Machine Intelligent Automatic Recognition of Critical Mobile Targets in Laser Radar Imagery

Richard L. Delanoy, Jacques G. Verly, and Dan E. Dudgeon

A variety of machine intelligence (MI) techniques have been developed at Lincoln Laboratory to increase the performance reliability of automatic target recognition (ATR) systems. Useful for recognizing targets that are only marginally visible (due to sensor limitations or to the intentional concealment of the targets), these MI techniques have become integral parts of the Experimental Target Recognition System (XTRS)—a general-purpose system for model-based ATR. Using laser radar images collected by an airborne sensor, the prototype system recognized a variety of semi-trailer trucks with high reliability, even though the trucks were deployed in high-clutter environments.

THE CONSTRUCTION of an automatic target recognition (ATR) system is a demanding task. ATR systems must be able to locate and identify specific targets that can be concealed intentionally through obscuration or camouflage, that are often designed to be nearly invisible in radar imagery, and that can be deployed in the midst of distracting signals. To gain tactical advantage, it is generally important that an ATR system be able to find a target from as far away as possible. Under such conditions, the selectively indicative signal features (signatures) associated with a target are often barely discernible from the background. Thus, inevitably, practical ATR systems must be able to discriminate targets from background in spite of weak, ambiguous, uncertain, variable, or even contradictory evidence.

ATR system development can be particularly difficult under certain mission constraints and when the costs of system error are high. One such ATR application is the use of airborne sensors to recognize strategic relocatable targets (SRT) such as the SS-25 ICBM of the former Soviet Union (Figure 1 [left]). An ATR system for recognizing SRTs must search through images generated by one or more sensors (laser radars, real- or synthetic-aperture radars, passive infrared imagers, and video cameras), requiring techniques of data fusion. The search is for a very small number of targets in a continent-sized area. The targets might be caught in the open, but more likely will be found along tree lines, perhaps partially occluded by foliage. Because of the nature of the targets, a high probability of detection is crucial. And yet the ATR system must generate few false alarms (FA) due to mission limits on the number of weapons that an aircraft can carry to destroy the SRTs, the flying time of the aircraft over the target area, and the processing capabilities of human operators who must decide which detections to pursue. A closely related mission is the detection of Scud launchers, such as those used by Iraq in the Persian Gulf War. A general term, critical mobile target (CMT), refers to all mobile missile launchers, including those used with SS-25 and Scud missiles.

FIGURE 1. Photographs of (left) mobile missile launcher carrying a strategic SS-25 ICBM of the former Soviet Union, as shown in a Soviet newspaper, and (right) tank truck used by Lincoln Laboratory as a substitute vehicle to develop and test an automatic target recognition (ATR) system for detecting critical mobile targets (CMT) such as the SS-25 launcher.

To achieve reliable detection and recognition performance in such demanding applications, we have developed several new machine intelligence (MI) techniques, including new approaches for model-based classification [1–3], automatic learning of models [4], knowledge-based signal processing [5, 6], selective attention [7], and pixel-level data fusion [7]. The Experimental Target Recognition System (XTRS) developed at Lincoln Laboratory [8–10] provides a framework for the application of these techniques and for the rapid prototyping of ATR systems. Though XTRS and these new MI techniques were intended specifically for ATR, they constitute a general-purpose approach to object recognition, with many potential applications. For example, XTRS has been applied successfully to the detection and tracking of hazardous weather phenomena, as described in the article "Machine Intelligent Gust Front Detection" by Richard L. Delanoy and Seth W. Troxel in this issue [11]. In the current article, we apply XTRS to the detection and recognition of CMTs, specifically, tank trucks (Figure 1 [right]) and logging trucks used as substitutes for missile launchers.

Low-Level Machine Intelligence

Computer vision systems have traditionally been designed in terms of a hierarchy of levels. Low-level vision works on a domain of pixel-level data. Because the associated image processing operations are highly repetitive and therefore relatively slow, low-level operations tend to be kept simple. And, because a scene can contain many objects of potential interest, low-level operations tend to be generic and relatively devoid of object-dependent knowledge. A typical low-level operation is edge detection. In high-level vision, the pixel-level data (for the edge-detection example, a pair of images showing the strength and orientation of edges) are transformed into symbolically described features. Object identification is then performed by matching these features against prior knowledge of object characteristics.

This basic organization of computer vision has often been used in the design of ATR systems. In the detection process, which is analogous to low-level vision, a threshold is applied to a set of signals. The signals can come directly from a sensor or they can be the result of a signal processing operation. For the threshold to be effective, the signals associated with targets must form a distribution that is distinguishable from the distribution of signals associated with other objects in the background (i.e., clutter). Figure 2(a) illustrates this point. In any realistic detection problem, there will exist some targets that have signals below the detection threshold; those targets will

not be detected. There will also exist instances of clutter associated with signals that are above the detection threshold; such instances will result in FAs.

Developers of ATR systems have generally followed the strategy of keeping low-level vision devoid of object-dependent knowledge, and the processing done in preparation for the application of thresholds is usually kept simple. As a result, the thresholds must be set fairly low to maximize the likelihood of detecting targets. The high-level recognition process is then responsible for suppressing as many FAs as possible. Usually, FA suppression is accomplished through the use of classifiers based on statistical techniques or MI techniques, including those involving expert systems, model-based matching, and neural networks. However, when a given ATR application involves only one or at most a few intended targets or classes of targets, the use of object-dependent knowledge for the detection process *is* feasible. And, in fact, because targets are often hidden, camouflaged, or otherwise only marginally visible, object-dependent knowledge can play an important role in enhancing detectability. In particular, detection performance can be improved by separating the distributions of targets and clutter, as shown in Figure 2(b). Although various techniques have been developed to increase this separation (see, for example, Reference 12), the techniques do not typically involve any detailed object-dependent knowledge.

Thus, in addition to using MI techniques in the conventional role of high-level classification and FA suppression, we have developed a set of new techniques for low-level MI. Thresholds are still an unavoidable part of detection but, when low-level MI is applied directly to the pixel-level data early in the detection process, the use of thresholds is in a relative sense postponed and, as a result, made more effective.

Interest Images

A key to implementing low-level MI in XTRS is the concept of *interest* and *interest image* [7]. By our definition, *interest* is a dimensionless quantity indicating the likelihood that a specific feature, indicative of a target or class of targets, is present at a given image pixel. A spatial map of such interest values, each constrained to the range [0,1], constitutes an *interest image.* Clusters of high interest values are used as a guide to focus computational resources on likely targets. In Figure 2(a), a threshold was applied to the quantity "signal strength." For simple ATR systems, the signal is typically the intensity of returned electromagnetic energy or a simple function thereof. Interest provides an alternative flexible metric to which thresholds can also be applied. The power of this approach is that the output of any sensor modality or feature detector can arguably be expressed as an interest image. Furthermore, the use of interest as a common denominator greatly simplifies the fusion of pixel-level data. Specifically, interest enables the use of simple arithmetic or fuzzy logic to fuse spatial evidence from a variety of sources.

There are three steps in low-level MI as used by XTRS in the detection process. First, relevant feature detectors are selected, given knowledge of the situational context. The context may include the intended set of targets, various sensor-related parameters, and identifiable environmental conditions affecting sensor performance and target appearance. Often the use

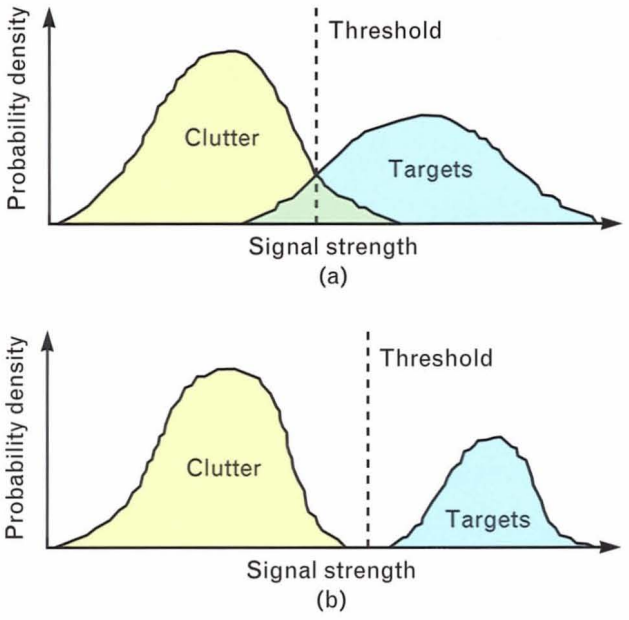


FIGURE 2. Discrimination of target and clutter signals through the application of a threshold: (a) typical overlap of targets and clutter distributions, and (b) illustration of how signal processing, either conventional or machine intelligence (MI) based, can improve detection performance by increasing the separation of the distributions.

of just one feature detector can accomplish adequate target-detection performance.

In the second step, the selected feature detectors are applied to the appropriately prepared input imagery. Each detector generates as its output an interest image that provides spatial evidence for the presence of particular target features. A targeted object may be represented by more than one detector; each detector looks, for example, for a distinct set of features or for an alternative target configuration.

The final step calls for the fusion of evidence, which is accomplished with a rule of combination prescribing how interest values from multiple interest images are to be combined. The rule of combination depends on the set of feature detectors selected. In the case of multiple feature detectors looking for alternative target configurations, the rule of combination could be the maximum (*fuzzy-or* in fuzzy set theory); i.e., at a specific pixel location, the maximum of the interest values across all interest images at that location could be used. In the case of several feature detectors looking for different vehicle features that are likely to be present all the time, the fusion of interest values might be done by an averaging process. Although not fully exercised in the CMT version of XTRS, the rule of combination could be arbitrarily complex to reflect knowledge of the variable reliability of different feature detectors under different viewing conditions.

For situations in which only targets return strong intensity signals, the intensity signal returns might provide a ready-made interest image. In practice, however, laser intensity can be an unreliable discriminant because the energy returned from a target surface depends on the specularity of the surface and its orientation relative to the incident laser beam. Also, high (or low) range values from a laser radar are usually unreliable predictors of target locations because targets are not customarily parked at the highest (or lowest) points in a locality. Thus the effective use of laser radar imagery requires that objects in the imagery be identified also on the basis of shape.

Functional Template Correlation

In studying the principal techniques for shape analysis, we found that the basic equations of cross-correlation and mathematical morphology (MM) [13] can be generalized into a single class of operations, which we have called *functional shape matching.* Briefly described, these shape-analysis tools all use kernels (structuring elements in MM), which are basically subimages that are looked for within the image to be probed. For probing a pixel in an image, the origin of a particular kernel is positioned over the pixel's location. A two-argument function is then applied to each kernel value and the corresponding value in the image. (For cross-correlation, the two-argument function is multiplication. For MM dilation and erosion, the functions are addition and subtraction, respectively.) Next, an arbitrary operator is applied to the function values obtained for the set of pixel locations on the kernel. (For cross-correlation, the operator is summation. For MM dilation and erosion, the operators are maximum and minimum, respectively.)

Eventually, we realized that functional shape matching not only includes the classic shape-analysis tools, but it also encompasses a variety of signal processing techniques that have never been tried before. From functional shape matching, we implemented a tool for generalized matched filtering called *functional template correlation* (FTC) [5]. Whereas the kernel of the classic techniques is a subimage indicating specific expectations of image values for a successful match, the kernel used in FTC is a set of indexes, each corresponding to a unique scoring function. Each of these scoring functions can define arbitrary expectations for image values at each pixel location on the kernel. The outputs of these scoring functions are scalar values, which are averaged and then "clipped" to the range [0,1]. (In the clipping process, those averaged scores which are less than zero are assigned a value of zero, while those averaged scores which are greater than one are assigned a value of one.) Comparable in spirit to the membership functions of fuzzy set theory, scoring functions provide a means of encoding uncertainties. But, in addition, scoring functions can be used to encode a surprising amount of knowledge of the physics of a matching problem. Using FTC, we can construct customized matching techniques that are more powerful than the classic shape-analysis operations. (Note: For a brief introduction to FTC, see the box, "Functional Template

Correlation," in the article "Machine Intelligent Gust Front Detection," by Delanoy and Troxel in this issue [11].)

Data Used in System Development and Testing

For the development and testing of a CMT version of XTRS, a large dataset of images in Maine was collected to simulate the detection of CMTs in a high-clutter, mainly forested environment. Semi-trailer trucks, which approximate the appearance of missile launchers, were positioned amidst both natural and man-made clutter. The simulated targets included the tank truck shown in Figure 1 (right), the same tank truck but under camouflage netting, a loaded logging truck, and an empty logging truck. This variety of vehicles was used to test XTRS's ability to discriminate between targets of similar shapes and sizes. The vehicles were positioned on or near roads, both in the open and along tree lines. The man-made, or cultural, clutter included residential neighborhoods and a logging camp (Figure 3) that contained heavy logging equipment such as other semi-trailer trucks.

Pixel-registered range and intensity images of the various vehicles were generated with the Hughes-Danbury GaAs Laser Linescanner carried aboard a Gulfstream G-1 aircraft (Figure 4). Characterized by a 0.85-μm wavelength, the linescanner has a range ambiguity of 10 m, a precision of ambiguous-range

FIGURE 3. Aerial photographs of man-made, or cultural, clutter represented in the laser-radar-image dataset. Contained in the photographs are railroad cars, fuel tanks, stacks of logs, empty logging trucks, other road vehicles, and heavy logging machinery. All of these objects, easily confused with the set of targets being sought, are potential sources of false alarms (FA).

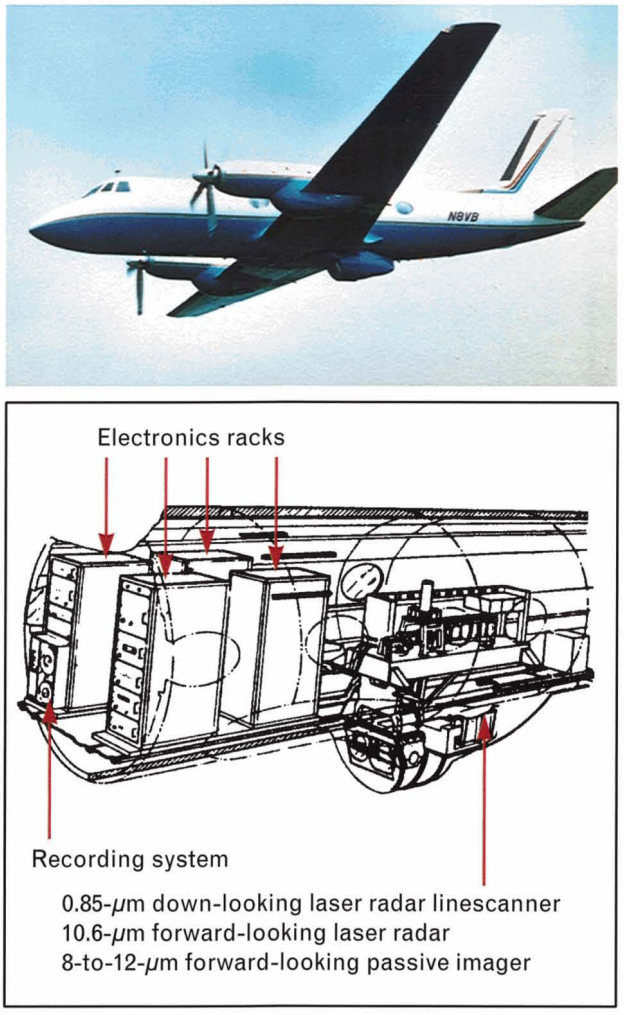


FIGURE 4. Lincoln Laboratory airborne sensor platform. The Gulfstream G-1 carries a 0.85-μm down-looking laser radar linescanner, a 10.6-μm forward-looking laser radar, and an 8-to-12-μm forward-looking passive imager. The images used in the experiments described in this article were collected with the 0.85-μm linescanner.

values of 15 cm, and an angular resolution of 1.0 mrad. Images were collected during the winter and summer of 1989 with a down-looking sensor platform that was operated at altitudes between 200 and 300 m by the Opto-Radar Systems Group of Lincoln Laboratory. The example range and intensity images shown in Figure 5 reveal the high resolution achieved with the 0.85-μm linescanner. Image widths were between 64 and 150 m; image lengths could be arbitrarily long because of the linescanner used. Long scans were subdivided into overlapping images ranging in length from 100 to 400 m. In total, the dataset collected contained 2303 image pairs (range and intensity) covering 17.13 km^2 of ground area under both winter (snow) and summer (dense foliage) conditions.

System Description

The architecture of the CMT version of XTRS consists of five modules (Figure 6): preprocessing, detection, extraction, decomposition, and matching. Each module has a standard structure (Figure 7) that consists of four main elements: (1) a parameter library—a collection of algorithms, numbers, and/or data structures that encode knowledge relevant to the current stage of processing; (2) a parameter selector—a rule-based expert, i.e., a collection of rules, that uses contextual information and previous results to choose parameters from the parameter library; (3) a generic processing engine; and (4) a rule-based feedback expert that evaluates the output of the processing engine and decides where control should be directed. In the complete system, the feedback expert of one module and the parameter selector of the subsequent module conceptually form a local-control node.

Preprocessing

The Hughes-Danbury GaAs Laser Linescanner produces pixel-registered range and intensity images that, in preparation for the detection process, require a number of data transformations.

First, because an aircraft's speed with respect to the ground varies depending on the wind velocity vector, the aspect ratios of targets that have been imaged by the linescanner can often become distorted. In an operational system, these distortions can be avoided by using an inertial navigation system either to regulate the linescan rate or to provide data that would allow the images to be corrected by interpolation. For the data used in this article, the interpolation was performed interactively to obtain the correct target aspect ratios.

Second, the ambiguous-range values are converted to absolute altitudes above some arbitrary reference altitude. Once the absolute altitudes have been determined, a map of altitudes for the local ground level can be computed with a technique based on morphological operations [14–16]. In the technique, only

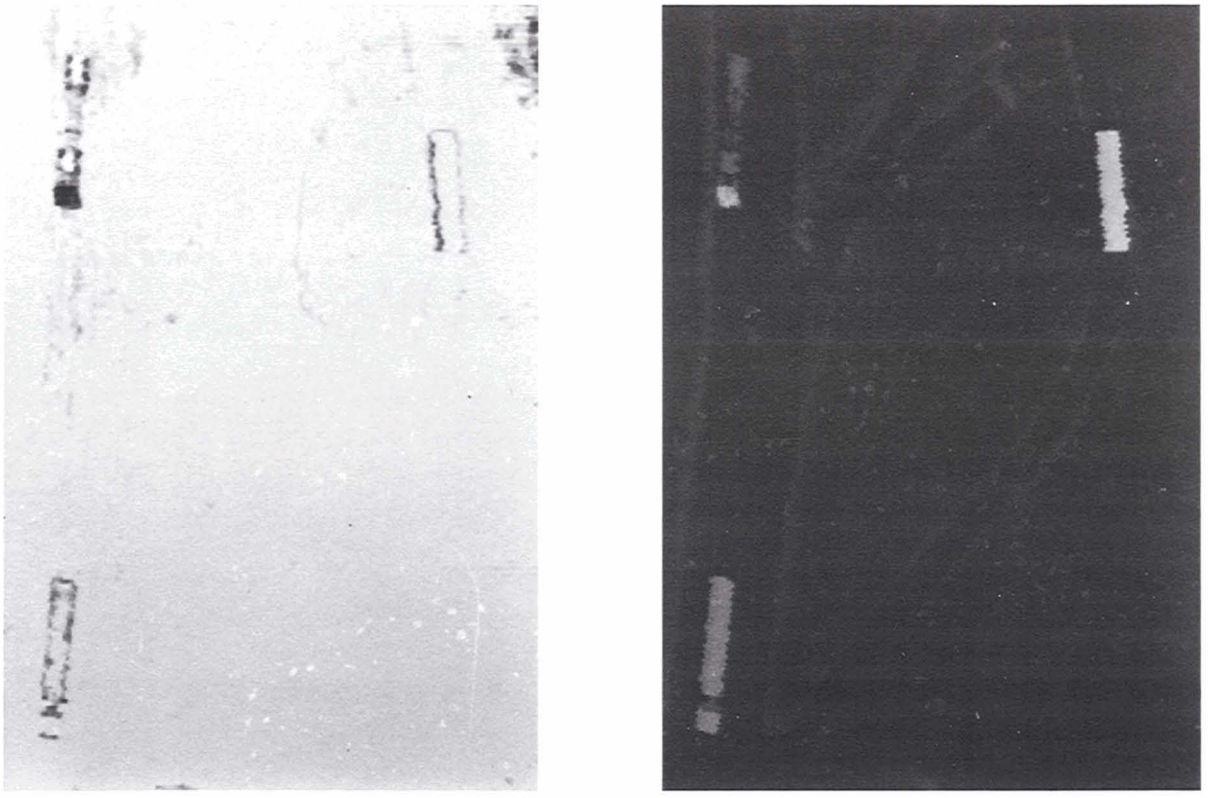
those surface shapes which are wider than the intended targets are retained as part of the local ground level. (Note: There exist other techniques for estimating the local ground level; see, for example, Reference 17.) When the altitudes of the local ground level are subtracted from the absolute altitudes, the resulting image will contain values that are the heights of small objects (including targets) above the local ground level. In this article, subsequent uses of the term *range image* will refer to this image of heights above the local ground level.

For the last step of preparation, the range and intensity images are scaled by linear interpolation to a resolution of 0.25 m per pixel side. Images of lower resolution (0.5 and 1.0 m per pixel side) are then generated by a subsampling of the data. The lower-resolution images are used for detection, while the high-resolution images are used for extraction and high-level matching.

(Note: During preprocessing, no attempt was made to clean the linescanner imagery of noise. Such a procedure was unnecessary due, in part, to both the high quality of the imagery and the noise-resistant properties of FTC.)

Detection

In the CMT version of XTRS, three-dimensional detection is essentially performed by four target detectors (i.e., feature detectors representing whole targets instead of individual features). The tank truck is represented by two alternative target detectors, one in which the truck is exposed, the other in which the truck is covered with camouflage netting. The logging truck is similarly represented by two target detectors, one in which the truck is empty, the other in which the truck is loaded with logs.



(a) (b)

FIGURE 5. Example (a) intensity and (b) range images taken during winter with the 0.85-μm down-looking laser linescanner shown in Figure 4. Note the tank truck (lower left), empty logging truck (upper left), and house trailer (upper right) in both images. The range image has been transformed such that each pixel value represents a height above an arbitrary reference altitude with lighter pixels indicating a greater height.

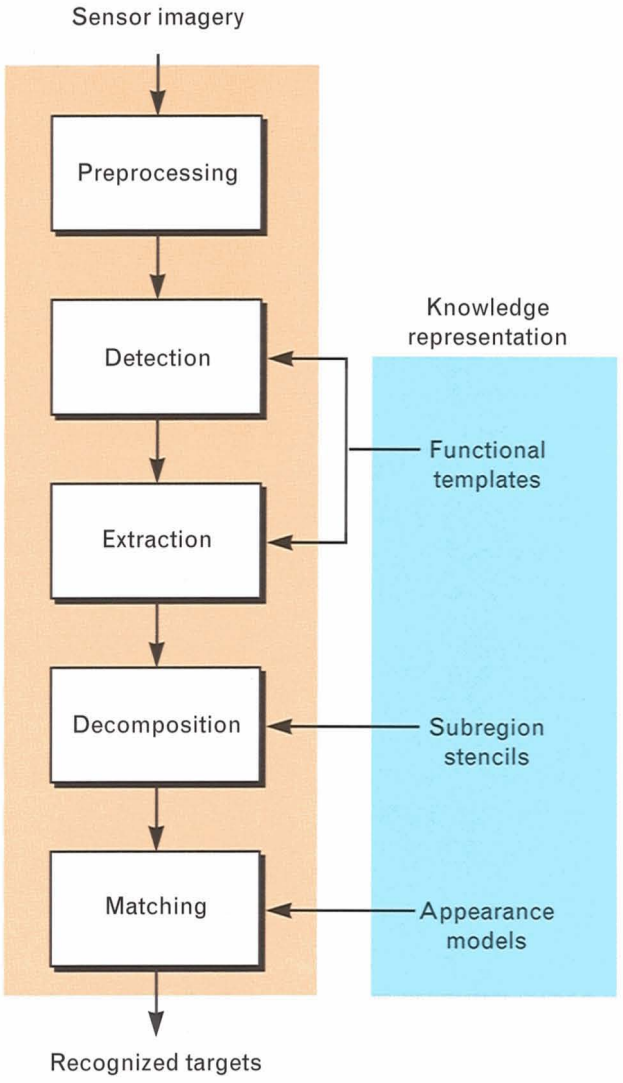


FIGURE 6. Architecture of Experimental Target Recognition System (XTRS), showing the five processing modules as well as the way in which knowledge is represented at each level.

The target detector for the exposed tank truck consists of the two functional templates shown in Figure 8. The first functional template encodes the expected appearance of the truck in range imagery (i.e., images of heights above the local ground level). In scoring function 1, which corresponds to the top surfaces of the cab and trailer, a maximal score of 1.0 is returned for heights from 2.5 to 3.5 m. The uncertainty comes from signal noise and inaccuracies in estimating the local ground level. The negative scores reflect the fact that tank trucks are opaque to laser illumination; i.e., the presence of ground-level heights where the cab or trailer is expected constitutes strong evidence that a target is not present at that location. On the other hand, heights greater than the expected interval of 2.5 to 3.5 m result in scores no less than 0.5, the level of ambiguity, because such heights could potentially indicate the presence of an occluding surface. In other words, the cab of a tank truck might or might not be present under an occluding surface that is at least 4.0 m high.

The other scoring functions work in the same manner, except that the expected interval of heights for the background in scoring function 0 is from 0.0 to 0.5 m, and the expected interval for the hitch area in scoring function 2 is from around 1.0 to 2.0 m. These scoring functions are tuned such that, when the template is applied to a patch of bare ground (zero height), the negative scores from scoring function 1 balance the positive scores generated by scoring functions 0 and 2, resulting in an overall score near 0.0. And, of course, an unobscured target should generate a score near 1.0.

The above functional-template design provides a simple means of minimizing the effects of occlusion. This capability, not easily obtained with cross-correlation or MM operations, is necessary for finding targets partially covered by foliage. In the extreme, a target that is completely occluded should generate a score of 0.5 (Figure 9).

In Figure 8, the second functional template for the exposed tank truck is designed for intensity imagery. Because the surface of the truck is smooth (specular) with regard to the laser wavelength, the reflected laser beam will tend either to miss the sensor (resulting in a low intensity value) or hit the sensor directly (resulting in a high value). Scoring function 4, corresponding to the trailer and cab, encodes these expectations by returning high scores for very low and very high intensity values, and low scores for intermediate intensity values. The hitch area returns the laser energy more diffusely; thus, scoring function 5, which corresponds to the hitch area, returns high scores for intermediate intensity values. The intensity values associated with surrounding ground areas are highly variable and unpredictable, except that they are seldom very low. Scoring function 3 encodes this expectation with highly negative scores for low intensity values, but *nil*

(i.e., no opinion) for intermediate and high intensity values.

The two functional templates shown in Figure 8 were applied simultaneously to the input range and intensity images, and overall scores were computed as the average of the scores returned from the six scoring functions. Because the orientations of targets are typically arbitrary and unknown *a priori,* an FTC score was computed for each of 36 uniformly spaced template rotations (10° increments) at each pixel location of the input imagery. For a particular pixel location in an input image, the score associated with the maximally scoring orientation was assigned to the corresponding pixel location in the output image. As previously indicated, each such output image is treated as an interest image, indicative of the likelihood of finding a target at any particular pixel.

In a similar way, output interest images were also generated for each of the three other target detectors, and the four interest images were combined by taking the maximal scores at each pixel location. The resulting combined interest image was then scanned for pixels having interest scores above a certain threshold (typically 0.75), and the above-threshold pixels were grouped into clusters.

Next, a box was placed around each cluster. The boxes were used to extract range and intensity subimages containing the interest cluster and thus the candidate target. The boxes were square, with sides 50% longer than the longest dimension of the targets being sought. Up to four boxes with above-threshold interest scores and a minimum of overlap with each other were constructed for each image. The cluster of above-threshold interest scores that led to the creation of a particular box was used to create a list of target hypotheses. At each pixel in a cluster, the interest score is always associated with the highest-scoring target detector at the highest-scoring orientation. Each pixel's hypothesis consisted of the highest-scoring target detector's name, pixel coordinates, orientation,

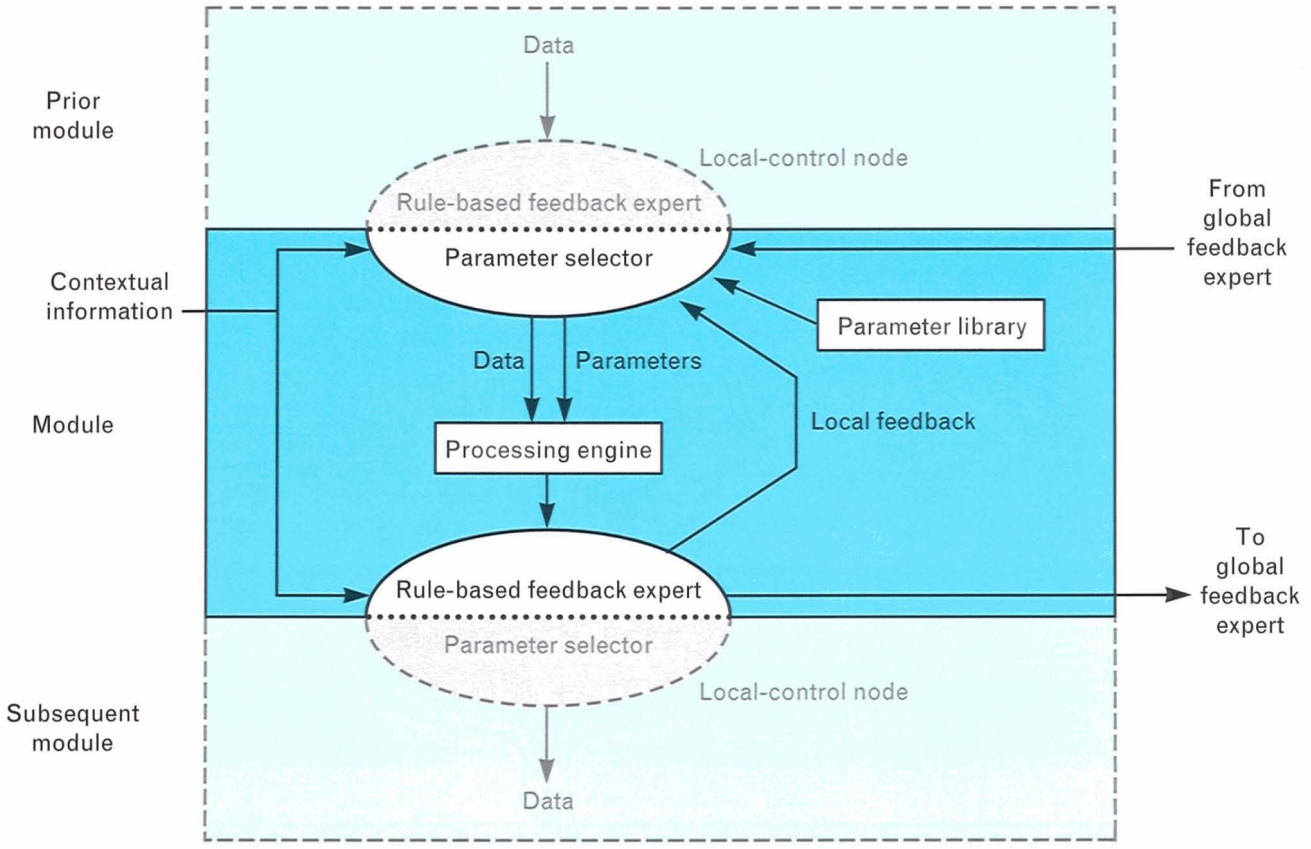


FIGURE 7. Standard structure of the XTRS processing modules shown in Figure 6. Note that the feedback expert of one module and the parameter selector of the subsequent module conceptually form a local-control node.

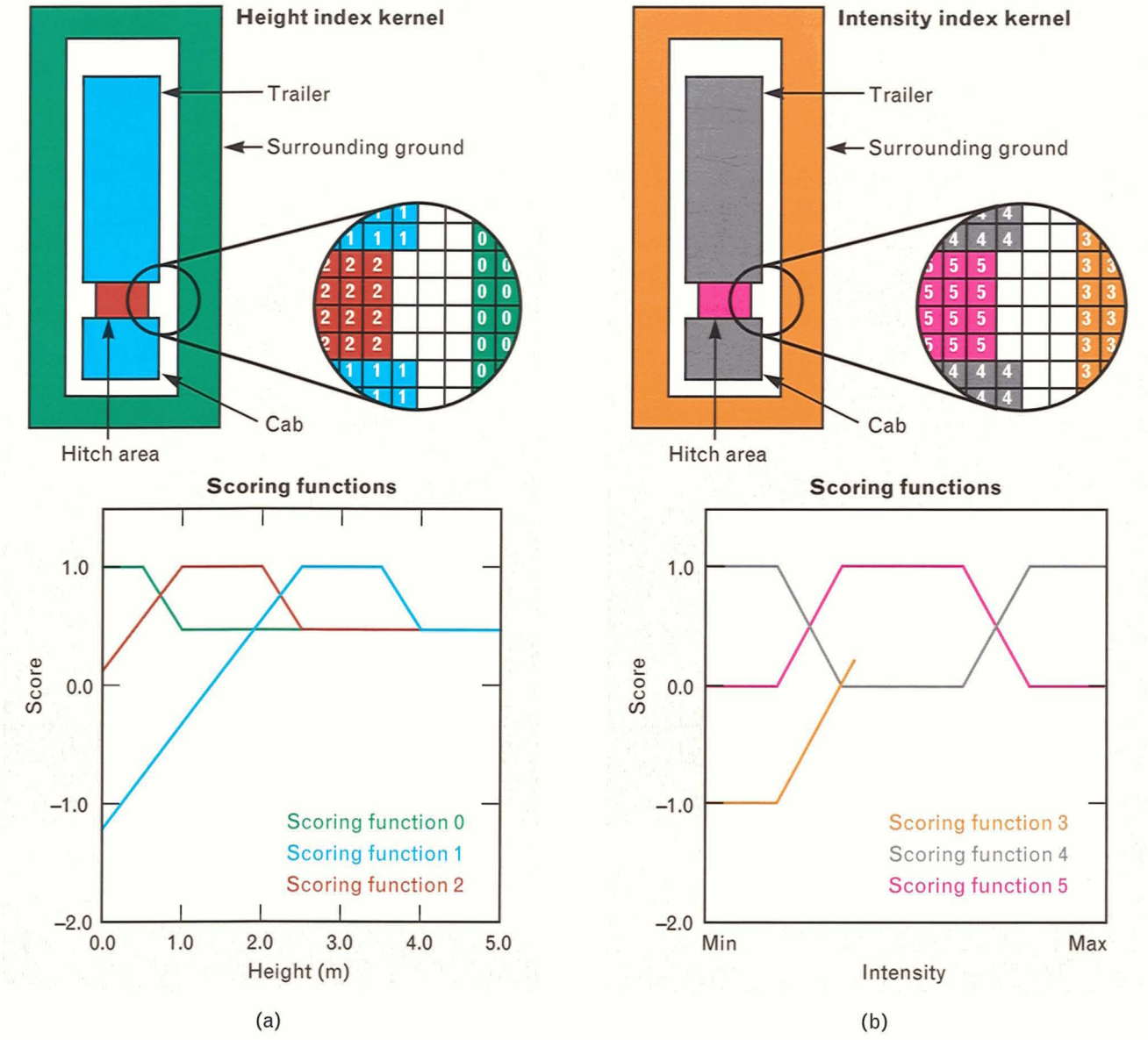


FIGURE 8. Functional template for the top view of the tank truck. Shown are the index kernels and indexed scoring functions for (a) range and (b) intensity images. Note that scoring function 3, i.e., the scoring function for the surrounding ground in intensity imagery, has no opinion above a certain intensity value.

and interest score, which was used to rank the hypotheses. For each box generated, information regarding the size and location of the box as well as hypotheses about what might be in the box was placed in a data structure referred to as a "window." Because the CMT version of XTRS uses the function maximum for the rule of combination of interest scores, the score achieved at any pixel location by the highest-scoring target detector is also the value stored at that same location in the window.

Extraction and Decomposition

Windows generated by the detection process are used as input to extraction. The position and size of each window are used to extract full resolution (0.25 m per pixel side) subimages of range, intensity, and interest. In the extraction module (Figures 6 and 7), the parameter selector chooses from the library an extractor corresponding to the highest-ranking hypothesis in a window. In the current implementation of our sys-

tem, a full-resolution functional template is created for the extractor by a zooming process that is applied to the corresponding detection template. The preliminary location and orientation information recorded in the hypothesis is then used to probe the window with the full-resolution template. The window is probed only at the pixels immediately surrounding the hypothesis location and only for orientations within 10° of the hypothesis orientation. Although the angular increment for FTC is 10° for detection, an increment of 1° is used for extraction.

At the location and orientation of the best full-resolution FTC match, a rectangular mask with outside dimensions that approximate the dimensions of the candidate target is positioned to isolate an image region. The isolated region then undergoes the application of height thresholds followed by a cleaning with MM, and the resulting region is subdivided with a stencil consisting of an array of rectangles (six or eight in our application), each marking the area limits of one subregion. For the tank truck, Figure 10(a) shows the eight idealized subregions, each of which is characterized with regard to a number of attributes such as length and width, and various texture measures such as a measure of the local variance in the subregion. The characterized object region and part subregions together with a list of candidate target identities extracted from the window hypotheses serve as input to the matching process. (Note: If the matching module fails to make an identification, control is directed back to the beginning of the extraction module. In such a case, the extraction process is repeated for the hypothesis that has the next highest interest score. The processing stops either when the target has been identified or when all hypotheses have been examined.)

Matching

Candidate targets are identified by matching the object region and part subregions against appearance

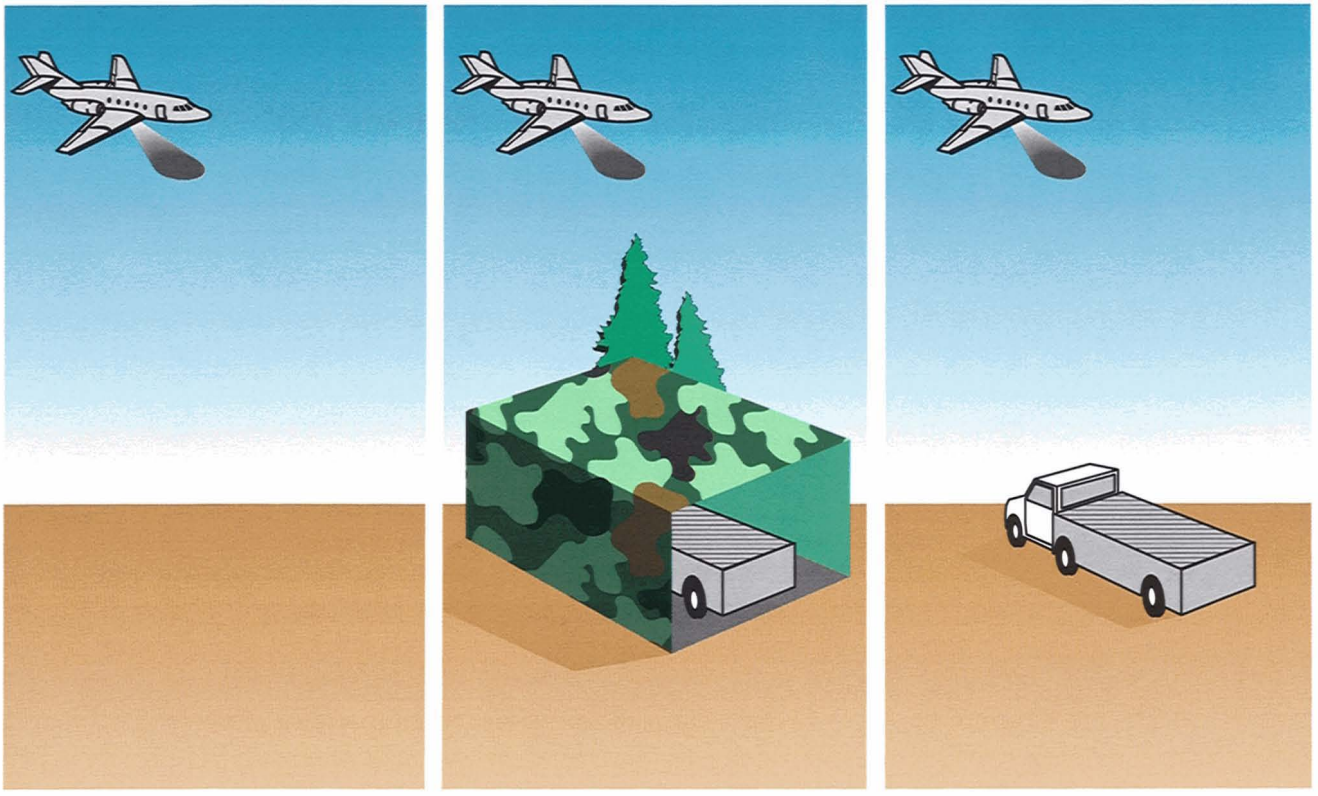
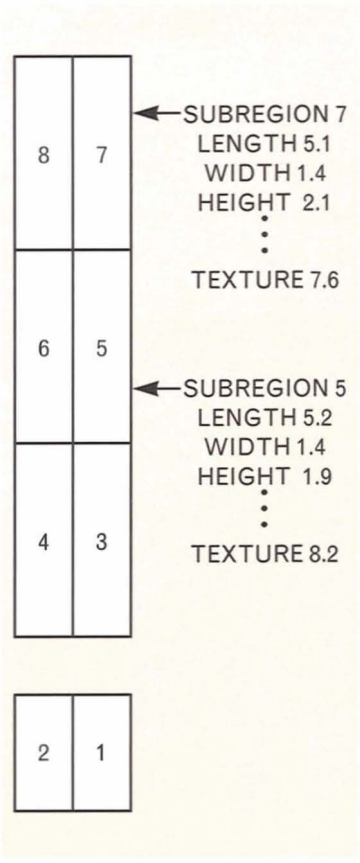
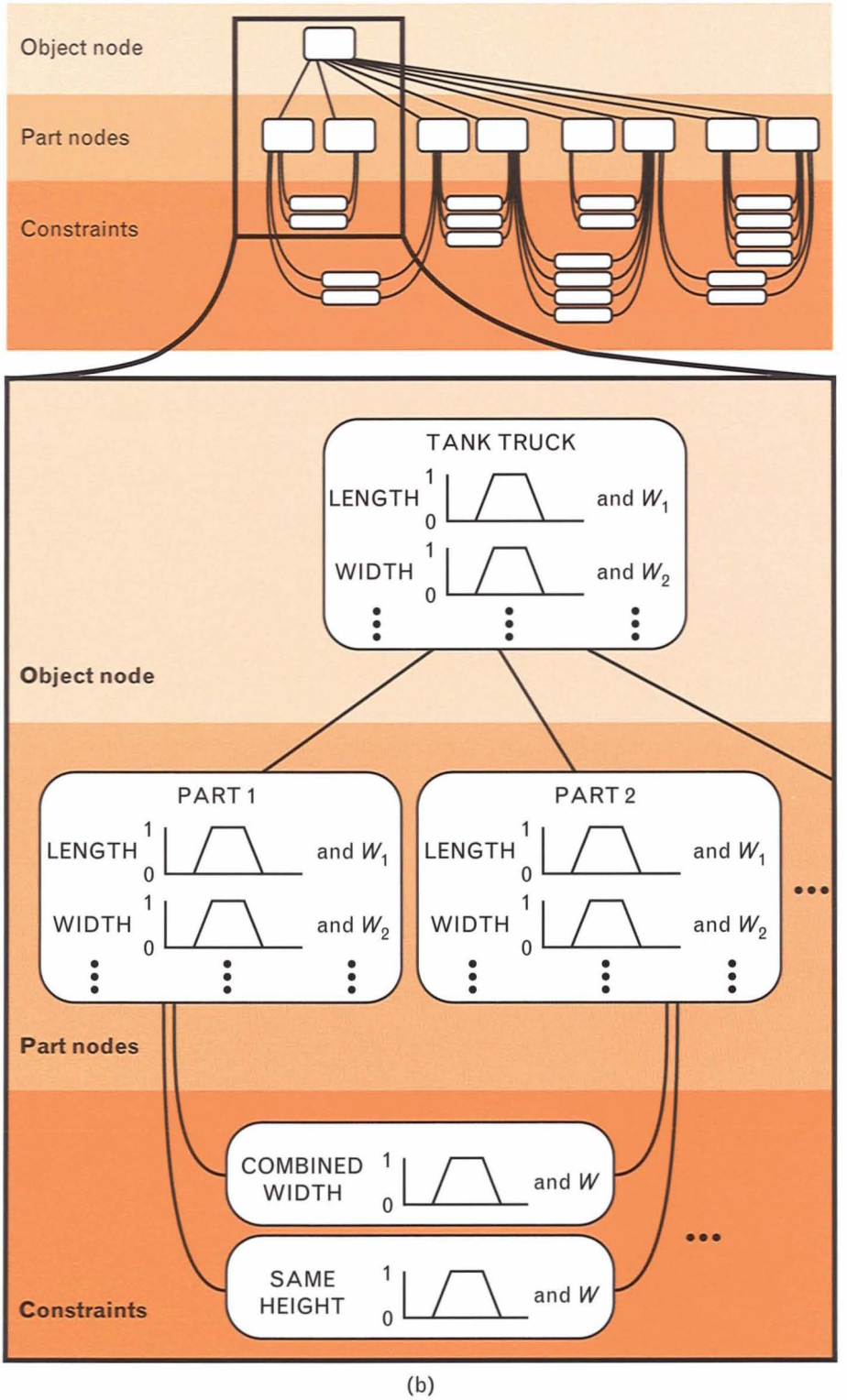


FIGURE 9. The functional template for a tank truck is applied to (a) bare ground, (b) a 100% occluded target, and (c) a fully exposed target. The corresponding expected scores for the three cases are 0.0, 0.5, and 1.0, respectively.



(a)



(b)

FIGURE 10. Appearance models: (a) top view of an idealized tank-truck region decomposed with a target stencil into eight characterized subregions, and (b) corresponding appearance model (AM) for the tank truck. Note that each of the eight subregions is characterized with regard to a number of attributes such as length and width, and various texture measures. In the AM, an object node (TANK TRUCK) is broken into eight part nodes (PART 1 through PART 8) corresponding to the eight subregions. Each object and part node contains a set of fuzzy predicates that define the allowable limits for computed values of the different attributes such as length and width. Each predicate has an associated weight W_i that is used to bias computed match scores. In a similar way, constraints (e.g., COMBINED WIDTH and SAME HEIGHT) specify the limits of the relationships between the different parts. The AM shown here has been simplified. In practice, the AMs of the modeled trucks have as many as 80 constraints between the different parts. Note, also, that for the sake of simplicity, the existence of constraints between certain parts (e.g., between PART 2 and PART 4) has not been shown.

models (AM) [1–3]. Figure 10(b) illustrates the general construction of an AM for the tank truck. Note that the AM consists of an object node (TANK TRUCK) and a series of part nodes (e.g., PART 1) that specify the limits of properties of the different parts. Attributes of the object region and part subregions can include length, width, aspect ratio, circumference, average image value, and texture measurements, among other quantities. Each object and part node contains a set of fuzzy predicates and associated weights that define the allowable limits for the computed values of the different attributes. There is typically one fuzzy predicate for each attribute. In a similar way, constraints (e.g., COMBINED WIDTH and SAME HEIGHT) specify the limits of the relations between parts.

By treating a computed attribute θ as the argument of the corresponding fuzzy predicate $f(x)$, we can easily obtain a score $f(\theta)$ for the computed value θ. The scores obtained from a set of fuzzy predicates together with the weights associated with those predicates can then be used to calculate a weighted average that provides an overall match score for each part. Similarly, a match score can be computed for each constraint. For example, the sum of the widths of PART 1 and PART 2 would be the input to the constraint COMBINED WIDTH shown in Figure 10(b). Match scores for each part and each constraint become pieces of evidence that can then be combined with the Dempster-Shafer theory of evidence. The output is a target identity, which may be *none* to indicate an unknown target type. (Note: References 1 through 3 provide a detailed description of matching based on AMs, including a description of the Dempster-Shafer theory of evidence.)

Using the above approach, we constructed five AMs, one each for the exposed tank truck, the camouflaged tank truck, the loaded logging truck, the empty logging truck, and the truck cabs. The cabs were modeled through a separate AM because, in several cases, the frame boundary of the images had occluded the trailers.

Through experience, we learned that the AMs that were more successful were generally more complicated. As the size and complexity of the AMs grew, however, it became apparent that we could not continue to construct AMs manually. Thus automatic construction techniques were needed.

Automatic model building requires example sets of the decomposed targets. For each attribute of each part of each target, fuzzy predicates can be constructed from the population of values found in the example set. Figure 11 shows a fuzzy predicate that has been constructed for the attribute LENGTH of the part node PART 1. The red dots at the top of the figure represent the population of length values from all PART 1s in the example set. During the construction of a fuzzy predicate, outlier (i.e., statistically inconsistent) values are discarded, and a cluster analysis is performed to determine the number of clusters that might best explain variances in the remaining values. For each cluster, the mean and standard deviation are computed, and an interval of maximum returned score (1.0) is established between the minimum and maximum lengths of each cluster. Outside this interval of maximum returned score, the fuzzy-predicate curve ramps down from 1.0 to 0.0 with a slope that is proportional to the standard deviation σ of the cluster population. The value of σ is multiplied by the coefficient β called the recognition tolerance, to determine the width of the ramping interval. For small values of β, the fuzzy predicate is relatively intolerant of lengths that are outside the already observed range of values, while high values of β result in a greater tolerance of such variations. The final fuzzy predicate is the maximum of the individual functions generated for each cluster. The weight associated with each fuzzy predicate is initialized to 0.1, a value chosen to allow an increase (and decrease) by at least an order of magnitude.

Fuzzy predicates are constructed for each attribute of each part. Not all attributes, however, are equally effective in discriminating targets from clutter. To determine which attributes are effective discriminants, we use a second phase of model building called *supervised discrimination learning*. In the process, weights associated with attributes that are weakly discriminating are decreased, while weights for attributes that are strongly discriminating are increased. Whether an attribute is discriminating or not is determined by individually reevaluating each attribute within the AMs of targets after an incorrect identification has been

made. If a fuzzy predicate returned a high score that contributed to the error, then the associated attribute is nondiscriminating and the corresponding weight is decreased. For example, consider the response to an FA in which some piece of clutter has been incorrectly identified as a target. In the identification process, fuzzy predicates were evaluated for the different attributes. If the score from a particular evaluation was greater than 0.5 (ambiguity), then that attribute contributed to the mistaken identity and is thus not discriminatory; consequently, the corresponding weight is decreased. On the other hand, if the score was less than 0.5, indicating that the attribute had correctly denied the mistaken identity but was outvoted by the other fuzzy predicates, then the associated weight is increased. (Note: Reference 4 contains specific equations and schedules for the weight adjustments, along with a more detailed description of supervised discrimination learning.)

Results

Much of the innovation of the CMT version of XTRS is in the development of techniques for low-level MI. To evaluate the effectiveness of these techniques, this section will present the detection results first, separate from the results of the overall system recognition performance.

Detection Performance

Each of the four FTC-based target detectors was tested individually for a range of interest thresholds. Figure 12 shows the probability of detection P_D plotted as a function of the false detection (FD) rate for the four implemented detectors. (Note: FD is distinct from FA, which is the false-alarm level for the *overall* system.)

For the tank-truck detectors, both exposed and camouflaged, the detection performance was quite good. In both cases, P_D was around 0.7 at the threshold level where the first FD occurred. Given the 17.13 km^2 of ground area covered by the dataset, the one FD resulted in a rate of 0.058 FD/km^2. For a P_D of 1.0, the associated minimum FD rate was approximately 2.0 FD/km^2. The target detector for the loaded logging truck performed slightly less well. Because the shape of the vehicle changed with each load of logs, the detector's functional template had to be constructed with more fuzziness; i.e., the template had to

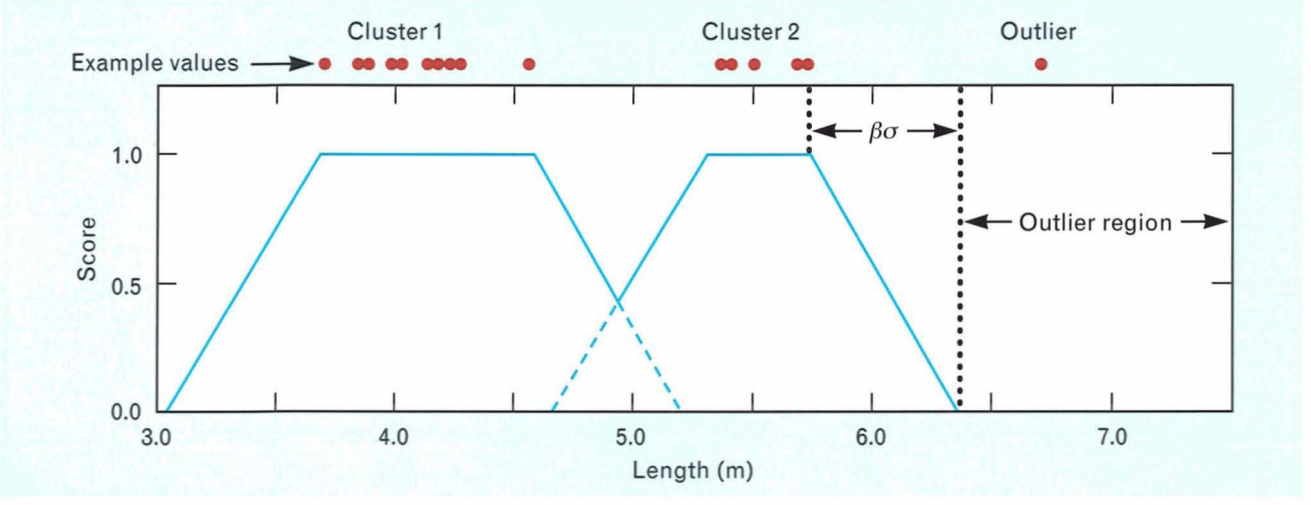


FIGURE 11. Automatic construction of a fuzzy predicate for the attribute LENGTH of the part node PART 1 of Figure 10(b). The red dots at the top of the figure represent the population of length values from all PART 1s in the example set. During the construction of a fuzzy predicate, outlier (i.e., statistically inconsistent) values are discarded, and a cluster analysis is performed to determine the number of clusters that might best explain variances in the remaining values. For each cluster, an interval of maximum returned score (1.0) is then established between the minimum and maximum lengths of that cluster. Outside this interval, the fuzzy-predicate curve for the cluster ramps down to 0.0 with a slope that is proportional to the standard deviation σ of the cluster population multiplied by the recognition tolerance β. The final fuzzy predicate is the maximum of the individual functions generated for each cluster.

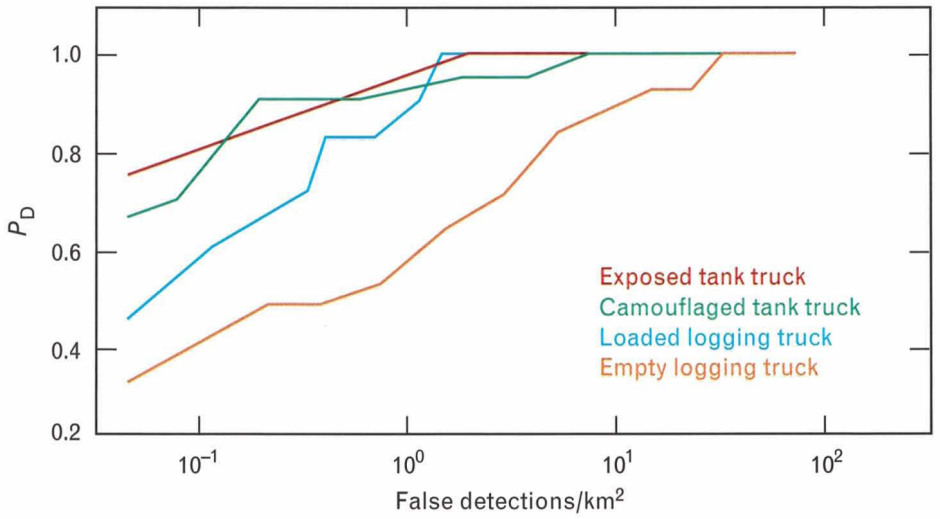


FIGURE 12. Probability of detection P_D plotted as a function of the false-detection rate for the four target detectors. Each point along any of the curves shown is associated with a particular value of the interest (or detection) threshold.

provide a greater tolerance for variations in shape. But the most difficult to represent as a functional template was the empty logging truck, because of the small size of the vehicle's trailer. With an elongate shape such as that of the tank truck, small uncorrected distortions in the length of the target did not have a serious effect. With the empty logging truck, however, the rear axles of the vehicle are the only reliably visible part of the trailer and, because this portion of the truck is short relative to the overall truck length, even a small distortion in the truck length can move the axles ahead or behind the patch of the functional template representing the axles. Consequently, the shape and appearance of the empty logging truck could not be defined as precisely as for the other targets. Fortunately, scoring functions can be modified easily to adjust the degree of tolerance to variations in shape and appearance. Although the detection rates for the two logging-truck configurations were lower for a given FD rate than for the tank truck, the performance was still respectable.

For a better understanding of the sources of FDs, the clutter data were divided into natural and man-made (cultural) clutter. Any image containing a large, man-made object (e.g., a building, non-target vehicles, or stacks of logs) was placed in the cultural-clutter group. Each of the target detectors was then applied to both divisions of the data. Figure 13 shows example results for the loaded logging truck. The probability density was computed as the percent of all detections found within each successive short interval of interest scores (0.016 in the range from 0.0 to 1.0).

Most instances of natural clutter (mainly trees and shrubs) tended to have interest scores around 0.65, with no interest scores above 0.8. The population of target interest scores (shown as red dots at the top of Figure 13) had scores ranging from 0.79 to 0.94. Thus the detector for the loaded logging truck achieved a perfect partition between targets and natural clutter (i.e., a threshold of 0.78 resulted in 100% detection with no FDs). In contrast, man-made objects were a more troublesome source of FDs because such objects generated a few interest scores that were as high as 0.9. Included in the high-scoring cultural objects were other semi-trailer trucks, such as the deployed tank truck, and stacks of logs similar in shape to the loads carried by the logging trucks. The results for the other target detectors were roughly the same as that for the loaded logging truck with the detectors for the tank truck providing a slightly better partition between targets and clutter, and the detector for the empty logging truck a slightly worse partition.

It can be argued that logging trucks provide more stressful testing than would arise from an actual CMT application. Because missile launchers are large and nonarticulated, their detection is less vulnerable to

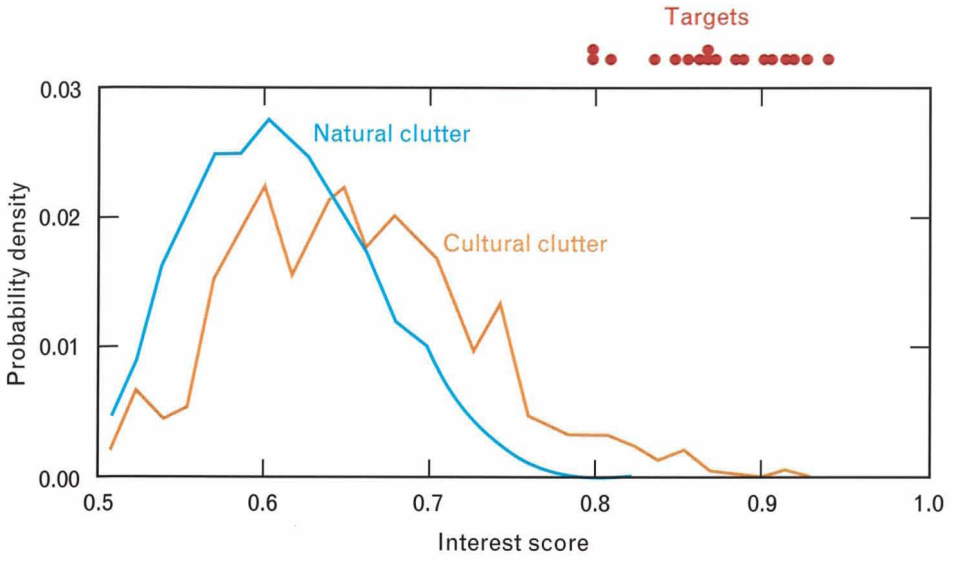


FIGURE 13. Distributions of interest (or detection) scores for natural and man-made (cultural) clutter for the loaded-logging-truck target detector. The red dots at the top of the figure indicate the interest scores for instances of the deployed loaded logging trucks. At an interest (or detection) threshold of about 0.78, the loaded-logging-truck target detector achieves perfect discrimination between deployed loaded logging trucks and natural clutter. There is no threshold, however, that would yield a perfect discrimination between these targets and cultural clutter.

the effects of distortion. Also, CMTs have only three basic shape variations: missile down for transport, missile erected for launch, and without a missile following launch. These three variations have precise known shapes, in contrast to the amorphous nature of log loads.

Occlusion Experiments

One of the primary motivations for the development of FTC was to overcome the way in which occlusion disrupted the more traditional shape-matching techniques. Various attempts to design an MM approach for detection and extraction failed with targets that had as little as 5% of their surface areas occluded by foliage. With functional templates, however, all targets could be detected readily without much challenge.

To explore the limitations of FTC in detecting targets occluded by foliage, we designed experiments in which targets that were cut out from one image were positioned along a tree line within another image. Beginning at locations where the target was completely unobscured, the target was incrementally moved under the foliage, with the vehicle's major (i.e., longitudinal) axis either perpendicular or parallel to the tree line. Figure 14 summarizes the results of the experiment with the target perpendicular to the tree line for target occlusions of 2%, 36%, and 66%.

The left frames in each row are range images of a tank truck that has been synthetically placed perpendicular to the tree line. The center frames show the location and orientation of the best match for the tank-truck functional template in the images, along with the corresponding interest scores. The pixels themselves indicate the scores returned from individual scoring functions for each location on the template: black, white, and gray pixels represent 0.0, 1.0, and intermediate values, respectively. The right frames show the results of recognition based on the matching of the decomposed target with AMs of the targets. It should be noted that occluded targets were included in the example set used to build the AMs.

The first row of Figure 14 shows that a target that is almost completely exposed (only 2% occlusion) results in a strong interest score and correct recognition. For a target occlusion of 36%, the interest score is barely above the interest threshold of 0.75, but the target is recognized correctly nonetheless. For an oc-

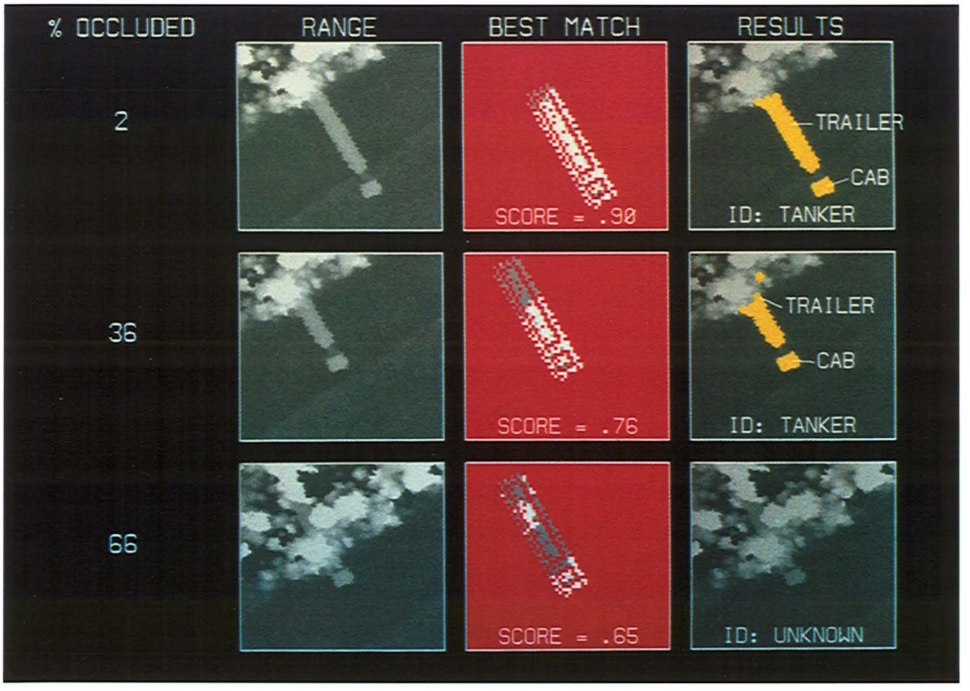


FIGURE 14. Summary of foliage occlusion experiment for a tank truck that has been synthetically positioned perpendicular to a tree line. The top, middle, and bottom rows are for the target with 2%, 36%, and 66%, respectively, of its surface area occluded by foliage. The left frames are range (height above ground) images, the center frames indicate the locations and orientations corresponding to the highest interest scores (indicated in the frames), and the right frames show the final AM-based recognition results.

FIGURE 15. Summary of foliage occlusion experiment similar to that of Figure 14, except the tank truck has been positioned parallel to the tree line.

clusion of 66%, the interest score is only 0.65, which is below the interest threshold of 0.75. Consequently, this target is not detected and therefore not processed by the matching module. (The UNKNOWN identification in the figure is the result of our dropping the interest threshold to below 0.63.) And yet, as shown in the center frame, the best match produced by FTC correctly determines the location and orientation of the target, despite the target's being more than half occluded. We have obtained similar results for the case in which the major axis of the tank truck is parallel to the tree line, as shown in Figure 15.

For the combined data of perpendicular and parallel target placements, Figure 16 contains a plot of interest score as a function of percent occlusion. The figure shows that the decrease in interest score as a function of percent occlusion conforms to an expected linear relationship: performance degrades gradually as occlusion increases, without any intervals of rapid degradation. At an interest threshold of 0.75, targets occluded up to around 36% are detected and recognized. Lowering the threshold would permit the detection and recognition of targets with an even higher percent of occlusion, but would also increase the FD rate.

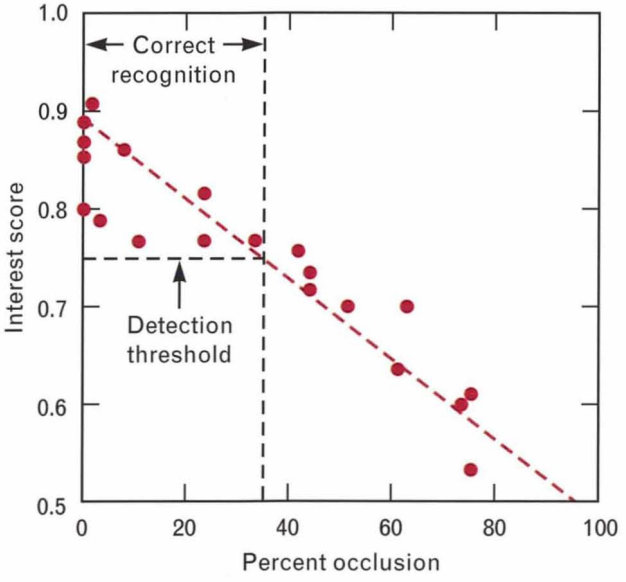


FIGURE 16. Summary of results for the experiments described in Figures 14 and 15. Note that at the detection interest threshold of 0.75, the system is able to detect and recognize targets with up to 36% of their surface areas occluded by foliage.

Overall System Performance

Although FTC can provide high detection rates with few FDs, the fewer FDs the better. Because FTC bases its interest scores solely on how well image values compare to expectations at different locations on the kernel, it does not exploit the known relationships of target parts. We have relied on the technique of AMs as a means of modeling such additional information and, by so doing, have provided the means for rejecting FDs and discriminating between multiple, similarly shaped target classes.

For initial testing, we used an interest threshold of 0.75 to detect clusters of high interest values in the Maine dataset. A total of 492 detections resulted, including all 63 deployed targets. These detected targets were then extracted, characterized, and matched against AMs, as described earlier.

Initially, when we built the AMs we used 50% of the targets as examples and a recognition tolerance β of 0.3. The remaining targets that were classified as UNKNOWN (i.e., insufficiently like any modeled target) were subsequently added to the example set as we refined the models. Eventually, 80% of the targets were included in the example set to reach a recognition performance of 100%. No supervised learning of weights was done for this test. Under these conditions, there were no FAs in all 17.13 km^2 (2303 image pairs) of data.

The above results include some targets deployed in the open, but they also include a number of very difficult cases. Figure 17 shows photographs and a map of a deployment in which the empty logging truck and the camouflaged tank truck were placed on a narrow dirt road with tall trees on either side. For the empty logging truck, Figure 18 contains the range and intensity images, an interest image highlighting pixels having above-threshold interest values and showing the selected windows, and an image showing the final recognition results. The truck, visible in the lower left corner, has been correctly recognized. An FD, triggered by a collection of shrubs having roughly the size and spacing of the parts of the empty logging truck, was correctly rejected during AM-based matching. Figure 19 shows the results for the camouflaged tank truck of Figure 17. Note that in this case the

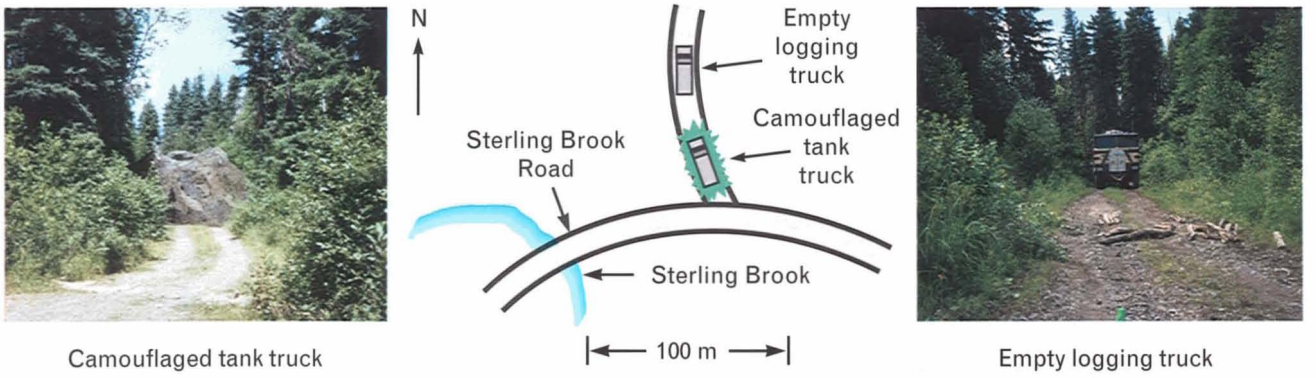


FIGURE 17. Photographs and deployment map of "hidden targets" used in the Portage, Maine, experiments.

target is represented in the range image mostly as a broad tent with only a few pixels having height values corresponding to the ground. Because the shape of the camouflage netting can vary from site to site, the scoring functions for the range functional template were constructed to incorporate considerable uncertainty and were thus weakly discriminating. In intensity images, however, the camouflage netting actually helps make the target stand out against the background, probably because of interference effects caused by some of the laser energy being reflected by the netting material and some being reflected by the ground [17]. The scoring function for the intensity functional template for the camouflaged tank truck exploits this phenomenon.

Generalization of AMs

In the course of evaluating the dataset, we discovered that, in addition to the one logging truck that Lincoln Laboratory personnel had deployed, there were six other empty logging trucks in the vicinity. Only two of these six trucks had interest scores greater than 0.75, and neither of the two was recognized as an empty logging truck. Figure 20 shows intensity and range images containing three of the six non–Lincoln Laboratory trucks, along with a road-mobile crane (top center). The truck in the image that had an interest score above threshold was classified as UNKNOWN. Initially, we were disappointed with this result until we realized that the discrimination made by XTRS was in fact reasonable and useful. Figure 21 (left) is a range image of the logging truck deployed by Lincoln Laboratory, and Figure 21 (right) is a range image of one of the six other trucks. Note that although both vehicles serve the same function and are called logging trucks, their appearances are in fact distinct. The truck deployed by Lincoln Laboratory has a tractor with the cab directly over the engine, while the other vehicle has a hooded tractor with the cab behind the engine. Also, although the overall lengths of the trailers are the same, the trailer of the Lincoln Laboratory truck is narrower and lighter in appearance. If the two vehicles are considered as two distinct objects, then XTRS did successfully discriminate between the two variants with 100% accuracy.

Suppose, however, that AMs we built using examples of one model within a target class were to be used to recognize a more general class of targets, including other models not represented in the example set. All six of the trucks not deployed by Lincoln Laboratory were detected by decreasing the interest threshold from 0.75 to 0.72. And all six vehicles were recognized as empty logging trucks by increasing the recognition tolerance β from 0.3 to 1.7. Thus, by using only two tunable parameters, we could adjust the generality of the recognition for the entire system. But the relaxation needed to generalize the recognition had an associated cost: the FA rate for the system as a whole increased from 0.0 to 1.7 FA/km^2. Of course, instead of generalizing the AMs to include similar related targets, we could have constructed additional functional templates and AMs.

Supervised Discrimination Learning of Model Weights

A common criticism of many research ATR systems is that, because of the limited availability of data, the

FIGURE 18. Detection and recognition results for the hidden empty logging truck of Figure 17. The truck is visible in the lower left corner of the images. A false detection, triggered by a collection of shrubs, has been correctly rejected as UNKNOWN.

FIGURE 19. Detection and recognition results for the hidden camouflaged tank truck of Figure 17.

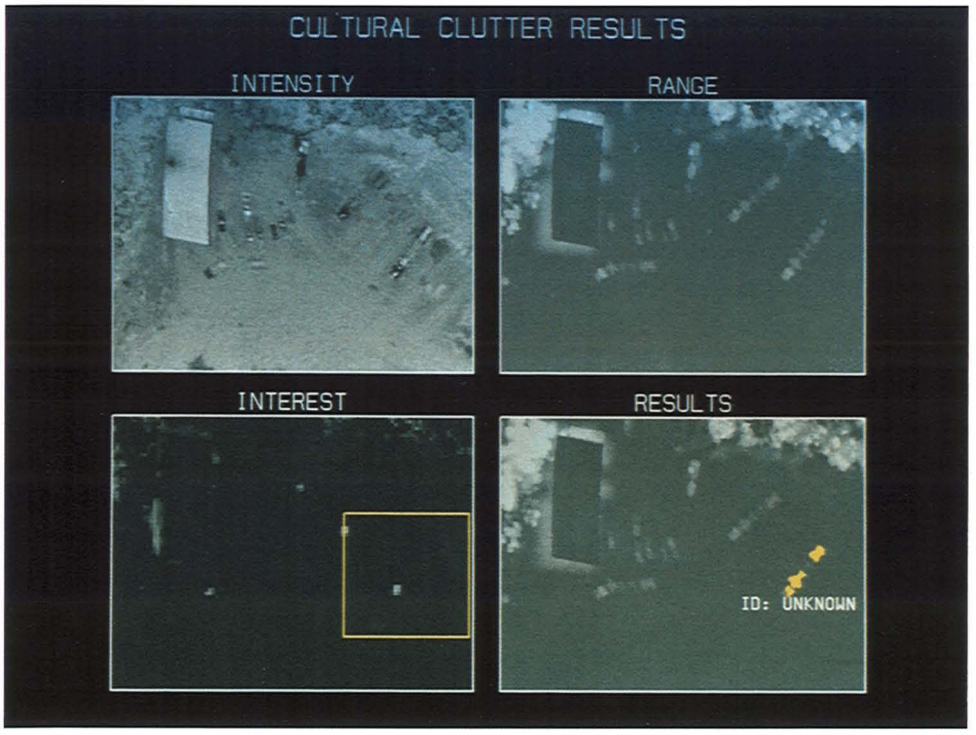


FIGURE 20. Detection and recognition results for three empty logging trucks that were not part of the test deployment. (Note: The images also contain other vehicles, including pickup trucks and a road-mobile crane, at the top center. Also note that in the range image the building in the upper left corner has incorrect height values due to an artifact in estimating the height of local ground for large objects.) Two of the three trucks received below-threshold interest scores and were thus not detected. The remaining truck had an interest score above threshold but was classified as UNKNOWN. The system failed to detect and recognize the three trucks because of differences between them and the logging truck deployed by Lincoln Laboratory (see Figure 21).

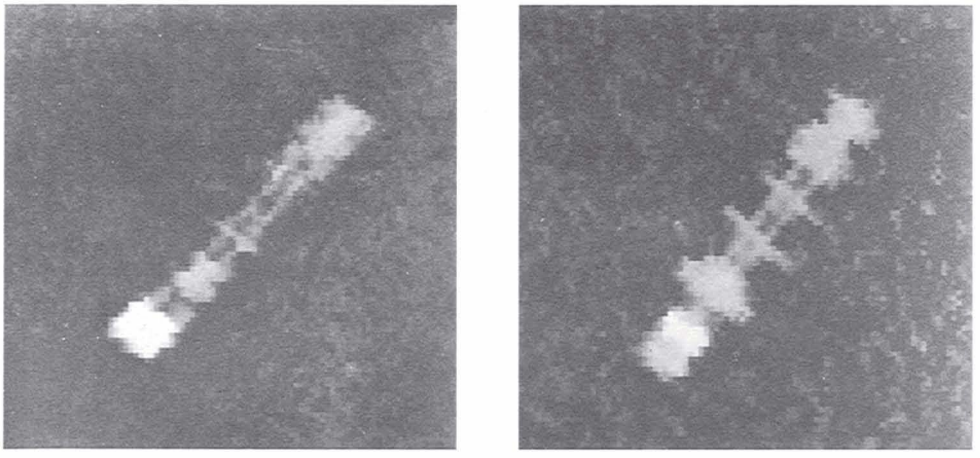


FIGURE 21. Enlarged range images for (left) empty logging truck deployed by Lincoln Laboratory, and (right) empty logging truck discovered in the image dataset. Although both vehicles are called logging trucks, their appearances are in fact distinct. The cab for the truck on the left is directly over the engine, while the cab for the other vehicle is behind the engine. Also, although the overall lengths of the trailers for the two trucks are the same, the trailer for the truck on the left is narrower and lighter in appearance.

dataset used to train a system is the same set used for testing. To a certain extent, we addressed this criticism by dividing our available data into two sets, one for training, the other for testing. Because at most only 50% of the targets were to be used in the example training set, we assumed that a large recognition tolerance β would be required to achieve a P_D of 1.0. With a large recognition tolerance, we also expected that the FA rate might be high. Consequently, the supervised learning of weights was used to suppress the FAs.

For training, we built the AMs with a high β value of 5.0 and an example set consisting of 1165 image pairs (range and intensity) containing 28 targets. AM weights were all initialized to 0.1. To establish a baseline FA rate, we did not use supervised discrimination learning to process the training data. The high β of 5.0 and a low interest threshold of 0.72 were selected so that enough FAs would be generated to promote opportunities for learning. The number of FAs under these conditions was 37 (4.3 FA/km^2). Next, supervised discrimination learning was initiated and, with each complete pass through the training data, the number of FAs generated during that pass was recorded. Figure 22 shows the results of 14 passes through the training data. Note that the number of FAs dropped from 37 to 21 during the first pass and stabilized to an average of 19 FAs (2.2 FA/km^2) by the fourth pass.

After the completion of training, testing was done on 27 targets in 1134 image pairs covering 8.4 km^2 of ground area. Using the weights learned from training,

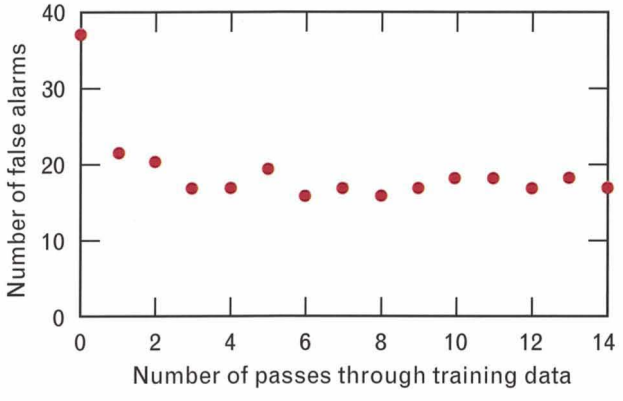


FIGURE 22. Learning curve showing the decrease in false alarms with training.

we built the AMs at progressively larger values of β ranging from 0.1 to 7.0. For each β value, the probability of correct recognition P_R was plotted as a function of the FA rate. Figure 23 shows the results. At the value of β in which the first FA occurs, the P_R was 0.44. For a P_R of 0.93, the FA rate was 1.42 FA/km^2. Two of the targets were not recognized (i.e., classification UNKNOWN) even with β set to 7.0. At some higher value of β, we do expect to achieve a P_R of 1.0, but we did not attempt to find that particular β value. In this test, as well as in the previously described tests, no targets were mislabeled as another target identity.

Out of a total of 63 targets, only 55 were used: 28 for training and 27 for testing. The reason for this intentional omission of eight targets was that there were only four images of the tank truck in the open and four images of an empty logging truck with the trailer completely occluded by the frame boundary of the image. With only two examples of a target for training, the resulting AMs were too restrictive to recognize any targets other than the two training examples. This finding highlights how the building of robust AMs depends on the proper selection of a training set. As with any learning system, a realistic and representative sampling of variations of object appearance is necessary to achieve robust performance.

Conclusions

The Experimental Target Recognition System (XTRS) provides a framework for applying machine intelligence (MI) techniques to the task of automatic target recognition (ATR). Based largely on aspects of fuzzy set theory, these MI techniques enable the representation of uncertainties and known variabilities in target appearance.

With rule-based experts and libraries of functions and data structures, XTRS can be organized to adapt automatically to environmental context and to reconfigure the search for alternative targets. Using multiple target detectors, XTRS can look simultaneously for different variations in target shape. The outputs of all target detectors are expressed as interest images, permitting the fusion of all sources of evidence into a single spatial map. Despite the apparent complexity of XTRS, system performance can be con-

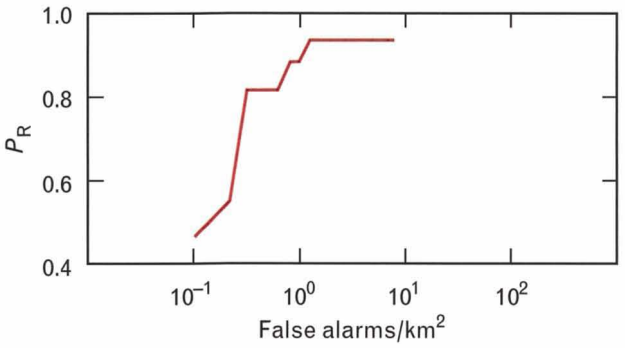


FIGURE 23. Probability of correct recognition P_R as a function of the false-alarm rate. Each point along the curve corresponds to a particular value of the recognition tolerance parameter β.

trolled effectively with only two tunable parameters: the interest threshold for controlling the output of low-level detection and the recognition tolerance for controlling the output of high-level matching based on appearance models (AM).

XTRS uses AMs to model how targets and their constituent parts appear in sensor imagery, thus providing an alternative to other classifiers, including those based on neural network, statistical, and other model-based approaches. Unlike other model-based approaches that encode the three-dimensional structure of an object, AMs define the observable appearance of targets in specific sensor data within the constraints of the likely target orientations. AMs provide a more controllable representation than neural networks. Because knowledge is represented in neural networks as a diffuse population of weights, it is difficult to identify which image features are being used. Not only are the attributes and weights of AMs easy to interpret, they can be modified by a user with predictable effects on recognition performance. Neural networks have gained in popularity as classifiers principally because of their ability to learn and encode discriminants automatically. As we have shown in this article, the automatic learning of class discriminants is also possible with AMs, but in a representation that is more amenable to understanding and selective editing.

Other techniques developed for XTRS embody what we call low-level MI. Most existing MI techniques used in computer vision rely on a preliminary abstraction of raw data into a symbolic form. But the process of abstraction necessarily reduces the amount of information available for decision making, thus handicapping an observer, no matter how intelligent. Tools for knowledge-based signal processing and pixel-level accumulation of evidence provide the intelligent means of using object- and context-dependent knowledge to guide the extraction of information directly from raw image data without the need for abstraction. In particular, functional template correlation (FTC) allows the construction of generalized matched filters that encode knowledge of the physics of a detection problem. Customized operations constructed with FTC are generally more powerful (i.e., more discriminating) than comparable traditional signal processing operations. In ATR versions of XTRS, we have used FTC as a one-step three-dimensional target recognizer. For other applications, we have developed knowledge-based fuzzy variations of standard image processing operations, including thin-line detection, smoothing operations, basic mathematical morphology (MM) operations, and pattern matching.

The need for FTC arose from a perceived inadequacy of the standard techniques of shape analysis. Although MM worked very well for unobscured targets, we could not devise a sequence of MM operations that would reliably detect and extract targets in high-clutter environments, especially when the target was partially occluded. We believe that our failure was due in part to the all-or-nothing nature of MM operations [18]. We have also investigated the use of normalized cross-correlation, the other commonly used tool for shape analysis. In its favor, cross-correlation does provide a variable degree of match that can be translated easily to interest values. But the matches generated by cross-correlation are too literal in that the interest scores are based on very specific, inflexible patterns of image values.

The repetitive evaluation of all scoring functions in a functional template—for all orientations for each pixel location—sounds computationally prohibitive. But the process becomes feasible when the input image values are scaled to some integer range (e.g., 0 to 255) and the scoring functions are implemented as a precomputed two-dimensional lookup table that is indexed by the scoring-function numbers and the integer image values. The use of such a lookup table is

Table 1. Performance of Prototype System

	Experiment			
	1	2	3	
			Training	Testing
Number of targets	63	69*	28	27
Ground area (km^2)	17.13	17.13	8.6	8.4
Detection threshold	0.75	0.72	0.72	0.75
Number of detections	492	1173	656	342
Number of targets in AM example set	50	50	28	—
Recognition tolerance	0.3	1.7	5.0	4.5
% correct recognition	100%	100%	100%	93%
FAs/km^2	0.0	1.7	2.2**	1.4

* Includes six empty logging trucks not intentionally deployed
** After discrimination learning

generally faster than multiplication, making FTC evaluation quicker than cross-correlation.

Low-level MI also allows XTRS to delay the application of thresholds. Instead of applying thresholds either to a single image consisting of raw data or to the output of some simple transformation of the raw data, we can apply the thresholds to maps of interest containing evidence that has been extracted from a variety of sources.

Unlike the AMs, the FTC-based target detectors were constructed and tuned manually. The development of a useful, operational ATR system that is able to adapt swiftly to different targets and mission scenarios requires a mechanism for constructing functional templates automatically. We have developed methods for building functional templates from statistics accumulated from example targets, but these methods have not yet been implemented. Functional templates might also be constructed by using the emerging techniques of genetic programming.

The success of our approach to ATR is indicated by the overall system performance of the prototype system, as summarized in Table 1. In experiment 1, in which we used strict tolerances for the automatic construction of the target AMs, we were able to achieve 100% correct target recognition in the available data with no mislabelings and no false alarms. It is important to note that the AMs are flexible and can be generalized to broader classes of vehicles by the manipulation of a single recognition tolerance. Experiment 2 demonstrates this flexibility and, in particular, the capability for generalization by increasing the recognition tolerances. Six empty logging trucks were found in the dataset that were somewhat different from the one logging truck that was intentionally deployed. These six were appropriately rejected as clutter in experiment 1. Suppose, however, that the additional six trucks were to be included in a broader class of empty logging trucks. By changing just the recognition tolerance from 0.3 to 1.7 in experiment 2, the system was able to recognize the six trucks as empty logging trucks. Of course, the cost of generalizing all models in this manner was that the FA rate increased from 0.0 to 1.7 FA/km^2. Experiment 3 shows that AMs constructed from more limited training sets can be used to recognize targets with reasonable reliability in a separate test set. The training sets were limited in size and did not provide a good representative sampling of vehicle appearances. Consequently, AMs were constructed with large recognition

tolerances in order to achieve high detection rates. The resulting elevated false-alarm rate was suppressed by roughly 50% through the use of supervised discrimination learning. Despite these limitations, reasonably good performance was evident in the separate test dataset.

In contrast, without the techniques of low-level MI and the automatic construction of complex AMs, we were unable to construct an ATR system for this application anywhere near as accurate, flexible, or robust as the one described in this article [18]. Of course, some credit for the performance of the system must go to the quality of the sensor images used. But images of good quality do not necessarily guarantee reliable detection performance. Even with an image of excellent quality, concealment and clutter can make target detection a challenging problem.

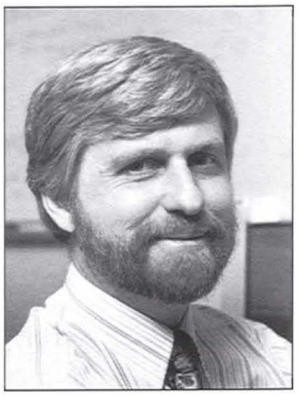
So far, XTRS has been applied to two other ATR problems: the recognition of armored vehicles both in forward-looking laser radar images [8] and in fully polarimetric synthetic-aperture radar images [9]. But XTRS provides the means of solving a more general class of object-detection problems. In addition to its use in recognizing military targets, XTRS has been applied successfully to the task of detecting and tracking hazardous weather phenomena in Doppler weather radars [11].

Acknowledgments

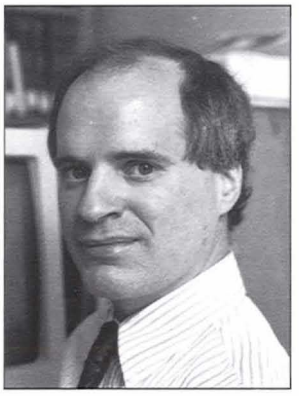
The authors thank their colleagues in the Opto-Radar System Group at Lincoln Laboratory for providing the extensive laser radar dataset used in this research. This work was sponsored by the Defense Advanced Research Projects Agency (DARPA).

REFERENCES

1. J.G. Verly and R.L. Delanoy, "Appearance-Model-Based Representation and Matching of 3-D Objects," *Proc. 3rd Intl. Conf. on Computer Vision* (Osaka, Japan, 4–7 Dec. 1990), p. 248.
2. J.G. Verly, B. Williams, and R.L. Delanoy, "Model-Based Pattern Recognition," U.S. Patent No. 5,123,057 (June 1992).
3. J.G. Verly, B. Williams, and R.L. Delanoy, private communication.
4. R.L. Delanoy, J.G. Verly, and D.E. Dudgeon, "Automatic Building and Supervised Discrimination Learning of Appearance Models of 3-D Objects," *SPIE* **1708,** 549 (1992).
5. R.L. Delanoy, J.G. Verly, and D.E. Dudgeon, "Functional Templates and Their Application to 3-D Object Recognition," *Proc. Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), San Francisco, 23–26 Mar. 1992,* p. III-141.
6. R.L. Delanoy and J.G. Verly, "Computer Apparatus and Method for Fuzzy Template Shape Matching Using a Scoring Function," U.S. Patent No. 5,222,155 (June 1993).
7. R.L. Delanoy, J.G. Verly, and D.E. Dudgeon, "Pixel-Level Fusion Using Interest Images," *Technical Report 979,* MIT Lincoln Laboratory (26 Apr. 1993).
8. J.G. Verly, R.L. Delanoy, and D.E. Dudgeon, "Machine Intelligence Technology for Automatic Target Recognition," *Linc. Lab. J.* **2,** 277 (1989).
9. J.G. Verly, R.L. Delanoy, and C. Lazott, "Principles and Evaluation of an Automatic Target Recognition System for Synthetic Aperture Imagery Based on the Use of Functional Templates," *SPIE* **1960** (1993), to be published.
10. J.G. Verly, R.L. Delanoy, and D.E. Dudgeon, "Model-Based System for Automatic Target Recognition from Forward-Looking Laser-Radar Imagery," *Opt. Eng.* **31,** 2540 (1992).
11. R.L. Delanoy and S.W. Troxel, "Machine Intelligent Gust Front Detection," in this issue.
12. L.M. Novak, M.C. Burl, R.D. Chaney, and G.J. Owirka, "Optimal Processing of Polarimetric Synthetic-Aperture Radar Imagery," *Linc. Lab. J.* **3,** 273 (1990).
13. J. Serra, *Image Analysis and Mathematical Morphology* (Academic Press, New York, 1982).
14. T.R. Esselman and J.G. Verly, "Applications of Mathematical Morphology to Range Imagery," *Technical Report TR-797,* MIT Lincoln Laboratory (Dec. 1987), DTIC #AD-189316.
15. T.R. Esselman and J. G. Verly, "Some Applications of Mathematical Morphology to Range Imagery," *Proc. Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)* **1,** *Dallas, Apr. 6–9 1987,* p. 245.
16. T.R. Esselman and J.G. Verly, "Feature Extraction from Range Imagery Using Mathematical Morphology," *SPIE* **845,** 233 (1987).
17. Private communication.
18. Private communication.



RICHARD L. DELANOY
is a staff member of the Machine Intelligence Technology Group. His work spans the fields of computer vision, machine learning, and construction of object-recognition systems. From 1980 to 1983, he was a research scientist at the University of Virginia Department of Psychology, where he investigated the biochemical correlates of learning and the effects of stress-related hormones on electrophysiological models of memory. Before joining Lincoln Laboratory in 1987, he worked for GE Fanuc Automation N.A., Inc., as a software engineer developing numerical and programmable controllers for manufacturing automation. Dick received a B.A. degree in biology from Wake Forest University in 1973, a Ph.D. degree in neuroscience from the University of Florida College of Medicine in 1979, and an M.S. degree in computer science from the University of Virginia in 1987. He was a National Science Foundation Predoctoral Fellow and a National Institute of Mental Health Postdoctoral Fellow.



JACQUES G. VERLY
received the Ingénieur Electronicien degree from the University of Liège, Belgium, in 1975. Through a sponsorship of the Belgian American Educational Foundation, he came to the United States that year and received an M.S. degree and a Ph.D. degree in electrical engineering from Stanford University in 1976 and 1980, respectively. At Stanford he performed doctoral research in the fields of image reconstruction and restoration. Upon graduation, he joined Lincoln Laboratory, where he has worked on (among other things) computer vision problems associated with laser radar data and, more recently, with fully polarimetric synthetic-aperture radar (SAR) imagery. He has over 40 publications in the areas of image reconstruction from projections (tomography), image processing, optical signal processing, distributed signal processing, mathematical morphology, computer vision, and automatic target recognition. Jacques is the coholder of two U.S. patents in the area of model-based object recognition, and is a fellow of the Belgian American Educational Foundation.

DAN E. DUDGEON
is a senior staff member in the Machine Intelligence Technology Group, where his focus of research has been in automatic target recognition. Before joining Lincoln Laboratory in 1979, he worked for Bolt, Beranek, and Newman, Inc., in Cambridge, Massachusetts. Dan received the following degrees from MIT: an S.B. in electrical science and engineering, and an S.M., an E.E., and a Sc.D. in signal processing. He was the corecipient of the 1976 IEEE Browder J. Thompson Prize, and is the coauthor of two books: *Multidimensional Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, New Jersey, 1984) and *Array Signal Processing* (Prentice-Hall, 1993). Because of his contribution in the field of multidimensional signal processing, Dan was elected an IEEE Fellow in 1987. He is also a fellow of the National Science Foundation.