

ARMY RESEARCH LABORATORY



Synthetic FLIR Signatures for Training and Testing Target Identification Classifiers

by Bruce A. Weber and Joseph A. Penn

ARL-TR-3451

March 2005

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Citation of manufacturers' or trade names does not constitute an official endorsement or approval of the use thereof.

DESTRUCTION NOTICE—Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Adelphi, MD 20783-1145

ARL-TR-3451

March 2005

**Synthetic FLIR Signatures for Training and Testing
Target Identification Classifiers**

Bruce A. Weber and Joseph A. Penn
Sensors and Electron Devices Directorate, ARL

Approved for public release; distribution unlimited.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) March 2005		2. REPORT TYPE Final		3. DATES COVERED (From - To) 3-01-06/03	
4. TITLE AND SUBTITLE Synthetic FLIR Signatures for Training and Testing Target Identification Classifiers			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Bruce A. Weber and Joseph A. Penn			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory Sensors & Electron Devices Directorate (ATTN: AMSRD-ARL-SE-SE) bweber@arl.army.mil Adelphi, MD 20783-1145			8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-3451		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) ARL 2800 Powder Mill Road Adelphi, MD 20783-1145			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT We performed a series of experiments to benchmark the performance of a target identification classifier trained on synthetic FLIR target signatures. Results show that the classifier, when trained on synthetic target signatures and tested on measured, real world, target signatures, can perform as well as when trained on measured target signatures alone. It is also shown that when trained on a combined database of measured plus synthetic target signatures, performance exceeds that when trained on either database alone. Finally, it is shown that within a large, diverse, database of signatures there exists a subset of signatures whose trained classifier performance can exceed that achieved using the whole database. These results suggest that for classification applications synthetic FLIR data can be used when enough measured data is unavailable or cannot be obtained due to expense or unavailability of targets, sensors, or site access.					
15. SUBJECT TERMS ATR, classifier, target identification, synthetic signatures, infrared, FLIR					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UNCLASSIFIED	18. NUMBER OF PAGES 28	19a. NAME OF RESPONSIBLE PERSON Bruce A. Weber
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED			19b. TELEPHONE NUMBER (Include area code) (301) 394-2501

Standard Form 298 (Rev. 8/98)

Contents

1. Introduction	1
2. Database Generation	2
2.1 Database Emulation.....	2
2.2 Synthetic-Signature Database Generation.....	3
2.3 Sequestered Test Database Generation	6
3. Experimental Procedures	6
3.1 Classifier Training, Testing, And Scoring.....	6
3.2 Scenario Relevant Classifiers.....	7
3.3 Description Of Experiments.....	7
3.3.1 Introduction	7
3.3.2 Experiment 1	8
3.3.3 Experiment 2	9
3.3.4 Experiment 3	9
3.3.5 Experiment 4	9
3.3.6 Experiment 5	11
3.3.7 Experiment 6	11
4. Results	13
4.1 Introduction	13
4.2 Results From Experiments Using Randomly Chosen Datasets.....	13
4.3 Results From Experiments Using Datasets Selected By Special Purpose Classifiers ..	16
5. Conclusions	18
Acknowledgments	19
References	20
Distribution list	21

Figures

Figure 1. This flow chart shows the iterative process for generating, training and testing, and validating the synthetic signature database. The white modules show the path for FLIR signature database generation. The gray modules show two different paths for classifier training, and one for testing and validation. The black paths show how a model can be modified after validation.....	3
Figure 2. Synthetic FLIR signatures: left to right: HMMWV, M60, T72, and M113. Top to bottom: three groups of these four targets are shown with increasing amounts of optical blur and line noise. Within each group of 16 images, from top to bottom, is shown simultaneous views of the front, right side, back, and left side. Gray scale normalization is applied to each view separately. Some variation in target centering is also seen. All such signatures are included in the synthetic database of 90,432 signatures.....	4
Figure 3. Visual comparison of real (top) and synthetic (bottom) signatures of a T72 tank for every 5-degrees of aspect starting with a front view at the top left and rotating clockwise as seen from above. Each group of signatures contains 8 rows and 9 columns.	5
Figure 4. Experiment 1: Training and testing of the SIG signature database. Full size databases are designated by curved modules, derived, subset databases by rectangular modules, and signature processing is designated by six sided modules. Classifier performance is designated by diamond shaped modules.	9
Figure 5. Experiment 2: Training and testing of the 5,156 SIG signature database augmented with synthetic signatures.	10
Figure 6. Experiment 3: Training and testing of the 90,432 synthetic signature database.	10
Figure 7. Experiment 4: Training and testing of the 90,432 synthetic signature database augmented with increasing amounts of signatures from the 5,156 SIG signature database.	11
Figure 8. Experiment 5: Training and testing of selected SIG-like and ROI-like signatures. The combined training database included approximately 8,500 SIG-like and 7,250 ROI-like synthetic signatures.....	12
Figure 9. Experiment 6: Training and testing of classifier trained on the combined SIG and ROI-TRAINING databases. The performance of this classifier benchmarks all previous experiments.	12
Figure 10. Performance comparison of K-Means classifier tested on 2,125 ROI sequestered signatures, as a function of the normalized number of training signatures used. From the bottom up, four curves represent results from individual and combined training databases: synthetic signatures alone, SIG alone, synthetic augmented with SIG, and SIG augmented with synthetic.....	14
Figure 11. This figure shows the number of codevectors as a function of the number of synthetic signatures used for training. The database included 22,608 target chips for each of four targets or 90,432 target chips in all.	15

Tables

Table 1. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on a database of the union of the ROI-LIKE and SIG-LIKE synthetic subsets. Overall probability for correct identification is 81%. 17

Table 2. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on a database of SIG images. Overall probability for correct identification is 82%. 17

Table 3. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on the union of SIG, SIG-LIKE and ROI-LIKE synthetic datasets. Overall probability for correct identification is 85%. 17

Table 4. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on a database of the union of SIG and ROI-TRAINING datasets. Overall probability for correct identification is 88%. 17

This page is intentionally left blank.

1. Introduction

Robust identification of targets in clutter is a difficult problem. This is particularly evident when classifiers trained on one set of data are tested on an unrelated set. This problem is due to the relationship of data to the limited conditions under which data is collected, and to the statistical nature of classifiers in representing data. To investigate this relationship and improve performance such classifiers need training data that is scenario relevant. But obtaining new signatures from field-testing, representative of new scenarios or new sensor systems, is both expensive and impractical. This paper will examine the potential to use synthetic data to supplement or replace measured data for the training of target classifiers.

The synthetic Forward Looking Infrared (FLIR) target signatures developed for this study are an improvement to those we provided a number of years ago for another comparison of real and synthetic data [1]. That study compared the use of synthetic, or measured, FLIR scenery with targets, as test data for an Automatic Target Recognizer (ATR) trained on measured data alone. Metrics were then determined to characterize ATR performance and signature differences. Additional target detection research, using synthetic FLIR imagery and human observers, examined target search prediction [2], and range performance prediction [3].

Our study tests an algorithm and not humans, ignores the detection phase of an ATR, and instead examines whether the classifier stage, trained on synthetic signatures alone and tested on measured signatures alone, can perform as well as when trained on measured signatures alone.

Our method is to train one classifier on synthetic signatures, another on measured signatures, test each trained classifier on a sequestered set of measured signatures, and then compare the results. Comparable performance validates the synthetic signatures for target identification applications and provides a tool to explore improving classifier performance. This validation differs from a one-to-one comparison of temperatures observed at specific locations on a target, and substitutes a weaker statistical measure focused on application needs.

The process of creating a useful synthetic database begins by creating signatures designed to approximate a real-world scenario, measuring the comparative performance of a classifier trained on the database, and then adjusting the database by adding new data in response to performance deficiencies. This improves performance until classifier state representations stabilize. A detailed description of this methodology is found in the section on database generation. This is followed by a description of the application of

such signatures to training a classifier, and a demonstration of enhanced classifier performance trained on a subset of synthetic signatures, selected from a larger database.

2. Database Generation

2.1 Database Emulation

For guidance in the synthesis of synthetic FLIR signatures we tried to emulate the kind of data available in the measured, real-world signatures of the COMANCHE FLIR database. This included the targets, the locations, and the environmental conditions as specified in the COMANCHE database ground truth.

The COMANCHE FLIR database consists of approximately 30,000 image scenes (720×480 pixels) containing 22,000, approximately target centered, signature instances (75×40 pixels) containing roughly equal numbers of 10 target types. Signatures were collected for three geographical locations, Yuma, Arizona, Hunter-Liggett, California, and Grayling, Michigan, two seasons including one summer and one spring or winter month, and a full diurnal cycle. Although ground-truth information did not specify vehicle exercise states, anecdotal information suggested that states varied from stopped with engine off, to stationary idling, to moving at low velocity. After extraction from the scenes, the 22,000-signature chips were separated into two databases designated SIG and ROI. The SIG database, of targets in uncluttered environments, included all 10 target types viewed from 72 evenly spaced rotational aspects spanning 0 to 355 degrees, while the ROI database of targets near clutter, included only 5 of the 10 target types viewed from eight, approximately evenly spaced, rotational aspects from 0 to 315 degrees. The ROI target signatures are more difficult to recognize.

For these experiments we modeled 4 of the 10 SIG or 3 of the 5 ROI target types. The 4 ground targets included the HMMWV, M60, T72, and the M113. Only the M113 was not part of the ROI database. This selection provided 5,156 signatures of 4 targets from the SIG database and 4,249 signatures of 3 targets from the ROI database. We chose these four targets because we wanted two tank, and two non-tank vehicles, one with tracks and one with wheels. We modeled each target in similar conditions and locations, and simulated realistic exercise routines to produce thermal signatures consistent with observed data. Nevertheless, specific synthetic signatures were not designed to match specific measured FLIR signatures since the measured signatures were uncalibrated and specific vehicle exercise information prior to data collection was unknown.

2.2 Synthetic-Signature Database Generation

Isothermal target model surface elements were calculated using the PRISM [4] commercial code, and IR-signatures were rendered using the U. S. Army Research Laboratory's (ARL) CREATION code [5]. Figure 1 shows a schematic of the methodology [6] for signature database generation, classifier training and testing, and model adjustment. The BRL-CAD/FRED/PRISM path (white modules) provides thermal prediction from thin, surface modeled, polygonal nodes. CREATION then images the distribution of thermal nodes from any location in the 3-dimensional space surrounding the object, adding wavelength and optical blurring sensor effects, atmospheric propagation loss, and noise. By comparing the performance and codevectors of synthetic, and measured data trained classifiers, the synthetic database was expanded by creating additional signatures representing missing exercise states, and model structural variations.

Figure 2 shows an example of the diversity of synthetic signatures. Four targets and four aspects are shown. From left to right: HMMWV, M60, T72, and M113. Top to bottom: three groups of these four targets are shown with increasing amounts of optical blur and line noise. Within each group of 16 images, from top to bottom, is shown simultaneous views of the front, right side, back, and left side. Gray scale normalization is applied to each signature separately. Some random variation in target centering, similar to that in the measured database, was also modeled. Any and all such signatures are included in the 90,432 synthetic FLIR database.

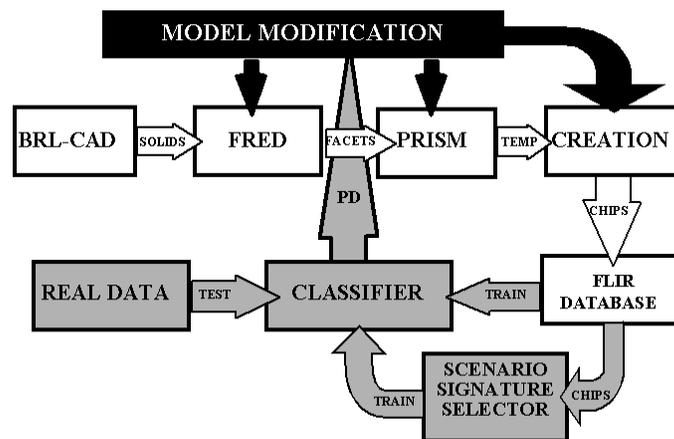


Figure 1. This flow chart shows the iterative process for generating, training and testing, and validating the synthetic signature database. The white modules show the path for FLIR signature database generation. The gray modules show two different paths for classifier training, and one for testing and validation. The black paths show how a model can be modified after validation.



Figure 2. Synthetic FLIR signatures: left to right: HMMWV, M60, T72, and M113. Top to bottom: three groups of these four targets are shown with increasing amounts of optical blur and line noise. Within each group of 16 images, from top to bottom, is shown simultaneous views of the front, right side, back, and left side. Gray scale normalization is applied to each view separately. Some variation in target centering is also seen. All such signatures are included in the synthetic database of 90,432 signatures.

Figure 3 shows a comparison of real (top) and synthetic (bottom) signatures of a T72 for every 5-degrees of aspect starting with a front view at the top left and rotating clockwise as seen from above. Each group of signatures contains 8 rows and 9 columns. Though for the signatures in Figure 3, we tried to model a specific COMANCHE signature sequence, in general we did not due to the lack of specific target temperature and vehicle exercise information. Some clutter, but no noise was added to this sequence so as to focus signatures on the target.



Figure 3. Visual comparison of real (top) and synthetic (bottom) signatures of a T72 tank for every 5-degrees of aspect starting with a front view at the top left and rotating clockwise as seen from above. Each group of signatures contains 8 rows and 9 columns.

2.3 Sequestered Test Database Generation

The sequestered database, which was used for all testing, was created by dividing the ROI database in half. Each database contained all target aspects, but not equal numbers of each aspect. One half designated the ROI-TESTING database, consisted of 2,125-signatures that were sequestered and used solely for testing. The other half, designated the ROI-TRAINING database, consisted of 2,124 signatures and was used for synthetic data subset selection and classifier baseline performance evaluation.

3. Experimental Procedures

3.1 Classifier Training, Testing, And Scoring

Figure 1 shows the procedure (gray modules) for training and testing the classifier. Two potential methods for classifier training are indicated. One method used the entire synthetic database and the other used a scenario dependent subset. These options will be discussed in more detail in the section describing the experiments.

The K-Means classifier, used in this research, was developed at ARL and is described as a minimized mean-squared-error (MSE) encoder [7]. All input target chips, both in the training and testing phase, were intensity scaled to zero mean and unity variance. The chips are 75 by 40 pixels in size and contain between 1,000 and 2,000 pixels on target, depending on target size, aspect angle, and range to target.

In the training phase a target size section was extracted from the center of the chip by a series of aspect dependent windows. It was then enlarged to a fixed size, and wavelet decomposed into four sub-bands. Training commenced by collecting like-aspect sub-band images from the same target, clustering these using the K-Means algorithm, and then averaging the sum to create codevectors. Codevectors were collected into a codebook representing the various signatures of a given target at a given pose. This process was repeated for all of the training chips representing each target thus creating a library of target and aspect specific codebooks.

For the experiments described below, we grouped all horizontally viewed aspects into 4 aspect dependent windows at 0, 90, 180, and 270 degrees respectively. For 4 targets this yields 64 codebooks (4 targets x 4 aspect windows x 4 wavelet sub-bands), with each codebook containing tens of codevectors representing the trained classifier states.

In the testing phase, an unknown target chip was similarly extracted, sized, and wavelet decomposed, and then compared with each of the codevectors of each of the codebooks for each of the learned targets. The target and aspect dependent, sub-band codevector with the lowest Mean-Square-Error (MSE) was declared the identity.

In scoring the algorithmic declarations, three measures were available: correct target identification, correct target aspect identification, and confidence. The confidence, given by a numerical value on the 0 to 1 interval, was the difference of the normalized inverse MSE values for the two highest value identities: the larger the difference the higher the confidence. This value could thus be used to threshold acceptable identifications depending on the application of the results. For the classification experiments described in this paper, correct identification alone was used to score declarations, and no threshold applied.

3.2 Scenario Relevant Classifiers

Classifiers are developed for many types of applications. Some require robust performance against a broad range of scenarios. Others, at the opposite end of the application spectrum, require near optimal performance against a narrowly defined scenario. Since classifier applications are often known in advance it is reasonable to use the knowledge of an expected scenario to help improve performance. In this section we describe a method that uses a classifier to select scenario relevant signatures from a large database, and then uses the selected signatures to train a scenario relevant classifier. The method is not intended to be optimal but rather simply as a proof of principle.

The methodology is shown schematically in Figure 1. The gray module, labeled SCENARIO SIGNATURE SELECTOR, is a filter, or series of filters, each consisting of a classifier trained on a scenario-labeled database of measured-data. Signatures correctly identified by this scenario-specific classifier were scored for database inclusion using a set of three criteria, one of which was determined by experimentation. The three criteria required that the targets be correctly identified with respect to target-ID, target-aspect, and that the confidence for correct ID should be 0.9 (on the 0 to 1 interval) or greater. The confidence threshold was determined by experimentation. Using these three scoring criteria is more restrictive than that used for scoring classifier performance described earlier. Experiment 5, described below, is an example of how this is done and shows that subset selection can significantly improve classifier performance.

3.3 Description Of Experiments

3.3.1 Introduction

In order to evaluate the usefulness of our synthetic signature database for target identification, we undertook a series of experiments that compared the performance of the K-Means classifier trained on databases containing mixtures of measured and synthetic signatures. In each experiment, the classifier performance was measured using the sequestered ROI-TESTING database.

Six separate database combinations were used to train the classifier. The first four include: 1) the entire SIG signature database alone, 2) the SIG database augmented with

increasing amounts of synthetic signatures, 3) the entire synthetic signature database alone, and 4) the entire synthetic database augmented with increasing amounts of SIG signatures. The fifth database combined two subsets of synthetic signatures that were selected to be similar to the signatures in the SIG and ROI-TRAINING databases respectively. The sixth database combined the entire SIG and the entire ROI-TRAINING databases, and was used to establish a baseline for classifier performance against the sequestered, ROI-TESTING, test set.

3.3.2 Experiment 1

The purpose of the first experiment was to demonstrate the performance of the classifier when trained on measured signatures taken from a different, but similar, data collection to that of the test data.

In this experiment, shown schematically in Figure 4, the K-Means classifier was trained on SIG signatures alone. Choosing a random set of 1.25% of the 5,156-SIG-signature database (requiring only that equal percentages of each of the four target types be represented) the classifier was trained, and then tested on the sequestered test database. This process was repeated seven times, using different percentages, creating seven individual training databases having 1.25, 2.5, 5, 10, 25, 50, and 100% of the signatures in the SIG database respectively. For each iteration additional signatures were selected randomly from the diminishing pool of SIG signatures and then added to the existing training database. Thus each successive, larger, database always included all of the signatures previously chosen plus an approximately equal number of new signatures. With each new, incrementally larger database the classifier was retrained from scratch using a list of signatures ordered by the sequence in which the increments were added. This ordering, which was the same for each training epoch, tended to bias initial codevector choices toward signatures that appeared earlier in the training process.

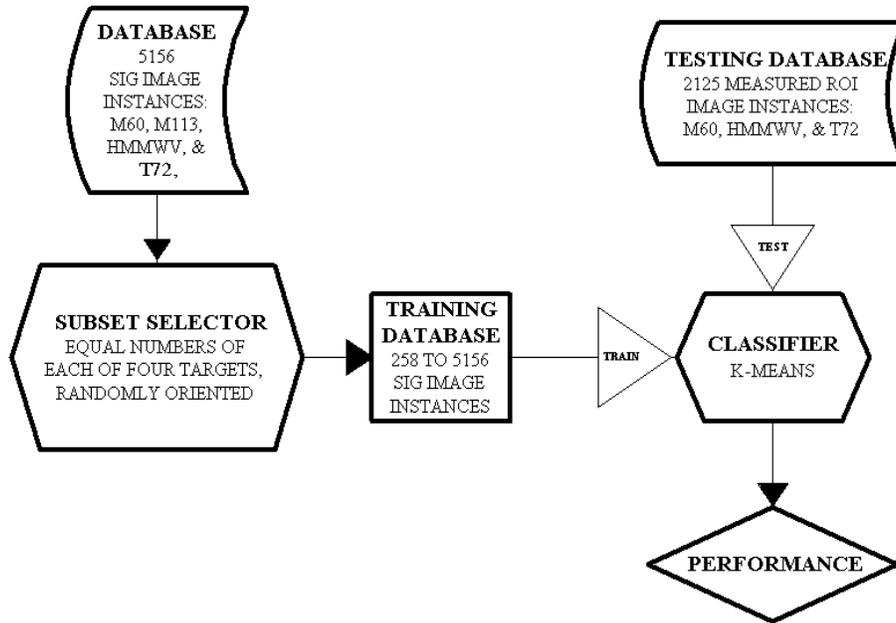


Figure 4. Experiment 1: Training and testing of the SIG signature database. Full size databases are designated by curved modules, derived, subset databases by rectangular modules, and signature processing is designated by six sided modules. Classifier performance is designated by diamond shaped modules.

3.3.3 Experiment 2

This experiment demonstrated how the performance of the classifier changed as the measured signature training set was combined with increasing amounts of synthetic signatures. For this experiment, shown schematically in Figure 5, the classifier was trained on all 5,156-SIG signatures combined with increasing amounts of synthetic signatures and then tested as in the first experiment. This process yielded seven new training databases, each containing all of the 5,156-SIG-signatures, plus 1.25, 2.5, 5, 10, 25, 50 and 100% of the 90,432-synthetic-signatures respectively. When fully combined the initial 5,156-signature database had increased in size to 95,588-signatures.

3.3.4 Experiment 3

This experiment demonstrated the performance of the classifier when trained on synthetic signatures alone. In this experiment, shown schematically in Figures 6, six databases of increasing size (2.5, 5, 10, 25, 50 and 100% of the 90,432-synthetic-signatures) were created containing synthetic signatures alone. A classifier was then trained on each of these databases and then tested as in the previous experiments.

3.3.5 Experiment 4

The purpose of this experiment was to demonstrate how the performance of the classifier changed as the synthetic signature training set was combined with increasing amounts of measured signatures. For this experiment, shown in Figure 7, the entire 90,432-synthetic

database was combined with increasing numbers of SIG signatures and then tested as in the previous experiments. This experiment is similar to that of the second experiment except that the roles of the SIG and synthetic databases are reversed. Again seven databases are created in successively increasing size starting at 90,432-signatures and ending with 95,588-signatures total.

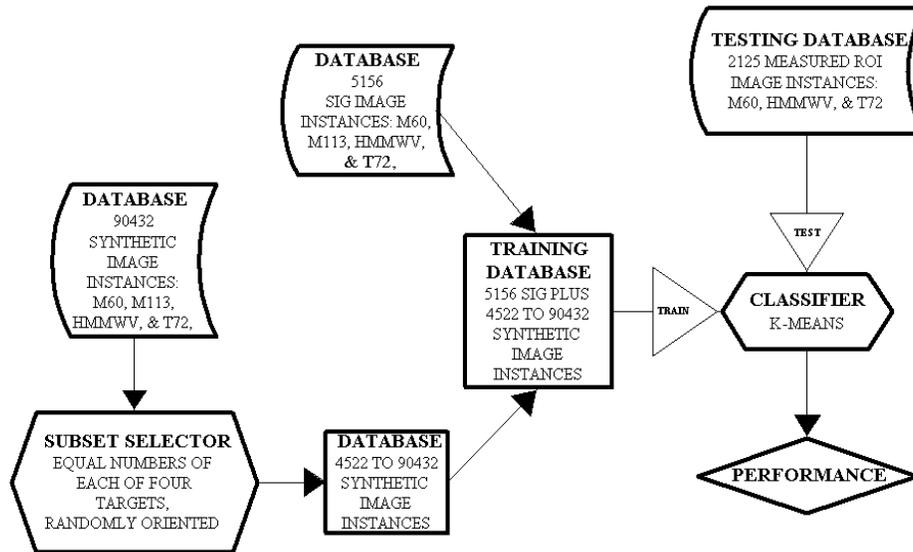


Figure 5. Experiment 2: Training and testing of the 5,156 SIG signature database augmented with synthetic signatures.

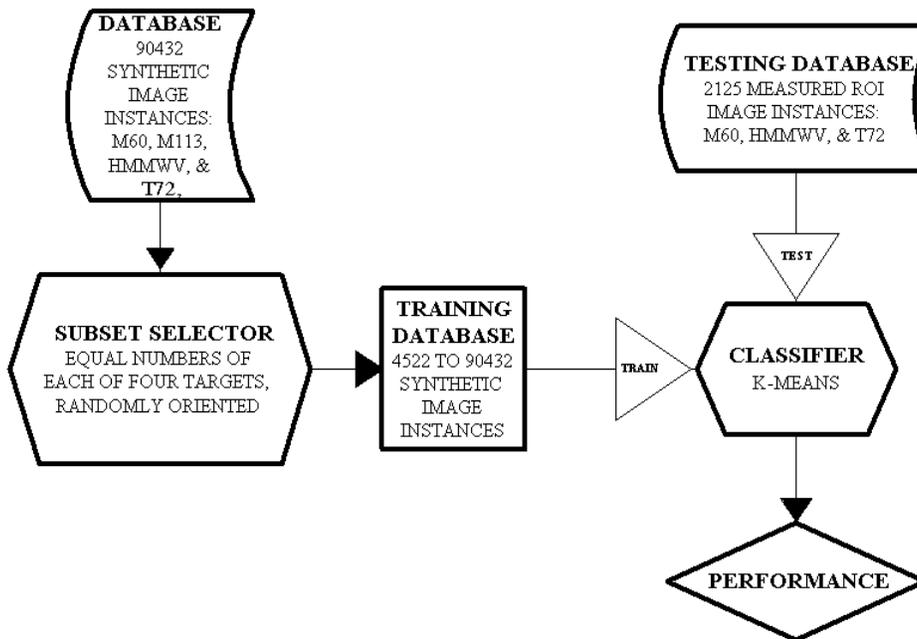


Figure 6. Experiment 3: Training and testing of the 90,432 synthetic signature database.

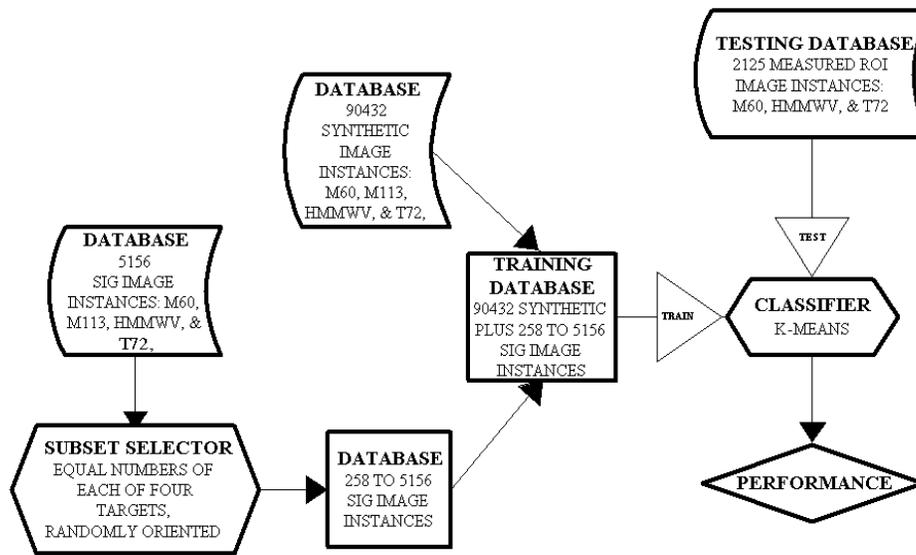


Figure 7. Experiment 4: Training and testing of the 90,432 synthetic signature database augmented with increasing amounts of signatures from the 5,156 SIG signature database.

3.3.6 Experiment 5

This experiment demonstrated that the performance of a classifier can be improved by training it on a subset of signatures selected from a large database, keeping signatures that are similar to those in the test set, and discarding signatures that are dissimilar. For this experiment, shown in Figure 8, two mutually exclusive subsets are selected from the 90,432-database of synthetic signatures, added together, and then used to train and test the classifier. The databases included: a SIG-LIKE database of 8,450-synthetic signatures and a ROI-LIKE database of 7,257-signatures.

The SIG-LIKE database was produced by training a classifier on the 5,156-signature SIG database and then using the trained classifier to identify similar signatures from the synthetic signature database. Correctly identified signatures were then selected for inclusion in the SIG-LIKE database if they met the three criteria mentioned earlier (correct identification, correct aspect, and confidence greater than or equal to 0.9). Similarly, the ROI-LIKE database was produced, as above, by first training a classifier on the 2,124-signature ROI-TRAINING database. Signatures common to both databases were arbitrarily removed from the SIG-LIKE database but retained in the ROI-LIKE database. We named the combined, 15,707 synthetic signature, database the ROI-LIKE_SIG-LIKE database.

3.3.7 Experiment 6

This experiment provided a baseline for classifier performance against which, the performance of other classifiers could be compared. For this experiment, shown in Figure 9, the SIG and ROI-TRAINING databases are combined to train the classifier. The performance of this classifier established a baseline since the ROI-TESTING database was the other, mutually exclusive, half of the original ROI database.

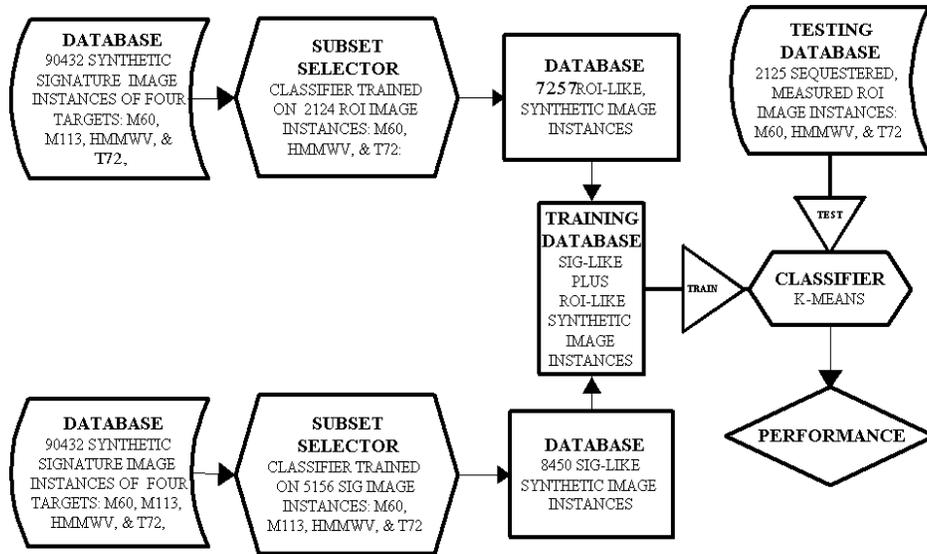


Figure 8. Experiment 5: Training and testing of selected SIG-like and ROI-like signatures. The combined training database included approximately 8,500 SIG-like and 7,250 ROI-like synthetic signatures.

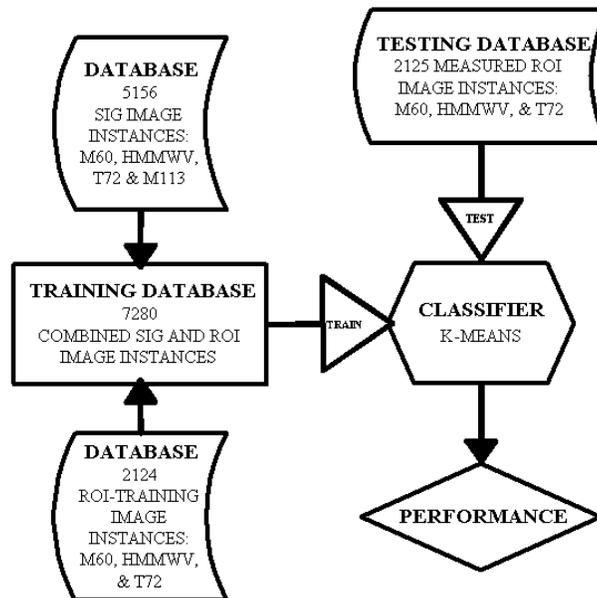


Figure 9. Experiment 6: Training and testing of classifier trained on the combined SIG and ROI-TRAINING databases. The performance of this classifier benchmarks all previous experiments.

4. Results

4.1 Introduction

The two sections that follow describe the results from the six experiments described in the previous section. All trained classifiers were tested on the sequestered, 3-target ROI-TESTING dataset alone. Classifier performance is specified as the probability (in percent) for correct identification. For overall performance against all targets a single number is specified; for performance against individual targets the Confusion Matrix tabular form is used. The Confusion Matrix table lists the Predicted ID down the left column and Actual ID along the top row. Correct identification probabilities are shown along the diagonal. Off diagonal elements represent the probability of misidentification. If a target is part of a training set but not part of the testing set no probability value can be assigned under that targets column. Such entries are indicated by dashes. For example: for performance in which Target 4 of the training set was not part of the test set, no performance value can be listed in the Target 4 column of the Confusion Matrix since Target 4 can not be the correct identification. Target 4 could, however, be incorrectly declared as one of the other three targets by placing a value in one of the first three spaces in row 4.

For the four-target datasets we investigated, targets 1 and 4, and targets 2 and 3 were of similar size, whereas targets 1 and 4 were smaller than targets 2 and 3. Target 1 was a HMMWV, Target 2 was an M60, Target 3 was a T72, and Target 4 was an M113. Target 4 was not part of the ROI-TESTING test set.

All training was accomplished using a K-Means classifier. The K-Means parameters were chosen to provide enough codevectors to produce good target identification performance.

4.2 Results From Experiments Using Randomly Chosen Datasets

Experiments 1 through 4 examined the effect on performance by the random selection of training data. Figure 10 shows four curves that summarize these experiments. The performance is specified as the probability for correct target identification (PD), as a function of the percentage of the total number of training signatures used. A PD of 1.0 corresponds to all targets being identified correctly.

In Figure 10 the curve designated by circles shows the performance of the SIG trained classifier as a function of the percentage of the 5,156-signatures used. PD saturation is not observed since the PD never levels off. This indicates that the SIG database contains mostly unique signatures with little redundancy. The maximum PD achieved was 82%.

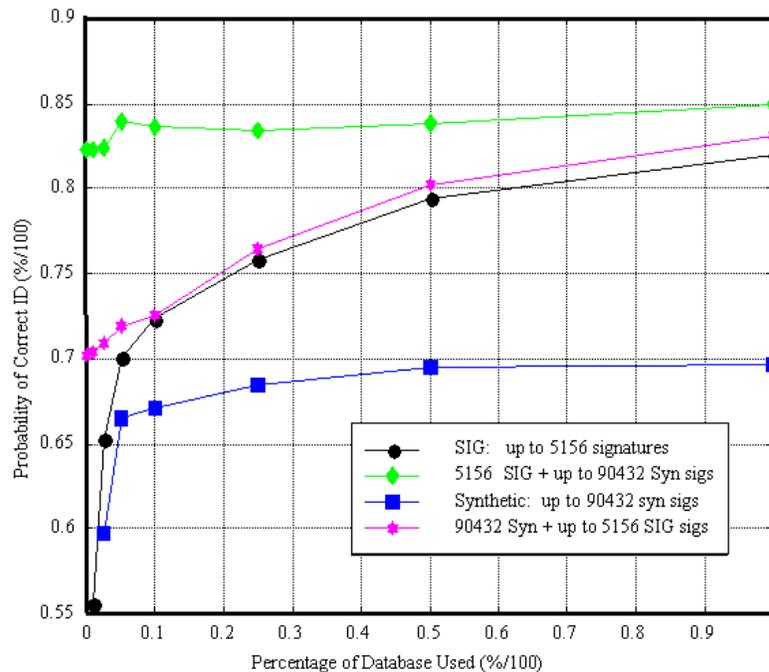


Figure 10. Performance comparison of K-Means classifier tested on 2,125 ROI sequestered signatures, as a function of the normalized number of training signatures used. From the bottom up, four curves represent results from individual and combined training databases: synthetic signatures alone, SIG alone, synthetic augmented with SIG, and SIG augmented with synthetic.

The curve designated by squares shows the performance of the synthetic data trained classifier as a function of the percentage of the 90,432-signatures used. For this case, PD saturation is observed: rolling over for a training set of approximately 9,000 signatures, and flattening out above 45,000 signatures. Two possibilities may have contributed to the PD saturation: either the number of codevectors, created in the training process, was not increasing, or the additional codevectors were not sufficiently different to improve performance. Figure 11 shows the increase in the number of codevectors as the number of files per target was increased. Clearly no saturation is observed in the number of codevectors being created. This suggests that the added codevectors were insufficiently different, and that a significant portion of the synthetic data is similar. The maximum PD achieved was 70%. This level of performance was 12% less than that produced by training the classifier on the SIG database alone.

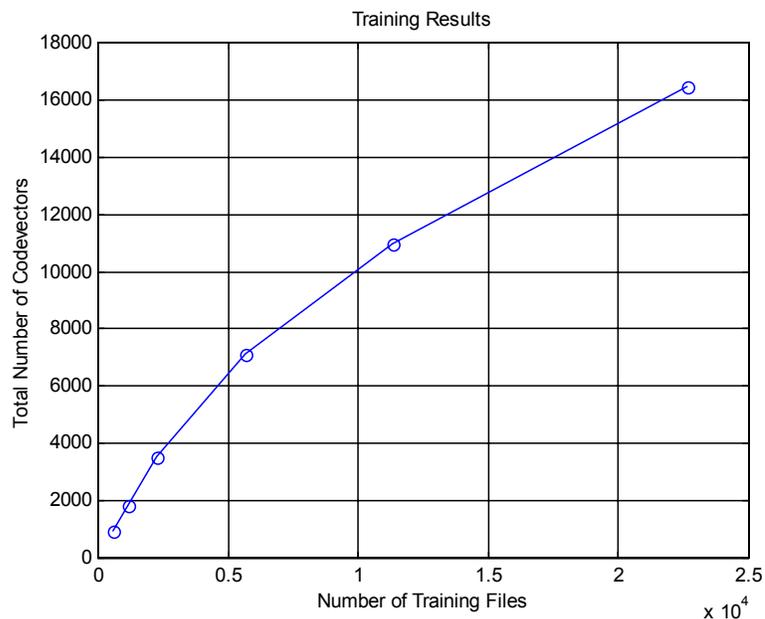


Figure 11. This figure shows the number of codevectors as a function of the number of synthetic signatures used for training. The database included 22,608 target chips for each of four targets or 90,432 target chips in all.

The curve designated by diamonds shows the improvement in PD as increasing percentages of the 90,432-signature synthetic database are added to the 5,156-signature SIG database. The combined performance was shown to increase 3% to a PD of 85% over using SIG signatures alone. This 3% increase is significant since the statistical uncertainty in recognizing files from the 2,245-signature ROI database is about 2%. The 1% difference in performance, for the leftmost point of the diamonds-curve compared to the rightmost point of the circles-curve (both for classifiers trained on the entire SIG dataset), is believed to be due to differences in the ordering of signatures in the training lists, and the bias, mentioned earlier, of codevector choice toward signatures that appear early in the training cycle. Though both training lists contained all of the SIG signatures the rightmost data point of the circles-curve used a scrambled SIG list derived by combining seven increments of randomly chosen signatures from the original SIG list, whereas the leftmost data point of the diamonds-curve used the original SIG list.

The curve designated by stars shows the improvement in PD as increasing amounts of the 5,156-signature SIG database is added to the 90,432-signature synthetic database. The performance is shown to improve to about a PD of 83%, an increase of 13% over the performance observed using the synthetic data alone, and a 1% increase over the performance observed using the SIG data alone. Note that the rightmost point of the diamonds-curve and the rightmost point of the stars-curve represent performance of classifiers trained on the same database (the sum of all the SIG and all the synthetic data), and yet there is a 2% difference in performance. As previously noted, we believe this difference in performance, is due to a difference in the ordering of the signatures in the two training lists. One list starts with the original, complete, list of all of the SIG data and then adds six increments of randomly selected synthetic signatures until all of the synthetic data is added, while the other list starts with the original, complete, list of all of the synthetic data and then adds six increments of randomly selected SIG signatures until all of the SIG data is added.

Comparing the curve represented by the star symbol (for classifier performance using all of the synthetic data plus increasing amounts of measured SIG signatures), with the curve represented by the circle symbol (for measured SIG signatures alone), we observe a statistically significant, but marginal, 3%, improvement in performance when more than 500 measured signatures are available, but when less than 250 measured signatures are available the improvement increases significantly. This shows that using a substantial, but randomly organized, synthetic FLIR database significant levels of classifier performance can be achieved even if little measured data is available.

4.3 Results From Experiments Using Datasets Selected By Special Purpose Classifiers

Experiment 5 examined classifier performance as trained by a scenario relevant classifier. Table 1, presented in Confusion Matrix form, summarizes the results of this experiment. Tables 2 through 4 are shown for comparison. The lighter, diagonal elements with bold numerals show the correct identifications.

Table 1. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on a database of the union of the ROI-LIKE and SIG-LIKE synthetic subsets. Overall probability for correct identification is 81%.

Probability of Correct Target Identification (%)				
Predicted/Actual	<i>Target 1</i>	<i>Target 2</i>	<i>Target 3</i>	<i>Target 4</i>
<i>Target 1</i>	90	9	14	-
<i>Target 2</i>	5	76	9	-
<i>Target 3</i>	4	14	76	-
<i>Target 4</i>	1	1	1	-

Table 2. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on a database of SIG images. Overall probability for correct identification is 82%.

Probability of Correct Target Identification (%)				
Predicted/Actual	<i>Target 1</i>	<i>Target 2</i>	<i>Target 3</i>	<i>Target 4</i>
<i>Target 1</i>	85	4	6	-
<i>Target 2</i>	4	81	8	-
<i>Target 3</i>	2	12	81	-
<i>Target 4</i>	9	3	5	-

Table 3. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on the union of SIG, SIG-LIKE and ROI-LIKE synthetic datasets. Overall probability for correct identification is 85%.

Probability of Correct Target Identification (%)				
Predicted/Actual	<i>Target 1</i>	<i>Target 2</i>	<i>Target 3</i>	<i>Target 4</i>
<i>Target 1</i>	85	2	4	-
<i>Target 2</i>	4	85	10	-
<i>Target 3</i>	2	8	80	-
<i>Target 4</i>	9	5	6	-

Table 4. A confusion matrix showing the probability for correct ROI-TESTING target identification for a classifier trained on a database of the union of SIG and ROI-TRAINING datasets. Overall probability for correct identification is 88%.

Probability of Correct Target Identification (%)				
Predicted/Actual	<i>Target 1</i>	<i>Target 2</i>	<i>Target 3</i>	<i>Target 4</i>
<i>Target 1</i>	97	2	2	-
<i>Target 2</i>	2	85	15	-
<i>Target 3</i>	2	13	81	-
<i>Target 4</i>	0	0	1	-

Table 1 shows the confusion matrix for the classifier trained on the union of the ROI-LIKE and SIG-LIKE synthetic datasets. This dataset was selected by a scenario relevant classifier. The overall PD was 81% indicating that the performance of this classifier is comparable to that of the SIG trained classifier as shown by the circles-curve in Figure 10. This result is in contrast with that obtained when the selection of synthetic-data was done incrementally and randomly. As shown by the squares-curve in Figure 10, incremental and random selection of synthetic training data produced a PD that never exceeded 70%, almost 11% less than that obtained by specially selecting the training data. This suggests that the incremental and random selection of training data is unlikely to produce optimal classifier performance, and that using advanced information about targets, locations and conditions can improve performance.

For comparison, Table 2 shows the confusion-matrix for the classifier trained on SIG data alone. The overall PD is 82%. Misidentifications of Target 2 as Target 3 and vice versa are due to the fact that both targets are tanks. Table 2 details the classifiers performance against the individual targets summarized by the rightmost point of the circles-curve in Figure 10.

Table 3 shows the confusion matrix for the classifier trained on the union of the SIG, SIG-LIKE and ROI-LIKE synthetic datasets. The overall probability for correct identification was 85%, an increase of over 3% from the single database results for either classifier in Tables 1 and 2. This demonstrates that classifier performance can be improved when synthetic and measured databases are joined, albeit marginally. Again misidentifications mix Target 2 with Target 3.

For comparison, Table 4 benchmarks the performance of the classier. It shows the confusion matrix for the classifier trained on the union of the measured SIG, and ROI-TRAINING datasets. Again the test set was the ROI-TESTING subset. The overall PD was 88%. We will use this performance level as a benchmark since it alone uses data taken from the same database from which the test set was chosen.

Using the results of Table 4 as the benchmark, the relative overall performance for classifiers represented in Tables 1, 2, and 3 are 93.5, 92.4, and 96.8 percent respectively.

5. Conclusions

We have shown that augmenting measured FLIR target signature data with synthetic data can produce performance that exceeds that of using either database alone. We have shown that synthetic sets can be competitive with real data sets, and so can be used when measured data sets cannot be obtained due to expense or unavailability of targets, sensors, or site access. We have shown that within a large, diverse, database of synthetic

FLIR signatures there exist subsets whose trained classifier performance can exceed that achieved using the whole database. We have shown that these subsets can be selected using scenario relevant classifiers, and that a classifier, trained with this data can perform as well as a classifier trained on measured images alone.

We have obtained these results with relatively low-resolution images, derived from extremely low-resolution target models. We have taken care to simulate physically reasonable target states commensurate with measured data scenarios and we have validated our data by comparing synthetic to measured data performance in the training and testing of target classifiers.

Acknowledgments

We wish to extend a special thanks to Lipchen Alex Chan and Sandor Der of ARL for many helpful discussions.

References

- [1] C. P. Walters, C. W. Hoover, Jr., and J. A. Ratches, "Performance of an automatic target recognizer algorithm against real and two versions of synthetic imagery," *Opt. Eng.*, 39 (8), 2279-2285 (2000).
- [2] T. C. Edwards, R. H. Vollmerhausen, and R. Driggers, "Use of Synthetic Imagery In Target Detection Model Improvement," *Proceedings of the SPIE - The International Society for Optical Engineering*, 4372, 194-205 (2001).
- [3] E. Jacobs, D. Wilson, "FLIR Range Performance Predictions Using Synthetic Imagery," *Proceedings of the SPIE - The International Society for Optical Engineering*, 4030, 124-132 (2000).
- [4] S. P. Sullivan, and W. R. Reynolds, "Validation of the Physically Reasonable Infrared Signature Model (PRISM)," *Proceedings of the SPIE - The International Society for Optical Engineering*, 890, 104-110 (1988).
- [5] H. M. Nguyen, J. A. Penn, T. A. Kipp, G. Huynh, and J. Colby, "The CREATION Scene Generation Program And Its Application To Development And Testing Of Automatic Target Recognizers," *ATRWG Proceedings*, 1, 56-73 (1996).
- [6] B. A. Weber, and J. A. Penn, "Synthetic Forward-Looking Infrared Signatures for Training and Testing Target Identification Classifiers," *Optical Engineering*, 43, 6, 1414-1423 (June, 2004).
- [7] A. L. Chan, and N. M. Nasrabadi, "An Application of Wavelet-Based Vector Quantization in Target Recognition", *International Journal on Artificial Intelligence Tools*, 6 (2), 165-178 (1997).

Distribution list

Admnstr
Defns Techl Info Ctr
ATTN DTIC-OCF (Electronic copy)
8725 John J Kingman Rd Ste 0944
FT Belvoir VA 22060-6218

DARPA
ATTN C Schwartz
ATTN IXO S Welby
ATTN R Hummell
3701 N Fairfax Dr
Arlington VA 22203-1714

Ofc of the Secy of Defns
ATTN ODDRE (R&AT)
The Pentagon
Washington DC 20301-3080

Army Rsrch Physics Div
ATTN AMSRD-ARL-RO-MM R Launer
PO Box 12211
Research Triangle Park NC 27709-2211

AVCOM
ATTN AMSAM-RD-WS-PL W Davenport
Bldg 7804
Redstone Arsenal AL 35898

US Army RDECOM AMRDEC
ATTN ATCD-B
FT Monroe VA 23651-5850

CECOM NVESD
ATTN AMSRD-CER-NV-D J Ratches
10221 Burbeck Rd Ste 430
FT Belvoir VA 22060-5806

Dir for MANPRINT Ofc of the
Deputy Chief of Staff for Prsnl
ATTN J Hiller
The Pentagon Rm 2C733
Washington DC 20301-0300

US Army Aberdeen Test Center
ATTN CSTE-DT-AT-WC-A F Carlen
400 Collieran Road
Aberdeen Proving Ground MD 21005-5009

US Army Aberdeen Test Center
ATTN CSTE-DTC-AT-TC-N D L Jennings
400 Collieran Road
Aberdeen Proving Ground MD 21005-5059

US Army Aviation & Mis Lab
ATTN AMSRD-AMR-SG-IP H F Anderson
Bldg 5400
Redstone Arsenal AL 35809

RTTC
ATTN CSTE-DTC-RT-TE-SA T C Edwards
Redstone Arsenal, AL 35898

US Army Avn & Mis Cmnd
ATTN AMSAM-RD-SG-IP R Sims
Bldg 5400
Redstone Arsenal AL 35898

US Army CERDEC, NVESD
ATTN AMSRD-CER-NV-xx J Hilger
ATTN AMSRD-CER-NV-xx P Perconti
ATTN AMSRD-CER-NV-xx R Driggers
10221 Burbeck Rd Ste 430
FT Belvoir VA 22060-5806

US Army Topographic Engrg Ctr
ATTN CEERD-TR-S R Rand
7701 Telegraph Rd
Alexandria VA 22315

US Army Natick RDEC
Acting Techl Dir
ATTN SBCN-TP P Brandler
Kansas Street Bldg 78
Natick MA 01760-5056

US Army Natl Ground Intllgnc Ctr
ATTN CE-MA D Rinald
2055 Boulders Rd
Charlottesville VA 22911-8318

US Army PM NV/RSTA
ATTN SFAE-IEW&S-NV D Ferrett
10221 Burbeck Rd
FT Belvoir VA 22060-5806

US Army RDECOM ARDEC
ATTN AMSRD-AAR-QES P Willson
Radiographic Laboratory, B.908
Picatinney Arsenal NJ 07806-5000

US Army ARDEC
ATTN AMSTA-AR-TD
Bldg 1
Picatinny Arsenal NJ 07806-5000

Commanding General
US Army Avn & Mis Cmnd
ATTN AMSRD-AMR W C McCorkle
Redstone Arsenal AL 35898-5000

US Army Tank-Automtv Cmnd RDEC
ATTN AMSRD-TAR-R G R Gerhart
ATTN AMSRD-TAR-R T Gonda
ATTN AMSRD-TAR-R D Thomas
Warren MI 48397-5000

US Army Soldier & Biological Chem Ctr
ATTN AMSSB-RRT-DP W Loerop
Edgewood Chem & Biological Ctr Bldg E-5544
Aberdeen Proving Ground MD 21010-5424

Commander
USAISEC
ATTN AMSEL-TD Blau
Building 61801
FT Huachuca AZ 85613-5300

AFRL/SNAA
ATTN M Jarratt
2241 avionics Circle Area B, Bldg 620
Wright Patterson AFB OH 45433-7321

CMTCO
ATTN MAJ A Suzuki
1030 S Highway A1A
Patrick AFB FL 23925-3002

SITAC
ATTN H Stiles
ATTN K White
ATTN R Downie
11981 Lee Jackson Memorial Hwy Suite 500
Fairfax VA 22033-3309

Director
US Army Rsrch Lab
ATTN AMSRD-ARL-RO-D JCI Chang
ATTN AMSRD-ARL-RO-EL W Sander
PO Box 12211
Research Triangle Park NC 27709-2211

US Army Rsrch Office
ATTN AMSRD-ARL-RO-PP R Hammond
PO box 12211
Research Triangle Park NC 27709-2211

US Army Rsrch Lab
ATTN AMSRD-ARL-CI-OK-T Techl Pub (2
copies)
ATTN AMSRD-ARL-CI-OK-TL Techl Lib (2
copies)
ATTN AMSRD-ARL-D A Grum
ATTN AMSRD-ARL-D J M Miller
ATTN AMSRD-ARL-SE J Pellegrino
ATTN AMSRD-ARL-SE J Rocchio
ATTN AMSRD-ARL-SE-S J Eicke
ATTN AMSRD-ARL-SE-SE B Weber
(30 copies)
ATTN AMSRD-ARL-SE-SE N Nasrabadi
ATTN AMSRD-ARL-SE-SE P Gillespie
ATTN IMNE-AD-IM-DR Mail & Records Mgmt