actual target. Normally, moments are used to select the predefined aimpoints of an acquired target. The semantic information block has several functions. It takes the input from the LADAR sensor and generates information regarding the type and aspect of the target. This helps to reduce aimpoint selection time and improves target classification accuracy. It takes its input from the interframe analysis and multimode tracking block. It allows the implementation and processing of any specific information regarding target type and mission scenario. For example, the assumed scenario may imply the need for processing only moving target information relative to several other classes.

When the range is small (100–500 m) there are two options for target classification using FLIR data (note MM wave images may be more useful at close ranges for detection and tracking, since they cover much larger FOVs than infrared images). Either the FLIR image can be demagnified with respect to some standard range so that the inside structure of the target is not visible in the image or structural information present in the image can be used. A simple example of structure utilization is shown in Fig. 8. The image in Fig. 8(a) has two targets, a tank and a truck. Fig. 8(b) shows the results of a zero-crossing edge operator and Fig. 8(c) shows the zero-crossing details of the tank target. Important parameters



(a) (b) (c)

Fig. 8. Use of structural information. (a) FLIR image. (b) Zero crossings of (a). (c) Zero crossings of tank target.

in the classification are the size of the wheels and their relative positions. Circles corresponding to the wheels are obtained using a least squares, RANSAC paradigm model fitting [71] to a set of points or the Hough transform. The classification result with its confidence measure is passed to the following classification block in the hierarchy. Structural information becomes very useful for missiles with fiber optics, where all the computing is done at the ground base.

The next examples are for the cases where a target is not detected by static image analysis, but by using motion analysis (temporal context) and other contextual cues. Fig. 9 shows an air-to-ground image. The potential targets detected by a double gate filter are pointed to by

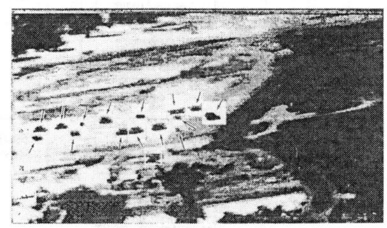
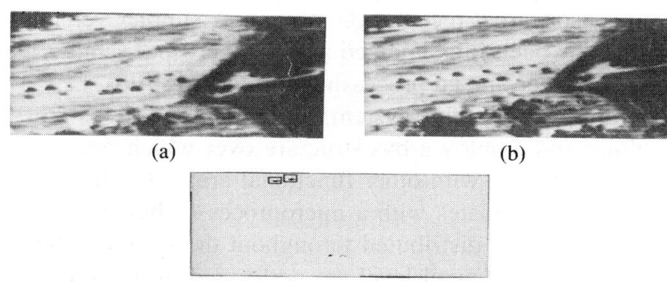


Fig. 9. Detection of targets using double gate filter. Moving target near top is not detected.

arrows [17]. However, it does not detect the moving target near the top of the image. In Fig. 10 two consecutive frames are shown of the same scenario as shown in Fig. 9. Using optical flow field analysis [47], the moving target is now detected. In Fig. 10(c) the predicted and the detected true targets are shown in



(a) (b) (c)

Fig. 10. Use of optical flow field analysis for detecting moving target near top of images. (a) Frame 1. (b) Frame 2. (c) Predicted and detected targets shown in boxes.

boxes. Fig. 11 shows another example where a large number of targets have been detected [17]. However, the three true targets, which are pointed to by arrows, are detected using the context, ''along the road side.'' Thus,

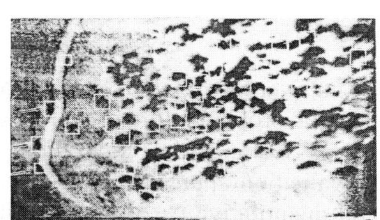


Fig. 11. Use of contextual information (''along the road side'') to detect and classify targets.

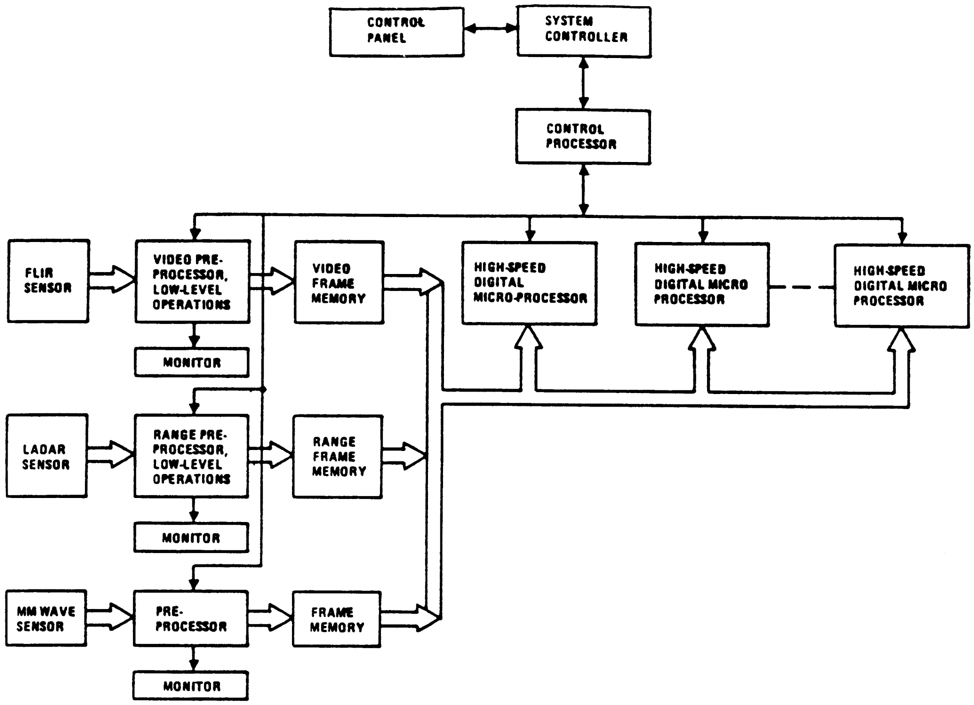


Fig. 12. Architecture of multisensor ATR system.

making use of context and semantics, detection and classification performances are improved.

System Architecture for the Multisensor System Implementation: The hardware and software implementation of the algorithms is an essential consideration to the solution of the autonomous targeting problem. Fig. 12 shows the architecture of the multisensor target recognition system. A simpler version of this has been implemented at Ford Aerospace under its modular video image processing system (MVIPS) program [72]. For this system, all hardware designs are modular and employ a bus structure over which the modules interact with other functional areas. Control of the system originates with a microprocessor-based system controller and is distributed throughout the system. The operations at the pixel level are performed in the three preprocessors whose inputs are from FLIR, LADAR, or MM wave sensors. Higher level processing (symbolic manipulation and control) is done in the expandable high-speed digital microprocessors. Frame memories are used to store the frame data (resolution 512 $\times$ 512) needed for interframe analysis and intermediate results. The system has an internal clock of 40 MHz and can accept video at a sample rate of 25 MHz.

VI. CONCLUSIONS

In this paper an overview of the ATR problem is presented. The major problems in the evaluation of algorithms are discussed and some solutions are suggested. The evaluation approaches are discussed in detail and several quantitative performance criteria are presented. Information measures of the image and some typical characteristics of FLIR images allow one to evaluate the scene with respect to the difficulty in detecting targets, and to set the requirements for the collection and organization of the data base. As sensor and VHSIC technology is progressing and data base collection efforts are being undertaken, new parallel algorithms are being developed and are being improved by utilizing semantic, contextual, structural information, and hierarchical reasoning in a multisensor system. This will improve the performance of an ATR system and even add the capability for countermeasures. Other available knowledge sources and techniques which should be incorporated in the algorithm design involve the use of appropriate context and integration of information from diverse sources. The sources of information are moving and stationary targets, location of the horizon and ground targets, targets moving in bulk, map and mission-specific information, etc. Some of the techniques to be used are the efficient focussing mechanism based on hypotheses testing, and on competition and cooperation among these hypotheses for locating targets from frame to frame, efficient search strategies for the storage and retrieval of data in classification, and classifier designs for different learning environment scenarios from supervised to vicissitudinous. The multisensor and other systems, as presented here, are attempts to make use of some of these concepts and techniques. For a logical sensor specification methodology, see [73] which describes a coherent and efficient treatment of the information received in a multisensor environment including fault tolerant capability. It is anticipated that a single algorithm may not be able to provide optimum performance under all scenarios. It is best to use different algorithm designs implemented such that they are compatible with the ATR system in which they are to be used. Depending upon the requirements the design can be selected to achieve maximum performance. All this leads to the development of recognition expert systems with flexible control strategies in a restricted domain and unification of pattern recognition and AI techniques.

ACKNOWLEDGMENTS

The author is grateful to Ford Aerospace & Communications Corporation, Newport Beach, Calif., where this work was performed.

REFERENCES

[1] Walker, P.F. (1983)
Smart weapons in naval warfare.
Scientific American, 248 (May 1983), 53–61.

[2] Scott, W.B. (1982)
Lantirn system flight testing begins.
Aviation Week and Space Technology (Sept. 20, 1982), 117–126.

[3] Barbe, D.F. (1981)
VHSIC systems and technology.
IEEE Computer (Feb. 1981), 13–22.

[4] Bhanu, B. (1983)
Evaluation of automatic target recognition algorithms. Architecture and algorithms for digital image processing.
In *Processing of SPIE, 435* (Aug. 1983), pp. 18–27.

[5] Lloyd, J.M. (1982)
Thermal Imaging Systems.
New York: Plenum Press, Oct. 1982.

[6] *Proceedings of The IRE*
Special issue on "Infrared Technology". (Sept. 1959).

[7] *IEEE Computer*
(Special issue on "Computer Architectures for Image Processing") Jan. 1983.

[8] Barr, A., and Feigenbaum, E.A. (1982)
The Handbook of Artificial Intelligence, Vol. 1, 2, 3.
William Kaufmann Inc., 1982.

[9] Huang, T.S. (Ed.) (1981)
Two-Dimensional Digital Signal Processing II, Transforms and Median Filters.
New York: Springer-Verlag, 1981.

10] Helland, A.R., Willet, T.J., and Tisdale, G.E. (1981)
Application of image understanding to automatic tactical target acquisition. Techniques and applications of image understanding.
In *Proceedings of the SPIE, 281* (1981), pp. 26–31.

[11] Soland, D.E., and Narendra, P.M. (1979)
Prototype automatic target screener: Smart sensors.
In *Proceedings of the SPIE, 178* (Apr. 1979), pp. 175–184.

[12] Narendra, P.M. (1981)
A separable median filter for image noise smoothing.
IEEE Transactions on PAMI (Jan. 1981), 20–29.

[13] Lo, C.M. (1980)
Forward looking infrared (FLIR) image enhancement for the automatic target cuer system: Image processing for missile guidance.
In *Proceedings of the SPIE, 238* (1980), pp. 91–102.

[14] Milgram, D.L., and Rosenfeld, A. (1978)
Algorithms and hardware technology for image recognition.
Final report to U.S. Army Night Vision and Electro-Optics Lab., Fort Belvoir, Va., Mar. 31, 1978.

[15] Burton, M., and Benning, C. (1981)
Comparison of imaging infrared detection algorithms: Infrared technology for target detection and classification.
In *Proceedings of the SPIE, 302* (1981), pp. 26–32.

[16] Schachter, B.J. (1982)
A survey and evaluation of FLIR target detection/segmentation algorithms.
In *Proceedings of DARPA Image Understanding Workshop*, Sept. 1982, pp. 49–57.

[17] Politopoulos, A.S. (1980)
An algorithm for the extraction of target-like objects in cluttered FLIR imagery.
IEEE Aerospace and Electronic Systems Society Newsletter (Nov. 1980), 23–37.

[18] Bhanu, B., Politopoulos, A.S., and Parvin, B.A. (1983)
Intelligent autocueing of tactical targets in FLIR images.
In *Proceedings of the IEEE Conference on CVPR* (1983), pp. 502–503; for a more detailed version of the paper see Architecture and algorithms for digital image processing.
In *Proceedings of the SPIE, 435* (Aug. 1983), pp. 90–97.

[19] Lacina, W.B., and Nicholson, W.Q. (1979)
Passive determination of three dimensional form from dynamic imagery.
In *Proceedings of SPIE, 186* (May 1979).

[20] Minor, L.G., and Sklansky, J. (1981)
Detection and segmentation of blobs in infrared images.
IEEE Transactions on Systems, Man, and Cybernetics (Mar. 1981), 216–232.

[21] Frey, R.L., et al. (1981)
ATAC autocuer modeling analysis.
Report Ft-80-70-1325, Hughes Aircraft Co. report for U.S. Army Night Vision Laboratory, Jan. 1981.

[22] Mitchell, O.R., and Lutton, S.M. (1978)
Segmentation and classification of targets in FLIR images: Image understanding systems and industrial applications.
In *Proceedings of the SPIE, 155* (1978), pp. 83–90.

[23] Rubin, L.M., and Tseng, D.Y. (1978)
Automatic target cueing.
In *Proceedings of the 4th International Conference on Pattern Recognition* (1978), pp. 774–779.

[24] Tisdale, G.E. (1980)
The AUTO-Q digital image processor for target acquisition, tracking, and handoff.
Presented at IRIS'80, Seattle, Wash., May 1980.

[25] Narayanan, K.A., and Rosenfeld, A. (1981)
Image smoothing by local use of global information.
IEEE Transactions on Systems, Man, and Cybernetics, 11 (1981), 826–831.

[26] Brown, D., and Frei, W. (1980)
Target segmentation in infrared images.
In *Proceedings of the 5th International Conference on Pattern Recognition* (1980), 957–960.

[27] Danker, A.J., and Rosenfeld, A. (1981)
Blob detection by relaxation.
IEEE Transactions on Systems, Man, and Cybernetics (Jan. 1981), 79–92.

[28] Rosenfeld, A., et al. (1979)
Blob extraction by relaxation.
In *Proceedings of the DARPA Image Understanding Workshop*, Apr. 1979, 61–65.

[29] Bhanu, B., and Faugeras, O.D. (1982)
Segmentation of images having unimodal distributions.
IEEE Trans. PAMI (July 1982), 408–419.

[30] Bhanu, B., and Holben, R.D. (1984)
Model based segmentation of FLIR images, Applications of digital image processing VII.
In *Proceedings of the SPIE, 504* (1984), pp. 10–18.

[31] Schneier, M. (1983)
Using pyramids to define local thresholds for blob detection.
IEEE Trans. PAMI (May 1983), 345–349.

[32] Burt, P., Hong, T.H., and Rosenfeld, A. (1981)
Segmentation and estimation of image region properties through cooperative hierarchical computation.
IEEE Transactions on Systems, Man, and Cybernetics (Dec. 1981), 802–809.

[33] Hartley, R.L., Kitchen, L.J., Wang, C.Y., and Rosenfeld, A. (1982)

Segmentation of FLIR images: A comparative study.
In *Proceedings of the DARPA Image Understanding Workshop*, Sept. 1982, pp. 323–341.

[34] Chen, C.H., and Yen, C. (1982)
Object isolation in FLIR images using Fisher's linear discriminant.
Pattern Recognition, 15, 3 (1982), 153–159.

[35] Yu, T.S., and Fu, K.S. (1983)
Recursive contextual classification using a spatial stochastic model.
Pattern Recognition, 16, 1 (1983), 89–108.

[36] Rounds, E.M., King, T., and Steffy, D. (1980)
Feature extraction from forward looking infrared (FLIR) imagery: Image processing for missile guidance.
In *Proceedings of the SPIE, 238* (1980), pp. 126–135.

[37] Cheatham, L., Casasent, D., and Fetterly, D. (1983)
Distortion invariant recognition using a moment feature space.
In *Proceedings of the IEEE Conference on CVPR* 1983, pp. 171–174.

[38] Smith, F.W., and Wright, M.H. (1971)
Automatic ship photo interpretation by the method of moments.
IEEE Transactions on Computers (Sept. 1971), 1089–1095.

[39] Parvin, B. (1984)
A structural classifier for ship targets.
In *Proceedings of the 7th International Conference on Pattern Recognition* (1984), pp. 550–552.

[40] Sethi, I.K., and Sarvarayudu, G.P.R. (1982)
Hierarchical classifier design using mutual information.
IEEE Transactions on PAMI, PAMI-4 (July 1982), 441–445.

[41] Ford Aerospace and Communications Corp. (1983)
Advanced Antishop Targeting Development (Technical Appendix).
Newport Beach, Calif.: Ford Aerospace & Communications Corp., Aug. 22, 1983.

[42] Bhanu, B., and Faugeras, O.D. (1984)
Shape matching of two-dimensional objects.
IEEE Transactions on PAMI (Mar. 1984), 137–156.

[43] Dasarathy, B.V., and Sheela, B.V. (1979)
A composite classifier design: concepts and methodology.
Proceedings of the IEEE, 67 (May 1979), 708–713.

[44] Narendra, P.M., Westover, B., and Panda, D.P. (1980)
Advanced target tracking by dynamic scene analysis: Electro-optical technology for autonomous vehicles.
In *Proceedings of the SPIE, 219*, pp. 148–155.

[45] Reischer, B. (1979)
Target tracking methodologies present and future.
In *Proceedings of the Workshop on Imaging Trackers and Autonomous Acquisition Applications for Missile Guidance*, Nov. 19–20, 1979, Redstone Arsenal, Ala., pp. 156–165.

[46] Ayala, I.L., et al. (1982)
Moving target tracking using symbolic registration.
IEEE Transactions on PAMI, Sept. 1982, pp. 515–520.

[47] Holben, R.D. (1980)
An MTI (moving target identification) algorithm for passive sensors.
In *Proceedings of the IEEE NAECON*, 1980, pp. 114–121.

[48] Dorrough, D.C., Pizzurro, V.F., Bhanu, B., and Pasek, J.R. (1982)
A multi-sensor multi-mode tracker approach to missile ship targeting.
In *Proceedings of the Tri-Service Workshop on Missile Ship Targeting*, Naval Research Laboratory, Washington, D.C., August 10–12, 1982.

[49] *IEEE Computer*
(Special issue on "Computer Analysis of Time-Varying Images") Aug. 1981.

[50] Gilbert, A.L., et al. (1980)
A real-time video tracking system.
IEEE Transactions on PAMI, Jan. 1980, pp. 47–56.

[51] Tseng, D.Y., et al. (1980)
Intelligent bandwidth compression.
Final report to U.S. Army Night Vision and Electro-Optics Lab., Fort Belvoir, Va., Feb. 1980.

[52] Kim, J.H., et al. (1983)
Investigation in context cueing (ICC).
Report to U.S. Army Night Vision and Electro-Optics Lab., Fort Belvoir, Va., Feb. 1983.

[53] Spiessbach, A.J., and Gilmore, J.F. (1983)
Context cueing techniques.
Report to U.S. Army Night Vision and Electro-Optics Lab., Fort Belvoir, Va., Feb. 1983.

[54] Drazovich, R.J., Lanzinger, F.X., and Binford, T.O. (1981)
Radar target classification.
In *Proceedings of the IEEE Conference on PRIP* (1981), pp. 496–501.

[55] Applications of artificial intelligence. (J.F. Gilmore, Editor)
In *Proceedings of the SPIE, 485*, 1984.

[56] Matsuyama, T. (1980)
A structural analysis of complex aerial photographs.
Ph.D. dissertation, Kyoto University, Japan, Apr. 1980.

[57] Dasarathy, B.V. (1980)
Information processing for target recognition from autonomous vehicles: Electro-optical technology for autonomous vehicles.
In *Proceedings of the SPIE, 219*, pp. 86–93.

[58] Gonzalez, R.C., et al. (1978)
A measure of scene content.
Proceedings of the PRIP (1978), 385–389.

[59] Kitchen, L., and Rosenfeld, A. (1981)
Edge evaluation using local edge coherence.
IEEE Transactions on Systems, Man, and Cybernetics (Sept. 1981), 597–605.

[60] Martelli, A., and Montanari, U. (1972)
Optimal smoothing in picture processing: An application to finger prints.
In *Proceedings of IFIP*, Congress 71, Amsterdam: North Holland, 1972.

[61] Goehrig, G.A., and Ledford, L.G. (1980)
Analysis of image segmentation approaches with emphasis on performance evaluation criteria: Smart sensors II.
In *Proceedings of the SPIE, 252* (1980), pp. 124–129.

[62] Parvin, B., and Bhanu, B. (1983)
Segmentation of images using a relaxation technique.
In *Proceedings of the IEEE Conference on CVPR* (1983), pp. 151–153.

[63] Bribiesca, E., and Guzman, A. (1979)
How to describe pure form and how to measure differences in shapes using shape numbers.
In *Proceedings of the IEEE Conference on PRIP* (1979), pp. 427–436.

[64] Weszka, J.S., and Rosenfeld, A. (1978)
Threshold evaluation techniques.
IEEE Transactions on Systems, Man, and Cybernetics (Aug. 1978), 622–629.

[65] Sethi, I.K. (1981)
A fast algorithm for recognizing nearest neighbors.
IEEE Transactions on Systems, Man, and Cybernetics, 11 (Mar. 1981), 245–248.

[66] Friedman, J.H., Bentley, J.L., and Finkel, R.A. (1977)
An algorithm for finding best matches in logarithmic expected time.
ACM Transactions on Mathematical Software, 3 (1977), 209–226.

[67] Garf, C., and Trimble, J. (1980)
Multisensor image processing integration for missile guidance.
Image Processing For Missile Guidance.
In *Proceedings of SPIE, 238* (1980), pp. 316–321.

[68] Kruse, P.W. (1980)
Why the military interest in near-millimeter wave imaging.
Millimeter Optics, SPIE, 259, 1980, pp. 94–99.

[69] Roeder, R. (1982)
Millimeter wave technology advances the science of force multiplication.
Military Electronics/Countermeasures (Dec. 1982), 24–26.

[70] Lin, Y.K., and Fu, K.S. (1983)
Automatic classification of cervical cells using a binary tree classifier.
Pattern Recognition, 16, 1 (1983), 69–80.

[71] Fischler, M.A., and Bolles, R.C. (1981)
Random consensus: A paradigm for model fitting with applications in image analysis and automated cartography.
Communication of the Association of Computer Machinery, 24 (June 1981), 381–395.

[72] Mack, H., et al. (1981)
Autonomous acquisition simulator system: A testbed for tactical image processing algorithm development. *Techniques and Applications of Image Understanding.*
In *Proceedings of the SPIE 281*, 1981, pp. 66–77.

[73] Henderson, T., and Shilcrat, E. (1984)
Logical sensor systems.
Journal of Robotic Systems, 1, 2 (1984), 169–193.

Bir Bhanu (S'76—M'82) received the B.S. degree with honors in electronics engineering from the Institute of Technology, BHU, India, in 1972, the M.E. degree with distinction in electronics engineering from the Birla Institute of Technology and Science, Pilani, India, in 1974, the S.M. and E.E. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1977, the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1981, and the M.B.A. degree from the University of California, Irvine, in 1984. He also received a diploma in German from Banaras Hindu University, India, in 1971.

During 1974 he was on the faculty of the Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani, India. In the spring of 1975 he was a graduate research student at the Rensselaer Polytechnic Institute, Troy, N.Y. At the Research Laboratory of Electronics, M.I.T., he worked on the problem of computation of complex cepstrum. At USC he was at the Image Processing Institute where he worked on the problems of recognition of 2-D and 3-D objects, segmentation, and motion analysis. In 1978 he worked with the computer science department of IBM Research Laboratory at San Jose in the image processing area. During the winter of 1980 he was a Research Fellow at the Institut National de Recherche en Informatique et en Automatique, France, where he worked on 3-D shape representation. From July 1981 to March 1984 he was an Engineering Specialist with the Aeronutronic Division of Ford Aerospace and Communications Corporation, Newport Beach, Calif. At Ford he worked on various aspects of the automatic target recognition problem and the industrial and tactical applications of image processing, pattern recognition, and artificial intelligence. Since March 1984, he has worked as an Assistant Professor of Computer Science at the University of Utah where he is engaged in teaching and research in the areas of computer vision, robotics, pattern recognition, and artificial intelligence.

Dr. Bhanu is a member of the Association for Computer Machinery, Sigma Xi, IEEE Computer Society, Pattern Recognition Society, and the Society of Photo-Optical Instrumentation Engineers.