An Overview of Automatic Target Recognition

Dan E. Dudgeon and Richard T. Lacoss

■ In this article we introduce the subject of automatic target recognition (ATR). Interest in ATR is increasing in the defense community as the need for precision strikes in limited warfare situations becomes an increasingly important part of our defense posture. We discuss reasons for the difficulty of the ATR problem and we survey the variety of approaches that try to solve the problem. We conclude by introducing the remaining articles in this special issue of the *Lincoln Laboratory Journal.*

AUTOMATIC TARGET RECOGNITION (ATR) generally refers to the use of computer processing to detect and recognize target signatures in sensor data. The sensor data are usually an image from a forward-looking infrared (FLIR) camera, a synthetic-aperture radar (SAR), a television camera, or a laser radar, although ATR techniques can be applied to non-imaging sensors as well. ATR has become increasingly important in modern defense strategy because it permits precision strikes against certain tactical targets with reduced risk and increased efficiency, while minimizing collateral damage to other objects. If computers can be made to detect and recognize targets automatically, the workload of a pilot can be reduced and the accuracy and efficiency of the pilot's weapons can be improved.

ATR technology can also be applied to non-military problems as well. For example, the problem of recognizing landmarks seen by a visual navigation system or a robotic system is related to the ATR problem. The recognition of particular objects or faces in photographs or video sequences is also related to ATR. We can think of the ATR problem as one part of the general problem of machine vision; namely, how can computers be made to do what we humans do so easily and naturally?

The fundamental problem of ATR is to detect and recognize objects of interest (targets) in an environment of clutter imaged by an imperfect sensor that introduces noise into the resulting signal. The definitions of target, noise, and clutter depend upon the application. Target categories can be coarse (e.g., a treaded ground vehicle) or fine (e.g., a specific type of tank or even a specific tank). Often the term *classification* is used for coarse categorization and the term *identification* is used for fine categorization, although they are also used synonymously with the term *recognition.* Unfortunately, usage is not consistent.

Clutter refers to real things that are imaged (buildings, cars, trucks, grass, trees, and other objects) but are not targets of interest. Sometimes a distinction is made between naturally occurring clutter (grass, trees, topographical features) and man-made cultural clutter (buildings, vehicles, and other works). Clutter tends to dominate the imagery simply because targets are generally sparse compared to the environment in which they operate. Noise refers to electronic noise in the sensor as well as inaccuracies introduced in the computations by a signal processor. Depending on the ATR application, the problem may be one of extracting a signal from noise or it may be one of separating a target from its surrounding clutter.

The distinction between detection and recognition is ill defined. We could argue that recognition is just the detection of a specific target type. But algorithms developed from this viewpoint tend to require prohibitive amounts of processing. For this reason, ATR systems generally include a front-end detection stage. The goal of the detection stage is to eliminate most of the sensor data from further consideration without

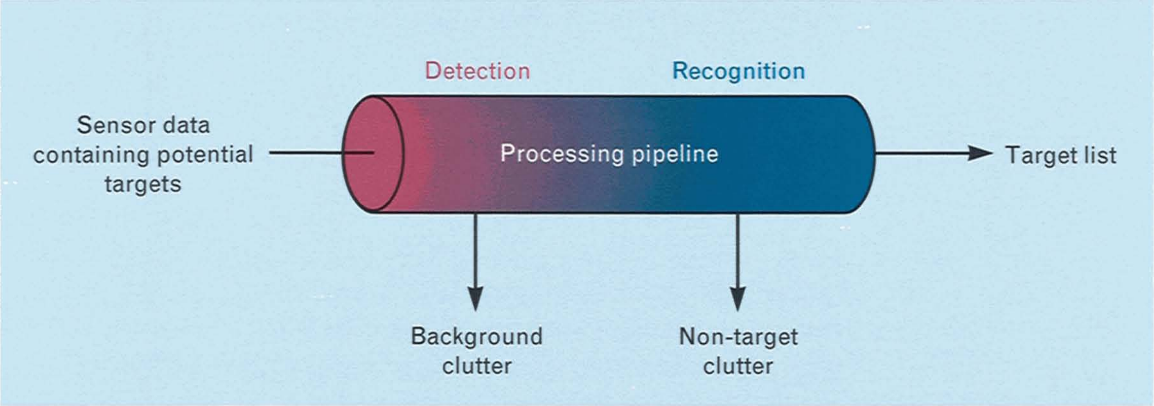


FIGURE 1. Conceptual data flow in automatic target recognition (ATR) systems. Simple detection algorithms are applied to all the sensor data to isolate small portions that might contain targets. More complex recognition algorithms then process the selected portions of the data to reject non-target clutter and classify targets. Ideally, all targets of interest pass through the pipeline and are included in the output target list.

eliminating any of the targets of interest. In this context, the term detection means that something interesting has been discovered, and this discovery requires further analysis. For example, a small cluster of bright pixels in an image could indicate the presence of an object. Computationally simple detection algorithms are required at this stage because all the sensor data in the input image must be examined.

Practical implementations of ATR systems can be viewed as pipeline processing systems, as illustrated in Figure 1. Ideally, all targets of interest pass through the pipeline and are included in the output target list. As data move through the pipeline, the processing algorithms become more target specific and computationally intensive for each data item, while the number of data items processed and the number of clutter false alarms each decreases. Even with this structure, the front-end detection stages of the processing pipeline often require the most computational power because the ATR system must search large amounts of imagery to find a few instances of the target.

For a specific ATR problem, both the target signatures and the clutter background can vary. In thermal images, for example, a tank can be hotter or colder than its background (causing positive or negative thermal contrast). It can also exhibit one shape when viewed from the side and another shape when viewed from the front; its turret can be rotated to any position and its gun barrel can take on a range of elevation angles. The background clutter can be a benign meadow, a treeline, a forest road, a featureless desert, or an urban area. Such complicated variations in both target signature and background clutter contribute significantly to the difficulty of the ATR problem.

ATR Technology

Many technologies and techniques are utilized in attempting to solve the problems of ATR, as illustrated in Figure 2. Sensor technology is critical because performance is ultimately limited by the quality of the information provided by the sensors. Processing hardware is also critical because ATR algorithms must process large quantities of data, often in real time, and because system development can be significantly hindered by the lack of processing capacity. Software, simulation, and evaluation methodologies and tools are also important elements. In addition, important but indirect contributions are made by neural and cognitive sciences, statistics, and sensor physics.

Figure 2 also identifies several algorithmic approaches to the problem of ATR; the multiplicity of these approaches indicates that no satisfactory single approach has yet been found. The most successful ATR systems will probably blend several algorithmic techniques to get satisfactory performance.

Detection Theory

ATR is based in part on detection theory and related statistical ideas that date back to the early days of radar processing. If a target signature is deterministic or if it is stochastic with well-defined stationary statistics, and if the clutter is well characterized as a stationary random process, then optimal detection techniques such as the matched filter can be derived. In classical detection-theory problems, a trade-off exists between the probability of detecting a target signature when the target is there, and the probability of declaring that a target is present when in fact it is not there (i.e., a false alarm). The objective of optimal detection processing is to separate the distributions of target signatures and clutter signatures so that they can be distinguished by a simple statistical test.

A seductive aspect of the detection-theory approach is that it provides a firm theoretical foundation for both the development of algorithms and the understanding of their performance. This approach to ATR, however, requires valid and analytically (or computationally) tractable statistical models for both targets and clutter. Such models are difficult to develop, and this approach is greatly complicated by signature and

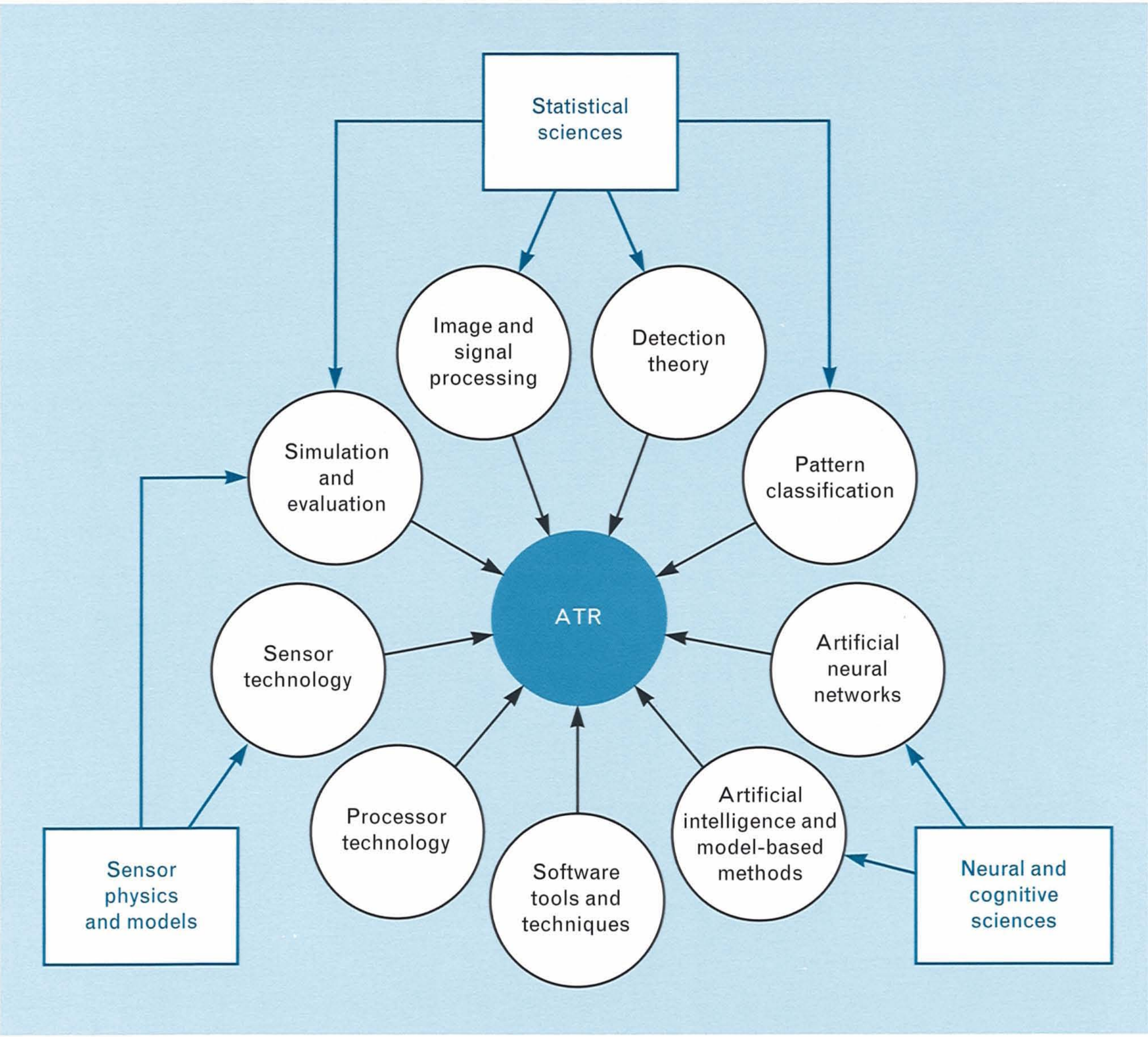


FIGURE 2. ATR technologies and processing methods. The circles identify some of the technologies and methods that are required to solve difficult ATR problems, although not all technologies and processing methods are used in all cases. The three boxes show some of the basic supporting research activities.

clutter variability. Detection theory is conceptually appealing, but it has had only limited success for ATR problems.

Pattern Recognition

Pattern recognition is the most mature approach used for ATR applications. Target-signature representation options range from two-dimensional image templates to lower-dimensional vectors of features that are designed to be differentially sensitive to targets and non-targets. Recognition depends on feature vectors (or templates thought of as vectors) for targets clustering together and being distinct from non-target clusters.

Potential targets (detections) are confirmed by comparing target images or feature vectors with a database of target and non-target exemplars. Recognition consists of selecting the best match between the target data and the exemplar database. The matching criteria may be ad hoc (e.g., mean-square differences between data and exemplar vectors), or they may be based on statistical assumptions that give the appearance of a more rational basis.

Pattern recognition is not disjoint from the detection-theory approaches discussed above or from the artificial neural network approaches discussed below. A close functional similarity exists between template (or feature vector) matching and matched filtering; differences between the two are at a more detailed level. Detection theory emphasizes optimal (or near-optimal) algorithms that are derived by using statistical models of raw data. Pattern recognition, which also relies heavily on statistics, includes more ad hoc approaches (e.g., spectral coefficients, fractal dimensions, and blob aspect ratios), especially in the definition and extraction of the features used to characterize targets.

Artificial Neural Networks

The neural network approach places even more emphasis upon experiential learning-by-example than does the pattern recognition approach. Neural networks represent a processing paradigm motivated by the human visual system, which remains the most flexible and robust target-recognition system for imagery that we know. The goal of neural network approaches is to develop an architecture that reproduces the flexibility and robustness of the human visual system in a piece of equipment.

In the nearer term, neural networks can have implementation advantages because they are highly parallel, and their ability to learn by example can make them capable of discovering and using signature differences that distinguish different types of targets. Neural networks are a type of nonlinear processing that could have advantages over classical detection-theoretic techniques or pattern-recognition techniques. One challenge for neural network approaches is to achieve good performance over the entire range of target signatures and background-clutter conditions, given limited training data.

Neural network learning can be categorized as supervised or unsupervised. Supervised networks are trained by using independently classified images or feature vectors. Training typically consists of computerized adjustment of parameters to optimize performance on the training set. Performance on new data depends on how well the training data are classified and how representative they are. Unsupervised networks define their own internal classifications of data, independent of the external source. Data are clustered together if they look similar to other members of the cluster. When new data are not similar enough to any existing cluster, a new cluster is formed.

Unsupervised networks, however, ultimately require supervised training to perform useful recognition. A target class must be associated with each internal cluster learned by the network. Typically, a cluster is assigned the target name corresponding to the majority (or even plurality) of training examples in the cluster. This training is often an iterative process in which parameters of the unsupervised network are adjusted by the ATR algorithm developer until each of its clusters tends to be dominated by only one target type. Multiple clusters of data can be assigned the same target type.

Model-Based Target Recognition

The genesis of model-based target-recognition methods is in artificial intelligence as it is applied to image understanding. Two characteristics typify the model-based approach: (1) matching of processed sensor data to predictions based on hypotheses concerning the target type, pose, and range; and (2) matching of processed sensor data on the basis of multiple localized features.

The difference between model-based target features and the feature vectors of other approaches is important. The elements of feature vectors are not features in the model-based sense. Corners, bright points, line segments, and small regions corresponding to identifiable parts of a target (e.g., the gun, turret, or body of a tank) might be features for a model-based system. Recognition depends on matching parts and interrelationships between the parts. Flexible and sophisticated matching techniques can be used, in principle, to design systems that are more robust with regard to obscuration and target variability. In contrast, the elements of feature vectors are usually global target measures such as the coefficients of a transform that represents the target image or a texture measure such as fractal dimension.

In some sense, every ATR algorithm is model based because every algorithm makes and uses *a priori* assumptions about target and clutter characteristics. The representation of the *a priori* information can vary widely, however, and it is conceptually different for model-based and non-model-based approaches. Generally, in model-based target recognition the appearance of the target in an image is modeled.

The search-and-match approaches of model-based methods also tends to distinguish them from other algorithmic approaches. The general paradigm is as follows: first form an initial set of hypotheses based on the sensor data (the *indexing* problem), then use the hypotheses to predict features and their relationships, and finally compare the predictions to features extracted from the data. This approach is quite different from conducting a computationally intensive exhaustive search for a best matching pattern.

Continuing areas of research in model-based target recognition include matching and evidence accumulation (i.e., what is the most robust way to match hypothesized signatures to the sensor data), indexing (how to generate a small number of hypotheses that contain the correct target), and modeling. In some systems the models are built by hand; that is, analysts examine sensor data, observe target signature characteristics, and encode them into some form of data representation that is convenient for matching. This approach is time consuming, and consequently there is a need to build models automatically from data. For example, the radar image-understanding system discussed in an earlier article in the *Lincoln Laboratory Journal* can build target models of reentry vehicles from radar data [1]. (Note that a system does not have to be a neural network to be capable of learning from data.)

Data Requirements for ATR System Development

ATR system development and evaluation requires an enormous quantity of data because of the variability in target signatures and background clutter. The assembly of large, realistic, experimental databases, however, is time consuming and expensive. As a result, we need to develop techniques that minimize data requirements or, to put it differently, utilize experimental data more effectively. Simulation is one such approach; limited experimental data can be used to develop and validate simulation models that can then be used to generate data for system development and evaluation.

If the statistics of a target signature can be modeled parametrically, then a small amount of experimental data might be used to determine the model parameters. Another approach is to develop physical models of image formation and use them to create synthetic target signatures in both real-clutter and synthetic-clutter backgrounds. This approach doesn't reduce the need for data; it simply allows us to create the data in a computer rather than collect real data with a real sensor observing real targets in real backgrounds in all their various states. Of course, the synthetic imagery must be realistic to be useful, and realism is determined by comparing the synthetic data to a real database. Fortunately, both model development and synthetic scene generation are done off line, so that real-time performance is not required. But the issue of how to develop and evaluate ATR systems in a reasonable amount of time with a reasonable amount of data is still open.

In This Issue

We are fortunate at Lincoln Laboratory to have several state-of-the-art sensors for collecting data that can be used to develop target detection and recognition systems. The in-house availability of these sensors is reflected in the articles included in this issue. The articles emphasize research that utilizes data from synthetic-aperture radars, laser radars, range-Doppler imaging radars, and air surveillance radars. Even when the work has not involved such sophisticated sensors, it is nevertheless experimentally oriented and it emphasizes the use of real sensor data. Realistic experimental data stimulate the work and provide a rich source of information for understanding and exploiting how targets and the clutter backgrounds in which they are embedded can be modeled and separated.

In this issue, we have collected nine other articles detailing various aspects of ATR. We begin by focusing on polarimetric synthetic-aperture radar imagery and the techniques used to detect, discriminate, and classify targets found in such imagery. The article by Leslie M. Novak et al. gives an overview of a classical pattern-recognition approach. The article by Daniel E. Kreithen et al. discusses the problem of discriminating targets from the natural-clutter objects (e.g., trees and bushes) that reflect enough radar energy to pass the detection stage. The article by Shawn M. Verbout et al. investigates some of the problems in recognizing three-dimensional targets from their two-dimensional SAR signatures, and develops a technique for incorporating three-dimensional information in the classification process.

The article by Allen M. Waxman et al. introduces the topic of neural network recognition systems. It discusses some of the motivation for neural network architectures that are derived from biological vision systems. It also presents the idea that information can be extracted from the changes in target signature from one viewpoint to the next as a target and a sensor move past each other. These neural-processing concepts are demonstrated on several different target-recognition problems.

The theme of neural networks continues in articles by Paul J. Kolodzy and by Murali M. Menon. The article by Kolodzy describes laser radar imagery, which simultaneously captures reflectivity information and range information. The article also addresses the problem of evaluating ATR algorithms. Because of the cost of collecting accurate training and test imagery spanning the breadth of possible ATR scenarios, the development of an electronic terrain board (which is a system that can generate synthetic imagery accurately and quickly) for testing and evaluating ATR algorithms may be more cost effective.

The article by Menon discusses the use of neural networks to implement an image-enhancement algorithm, based on Markov random fields, for laser radar imagery. The effect of this preprocessing step on the performance of the target classifier is demonstrated.

The article by Richard L. Delanoy et al. also demonstrates the use of laser radar measurements for target recognition, but here the authors use a model-based approach. Detection is done by several algorithms that operate in parallel and are keyed to look for various image features by using a process called *functional template correlation.* Multiple functional-template outputs are combined to form a simple interest image to focus subsequent processing. Target information is extracted and decomposed into features that are matched against stored appearance models. The weights used to govern the matching process can be adaptively learned on training data. The appearance models themselves can also be built semiautomatically from training data.

The article by Delanoy and Seth W. Troxel demonstrates the flexibility and power of functional template correlation. This article discusses how functional templates are used to detect and recognize gust fronts in weather radar data. The algorithm was recently implemented in real time and demonstrated on the ASR-9 surveillance radar system at Orlando International Airport in Florida.

The issue concludes with an article by Su May Hsu on research in ballistic missile defense. This article discusses the use of image processing techniques to extract target features for classification of reentering objects.

The Future of ATR

ATR continues to be an important defense technology. While great progress has been made in sensor technology and processing-hardware technology, the development of recognition algorithms is far from complete. Efforts are continuing to accelerate the development of these algorithms. These efforts are sponsored by ARPA, the Army Night Vision Electronic Sensors Directorate (formerly the Night Vision Lab), the Air Force Wright Laboratory, the Naval Air Warfare Center (China Lake), and other government laboratories. Common databases and evaluation techniques are being developed so that ATR approaches can be objectively compared.

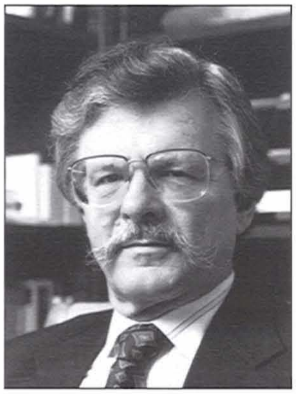
Because of the amount of data required to develop and test ATR systems for robustness, and the cost of mounting realistic data collection efforts, the development of accurate synthetic data-generation tools is also being addressed. Organizations such as the ATR Working Group (a joint industry-government working group) and the Department of Defense Working Group on ATR bring together government, industry, and academic laboratories to exchange views on the problems and potential solutions in ATR technology.

Just as the speech-recognition problem required decades of serious research before simple applications could be brought to market, the ATR problem needs a wide range of research before we can expect practical systems. Most people believe the ATR problem will not be solved by a single brilliant idea. The solution will probably require a combination of improved sensors, faster computers, and better algorithms. From our point of view, the area is extremely interesting because it involves so many technical disciplines (material science, sensor technology, signal and image processing, detection theory, algorithm development, computer technology, data representation, pattern classification, and neural networks) and because it has so many potential applications.

REFERENCES

1. A.M. Aull, R.A. Gabel, and T.J. Goblick, "Real-Time Radar Image Understanding: A Machine Intelligent Approach," *Linc. Lab. J.* 5, 195 (1992).

DAN E. DUDGEON
is a senior staff member in the Machine Intelligence Technology group. His research specialities are in image processing, computer vision, and computer architectures for algorithm development and real-time implementations. Before joining Lincoln Laboratory in 1979, he worked at Bolt, Beranek, and Newman, Inc. in Cambridge, Massachusetts. Dan received S.B., S.M., and E.E. degrees in electrical engineering, and a Sc.D. degree in signal processing, all from MIT. In 1976 he was corecipient of IEEE's Browder J. Thompson Memorial Prize for best paper by an author under the age of 30. He is the coauthor of *Multidimensional Digital Signal Processing* and *Array Signal Processing*. Dan was elected an IEEE Fellow in 1987, and he was named a Distinguished Lecturer of the ASSP Society in 1988. He is currently taking ice skating lessons in a futile attempt to keep pace with his daughter's athletic endeavors.



RICHARD T. LACOSS
is the leader of the Machine Intelligence Technology group. He received a B.A. degree and a B.S. degree in electrical engineering from Columbia University, and a Ph.D. degree in electrical engineering from the University of California at Berkeley. From 1965 to 1978 he worked in seismic discrimination and signal processing research at Lincoln Laboratory. His research now includes image understanding, expert systems, seismo-acoustic surveillance, neural networks, and parallel processing. He has taught in both the Earth and Planetary Sciences Department and the Electrical Engineering Department at MIT, and he is the author of many articles and reports in seismology and signal processing. His two-year old son and his three-year-old daughter also keep him very busy these days.