Piecewise Linear Approach:
a New Approach in Automatic Target Recognition

Yujing, Zeng , Janusz Starzyk
Department of Electrical and Computer Engineering, Ohio University

ABSTRACT

Automatic Target Recognition (ATR) of moving targets has recently received increased interest . High Range Resolution (HRR) radar mode provides a promising approach which relies on processing high-resolution 'range profiles' over multiple look angles. To achieve a robust, reliable [1]and cost effective approach for HRR-ATR, a model-based approach is investigated in this paper. A subset of the Moving and Stationary Target Acquisition and Recognition (MSTAR) data set was used to study robustness and sensitivity issues related to 1D model-based ATR development and performance. The model is built based on the statistic analysis of the training data and the dependence of the HRR signature on the azimuth is considered. The dependence is approximated by a linear regression algorithm to construct the templates of targets, which gives this approach the name of piecewise linear approach (PWL). Compared with the 1D model-based ATR approach developed by the Wright Laboratory, results are presented demonstrating an increase of about 10% in the correct identification probability of known targets when declaration probability P_{dec} is above 85% while maintaining a low time-cost.
Keyword: ATR, HRR, PWL, piecewise linear

1. BACKGROUD

The SHARP objective is to design, evaluate, test, and demonstrate air-to-ground ATR on the Air Force reconnaissance and surveillance platforms. Target classification is one of the important tasks in target tracking, detection, and recognition within a battlefield scenario. SAR images of moving targets are based on the Doppler effect and quite often are not as clear as the images of stationary targets. Variable terrain and variances of the observed objects make this task more difficult. The problem becomes more challenging for the HRR signatures used in our program, based on following facts:

- a lack of training data
- an inherent variability of data signals due to the nature of radar return signals

SAR and HRR images are collected by reconnaissance and surveillance platforms (like Global Hawk). GMTI (ground moving target indicator) system scans a selected strip of the ground and collects SAR images of the moving targets. In HRR mode SAR waveforms are processed to obtain high range resolution target signatures. Target chips that result from processing Doppler effect are collected for various observation angles. Alternatively a synthetic, model-based images can be used to describe the targets.

A detailed overview of moving target data collection and target modeling issues are presented in [7]. In this research the Moving Target Acquisition and Recognition (MSTAR) program data collections 1 & 2, scene 1 were used. These data were produced using X-band 1x1 foot resolution SAR images recorded at 15^o and 17^o depression angles with 360^o coverage in aspect angle. Each 17^o data consists of approximately 250 SAR chips per target and 15^o data of approximately 195 SAR chips per target. Data collected at 15^o were used for testing and 17^o for training. 10 target classes were considered with two target classes (BMP2 and T72) having several variants.

SAR chips are organized, in the data collection, by target depression and azimuth angles. The HRR profiles are obtained by using image segmentation from clutter to fit a target area. The mask applied to the original image is rotated to the target orientation and zero padded to the chip dimension. The cross-range inverse FFT is applied to obtain range versus angle information. The resulting data are dewhitened in cross-range using an inverse Taylor window over the valid data. After taking the magnitude of the result data each range bin is normalized by mean power in the signature to remove the gain control and range effects. A power transform is used to each range bin to transfer data distribution close to Gaussian. Finally the pixels are averaged in azimuth to form HRR profile.

[1] This work was supported by the grant from the Air Force Office of Scientific Research, grant # GAIF012.

The resulting range profiles are collected in chips that span over 3.5° in azimuth angle. FIGURE 1 shows an example of one-chip profiles.

Subsequently, center 8 profiles are extracted from each chip and power transform is used. Templates are then formed from the average of these 8 profiles. FIGURE 2 shows the averaged profile of the same target at the same azimuth angle.

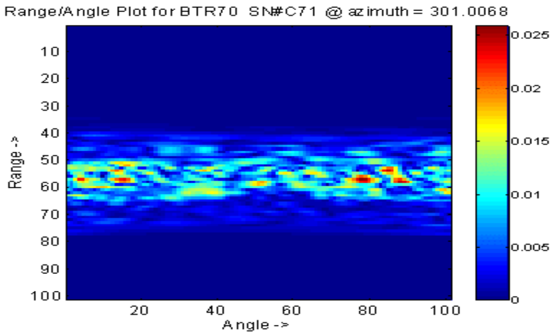


FIGURE 1 CHIP DATA FOR BTR60 AT AZIMUTH 301.0068 DEGREES

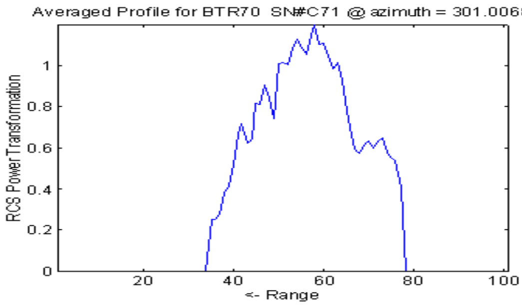


FIGURE 2 AVERAGE PROFILE FOR BTR60 AT AZIMUTH 301.0068 DEGREES.

2. SUMMARY OF THE PRIOR RESEARCH

ATR systems for target recognition based on SAR images were in development for several years. Irving demonstrated that multiresolution processing could be used to discriminate the radar signatures from clutter [2]. An approach to produce the moving target images from X-band 1x1 foot data was described in [1]. These images were used to produce data for the MSTAR program.

Obtained HRR signatures were used in the baseline algorithm to formulate templates and to perform signal classification. The baseline approach is a robust model-based automatic target recognition approach. It uses profiles of training data to construct one-dimension template for each degree per target. After using fist aligning and normalizing the templates and observations, this approach uses a mean-square statistic to find out the preliminary target classification. Upper bounds on the preliminary target classification mean-square error are used to reject unknown targets, which are targets not in the training set.

Since our aim was to develop methods that improve the technology, which is currently applied to target recognition, we referred in our work to the baseline approach which proved to be a very robust and successful method for moving target ATR as reported in [5]. Details of the baseline approach were presented in [8] and in internal document *'A Framework for 1D HRR ATR Evaluation'*. This document specified basic definitions for MSTAR program, metrics of performance, decision thresholds, leave-one out method used for the unknown target problem, and basic receiver operating conditions (ROC) curves. Although another technique (StaF classifier) presented in [5] reported a significant improvement over the baseline, it used test data information to set up the classification thresholds and therefore it couldn't be considered as a reference for our work.

3. PWL APPROACH

3.1 The Basic Idea

The baseline classification makes use of the mean value of a chip signatures as the template for a specified azimuth and target class. Let us represent the training data in a chip under consideration by $X \in R^{m \bullet n}$, where m is the number of signatures in this chip and n is the dimensionality of each signature. In addition, let us assume that an example chip data for a selected range value is distributed as shown in FIGURE 3.

The template method uses a constant value based on the mean of the chip data at a specified azimuth window

$$f(azimuth) = mean(X) \quad . \tag{1}$$

As a result, for each range value, a horizontal line effectively estimates all signatures X in this chip. Even though, data distribution varies with respect to the azimuth, the estimate used in the template approach does not make an explicit use of this variation. If the relationship between the distribution of the training samples and the azimuth value can be estimated and used to construct the training model, the classifier performance would be improved. Suppose that a training data X varies as a function of azimuth as follows:

$$f'(azimuth) = g(X, azimuth) + c(X) \quad (2)$$

where *c(X)* is the constant related only to the training signature X, and *g(X,azimuth)* is a complicated function and it is expensive and difficult to get its exact estimation.

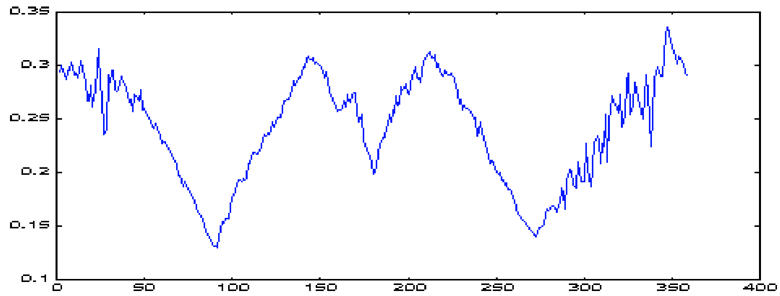


FIGURE 3   AVERAGE OF PROFILE FEATURES W.R.T. AZIMUTH FOR 2S1-B01

Within a specified range of the azimuth angles, this function can be approximated by a known functional. The simplest form of this functional is the linear function, that is

$$g'(X, azimuth) = \alpha(X) \bullet azimuth$$
$$where \quad \alpha(X) \text{ is a constant determined by the distribution of } X \quad (3)$$

Substituting (3) into (2), we get

$$f'(azimuth) = \alpha(X) \bullet azimuth + c(X) \quad (4)$$

So, instead of constants, we use linear functions to express the distribution of training signatures at specified azimuth angles, as shown by the dashed line in FIGURE 4. Furthermore, the new template model of one target is represented by a group of lines instead of a group of constant values used in the template approach. We call it a piecewise-linear (PWL) template representation.

The main principle behind the piecewise linear template approximation is to make an explicit use of the training data variation with the azimuth angle by using localized linear regression fits. The corresponding algorithm is described next as the PWL algorithm.

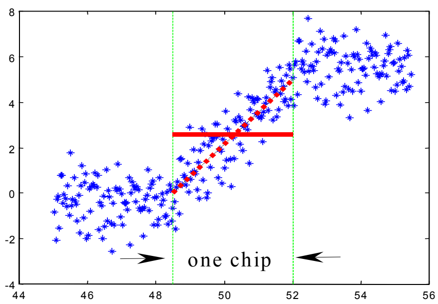


FIGURE 4   TWO TEMPLATE REPRESENTATIONS USING THE TRAINING DATA

3.2 The PWL algorithm

3.2.1 Linear regression fit to training data

Suppose that all the signatures are partitioned into several intervals according to their azimuth angle. In each range bin of each interval, we have a group of scalars related to the azimuth angle. Let $\vec{X}_i$ denotes the vector of i-th range bin values for various azimuth angles, and x_{ij} be j-th component of the vector $\vec{X}_i$. The corresponding azimuth vector is represented by $\vec{A}$.

$$\vec{X}_i = [x_{i1}, \cdots, x_{im}]'$$
$$\vec{A} = [a_1, \cdots, a_m]'$$

where m is the number of signatures in the group

The expected multidimensional line projected onto this range bin dimension minimizes the following equation:

$$\left\| \begin{bmatrix} \vec{A} & \vec{I} \end{bmatrix} * \begin{vmatrix} k_i \\ b_i \end{vmatrix} - \vec{X}_i \right\| \quad (5)$$

where $\vec{I} = [1, \cdots, 1] \quad R^m$

The solution of this minimization problem yields 2×2 linear equation that provides two linear regression fit coefficients k_i and b_i for each range bin value:

$$\begin{bmatrix} \| \vec{A} \| & \sum_{j=1}^{m} a_j \\ \sum_{j=1}^{m} a_j & m \end{bmatrix} \bullet \begin{bmatrix} k_i \\ b_i \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m} a_j \bullet x_{ij} \\ \sum_{j=1}^{m} x_{ij} \end{bmatrix} \quad (6)$$

FIGURE 5 gives an example of linear regression fit to the real HRR data of target 2s1b01.

Repeating this calculation for various range bins, we can get two vectors, $\vec{K}$ and $\vec{B}$, which describe the multidimensional linear regression fit to the training data. These vectors are respectively,

$$\vec{K} = [k_1, \cdots, k_n]$$
$$\vec{B} = [b_1, \cdots, b_n]$$

where n is the number of range bin

As a result, training data is described by two constant vectors $\vec{K}$ and $\vec{B}$ rather than a single one used in the template approach.

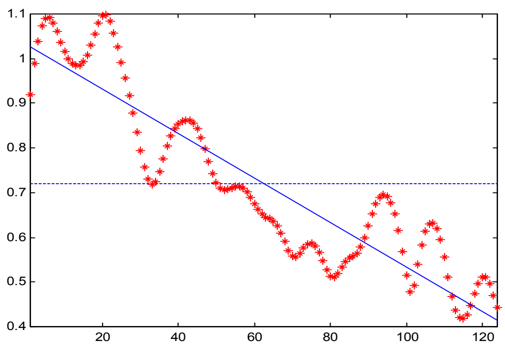
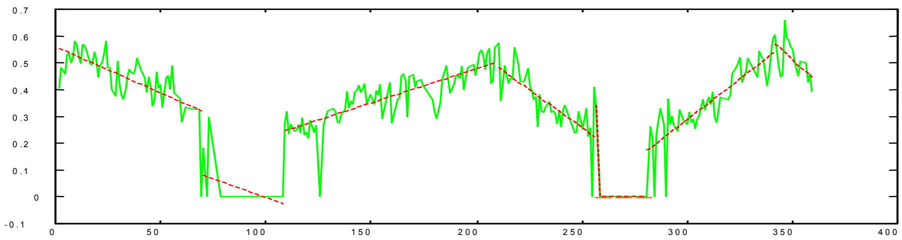


FIGURE 5 TWO TEMPLATE REPRESENTATIONS FOR AN EXAMPLE HRR DATA (TARGET 2S1_B01, FEATURE 58 WITH AZIMUTH =2.2248)

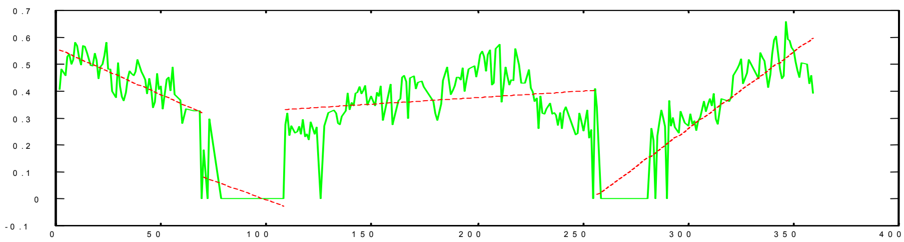
3.2.2 Interval selection for the linear regression fit

Based on the process described in the prior section, we can see that the selection of the azimuth interval is critical to the accuracy and efficiency of the estimate. If the selected interval is too large, it will result in a large estimation error that can ruin the classification performance. On the other hand, too small intervals result in a large number of PWL lines, which increases the computation cost of the algorithm. Moreover, the statistical estimate in a small interval is very sensitive to the local noise in the original signatures, so too small intervals may increase computational cost without increasing the classification performance.

The best choice is to use an adaptive linear regression approach. That is, the intervals are assigned based on the distribution of all the signatures. The area where the signatures have large variance is partitioned into smaller intervals to maintain high accuracy of the estimate; in the low variance area, a large interval is taken to reduce the computation cost of the classification algorithm which uses this piecewise linear template model. The adaptive partition of intervals is controlled by the mean squared error of the estimate. When the obtained estimate has a mean squared error smaller than the accepted error threshold, the partitioning process is stopped. A recursive algorithm that subdivides successive intervals hierarchically realizes this process. FIGURE 6 shows the result of adaptive linear regression algorithm on the HRR data with two different error thresholds.



a)



b)

FIGURE 6 ADAPTIVE LINEAR REGRESSION FIT RESULTS FOR HRR DATA (TARGET 2S1_B01, THE 35TH RANGE BIN)
A) ERROR THRESHOLD = 0.01; B) ERROR THRESHOLD = 0.005

3.2.3 Profile and template line alignment

In the baseline method, each template is just a single vector. The alignment algorithm employs L2 minimum-norm discriminant function to match the template and the observation vectors, which may be posed as the following problem:

$$
\begin{aligned}
& Error_B = \underset{a,b\in R}{minimize} \left\| \vec{P} - \left(\lambda_1 \bullet \vec{T} + \lambda_2 \bullet \vec{I}\right) \right\| \\
& where \quad \vec{T} = [t_1, \cdots, t_n] \in R^n \text{ the template vector}, \\
& \qquad \vec{P} = [p_1, \cdots, p_n] \in R^n \text{ the profile vector}, \\
& \qquad \lambda_1 \text{ and } \lambda_2 \text{ are the alignment parameters}
\end{aligned} \tag{7}
$$

The PWL method uses a line equation as a template, which can be expressed by the vectors $\vec{K}, \vec{B}$ on both sides of this line segment as

$$
\vec{T}(\alpha) = \alpha \bullet \vec{K} + \vec{B} \qquad \text{for} \quad \alpha_1 \le \alpha \le \alpha_m \tag{8}
$$

Therefore, the alignment problem is to find the point from this line and the corresponding alignment parameters, λ_1 and λ_2, such that the discriminant function (an error function) of the aligned template and the profile has a minimum norm. For the convenience of calculation, the three alignment parameters are combined and the norm of the error function can be calculated as:

$$
\begin{aligned}
& Err\_P_1 = \underset{\beta_1,\beta_2,\beta_3\in R}{minimize} \left\| \vec{P} - \left(\beta_1 \bullet \vec{K} + \beta_2 \bullet \vec{B} + \beta_3 \bullet \vec{I}\right) \right\| \\
& \qquad where \quad \beta_1 = \alpha \bullet \lambda_1; \quad \beta_2 = \lambda_1 \quad \beta_3 = \lambda_2
\end{aligned} \tag{9}
$$

To satisfy these equations we must solve the following:

$$\begin{bmatrix} \|\vec{T}_1\| & \sum_{i=1}^{n} t_{1i} \bullet t_{2i} & \sum_{i=1}^{n} t_{1i} \\ \sum_{i=1}^{n} t_{1i} \bullet t_{2i} & \|\vec{T}_2\| & \sum_{i=1}^{n} t_{2i} \\ \sum_{i=1}^{n} t_{1i} & \sum_{i=1}^{n} t_{2i} & n \end{bmatrix} \bullet \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} t_{1i} \bullet p_i \\ \sum_{i=1}^{n} t_{2i} \bullet p_i \\ \sum_{i=1}^{n} p_i \end{bmatrix} \tag{10}$$

Equation (10) is used to determine β_1, β_2 and β_3, and subsequently λ_1, λ_2 and α values. This calculation is repeated for each shift of the profile vector, and the minimum distance determined for each one, using distance from a profile vector to the multidimensional line segment.

3.2.4 Observation group alignment

In the previous section, we discussed the alignment between the observation profile and the template line. In that approach, information about the signature's azimuth is used in the estimate of template construction, but it is still ignored in matching the observation to the template. To improve our approach, a group of observation signatures is used to find the best matching segment of the PWL template line. So the alignment will be performed between two groups of signatures.

As we know, the estimate of signal distribution by a known function, such as the linear function, introduces the estimation noise. To improve the recognition accuracy, the individual observation signatures vs. azimuth are not estimated. Instead, we group all the observation signatures in a large vector, by attaching them one by one:

$$\begin{aligned} &\overline{\mathrm{P}} = [\, \vec{u}_1, \cdots, \vec{u}_l \,] \in R^{l \bullet n} \\ &\textit{where} \quad \vec{u}_i = [u_{i1}, \cdots, u_{in}] \textit{ is a single observation signature with azimuth } \omega_i \\ &\qquad\quad \vec{\omega} = [\omega_1, \cdots, \omega_l] \textit{ is the azimuth of each signature in the group} \\ &\qquad\quad l \textit{ is the number of the signatures in the observation group} \end{aligned}$$

For convenience, we set $\omega_1 = 0$. Correspondingly, the match point in the template line is represented as:

$$\begin{aligned} &\overline{T}(\alpha) = \alpha \bullet \overline{K} + \overline{B} \qquad\qquad a_0 \le \alpha \le a_m - (\omega_l - \omega_1) \\ &\textit{where} \quad \overline{K} = [\vec{K} \bullet \omega_1, \cdots, \vec{K} \bullet \omega_l] = [\bar{k}_1, \cdots, \bar{k}_{l \bullet n}] \in R^{l \bullet n} \\ &\qquad\quad \overline{B} = [\vec{B}, \cdots, \vec{B}] = [\bar{b}_1, \cdots, \bar{b}_{l \bullet n}] \in R^{l \bullet n} \end{aligned} \tag{11}$$

The solution of this optimization problem is obtained by solving the following equation with respect to γ_1, γ_2 and γ_3

$$\begin{bmatrix} \|\overline{K}\| & \sum_{i=1}^{l \times n} \bar{k}_i \bullet \bar{b}_i & \sum_{i=1}^{l \times n} \bar{k}_i \\ \sum_{i=1}^{l \times n} \bar{k}_i \bullet \bar{b}_i & \|\overline{B}\| & \sum_{i=1}^{l \times n} \bar{b}_i \\ \sum_{i=1}^{l \times n} \bar{k}_i & \sum_{i=1}^{l \times n} \bar{b}_i & n \times l \end{bmatrix} \bullet \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{l \times n} \bar{k}_i \bullet \bar{p}_i \\ \sum_{i=1}^{l \times n} \bar{p}_i \bullet \bar{b}_i \\ \sum_{i=1}^{l \times n} \bar{p}_i \end{bmatrix} \tag{12}$$

Once the optimized parameters γ_1, γ_2 and γ_3 are found the minimum alignment error is determined by (12) and classification decision is based on the class of PWL template with the smallest minimum alignment error.

3.3 Application Results

In the application of PWL approach, there are several parameters whose values are crucial to the classification performance. One is the interval size of linear regression fit, as mentioned in 3.2.2. All the parameters are dependent on the data source to classify. Therefore, we determine that empirically [6].

The performance results of the PWL classifier is described by the confusion matrix and receiver operating conditions (ROC) curves. The ROC curves show the tradeoff between the probability of false alarm P_{fa}, probability of declaration P_{dec}, and the probability of correct identification P_{cc} comprehensively, while the confusion matrix summarizes the performance estimates at a single operating point. We break the three parameters in ROC curves into three sets, represented by three two-dimensional ROC curves. The first ROC curve relates conditional P_{cc} given P_{dec} as a function of P_{dec} and will be refereed to as ROC1. The second ROC curve, referred to as ROC2, relates P_{fa} as a function of P_{dec}. The last ROC curve, referred to as ROC3, shows P_{dec} vs. P_{fa}.

FIGURE7 -9 summarize the performance of PWL classifier for the 10 target classes, compared to Baseline approach. In both methods a single-look was used for MSTAR public targets from collections 1 & 2, scene 1. The overall average results obtained for the Baseline approach can be summarized by P_{id}=78.7 %, P_{dec}=97.2 %, P_{fa}=66 %, while for the PWL method the corresponding results were P_{id}=85.7% (increase by 7%), P_{dec}=96.8% (decrease by 0.4%), P_{fa}=58% (decrease by 8%), as shown in FIGURE 9. PWL method used 10 degrees PWL templates with observations extracted from 5-point averaged chip data. And the ROC curves in FIGURE7 and FIGURE8 show that the PWL classifier outperforms the Baseline in most operating points.

The PWL classifier with this configuration has a similar computation cost with the Baseline approach but a better performance. The process for 5 targets takes both classifiers about 100 minutes in PC (233MHz). Moreover, we can get a faster algorithm with losing a little performance by adjusting the parameters inside the program, such as the number of observation data. This scalability of PWL is helpful for the future research and improvement of the developed approach.

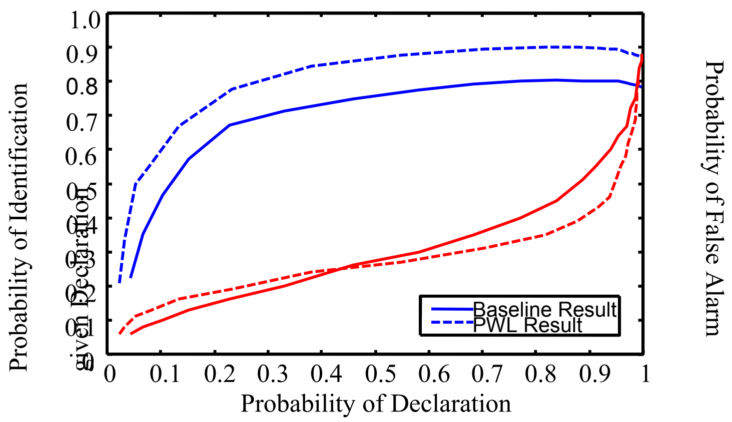


FIGURE 7 ROC CURVES 1 AND 2 OF PWL APPROACH FOR 10 TARGETS

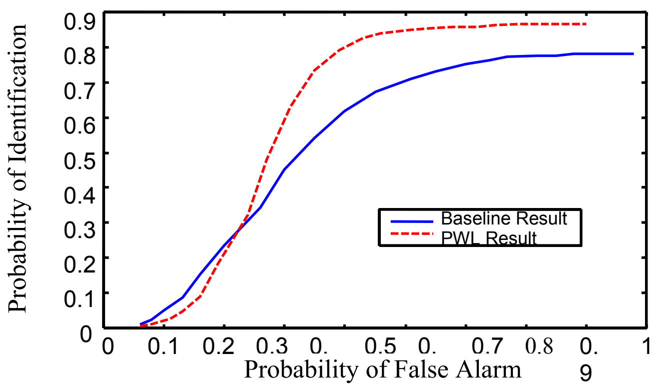


FIGURE 8 ROC 3 OF PWL APPROACH FOR 10 TARGETS

4. CONCLUSION

This paper describes a study of practical methods developed to improve classification performance of HRR based target recognition for air-to-ground images of moving targets. The piecewise linear method uses a localized linear-regression fit to the training data. It uses signal alignment and normalization which minimize least square error between the observation signals and PWL templates. PWL intervals can be selected automatically to minimize the linear regression fit error to the training data. The size of PWL intervals was optimized for classification performance. The method shows a noticeable improvement over the Baseline algorithm on the test data.

REFERENCES

[Car95] W. G. Carrara, R. S. Goodman, and R. M. Malewski, *Spotlight Synthetic Aperture Radar: Signal Processing Algorithms*, Artech House, Boston, 1995.

[Irv95] W.W. Irving, L.M. Novak, and A. S. Willsky, *A multiresolution approach to discriminating targets from clutter in SAR imagery*, Proc. SPIE, 1995.

[Fuk90] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Inc., 1990.

[Mit94] R.A. Mitchell and R. DeWall, *Overview of high range resolution target identification*, Proceedings of the ATR Working Group Systems and Technology Symposium, Monterey, CA, 1994.

[Mit98] R.A. Mitchell and J. Westercamp, *A Statistical Feature Based Classifier for Robust Hight Range Resolution Radar Target Identification*, Submitted to IEEE Trans. on Aerospace and Electronic Systems.

[Sta99] J.A. Starzyk, *Hybrid Distance and Picewise Linear Approaches*, Final Report to Air Force Research Laboratory, Sensors Directorate, AFRL/SN, Feb. 1999.

[Wes98]J. Westercamp, *Automatic Target Recognition Using High Range Resolution Radar Signature Features*, Final Report to Air Force Research, Lab Sensors Directorate, AFRL/SN, Sep. 1998.

[Wil97] R. Williams, J. Westercamp, D. Wardell, S. Worrel, and M. Ressler, *Robustness issues for 1D ATR*, Proceedings of the Sixth ATR Working Group Systems and Technology Symposium, Redstone , Alabama, October 1997.

Target	Serial Number	BMP2 c21	BTR60 k10+	BTR70 c71	T72 132	2S1 b01	BRDM2 e-71	D7 92v+	T62 a51	ZIL131 e12	ZSU23/4 d08	Threshold Reject
BMP2	c21	**80**	1	4	0	7	1	0	1	2	2	3
BTR60	k10+	0	**77**	9	0	7	5	0	0	0	0	2
BTR70	c71	0	9	**74**	0	10	4	0	0	0	0	2
T72	132	1	0	0	**86**	2	0	0	4	0	1	5
2S1	b01	2	3	6	0	**85**	2	0	0	0	0	3
BRDM2	e-71	0	3	2	0	2	**91**	0	0	0	0	2
D7	92v+	0	0	0	0	0	0	**97**	0	1	0	2
T62	a51	1	0	1	1	0	0	0	**89**	2	1	4
ZIL131	e12	1	0	0	0	2	0	0	0	**87**	2	8
ZSU23/4	d08	1	1	0	0	1	0	0	3	2	**91**	1
Unknown	NA	3	10	13	1	13	3	0	7	5	3	42

a)

Target	Serial Number	BMP2 c21	BTR60 k10+	BTR70 c71	T72 132	2S1 b01	BRDM2 e-71	D7 92v+	T62 a51	ZIL131 e12	ZSU23/4 d08	Threshold Reject
BMP2	c21	**75**	2	2	0	10	1	0	0	3	2	4
BTR60	k10+	4	**64**	6	0	8	1	0	2	5	7	3
BTR70	c71	3	8	**63**	0	11	0	0	1	5	5	4
T72	132	1	1	0	**88**	0	0	0	5	0	1	3
2S1	b01	7	4	6	0	**70**	1	0	1	6	4	2
BRDM2	e-71	0	0	1	0	0	**87**	4	0	6	0	1
D7	92v+	0	0	0	0	0	1	**96**	0	0	0	2
T62	a51	1	1	3	3	2	0	0	**80**	1	3	5
ZIL131	e12	2	2	2	0	3	4	1	0	**81**	3	2
ZSU23/4	d08	1	2	2	0	6	1	0	1	2	**83**	2
Unknown	NA	7	7	9	2	11	3	2	7	12	7	34

b)

FIGURE 7 Confusion Matrices of PWL And Baseline for 10 Targets ($P_{dec.}$ =97%)
a) Confusion Matrix of PWL; b) Confusion Matrix of Baseline