The Automatic Target-Recognition System in SAIP

Leslie M. Novak, Gregory J. Owirka, William S. Brower, and Alison L. Weaver

Lincoln Laboratory has developed a new automatic target recognition (ATR) system that provides significantly improved target-recognition performance compared with ATR systems that use conventional synthetic-aperture radar (SAR) image-processing techniques. We achieve significant improvement in target-recognition performance by using a new superresolution image-processing technique that enhances SAR image resolution and image quality prior to performing target recognition. A computationally efficient two-stage template-based classifier is used to perform the target-recognition function. This article quantifies the improvement in target-recognition performance achieved by using superresolution image processing in the new ATR system.

The goal of automatic target recognition (ATR) is to detect and recognize objects, such as tanks, in images produced by a laser radar, a synthetic-aperture radar (SAR), or an infrared or video camera. Current ATR methods are not fully automatic, relying instead on a partnership between computer processing and human analysis. Although computer algorithms accomplish much preliminary ATR work of sifting through millions of data pixels per image to identify candidate targets, image analysts play a crucial role in the recognition process by reviewing the ATR algorithm results. In a typical conflict situation, seven image analysts and one supervisor perform image interpretation for about thirty minutes to produce a target report that details the presence of missile launchers and the ground order of battle vehicles such as tanks, armored personnel carriers, and trucks.

The advent of low-cost, wide-area sensor platforms on unmanned air vehicles (UAV) will generate a profusion of data that dramatically increases the demands on image analysts [1]. For example, Teledyne Ryan Aeronautical's Global Hawk UAV is designed to provide sustained high-altitude surveillance and reconnaissance at large standoff ranges. Its mission goals include an operating range of 3000 nautical miles and the ability to loiter over the target area for 24 hours at altitudes of 65,000 feet. In addition to electro-optical and infrared sensors, this UAV will carry a SAR sensor that is projected to collect in one day enough data sampled at a resolution of 1.0 m × 1.0 m to cover 140,000 km^2 (roughly the size of North Korea).

To analyze the growing quantities of image data, the Defense Advanced Research Projects Agency (DARPA) has been developing a project called Semi-Automated IMINT (image intelligence) Processing, or SAIP [2]. This system, which currently processes SAR data, requires fewer image analysts to examine data in near real time. With SAIP, two image analysts and a supervisor can produce target reports within five minutes of receiving the data.

The SAIP system combines advanced ATR algorithms and robust false-alarm mitigation techniques with commercial off-the-shelf computer hardware to filter out natural and cultural clutter, and to recognize and prioritize potential targets. It annotates image areas containing potential targets with target cues for the image analysts, who use visualization tools provided by the SAIP human-computer interface to establish the true target types and produce a target report to pass to the battlefield commander. The image

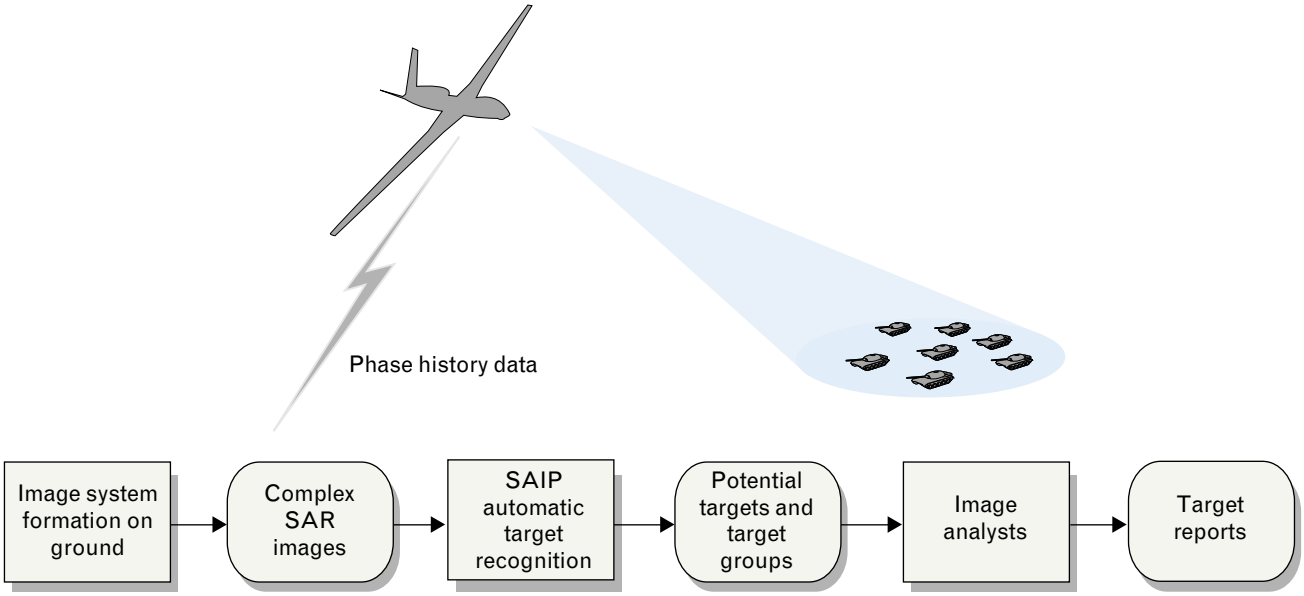

FIGURE 1. Process flow for phase history data gathered by an unmanned air vehicle. The phase history data gathered by the synthetic-aperture radar (SAR) sensor are sent to the Semi-Automated IMINT (image intelligence) Processing (SAIP) ground station, where the complex SAR image is formed and passed to the SAIP automatic target recognition (ATR) algorithms. These SAIP ATR algorithms cue the analysts to areas of the SAR image containing potential targets and groups of targets. SAIP integrates with existing image intelligence systems to provide improved wide-area surveillance and to reduce the work load of the image analysts.

analysts can also retask the SAR sensor for a high resolution (0.3 m × 0.3 m) revisit. Figure 1 shows the deployment of the SAIP system in a typical battlefield situation and the flow of data in the image-gathering process. The phase history data gathered by the SAR sensor are sent to the SAIP ground station, where the complex SAR image is formed and passed to the SAIP ATR algorithms. The ATR algorithms cue the analysts to areas of the SAR image containing potential targets and groups of targets.

Lincoln Laboratory has led the development of the human-computer interface and ATR components of the SAIP system. This article focuses on the classification components of the ATR that consist of a super-resolution processing technique known as high-definition vector imaging (HDVI) and mean squared error (MSE) classifier stages [3, 4].

The basic SAIP design combines existing Lincoln Laboratory ATR algorithms with robust false-alarm mitigation techniques such as terrain delimitation, object-level change detection, cultural-clutter identification, and spatial clustering, which includes force structure analysis. Figure 2 illustrates the target-recognition process implemented in the SAIP system. The Lincoln Laboratory ATR algorithms, highlighted in blue, consist of a constant false-alarm rate (CFAR) detector, a feature extractor, a discrimination-thresholding module, and the HDVI-MSE template-based classifier. Once the SAR image data have entered the SAIP system, the CFAR detector locates objects called candidate targets in the image having high radar contrast. A small area, called a chip, around each target candidate is extracted from the large SAR image. The candidate targets, consisting of true targets and clutter false alarms, are analyzed in the feature extractor, which calculates a set of features such as length, width, and diameter that characterize each object.

The feature extractor produces a scored measure of detection characterization, or target likeness, from the calculated feature values. The false-alarm mitigation modules—for example, terrain delimitation and object-level change detection—also provide target-likeness scores. These various scores are weighted and combined into an overall measure of target likeness, in the discrimination-thresholding module. The can-

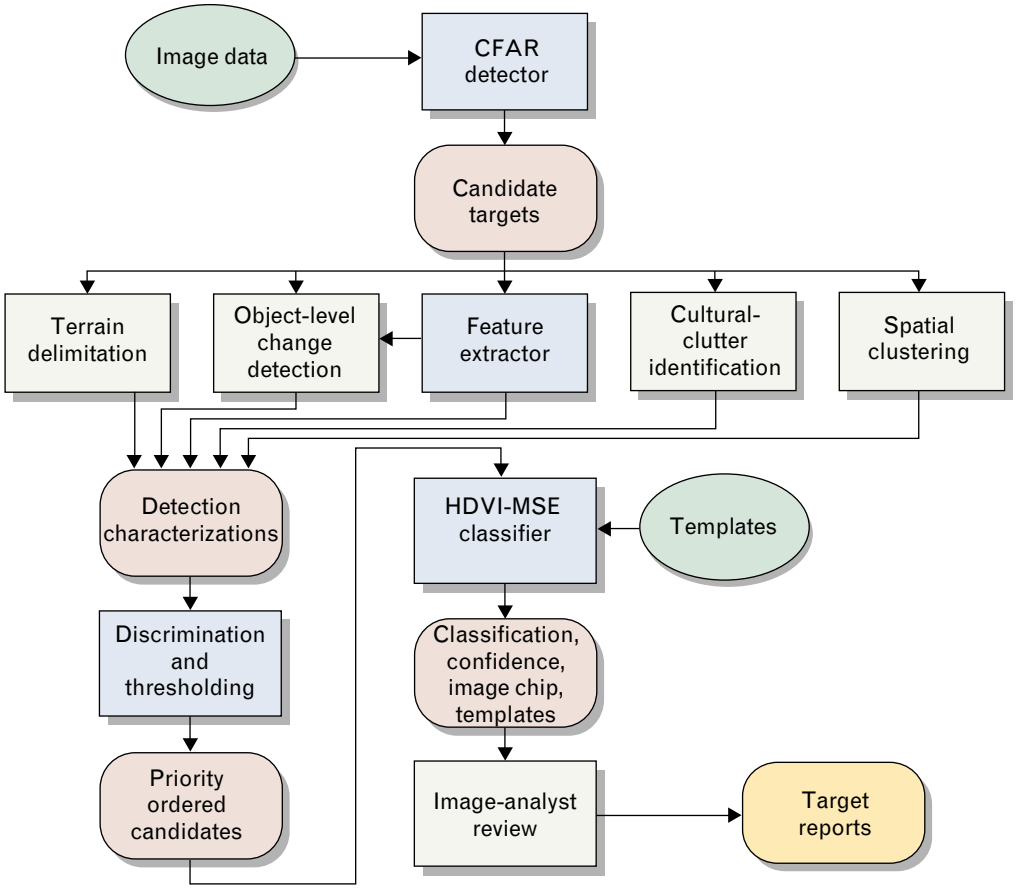

FIGURE 2. Process flow of SAIP system. The boxes highlighted in blue represent the Lincoln Laboratory ATR algorithms, which consist of a constant false-alarm rate (CFAR) detector, a feature extractor, a discrimination-thresholding module, and the high-definition vector imaging (HDVI) mean squared error (MSE) template-based classifier.

didate targets are then ranked according to their scores and appropriately thresholded and prioritized for further processing in the HDVI-MSE classifier.

Superresolution (HDVI) processing is applied to the complex SAR images of the priority-ordered candidate targets to enhance the resolution of the SAR data prior to performing classification. Each candidate target is compared to a set of stored reference templates. Each template is an average representation of a particular target type at a particular aspect angle. The MSE between the candidate target and each of the templates is calculated. For each target type, the minimum MSE score over aspect angle is converted to a confidence value for that target type. The template-based target classification and confidence values are reviewed by image analysts, who are also provided the HDVI-processed image of the candidate target, and the corresponding classifier templates that provided the best matches.

This article quantifies the improved target recognition that was achieved by using superresolution HDVI image processing prior to performing target recognition [5]. We introduce the two-stage HDVI-MSE classifier implemented in the SAIP system and show that this classifier implementation improves the computational speed by a factor of thirty, while suffering only a marginal decrease in performance compared with the classifier implementation that uses only HDVI-processed imagery. Finally, we present an end-to-end ATR system performance evaluation by using the SAIP detection, discrimination, and HDVI-MSE classifier stages.

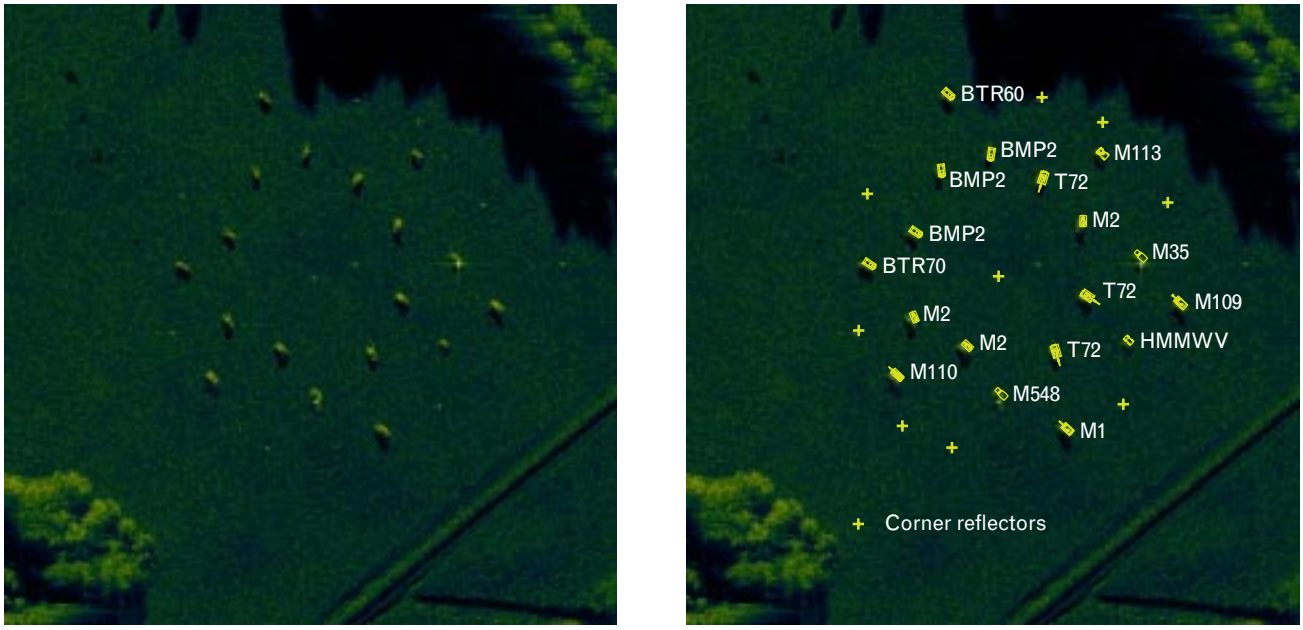

FIGURE 3. SAR image of the MSTAR target array (left) at Redstone Arsenal in Huntsville, Alabama, and with ground truth superimposed (right). The radar illumination is from the top; therefore, radar shadows appear below tall objects. A road can be seen in the lower right corner of the image. The plus signs represent corner reflectors used for image registration.

Data Description

The SAR imagery used in these studies was provided to Lincoln Laboratory by Wright Laboratories of Wright Patterson Air Force Base in Dayton, Ohio. These data were gathered in the fall of 1995 at the Redstone Arsenal in Huntsville, Alabama, by the Sandia X-band (9.6 GHz) HH-polarization SAR sensor, in support of the DARPA-sponsored Moving and Stationary Target Acquisition and Recognition (MSTAR) program [6]. The data comprise military targets imaged over 360° of aspect angle in spotlight mode, and approximately 30 km^2 of natural and cultural clutter gathered in stripmap mode. Figure 3 shows a typical SAR spotlight image (0.3-m × 0.3-m resolution) of the target array and a key that identifies each target. Figure 4 shows photos of the eighteen-target set, which includes three versions each of the BMP2 armored personnel carrier, the M2 infantry fighting vehicle, and the T72 main battle tank. The T72 tanks differ noticeably from each other—T72 #2 has fuel drums mounted on the rear of the target, while T72 #3 lacks skirts along the side of the target. The BMP2 and M2 infantry fighting vehicle have minor differences in target-to-target configuration.

We used images of ten targets to train the target classifier and then evaluated the ability of the classifier to recognize images of the eight remaining targets. Initially, we used non-HDVI-processed data in our evaluations to establish a baseline for comparing the classifier performance with HDVI-processed data. Then we evaluated the classifier performance with HDVI-processed data. Performance results are presented in terms of classifier confusion matrices that show the number of correct and incorrect classifications achieved. We also summarize the performance in terms of a probability of correct classification (P_{cc}), which is based on the total number of targets entered into the system, and includes the number of confuser vehicles correctly identified as unknown.

We trained the MSE classifier by constructing classifier templates with SAR images of ten of the targets shown in Figure 4: BMP2 #1, M2 #1, T72 #1, BTR60, BTR70, M1, M109, M110, M113, and M548. We constructed 72 templates per target that covered 360° of aspect per target. Initially, we tested the classifier on the training images to check the performance of the algorithm code, which we expected to be nearly perfect. The classifier was then tested

FIGURE 4. The eighteen-vehicle MSTAR tactical-target data set. There are three different vehicles for the BMP2, M2, and T72 targets, which allows for evaluation of serial-number variability. Note that T72 #2 has fuel barrels mounted on the rear. Ten vehicles in the data set were used for training the classifier, and the other eight vehicles were used to test the classifier. Of the eight test vehicles, the HMMWV and M35 were used as confuser vehicles. In our presentation of results, the confuser vehicles should be classified as unknown.

Table 1. Classifier Performance Results for Conventionally Processed 0.3-m × 0.3-m-Resolution Data

*Target**	*Number of Targets Classified As*										
	BMP2	BTR60	BTR70	M109	M110	M113	M1	M2	M548	T72	Unknown
BMP2 #2	**251**	—	—	—	—	—	—	1	—	—	3
BMP2 #3	**251**	—	—	—	—	—	—	1	—	2	2
M2 #2	—	—	—	—	—	—	—	**251**	—	—	4
M2 #3	—	—	—	—	—	—	—	**252**	—	—	3
T72 #2	—	—	—	—	—	—	—	—	—	**216**	39
T72 #3	—	—	—	—	—	—	—	—	—	**247**	4
HMMWV	12	6	—	—	—	37	—	2	—	—	**187**
M35	—	—	—	—	—	—	—	—	4	—	**251**

* Darker tone in table indicates independent test vehicles; HMMWV and M35 are confuser vehicles (not in training set).

with the remaining eight targets that were not used in training the classifier: BMP2 #2, BMP2 #3, M2 #2, M2 #3, T72 #2, T72 #3, HMMWV, and M35. Of these eight targets, the HMMWV and M35 were used as confuser vehicles, not included in the ten-target training set, and should be classified as unknown.

Performance Results for Non-HDVI-Processed Data

After training the MSE classifier, we evaluated its performance with conventional SAR image-formation processing and determined the SAR resolution required to achieve reliable P_{cc} for this target set. The SAR data were conventionally processed by using two-dimensional fast Fourier transform (FFT) processing with Hamming weighting to produce three nominal resolutions: 0.3 m × 0.3 m, 0.5 m × 0.5 m, and 1.0 m × 1.0 m. At each resolution, we implemented an MSE classifier by constructing 72 templates per target. We then tested each classifier by using target imagery at the appropriate resolution. The results of these evaluations are summarized in Tables 1, 2, and 3.

Table 1 is the classifier confusion matrix for the 0.3-m × 0.3-m-resolution data. When we tested the classifier by using independent test data, we achieved nearly perfect classifier performance. The P_{cc} for Table 1 data, calculated by using the independent test vehicles and the confuser vehicles, is 93.9%. Note, however, that the performance for T72 #2, which had fuel drums mounted on the rear of the tank, resulted in the classifier declaring 39 of the 255 images unknown. The performance for T72 #3, which did not have skirts along the sides of the target as did T72 #1 used in training, was nearly perfect: the classifier declared only 4 of the 251 images unknown, with the other 247 images classified as a T72 tank. The classifier rejected a large number of confuser vehicles: 438 of the 499 images. At this resolution, 37 of the 244 HMMWV test images were misclassified as an M113 armored personnel carrier, and 4 of the 255 M35 test images were misclassified as an M548 truck.

Table 2 is the classifier confusion matrix for the 0.5-m × 0.5-m-resolution data. The P_{cc} for these data, calculated by using only the independent test vehicles and the confuser vehicles, is 84.1%. At this resolution, the M35 truck was misclassified 13 out of 255 times. The HMMWV, however, was misclassified most of the time: only 61 out of 244 HMMWV images were declared unknown.

Table 3 shows the classifier confusion matrix for the 1.0-m × 1.0-m-resolution data. We observe at this

Table 2. MSE Classifier Performance Results for Conventionally Processed 0.5-m × 0.5-m-Resolution Data

*Target**	*Number of Targets Classified As*										
	BMP2	BTR60	BTR70	M109	M110	M113	M1	M2	M548	T72	Unknown
BMP2 #2	**235**	1	—	—	—	—	—	8	—	7	3
BMP2 #3	**237**	1	—	—	—	1	—	11	—	4	3
M2 #2	—	—	—	4	5	—	—	**233**	—	4	9
M2 #3	—	—	—	2	5	1	—	**239**	—	4	5
T72 #2	—	—	—	—	3	—	8	3	—	**217**	24
T72 #3	—	—	—	1	—	—	2	1	—	**241**	6
HMMWV	29	8	1	9	3	115	—	14	3	1	**61**
M35	1	—	—	3	3	1	—	—	5	—	**242**

* Darker tone in table indicates independent test vehicles; HMMWV and M35 are confuser vehicles (not in training set).

resolution a large degradation in classifier performance. The probability of correct classification P_{cc} degrades to 45.4% with poor confuser rejection. From these initial classifier studies, we concluded that SAR resolution of 0.5 m × 0.5 m or better is needed to achieve a reliable probability of correct classification for this target set. The probability of correct classification degrades rapidly for SAR resolutions coarser than 0.5 m × 0.5 m.

Performance Results for HDVI-Processed Data

Next we studied the ATR performance achieved by the MSE classifier with HDVI-processed SAR imagery. In these studies, the target data that were conven-

Table 3. Classifier Performance Results for Conventionally Processed 1.0-m × 1.0-m-Resolution Data

*Target**	*Number of Targets Classified As*										
	BMP2	BTR60	BTR70	M109	M110	M113	M1	M2	M548	T72	Unknown
BMP2 #2	**129**	15	21	8	1	39	3	15	—	18	6
BMP2 #3	**135**	7	10	5	2	43	—	25	—	18	11
M2 #2	14	3	2	37	14	36	5	**121**	—	16	7
M2 #3	21	6	—	18	6	47	9	**118**	—	27	4
T72 #2	7	12	3	19	9	23	17	32	1	**94**	38
T72 #3	19	8	6	23	5	12	14	19	—	**127**	18
HMMWV	48	7	4	4	—	145	—	13	1	—	**22**
M35	8	—	—	12	9	29	4	4	13	1	**175**

* Darker tone in table indicates independent test vehicles; HMMWV and M35 are confuser vehicles (not in training set).

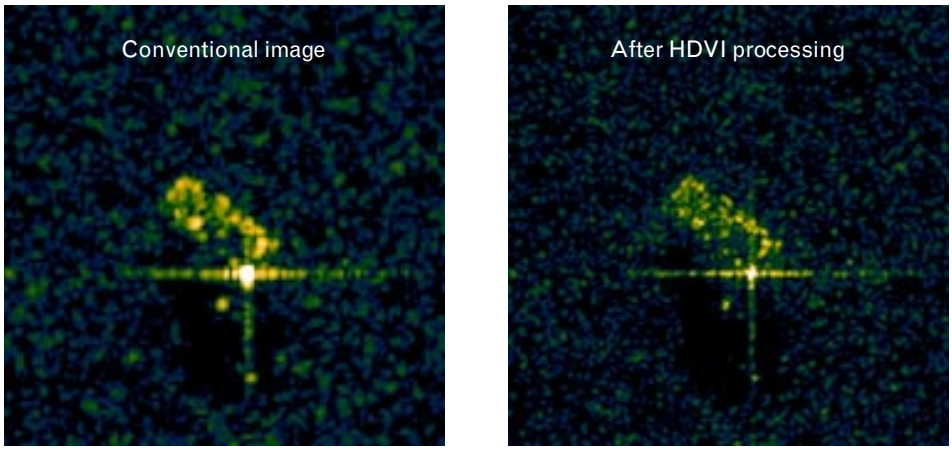

FIGURE 5. SAR images of an M35 truck. The left image has been processed with conventional two-dimensional fast Fourier transform (FFT) techniques and has 0.3-m × 0.3-m resolution. The right image has been HDVI processed to an approximate 0.15-m × 0.15-m resolution. The bright return from the lower right corner of the M35 truck is caused by the trihedral structure formed by the two walls and the floor of the truck bed. This bright return is sharper in the HDVI-processed image.

tionally processed into 0.3-m × 0.3-m and 1.0-m × 1.0-m-resolution SAR imagery were then reprocessed with superresolution (HDVI) processing. The resolutions achieved after HDVI processing were approximately 0.15 m × 0.15 m and 0.5 m × 0.5 m, respectively. Figure 5 compares side-by-side 0.3-m × 0.3-m-resolution images of an M35 truck, with and without HDVI superresolution processing. Figure 6 shows a similar image comparison for the conventionally processed 1.0-m × 1.0-m-resolution image versus the HDVI-processed image.

With HDVI-processed training imagery, we constructed classifier templates and tested each classifier by using the appropriate HDVI-processed test imagery. Table 4 shows the classifier confusion matrix for the HDVI-processed 0.3-m × 0.3-m-resolution data. Comparing the results of Table 4 with those of Table 1 shows somewhat improved MSE-classifier performance. The probability of correct classification with HDVI-processed data increased to 96.4%, an improvement of 2.5% over the conventionally processed data. With HDVI-processed 0.3-m × 0.3-m data, the

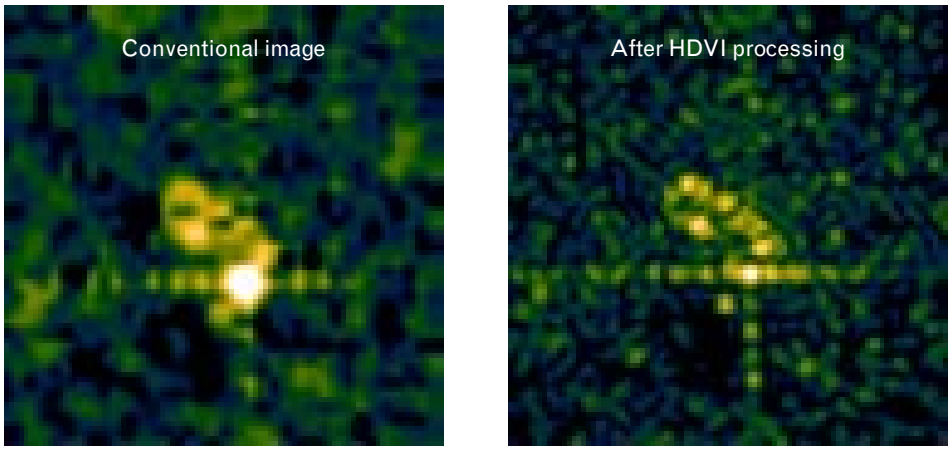

FIGURE 6. SAR images of an M35 truck. The left image has been processed with conventional two-dimensional FFT techniques and has 1.0-m × 1.0-m resolution. The right image has been HDVI processed to approximately 0.5-m × 0.5-m resolution.

Table 4. Classifier Performance Results for HDVI-Processed 0.3-m × 0.3-m-Resolution Data

*Target**	*Number of Targets Classified As*										
	BMP2	BTR60	BTR70	M109	M110	M113	M1	M2	M548	T72	Unknown
BMP2 #2	**252**	—	—	—	—	—	—	—	—	—	3
BMP2 #3	**255**	—	—	—	—	—	—	1	—	—	—
M2 #2	—	—	—	—	—	—	—	**251**	2	—	2
M2 #3	—	—	—	—	—	—	—	**248**	3	—	3
T72 #2	—	—	—	—	—	—	—	—	—	**232**	23
T72 #3	—	—	—	—	—	—	—	—	—	**245**	6
HMMWV	4	1	—	—	—	13	—	3	—	—	**223**
M35	—	—	—	—	—	—	—	1	6	—	**248**

* Darker tone in table indicates independent test vehicles; HMMWV and M35 are confuser vehicles (not in training set).

classifier rejected a larger number of confuser vehicles—471 out of the 499 images.

Table 5 shows the classifier confusion matrix for the HDVI-processed 1.0-m × 1.0-m-resolution data. After HDVI processing, the data resolution is approximately 0.5 m × 0.5 m. Comparing the results of Table 5 with the results of Table 3 shows a dramatic improvement in MSE-classifier performance. The probability of correct classification with HDVI-processed data increases significantly over that achieved with conventionally processed 1.0-m × 1.0-m-resolution data, from 45.4% to 73.4%. With HDVI-processed 1.0-m × 1.0-m data, the number of rejected confuser vehicles increases from 197 to 323 out of 499 images.

A Computationally Efficient Two-Stage Classifier

We developed a computationally efficient implementation of the HDVI-MSE classifier for the SAIP system to provide significantly increased speed in the ATR function with only a marginal loss in ATR performance. The HDVI-MSE classifier is preceded by a preclassifier stage that performs a coarse MSE classification on non-HDVI-processed 1.0-m × 1.0-m-resolution data. This reduced-resolution MSE preclassifier provides an estimate of the pose (aspect angle) of the target and an estimate of the target's true class. This information is passed to the more computationally intensive HDVI-MSE classifier and is used to limit the search space over target aspect and target type, which results in a more computationally efficient ATR algorithm [7].

Figure 7 compares the single-stage HDVI-MSE classifier with the computationally efficient two-stage approach. The figure shows the probability of correct classification results by using the 15° depression-angle SAR target data to test both classifiers. The classifier templates were constructed from 17° depression-angle target data. The upper diagram in Figure 7 shows the performance results for the HDVI-MSE classifier, which does a brute-force search over the ten-target template set at all possible pose angles. Also, x and y offsets of ±12 pixels are shifted over to account for small alignment errors between the test input chip and the templates. The probability of correct classification is 73.4% for the baseline single-stage HDVI-MSE classifier.

The lower diagram in Figure 7 shows the first-stage, ten-target MSE preclassifier that uses conventionally processed 1.0-m × 1.0-m-resolution data. This stage is used to determine the best five classifica-

Table 5. Classifier Performance Results for HDVI-Processed 1.0 m × 1.0 m-Resolution Data

*Target**	*Number of Targets Classified As*										
	BMP2	BTR60	BTR70	M109	M110	M113	M1	M2	M548	T72	Unknown
BMP2 #2	**202**	4	2	—	—	1	—	17	—	10	19
BMP2 #3	**195**	2	4	2	—	4	—	22	—	17	10
M2 #2	4	—	2	13	11	2	1	**188**	—	24	10
M2 #3	8	1	2	16	5	3	1	**180**	—	27	13
T72 #2	3	—	—	5	—	1	7	10	—	**190**	39
T72 #3	5	1	—	3	4	—	3	10	—	**211**	14
HMMWV	25	8	8	1	—	101	—	8	—	—	**93**
M35	1	—	—	3	1	6	—	—	14	—	**230**

* Darker tone in table indicates independent test vehicles; HMMWV and M35 are confuser vehicles (not in training set).

tion scores, a pose estimate for the top five classes, and a refinement of the x and y offsets for the top five classes. This information is then used to reduce the search space of the HDVI-MSE classification stage. The final classification decision is made with HDVI-processed data. The corresponding probability of correct classification is 70.0% for the two-stage classifier. Using the non-HDVI-MSE classifier on 1.0-m × 1.0-m-resolution data to prune the search space of the HDVI-MSE classifier introduces some small errors into the system that lead to a 3.4% loss in P_{cc}.

Figure 8 shows a cumulative error probability curve of the MSE-classifier pose-estimation error in degrees for 1.0-m × 1.0-m-resolution target data. Because each template represents 5° of aspect angle, a pose error of 20° angle corresponds to ±4 templates. The curve in Figure 8 indicates that approximately 95% of the time the correct pose is contained in the ±4 template search space. Note that a 180° ambiguity is included in the pose estimates because these targets are nearly symmetric when facing forward or backward. Therefore, a ±4 template pose estimate with the

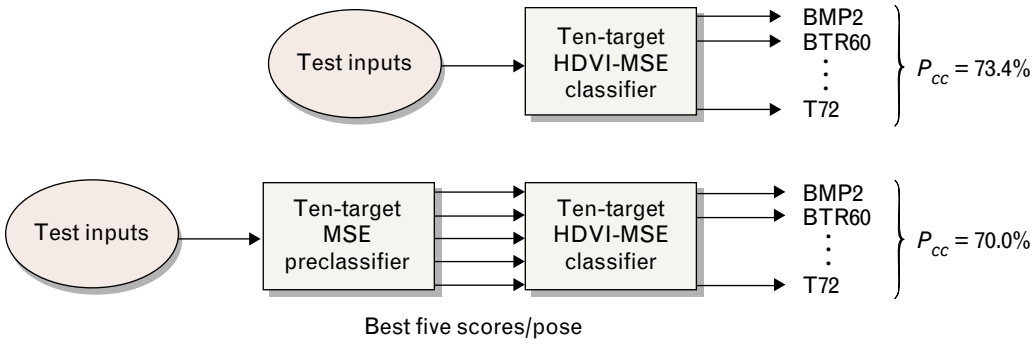

FIGURE 7. Two HDVI-MSE classifiers. The upper diagram shows the single-stage HDVI-MSE classifier that performs a brute-force search of all SAR image data. The lower diagram shows the reduced (1.0 m × 1.0 m) resolution MSE preclassifier that is used to prune the search space, followed by the enhanced (0.5 m × 0.5 m) resolution HDVI-MSE classifier that makes the final classification decision.

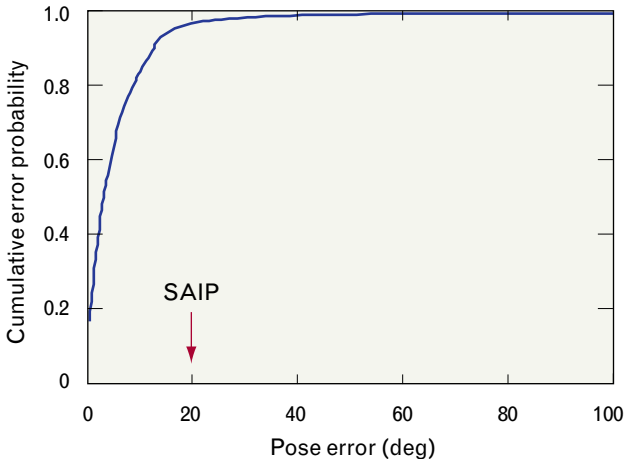

FIGURE 8. Cumulative error probability versus pose error with 1.0-m × 1.0-m-resolution target data. Because each template represents 5° of aspect, a pose error used by SAIP of 20° corresponds to ±4 templates. The correct pose is contained in the ±4 template search space approximately 95% of the time.

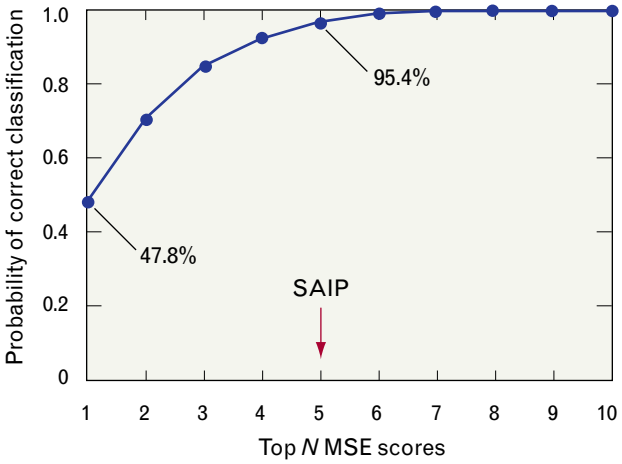

FIGURE 9. The probability that the correct target class is contained in the top *N* MSE scores with 1.0-m × 1.0-m-resolution target data. For this study, the correct target class was found in the top five SAIP-generated scores approximately 95.4% of the time.

180° ambiguity yields a total of eighteen templates to be searched at higher resolution. Thus the higher-resolution MSE classifier does not have to search all 72 templates per target; rather, it searches a much smaller subset of the HDVI template set.

Figure 9 presents a plot of the probability that the correct target class is contained in the top *N* MSE scores for 1.0-m × 1.0-m-resolution imagery. For this study, the top score gave the correct class only 47.8% of the time. The correct class, however, was contained in the top five scores approximately 95.4% of the time.

The curves in Figures 8 and 9 show why the 20° pose error angle and the top five scores from the preclassifier can be used to prune the HDVI-MSE classification space with only a small degradation in performance. From these evaluations, we also determined that by using the refinement of the *x* and *y* offsets from the preclassifier, we could reduce the shifting—moving the test input relative to the template to find the best match—in the HDVI-MSE classifier from ±12 to ±4 pixels without degrading performance

Table 6 shows the classifier confusion matrix for the two-stage classifier. For this study, the probability of correct classification is 70.0% for all eight targets, including the HMMWV and M35 confuser vehicles.

Computational Efficiency

The previous section compared the performance of an HDVI classification system and a two-stage HDVI-MSE classification system that uses the conventional imagery for preclassification, then applies HDVI processing followed by MSE on a reduced data set. The results showed that the two-stage classification system provided performance that was only slightly degraded (from 73.4% to 70%). We now quantify the benefits of the two-stage classification system from a standpoint of computational efficiency.

HDVI processing enhances image resolution by approximately a factor of two or more, depending on the quality of the initial imagery [5]. To exploit this resolution enhancement, we must sample the HDVI-processed imagery at a higher pixel rate than the original imagery. Our studies show that a factor of three in both range and cross-range pixel sample spacing provides good results. A conventional test-image chip of size $N \times N$ produces an HDVI image of size $3N \times 3N$ (each pixel in a conventional image is replaced with nine HDVI pixels). This factor of nine is present in both HDVI-processed test images and HDVI-MSE reference templates. Another factor of nine comes from the amount of spatial shifting required to properly align the test and template image.

Table 6. Multiresolution Classifier Performance for 1.0 m × 1.0 m-Resolution Data*

*Target***	*Number of Targets Classified As*										
	BMP2	BTR60	BTR70	M109	M110	M113	M1	M2	M548	T72	Unknown
BMP2 #2	**131**	2	8	—	—	4	—	16	—	21	14
BMP2 #3	**122**	3	3	1	1	2	—	25	—	22	17
M2 #2	—	—	—	13	10	2	1	**133**	—	16	21
M2 #3	4	1	—	9	6	2	—	**145**	—	15	14
T72 #2	3	1	—	1	1	1	5	7	—	**133**	43
T72 #3	—	1	1	2	2	—	1	9	—	**152**	23
HMMWV	15	9	7	3	—	71	—	5	—	—	**84**
M35	—	—	—	1	—	2	—	—	8	—	**185**

*17° depression angle for training data and 15° depression angle for test data.
** Darker tone in table indicates independent test vehicles; HMMWV and M35 are confuser vehicles (not in training set).

Consequently, the HDVI-MSE classifier requires eighty-one times more computations than the conventional 1.0-m × 1.0-m-resolution MSE classifier. Therefore, if T seconds are needed to classify an HDVI-processed test image, T/81 seconds are needed for preclassification with conventional imagery.

In the two-stage system, the final classification decision is made with the HDVI-MSE classifier, in a reduced search space. The computations required to perform the reduced-search-space HDVI-MSE classifier must be calculated and added to the preclassification computations to obtain the overall system computations. Again, we can calculate the computations required for the reduced-search-space HDVI-MSE classifier relative to the HDVI-MSE classifier alone. For the reduced-search-space classifier, the number of computations is 7290 (5 classes times 18 poses times 81 shifts). For the HDVI-MSE classifier, the number of classifications is 450,000 (10 classes times 72 poses times 625 shifts). Therefore, if T seconds are required to classify an HDVI-processed image with just the HDVI-MSE classifier, only about T/50 seconds are required to classify a reduced-search-space HDVI test input (450,000 divided by 7290 approximately equals 50).

By adding the computations required by the preclassifier and the reduced-search-space HDVI-MSE classifier, we find that this two-stage approach is approximately thirty times faster than the full-search-space HDVI-MSE classifier alone. This result is significant from a standpoint of real-time system implementation. We also note that the two-stage classification templates require only 11% more memory than the HDVI-MSE templates.

End-to-End Performance

Our system implementation consisted of a two-parameter CFAR detector, a discrimination stage with discrimination thresholding, and the two-stage template-based classifier. These ATR modules are highlighted in blue in the SAIP-system process diagram illustrated in Figure 2. The performance results are presented in the form of receiver operating characteristic (ROC) curves showing probability of detection versus false-alarm density (false alarms/km^2). The data sets used for these performance evaluations consisted of the 30-km^2 clutter set and the 1528 independent test target images. These data were processed to have 1.0-m × 1.0-m resolution as the input to the ATR system.

Detector

We used a two-parameter CFAR detector as a prescreener to eliminate large areas of clutter and to select objects that produce high-radar-contrast SAR returns. CFAR detections (bright pixels) that are spatially close were clustered into target-sized objects and centroided (the location of the target center was estimated by using a SAIP algorithm). Then a region of interest (ROI) chip was extracted around each candidate target. Figure 10 shows the ROC curve characterizing the CFAR detector (black curve). This curve characterizes the detection probability versus false-alarm density as the CFAR threshold is varied. We selected a CFAR threshold of 3.5, which resulted in a 0.95 detection probability and 30 false alarms/km^2 (905 false objects detected in the 30-km^2 clutter). This point on the CFAR ROC curve corresponding to the CFAR threshold of 3.5 was the operating point for the detector. All the detected ROI chips, including targets and false alarms, that exceed this threshold were then passed to the discrimination stage.

Discrimination

To study discrimination performance, we implemented two SAIP discrimination algorithms that measure different target characteristics. The target-likeness scores from these two discrimination algorithms were appropriately combined and thresholded in a simplified discrimination-thresholding stage.

One discrimination algorithm, denoted as the quadratic distance (QD) discriminator, calculates a set of features for each ROI chip entering the discrimination stage. Four features are calculated—object diameter, inertia, percent-bright CFAR, and fractal dimension [4]—and are placed into a four-dimensional feature vector. The distance between the feature vector and the target training data is the QD score, which provides a measure of target likeness. Figure 11 shows distributions of the QD score for the 905 false alarms and the 1451 detected targets out of 1528 test targets. If we select the QD threshold to be 13.1, the discrimination algorithm provides a 0.95 detection probability and the false-alarm density is reduced to 21.4 false alarms/km^2 (263 out of 905 false alarms were rejected).

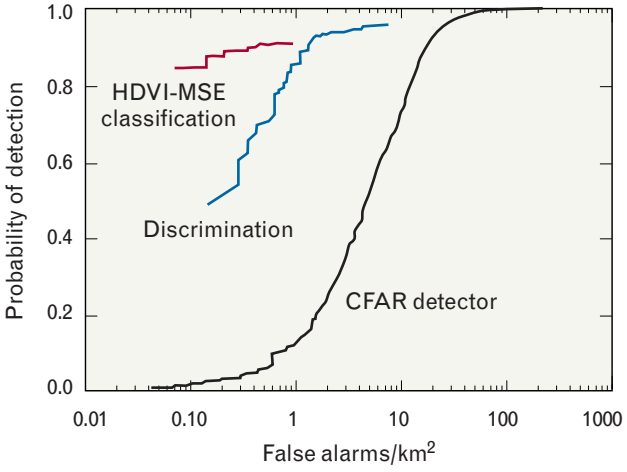

FIGURE 10. Receiver operating characteristic (ROC) performance curves for each element of the end-to-end ATR system. The HDVI-MSE classifier provided a probability of detection of greater than 0.8 with 0.1 false alarms/km^2.

The second discrimination algorithm is known as the quadratic polynomial discriminator (QPD) [8]. The QPD uses circularly symmetric kernels constructed from one-dimensional gamma functions to estimate local statistics in concentric bands around the center pixel of the ROI chip. First, the ROI chip is normalized by removing the local clutter mean. Next,

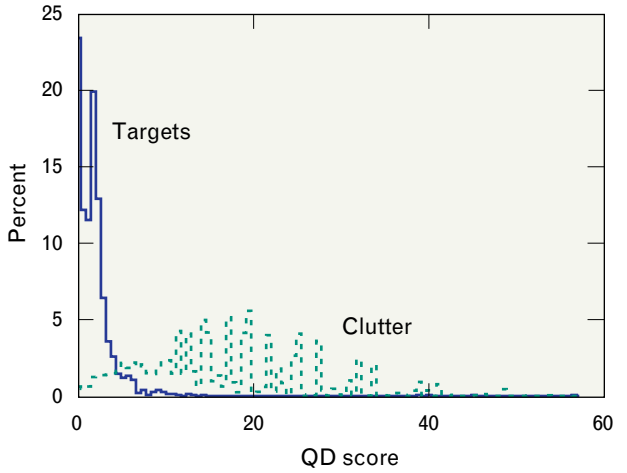

FIGURE 11. Distribution of quadratic distance (QD) scores for targets and clutter false alarms. The target distribution is to the left of the clutter distribution because the QD metric represents the distance measured to the target training set. A lower score indicates a target. A QD threshold of 13.1 provides a 0.95 detection probability with a false-alarm density of 21 false alarms/km^2. There is good separability between the two distributions, although some overlap remains.

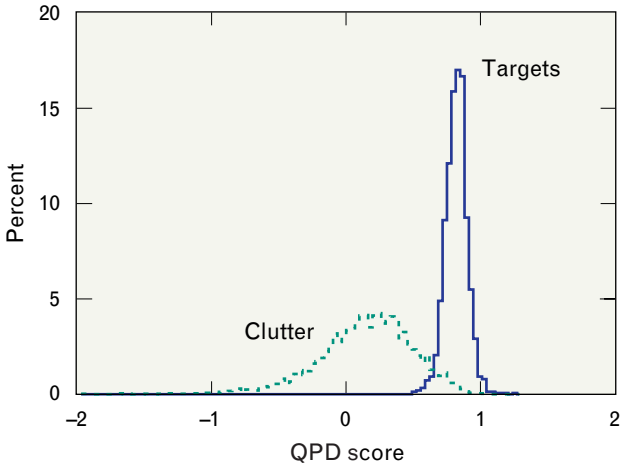

FIGURE 12. Distributions of quadratic polynomial discriminator (QPD) scores for targets and clutter false alarms. There is good separability between the two distributions, but some overlap remains. A QPD threshold of 0.26 would provide a 0.95 detection probability with a false-alarm density of 12 false alarms/km^2.

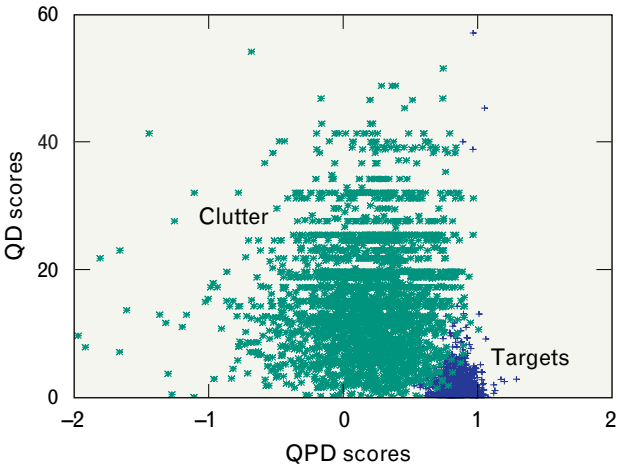

FIGURE 13. Two-dimensional scatter plots of the QD versus QPD scores for targets and clutter false alarms. This figure shows that a nonlinear threshold function can be used to provide better separability than either the QD score or the QPD score.

eight features are extracted by taking inner products of the image and the image squared with the kernels. Finally, the inner product of the feature vector and the QPD weight vector (calculated *a priori* by using training data of representative targets and clutter false alarms) is used to calculate the QPD score. Figure 12 shows distributions of the QPD score for the 905 false alarms and the 1451 detected targets. If we select the QPD threshold to have the value 0.26, this algorithm provides a 0.95 detection probability and the false-alarm density is reduced to 12 false alarms/km^2 (537 out of 905 false alarms were rejected).

The SAIP discrimination-thresholding module combines information from the other SAIP modules (including scores from the QD and QPD discriminators, terrain delimitation, and object-level change detection). In these studies we implemented a simplified discrimination-thresholding algorithm by using the QD and QPD scores only. The two-dimensional distributions of the QD and QPD scores were examined. Figure 13 shows these distributions for the detected targets and clutter false alarms, which demonstrate that the QD and QPD scores are essentially uncorrelated and therefore should provide improved discrimination performance when appropriately combined and thresholded. A simple quadratic classifier combined the QD and QPD scores into an overall discrimination statistic.

Figure 10 shows the resulting discrimination ROC curve (blue curve) with the combined QD and QPD scores. The selected operating point for the output of the simplified discrimination-thresholding stage is at the detection probability of 0.9, where the false-alarm density is reduced to approximately 1 false alarm/km^2 (875 out of 905 false alarms were rejected).

Classification

The two-stage template-based MSE classifier described previously was implemented as the final stage of the ATR system. This computationally efficient version of the HDVI-MSE classifier was shown to achieve a probability of correct classification of 70.0% against the eight targets, including the two confuser vehicles (HMMWV and M35). In these studies, we input the detected targets and clutter false alarms from the discrimination-thresholding stage and evaluated the overall end-to-end detection performance for the ATR system. The corresponding HDVI-MSE classifier ROC curve is shown in Figure 10 (red curve). At the detection probability of 0.8, the two-stage classifier rejected all but three clutter false alarms (resulting in a false-alarm density of approximately 0.1 false alarms/km^2).

Summary

This article examines the performance of a new ATR system developed at Lincoln Laboratory. The SAIP ATR system provides significantly improved target-recognition performance by using a new superresolution imaging technique known as HDVI processing prior to performing a template-based MSE classification. Our study was extensive, including a data set that covered approximately 30 km^2 of clutter and a significant number of tactical military targets (1528 images) to evaluate the performance of the system.

We first examined the performance of a ten-target template-based MSE classifier as a function of SAR image resolution. These initial studies used conventional two-dimensional FFT-processed SAR imagery. Performance of the classifier was evaluated by using 0.3-m × 0.3-m, 0.5-m × 0.5-m, and 1.0-m × 1.0-m-resolution imagery. These studies showed that a SAR image resolution of 0.5 m × 0.5 m or better is required to achieve reliable ATR performance (for the ten-target classifier we implemented). The P_{cc} degraded significantly at 1.0-m × 1.0-m resolution. For these three resolutions, the observed P_{cc} were 93.9%, 84.1%, and 45.4%, respectively.

We then demonstrated that when we applied HDVI superresolution processing to the 0.3-m × 0.3-m-resolution and 1.0-m × 1.0-m-resolution complex imagery, the observed P_{cc} were 96.4% and 73.4%, respectively. Also, the resolutions achieved after HDVI processing were approximately 0.15 m × 0.15 m and 0.5 m × 0.5 m, respectively. These initial studies established that 1.0-m × 1.0-m-resolution complex SAR imagery could be reliably superresolved to a resolution of 0.5 m × 0.5 m and that ATR performance was improved significantly.

In addition, a two-stage implementation of the template-based classifier was developed for the SAIP system. We showed that a lower-resolution 1.0-m × 1.0-m classifier can be used as an initial preclassifier stage, followed by an HDVI-MSE classifier. This implementation was shown to be computationally efficient, providing a significant increase in speed in performing the ATR function with only a marginal loss in performance. The correct classification performance for the ten-target classifier was reduced from 73.4% to 70.0% when the two-stage classifier was tested. The speed, however, was increased by a factor of thirty.

Finally, the end-to-end performance of the ATR system was evaluated by using a 1.0-m × 1.0-m-resolution data set including 30 km^2 clutter and 1528 test target images. The system implementation included only the CFAR detector, discriminator, and two-stage template-based classifier stages. We observed that at a false-alarm density of 0.1 false alarms/km^2, the system detection probability was better than 80% against the tactical military targets in the target set.

REFERENCES

1. W.P. Delaney, "The Changing World, the Changing Nature of Conflicts: A Critical Role for Military Radar," *1995 IEEE Natl. Radar Conf., Alexandria, Va., 8–11 May 1995,* pp. 11–15.
2. D.A. Fulghum, "DARPA Looks Anew at Hidden Targets," *Aviation Week Space Technol.,* 6 Jan. 1997, pp. 56–57.
3. G.R. Benitz, "High-Definition Vector Imaging for Synthetic Aperture Radar," *Proc. 31st Asilomar Conf. on Systems, Signals & Computers* **2,** *Pacific Grove, Calif., 2–5 Nov. 1997*, pp. 1204–1209.
4. L.M. Novak, S.D. Halversen, G.J. Owirka, and M. Hiett, "Effects of Polarization and Resolution on the Performance of a SAR ATR System," *Linc. Lab. J.* **8** (1), 1995, pp. 49–68.
5. L.M. Novak, G.R. Benitz, G.J. Owirka, and L.A. Bessette, "ATR Performance Using Enhanced Resolution SAR," *SPIE* **2757,** 1996, pp. 332–337.
6. E.R. Keydel and S.W. Lee, "MSTAR Extended Operating Conditions: A Tutorial," *SPIE* **2757,** 1996, pp. 228–242.
7. G.J. Owirka, A.L. Weaver, and L.M. Novak, "Performance of a Multi-Resolution Classifier Using Enhanced Resolution SAR Data," *SPIE* **3066,** 1997, pp. 90–100.
8. J.C. Principe, A. Radisavljevic, M. Kim, J. Fisher III, M. Hiett, and L.M. Novak, "Target Prescreening Based on 2D Gamma Kernels," *SPIE* **2487,** 1995, pp. 251–258.

LESLIE M. NOVAK
is a senior staff member in the Surveillance Systems group, where he develops target-detection, acquisition, and classification algorithms for millimeter-wave synthetic-aperture radar systems. He is also studying polarimetric radar signal processing algorithms and superresolution signal processing algorithms. Les joined Lincoln Laboratory in 1977, after working from 1972 to 1977 at Raytheon Company, Bedford, Massachusetts, where he helped design the digital signal processor for the Patriot system and develop algorithms for Raytheon's pulse Doppler map-matching system. Before that, he developed extended Kalman tracking-filter algorithms for the TPQ-36 and TPQ-37 artillery and mortar-locating radar systems for Hughes Aircraft Company, Fullerton, California. From 1961 to 1968, he performed analytical studies of radar systems for Autonetics in Anaheim, California.

He holds a B.S.E.E. degree from Fairleigh Dickinson University, an M.S.E.E. degree from the University of Southern California, and a Ph.D. degree in electrical engineering from the University of California, Los Angeles.

GREGORY J. OWIRKA
is an associate staff member in the Surveillance Systems group, developing algorithms for automatic target recognition with synthetic-aperture data. He joined Lincoln Laboratory in 1987, after receiving a B.S. degree in applied mathematics from Southeastern Massachusetts University. Greg is pursuing an M.S. degree in electrical engineering at Northeastern University.

WILLIAM S. BROWER
is a staff member in the Surveillance Systems group, developing algorithms for automatic target recognition with synthetic-aperture radar and inverse synthetic-aperture radar imagery. He joined Lincoln Laboratory in 1996 after receiving a Ph.D. degree in experimental particle physics at the University of California, San Diego. He also holds a B.S. degree in physics from the Georgia Institute of Technology.

ALISON L. WEAVER
is an assistant staff member in the Surveillance Systems group, programming algorithms for target detection, discrimination, and classification. She joined Lincoln Laboratory in 1984, after receiving a B.A. degree in mathematics from Mount Holyoke College.