TARGET RECOGNITION USING LINEAR CLASSIFICATION OF HIGH RANGE RESOLUTION RADAR PROFILES

THESIS

Ricardo A. Diaz, Captain, USAF

AFIT/GE/ENG/04-06

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/GE/ENG/04-06

TARGET RECOGNITION USING LINEAR CLASSIFICATION OF HIGH RANGE RESOLUTION RADAR PROFILES

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Electrical Engineering

Ricardo A. Diaz, BSEE

Captain, USAF

March 2004

AFIT/GE/ENG/04-06

TARGET RECOGNITION USING LINEAR CLASSIFICATION OF HIGH RANGE RESOLUTION RADAR PROFILES

Ricardo A. Diaz, BSEE
Captain, USAF

Approved:

/signed/

Steven C. Gustafson, Ph.D. (Chairman)	Date

/signed/

Maj Matthew E. Goda, Ph.D. (Member)	Date

/signed/

Michael A. Temple, Ph.D. (Member)	Date

In memory of Maj William D. Wood, PhD, whose guidance was generously extended and well received during a period of personal difficulty.
I am forever grateful.

Table of Contents

Table of Figures

Abstract

High Range Resolution (HRR) radar profiles map three-dimensional target characteristics onto one-dimensional signals that represent reflected radar intensity along target extent. In this thesis, second through fourth statistical moments are extracted from HRR profiles and input to Fisher Linear Discriminant (FLD) classifiers. Each FLD implements dimensionality reduction by projecting features onto the line yielding the greatest separation between classes. Features extracted from HRR profiles for unknown targets are projected onto this line and classified as belonging to the nearest target (i.e., the relative probability that the projection is from the probability density of this target is maximized).

An iterative classification process is applied that gradually minimizes required a priori knowledge about the target data. First, an HRR profile from a known target class and a known aspect angle is compared to an HRR profile from one of two target classes at different known aspect angles. A determination is then made as to whether the second HRR profile is from a "same" or "different" target. Second, a single HRR profile is extracted from a 360-degree aspect angle window and classified into one of two target classes with no a priori knowledge about the aspect angle. Finally, a single HRR profile is extracted from an aspect angle window and classified into one of three target classes.

It is found that the second through fourth statistical moments of HRR profiles are useful features in the FLD classification of dissimilar targets and they provide reasonable discrimination of similar targets. Greater than 69% correct classification for two-target scenarios and greater than 60% correct classification for three-target scenarios is obtained using a single HRR profile extracted from a full 360-degree aspect angle window. A key contribution of this thesis is the demonstration that simple statistical moment features and simple linear classifiers can be used to effectively classify HRR profiles.

TARGET RECOGNITION USING LINEAR CLASSIFICATION OF HIGH RANGE RESOLUTION RADAR PROFILES

I. Introduction

1.1 Background

Airborne radar is a primary means of target identification in the absence of active emitters such as Identification Friend or Foe (IFF) systems or tactical data link communication systems. High Range Resolution (HRR) radar profiles map the three-dimensional physical characteristics of a target onto a one-dimensional signal. The one-dimensional signal represents the reflected radar intensity versus range along the target extent. Using HRR profiles for classification is problematic because they are highly variable and depend on target aspect and elevation angles with respect to the airborne platform.

1.2 Previous Work

Target classification research using HRR profiles requires the following. First, useful HRR profile features must be identified and extracted for input to a classifier. Second, a classification scheme must be selected for training on the extracted features. Some possible feature sets and classification schemes are indicated below.

1.2.1 HRR Profile Features

In past work, numerous HRR profile features have been employed for classification. Recent research has used target HRR profile range bins as the primary feature set [15] [17]. Other research has focused on the scattering geometry [22] and polarization of the target signature [15]. Some work has used the target major intensity peaks [8] [10], wavelet transformations [5] and fractal dimension [18].

1.2.2 HRR Profile Classification

Classification methods using HRR profile features have employed, for example, clustering [12] [19], correlation filters in the time and frequency domains [17], and hidden Markov models [4]. Classification has also used a variety of neural networks, including radial basis function [17], multilayer perceptron [8], and adaptive time delay [8] designs.

1.2.3 Literature Review

The Defense Advanced Research Projects Agency (DARPA) sponsored the Moving Target Features and Phenomenology (MTFP) program to discover, analyze, and evaluate HRR target features [1]. The MTFP program measured HRR signature data from seven moving targets taken from an airborne radar platform at aspect angles from 1 degree to 360 degrees.

Under the MTFP program a nonlinear scheme to classify moving targets was developed [1]. It used two HRR profiles as input and calculated the probability that the two signals belonged to the "same" or "different" vehicles when compared to two

previously identified HRR profiles. The scheme used a machine learning approach to track a target and attempted to discriminate HRR profiles independent of vehicle type. Two test hypotheses were of particular interest: relative and absolute. For the relative test hypothesis two HRR profiles were identified. Subsequently, two new HRR profiles were randomly selected from a sample set bounded by an aspect angle differential of less than 10 degrees from the first input HRR profiles. The two new input HRR profiles were matched uniquely as either a "same" or "different" target relative to the previously identified HRR profiles. The classifier had a 93 percent correct decision rate given these parameters [1]. For the absolute test hypothesis the second HRR profiles were evaluated independent of the competing hypothesis used in the relative test case and thus unique matching to the previously identified tracks was not required. This procedure greatly reduced the processing time, but classifier performance dropped significantly to a 40 percent correct decision rate when multiple targets were considered [1].

Classification using Multinomial Pattern Matching achieved a 74% success rate with HRR profiles from three airborne target classes [21]. Sampling for classification was restricted to a five-degree-by-five-degree aspect and elevation angle window.

1.3 Problem Statement

The performance of target HRR profile classifiers is highly dependent on both the type of features extracted and the classifier chosen. As stated above, there has been extensive research involving numerous feature sets and classification schemes. This thesis develops a linear target HRR profile classifier motivated by the scenario described in [1]. A Fisher linear classifier is trained using a subset of measured data from the MTFP

program dataset and tested using a different subset of the same data set. The second through fourth moments of the target HRR profiles are analyzed for effectiveness as classification features.

1.4 Objectives

A simplified approach to target HRR profile classification is investigated. Effective neural network classifiers typically have an associated credit assignment problem: it is not possible to determine the relationship among the input features that it employs to implement the classification [4]. Linear classification schemes offer the considerable advantage that these relationships are readily determined. In this thesis the second through fourth moments of each HRR profile is used as the feature set. Through an iterative approach, feature extraction and the application of a linear classifier culminates in the ability to identify a randomly selected HRR profile and associate it with one of three possible targets. The HRR profile sample is extracted from a set of measured data on three known targets, over a 1-degree to 360-degree range of aspect angles. The goal is to show that properly selected features processed with an appropriate linear classifier yields results comparable to or exceeding those obtained using more complex features and classifiers.

1.5 Materials and Equipment

The MTFP dataset is used exclusively as the data source. The Application Program Interface (API) is used to extract and sort data from the MTFP dataset [7]. Algorithms are developed on a personal computer with Matlab® Version 6.5 software using the signal processing and statistical toolboxes.

1.6 Thesis Organization

Chapter II provides an overview of relevant pattern recognition concepts, radar HRR profiles, and the MTFP dataset. Chapter III discusses the methodology, and Chapter IV contains the experiment results and analysis. The last chapter presents conclusions and suggestions for future research.

II. Background

2.1 Pattern Recognition

Pattern recognition conventionally involves calculating probability densities for each class of data so that any point in the feature space has a greater a posterior probability of being in one class than other classes. Data used for target recognition contains numerous features, some of which are readily apparent to the human eye, while others are less evident. Pattern recognition systems attempt to identify distinct classes through the extraction of useful and significant features. Less helpful or irrelevant features are discarded as noise. This thesis uses HRR profiles exclusively as input for classification.

2.1.1 Feature Extraction

Pattern recognition requires extraction of attributes that will be most advantageous for good classification. For HRR profiles the Doppler data is not applicable and is discarded, but the magnitude of the signal is retained. After constructing the HRR profile, specific HRR profile features are extracted, further simplifying the representation of the original radar return signal. These features serve as input to the designated target recognition classifier. The number of features must be sufficient to yield separable classes at the output. Selecting effective features reduces processing time and can greatly enhance classifier performance, and it is perhaps the most critical step in the pattern recognition process.

2.1.2 Classification

Classification of the extracted features is the final step in pattern recognition and is ordinarily performed within a statistical framework. The extracted features are analyzed for patterns or trends that categorize the input data into a specific class.

Statistical pattern recognition often encounters the "curse of dimensionality," in that large numbers of features reduce classification performance. The selected features constitute a feature space that can easily exceed the three dimensions which humans can visualize. The number of dimensions in this feature space may be reduced by projection into a lower dimensional space. There are several common methods for dimension reduction. Two will be examined further: Principal Components Analysis (PCA) and Fisher Linear Discriminant Analysis (FLD).

2.1.3 Principal Component Analysis

PCA, also known as the Karhunen-Loeve transform in function space, is frequently used for dimension reduction. Assume data representing measured HRR profiles exists for one class is extracted from a limited range of aspect angles (less than 90 degrees) as viewed from the airborne platform. Further assume that the samples in the class are represented by n vectors $\{x_1, x_2, \ldots, x_n\}$, where each vector represents k distinctive measurements or features. In its current state, this sample is characterized in k-dimensional space, but PCA can represent the entire class of n vectors in $k' < k$ dimensional space to optimally characterize the class. Although the dimension reduction may be useful, PCA does not necessarily make it possible to discriminate *between* classes, and it may discard useful *discriminating* dimensions. A popular example

involves performing PCA on features that characterize data consisting of samples of the letter O and the letter Q. A reduced dimension characterization keeps the coarse features of the letters intact but omits the tail from the Q, which makes it impossible to distinguish the O from the Q in the new feature space [4]. Therefore, the use of a class *discriminator* is more advantageous.

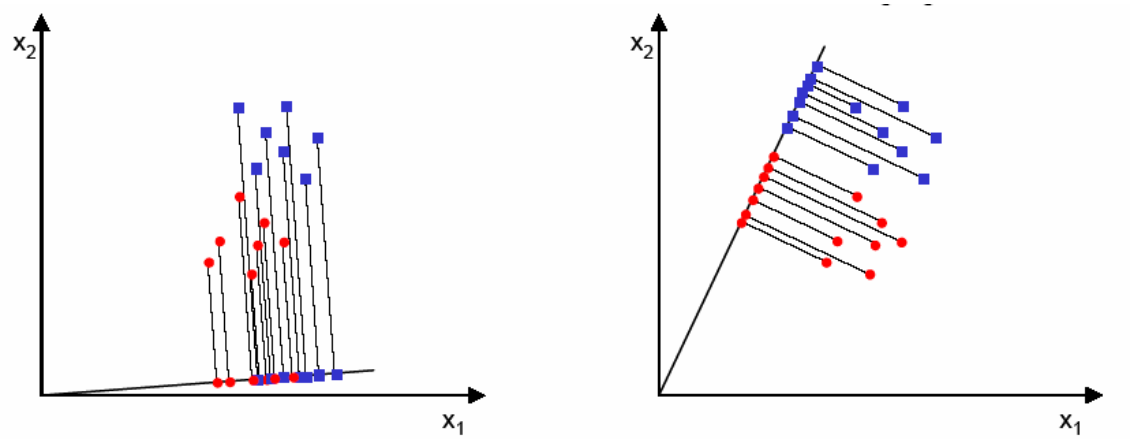
2.1.4 Fisher Linear Discriminant Analysis

FLD analysis not only characterizes a class of data through dimension reduction, but it also projects the data onto a line that yields the greatest separation between classes. The challenge lies in orienting the line so that the data projected onto it are clearly separated [6].

For each vector $x = \{x_1, x_2, ...,x_N\}$ of D-dimensional features, where N_1 of the vectors are in class ω_1 and N_2 are in class ω_2 and $N = N_1 + N_2$, FLD obtains a scalar value y by projecting x onto a line,

$$y = w^T x, \tag{1}$$

where the vector w specifies the slopes of the line. Infinitely many lines are possible, but FLD finds the line that maximizes the separability of between-class scalar values as shown in Figure 1.



(a) (b)

Figure 1. Fisher Linear Discriminator Example

A two-dimensional case demonstrating choices of a discriminating line. (a) Classes not clearly separated (b) Classes clearly separated by the FLD line. Figure from [6].

One measure of separation between the projections is the difference between the means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ of the feature vectors in each class after projection to the line, so that the function maximized is

$$J(w) = \left|\tilde{\mu}_1 - \tilde{\mu}_2\right| = \left|w^T(\mu_1 - \mu_2)\right|. \tag{2}$$

where the D-dimensional sample mean is defined by

$$\mu_i = \frac{1}{n_i}\sum_{x \in D_i} x_i . \tag{3}$$

However, the distance between the projected means does not account for the standard deviation within the classes. Therefore, maximizing the distance between the means could still yield minimal class separation, as demonstrated in Figure 2.

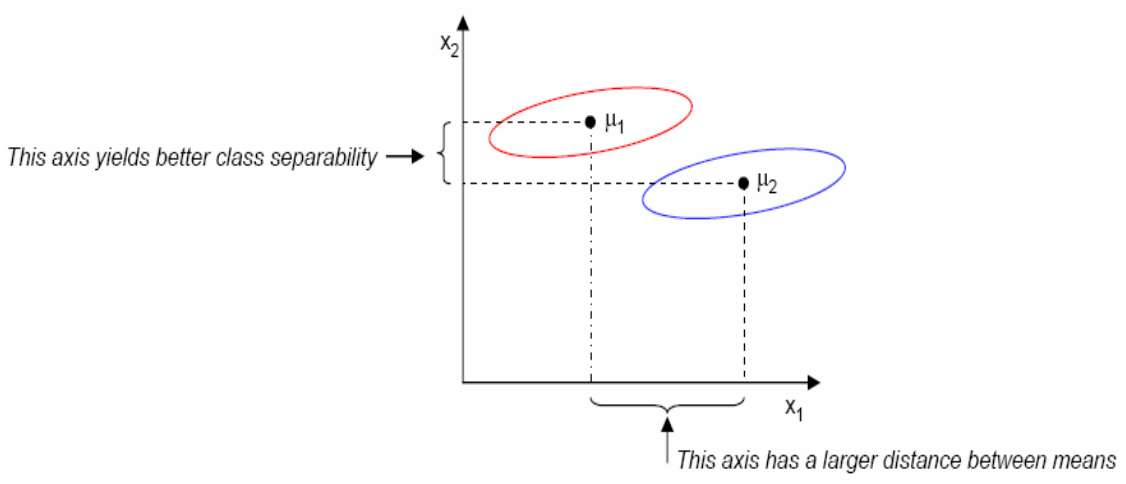


Figure 2. Comparison of Class Separability Along Different Axis

Maximizing the projected distance between class means does not guarantee greater class separability. For the case above, the projection onto the axis of greatest class mean difference yields significant between class overlap. Figure from [6].

The sample mean for the projected points is given by

$$\begin{aligned} \tilde{\mu}_i &= \frac{1}{n_i} \sum_{y \in y_i} y \\ &= \frac{1}{n_i} \sum_{x \in D_i} w^T x \\ &= w^T \mu_i \end{aligned} \tag{4}$$

and is simply the projection of $\boldsymbol{\mu}_i$.

The distance between the projected means from equation (2) is

$$\left| \tilde{\mu}_1 + \tilde{\mu}_2 \right| = \left| w^T (\mu_1 - \mu_2) \right|. \tag{5}$$

The difference between the means should be large relative to some measure of the standard deviations for each class. Therefore, define the scatter for the projected samples by

$$\tilde{s}_i^2 = \sum_{y \in Y_i} (y - \mu_i)^2 . \tag{6}$$

FLD seeks to maximize the separation between projected classes by maximizing the criterion function,

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2}, \tag{7}$$

where $\tilde{s}_1^2$ and $\tilde{s}_2^2$ are proportional to the variances of the projected points in each class about their means. Thus, the ratio of squared inter-class separation to the sum of the squares of the intra-class variance separations is maximized as shown in Figure 3.

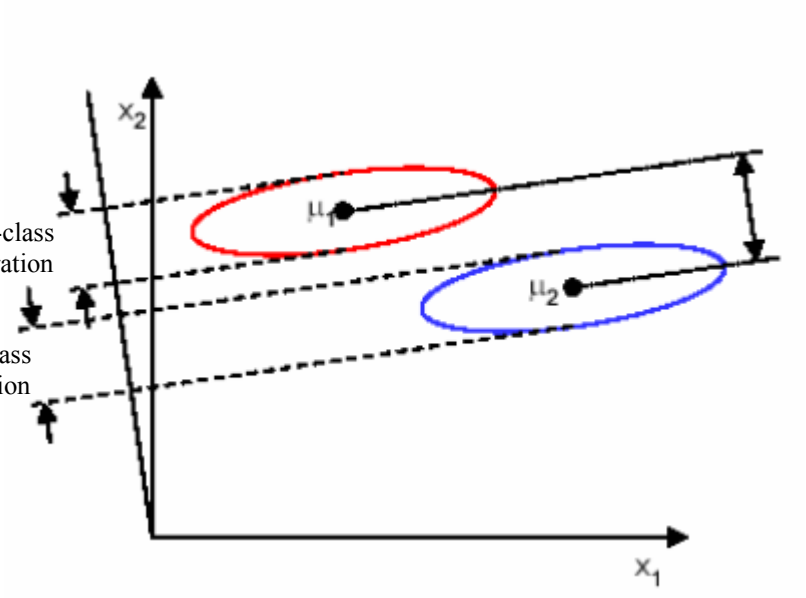


Figure 3. Fisher Linear Discriminant Projection Example

The Fisher linear discriminant identifies a projection where intra-class samples are projected close together while the inter-class projections are projected far apart [6].

It is necessary to express *J(w)* as an explicit function of *w* in order to find *w*.

Therefore, define the scatter matrices equivalent to the scatter in the projection as

$$S_i = \sum_{x \in D_i} (x - \mu_i)(x - \mu_i)^T \tag{8}$$

and

$$S_W = S_1 + S_2 . \tag{9}$$

The matrix S_W is proportional to the sample covariance matrix and is called the total within-class scatter matrix.

The scatter of the projection is expressed as a function of the scatter matrix in the *x* feature space.

$$\begin{aligned} \tilde{s}_i^2 &= \sum_{x \in D_i} (w^T x - w^T \mu_i)^2 \\ &= \sum_{z \in D_i} w^T (x - \mu_i)(x - \mu_i)^T w \\ &= w^T S_i w; \end{aligned} \tag{10}$$

therefore the sum of these scatters can be written

$$\tilde{s}_1^2 + \tilde{s}_2^2 = w^T S_W w . \tag{11}$$

Similarly, the separations of the projected means obeys

$$\begin{aligned} (\tilde{\mu}_1 - \tilde{\mu}_2)^2 &= (w^T \mu_1 - w^T \mu_2)^2 \\ &= w^T (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T w \\ &= w^T S_B w, \end{aligned} \tag{12}$$

where the matrix S_B is the between-class scatter matrix defined as

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T . \tag{13}$$

Combining yields,

$$J(w) = \frac{w^T S_B w}{w^T S_w w}. \tag{14}$$

Maximizing the between-class scatter (numerator) and minimizing the within-class scatter (denominator) for optimal class discrimination is achieved as follows:

$$\begin{aligned}
\frac{d}{dw}\left[J(w)\right] &= \frac{d}{dw}\left[\frac{w^T S_B w}{w^T S_w w}\right] = 0 \\
&\Rightarrow \left[w^T S_w w\right]\frac{d\left[w^T S_B w\right]}{dw} - \left[w^T S_B w\right]\frac{d\left[w^T S_W w\right]}{dw} = 0 \\
&\Rightarrow \left[w^T S_W w\right]2S_B w - \left[w^T S_B w\right]2S_W w = 0 \\
&\Rightarrow \frac{\left[w^T S_W w\right]}{\left[w^T S_W w\right]} S_B w - \frac{\left[w^T S_B w\right]}{\left[w^T S_W w\right]} S_W w = 0 \\
&\Rightarrow S_B w - J S_W w = 0 \\
&\Rightarrow S_W^{-1} S_B w - J w = 0,
\end{aligned} \tag{15}$$

where solving the generalized eigenvalue problem $(S_W^{-1} S_B w = Jw)$ yields

$$w = \arg\max\left\{\frac{w^T S_B w}{w^T S_W w}\right\} = S_w^{-1}(\mu_1 - \mu_2). \tag{16}$$

For the two-class case C = 2 there is only one discriminant function that projects the D-dimensional space onto the C-1 dimensions. However, in the general case where the number of classes is C > 2, FLD reduces the dimensionality from D-dimensions to C-1 dimensions [4]. In this case, the projection vectors placed as columns in a matrix **W**

$$\begin{aligned}
y_i &= w^T x \Rightarrow y \Rightarrow W^T x \\
W &= \left[w_1 \quad w_2 \quad \ldots \quad w_{c-1}\right],
\end{aligned} \tag{17}$$

The generalized total within-class scatter matrix is now

$$\begin{aligned} S_i &= \sum_{x \in D_i} (x - \mu_i)(x - \mu_i)^T \Rightarrow \\ S_W &= \sum_{i=1}^{c} S_i, \end{aligned} \tag{18}$$

where D_i is the set of samples drawn from class *i*. The generalized case for the total between-class scatter matrix S_B is defined as

$$S_B = \sum_{i=1}^{c} N_i (\mu_i - \mu)(\mu_i - \mu)^T, \tag{19}$$

where N_i is the total number of samples from class *i* and $\boldsymbol{\mu}$ is the mean of individual class means $\boldsymbol{\mu}_i$ restated here from equation (3)

$$\mu_i = \frac{1}{n_i} \sum_{x \in D_i} x_i \tag{20}$$

Transforming these matrices to *y-space* yields

$$\begin{aligned} \tilde{S}_B &= \sum_{i=1}^{c} N_i (\tilde{\mu}_i - \tilde{\mu})(\tilde{\mu}_i - \tilde{\mu})^T \\ &= W^T S_B W, \end{aligned} \tag{21}$$

$$\begin{aligned} \tilde{S}_W &= \sum_{i=1}^{c} \sum_{y \in Y_i} (y - \tilde{\mu}_i)(y - \tilde{\mu}_i)^T \\ &= W^T S_W W, \end{aligned} \tag{22}$$

The criterion function is then

$$J(W) = \frac{\left|\tilde{S}_B\right|}{\left|\tilde{S}_W\right|} = \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|}, \tag{23}$$

maximizing the between-class matrix S_B and minimizing the within-class matrix S_W yields

$$W = \left[w_1 \;\; w_2 \ldots w_{c-1}\right] = \arg\max\left\{\frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|}\right\}. \tag{24}$$

Equivalently, the optimal projection matrix W consists of columns of eigenvectors corresponding to the largest eigenvalues in the eigenvalue problem [13]

$$S_B w_i = \lambda_i S_W w_i \tag{25}$$

This problem is addressed by first computing the eigenvalues as the roots of the characteristic polynomial $|S_B - \lambda_i S_W| = 0$. Solving for the w_i yields

$$(S_B - \lambda_i S_W) w_i = 0 \tag{26}$$

See Figure 4 for an illustration that uses three classes.

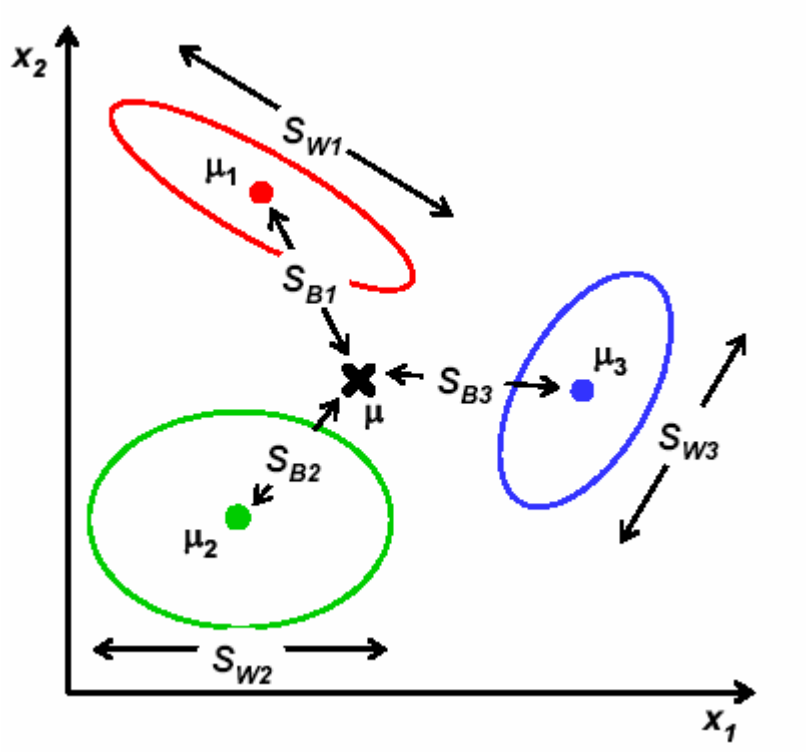
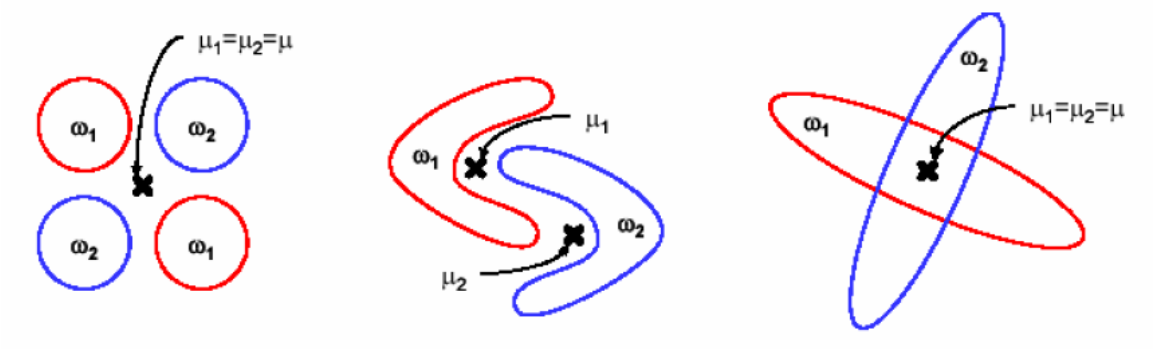


Figure 4. Example of Three-Class Fisher Linear Discrimination

Optimal results are obtained by minimizing the within-class matrix S_W and maximizing the between-class matrix S_B [6].

There are some limitations to using FLD. The discriminatory information must lie in the mean for FLD to succeed, and highly overlapping multi-model distributions reduce classification performance as illustrated in Figure 5 [13].



(a) (b) (c)

Figure 5. Fisher Linear Discrimination Examples

Fisher Linear Discriminant analysis fails if the class discriminatory information is not found within the mean. (a) The mean is identical in both classes and discrimination is not possible, (b) and (c), Class overlap prevents FLD discrimination [6].

2.1.5 Parzen Windows Probability Density

Parzen windows estimation is a nonparametric probability density estimation technique often used in classification. It approximates a density for a set of data points using a linear combination of a given kernel function centered on each of the points [13]. In this research, the kernel is the Gaussian radial basis function

$$K(x, x_i) = \exp\left(\frac{-(x - x_i)^2}{2\sigma^2}\right), \tag{27}$$

where x_i is the value of the i^{th} sample. Gaussian kernels force the weighted effect of a kernel on a neighbor to decrease exponentially with the square of the distance, rendering distant points irrelevant. The width σ of the chosen Gaussian controls the smoothness of the estimated density and were modified to approximate marginal unimodal densities for the target classes. An example of a Parzen windows estimation is shown in Figure 6.

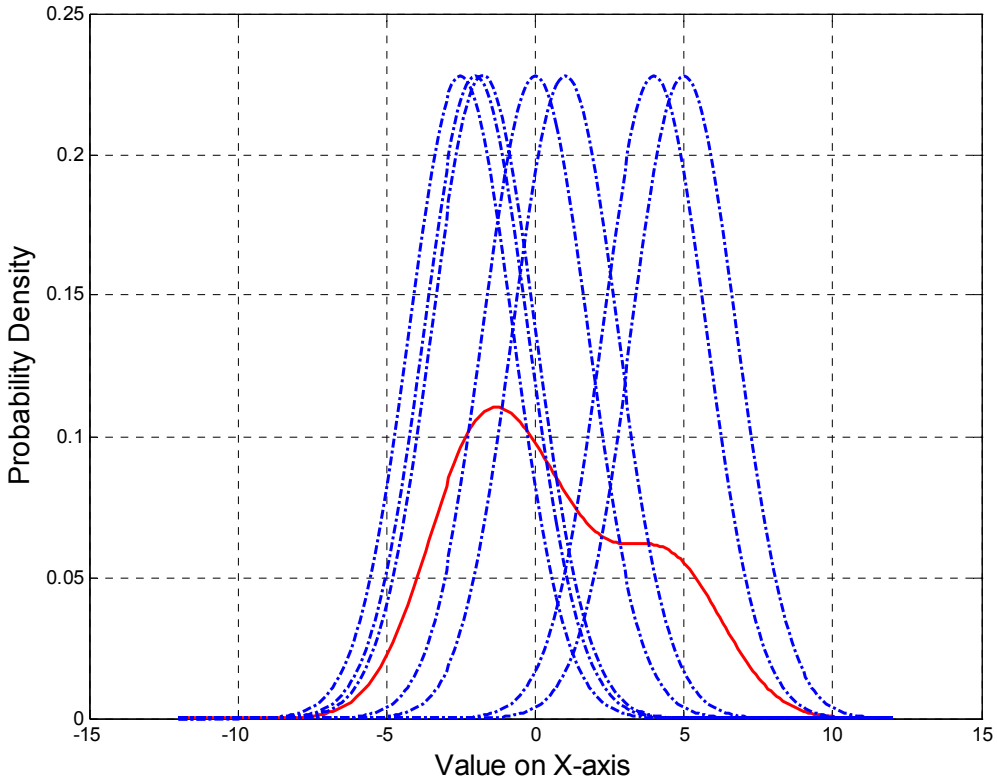


Figure 6. Parzen Window Example

A Gaussian density is placed on every data point projection on the FLD line and the densities are summed and divided by the total number of data points. In this example, sigma is chosen to create a bi-modal density, which means that there are two points where it's first derivative is zero.

2.2 High Range Resolution Radar

The review below focuses on topics relevant to High Range Resolution Radar as applied in this thesis.

2.2.1 High Range Resolution Profiles

A HRR profile used in this thesis is an n-dimensional vector $x = (a_1, a_2, ..., a_n)$, where $a_i \in \{0,1,...,199\}$. The position relative to the line of sight from the radar to the target is divided into range bins. Each element a_i in the HRR profile vector x is the magnitude of the total radar return in the i^{th} range bin. This thesis assumes that detection of the target has occurred and focuses on classification of a set of target HRR profiles.

For any target there are many possible HRR profiles due to the fact that target HRR profiles are not persistent throughout detection. Targets contain thousands of radar scatter points, and their net effect on the echo detected at the radar receiver changes significantly with the aspect and elevation angles of detection. Figure 7 illustrates a window of possible data extraction for an airborne target. Figure 8 reveals how a simple aspect angle change occurring in a fraction of a second can affect the HRR profile measured by a radar receiver.

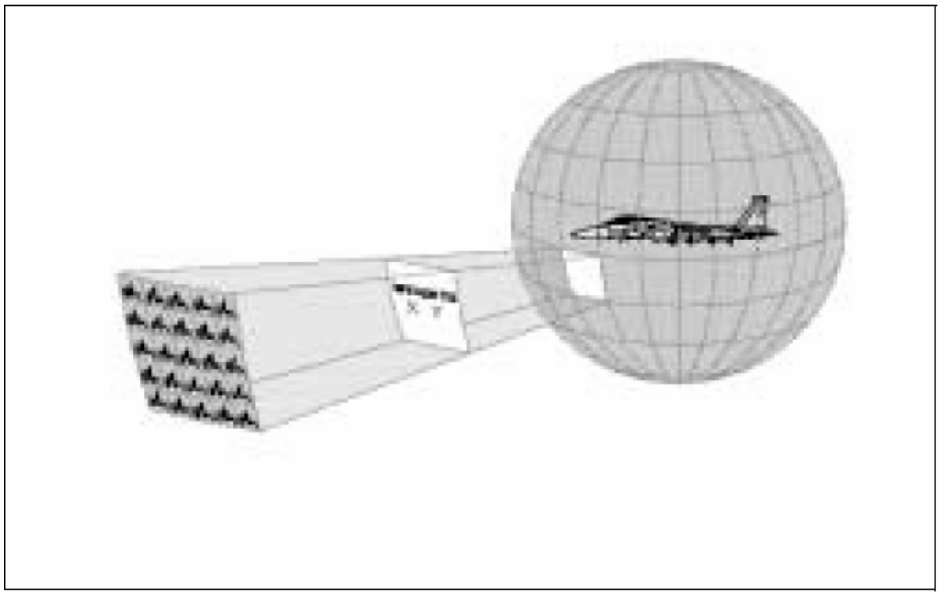


Figure 7. Window for Extraction of High Range Resolution Profile

Example of one viewing window for High Range Resolution profile extraction. It is possible to view an airborne target from many combinations of aspect and elevation angles [11].

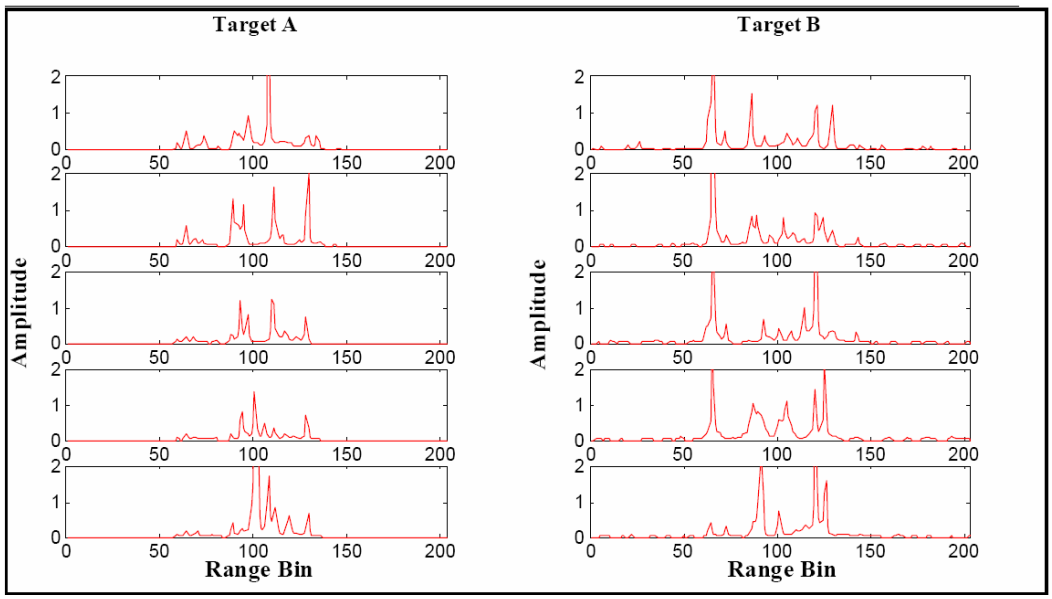


Figure 8. High Range Resolution Profile Variation

Example of High Range Resolution profiles extracted from a moving target over a period of one second. Each profile is taken 200 milliseconds apart [11].

Although many HRR profiles exist for any one target, some features can persist for at least some range of aspect angles. A target radar echo return is the accumulation of thousands of scatter centers, and it is normal for some scatter centers to dominate the total return at the receiver when viewed from certain angles. For example, a tank turret may generate a distinctive HRR profile when viewed from a 0-degree (head-on) position or from a 90-degree (side view) position. However, the same tank viewed from a 180-degree position (tail-on) may not have any noticeable return from the turret because the rest of the tank may obscure the turret.

Target recognition using HRR profiles relies on extracting persistent features for classification. In the example above, it may be easy to distinguish a tank from a civilian target such as a minivan by using the HRR profile feature due to the tank turret as a distinguishing characteristic. Unfortunately, the differences between military targets are usually subtle and more difficult to distinguish. This thesis examines the application of Fisher linear discrimination for target classification using moment features of HRR profiles.

2.3 HRR Profile Moments

HRR profiles for targets are complex and difficult to predict. However, certain characteristics may be persistent for at least small windows of aspect angle about the line-of-sight. Since the HRR profile length can be used to distinguish small targets from large targets, it is natural to speculate that the extraction of statistical characteristics for the length and shape of the HRR profile may provide good features for target classification.

Moments capture geometric information, and normalized HRR profiles can be viewed as one dimensional probability density functions (PDF). The first through fourth moments are used in this work. When applied to a one-dimensional probability density, the first moment, or mean, measures location. The mean position of a set of samples in units of sample number is

$$\mu_1 = \frac{\sum_{i=1}^{n} i y_i}{\sum_{i=1}^{n} y_i}, \tag{28}$$

where n is the total number of samples and y_i is the value of the i^{th} sample.

The second moment, or variance, of a probability density measures its extent. A small variance implies that the samples are clustered near the mean and a large variance indicates that the density is either spread out thinly or clustered at a position some distance from the mean. The variance of sample number is

$$\mu_2 = \frac{\sum_{i=1}^{n} (i - \mu_1)^2 y_i}{\sum_{i=1}^{n} y_i}, \tag{29}$$

The third moment, or skewness, measures the asymmetry of a density; see Figure 9. The skewness of sample number is

$$\mu_3 = \frac{\sum_{i=1}^{n} (i - \mu_1)^3 y_i}{\mu_2^{3/2} \sum_{i=1}^{n} y_i} \tag{30}$$

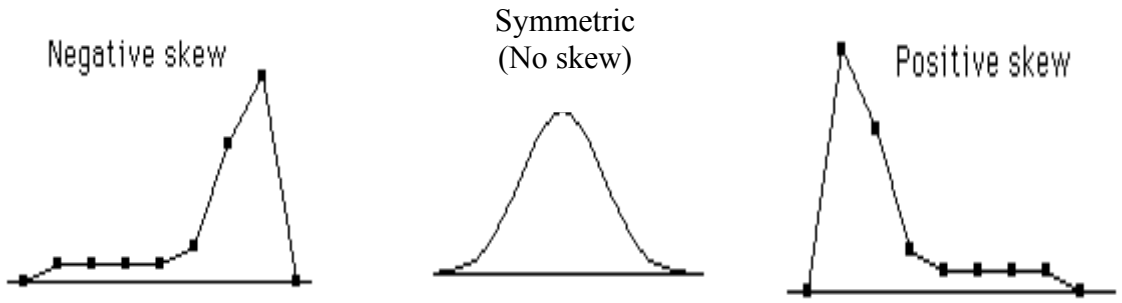


Figure 9. The Third Statistical Moment--Skewness

A positive skew indicates that the density has more probability in the positive direction while a negative skew indicates probability concentrated in the negative direction [9].

The fourth moment, or kurtosis, is a measure of the size of the density tails. Densities with larger tails are leptokurtic and densities with small tails are platykurtic. Kurtosis is defined by

$$\mu_4 = \frac{\sum_{i=1}^{n} (i - \mu_1)^4 y_i}{\mu_2^2 \sum_{i=1}^{n} y_i}. \tag{31}$$

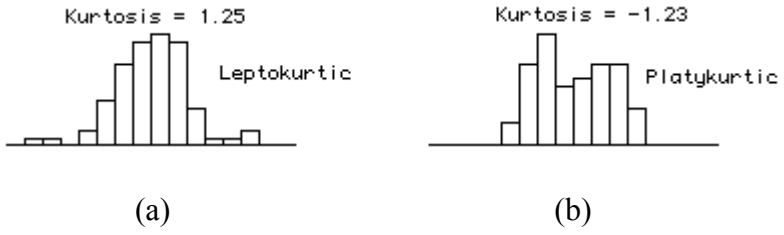


Figure 10. The Fourth Statistical Moment--Kurtosis

The two densities above have the same variance and skew but different kurtosis [9].

2.4 Moving Target Features and Phenomenology (MTFP) Program

Data from the DARPA Moving Target Features and Phenomenology (MTFP) program [1] are used for both training and testing the linear classifier. There are twelve target configurations using seven platforms: T-72 Tank, BTR-80 Armored Personnel Carrier, Mobile SCUD Launcher, M-2 Infantry Fighting Vehicle, ZIL-131 Truck, 2S1 Self Propelled Howitzer, and ZSU-23 Anti-Aircraft Gun. The twelve configurations are shown in Figure 11 and resulted from adding antennas, skirts, or reactive armor to the platforms.

Figure 11. Target Configurations for the MTFP Program

Measured data is collected from twelve target configurations using seven targets for the DARPA Moving Target Features and Phenomenology (MTFP) program [14].

The MTFP test setup uses a moving target on a 200-meter diameter track, and data is collected for four revolutions of each target configuration. The airborne platform has a side-looking two-channel Synthetic Aperture Radar operating in X-band (9.66 GHz) with a 1953 Hz pulse repetition frequency and one-foot resolution. Data is collected at the depression angles, slant ranges, and altitudes indicated in Figure 12.

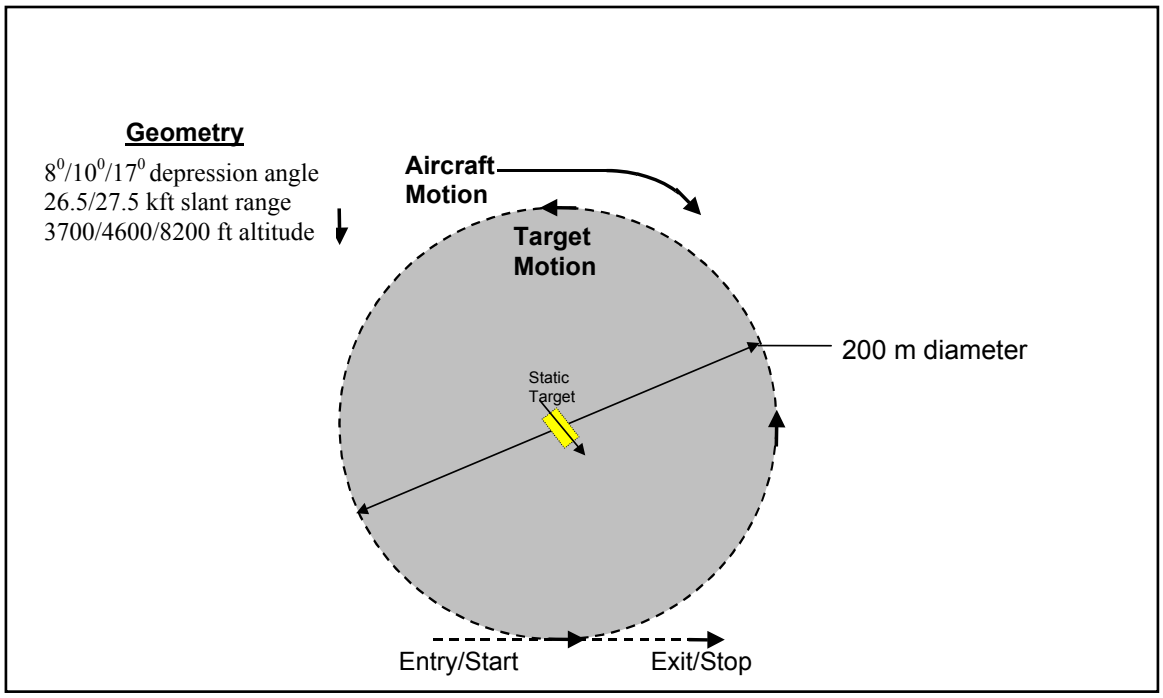


Figure 12. MTFP Test Set Up

A moving target circles a 200-meter circular track while the airborne platform flys in the opposite direction [14].

III. Methodology

3.1 Introduction

There are five steps in this research. The first step involves extracting data from the Moving Target Features and Phenomenology (MTFP) program data set. The second step involves pre-processing the data for further use in training and testing. The third step requires training and testing a linear classifier using significant a priori knowledge. The fourth step expands on results of the third step and attempts to recognize two target classes with minimal a priori knowledge. The final step expands results from step four to include three-target classification. This thesis effort generated several thousand lines of Matlab® code; the final Matlab® software product allows for operator initiated options and is easily adaptable for future research.

3.2 Data Extraction

The measured MTFP data contains ancillary information, or attributes, that explain test measurement conditions such as elevation, radar operating frequency, pulse repetition frequency (PRF), aspect angle, depression angle, aircraft speed and target speed, etc. The Data Object Application Program Interface (API) [7] facilitates the standardized storage of this data and is the primary interface used in this research. The measured data is queried and sorted by aspect angle for each target under consideration. The depression angle and the target configuration are not considered during sorting to permit less restricted data set (i.e., all depression angles and target configurations are used). After data is extracted it is split into separate training and testing sets. The primary data object

of interest is the HRR profile and the attribute of primary interest is the aspect angle at the time of measurement.

3.3 Data Pre-Processing

The measured data requires pre-processing prior to classification. The HRR profiles are extracted and input to Matlab® as a vector where the element number is x-axis position (1:200) and the element value is the y-axis decibel value. This assignment creates a range versus magnitude plot as seen in Figure 13 (a). The start range bin varies from target to target and profile to profile and depends on the true distance from the airborne platform to the target at the time of measurement. Among the attributes available in the MTFP data set is a target mask for each profile, Figure 13 (b). The target mask cuts the true target HRR profile from the measured data and zeros out all non-target values as shown in Figure 13 (c). Normalization along the y-axis scales the magnitude values in each target bin by a constant that forces the total area under the profile to equal one, creating a normalized power spectral density representation of the radar return. Although there is no associated random variable, the normalized HRR profile is regarded as a probability density function; Figure 13 (d).

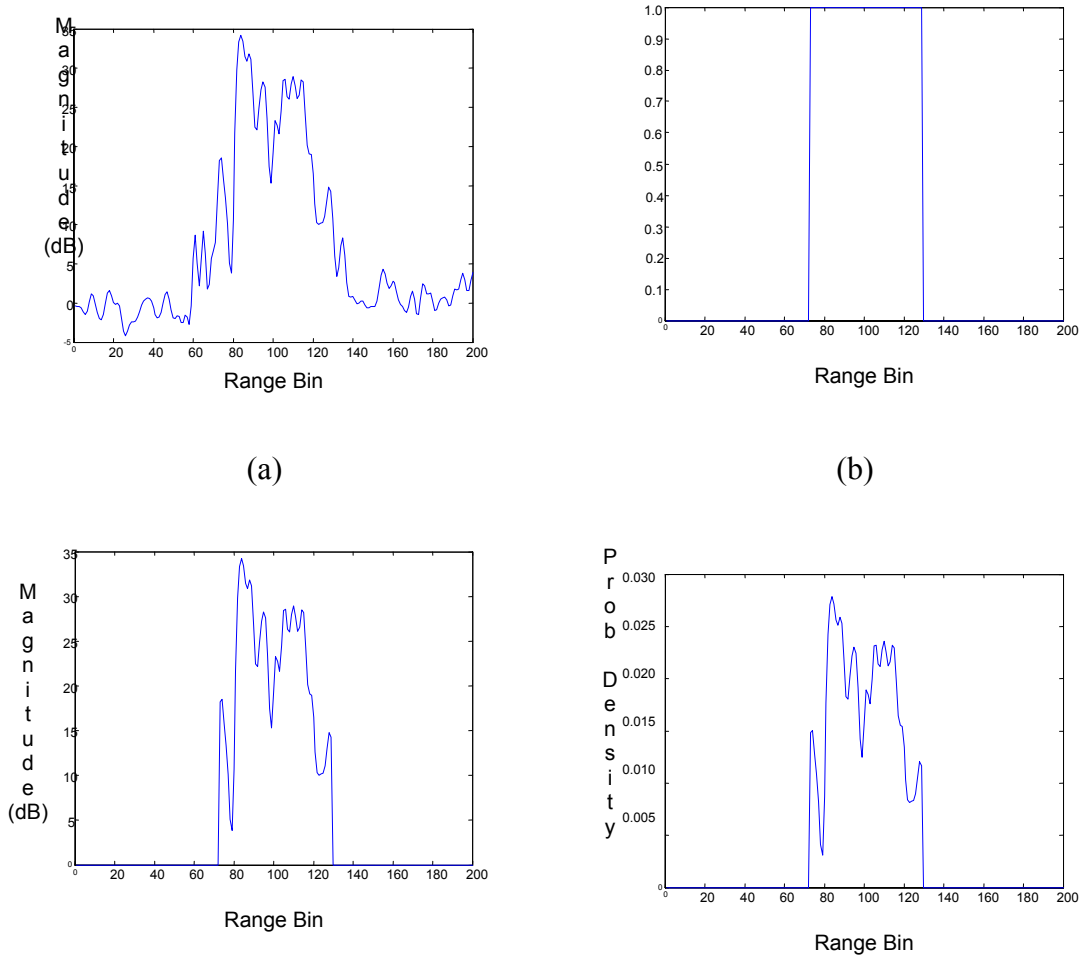
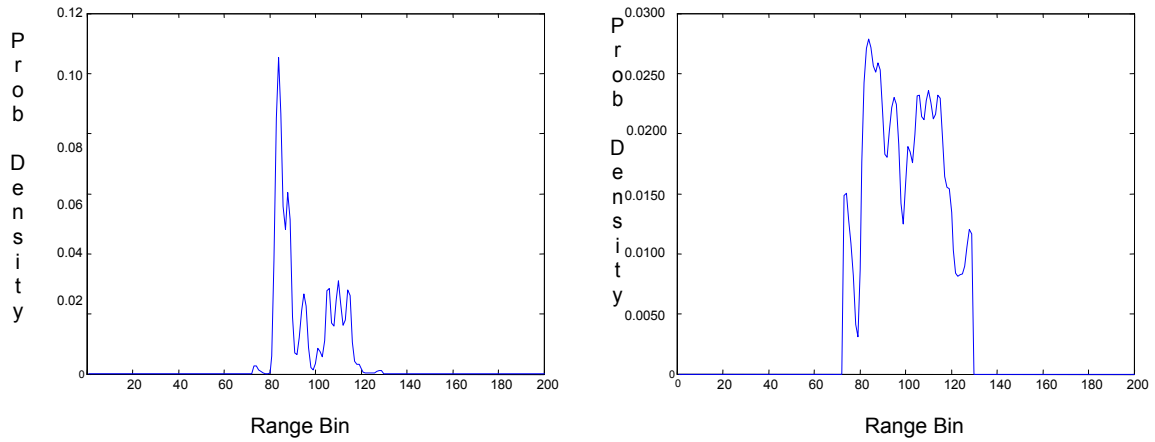


Figure 13. Application of Target Mask to HRR Profile Data

Here (a) is an example of measured data HRR profiles extracted from the MTFP data set, (b) is a target mask which identifies the range bins in which the target is located, (c) shows the resulting masked profile with extraneous noise removed, (d) is the profile normalized to form a PDF.

3.3.1 HRR Profile Scaling

A typical PDF normalized from a linear scale HRR profile, Figure 14 (a), is compared to a PDF normalized from a decibel scale HRR profile, Figure 14 (b). The decibel or logarithmically scaled profiles retained and highlighted distinguishing characteristics better than the linear scale profiles. Generally, as discussed in Chapter IV, classification using PDFs normalized from decibel-scaled HRR profiles yields the best results.



(a) (b)

Figure 14. Comparison of Linear and Decibel Scaled PDF

(a) PDF normalized from a linearly scaled HRR profile, (b) PDF normalized from decibel scaled HRR. Both PDFs originated from the same measured data.

3.3.2 Moment Extraction

The moments for each profile replace the 200 element HRR profile vector with a four element feature set vector consisting of the first through fourth moments. Since only these moments are used for classification, the result is a three-dimensional feature set representation of the HRR profiles. Figure 15 shows a comparison of the three-dimensional feature set for two similar targets, and Figure 16 shows a two-dimensional comparison of the 3rd and 4th moment feature set for four targets.

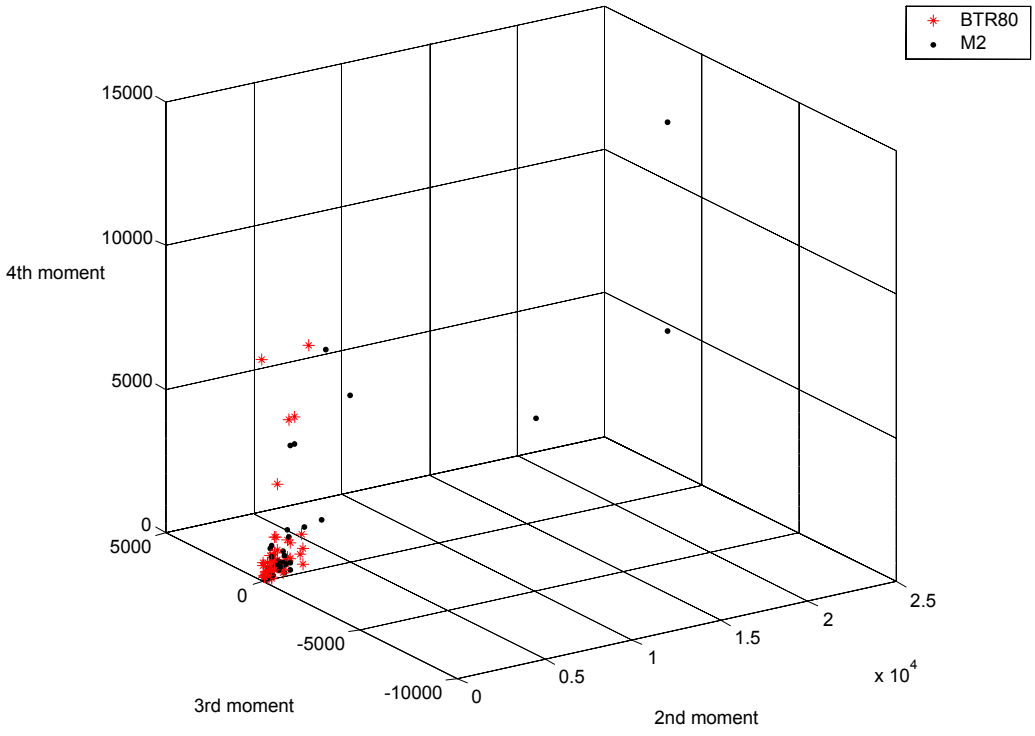


Figure 15. Three-Dimensional Comparison of Moment Feature Set

The moment feature set for a BTR-80 and an M2. With the exception of a few outliers, the majority of the data points are clustered closely and classification is difficult.

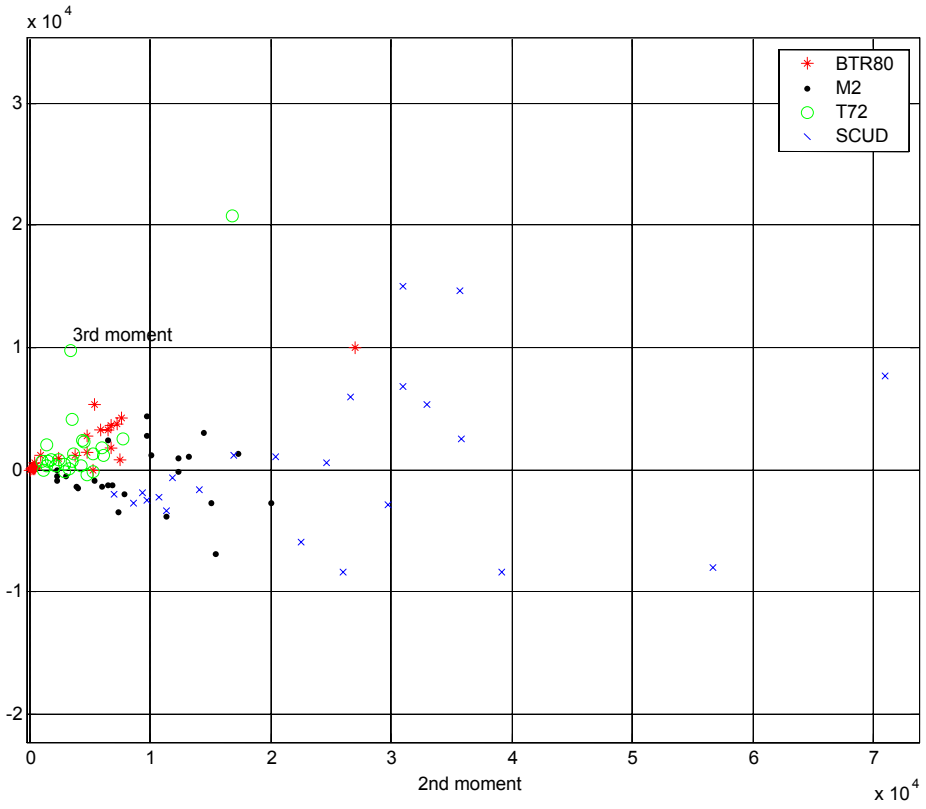


Figure 16. Four Target Moment Comparison

The third and second moment distributions for four targets. There is little noticeable distinction between three of the four targets.

3.4 Classification Training Matrix Development

Fisher Linear Discriminant classification projects data from a multi-dimensional space onto a line that maximizes separation between-classes and minimizes separation within-classes. Thus the scalar values of the classes along the line yield the greatest separation between them. The projected values may be obtained directly, without coordinate transformations that require finding eigenvalues and eigenvectors, by making a linear combination of the training points that approximates different constants for different training classes using least squares[2].

For a two-class classification, the feature matrix is

$$X = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} & \cdot & \cdot & \cdot & \alpha_{1n} \\ 1 & \alpha_{21} & \alpha_{22} & \cdot & \cdot & \cdot & \alpha_{2n} \\ & & & \cdot & & & \\ & & & \cdot & & & \\ 1 & \alpha_{N1} & \alpha_{N2} & \cdot & \cdot & \cdot & \alpha_{Nn} \\ 1 & \beta_{11} & \beta_{12} & \cdot & \cdot & \cdot & \beta_{1n} \\ 1 & \beta_{21} & \beta_{22} & \cdot & \cdot & \cdot & \beta_{2n} \\ & & & \cdot & & & \\ & & & \cdot & & & \\ 1 & \beta_{N1} & \beta_{N2} & \cdot & \cdot & \cdot & \beta_{Nn} \end{bmatrix}, \tag{32}$$

where α and β indicate two classes, the first subscript is sample number, and the second subscript is feature number.

Also, the two-class vector is

$$\mathbf{y} = \begin{bmatrix} -1_1 & \cdots & -1_N & 1_1 & \cdots & 1_N \end{bmatrix}^T. \tag{33}$$

The elements of this vector need not be assigned a ± 1 value, and are in fact modified significantly in the three-target classification scenario. The coefficient vector is

$$\mathbf{c} = \mathbf{X}^{-1}\mathbf{y}, \tag{34}$$

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_{n+1} \end{bmatrix} = \mathbf{X}^{-1}\mathbf{y} = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} & \cdot & \cdot & \cdot & \alpha_{1n} \\ 1 & \alpha_{21} & \alpha_{22} & \cdot & \cdot & \cdot & \alpha_{2n} \\ & & & \cdot & & & \\ & & & \cdot & & & \\ 1 & \alpha_{N1} & \alpha_{N2} & \cdot & \cdot & \cdot & \alpha_{Nn} \\ 1 & \beta_{11} & \beta_{12} & \cdot & \cdot & \cdot & \beta_{1n} \\ 1 & \beta_{21} & \beta_{22} & \cdot & \cdot & \cdot & \beta_{2n} \\ & & & \cdot & & & \\ & & & \cdot & & & \\ 1 & \beta_{N1} & \beta_{N2} & \cdot & \cdot & \cdot & \beta_{Nn} \end{bmatrix}^{-1} \begin{bmatrix} -1_1 \\ \cdot \\ \cdot \\ -1_N \\ 1_1 \\ \cdot \\ \cdot \\ 1_N \end{bmatrix}, \tag{35}$$

where there are (n + 1) coefficients because of the bias constant 1 added to the feature matrix. Solving for the coefficient vector completes the first part of classifier training. This vector is then multiplied by a feature matrix to obtain a training projection vector, the elements of which are the values of the training points projected onto the FLD line. Parzen windows analysis is performed on the projected points and a decision boundary is assigned, see Figure 17.

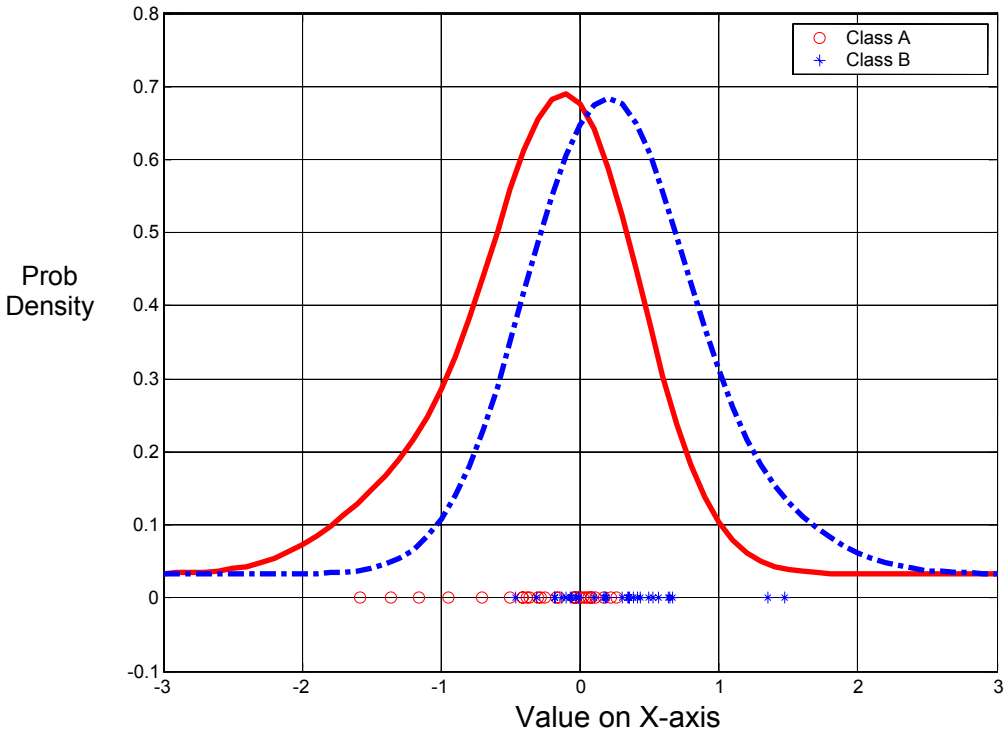


Figure 17. Example of a Training Matrix Projection Output

The training matrix output is a coefficient vector. This vector is multiplied by a feature matrix to obtain a vector of values projected onto the FLD line, which here is the x-axis. Gaussian Parzen window density estimate is used to produce a unimodal density from the projected points for each class. Zero would be an appropriate decision boundary.

3.5 Classification Testing Matrix Development

A testing matrix is configured as the training matrix and uses feature samples from the classes of interest. The inner product of the calculated coefficient vector and the test matrix **X** generates values projected onto the Fisher line. These values may separate the classes sufficiently.

3.6 Hypothetical Two-Target "Same" or "Different" Scenario

Two known target classes are examined, Target A and Target B. The feature set of moments from an HRR profile of Target A is compared to the moments from a second HRR profile of Target A or Target B. The first input HRR profile feature set is from Target A at a known aspect angle of θ degrees. The second input HRR profile feature set is from a known rotated aspect angle ($\theta + \beta$) degrees from one of two target classes. The linear classifier compares the two inputs and determines if the second HRR profile is the "same" as or "different" from the first HRR profile. Figure 18 illustrates this scenario.

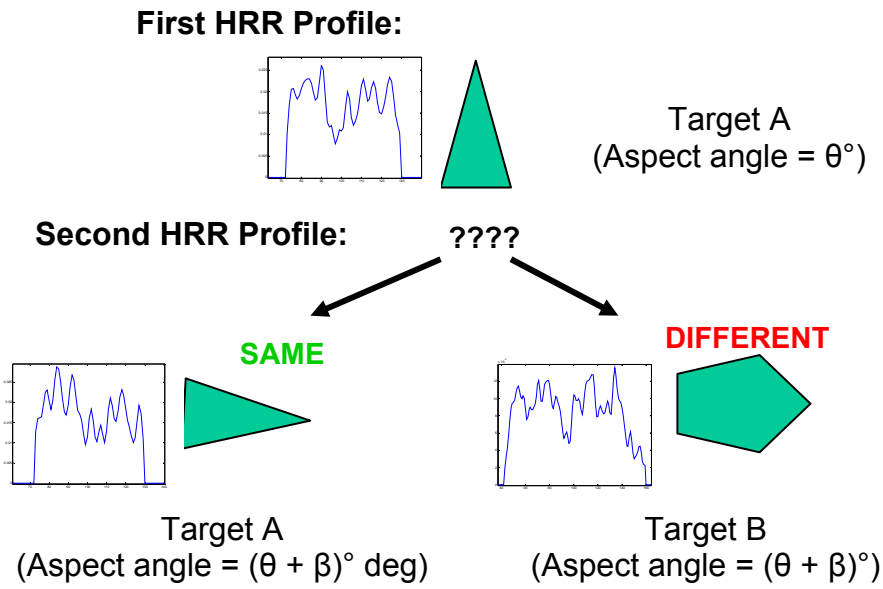


Figure 18. Hypothetical Scenario for Moving Target Classification

The first HRR profile is extracted from a target at a known aspect angle and then compared to an HRR profile from either a "same" or "different" target at a known rotated aspect angle.

Samples are collected from the MTFP data set HRR profiles and used to form two *teams*: "same" and "different". The "same" teams consist of a moment feature set extracted from a target HRR profile at an angle θ combined with another moment feature set extracted at an angle $(\theta + \beta)$ from the same target. Target A and B both have "same" teams formed. The "different" team contains a moment feature set extracted from a target HRR profile at an angle θ for Target A combined with a moment feature set extracted at an angle $(\theta + \beta)$ from Target B. Therefore, each team, whether "same" or "different", consists of three moments from the initial angle HRR profile and three moments from the rotated angle HRR profile, resulting in a six dimension moment feature space. Figure 19 shows the distribution of the sample data into the teams.

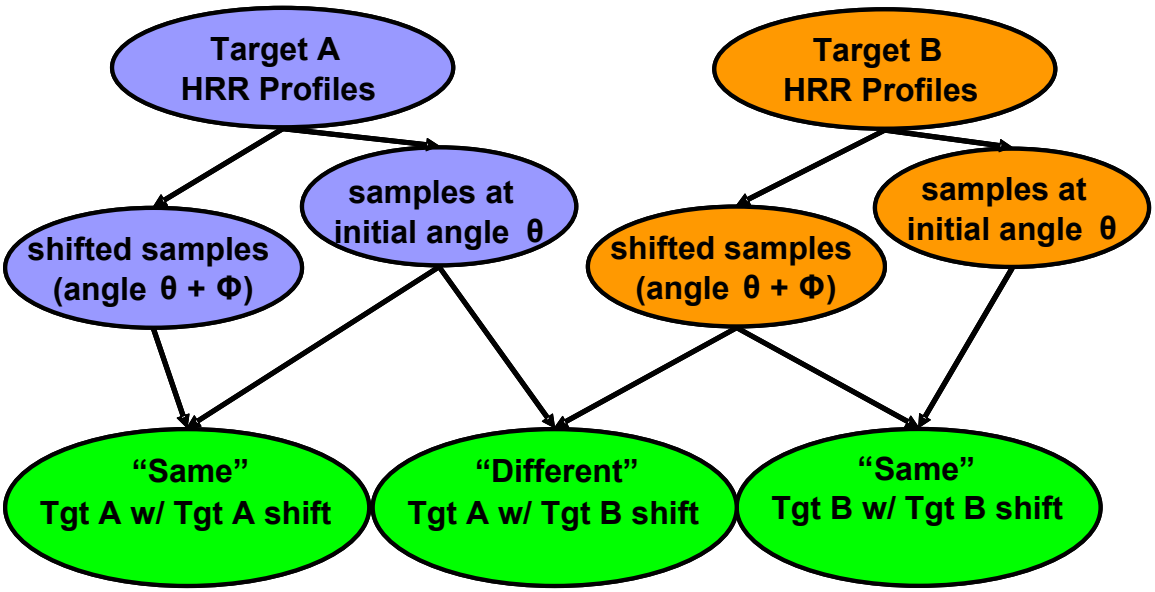


Figure 19. HRR Profile Sample Distribution Tree

HRR profiles are extracted from two targets and separated for training and testing. The individual "same" teams from each target consist of half the number of samples used to form the "different" team.

A "rule of thumb" for training data is to use 30 samples of each class of data. This "rule of thumb" is generally adhered to in this research, although increased sampling is used if it improves classifier performance.

For the "same" versus "different" data, the availability of the minimum required samples depends on the predetermined sample window θ, the target rotation angle β, and the target sets used for comparison. To avoid overlapping samples used for training and testing, sufficient samples are sought within the specified parameters for each target. If the predetermined sample window is not sufficient for the number of samples required, the sample window size is increased incrementally until the minimum sampling requirements are met.

For a two-class classification using the "same" and "different" teams, the feature matrix is

$$X = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \alpha_{15} & \alpha_{16} \\ & & & \vdots & & & \\ 1 & \alpha_{\frac{N}{2}1} & \alpha_{\frac{N}{2}2} & \alpha_{\frac{N}{2}3} & \alpha_{\frac{N}{2}4} & \alpha_{\frac{N}{2}5} & \alpha_{\frac{N}{2}6} \\ 1 & \beta_{11} & \beta_{11} & \beta_{11} & \beta_{11} & \beta_{11} & \beta_{11} \\ & & & \vdots & & & \\ 1 & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} \\ 1 & \alpha_{11} & \alpha_{12} & \alpha_{13} & \beta_{14} & \beta_{15} & \beta_{16} \\ & & & \vdots & & & \\ 1 & \alpha_{N1} & \alpha_{N2} & \alpha_{N3} & \beta_{N4} & \beta_{N5} & \beta_{N6} \end{bmatrix}, \quad (36)$$

where α and β represent Target A and B respectively. The first subscript is the sample value while the second subscript designates the HRR PDF moment value, and *N* is the number of minimum samples required (usually 30). The second through fourth moments from the initial angle HRR PDF are represented by the second subscripts 1, 2, and 3. The second through fourth moments from the rotated angle HRR PDF are represented by subscripts 4, 5, and 6. The upper half of the *feature matrix* contains "same" teams representing Target A and Target B. Only half of the minimum number of samples is necessary for each "same" team because two teams are used. A constant column with a value of "1" is added to the feature matrix to represent the y-intercept, increasing the

feature dimension by one (to seven). The "same" and "different" classes are pre-assigned a scalar value on the projected FLD line. Each class has a scalar value of negative 1 ("same" team) or a positive 1 ("different team") on the line, which forces each class to a separate point on the FLD line. Thus the projection vector is

$$\mathbf{y} = \left[-1_{\text{A–A}} \quad \cdots \quad -1_{\text{A–A}} \; -1_{\text{B–B}} \quad \cdots \quad -1_{\text{B–B}} \quad 1_{\text{A–B}} \quad \cdots \quad 1_{\text{A–B}} \right]^{\text{T}}, \tag{37}$$

and the coefficient vector is

$$\mathbf{c} = \mathbf{X}^{-1}\mathbf{y}, \tag{38}$$

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{bmatrix} = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \alpha_{15} & \alpha_{16} \\ & & & \vdots & & & \\ & & & \vdots & & & \\ 1 & \alpha_{\frac{N}{2}1} & \alpha_{\frac{N}{2}2} & \alpha_{\frac{N}{2}3} & \alpha_{\frac{N}{2}4} & \alpha_{\frac{N}{2}5} & \alpha_{\frac{N}{2}6} \\ 1 & \beta_{11} & \beta_{11} & \beta_{11} & \beta_{11} & \beta_{11} & \beta_{11} \\ & & & \vdots & & & \\ & & & \vdots & & & \\ 1 & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} & \beta_{\frac{N}{2}1} \\ 1 & \alpha_{11} & \alpha_{12} & \alpha_{13} & \beta_{14} & \beta_{15} & \beta_{16} \\ & & & \vdots & & & \\ & & & \vdots & & & \\ 1 & \alpha_{N1} & \alpha_{N2} & \alpha_{N3} & \beta_{N4} & \beta_{N5} & \beta_{N6} \end{bmatrix}^{-1} \begin{bmatrix} -1_{\text{A–A}} \\ \vdots \\ -1_{\text{A–A}} \\ -1_{\text{B–B}} \\ \vdots \\ -1_{\text{B–B}} \\ 1_{\text{A–B}} \\ \vdots \\ \vdots \\ 1_{\text{A–B}} \end{bmatrix} \tag{39}$$

3.7 Classifier Simulation Loop

The aspect angle window is the range of aspect angles from which samples are collected for any one training or test scenario. For example, 30 samples may be collected from a narrow sampling such as 0° to 10° aspect angle for both targets. The window may be large, such as a 0° to 180°, and its center is also considered. For example, given a 10° aspect angle window from which to extract data samples, the classifier increments through 10° aspect angle windows with the center of the window ranging from 1° to 356° in 5° increments. Classifier performance is evaluated for robustness at different window sizes and at all aspect angles around the target.

3.8 Two-Target and Multi-Target Scenarios

The classifier compares two and three targets with no a priori knowledge of target rotation. Training and testing is similar to that described in the "same" versus "different" scenario above.

IV. Results

4.1"Same" Versus "Different" Team Results

The "same" versus "different" team hypothesis is applied to a combination of six targets, two targets at a time: BTR-80, T-72, ZSU-23, ZIL-131, M-2, and SCUD Mobile Launcher. The 2S1 data is unreliable and is not used. First, the minimum number of samples for each target is initially set at 30 and subsequently raised to 45, 60, and then 90. The results obtained with greater than 45 samples remained consistent and only served to increase processing time. Therefore, this number is used to obtain all "same" versus "different" results. The "same" versus "different" hypothesis is as follows, the first HRR profile is from known Target A at aspect angle θ and the second HRR profile is from either Target A or Target B rotated at an aspect angle ($\theta + \beta$) from the original HRR profile. In training, the minimum number of samples from each target are collected from a window centered at initial angle θ, and samples are also collected for the rotated target angle ($\theta + \beta$). If necessary, the sample window automatically expands until the minimum number of samples desired are retrieved from the data set. The center of the window is incremented around the 360° aspect angle view of the targets to ensure full coverage, which validates classifier robustness at different aspect angles and relative rotations. To establish a baseline parameter, all two-target combinations are compared with a 0° target rotation β value. Figures 20-25 show some baseline results. A complete listing is in Appendix A.

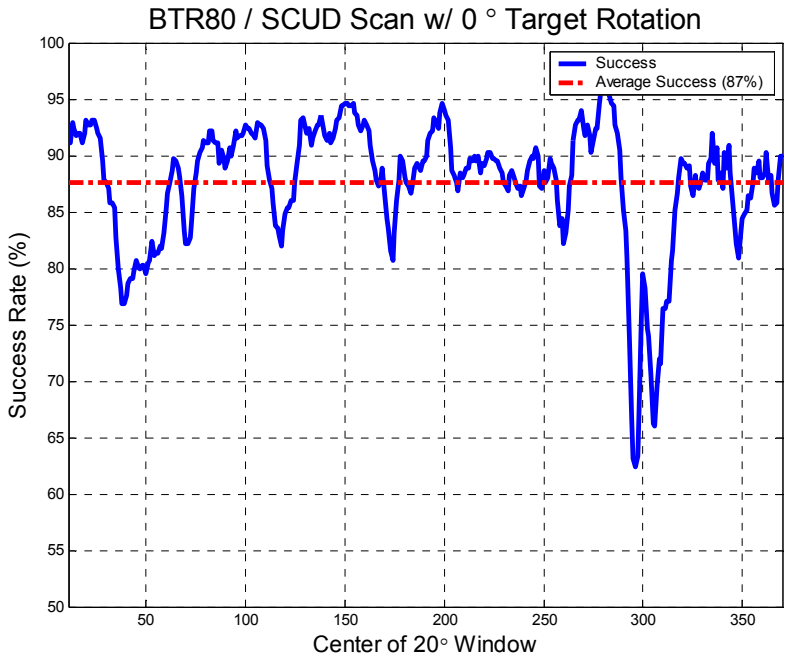


Figure 20. "Same" versus "Different" Comparison of BTR-80 and SCUD

"Same" versus "Different" comparison of BTR-80 and SCUD with 0° target rotation. This result baselines the best possible classifier performance using these two targets.

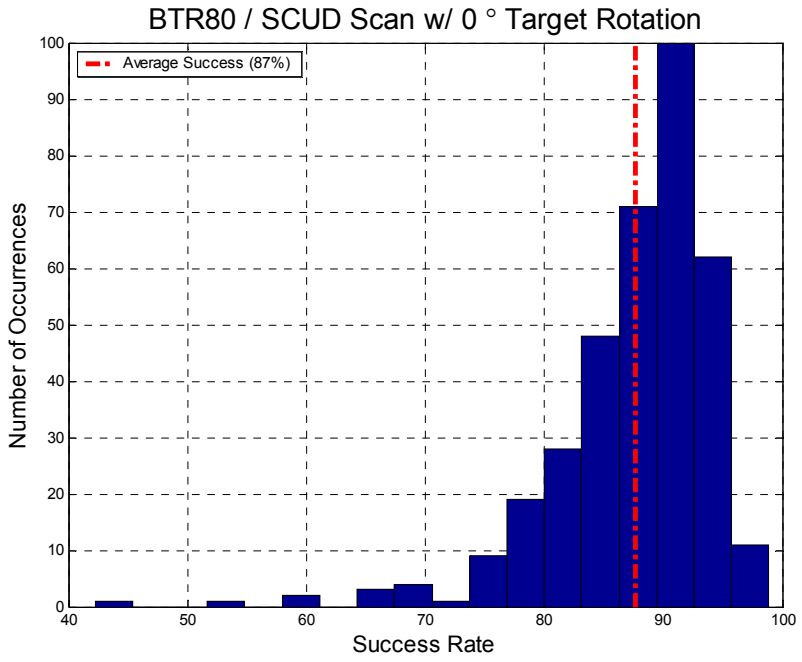


Figure 21. Histogram of BTR-80 and SCUD with 0° Target Rotation

The average 87% classification rate is highlighted in the dotted-dash line. This high success rate is expected for dissimilar targets.

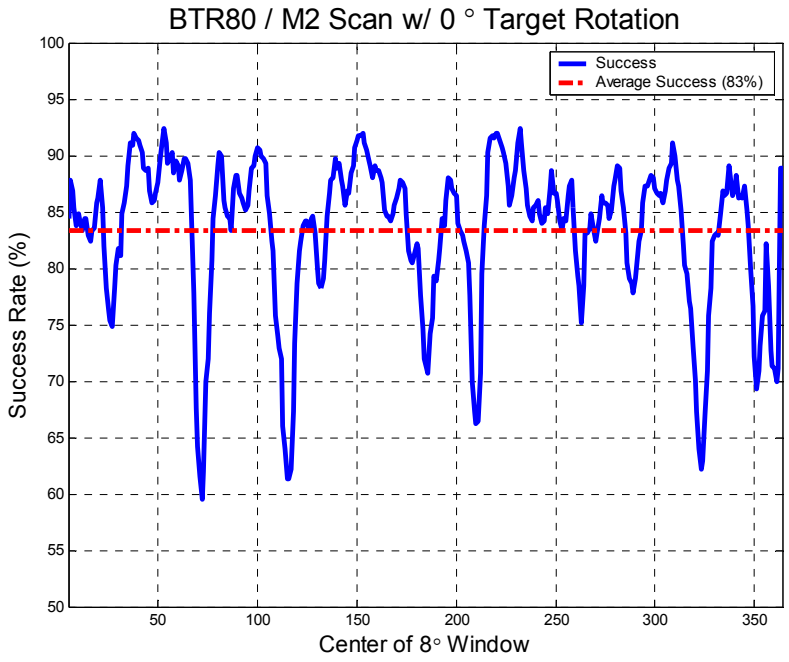


Figure 22. "Same" versus "Different" Comparison of BTR-80 and M2

"Same" versus "Different" comparison of BTR-80 and M2 with 0° target rotation. This result baselines the best possible classifier performance using these two targets.

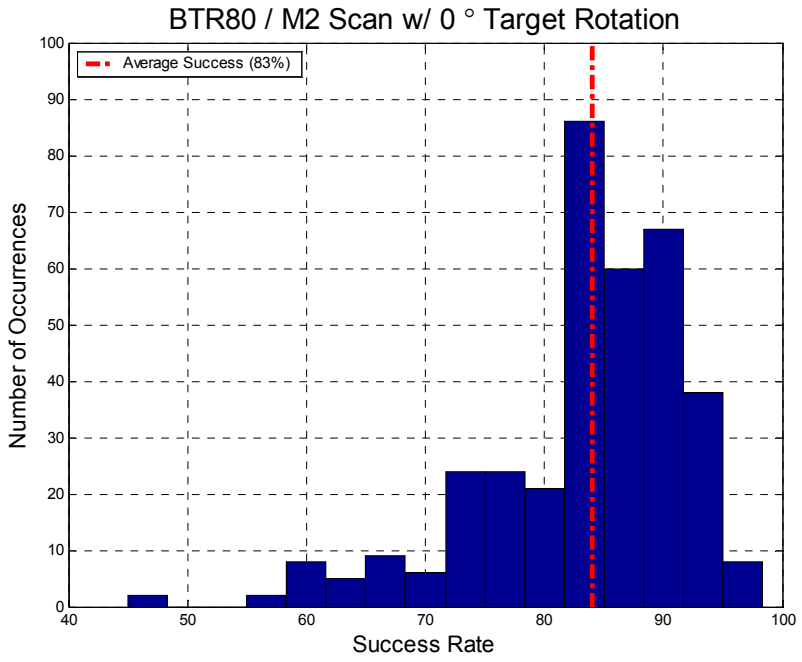


Figure 23. Histogram of BTR-80 and M2 with 0° Target Rotation

The average classifier 83% success rate shows good separation between targets on the FLD line.

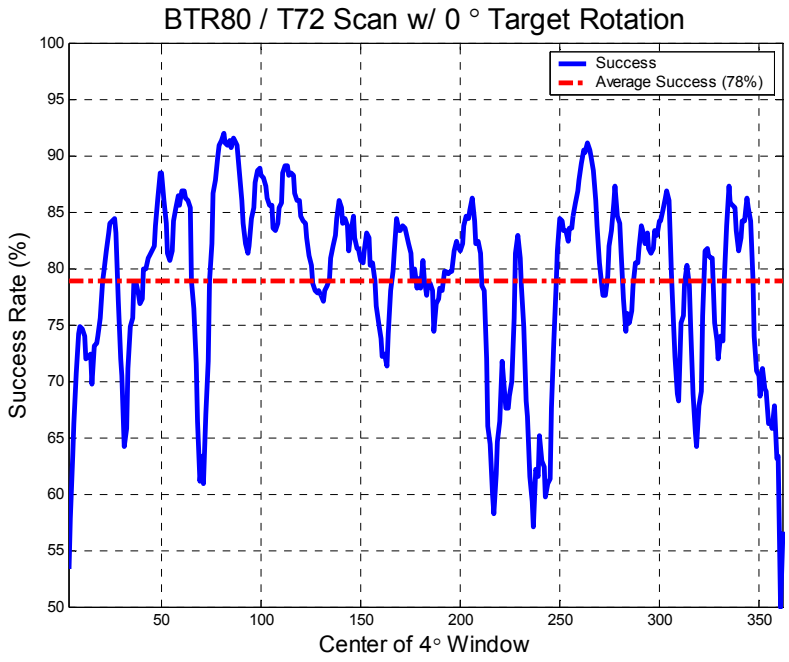


Figure 24. "Same" versus "Different" Comparison of BTR-80 and T-72

"Same" versus "Different" comparison of BTR-80 and T-72 with 0° target rotation. This result baselines the best possible classifier performance using these two targets.

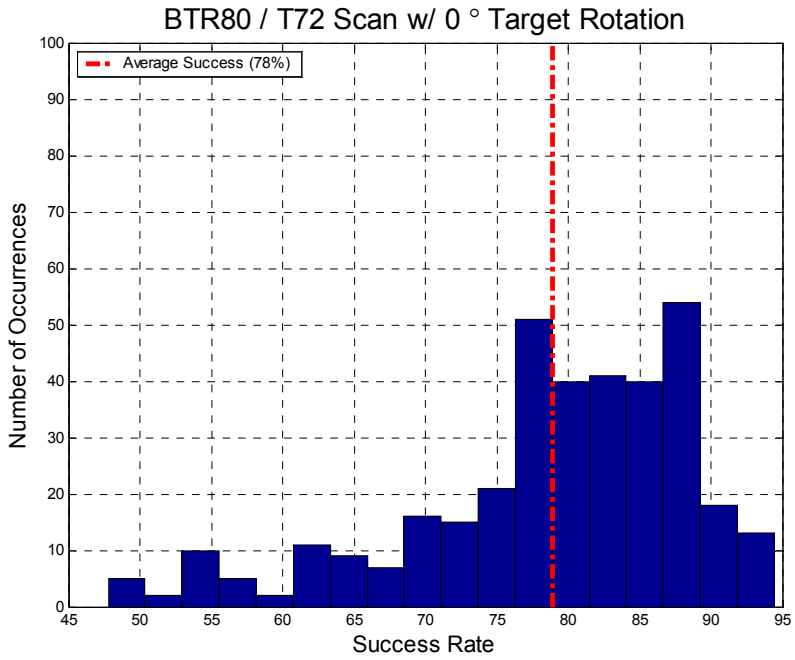


Figure 25. Histogram of BTR-80 and T-72 with 0° Target Rotation

The average classifier 78% success rate highlights the increased similarity between targets and the diminished classifier performance.

The best performance occurs when dissimilar targets are compared, such as the BTR-80 versus the SCUD (Figures 20 and 21). Classifier performance diminishes when similar targets are compared such as the BTR-80 and M-2 (Figures 22 and 23) or the BTR-80 and T-72 (Figures 24 and 25). There is significant difference in classifier performance when target rotation β is greater than 0°. However, overall classifier performance is independent of the actual rotation angle. Figures 26 and 27 show the classifier performance for the BTR-80 and T-72 at angle rotations β = 10, 45, 60, and 90°.

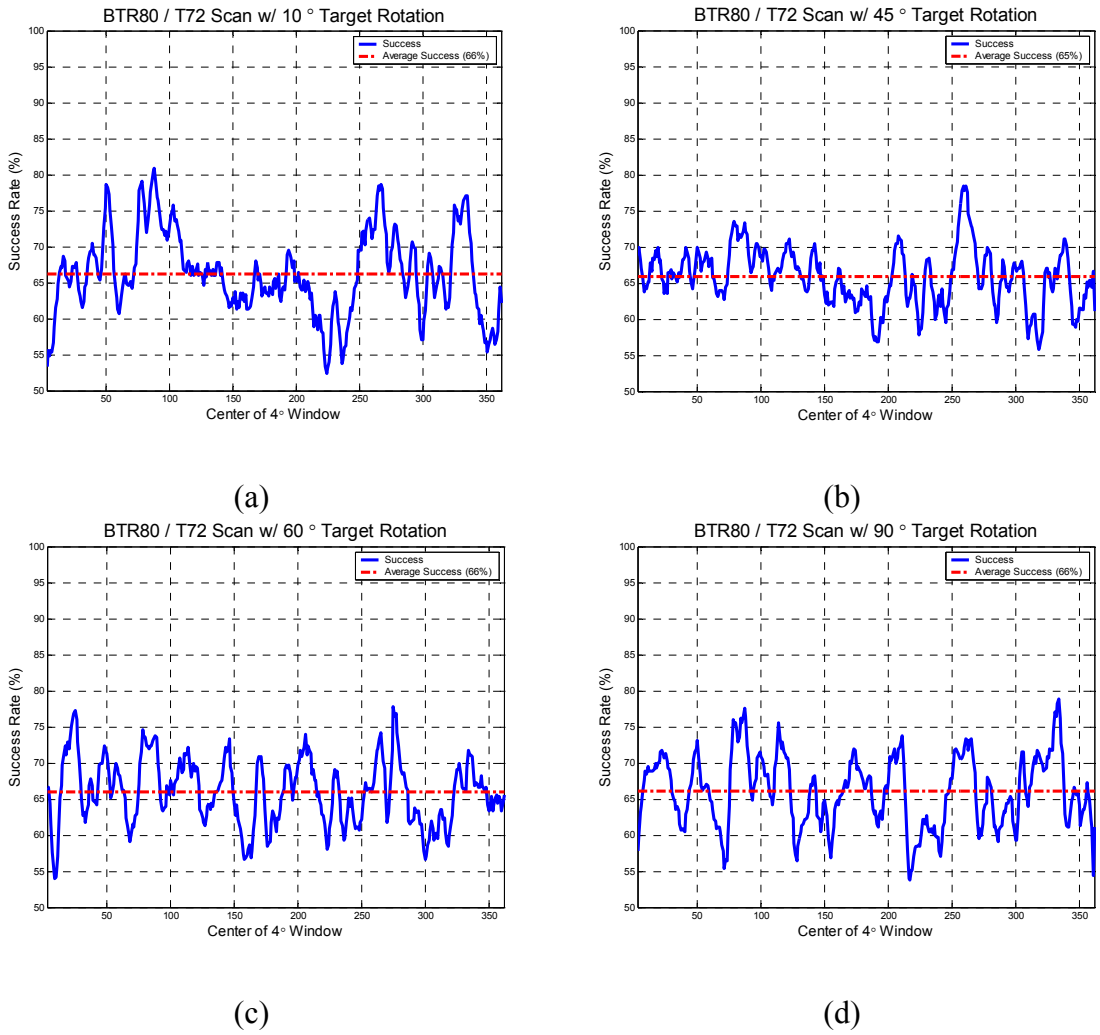


Figure 26. Comparison of BTR-80 and T-72 with Different Rotation Angles

Plots of BTR-80 and T-72 classification with target rotation equal to (a) 10°, (b) 45°, (c) 60°, and (d) 90°. Performance does not vary significantly with increased rotation because training and testing discriminate between pairs of moments at different angles.

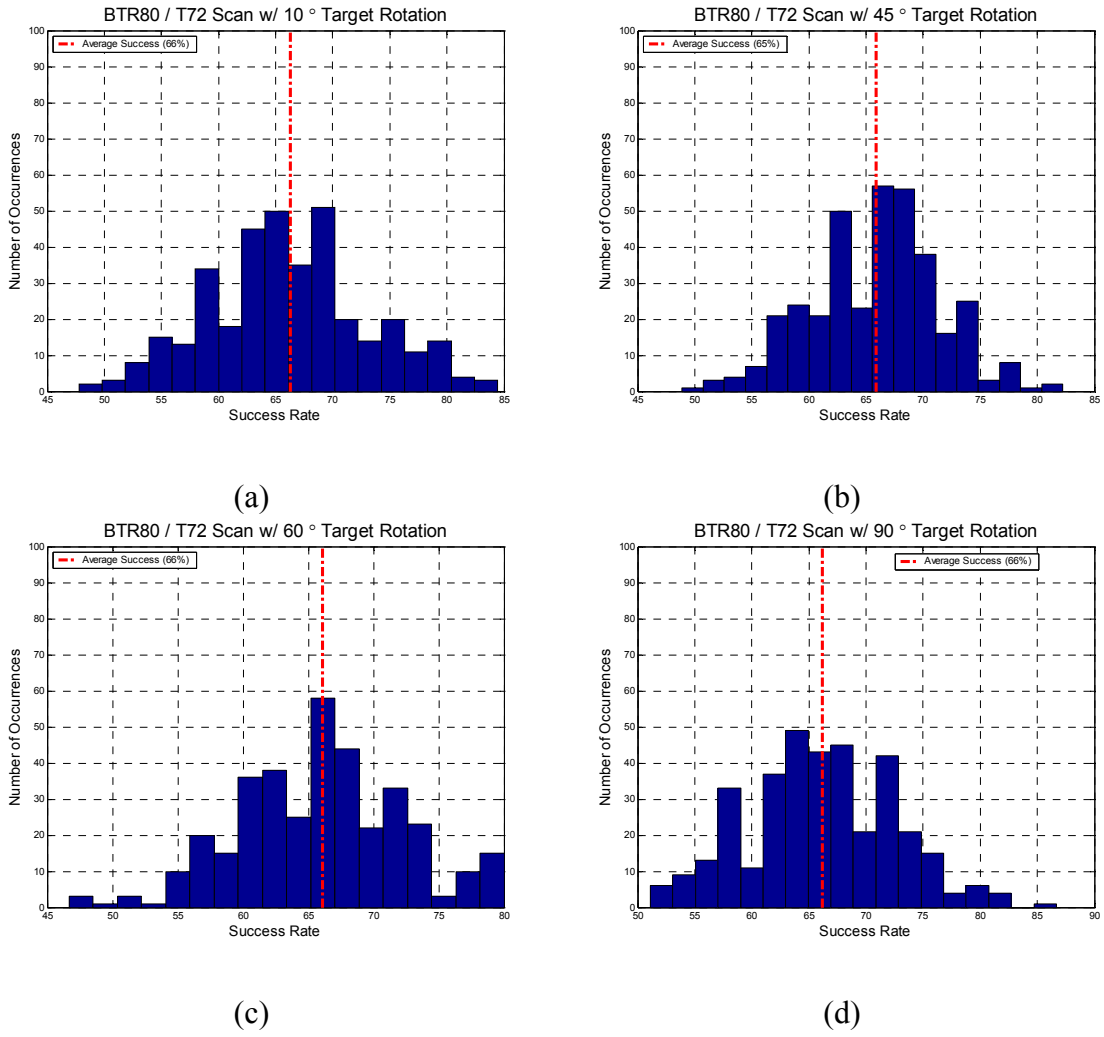


Figure 27. Histograms of BTR-80 and T-72 at Different Rotation Angles

Histograms of BTR-80 and T-72 classification simulation with target rotation equal to (a) 10°, (b) 45°, (c) 60°, and (d) 90°. Performance does not vary significantly with increased angle rotation.

The classifier performs well with target rotation introduced. However, this result simply reflects the accuracy of a priori knowledge introduced in training. The relative rotation angle does not change results significantly because the classifier discriminates between two sets of moments: those in the initial HRR profile aspect angle (θ) and those in the rotated HRR profile aspect angle ($\theta + \beta$). The FLD line is recalculated with every

simulation and only determines a discrimination for a "same" or "different" target for those samples. The results indicate that applying FLD analysis with HRR profile moments provides a solid foundation for classification. As expected, the classifier performs better for dissimilar target teams. However, the a priori knowledge available in this hypothesis is unrealistic in a typical scenario, and a more realistic classification scheme is necessary.

4.2 Two-Target Classification

This scheme classifies two targets with minimal a priori knowledge. The classifier uses moments from a two-target data set of HRR profiles within a predetermined sample aspect angle window. The center of the window is incremented through a 360° target aspect angle view. Training and testing is performed and results are recorded for every increment of the center of aspect angle window. The classifier then expands the sample aspect angle window and repeats the process. The "same" versus "different" matrix is simplified in that *n* is now three since an a priori rotation is no longer used and only the second, third, and fourth moments from each team are used for classification. Using a leave-one-out methodology during initial training and testing, the results reveal that using only two moments for classification yields approximately the same performance as using three moments. Further analysis shows that the third and fourth moments provide slightly better discrimination in some two-target comparisons. Thus the two-target *feature matrix is*

$$X = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} \\ 1 & \alpha_{21} & \alpha_{22} \\ & \vdots & \\ 1 & \alpha_{N1} & \alpha_{N2} \\ 1 & \beta_{11} & \beta_{12} \\ 1 & \beta_{21} & \beta_{22} \\ & \vdots & \\ 1 & \beta_{N1} & \beta_{N2} \end{bmatrix} \tag{40}$$

and the two-target coefficient vector is

$$c = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} \\ 1 & \alpha_{21} & \alpha_{22} \\ & \cdot & \\ & \cdot & \\ 1 & \alpha_{N1} & \alpha_{N2} \\ 1 & \beta_{11} & \beta_{12} \\ 1 & \beta_{21} & \beta_{22} \\ & \cdot & \\ & \cdot & \\ 1 & \beta_{N1} & \beta_{N2} \end{bmatrix}^{-1} \begin{bmatrix} -1_1 \\ -1_2 \\ \cdot \\ \cdot \\ -1_N \\ 1_1 \\ 1_2 \\ \cdot \\ \cdot \\ 1_N \end{bmatrix} \tag{41}$$

The two-target classification objective is to identify a coefficient vector in training that can be successfully applied to identify a single HRR profile from within a large aspect angle window. Ideally, there is a coefficient vector that can be applied to any HRR profile from a two-target class data set containing samples from a 360° aspect angle field-of-view and no a priori knowledge of a previous HRR profile is required. The results show that the FLD classification method has a 65% success rate in a two-target class environment, where HRR profiles are extracted from up to a 360° aspect angle window.

Classification training occurs similarly to the method used in the "same" versus "different" team hypothesis. Two targets are selected for comparison. The HRR profiles are selected from an aspect angle window that expands from 10° to 360° in 10° increments. For each successive expansion of the sample aspect angle window, the center of the sample window pivots around the 360° aspect angle view in 10° increments.

The coefficient vector is calculated for each pivot of the center of the sample window, and the mean results for each aspect angle window is calculated upon completion of the pivot. The training produces a mean coefficient matrix where the rows represent the respective aspect angle window coefficient vector sequentially from the smallest window to the largest. Thus, the output from two-target classification training is the coefficient matrix

$$C = \begin{bmatrix} c_{1-10} & c_{2-10} & c_{3-10} \\ c_{1-20} & c_{2-20} & c_{3-20} \\ & \cdot & \\ & \cdot & \\ c_{1-360} & c_{2-360} & c_{3-360} \end{bmatrix} \tag{42}$$

Testing is performed by extracting samples from the same aspect windows and pivots described above, forming a two-target test matrix, and obtaining projections on the FLD.

$$y = \begin{bmatrix} y_{11} \\ y_{12} \\ \cdot \\ \cdot \\ y_{1N} \\ y_{21} \\ y_{22} \\ \cdot \\ \cdot \\ y_{2N} \end{bmatrix} = Xc = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} \\ 1 & \alpha_{21} & \alpha_{22} \\ & \cdot & \\ & \cdot & \\ 1 & \alpha_{N1} & \alpha_{N2} \\ 1 & \beta_{11} & \beta_{12} \\ 1 & \beta_{21} & \beta_{22} \\ & \cdot & \\ & \cdot & \\ 1 & \beta_{N1} & \beta_{N2} \end{bmatrix} \begin{bmatrix} c_{1-360} \\ c_{2-360} \\ c_{3-360} \end{bmatrix}. \tag{43}$$

The coefficient vector from the 360° window should provide the greatest variability in classifying HRR profiles. Testing is accomplished with twice as many HRR profiles extracted from the MTFP data set as are used to perform training. The success rate for each increment around the 360° target aspect angle window is calculated and tabulated. The results are presented in two types of figures. The first type shows average classifier success rate with respect to the center of the aspect angle window. The second type shows average classifier success rate as the aspect angle window increases in size. Results not presented in this chapter are included in the Appendix.

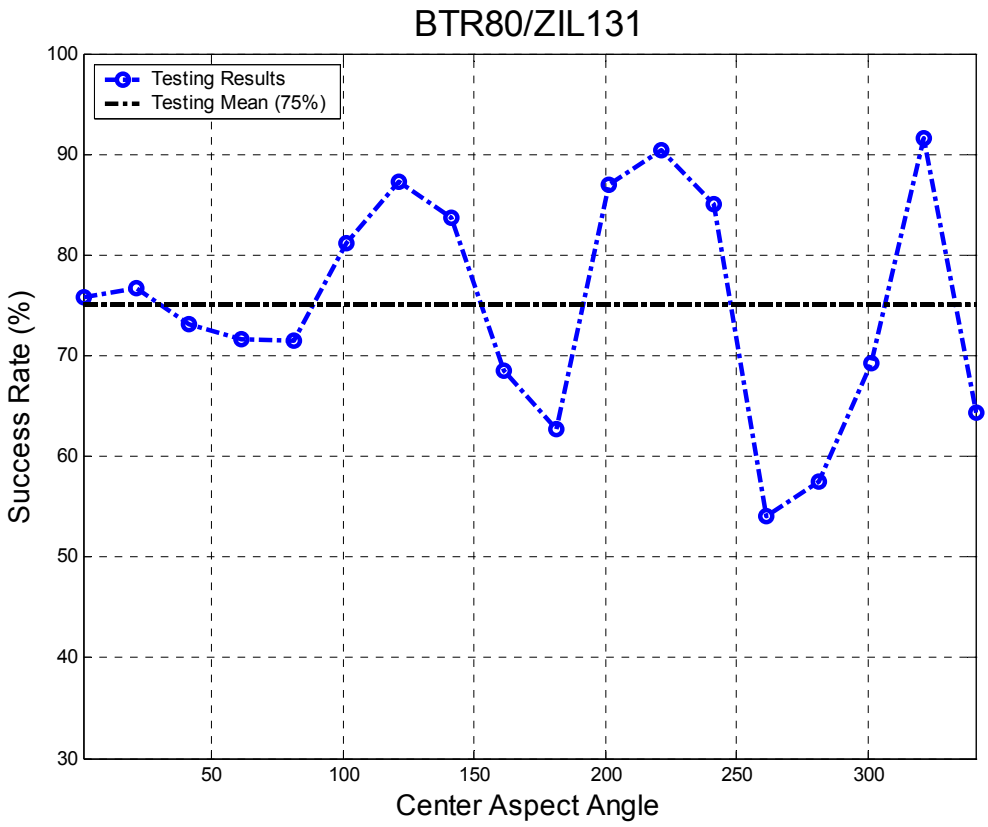


Figure 28. Aspect Angle View Classification Results for BTR-80 and ZIL-131

Classification results for comparison of BTR-80 and ZIL-131. The classifier pivots a sample window around a 360° aspect angle view of the targets. Results show overall 75% average successful classification and highlight areas where the targets look similar in moment feature space. In the example above, the targets are difficult to classify at a 260° aspect angle.

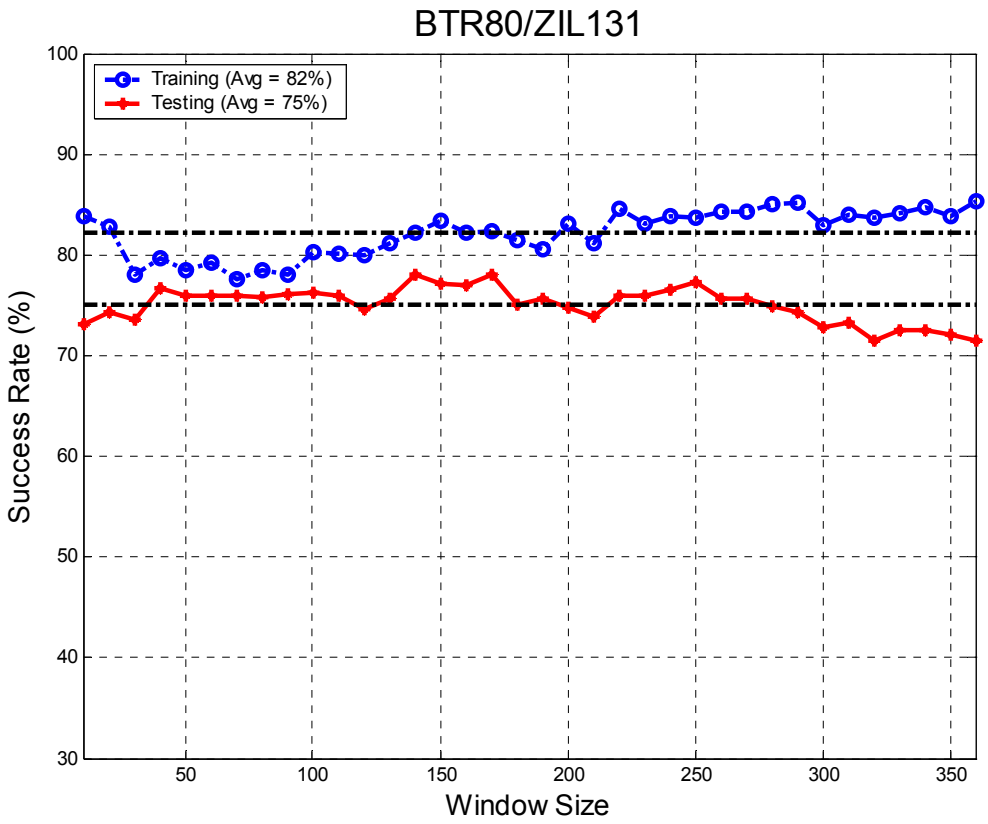


Figure 29. Sample Window Size Classification Results for BTR-80 and ZIL-131

Classification results for comparison of BTR-80 and ZIL-131. The classifier gradually expands the aspect angle window from which HRR profiles are selected for comparison. These results show that classification success remains steady for all window sizes.

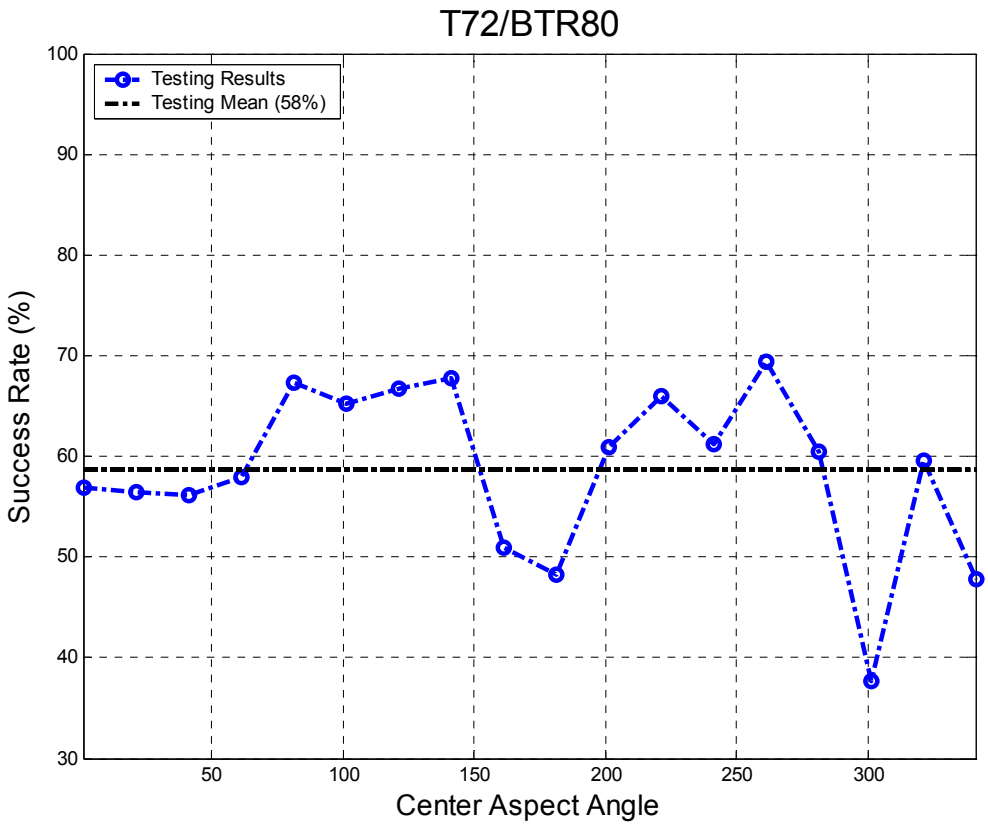


Figure 30. Aspect Angle View Classification Results for T-72 and BTR-80

Classification results for comparison of T-72 and a BTR-80. The classifier has an overall 58% average successful classification rate. The figure highlights classification difficulty at aspect angles of 300° and 360°.

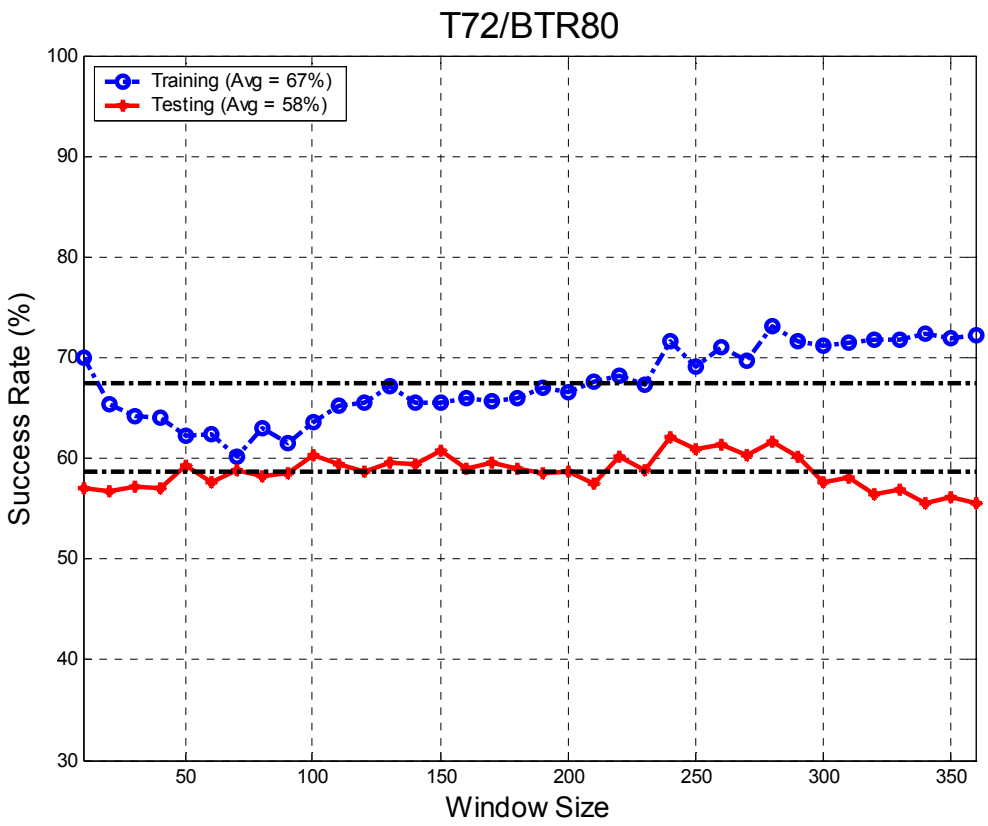


Figure 31. Sample Window Classification Results for T-72 and BTR-80

Classification results for comparison of T-72 and BTR-80. Once again, the classification remains steady for all window sizes. However, the targets are very similar in moment feature space and nearly indistinguishable from one another.

Target	Training	Testing
ZSU23	72.2	67.2
ZIL131	79.0	71.8
BTR80	75.2	67.8
SCUD	83.8	78.4
M2	74.6	65.4
T72	70.4	63.4
Average:	75.9	69.0

Table 1. Average Classification Results Using FLD in Two-Target Scenario

Two-target classification has a 69% average correct classification rate across all target combinations, which demonstrates the robustness and applicability of moment features for FLD classification.

4.3 Three-Target Classification

A three-target classification scheme is developed and tested. The three-target hypothesis follows the two-target classification method described above with one exception. For two-target classification, one target is forced to a -1 on the FLD line and the other target is forced to +1 on the FLD line. The natural pre-determined decision point becomes 0 by default. However, for a three-target classification, the classifier must select different projected values for each target, since symmetry about the origin of the FLD line is no longer available. Additionally, two decision points between the targets are now determined. The three-target classifier problem has a feature matrix

$$X = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} \\ & \cdot & \\ 1 & \alpha_{N1} & \alpha_{N2} \\ 1 & \beta_{11} & \beta_{12} \\ & \cdot & \\ 1 & \beta_{N1} & \beta_{N2} \\ 1 & \eta_{11} & \eta_{12} \\ & \cdot & \\ 1 & \eta_{N1} & \eta_{N2} \end{bmatrix}, \tag{44}$$

and the three-target coefficient vector is

$$c = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = X^{-1}y = \begin{bmatrix} 1 & \alpha_{11} & \alpha_{12} \\ & \cdot & \\ 1 & \alpha_{N1} & \alpha_{N2} \\ 1 & \beta_{11} & \beta_{12} \\ & \cdot & \\ 1 & \beta_{N1} & \beta_{N2} \\ 1 & \eta_{11} & \eta_{12} \\ & \cdot & \\ 1 & \eta_{N1} & \eta_{N2} \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \cdot \\ \rho_N \\ \varsigma_1 \\ \cdot \\ \varsigma_N \\ \omega_1 \\ \cdot \\ \omega_N \end{bmatrix}, \tag{45}$$

where ρ, ς, and ω are adjustable projection values.

This method is an extension of the two-target classifier with some notable exceptions. Three-target classification uses more samples per iteration and per target than two-target classification. Classifier performance improves significantly with increased samples for training and testing. On average, the three-target classifier is tested with over 135 samples compared with 45 samples for training. Also, the coefficient vector is applied to a set of samples near the samples used to calculate the coefficient vector. This procedure yields a theoretical performance measure and is applied until an 85% success rate for all three targets is achieved. This process is necessary because classifier performance during testing degrades by approximately 20% from the predicted training average.

The coefficient vector is calculated using the first set of moment features and with initial bias values of -1, 0, and 1 in the projection vector. If histograms indicate that classes are overlapping, the bias values are modified to reduce the overlap. Different bias

values are tested until class separation approaches a 85% success rate for all three targets during training.

Lessons learned from the two-target classifier are applied to three-target classification. Recall that two-target classifier testing results are realized using a constant 360° sample window. This wide sample window provides more robust results than smaller and more focused sample windows. Thus, training no longer attempts to determine coefficient vectors for sample windows less than 360°. As in the two-target classifier, the training and testing loops pivot the 360° sample window incrementally around the 360° target aspect angle view for data collection.

Three-target classifier results are presented below. Results not covered in this chapter are in the Appendix.

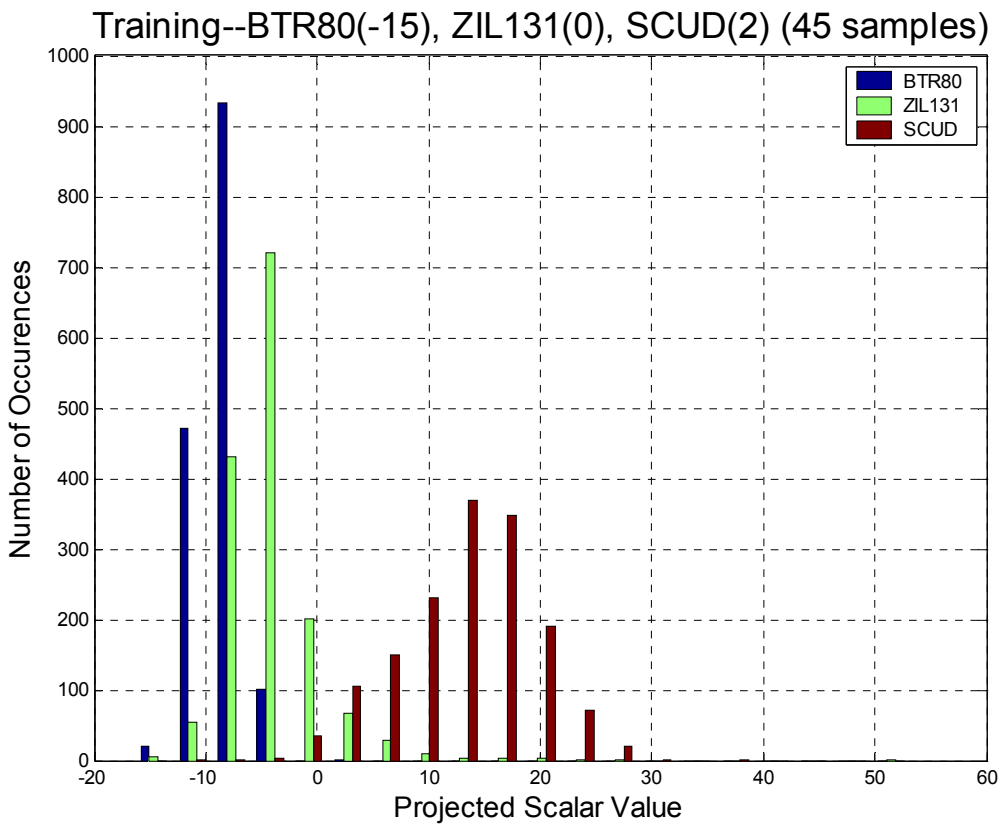


Figure 32. Three-Target Classification Training Histogram (BTR, ZIL, SCUD)

The bias values (number in parenthesis) used in this iteration of training yield good separation between the three target classes.

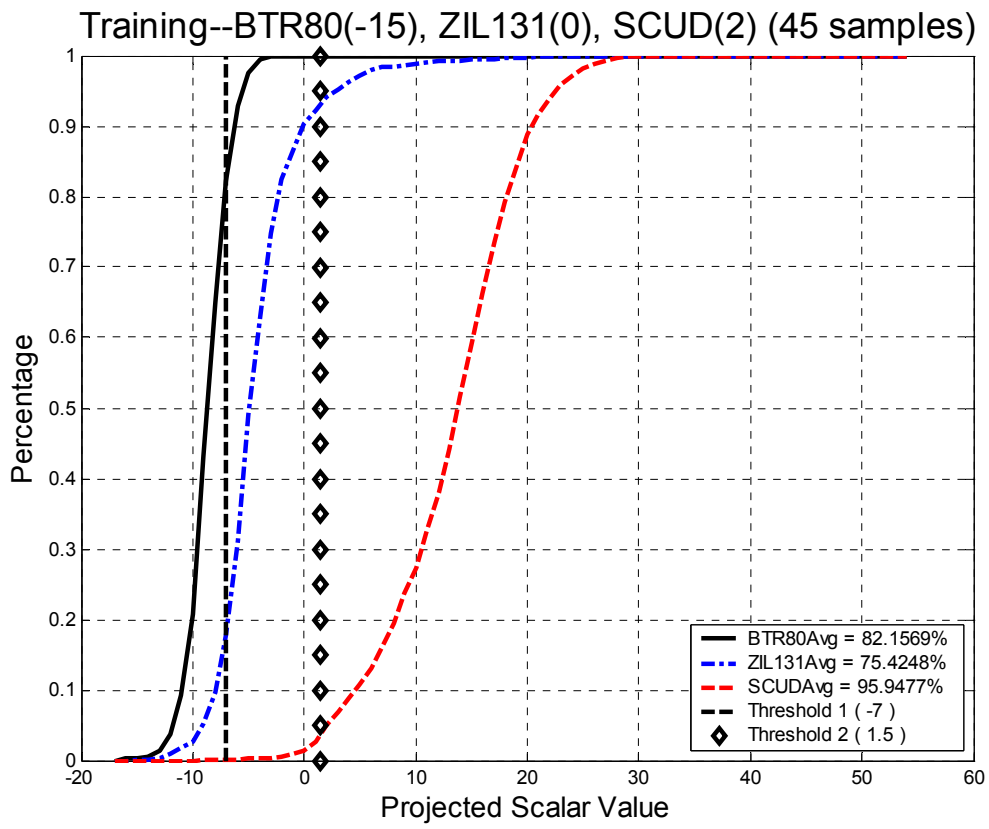


Figure 33. Three-Target Classification Training Plot (BTR, ZIL, SCUD)

These curves demonstrate the percentage of the target under the curve versus the projected value on x-axis. The bias values used for this iteration are the numbers in parenthesis after the respective target. The task is to determine threshold values that allow 85% successful classification.

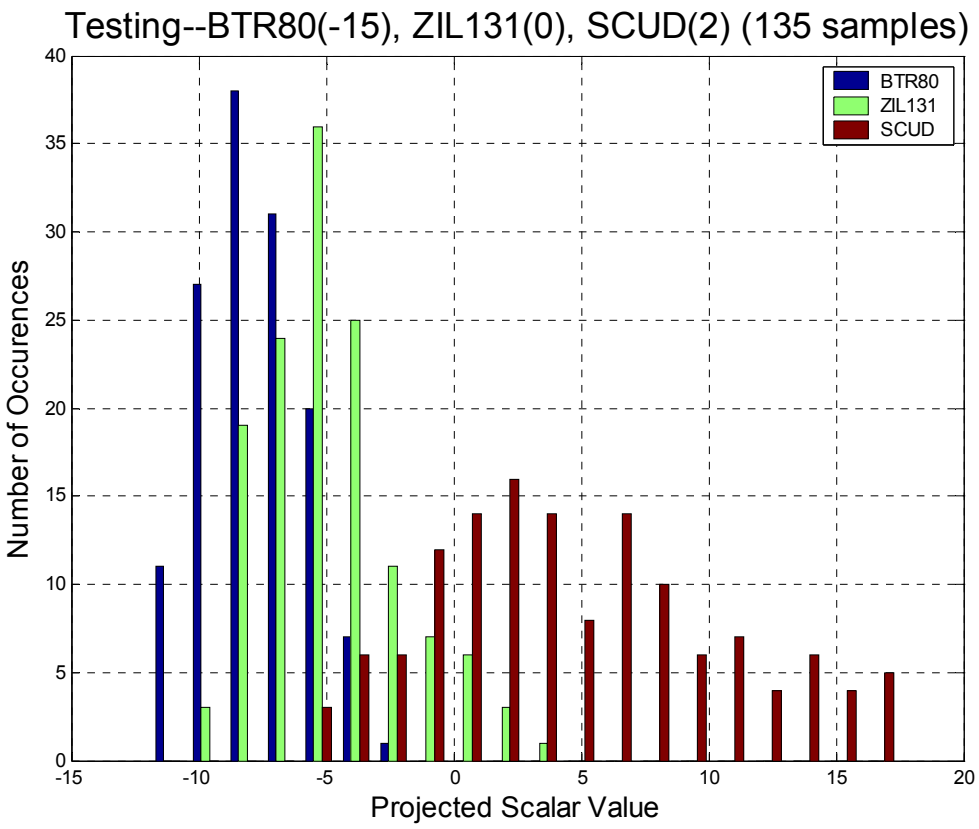


Figure 34. Three-Target Classification Testing Histogram (BTR, ZIL, SCUD)

The calculated coefficients determined during training yield the above result. Separation between the target classes diminishes slightly from the training result. Three times as many samples are used in testing to validate the training results.

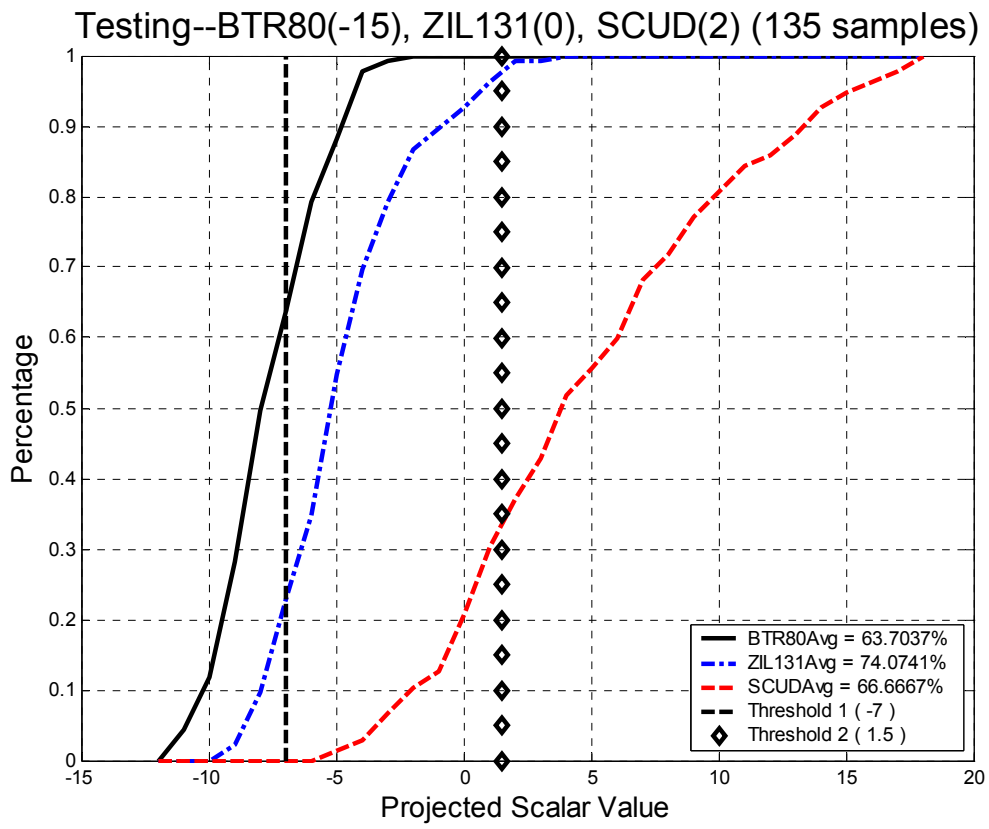


Figure 35. Three-Target Classification Testing Plot (BTR, ZIL, SCUD)

The threshold values determined during classifier training result in 68% successful classification in testing.

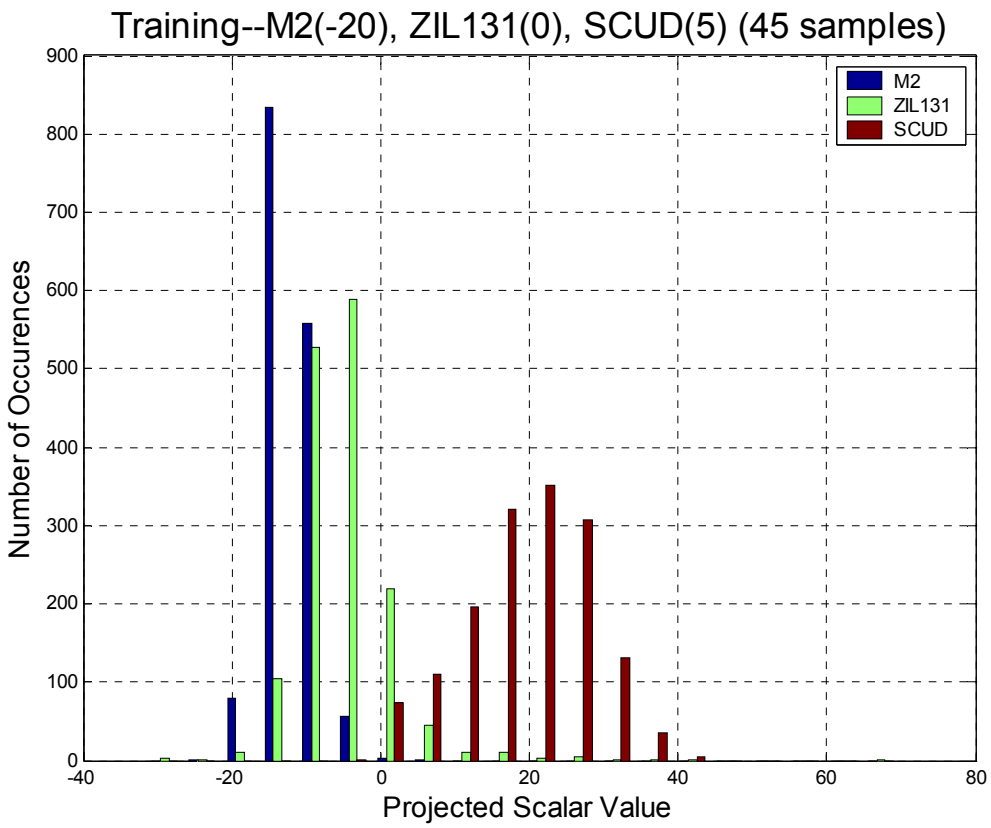


Figure 36. Three-Target Classification Training Histogram (M2, ZIL, SCUD)

Classifier training shows good separability with this three-target combination.

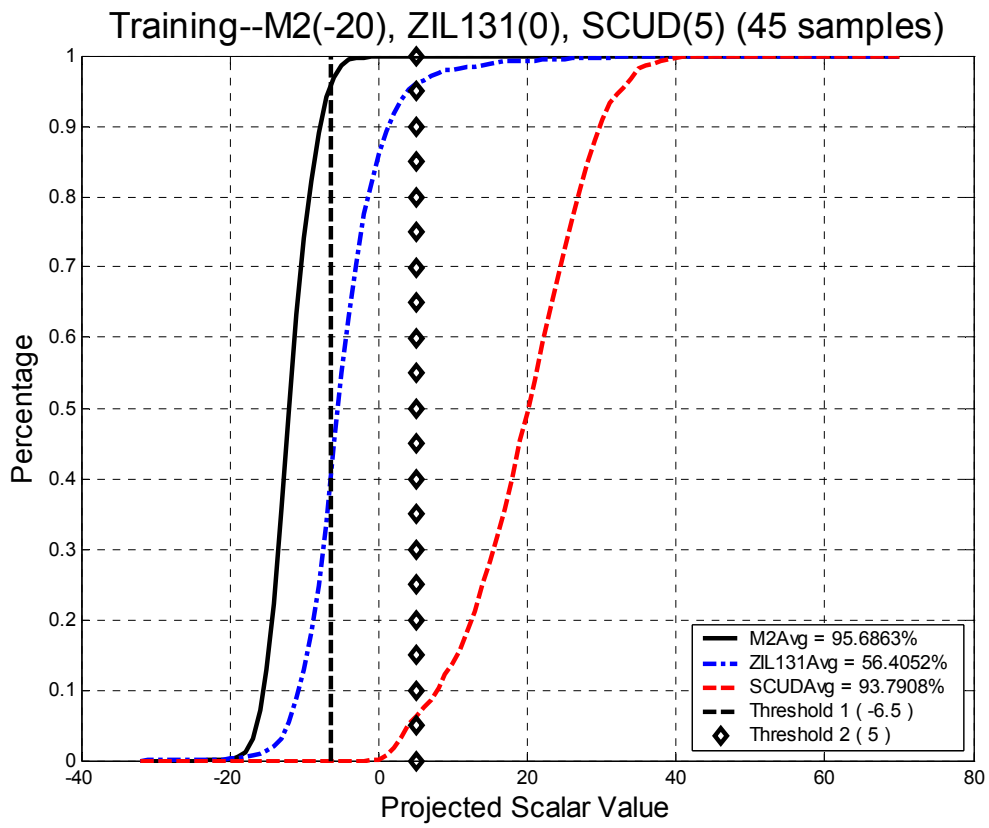


Figure 37. Three-Target Classification Training Plot (M2, ZIL, SCUD)

The threshold values determined during classifier training can be changed to adjust predicted success in testing.

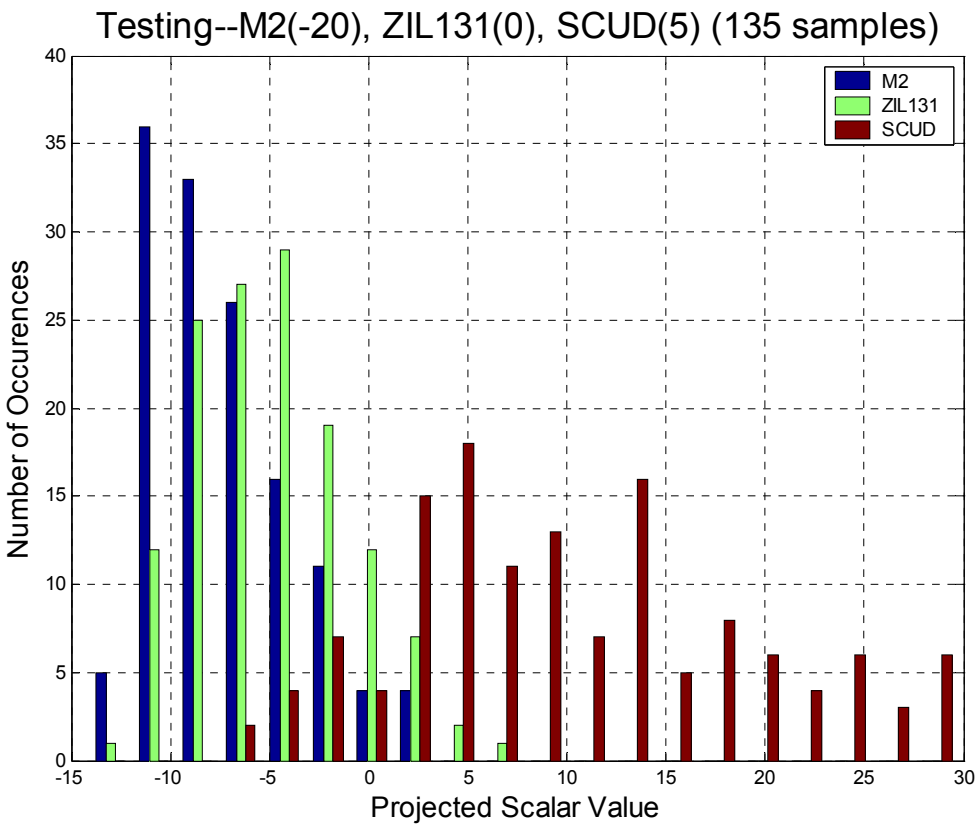


Figure 38. Three-Target Classification Testing Histogram (M2, ZIL, SCUD)

Classifier testing shows some overlap between M2 and ZIL-131; however, good threshold selection can differentiate the targets.

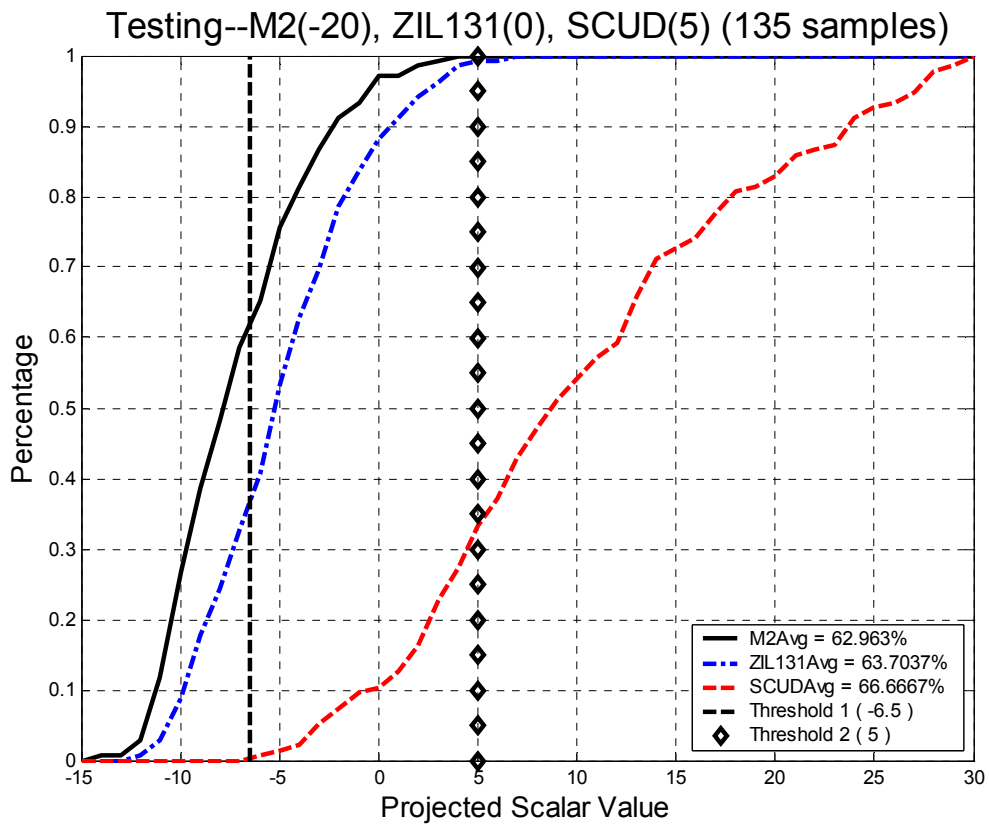


Figure 39. Three-Target Classification Testing Plot (M2, ZIL, SCUD)

The threshold values result in an average rate of approximately 64%, therefore, the classifier improves target identification significantly.

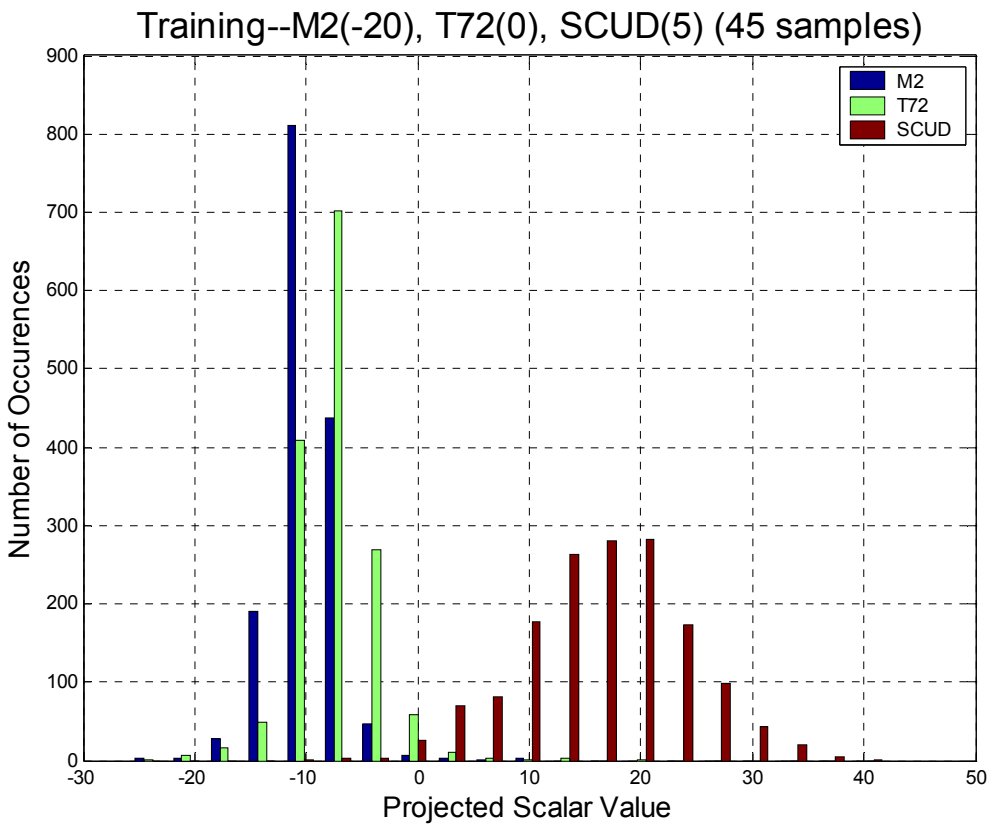


Figure 40. Three-Target Classification Training Histogram (M2, T72, SCUD)

Classifier training shows that separability is possible with this three-target combination.

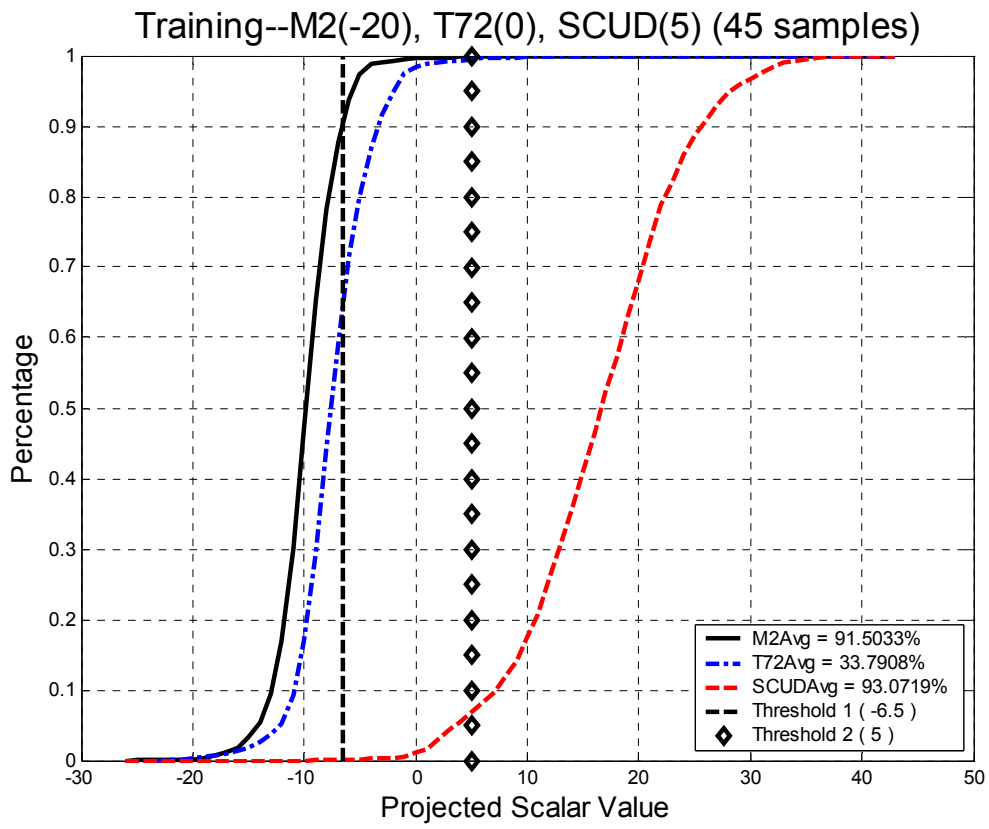


Figure 41. Three-Target Classification Training Plot (M2, T72, SCUD)

The threshold values determined during classifier training can be changed to adjust the predicted success rate in testing.

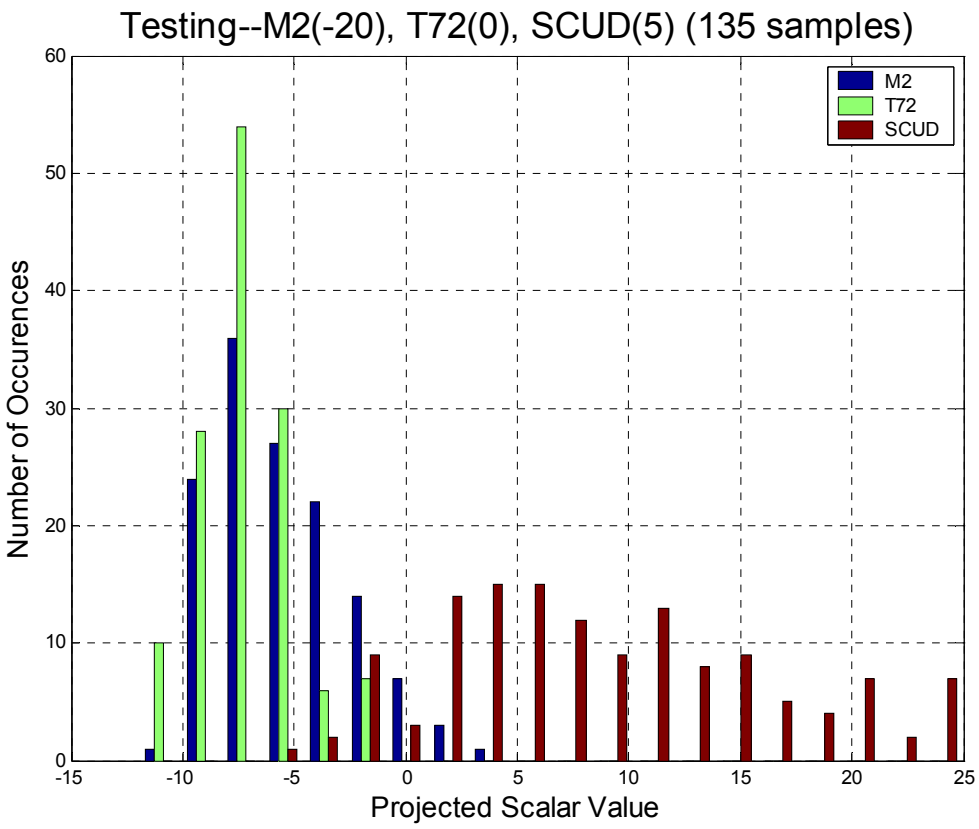


Figure 42. Three-Target Classification Testing Histogram (M2, T72, SCUD)

Classifier testing shows some overlap between M2 and T72, however, good threshold selection can differentiate the targets.

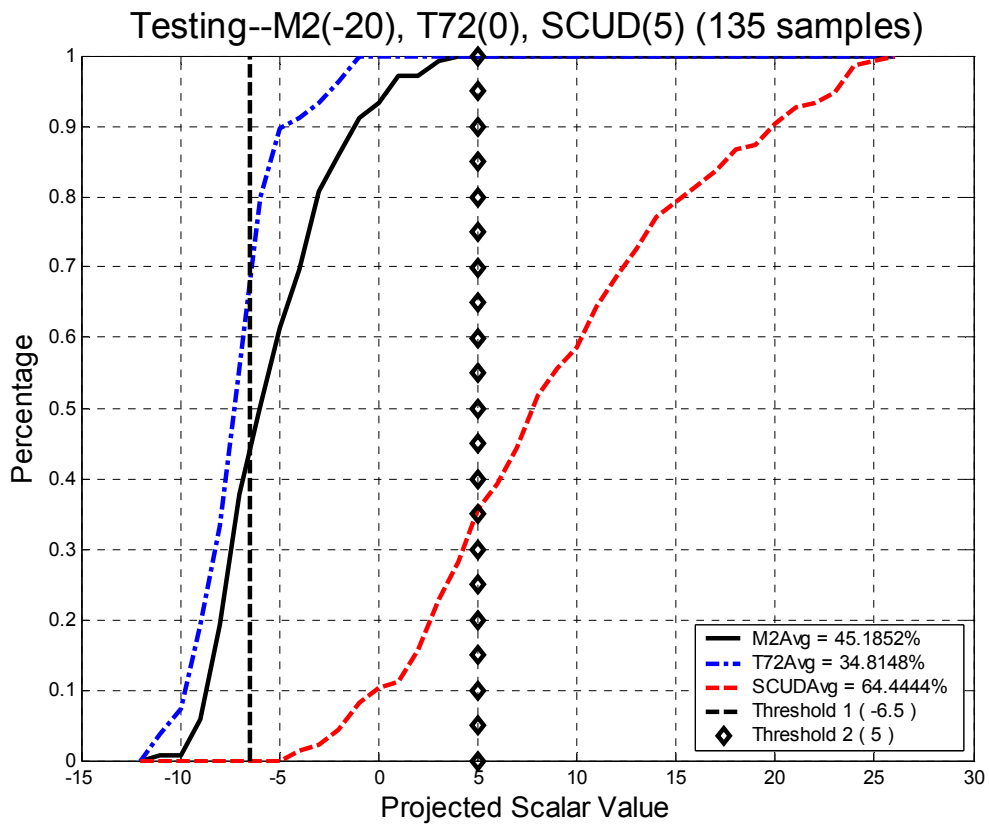


Figure 43. Three-Target Classification Testing Plot (M2, T72, SCUD)

The threshold values result in an average selection rate of under 50 %; however, random results yield a correct selection rate of about 33 %, therefore, the classifier does improve target identification.

Results show that FLD is a good method for classification of three targets. The adjustable biases and decision thresholds maintain the performance results found for two-target classification. The three-target classifier relies heavily on the targets chosen for classification and does not perform well with three very similar targets.

V. Conclusions

5.1 Summary

The research presented here evaluates the application of a Fisher Linear Discriminator for target classification using statistical moments extracted from High Range Resolution profiles. The research concludes that 1) HRR profile moments are suitable features for classification, 2) Fisher Linear Discrimination is a suitable method for target classification. Thus a simple approach to target classification yields good results if appropriate features are used.

5.1.1 Performance of HRR Profile Moments as Classification Features

The first four moments of HRR profiles offer significant and distinguishing information for use in classification. Pre-processing the moment data is critical to successful implementation. It is unlikely that using higher order moments will aid classification because performance results peak with only two moments.

5.1.2 Performance of Fisher Linear Discrimination as a Classification Method

Classification with Fisher Linear Discrimination is appropriate when used with greatly dissimilar targets, and it performs adequately for like targets. The overall method has reduced processing time compared to neural network techniques, and it provides a greatly simplified classification alternative. Unlike most neural net models, the FLD model avoids the credit assignment problem, i.e., determining the extent to which each input feature contributes to the classification output. Multinomial Pattern Matching (MPM) achieved a 74% correct classification rate for HRR profiles from three airborne targets

sampled from narrow five-degree-by-five-degree aspect and elevation angle windows [21]. FLD achieved comparable results for a full 360-degree aspect angle window with an unbound elevation angle input parameter.

5.1.3 Linear Classification Performance

An appropriately applied simple classification scheme can meet or exceed the performance of more complex target recognition methods. The FLD has a greater than 60% success rate for two and three target scenarios with a single HRR profile extracted from a 360° sample aspect angle window.

5.2 Recommendations for Future Work

5.2.1 Expand Fisher Linear Discrimination to Four-Target Scenario

Since performance remained steady for classification of both two and three target scenarios, it is natural to assume that a four-target scenario may also benefit from this classification technique and the use of moments as features.

5.2.2 Select a Different Feature Set for Use with FLD Classification

Wavelet transform features have been applied successfully in HRR profile classification. They may provide a more powerful feature set for use with Fisher Linear Discrimination.

5.2.3 Non-linear Classification Using Moment Features

Non-linear classification may further enhance the use of a moment feature set for target recognition. Of particular interest are non-parametric methods and clustering techniques.

5.2.4 Train on Synthetic Data and Test on Measured Data

Synthetic data is often used for classification training since measured data is difficult and expensive to obtain. Complicated classifiers are highly sensitive to even slight discrepancies between synthetic and measured data. Since moment features provide a course representations of HRR profiles with little regard to details, research extending the applicability of moment features for this purpose may be useful.

Bibliography

1. Alphatech Inc*, Moving Target Features and Phenomenology (MTFP) for Track Maintenance*. Final Report, 2003.

2 Bishop, C. M., *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

3. DeWitt. Mark R. *High Range Resolution Radar Target Identification Using the Prony Model and Hidden Markov Models*. MS thesis. Air Force Institute of Technology, 1992.

4. Duda, et al., Pattern Classification, 2nd Edition. New York: John Wiley & Sons, Inc. 2001.

5. Eisenbies, Christopher Lawrence. *Classification of Ultra High Range Resolution Radar Using Decision Boundary Analysis.* MS thesis, Air Force Institute of Technology, 1994.

6. Gutierrez-Osuna, Ricardo. Class handout, CS 790, Selected Topics in Computer Science. Russ School of Engineering, Wright State University OH. Dec 2002.

7. Hawley, Robert. Presentation, MRC's Data Object API. Mission Research Corporation. Oct 2003.

8. Huaitie, Xiao, et al. "On Notions of Information Transfer in VLSI CIrcuits." *Proceedings of the IEEE 1997 National Aerospace and Electronics Conference*. 1997.

9. HyperStat Online Contents. 5 Jan 2004 http://davidmlane.com/hyperstat/A53638.html

10. Mitchell, Richard A. and John J. Westerkamp. "A Statistical Feature Based Classifer for Robust High Range Resolution Target Identification." Submission to IEEE Transactions on Aerospace and Electronic Systems, Nov 1997.

11. Nelson, Dale E. *High Range Resolution Radar Target Classification: A Rough Set Approach.* Russ College of Engineering and Technology, Ohio University, June 2001.

12. Pham, Dzung Tri. *Applications of Unsupervised Clustering Algorithms to Aircraft Identification Using High Range Resolution Radar.* MS thesis, Air Force Institute of Technology, 1997.

13. Polikar, Robi. Class handout, ECE 504, Theory and Application of Pattern Recognition. School of Engineering, Rowan University, Glasboro, NJ, Dec 2003.

14. Shmitz, James. Presentation Slides. MTFPTM Data Collection and Management Overview. Veridian Engineering Dayton OH. 2003.

15. Songhua, He, et al., "Target Discrimination and Recognition Using High Resolution Range Features." *Proceedings of the IEEE 1992 National Aerospace and Electronics Conference.* 1992.

16. Stimson, George W. Introduction to Airborne Radar, 2nd Edition. Mendham NJ: SciTech Publishing Inc. 1998.

17. Tang, et al.,. "Comparison Study on High Resolution Radar Target Recognition." *Proceedings of the IEEE 1996 National Aerospace and Electronics Conference.* 1996.

18. Tang, et al., "High Resolution Radar Detection Based on Fractal Dimension." *Proceedings of the IEEE 1996 National Aerospace and Electronics Conference.* 1996.

19. Ulug, et al. "Efficient ATR Using Compression." *IEEE Transactions on Aerospace and Electronic Systems.* 1997.

20. Zahirniak, Daniel R. Characterization of Radar Signals Using Neural Networks. MS thesis, AFIT/GE/ENG/90D-69, School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH. Dec 1990.

21. Zumwalt, Michel P. *Robust High Range Resolution Radar for Target Classification*. MS thesis, AFIT/GE/ENG/00M-19, School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2000.

22. Zun, Zhang. et al. "Aircraft Target Recognition Using Adaptive Time-Delay Nueral Network." *Proceedings of the IEE 1997 National Aerospace and Electronics Conference.* 1997.

Appendix-A

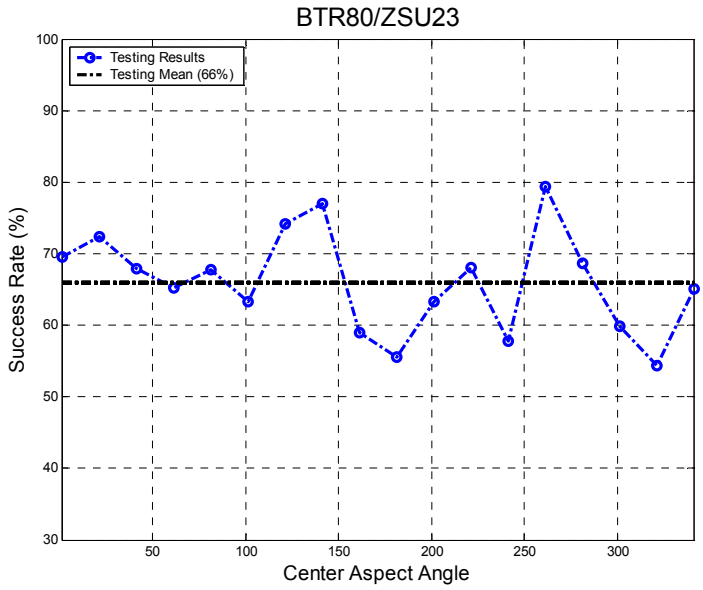


Figure 44. Aspect Angle View Classification Results for BTR-80 and ZSU-23

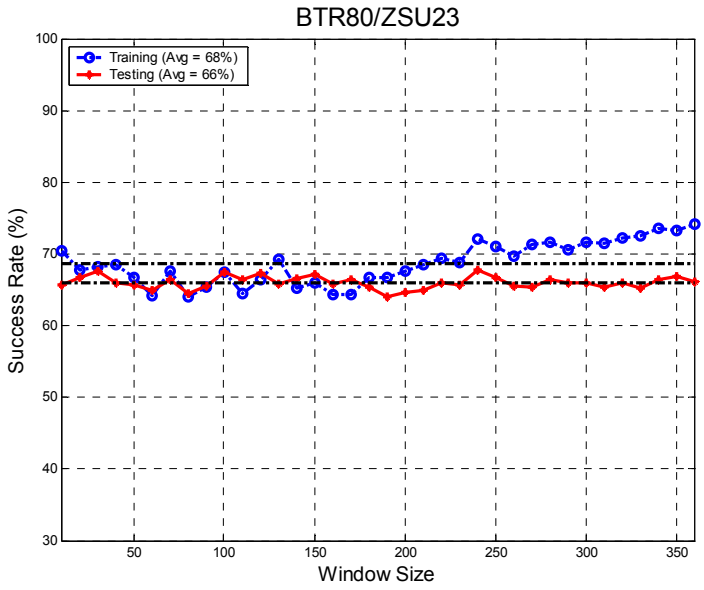


Figure 45. Sample Window Classification Results for BTR-80 and ZSU-23

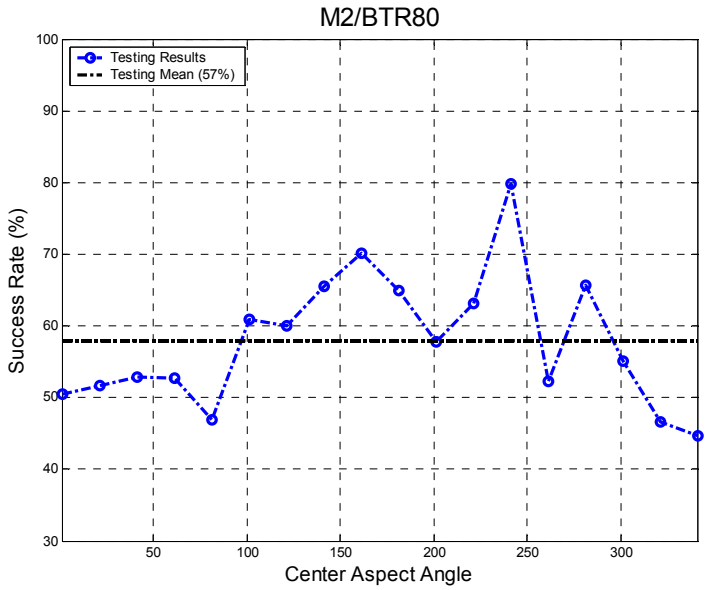


Figure 46. Aspect Angle View Classification Results for M-2 and BTR-80

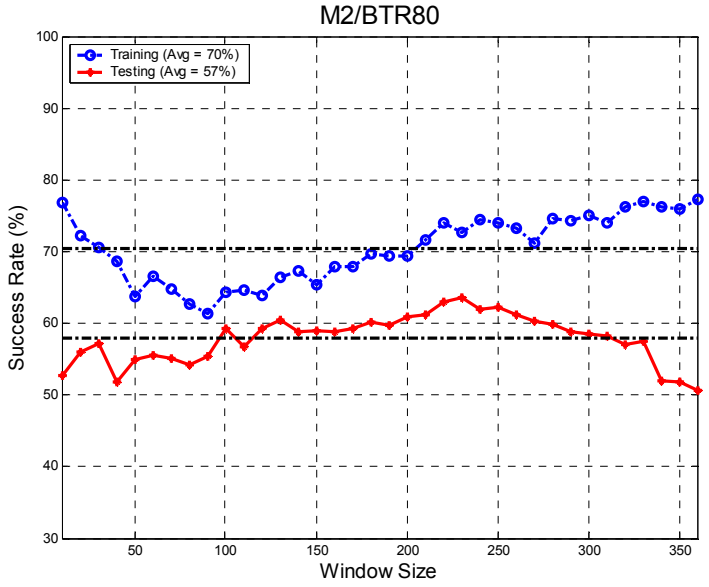


Figure 47. Sample Window Classification Results for M-2 and BTR-80

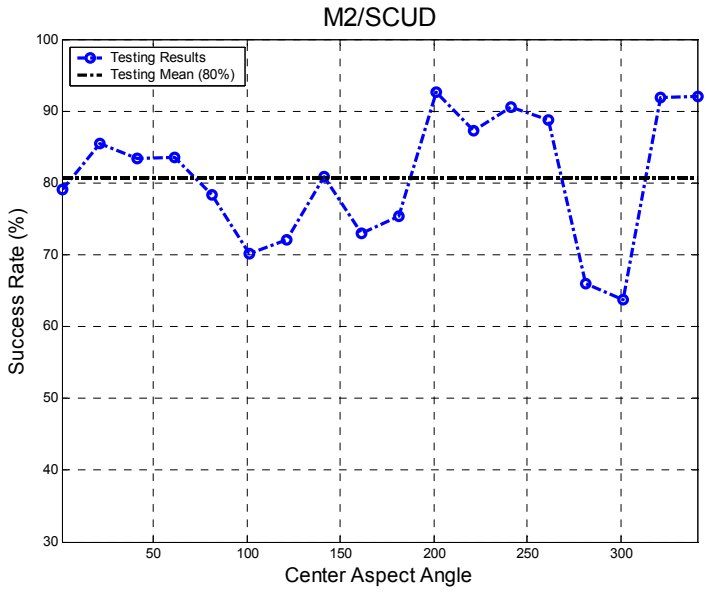


Figure 48. Aspect Angle View Classification Results for M-2 and SCUD

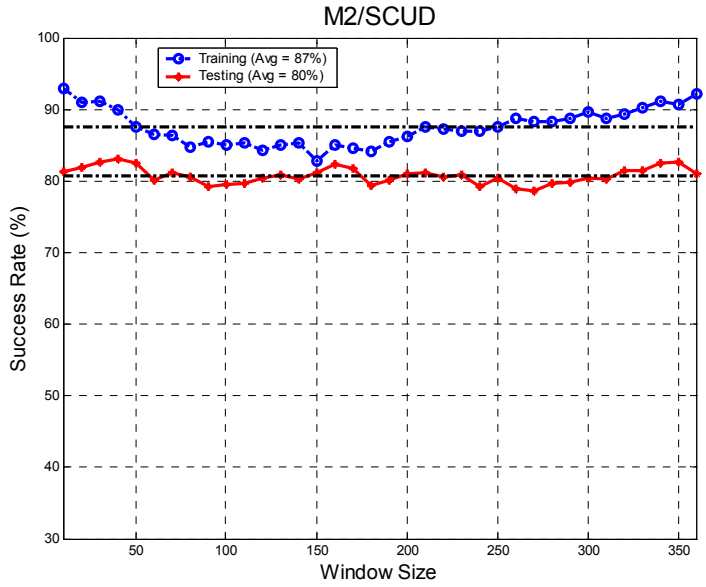


Figure 49. Sample Window Classification Results for M-2 and SCUD

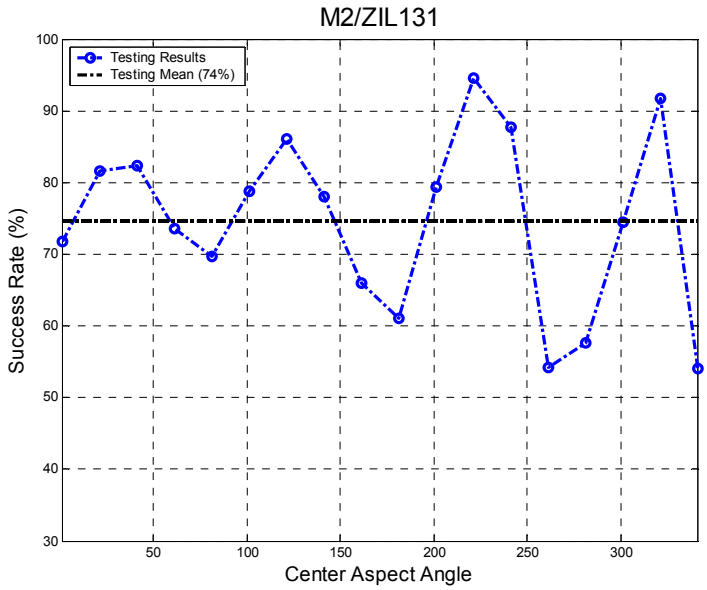


Figure 50. Aspect Angle View Classification Results for M-2 and ZIL-131

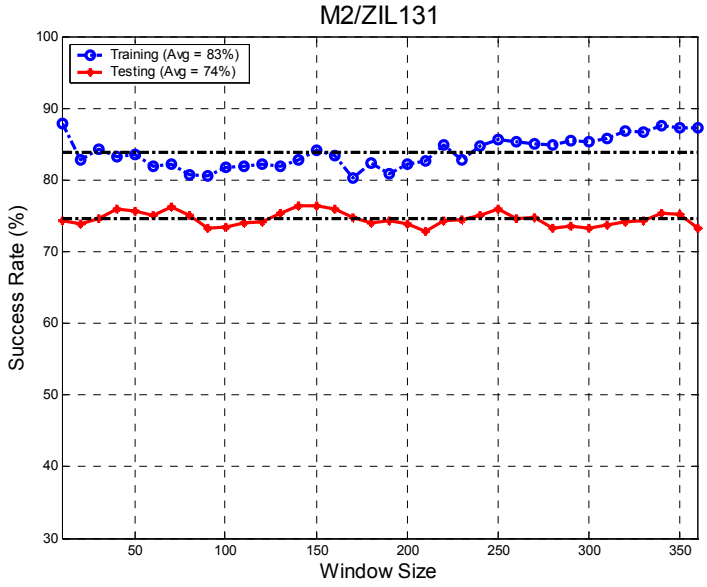


Figure 51. Sample Window Classification Results for M-2 and ZIL-131

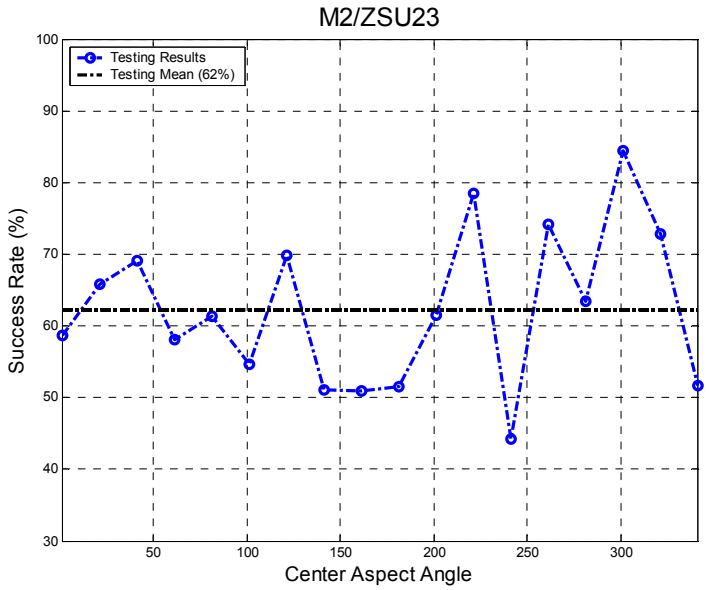


Figure 52. Aspect Angle View Classification Results for M-2 and ZSU-23

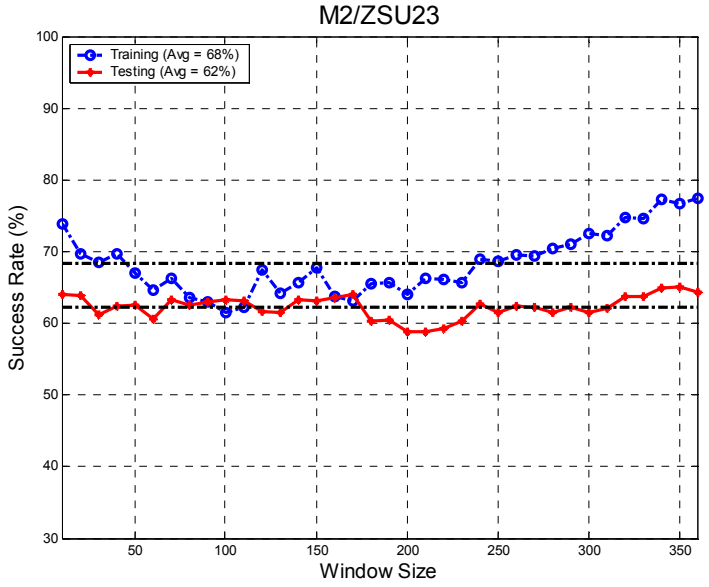


Figure 53. Sample Window Classification Results for M-2 and ZSU-23

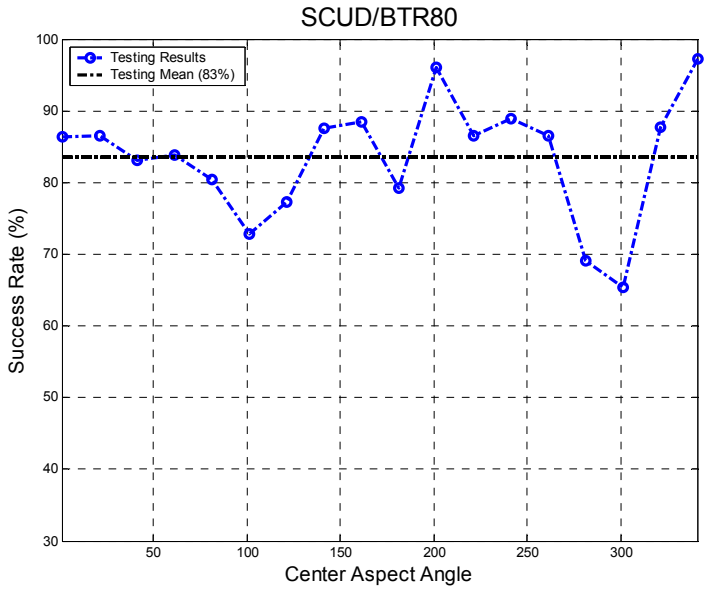


Figure 54. Aspect Angle View Classification Results for SCUD and BTR-80

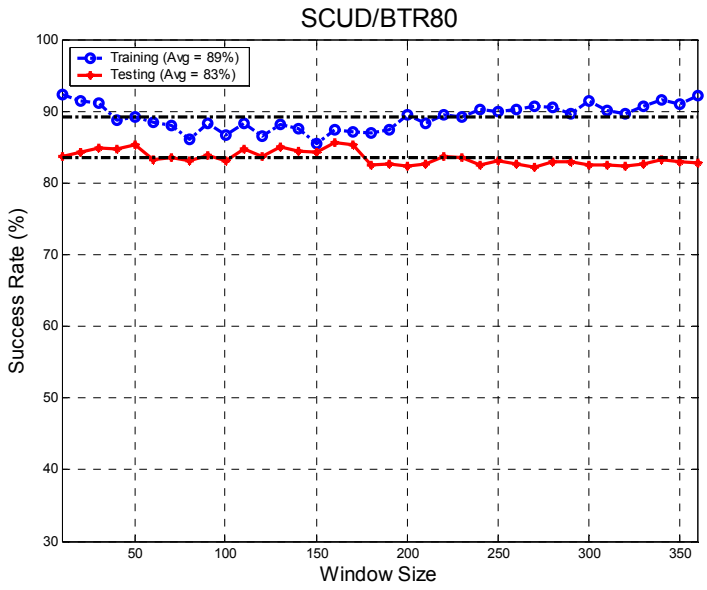


Figure 55. Sample Window Classification Results for SCUD and BTR-80

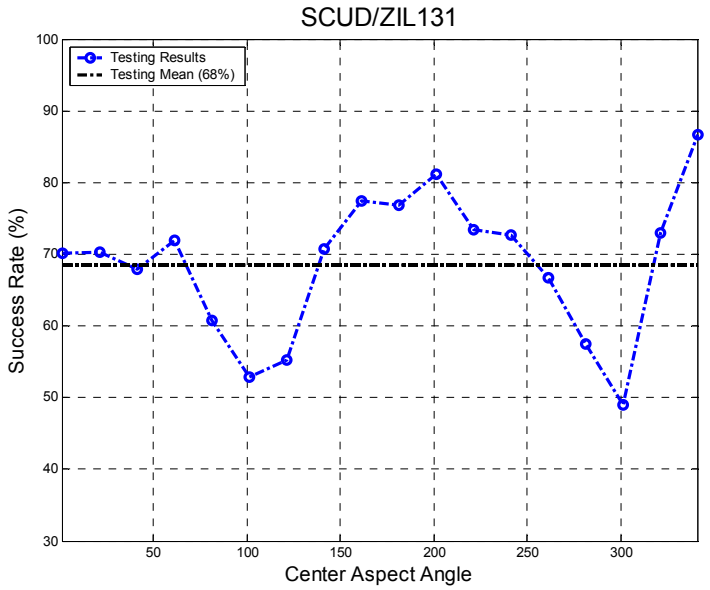


Figure 56. Aspect Angle View Classification Results for SCUD and ZIL-131

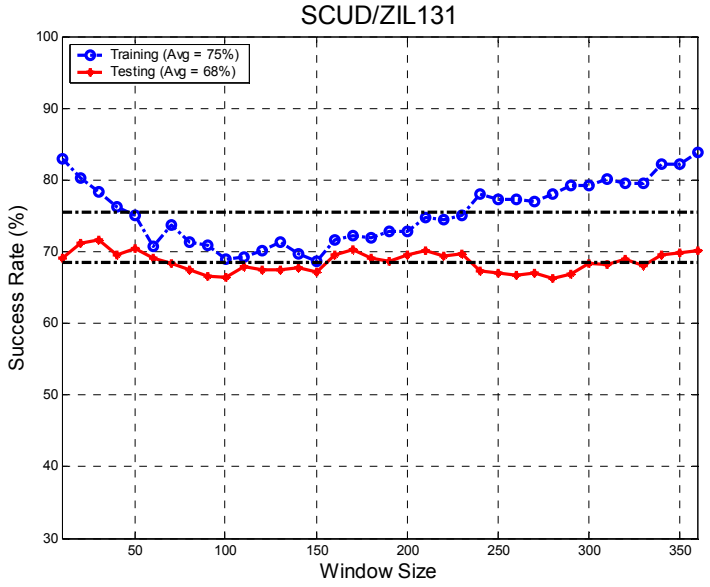


Figure 57. Sample Window Classification Results for SCUD and ZIL-131

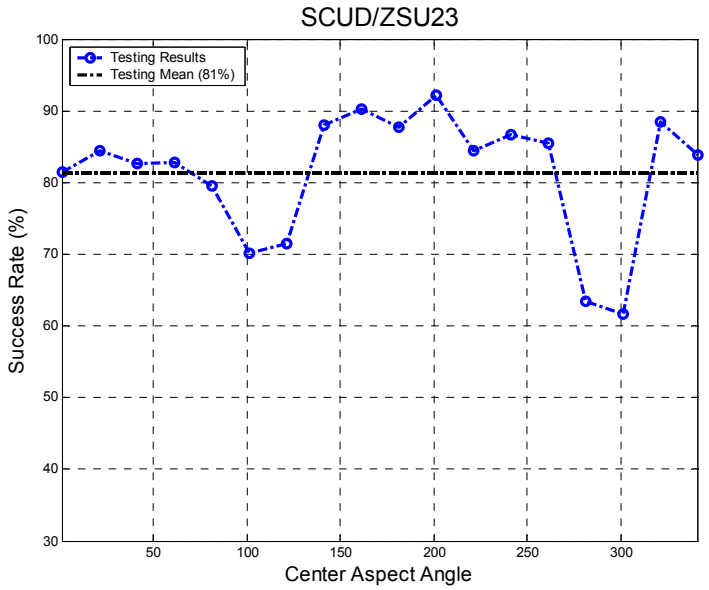


Figure 58. Aspect Angle View Classification Results for SCUD and ZSU-23

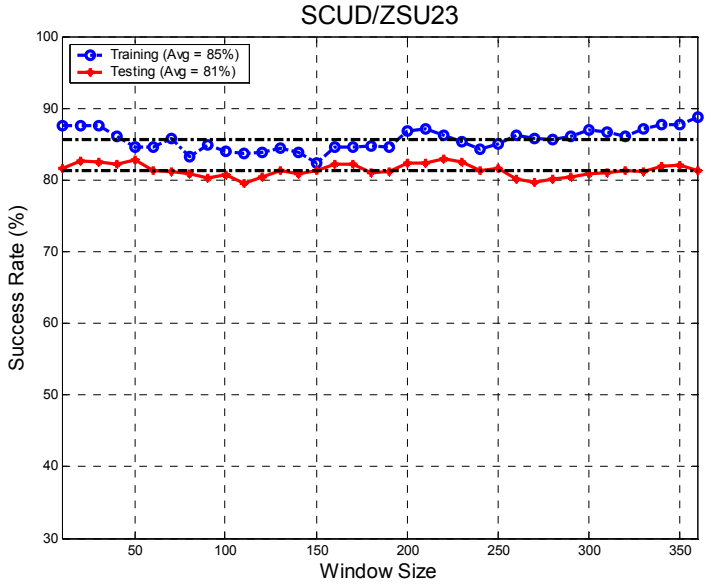


Figure 59. Sample Window Classification Results for SCUD and ZSU-23

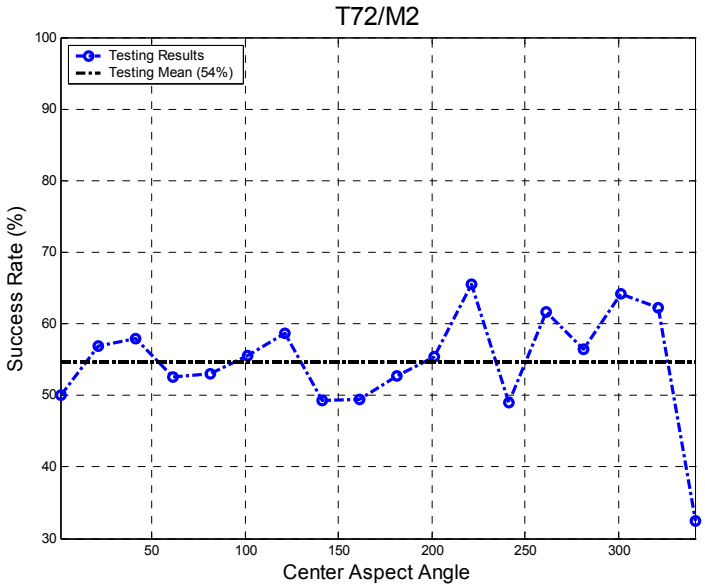


Figure 60. Aspect Angle View Classification Results for T-72 and M-2

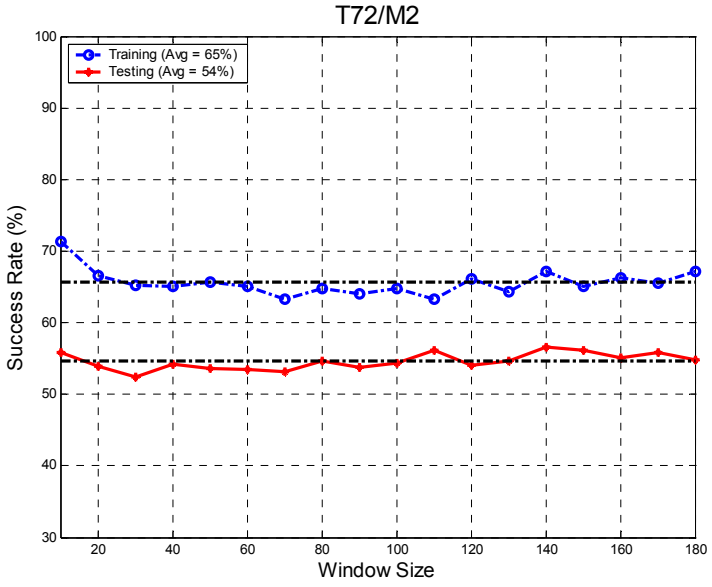


Figure 61. Sample Window Classification Results for T-72 and M-2

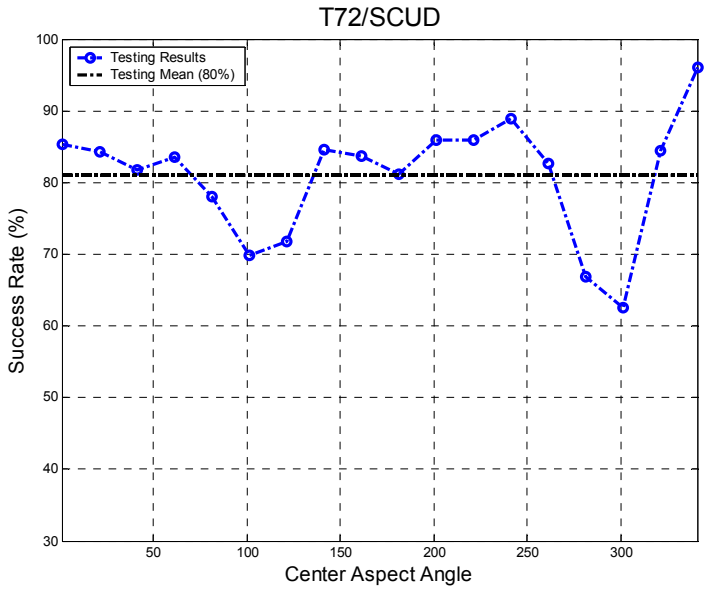


Figure 62. Aspect Angle View Classification Results for T-72 and SCUD

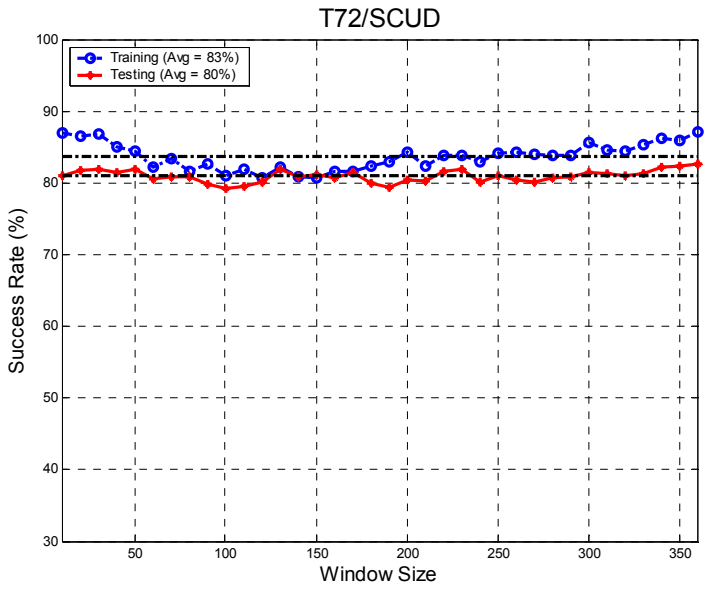


Figure 63. Sample Window Classification Results for T-72 and SCUD

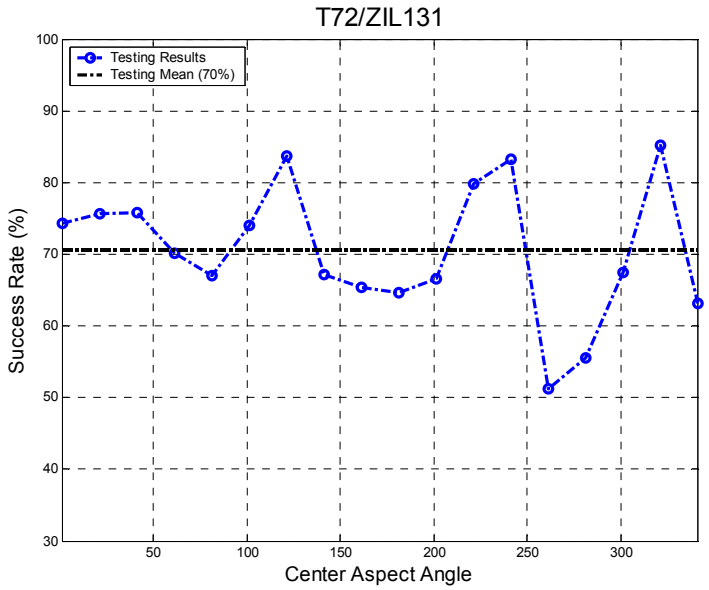


Figure 64. Aspect Angle View Classification Results for T-72 and ZIL-131

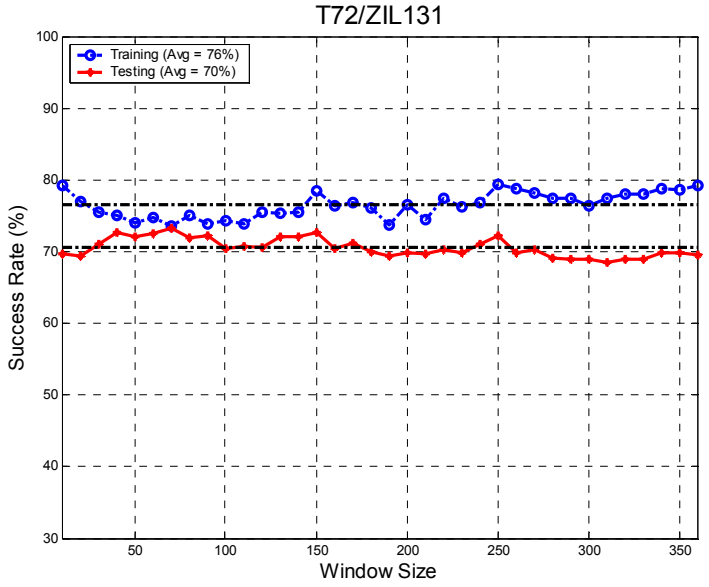


Figure 65. Sample Window Classification Results for T-72 and ZIL-131

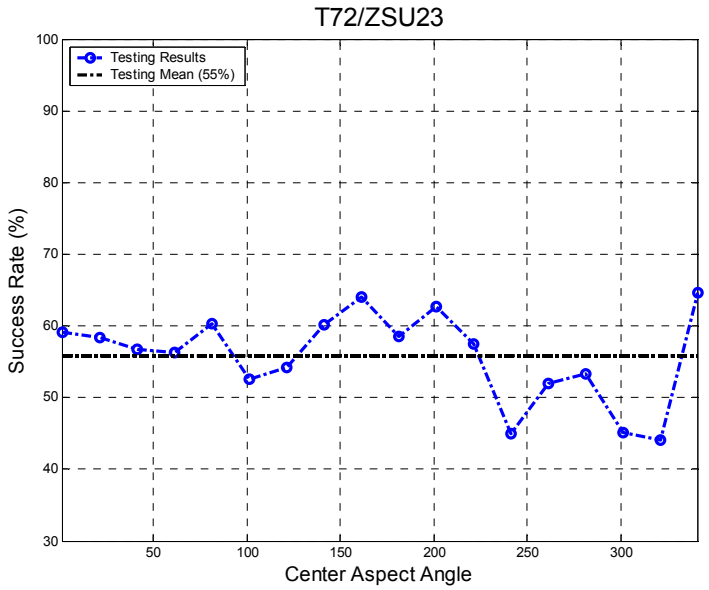


Figure 66. Aspect Angle View Classification Results for T-72 and ZSU-23

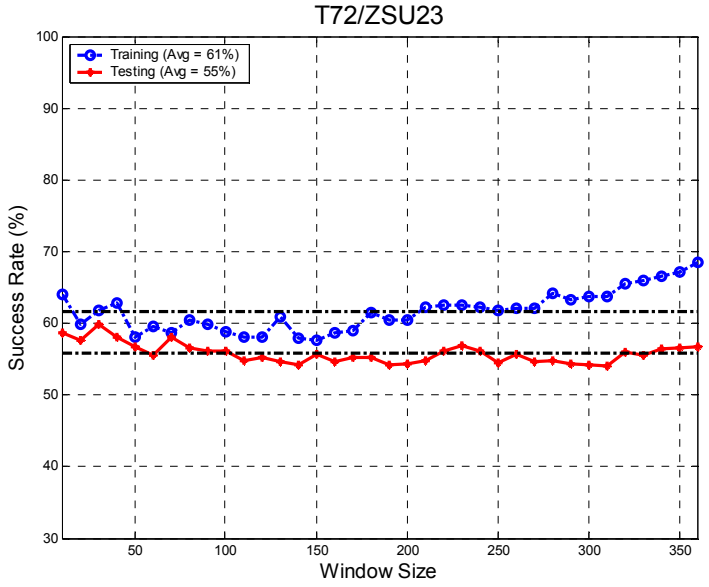


Figure 67. Sample Window Classification Results for T-72 and ZSU-23

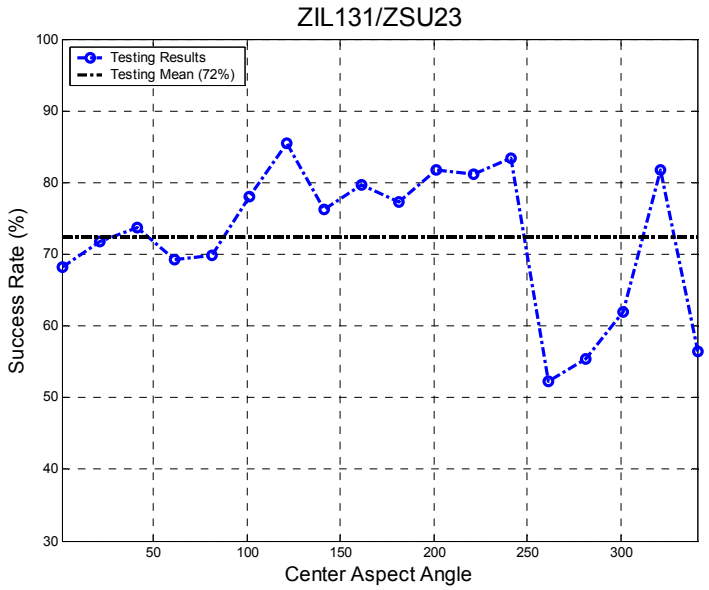


Figure 68. Aspect Angle View Classification Results for ZIL-131 and ZSU-23

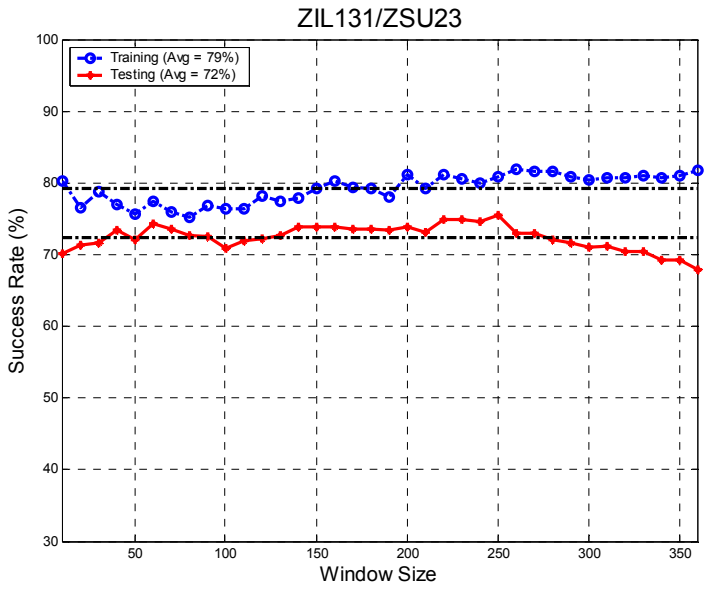


Figure 69. Sample Window Classification Results for ZIL-131 and ZSU-23

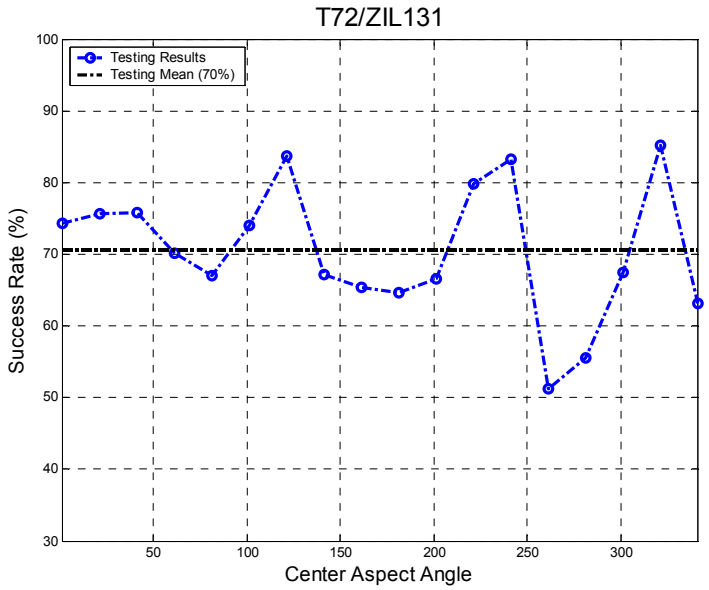


Figure 70. Aspect Angle View Classification Results for T-72 and ZIL-131

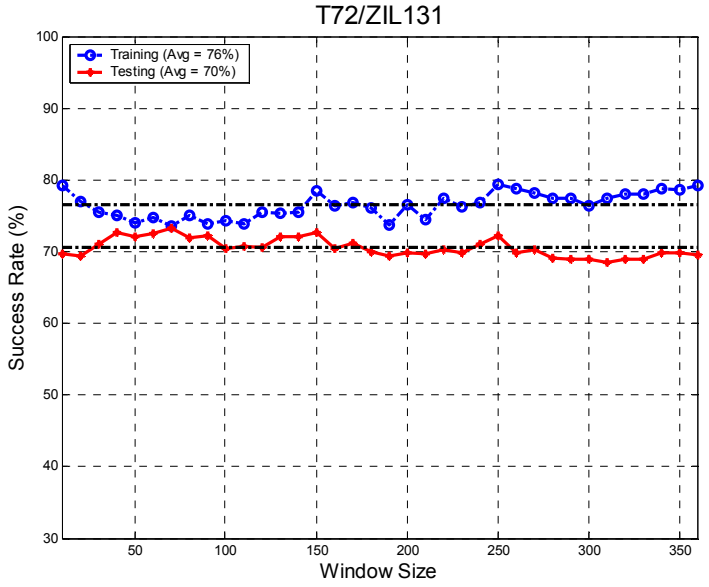


Figure 71. Sample Window Classification Results for T-72 and ZIL-131

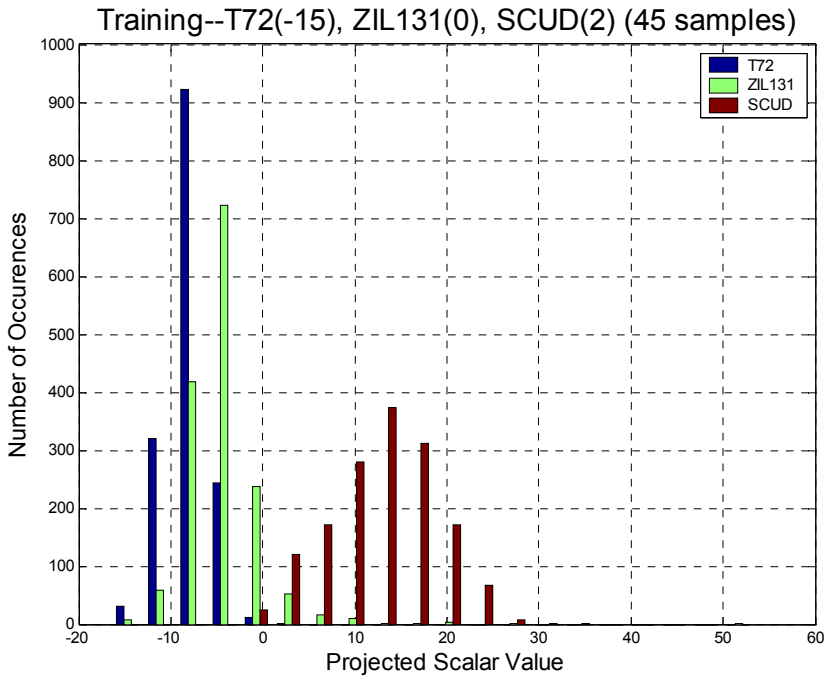


Figure 72. Three-Target Classification Training Histogram (T72, ZIL, SCUD)

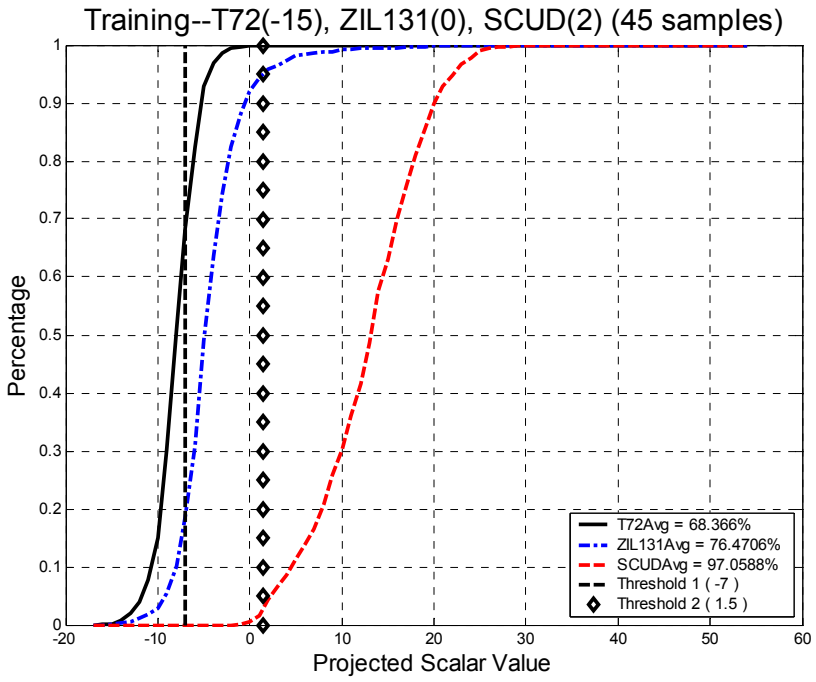


Figure 73. Three-Target Classification Training Plot (T72, ZIL, SCUD)

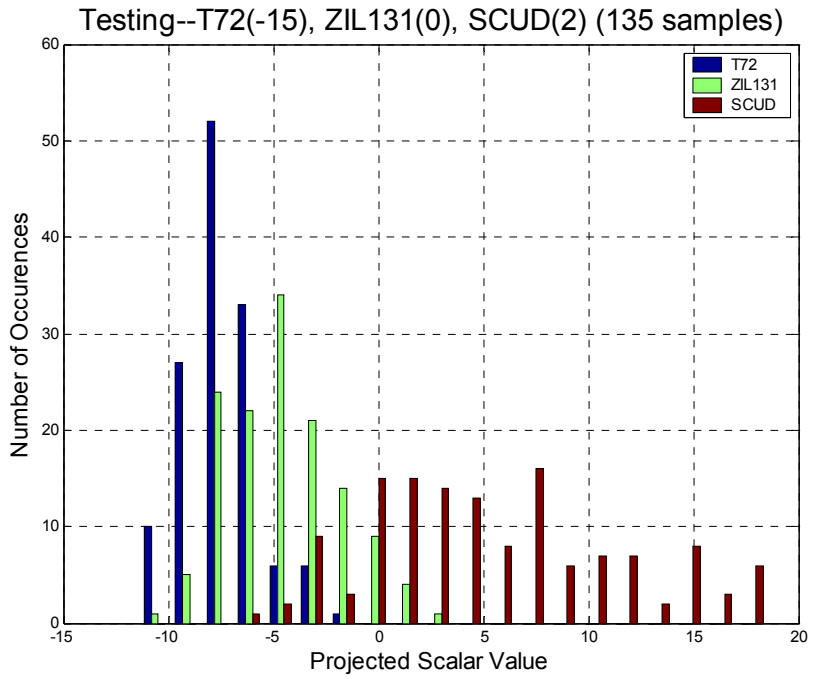


Figure 74. Three-Target Classification Testing Histogram (T72, ZIL, SCUD)

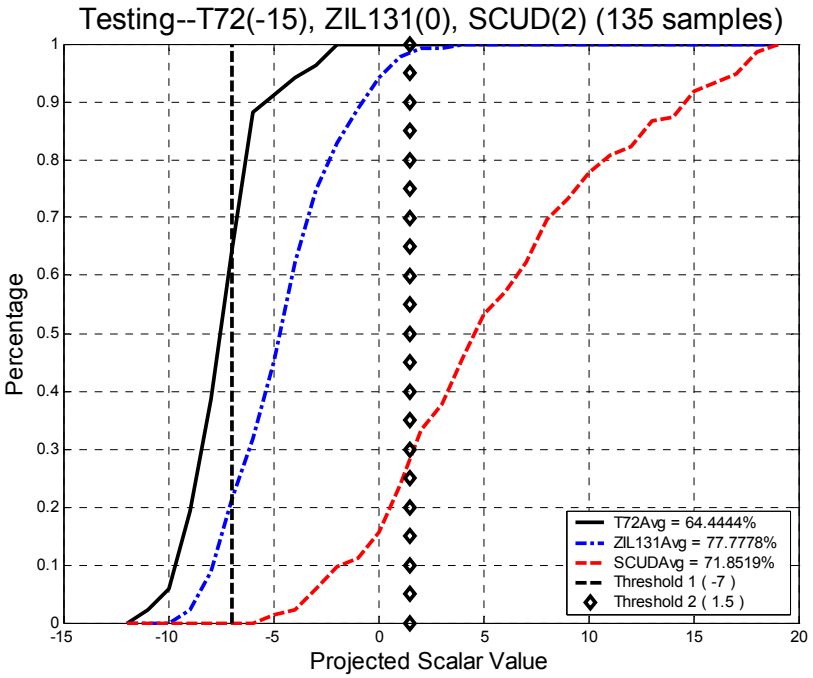


Figure 75. Three-Target Classification Testing Plot (T72, ZIL, SCUD)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
23-03-2004	**Master's Thesis**	Jun 2003 – Mar 2004

4. TITLE AND SUBTITLE

TARGET RECOGNITION USING LINEAR CLASSIFICATION OF HIGH RANGE RESOLUTION RADAR PROFILES

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Diaz, Ricardo. A., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GE/ENG/04-06

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFRL/SNAT
Attn: Dr. Devert Wicker
AFRL/SNAT (AFMC)
Bldg 620, 2241 Avionics Circle
Wright Patterson AFB, OH 45433-7321
DSN: 785-1115 ext 4250 e-mail: Devert.Wicker@wpafb.af.mil

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

High Range Resolution (HRR) radar profiles map three-dimensional target characteristics onto one-dimensional signals that represent reflected radar intensity along target extent. In this thesis, second through fourth statistical moments are extracted from HRR profiles and input to Fisher Linear Discriminant (FLD) classifiers. An iterative classification process is applied that gradually minimizes required a priori knowledge about the target data.

It is found that the second through fourth statistical moments of HRR profiles are useful features in the FLD classification of dissimilar targets and they provide reasonable discrimination of similar targets. Greater than 69% correct classification for two-target scenarios and greater than 60% correct classification for three-target scenarios is obtained using a single HRR profile extracted from a full 360-degree aspect angle window. A key contribution of this thesis is the demonstration that simple statistical moment features and simple linear classifiers can be used to effectively classify HRR profiles.

15. SUBJECT TERMS
Signal Classification, Target Classification, Target Discrimination, Target Recognition, Pattern Recognition, Radar, Linear Classification, High Resolution, Radar Signatures

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Steven C. Gustafson
U	U	U	UU	110	19b. TELEPHONE NUMBER *(Include area code)* (937) 785-3636 x4598; e-mail: steven.gustafson@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18