

Article

# Sparse Representation-Based SAR Image Target Classification on the 10-Class MSTAR Data Set

Haibo Song, Kefeng Ji \*, Yunshu Zhang, Xiangwei Xing and Huanxin Zou

Received: 30 September 2015; Accepted: 8 January 2016; Published: 20 January 2016

Academic Editor: Antonio Fernández-Caballero

College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China; songhaibo@nudt.edu.cn (H.S.); zyshuu@sina.com (Y.Z.); xingxiangwei@nudt.edu.cn (X.X.); hxzou2008@163.com (H.Z.)

\* Correspondence: jikefeng@nudt.edu.cn; Tel./Fax: +86-731-8457-6350

**Abstract:** Recent years have witnessed an ever-mounting interest in the research of sparse representation. The framework, Sparse Representation-based Classification (SRC), has been widely applied as a classifier in numerous domains, among which Synthetic Aperture Radar (SAR) target recognition is really challenging because it still is an open problem to interpreting the SAR image. In this paper, SRC is utilized to classify a 10-class moving and stationary target acquisition and recognition (MSTAR) target, which is a standard SAR data set. Before the classification, the sizes of the images need to be normalized to maintain the useful information, target and shadow, and to suppress the speckle noise. Specifically, a preprocessing method is recommended to extract the feature vectors of the image, and the feature vectors of the test samples can be represented by the sparse linear combination of basis vectors generated by the feature vectors of the training samples. Then the sparse representation is solved by  $l_1$ -norm minimization. Finally, the identities of the test samples are inferred by the reconstructive errors calculated through the sparse coefficient. Experimental results demonstrate the good performance of SRC. Additionally, the average recognition rate under different feature spaces and the recognition rate of each target are discussed.

**Keywords:** synthetic aperture radar; classification; sparse representation

---

## 1. Introduction

Synthetic Aperture Radar (SAR), an active sensor, has been widely applied in many areas such as disaster assessment and military defense, due to its ability to work against 24 h a day and severe weather. With the high-resolution SAR coming to work, it is hard to carry on manual interpretation, thus making the Automatic Target Recognition (ATR) popular. ATR uses the signal processing techniques to recognize the unknown targets. Moreover, a typical ATR in SAR has three separate stages: detection, discrimination and classification [1]. Detection focuses on finding local Regions of Interest (ROI) which include targets and numerous false alarms. Then the following stage, discrimination, is to filter natural clutter out. Finally, the identities of targets are predicted and the man-made clutter is rejected. The classification of unknown targets is discussed in this article.

The SAR images contain coherent speckle noise, which lowers the images quality significantly, so it is very difficult to interpret them. Despite that, many works have been done in researching SAR. Firstly, template matching-based algorithms are utilized to achieve classification [2]. The methodology usually defines “distances” between the test sample and the templates generated by the training set. Then the identity of the test sample will be assigned to the class to which the matched template selected by the “distances” belongs. So the addition of an object requires creating an additional set of templates, thus causing burdensome calculation. To improve the performance, a number of methods called correlation pattern recognition have been presented [3], and they accomplish the training of the

filter through minimizing the correlation between the filter and the spectral envelope of the training set in the frequency domain. Beyond that, a Learning Vector Quantization (LVQ) has been used for learning the templates for classification in [4]. Then, the nonlinear classifier is used for the SAR ATR, e.g., [5], and its performance is better than the conventional template-based approaches. All these algorithms reduce their complexity by applying a set of classifiers trained by the training samples over a given range of aspect angles, so the performances of the algorithms are limited by the accuracy of aspect angle estimate.

Recently, a considerable resurgence of interest in the sparse representation of signals has been witnessed [6–10]. Much of its effectiveness is caused by the fact that signals can be well represented by a set of basis vectors in most cases. Given the basis vectors, finding the sparse representation of a signal is processed as an optimal problem. The optimal problem, called  $l_0$ -norm minimization, is known to be a NP-hard problem, which can be solved by the Greedy algorithm [11,12] or the convex relaxation algorithm [13]. The Sparse Representation-based Classifications (SRC) are recommended firstly in [14], and then the framework is extensively utilized in many classifications [6,7]. The framework is re-explained from the perspective of class manifolds in [1]; meanwhile, it is applied in the classification on three-class MSTAR targets. Then SRC on three-class MSTAR targets (BMP2, BTR70 and T72) is further studied in [8–10], which focuses on extracting effective features to improve the classification performance. Simultaneously, a better classification performance of SRC over SVM has been verified. It is noticeable that the MINACE filter and SVM are used in recognition of the three-class targets, respectively [3,5]. Then, the two methods are extended to the recognition of the 10-class targets [15,16]. Therefore, in this paper, SRC is extended to classify the 10 targets, which really is more challenging than three-class targets.

Because the raw MSTAR images have high dimensionality and speckle noise, thus making the classification more complicated and influencing the recognition performance worse, the sizes of the raw MSTAR images are not unified, and a preprocessing of MSTAR images is recommended in this paper to maintain the useful information, suppress the speckle noise and reduce the dimensionality. Meanwhile, conventional and unconventional feature extraction methods are utilized to extract features and reduce the dimensionality further. Then, the processed images are inputted into the framework, and the sparse coefficient of the test sample is worked out. So the identity of the test sample is inferred by the reconstructive errors calculated through the sparse coefficient. The comparisons between SRC and SVM are carried out to verify the excellent performance of SRC. Besides the average recognition accuracy, the recognition rate of each target is discussed.

The organization of the paper is as follows: we give a brief review of the sparse representation-based classification in Section 2. Section 3 describes the preprocessing of the MSTAR images before putting them into the classifier. The results compared with SVM will be presented in Section 4. Finally, the paper is concluded in the Section 5.

## 2. Sparse Representation-Based Classification

As described in the previous section, the training set has been given and the goal of the classification is to decide the class of the unknown samples correctly. Assume that there are  $K$ -class targets, and the  $i$ -th class has  $n_i$  samples. If every sample is represented in the form of a vector, the  $n_i$  samples can be concatenated into a matrix  $\phi_i \in R^{m \times n_i}$ :

$$\phi_i = [\phi_{i,1}, \dots, \phi_{i,n_i}] \quad (1)$$

where  $m$  is the dimension of the vector, and  $\phi_{i,j}$  ( $j = 1, \dots, n_i$ ) stands for  $j$ -th sample in the class  $i$ . When the number of samples in class is sufficient, the test sample  $y$  belonging to class  $i$  can be well approximated by the linear combination of these samples:

$$y = c_{i,1}\phi_{i,1} + \dots + c_{i,n_i}\phi_{i,n_i} \quad (2)$$

where  $c_{i,j}$  ( $j = 1, \dots, n_i$ ) is the contribution of the  $j$ -th sample to reconstructing sample  $y$ . However, the class to which the test sample  $y$  belongs is unknown, in fact. All matrices  $\phi_i$  ( $i = 1, \dots, K$ ) need to be concatenated to a big matrix  $\phi = [\phi_1, \dots, \phi_K] \in R^{m \times N}$  named dictionary, where  $N$  represents the number of training samples. So, the unknown sample  $y$  can be represented by all training samples:

$$y = \phi x \tag{3}$$

where  $x = [c_{1,1}, \dots, c_{1,n_1}, \dots, c_{i,1}, \dots, c_{i,n_i}, \dots, c_{K,1}, \dots, c_{K,n_K}]$  is the weight vector. Thanks to the sufficient samples in every class, the weight vector  $x$  is expected to be  $x = [0, \dots, 0, c_{i,1}, \dots, c_{i,n_i}, 0, \dots, 0]$ . In other words, the test sample  $y$  can be well approximated by the samples, which belong to the same class as  $y$ , and additionally,  $y$  has no relationship with the other samples, ideally.

Considering the formulation in Equation, the test sample  $y$  and dictionary  $\phi$  are given, and the weight vector  $x$  is to be solved. Moreover, the solution  $x$  varies with the dictionary  $\phi \in R^{m \times N}$ . When  $m > N$ , the equality is an overdetermined system, so the unique solution  $x$  can be obtained by solving the formulation. In most cases, however, the equality is an underdetermined system, i.e.,  $m < N$ , thus there are many solutions meeting the formulation. Therefore, a regularization constraint is needed to sift the optimal solution. Since a test sample  $y$  can be well represented by the training samples from the same class, the weight vector  $x$  is expected to have a lot of zero entries, i.e.,  $x$  is sparse. So, the regularization constraint can be defined as  $l_0$ -norm minimization, that is:

$$\min_x \|x\|_0 \quad \text{s.t.} \quad y = \phi x \tag{4}$$

where  $\|x\|_0$  counts the number of non-zero entries in vector  $x$ . Given that the test sample  $y$  is represented approximately by the training samples, the equality constraint in Equation (4) allows a small error. So the expression in Equation (4) can be rewritten as:

$$\min_x \|x\|_0 \quad \text{s.t.} \quad \|y - \phi x\|_2 \leq \varepsilon \tag{5}$$

where  $\varepsilon$  is the allowed error, which is a small positive integer. Solving the  $l_0$ -norm minimization in Equations(4) or Equation (5), however, is a NP-hard problem, which can be solved by OMP [11], and only an approximate solution will be acquired. Fortunately, it has been proved that if  $x$  is sparse enough, the optimization problem can be relaxed to  $l_1$ -norm minimization:

$$\begin{aligned} \min_x \|x\|_1 \quad \text{s.t.} \quad y = \phi x \\ \min_x \|x\|_1 \quad \text{s.t.} \quad \|y - \phi x\|_2 \leq \varepsilon \end{aligned} \tag{6}$$

In this article, the second formulation in Equation (6) is taken into consideration, and the method to solve the minimization adopted here is second-order cone programs (SOCP) [13].

After the optimal solution  $x$  is settled, the class of the test sample  $y$  will be worked out by the minimal reconstruction error:

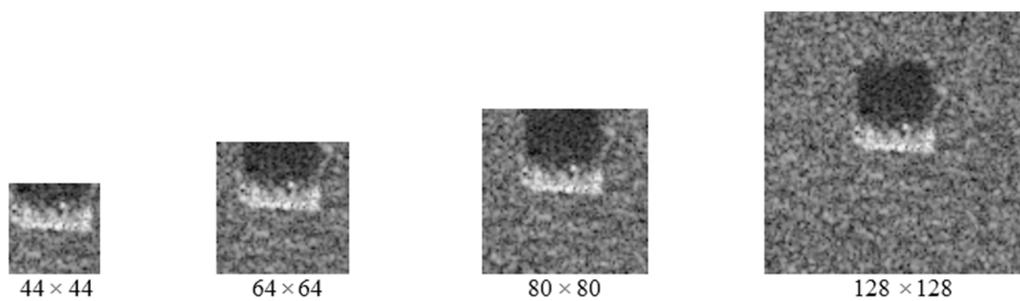
$$\operatorname{argmin}_i \|y - \hat{y}_i\| = \operatorname{argmin}_i \|y - \phi \delta_i(x)\| \tag{7}$$

where  $\delta_i(\cdot)$  is an operator that selects the weight coefficients belonging to class  $i$  and sets all the other coefficients to be zero. So  $\hat{y}_i$  is a linear combination of the training samples belonging to class  $i$ . Finally, the  $i$  meeting the objective function is the class label of  $y$ .

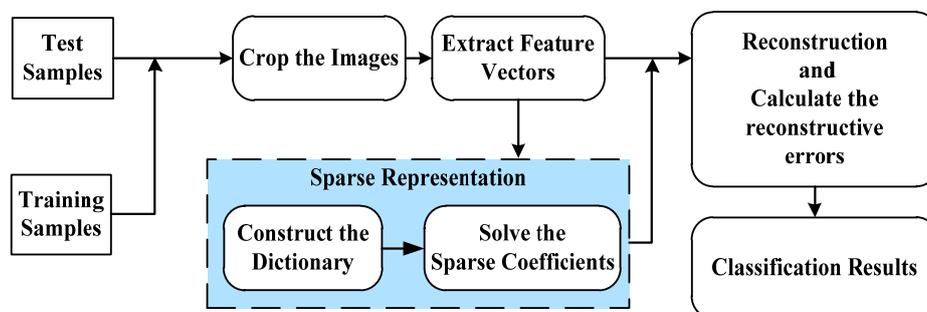
### 3. Classification on the 10-Class MSTAR Targets

The standard moving and stationary target acquisition and recognition (MSTAR) database [17] has 10 SAR targets, which will be discussed in the next section. Instead of putting the MSTAR images into the SRC directly, some measures should be taken to reduce the dimensions of the MSTAR images.

The reasons are that the dimensions of the MSTAR images are so large that it makes the classifier take too much time, and simultaneously, the speckle noise in the MSTAR images will influence the performance of the classifier; moreover, the dimensions of the raw SAR images are not consistent with each other, e.g., images of BTR70 are  $128 \times 128$  pixels, 2S1 are  $158 \times 158$  pixels, T62 are  $172 \times 173$  pixels. So it is necessary to reduce the dimensions, and before that, we need to unify the sizes of the images. There is a registration method in [16] to center the training and test images. Then the images are cropped into the same size according to the “center”. However, we do not think the method should be adopted here to achieve the unification. The shadow of the target in the SAR image is useful for classification, but the method, however, calculates the “center” according to the intensity of the image, and then it crops the “center” in a  $44 \times 44$  pixel patch, which barely contains the shadow as seen in Figure 1. The reason is that the intensity of the shadow is even lower than the speckle noise, thus making the “center” located between the target and speckle noise. The center  $64 \times 64$  pixel patch adopted in [10] does not contain enough information of the shadow. Therefore, the center  $80 \times 80$  pixel patch, whose center is located according to the size of the raw image, is adopted here because it contains nearly the whole shadow and excludes most speckle noise.



**Figure 1.** The  $128 \times 128$  pixel patch is one of the BTR70 raw images, and the other three patches are cropped patches from the  $128 \times 128$  pixel patch.



**Figure 2.** The flow-process diagram of SRC.

After the unification of all images, conventional and unconventional feature extraction methods are applied to extract the features and reduce the dimensions of the cropped images. Then, the feature vectors are inputted into the SRC to achieve the classification. The overall procedure of the classification is summarized in Figure 2. The two successive steps, cropping the images and extracting feature vectors, are applied to both test and training samples to generate the test and training vectors, respectively. Then, the feature vectors are inputted into the Sparse-Representation module, which constructs the dictionary and solves the sparse coefficients. The dictionary will be achieved by the training vectors, and it is combined with the test vectors to solve the sparse coefficients. The dictionary and sparse coefficients are to be utilized to reconstruct the test vectors. Then the reconstructive errors are calculated to infer the identities of the test samples.

### 4. Experiment Results

The MSTAR database [17] has 10 SAR targets, which are BMP2, T72, T62, BTR70, BTR60, 2S1, BRDM2, D7, ZIL131 and ZSU234. The images of these targets are collected at two different depression angles (15° and 17°) over a full 0°–360° range of aspect view. The experimental procedure adopted in this article follows the guideline [18], which states that classifiers using this database are to be trained using targets at the 17° depression angle and tested on targets at the 15° depression angle. Simultaneously, the targets BMP2 and T72 have different variants (BMP2-9563, BMP2-9566, BMP2-C21, T72-132, T72-812 and T72-S7), which should be included in the test set. The training set, however, only has variants BMP2-C21 and T72-132, because the classifier should recognize the variants not present in the training set. The number of each target available at the 17° training set and the 15° test set depression angles are listed in Table 1. Meanwhile, the optical images of the 10 targets are displayed in Figure 3.

Table 1. The number of images of each object at different depression angles.

Targets	BMP2	BTR70	T72	BTR60	2S1	BRDM2	D7	T62	ZIL131	ZSU234
17°	233	233	232	256	299	298	299	299	299	299
15°	587	196	582	195	274	274	274	273	274	274



Figure 3. The optical images of the 10 targets. Among them, the T62 and T72 are tanks. The BRDM2, BMP2, BTR60 and BTR70 are armored personnel carriers. The 2S1 is a rocket launcher, the D7 is a bulldozer, the ZIL131 is a truck and the ZSU234 is an Air Defense Unit.

Table 2. The recognition rates of the cropped images using SRC.

Patch	44 × 44 Pixel	64 × 64 Pixel	80 × 80 Pixel
Accuracy	34.1%	64.3%	75.1%

The preprocessing of the SAR images was discussed in last section. Table 2 shows the recognition rates of the three cropped images under SRC. From the table, the 80 × 80 pixel patch indeed contains more effective information than the other patches. After unifying the sizes of the images, the unconventional methods, down-sample and Principal Component Analysis (PCA), are utilized to extract the features of the cropped images. The down-sample is achieved by bilinear interpolation here, and the output pixel value is a weighted average of pixels in the nearest two-by-two neighborhood. In comparison with the down-sample, the dimensions of the feature vector extracted by PCA are the same as the former. Meanwhile, the conventional features, baseline features [19], are adopted here to make a comparison with the unconventional features. In order to achieve excellent performance of SVM [20], the multiclass classifier, based on the one-against-one approach, adopts the feature normalization method if there is no additional stress. Simultaneously, because the different features are

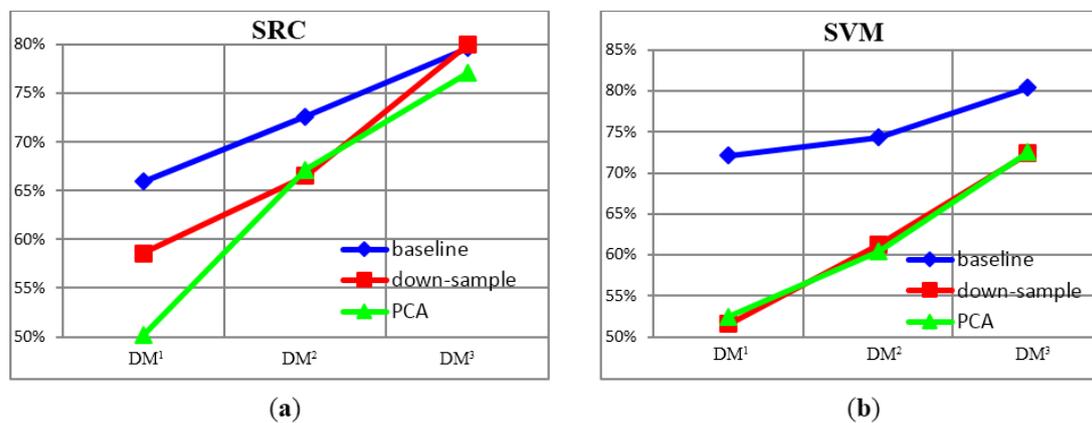
employed here, SVM uses a linear method to deal with down-sampling features and baseline features, and uses the Gaussian kernel method to process PCA features.

#### 4.1. Recognition Accuracy of Two Classifiers

Owing to the finite dimensions of the baseline features, the comparison under the two classifiers between the conventional features and the unconventional features is listed firstly. Then, the dimensions of the unconventional features are enlarged to verify the performances of the two classifiers further.

##### 4.1.1. Comparison between Conventional Features and Unconventional Features

According to [19,21] and the down-sampling method used here, the dimensions of the baseline features adopted are sizes of 16, 36 and 81, which are represented by  $DM^1$ ,  $DM^2$  and  $DM^3$ , respectively. In order to compare them to the conventional features, the unconventional feature extraction methods reduce the dimensions of the cropped image to the same size as the former. Considering the big differences within the baseline features, the SVM does not adopt the feature normalization method in this experiment.



**Figure 4.** The recognition accuracy of the two classifiers under different feature spaces; (a) and (b) refer to the recognition accuracy achieved by SRC and SVM, respectively.  $DM^1$ ,  $DM^2$  and  $DM^3$  stand for the dimensions of feature spaces adopted here that have sizes of 16, 36 and 81, respectively.

Figure 4 shows the recognition accuracy of the two classifiers under different feature spaces. Figure 4a plots the recognition accuracy under different feature spaces of SRC, and Figure 4b plots that of SVM. Compared with the unconventional features, the recognition rates of the two classifiers under baseline features are nearly highest. So in the low-dimensional feature spaces, baseline features are better for describing the targets. Simultaneously, the highest recognition rates of SRC and SVM are almost identical, but SRC implies that the precise choice of feature space is no longer critical [14]. That is to say that when the dimensions of the feature space surpasses a certain threshold, unconventional features perform just as well as conventional features towards SRC. From Figure 4a, the dimension threshold in this experiment is near 80.

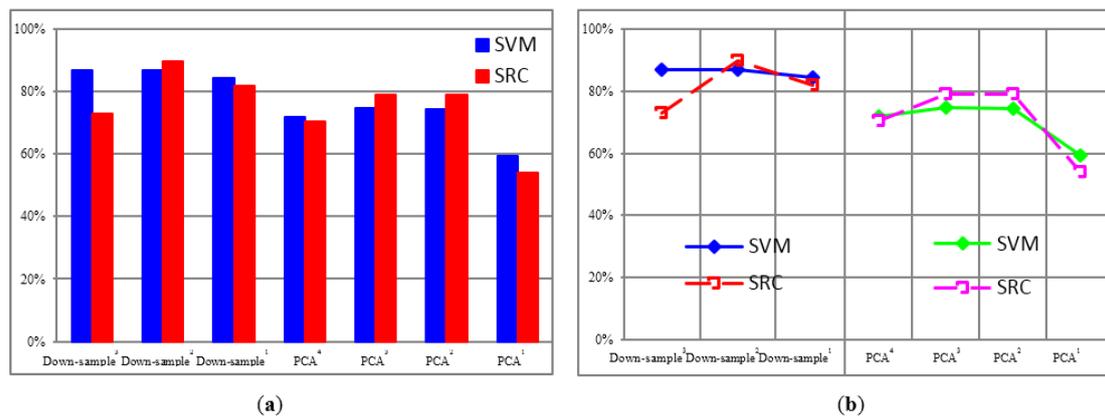
##### 4.1.2. Recognition Accuracy under Unconventional Feature Spaces

The recognition rates of the two classifiers under different unconventional feature spaces are listed in Table 3. Down-sample<sup>1</sup>, Down-sample<sup>2</sup> and Down-sample<sup>3</sup> stand for reducing the dimensions of the cropped image by down-sample to sizes of 100, 400 and 1600, respectively. PCA<sup>1</sup>, PCA<sup>2</sup> and PCA<sup>4</sup> represent that the dimensions of the cropped image are reduced by PCA to sizes of 25, 100 and 400. Additionally, PCA<sup>3</sup> refers to reducing the dimensions to the extent where the accumulative contribution rate approximately equals 80%, because the empirical value guarantees the reduced

dimensions will cover all the primary information. In this experiment, PCA<sup>3</sup> stands for reducing the dimensions to a size of 125. From Table 3, SRC achieves the highest recognition rates in both feature spaces. The recognition rates of SRC are 3.03%, 4.18% and 4.46% better than SVM in Down-sample<sup>2</sup>, PCA<sup>3</sup> and PCA<sup>2</sup>, respectively. Figure 5 is plotted here to study the tendency of the recognition accuracy.

**Table 3.** The recognition accuracy under different unconventional feature spaces.

Features	Down-Sample <sup>3</sup>	Down-Sample <sup>2</sup>	Down-Sample <sup>1</sup>	PCA <sup>4</sup>	PCA <sup>3</sup>	PCA <sup>2</sup>	PCA <sup>1</sup>
SVM	86.79%	86.73%	84.39%	71.84%	74.71%	74.40%	59.19%
SRC	72.93%	89.76%	81.83%	70.34%	78.89%	78.86%	54.04%



**Figure 5.** The recognition rates under different unconventional feature spaces; (a,b) are two forms of the recognition rates. Down-sample<sup>3</sup> refers to reducing the dimensions of the cropped image to a size of 1600 by down-sample. Similarly, Down-sample<sup>2</sup>, Down-sample<sup>1</sup>, PCA<sup>4</sup>, PCA<sup>3</sup>, PCA<sup>2</sup> and PCA<sup>1</sup> stand for sizes of 400, 100, 400, 125, 100 and 25, respectively.

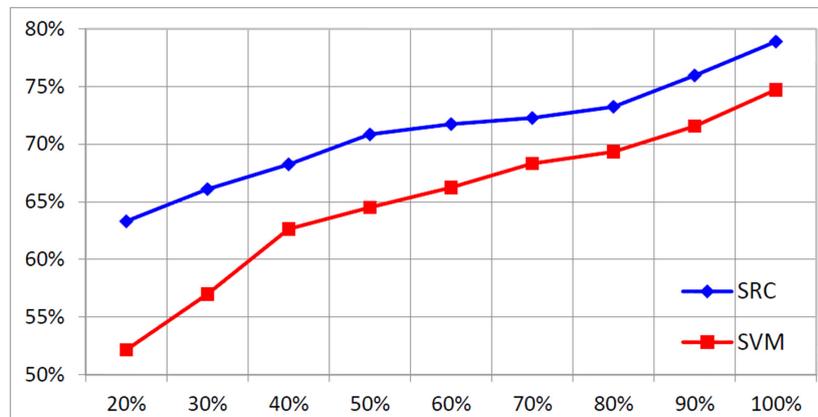
Figure 5 shows the recognition rates under different unconventional feature spaces. Figure 5a,b are two forms of the recognition rates. Figure 5a is a column diagram, and Figure 5b is a line chart. The two feature spaces are divided in Figure 5b. Meanwhile, different colors are adopted here to represent different classifiers. Obviously, SRC is indeed superior to SVM in Down-sample<sup>2</sup>, PCA<sup>3</sup> and PCA<sup>2</sup>. However, with increments of decreases of the dimensions, the performances of both classifiers tend to descend. The reason is that speckle noise will be added into the data with increasing dimensions [10]. Simultaneously, when the dimensions of the data decrease, the data is too small to cover all the primary information. The performance of SRC descends a little faster than SVM. Additionally, the performances of both classifiers in the down-sampling spaces are better than that in the spaces of PCA. That is to say, down-sample contains more effective information than PCA in this experiment.

#### 4.2. Recognition Accuracy under Incomplete Training Samples

In the last subsection, all samples at the 17° depression angle over a full 0°–360° range of the aspect view are used to train the classifiers. In the real-world tasks, the data cannot cover all conditions of the targets. In this experiment, a certain percentage of samples at the 17° depression angle are randomly selected to construct the training set, and all the samples at the 15° depression angle are tested. The feature space, PCA<sup>3</sup>, is adopted here, which is to say that the dimensions of samples are reduced to the extent where the accumulative contribution rate is approximately equal to 80% before classification.

Figure 6 shows the recognition rates under different percentages of training samples. The recognition rates of the two classifiers rise with the increasing percentage of the training samples.

In other words, both the classifiers are sensitive to the number of samples in the training set. Additionally, with the decreasing percentage of the training samples, the recognition accuracy of SVM descends faster than that of SRC, so SRC is more robust than SVM in this experiment.



**Figure 6.** The recognition rates under different percentages of training samples. The horizontal coordinate stands for the percentage of training samples, and the vertical coordinate represents the recognition accuracy.

#### 4.3. Recognition Rate of Each Object

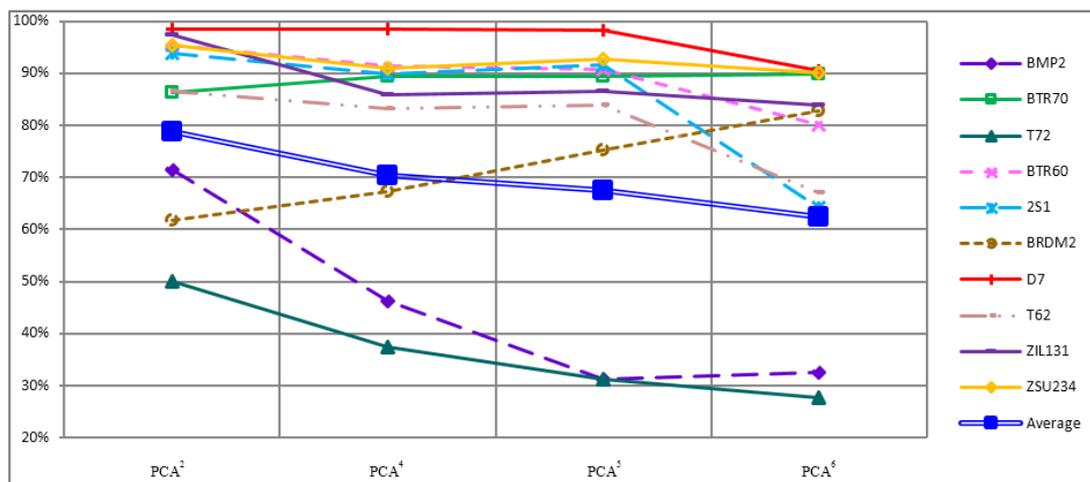
The average recognition rates of the classifiers are an objective evaluation index of their performance. Simultaneously, further information is found by listing the recognition rate of each target here. The confusion matrix is a good reflection of the recognition rate of each object. However, not all confusion matrices in the feature spaces of the two classifiers are needed to be listed here, because the information they try to convey is similar. Table 4 lists the 10-class target confusion matrices of the two classifiers under the feature spaces of Down-sample<sup>2</sup>, PCA<sup>4</sup> and PCA<sup>2</sup>, which are explained above.

Firstly, it is common that the recognition rates of BMP2 and T72 are very low. The reason is that there are many variants of the two targets, which are not included in the training set, in the test set. So the variant has a more important influence in classification than the depression angle in this experiment for only 2° difference in the depression angle. With an increasing difference in depression angle, the recognition rates indeed descend rapidly [9]. Then, the recognition rates of BRDM2 are obviously low in the feature spaces of PCA, because the dimensionality of PCA is too small to contain all the primary information. To illustrate it, Figure 7 is plotted. The figure shows the average recognition rate and the recognition rate per target of SRC under different PCA feature spaces. PCA<sup>2</sup> and PCA<sup>4</sup> are described above. PCA<sup>5</sup> and PCA<sup>6</sup> refer to reducing the dimensions of the cropped image to 900 and 1600 by PCA, respectively. From Figure 7, the recognition rate of BRDM2 indeed rises with the increasing dimensions of PCA. Similarly, BTR70 enjoys the same trend, which is more moderate than BRDM2. Contrarily, the recognition rates of other targets tend to descend with the increasing dimensions, and the shape of the average recognition rate in Figure 7 is a response to that. Therefore, it is not wise to improve the recognition accuracy of one object, sacrificing the whole performance. Last but not least, the classifiers both tend to classify other targets into 2S1 incorrectly, which is mainly caused by its structures. The structures, such as the barrel and upper surface, make it similar to tanks (T72, T62) and armored personnel carriers (BMP2, BTR60, BTR70 and BRDM2). Simultaneously, it is no wonder that the classifiers classify 2S1 into other targets.

**Table 4.** The 10-class target confusion matrices of two classifiers under different feature spaces.

SVM		Features	Targets	BMP2	BTR70	T72	BTR60	2S1	BRDM2	D7	T62	ZIL131	ZSU234	SRC											
SVM	a	BMP2	83.6	1.5	8.5	2.0	4.1	0.0	0.0	0.2	0.0	0.0	0.0	a	BMP2	79.4	4.1	10.2	3.6	0.3	0.7	0.5	0.3	0.0	0.9
		BTR70	2.0	93.9	0.0	0.0	4.1	0.0	0.0	0.0	0.0	0.0	0.0		BTR70	0.0	98.5	0.5	1.0	0.0	0.0	0.0	0.0	0.0	0.0
		T72	9.6	0.7	73.0	0.0	5.5	0.0	0.0	10.3	0.9	0.0	0.0		T72	8.1	5.2	77.8	7.6	0.0	1.0	0.0	0.2	0.0	0.2
		BTR60	1.0	1.5	0.0	95.4	0.0	0.5	0.0	0.5	0.0	1.0	1.0		BTR60	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
		2S1	0.4	0.0	1.5	1.1	92.3	0.7	0.0	3.3	0.7	0.0	0.0		2S1	0.4	0.0	1.1	0.7	95.6	2.2	0.0	0.0	0.0	0.0
		BRDM2	1.1	2.2	0.4	4.4	3.3	85.0	0.0	0.0	3.3	0.4	0.4		BRDM2	0.4	2.9	5.5	3.3	1.1	86.5	0.4	0.0	0.0	0.0
		D7	0.0	0.0	0.0	0.0	0.4	0.4	97.4	0.0	0.0	1.8	0.0		D7	0.0	0.0	0.7	0.0	0.0	0.0	99.3	0.0	0.0	0.0
		T62	0.4	0.0	0.7	0.0	3.3	0.0	0.7	79.5	9.5	5.9	0.0		T62	0.7	0.0	0.7	1.1	0.0	0.7	0.4	95.6	0.0	0.7
		ZIL131	0.0	0.0	0.0	0.0	0.4	0.0	0.4	2.2	96.4	0.7	0.0		ZIL131	0.0	0.0	0.0	2.2	0.4	0.0	0.0	0.4	97.1	0.0
		ZSU234	0.0	0.0	0.0	0.0	0.0	0.0	4.4	0.4	1.1	94.2	0.0		ZSU234	0.0	0.0	0.0	0.4	0.4	0.0	0.7	0.0	0.0	98.5
	b	BMP2	71.7	8.2	8.3	4.3	2.6	2.6	0.2	0.9	0.3	1.0	1.0	b	BMP2	46.2	1.7	0.9	0.3	7.3	0.3	16.0	10.6	4.1	12.6
		BTR70	1.5	91.3	0.0	4.6	1.5	1.0	0.0	0.0	0.0	0.0	0.0		BTR70	0.0	89.3	0.0	0.0	2.6	0.0	2.0	3.6	0.5	2.0
		T72	12.2	6.5	49.0	2.1	13.4	2.7	0.7	7.9	5.2	0.3	0.3		T72	0.0	1.5	37.5	0.2	10.1	0.0	14.9	17.0	7.6	11.2
		BTR60	1.0	5.6	1.0	85.1	2.6	2.6	0.0	0.5	0.0	1.5	1.5		BTR60	0.0	0.0	0.0	91.3	2.1	0.0	4.1	0.5	0.0	2.1
		2S1	1.8	8.0	4.4	1.5	74.8	0.4	0.0	4.7	3.3	1.1	1.1		2S1	0.0	0.0	0.0	0.0	89.8	0.0	3.3	3.3	0.7	2.9
		BRDM2	4.7	1.5	1.5	3.6	3.6	73.7	0.0	4.7	4.0	2.6	2.6		BRDM2	0.0	0.0	0.0	0.4	4.0	67.2	7.3	7.3	6.6	7.3
		D7	0.0	0.0	0.4	0.4	0.0	1.5	89.8	1.5	1.1	5.5	5.5		D7	0.0	0.0	0.0	0.0	0.4	0.0	98.5	0.7	0.4	0.0
		T62	1.5	2.9	4.4	0.4	12.8	1.8	1.5	63.7	6.2	4.8	4.8		T62	0.0	0.0	0.4	0.4	3.3	0.0	7.3	83.2	3.3	2.2
		ZIL131	0.4	0.0	0.4	1.1	10.6	0.0	1.5	5.1	79.2	1.8	1.8		ZIL131	0.4	0.0	0.0	0.4	2.6	0.0	7.7	2.6	85.8	0.7
		ZSU234	0.7	0.4	1.1	0.4	0.4	0.7	9.5	6.6	5.1	75.2	75.2		ZSU234	0.0	0.0	0.0	0.0	0.4	0.0	7.3	1.5	0.0	90.9
	c	BMP2	75.6	5.8	9.4	3.1	2.6	2.7	0.0	0.3	0.0	0.5	0.5	c	BMP2	71.6	2.9	1.5	0.5	6.0	0.7	2.7	5.8	1.2	7.2
		BTR70	2.0	86.7	2.6	3.1	5.6	0.0	0.0	0.0	0.0	0.0	0.0		BTR70	0.0	86.2	0.5	1.5	4.6	0.0	1.5	1.0	2.0	2.6
		T72	11.3	4.1	62.0	0.7	11.0	4.1	0.3	4.1	2.2	0.0	0.0		T72	0.3	2.1	50.0	1.0	5.7	0.2	6.9	22.0	4.8	7.0
		BTR60	2.6	5.6	1.0	84.6	1.0	3.6	0.0	1.0	0.0	0.5	0.5		BTR60	0.0	1.0	1.0	95.4	0.0	0.5	1.0	1.0	0.0	0.0
		2S1	3.6	4.7	7.3	1.5	70.4	3.6	0.0	4.7	3.6	0.4	0.4		2S1	0.0	0.0	0.7	0.0	93.8	0.0	0.4	3.3	0.4	1.5
		BRDM2	2.2	4.4	5.8	5.5	8.4	63.9	0.0	2.9	5.5	1.5	1.5		BRDM2	0.0	0.4	0.0	2.9	2.6	61.7	5.1	7.3	11.7	8.4
		D7	0.0	0.0	0.0	0.7	0.4	0.0	96.0	0.4	0.7	1.8	1.8		D7	0.0	0.0	0.0	0.4	0.0	0.0	98.5	0.0	1.1	0.0
		T62	0.7	1.1	1.8	0.4	13.2	7.0	0.7	61.9	8.4	4.8	4.8		T62	0.0	0.0	1.1	0.4	1.8	0.0	3.3	86.4	5.5	1.5
		ZIL131	0.4	0.4	1.5	0.0	1.8	1.8	0.7	4.0	89.1	0.4	0.4		ZIL131	0.0	0.0	0.0	0.0	0.7	0.0	0.7	0.7	97.4	0.4
		ZSU234	2.6	0.0	0.0	2.6	0.0	7.3	6.9	5.1	2.9	72.6	72.6		ZSU234	0.0	0.0	0.0	0.0	0.0	0.0	2.2	2.2	0.4	95.3

\*a, b, c refer to the confusion matrices generated under Down-sample<sup>2</sup>, PCA<sup>4</sup> and PCA<sup>2</sup>, respectively. Down-sample<sup>2</sup> refers to reducing the dimensions of the cropped image to a size of 400 by down-sample; similarly, PCA<sup>4</sup> and PCA<sup>2</sup> stand for sizes of 400 and 100 by PCA, respectively.



**Figure 7.** The recognition rates of SRC under PCA feature spaces; PCA<sup>2</sup> refers to reducing the dimensions of the cropped image to a size of 100 by PCA; similarly, PCA<sup>4</sup>, PCA<sup>5</sup> and PCA<sup>6</sup> stand for sizes of 400, 900 and 1600, respectively.

## 5. Conclusions

In this paper, sparse representation-based classification has been applied to the 10-class MSTAR data set. A preprocessing method of the MSTAR images is introduced to unify the sizes of the images, reduce the complexity and improve the performance of the classifiers. Then, the identity of the test sample is inferred by the framework of SRC. Compared with SVM, SRC also achieves good performance of recognition accuracy, and when the dimensions of the feature space surpass a certain threshold, the performances of the classifiers under conventional and unconventional features converge. Additionally, the recognition rate of each target also has been discussed. The performance is badly influenced by the variation of the target configurations and articulations, the different depression angles in the sensor and the aspect angles of the targets, which is an open problem called over-fitting. Simultaneously, the wrong classification of the specific targets BRDM2 and ZS1 is discussed further. The main reason is that the feature extraction methods cannot describe the targets effectively.

Taking the minor flaws into consideration, on one hand, the future work will focus on improving the robustness of SRC. In other words, a more efficient framework of SRC should be explored. On the other hand, we expect to find more effective features to describe the targets. Although baseline features are studied here, conventional features need to be enriched to describe targets more effectively and so do the unconventional features.

**Acknowledgments:** This work is partially supported by the National Natural Science Foundation of China under Grant 61372163 and 61240058.

**Author Contributions:** Haibo Song performed the simulations and algorithm analysis, and contributed to a main part of manuscript writing. Kefeng Ji contributed in conceiving and the structure of the manuscript. All authors contributed to the writing of the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Thiagarajan, J.; Ramamurthy, K.; Knee, P.P.; Spanias, A.; Berisha, V. Sparse representation for automatic target classification in SAR images. In Proceedings of Communications, Control and Signal Processing (ISCCSP), 2010 4th International Symposium on, Limassol, Cyprus, 3–5 March 2010.
2. Owirka, G.J.; Verbout, S.M.; Novak, L.M. Template-based SAR ATR performance using different image enhancement techniques. *Proc. SPIE* **1999**, *3721*, 302–319.

3. Patnaik, R.; Casasent, D.P. MSTAR object classification and confuser and clutter rejection using minace filters. In Proceedings of Automatic Target Recognition XVI, Orlando, FL, USA, 17 April 2006.
4. Marinelli, M.P.; Kaplan, L.M.; Nasrabadi, N.M. SAR ATR using a modified learning vector quantization algorithm. In Proceedings of the SPIE Algorithms for Synthetic Aperture Radar Imagery VI, Orlando, FL, USA, 5 April 1999.
5. Yuan, C.; Casasent, D. A new SVM for distorted SAR object classification. In Proceedings of the SPIE Conference on Optical Pattern Recognition XVI, Orlando, FL, USA, 28 March 2005.
6. Huang, J.B.; Yang, M.H. Fast sparse representation with prototype. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2010), San Francisco, CA, USA, 13–18 June 2010; pp. 3618–3625.
7. Shin, Y.; Lee, S.; Woo, S.; Lee, H. Performance increase by using an EEG sparse representation based classification method. In Proceedings of the IEEE International Conference on Consumer Electronics, Las Vegas, NV, USA, 2013; pp. 201–203.
8. Huang, X.; Wang, P.; Zhang, B.; Qiao, H. SAR target recognition using block-sparse representation. In Proceedings of the International Conference on Mechatronic Sciences, Electric Engineering and Computer, Shengyang, China, 20–22 December 2013; pp. 1332–1336.
9. Dong, G.; Wang, N.; Kuang, G. Sparse representation of monogenic signal: With application to target recognition in SAR images. *IEEE Signal Process. Lett.* **2014**, *21*, 952–956.
10. Xing, X.; Ji, K.; Zou, H.; Sun, J. SAR vehicle recognition based on sparse representation along with aspect angle. *Sci. World J.* **2014**, *2014*. [[CrossRef](#)] [[PubMed](#)]
11. Pati, Y.C.; Rezaiifar, R.; Krishnaprasad, P.S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In Proceedings of the Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 1–3 November 1993; pp. 40–44.
12. Chen, S.; Billings, S.; Luo, W. Orthogonal least squares methods and their application to non-linear system identification. *Int. J. Control* **1989**, *50*, 1873–1896. [[CrossRef](#)]
13. Candes, E.; Romberg, J.  $l_1$ -magic: Recovery of sparse signals via convex programming. California Institute of Technology. 2005. Available online: [http://thesis-text-naim-mansour.googlecode.com/svn/trunk/Thesis Text/Thesis%20backup/Literatuur/Optimization%20techniques/l1magic.pdf](http://thesis-text-naim-mansour.googlecode.com/svn/trunk/Thesis%20Text/Thesis%20backup/Literatuur/Optimization%20techniques/l1magic.pdf) (accessed on 10 December 2014).
14. Wright, J.; Yang, A.Y.; Ganesh, A.; Sastry, S.S.; Ma, Y. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 210–227. [[CrossRef](#)] [[PubMed](#)]
15. Patnaik, R.; Casasent, D.P. SAR classification and confuser and clutter rejection tests on MSTAR 10-class data using minace filters. In Proceedings of the Optical Pattern Recognition XVIII, (SPIE 6574, 657402), Orlando, FL, USA, 9 April 2007.
16. Yuan, C.; Casasent, D. MSTAR 10-class classification and confuser and clutter rejection using SVRDM. In Proceedings of the Optical Pattern Recognition XVII, (SPIE 6245, 624501), Orlando, FL, USA, 17 April 2006.
17. Moving and Stationary Target Acquisition and Recognition (MSTAR) Public Release Data. Available online: <https://www.sdms.af.mil/datasets/matar/> (accessed on 24 August 2015).
18. Ross, T.; Worrell, S.; Velten, V.; Mossing, J.; Bryant, M. Standard SAR ATR evaluation experiments using the MSTAR public release data set. In Proceedings of the SPIE on Algorithms for SAR, Orlando, FL, USA, 13 April 1998; pp. 566–573.
19. El-Darymli, K.; McGuire, P.; Gill, E.; Power, D.; Moloney, C. Characterization and statistical modeling of phase in single-channel synthetic aperture radar imagery. *IEEE Trans. Aerospace Electron. Syst.* **2015**, *51*, 2071–2092. [[CrossRef](#)]
20. Chang, C.; Lin, C. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 389–396. [[CrossRef](#)]
21. Mathworks. Regionprops. Available online: <http://tinyurl.com/k58dlqf> (accessed on 5 December 2015).

