Shape-based Recognition of Targets in Synthetic Aperture Radar Images using Elliptical Fourier Descriptors

Louis P. Nicoli, Georgios C. Anagnostopoulos*
Department of Electrical & Computer Engineering, Florida Institute of Technology, 150 West University Boulevard, Melbourne, Florida 32901-6975, USA.

ABSTRACT

This paper primarily investigates the use of shape-based features by an Automatic Target Recognition (ATR) system to classify various types of targets in Synthetic Aperture Radar (SAR) images. In specific, shapes of target outlines are represented via Elliptical Fourier Descriptors (EFDs), which, in turn, are utilized as recognition features. According to the proposed ATR approach, a segmentation stage first isolates the target region from shadow and ground clutter via a sequence of fast thresholding and morphological operations. Next, a number of EFDs are computed that can sufficiently describe the salient characteristics of the target outline. Finally, a classification stage based on an ensemble of Support Vector Machines identifies the target with the appropriate class label. In order to experimentally illustrate the merit of the proposed approach, SAR intensity images from the well-known Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset were used as 10-class and 3-class recognition problems. Furthermore, comparisons were drawn in terms of classification performance and computational complexity to other successful methods discussed in the literature, such as template matching methods. The obtained results portray that only a very limited amount of EFDs are required to achieve recognition rates that are competitive to well-established approaches.

Keywords: Automatic Target Recognition, Synthetic Aperture Radar, MSTAR, Elliptical Fourier Descriptors, Support Vector Machines

1. INTRODUCTION

The task of Automatic Target Recognition (ATR) based on Synthetic Aperture Radar (SAR) images is to first detect the presence of a target in a recorded image and then, subsequently, identify the specific type of the target. This particular problem domain has been studied by several researchers and many approaches have been suggested in the recent literature. Furthermore, the *Moving and Stationary Target Acquisition and Recognition* (MSTAR) data set, a collection of SAR images taken of soviet made military vehicles[1], has been used extensively as a benchmark to test performance of various recognition schemes. In particular, two noteworthy ATR approaches that have been successfully applied in the past are *template matching* and *high-level feature-based recognition*.

Templates are any reference images in either the spatial or frequency domain that are constructed from the training set. In the case of frequency domain templates, there are an equal number of harmonic coefficients, as there are pixels in the training images. Typically, several templates are identified or constructed for each target class to account for intra-class distortion due to variations in pose and illumination. A test image is classified depending on the template that most closely resembles it. Here, resemblance is usually quantified by an appropriately chosen distance metric and the classification model utilized is a simple 1-Nearest Neighbor Rule[2]. Overall, template-based approaches are straightforward to implement and have produced very high rates of classification accuracy[3,4,5]. Template matching has been accomplished using filter templates[5,6,7] and minimum squared error templates/maximum likelihood[3,4,8,9].

Filter templates types, that have been studied with the MSTAR data set, are *Maximum Average Correlation Height* (MACH) combined with a *Distance Correlation Classifier Filter* (DCCF)[6], *Extended Maximum Average Correlation Height*[7], *Polynomial Distance Correlation Classifier Filter* (PDCCF)[7], and *Minimum Noise and Correlation Energy* (MINACE)[5]. The MACH/DCCF combination was the first study to propose a filter template method. It investigated the use of the filters to classify three target classes from the MSTAR dataset, namely BTR70, T72, and BMP2. In the MACH/DCCF method, a class template is generated by averaging all training images that belong to that class. A DCCF

*georgio@fit.edu; phone +1 321 674-7125; fax +1 321 674-8192; my.fit.edu/~georgio

Automatic Target Recognition XVIII, edited by Firooz A. Sadjadi, Abhijit Mahalanobis
Proc. of SPIE Vol. 6967, 69670G, (2008) · 0277-786X/08/$18 · doi: 10.1117/12.777806

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

filter is then designed that maximizes the interclass distance and is then applied to all the templates. The percent of test images that were correctly classified using this approach was 98.2%. Thus, for the three class problem, the MACH/DCCF approach was very accurate. Extensions to the MACH/DCCF were also studied. The *Extended Maximum Average Correlation Height Filter* (EMACH) was an improved version of the MACH filter[7]. The EMACH approach was applied to a 10-class instead of a 3-class problem so the improvement due to the extension cannot be directly seen from the results. For the ten classes, the EMACH approached produce a classification accuracy of 95%. An extension to the DCCF called polynomial distance correlation classifiers (PDCCF) filters was developed as well[7]. The PDCCF extended the DCCF method by first raising the individual features by an integer power before applying the distance correlation operation. When applied to the 3-class target problem, the classification accuracy reported for PDCCF was 99.1%. MINACE filters were also investigated as a template-based method for the 3-class ATR problem[5]. The construction of the MINACE filters is very different from the MACH/DCCF approach. In particular, they are synthesized iteratively from the training set. The MINACE approach exhibited a classification accuracy of 86.4% on the 3-target test set that had been augmented with clutter-only and confuser images.

The second family of template-based methods consists of maximum likelihood classifiers. Three such models that have been investigated are the *Quarter Power*, *Log Magnitude* and *Conditionally Gaussian*[3]. Each of these three approaches models the intensity (magnitude) of every pixel in the image as an independent identically-distributed random variable. The Quarter Power model assumes a gamma distribution, the log magnitude model assumes a Log-Normal model, and the conditionally Gaussian model assumes a complex normal distribution. With these methods, each pixel is used as a feature in a class membership likelihood calculation. For the quarter power and log magnitude approaches, the likelihood equation is mathematically equivalent to finding the template with the minimum squared error to the test image. As discussed earlier, each of the approaches used a total number of 720 templates. For best results, the images were normalized before training and classification. The log magnitude method had the worst accuracy of the three methods with a classification accuracy of 95%. The conditionally Gaussian classifier had the best accuracy at about 97% using an image size of 96×96 pixels. The quarter power method also had a similar result of about 97% using an image size of 128×128 pixels. These results indicate fairly good classification accuracy is achievable using template matching methods.

Nevertheless, one disadvantage of these approaches is that they tend to feature spatial complexity of the order O(*CAP*), which is high in typical cases. This is because these models are required to store a large number of *P*-pixel reference images taken from the *C* target classes and *A* poses (azimuths). The side-effect of these storage demands result in recognition that is computationally intensive. In more detail, the SAR image of each target varies with azimuth (pose) making the images non-invariant to rotation. Therefore, each target class must be composed of an ensemble of sub-classes (components of a mixture distribution) that include models of an individual target at several orientations. In the past, for optimum classification utilizing template matching, up to 72 orientation sub-classes (one for every 5°) have been used[3]. With 72 orientation subclasses, a 10-target classifier requires a minimum of 720 templates. Since there are a large number of sub-classes that need to be accounted for, the quantity of samples available for training is rather not sufficient. When divided into 720 subclasses, the portion of samples that belong to a particular vehicle and pose subclass is small. Case in point, for the *Maximum Likelihood* model[3] there is only an average of 10.2 sample images per subclass used in training. Even worse, the dimensionality *P* of these patterns is also high, because almost all of the template matching methods use each image as a single feature vector, i.e. each pixel in the images is considered a feature. For example, a 128×128 image would contain 16384 features. Another disadvantage with respect to template matching is that the entire scene is considered by the model. This includes the clutter region of the MSTAR chips. Obviously, the clutter region contains no information about the target and should, theoretically, only add noise to the distribution of the data.

In order to address these high storage demands and the disproportion between the relatively small number of available training patterns and the high dimensionality of the resulting feature space, another family of approaches investigated in the relevant literature is the ones that are based on high-level features of the SAR images. Rather than considering the raw image intensity values as individual features, these methods preprocess the images and extract a relatively small amount of features that are capable of capturing important information about a target class. For instance, the pose of the target has already been shown to convey information that has a large impact on the classification accuracy of an ATR system[3,6,10]. Even though the estimates of the pose are imperfect, their use in the ATR systems reduces complexity and can improve performance[10]. Other heuristic features from the problem domain have been investigated for use in SAR ATR. Among the most obvious choices for high level features are some measures of length and width of the target.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

Other high level features specific to radar are the average radar cross section (RCS) and log standard deviation (LSD)[11]. The radar cross section is a measure of how well a target reflects radar waves. The average radar cross section is therefore a feature that is related to the overall reflectivity of the target. The LSD measures the intra-pixel variation in intensity in the target region. Using only the RCS, LSD, length and width, a classification accuracy of 92% was achieved[3]. Nevertheless, the pose was assumed known beforehand, which inherently simplifies the ATR problem. Still, the use of high level features was shown to be a useful tool in ATR, considering that the smallest number of features used in the template approaches was 2304 (48×48 pixels)[3] as opposed to the four in the high level feature approach[11]. Another common feature used specific to radar is the location and magnitude of peaks[11,12,13,14,15]. In addition to the aforementioned heuristic features, additional ones, common to general image processing, are also used in ATR systems. Simple optical features such as edges and corners have been investigated[11,16]. Moreover, one of the studies focused on deriving features from the shape of the target region. Examples of such features are the *Hu Moments*[17], which are rotation- and scale-invariant. However, since the target images are not rotationally invariant, this advantage is nullified. Using a 3-nearest neighbor classifier and Hu Moments, a classification accuracy of 76.85% on a set of 7 vehicle classes was achieved. Also in the same paper, *Principle Component Analysis* (PCA) and *Independent Component Analysis* (ICA) were considered[17]. Using a 3-nearest neighbor classification rule, a performance of 96.47% was achieved using PCA on a test set of 679 images. While PCA's results are very good, it requires that the eigen-decomposition of the training set's covariance matrix be calculated. The 7-target class set yields a 4225×4225 covariance matrix, whose estimation and decomposition is quite computationally demanding. Nevertheless, it is only performed once during training and, subsequently, classifying a test pattern is relatively quick and inexpensive. ICA on the same images yielded a classification rate of 88.3%, which is much worse than the simpler PCA.

The purpose of our study is to investigate the explicit use of target outline descriptions as high-level features to address the ATR problem at hand. To the best of our knowledge, target outline descriptors have never been considered before in this context. In our work a parametric representation of the target outline is used as a feature set for classifying the targets. Additionally, the particular parameters used in this representation are *elliptical Fourier descriptors* (EFDs)[18]. An important characteristic of EFDs is that they can compactly describe any smooth, closed contour. By using them, the computational complexity of the proposed ATR system is far lower than template-based methods as we will show. Before venturing into the specifics of the proposed approach, some background information is due. The rest of the manuscript is organized as follows: Section 2 provides the necessary background with respect to EFDs and explains our approach to extracting an outline description of the targets. Section 3 contains all our experimental findings; additionally, it also shows a series of comparative results featuring our method versus some well established, template-based alternatives. Finally, Section 4 summarizes these findings and states a few important conclusions.

2. PROPOSED APPROACH

Overall, the classification of a SAR target image is accomplished in a three step approach. A segmentation stage first determines the target outline via a sequence of thresholding and morphological operations. Next, the EFD coefficients are computed. Finally, a classification stage based on an ensemble of Support Vector Machines (SVMs) identifies the target with the appropriate class label. Each of these stages is discussed in one of the following subsections.

2.1 Segmentation

The first step in the process, segmentation, isolates the target region from the clutter and shadow regions within the intensity images. The MSTAR images can be separated into three distinct regions: *target region*, *shadow region*, and *clutter region*. Fig. 1 shows each of the three regions in a sample image. Typically, the target region features brighter pixels called *peaks*; they represent locations on the target that produce strong radar returns. The shadow region features low intensity pixels, which correspond to the void behind the target, where the radar pulse does not reach, because it is blocked by the target's volume. The clutter region consists of pixels that correspond to noisy radar reflections from the ground surrounding the target.

Segmentation of the target region was accomplished through two basic steps of thresholding and morphological operations in order to keep the implementation simple and computationally efficient. First, a threshold of 10% of the magnitude of the brightest pixel was used to produce a grossly approximated target region.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

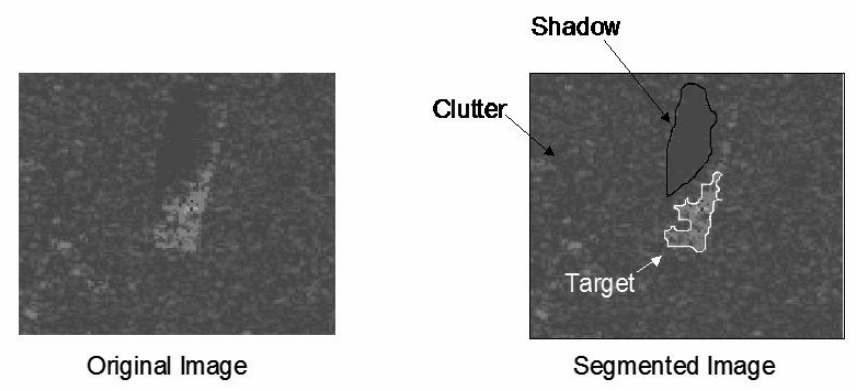

Fig. 1: The SAR images can be segmented into three regions: Target, Clutter, and Shadow. The shape and size of the shadow and target regions depend heavily on the azimuth of the depicted target in the SAR image. A sequence of computationally inexpensive thresholding and morphological close operations are used to identify the potential target region.

Then, a morphological close operation was applied to the image using a square structuring element. If the segmentation produced multiple disconnected regions, further processing was required. Segmented regions that did not contain a pixel that was brighter than 40% of the magnitude of the brightest pixel were discarded. Then, the close operation was applied again with an increasingly large structuring element until a single target region was produced, so that a single target region outline is obtained.

Table 1: 48×48 Quarter Power Template with Target Region Segmentation. The table depicts the confusion matrix for the Quarter Power, template-based, classification method using the segmented images as inputs. Each data entry of the table, except the ones of the last column and last row, corresponds to the number of instances of the target class designated by the row identifier that were classified as the target class specified by the column identifier. Off-diagonal entries correspond to misclassifications. Ideally, all off-diagonal entries should be 0 to achieve an error rate of 0%. The table shows that the most frequent confusions (24 of them) were T62 targets erroneously identified as T72s.

		Estimated Class										Classification Accuracy
		2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU23 4	
Actual Class	2S1	252	4	6	0	3	0	3	2	4	0	91.97%
	BMP2	1	559	7	2	4	0	0	9	2	3	95.23%
	BRDM2	0	2	246	8	9	0	0	0	9	0	89.78%
	BTR60	0	0	0	194	0	0	0	1	0	0	99.49%
	BTR70	0	1	5	0	190	0	0	0	0	0	96.94%
	D7	0	0	0	1	1	271	1	0	0	0	98.91%
	T62	3	1	0	2	2	0	227	24	9	5	83.15%
	T72	0	14	0	0	4	0	7	557	0	0	95.70%
	ZIL131	1	1	0	2	6	0	1	6	257	0	93.80%
	ZSU23 4	0	1	0	1	0	2	0	9	0	261	95.26%
Estimate Accuracy		98.05%	95.88%	93.18%	92.38%	87%	99.27%	95%	91.61%	91.46%	97.03%	
		Total Classification Accuracy				**94.10%**						

In order to test the quality of the segmentation process employed, the Quarter Power classification method was applied on the segmented images. Before being classified, from each segmented image, the regions that were identified as non-target (i.e. shadow and clutter) were set to a magnitude of zero, since non-target regions should be irrelevant to the ATR task. In this manner, if the utilized segmentation process retained all important target discrimination information, the

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

classification results should match (or, at least, approximately match) the results obtained, when using the original raw SAR images. These classification results for images of size of 48×48 pixels are shown in Table 1. The overall accuracy obtained is 94.1%, which is very close to the classification accuracy of 95.5%, which is achieved, when using the raw images. This result indicates that the employed segmentation method rejects most of the irrelevant information from the image and produces a very reliable estimate of the target region.

2.2 Elliptical Fourier Descriptors

The goal of this work was to investigate the utility and resulting discriminatory power of target outlines in the ATR task based on SAR images. Toward this end the target region contour needs to be suitably described. These quantities, which are usually referred to as *descriptors*, are then used as features for a classification model. There are many approaches to outline description in the literature. Among a vast plethora of choices, which will not be mentioned here, we opted to use *Elliptical Fourier descriptors*[18] (EFDs), because (i) they always guarantee a closed curve regardless of their values and (ii) in our preliminary experimentations we observed that EFDs were more economic (less number of coefficients gave satisfactory contour approximation) in describing complicated target outlines than other alternatives (e.g. various other Fourier descriptors).

The EFD parametric representation is based on harmonically related ellipses, as we shall see. If C is a smooth, closed curve on the plane represented by a vector $v(t) = [x(t)\ y(t)]^T$ parameterized by $t \in [0,2\pi)$, then $v(t)$ is a periodic function of t and, thus, it can be represented as a Fourier Series according to the following synthesis equation:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \sum_{k=0}^{\infty} F_k \begin{bmatrix} \cos(kt) \\ \sin(kt) \end{bmatrix} \tag{1}$$

The coefficients $F_k \in R^{2\times 2} \quad k = 0,1,2,...$ are 2-by-2 real-valued matrices and constitute the EFDs. Furthermore, they can be calculated via the associated analysis equation:

$$F_k \triangleq \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} = \frac{1}{2\pi} \int_0^{2\pi} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}^T \begin{bmatrix} \cos(kt) \\ \sin(kt) \end{bmatrix} dt \tag{2}$$

F_o, the zero[th] EFD in the series expansion, is of diagonal form; in this case, the vector $[a_o\ c_o]^T$ describes the contour's barycenter. Each term in the summation depicted in Eq. 1 describes an ellipse on the plane. Moreover, the reconstruction of the contour can be viewed as the superposition of phasors and is accomplished by computing Eq. 2 over the interval $(0,2\pi)$. Each phasor $[\cos(kt)\ \sin(kt)]^T$ rotates at a rate proportional to its harmonic number k. While rotating, it traces the outline of an ellipse, whose major and minor axes sizes are determined by the coefficient matrix. In other words, a contour, whose EFDs are assumed to be known, can be described as a superposition of simultaneously rotating phasors, each one tracing an ellipse determined by the appropriate EFD, at angular speeds that are multiples of a fundamental one. The advantage of these elliptical orbits is that they allow for a good estimate of the target's shape, even if only a limited amount of low-order terms are employed. A graphical representation of the reconstruction of an outline using EFD's is shown in Fig. 2.

It is interesting to observe that the first (k=1) EFD describes an ellipse, whose orientation and size roughly resemble the ones of the target region. Fig. 3 shows the first ellipse superimposed over a target image. As demonstrated in the figure, the length of the semi-major axis is roughly half the length of the target region. Moreover, the length of the semi-minor axis is roughly half the width of the target region. Also, the angle of the semi-major axis gives an indication as to the pose of the target. Here it is important to mention that, since these descriptors inherently contain an estimate of the pose, there is no need to estimate the pose in a separate step as is done by other ATR systems[6,10]. Most of this information is contained within the first EFD.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

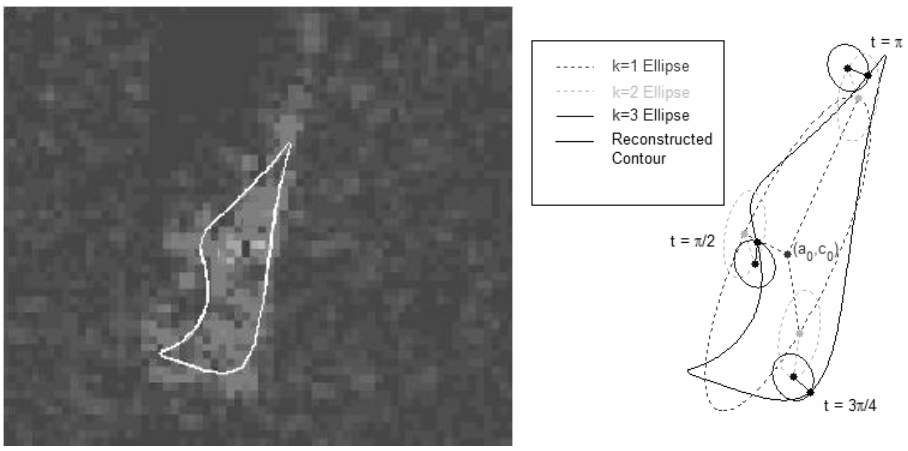

Fig. 2: Graphical representation of the reconstruction of a target outline. The outline of the target region shown in the left panel has been reconstructed using only the first 3 EFDs (k=1,2,3). The right panel shows the 3 corresponding ellipses and phasors that take part in the reconstruction. A snapshot of each phasor is represented as a straight line from the center of an ellipse to its corresponding edge. As the reader can witness, even a small amount of EFDs generate a contour with sharp parts.

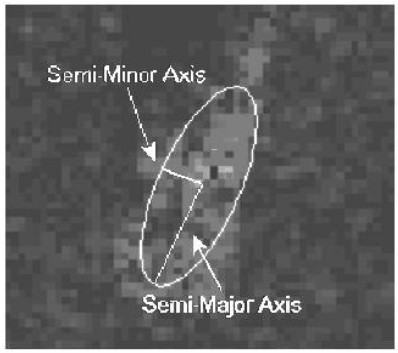

Fig. 3: Here we show the same SAR image as in Fig. 2 and we overlay the ellipse traced by the first phasor (k=1). As seen, the semi-minor and semi-major axes are approximately proportional to the width and length of the target region respectively. Additionally, the orientation of the axes with respect to the image gives a relatively good indication of the pose of the target.

2.3 The Classification Stage

The final step in our approach is the classification of target images based on the EFD features discussed in the previous subsection. For every segmented image we extracted the first 4 EFDs along with the EFD of k=0, which resulted in a single 18-dimensional feature vector per image (4 values for each EFD, except for k=0, where we use only its diagonal values; the off-diagonal values are always equal to zero for k=0). These feature vectors were then used to train a *Support Vector Machine* (SVM) model using the *Sequential Minimization Optimization* algorithm[19]. The SVM classifier was introduced by Vapnik[20]. Please refer to the previous reference for a complete description of the theory behind SVMs. A search of the recent literature reveals that SVMs have already been applied to the classification of MSTAR images[16]. This study has opted for pursuing SVMs for the ATR classification stage for two reasons: (i) As we have discussed earlier, there are only a few available patterns (images) for each subclass (about 5.1 on average) with respect to the input space dimensionality (18 features) and SVMs have good generalization properties, when dealing with difficult problems like the one at hand. (ii) SVM's are popular models and have been yielding good results in a variety of settings and application domains over the last decade. In that sense, they seem to represent a trusted choice among other alternatives. We need to note here, that a SVM is only a binary classifier, that is, it can only handle 2-class problems. In order to handle the multi-class recognition problem at hand, a method of combining the output of several binary SVMs was

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

needed. Two of the methods considered in this study were *Max Wins*[21] and *Directed Acyclic Graph SVM* (DAGSVM)[22]. In the next section we discuss the experimental results we obtained with our approach.

3. EXPERIMENTAL RESULTS

In this section, we showcase the majority of experimental results we have obtained. In specific, the classification accuracy using different numbers of EFD's and classifiers is presented. Then, the confusion matrix of our method's best performance is shown. Additionally, we compare our process to a few selected, existing approaches in terms of classification performance and computational complexity.

As mentioned in the Introduction, the data we used for benchmarking purposes was the MSTAR collection[1], which the United States Air Force released to the public for research purposes. In this collection, each image contains a single target vehicle. For the purposes of classification, it is assumed that all images belong to one of ten classes and that the target is centered in the image. Table 2 lists the number of training and test images from the dataset that were used. The third column in Table 2 shows the average number of training samples that were available for each class.

Table 2: Number of images per class included in the training and test sets.

Vehicle Class	Number of Training Images	Number of Testing Images
2S1	299	274
BMP2	697	587
BRDM2	298	274
BTR60	256	195
BTR70	233	196
D7	299	274
T62	299	273
T72	691	582
ZIL131	299	274
ZSU23 4	299	274
Entire Set	**3670**	**3203**

In order to cope with the small dataset sizes involved, we used images that were recorded at a 17° compression angle for training, and the ones recorded at 15° compression angle for testing purposes. The legitimacy of this experimental setup stems from the fact that the differences between corresponding SAR images at 15° and 17° are assumed negligible, as has been done in prior works[3].

The classification accuracies for the 10-class ATR problem using k=2 to 20 EFDs and either the DAGSVM or the Max Wins SVM classification schemes are graphically shown in Fig. 4. The best classification rate of 86.67% was achieved using 7 EFDs and the Max Wins scheme, as shown in Table 3. Using 20 EFDs yielded an accuracy of 86.38%, which is nearly the same classification rate as in the optimal case. With only 5 EFDs, the classification rate was slightly lower, that is, 85.49%. Moreover, we observed that the rate decreased rapidly as we decreased the number of EFDs. The conclusions to be drawn are as expected: crude characteristics of the target region's outline suffice to classify a target; additional contour detail does not add any significant discriminatory information.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

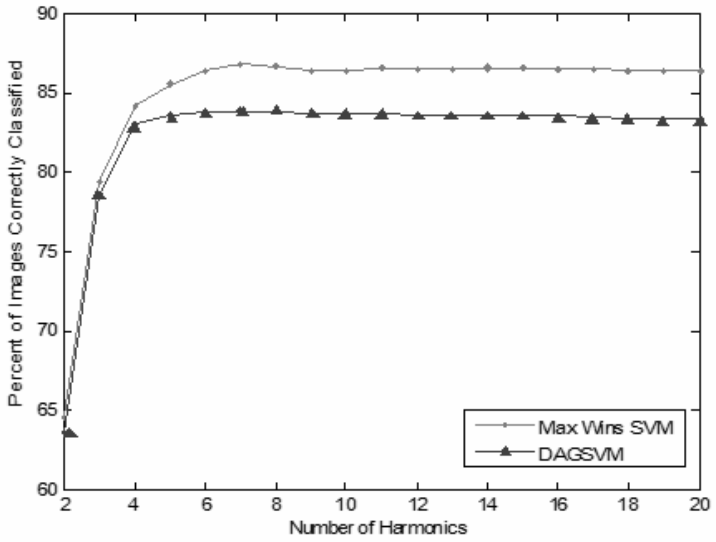

Fig. 4: Performance of our approach in terms of classification accuracy versus the number of EFD coefficients utilized as features. Results are shown for both Max Wins and DAGSVM classification schemes. In our experiments the Max Wins scheme yielded the best results, when 7 EFDs were used.

Table 3: Confusion matrix, when using Max Wins SVM with 7 EFDs (26 features). The most frequent errors occurred again, when a large proportion of T62 targets was misclassified as T72s.

		Estimated Class										Classification Accuracy
		2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU23 4	
Actual Class	2S1	212	32	4	1	5	0	7	6	6	1	77.37%
	BMP2	1	553	2	0	2	0	2	24	2	1	94.21%
	BRDM2	0	17	197	7	10	0	0	2	2	0	83.83%
	BTR60	2	12	4	150	5	0	2	2	5	4	80.65%
	BTR70	5	24	4	1	153	0	1	1	2	0	80.10%
	D7	0	8	0	0	0	264	0	0	0	2	96.35%
	T62	9	6	1	6	1	0	197	42	8	3	72.16%
	T72	2	16	0	3	1	0	7	549	3	1	94.33%
	ZIL131	7	17	0	13	13	0	1	15	204	4	74.45%
	ZSU23 4	1	9	1	0	0	1	0	5	6	251	91.61%
Confidence		88.70%	79.68%	92.49%	82.87%	80.53%	99.62%	90.78%	84.98%	85.71%	94.01%	
		Total Classification Accuracy				**86.67%**						

Table 4 and Table 5 show the performance comparison of the EFD-based classification scheme to those of previous works. The tables also list the amount of time taken to perform the calculations during the training and testing periods. It is important to note that the calculation times listed include the time it takes to read a SAR image from a hard drive. All calculations were performed in Matlab™ on a 1.6 GHz Intel Core2Duo Processor. The SVM classification schemes were implemented using a freely available package for Matlab™ called the *Support Vector Machine Toolbox Version 0.5*[23]. In general, it was difficult to compare the EFD-based classification results with the ones of previous studies, as many of them used subsets of the available images in the MSTAR dataset for both 3-target and 10-target ATR problems[11,12,24]. Some studies considered the pose (azimuth) to be known a priori, which significantly simplified the ATR problem (see for example Cetin [9] and others[11]) and only a few did otherwise[3,5,10,25]. Table 4 shows performance comparisons of the methods used by DeVore[3] to the performance of the ATR system using 7 EFDs and the Max Wins scheme. Obviously,

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

the use of EFD's significantly reduces the number of features used by the classifier; n EFD coefficients produce $2(2n\text{-}1)$ features, since each EFD is described by four coefficients except for the zeroth EFD, which is a diagonal matrix. For template-based methods, the intensity of each pixel is considered to be a feature. Although using EFDs has a classification accuracy of about 10% lower than the template-based methods, it employs two orders of magnitude (88.6 times, to be exact) less features than the template-based methods, when using 48×48 images. A graphical representation of Table 4 is shown in Fig. 5. For the 10-class ATR problem, our EFD-based approach still seems to be less accurate and slower (by an order of magnitude, when training) compared to other methods listed in Table 4. However, testing times seem promisingly competitive due the fast feature extraction procedure (segmentation, target region outline identification, EFD computations) and due to the smaller number of features entering the final classification stage.

Table 4: 10-Class ATR Performance Comparisons in terms of features, classification accuracy and computational complexity, which is measured in running time.

Feature Set	Classifier	Number of Features	Classification Accuracy	Training Time (sec)	Testing Time (sec)
7 Elliptical Fourier Descriptors	DAGSVM	26	83.78%	258.55	287.48
7 Elliptical Fourier Descriptors	Max Wins	26	86.67%	775.86	331.41
QP Normalized Image (128x128)[3]	Nearest Neighbor	16384	97.17%	138.95	2005.3
QP Normalized Image (48x48)[3]	Nearest Neighbor	2304	95.50%	98.84	378.08
Conditionally Gaussian Full Resolution (96x96)[3]	Max Likelihood	9216	97.70%		

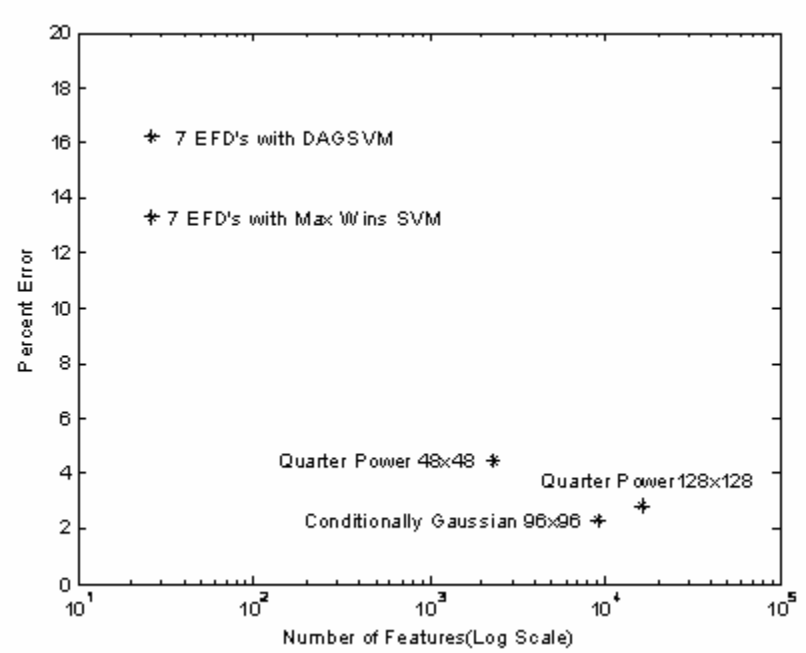

Fig. 5: 10-Class ATR Performance Comparisons of percent misclassification error versus number of features for the methods listed in Table 4.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

The performance of our EFD-based approaches improves, when limited to discriminating between only three target classes. Table 5 compares the accuracy of EFDs to a few, well-known alternative methods. The accuracy of the EFD-based schemes is approximately 3% higher than the raw magnitude SVM[25] and MINACE implementations[5]. In addition, although the accuracy is about 6% less than the *AdaBoost* method, it uses 246 times less features. Furthermore, it can almost surely be assumed that the AdaBoost-based scheme's training algorithm is much more computationally complex than the combined extraction of EFD features and subsequent training of the Max Wins SVM classifier ensemble. The EFD-based method also has the advantage that the pose does not need to be estimated explicitly as it is in the AdaBoost method. A graphical representation of Table 5 is shown in Fig. 6. The evaluation is relatively fast. 3203 test images were classified in about 287 seconds with the Max Wins method. That is about 11 images per second.

Table 5: 3-Class ATR Performance Comparisons in terms of feature set size and classification accuracy.

Feature Set	Classifier	Number of Features	Classification Accuracy	Training Time (sec)	Testing Time (sec)
7 Elliptical Fourier Descriptors	Max Wins SVM	26	93.82%	258.55	287.48
7 Elliptical Fourier Descriptors	DAGSVM	26	93.46%	162.34	144.48
80x80 Normalized Image[25]	SVM	6400	90.99%		
MINACE[5]	Max Value	4096	90.60%		
AdaBoost[10]	RBF Neural Net	12800Train (6400 Test)	99.63%		

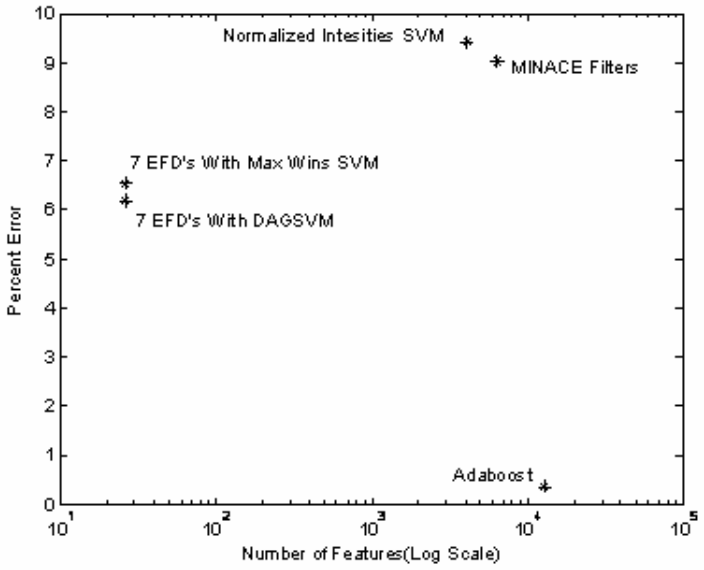

Fig. 6: 3-Class ATR Performance Comparisons of percent misclassification error versus number of features for the methods listed in Table 5.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

4. SUMMARY & CONCLUSIONS

Even with the high levels of distortion and variation in the MSTAR images, many ATR systems have produced classification rates from 80% to close to 100%. Nevertheless, even the most successful ones utilize the entire image as a feature vector, which even includes sub-regions of the image with no class-relevant information, such as ground clutter. The increased dimensionality of the associated feature space makes classification of unlabeled samples computationally expensive, at least to some degree. In this paper we have reported a straightforward and effective segmentation procedure which, as we have shown, manages to retain almost all discriminatory information in the identified target region. Furthermore, we have demonstrated that Elliptical Fourier descriptors (EFDs) provide efficient means to parameterize the closed contour outlines of MSTAR vehicles. In specific, because of the large length-to-width ratio of many targets, the EFDs allow for an excellent representation of the target outline employing only very few descriptors. Also, the implicit estimate of target pose from the orientation of the semi-major axis of the first EFD eliminates the need to make a separate estimate of target pose as is required by other methods. As illustrated in the paper, the descriptors can be used by powerful classification methods, such as Support Vector Machines, to determine target class with an accuracy of 86.7%, when classifying ten different types of vehicles of the MSTAR dataset. This demonstrates that the target region outline description can be very useful and effective as a high-level classification feature. To the best of our knowledge, the accuracy rate using EFDs is much higher than previous attempts to use other high-level features for classification purposes. The studied approach still lags by roughly 10% for a 10-class problem or leads by just a 3% for a 3-class problem in terms of classification accuracy, when compared to a few models championed in the recent literature. Nevertheless, a clear advantage of the proposed scheme is the small amount of features utilized and the advantage it presents in terms of computations, when attempting to classify a previously unseen test image. The encouraging results obtained through these reported results prompt for future efforts that ought to focus on increasing the recognition accuracy of the overall ATR task through a methodical and extensive evaluation of more classification schemes based, once again, on EFD features.

ACKNOWLEDGEMENTS

This material is based upon work/research supported in part by the National Science Foundation under Grant No. 0647120, Grant No. 0647018, Grant No. 0717680 and Grant No. 0717674. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] *Moving and Stationary Target Acquisition and Recognition(MSTAR) Public Dataset*. https://www.sdms.afrl.af.mil/datasets/mstar/index.php

[2] R. Duda, P. Hart, and D. Stork. *Pattern classification, 2nd edition*. John Wiley & Sons. Inc., New York, 2001.

[3] M. D. DeVore, and J. A. O'Sullivan. "A performance-complexity study of several approaches to automatic target recognition from synthetic aperture radar images." IEEE Transactions on Aerospace Electronic Systems **38**(2), 632-648 (2002).

[4] L.M. Kaplan. "Analysis of multiplicative speckle models for template-based SAR ATR." IEEE Transactions on Aerospace Electronic Systems, **37**(4), 1424-1432 (2001).

[5] R. Patnaik and D. Casasent. "MINACE filter classification algorithms for ATR using MSTAR data." *Proceedings of the SPIE, Automatic Target Recognition XV*, 5807, 101-111 (2005).

[6] A. Mahalanobis, D. Carlson, and B.V. Kumar. "Evaluation of MACH and DCCF correlation filters for SAR ATR using the MSTAR public database." *Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery V*, 3370, 460-468 (1998).

[7] R. Singh and B.V. Kumar. "Performance of the extended maximum average correlation height (EMACH) filter and the polynomial distance classifier correlation filter (PDCCF) for multiclass SAR detection and classification." *Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery IX*, 4727**,** 265-276 (2002).

[8] R. Williams, J. Westerkamp, D. Gross, A. Palomino and T. Fister. "Automatic target recognition of time critical moving targets using ID high range resolution (HRR) radar." IEEE AES System Magazine, 37-43 (2000).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

[9] M. Cetin, W.C. Karl and D.A. Castanon. "Feature Enhancement and ATR Performance Using Nonquadratic Optimization-Based SAR Imaging." IEEE Transactions on Aerospace and Electronic Systems, **39**(4), 1375-1395 (2003).
[10] S. Yijun, Z. Liu, S. Todorovic, and J. Li. "Synthetic Aperture Radar Automatic Target Recognition Using Adaptive Boosting." *Proceedings of the SPIE, Algorithms for Synthetic Aperture Radar Imagery XII*, 5808, 282-293 (2005)
[11] J. Saghri and C. Guilas. "Hausdorff Probabilistic Feature Analysis in SAR Image Recognition." *Proceedings of SPIE, Applications of Digital Image Processing XXVIII*, 5909**,** 21-32 (2005).
[12] B. Ravichandran, A. Gandhe, and R. Smith. "XCS for Robust Automatic Target Recognition." *Proceedings of the 2005 Conference on Genetic and Evolutionary Computation*, 1803-1810 (2005).
[13] T. Pink and U. Ramanathan. "Intelligent Selection of Useful Features for Optimal Feature-Based Classification." *IGARSS*, 7, 3012-3014 (2000).
[14] B. Bhanu and G. Jones III. "Exploiting Azimuthal Variance of Scatterers for Multiple Look SAR Recognition." *Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery IX*, 4727, 290-298(2002).
[15] A. Kim, S. Dogan, J. Fisher III, R. Moses, and A. Willsky. "Attributing Scatterer Anisotropy for Model Based ATR." *Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery VII*, 4053**,** 176-188 (2000).
[16] K. Krawiec and B. Bhanu."Visual Learning by Coevolutionary Feature Synthesis." IEEE Transactions on Systems, Man, And Cybernetics—Part B: Cybernetics, 35(3), 409 – 425 (2005).
[17] Y. Yang, Y. Qiu, and C. Lu, "Automatic Target Classification Experiments on the MSTAR SAR Images." *SNPD/SAWN'05*, 2-7 (2005).
[18] F. Kuhl and C. Giardina. "Elliptic fourier features of a closed contour." Computer Graphics and Image Processing, **18**(3), 236-258 (1982).
[19] J. Platt. *Advances in Kernel Methods: Support Vector Learning*. MIT Press, Massachusetts, 1998.
[20] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 1995.
[21] J. Friedman. "Another approach to polychotomous classification." Technical Report, Stanford Department of Statistics, 1996.
[22] J. Platt, N. Cristianini, and J. Shawe-Taylor. *Advances in Neural Information Processing Systems 12th Edition*, MIT Press, Massachusetts, 2000.
[23] G. Cawley. http://theoval.sys.uea.ac.uk/svm/toolbox/, 2000.
[24] Q. Pham, T. Brosnan, and M. Smith. "Multistage algorithm for detection of targets in image data." *Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery IV*, 3070**,** 66-75 (1997).
[25] M. Bryant and F. Garber. "SVM classifier applied to the MSTAR public data set." *Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery VI*, 3721, 355-360 (1999).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 06/12/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx