Neurocomputing 113 (2013) 97–104

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Decision fusion of sparse representation and support vector machine for SAR image target recognition

Haicang Liu, Shutao Li*

College of Electrical and Information Engineering, Hunan University, Changsha, Hunan 410082, China

ARTICLE INFO

Article history:
Received 7 April 2012
Received in revised form
7 December 2012
Accepted 3 January 2013
Communicated by: Shiguang Shan
Available online 5 March 2013

Keywords:
Sparse representation
SVM
Decision fusion
SAR
Target recognition

ABSTRACT

We propose a decision fusion method of Sparse Representation (SR) and Support Vector Machine (SVM) for Synthetic Aperture Radar (SAR) image target recognition in this paper. First, a fast SR classifier (FSR-C) with Matching Pursuit (MP) solution is proposed. In the FSR-C, the dictionary is composed of training images. Just one nonzero element in SR coefficient of the testing image is found out based on MP, and the testing image is classified through the location of the nonzero element. To further improve the recognition accuracy, the SVM classifier (SVM-C) is selected. In SVM-C, PCA feature is extracted, and for seeking the linear separating hyperplane, the RBF kernel function is used in mapping the training vectors into high dimensional space. The results of the FSR-C and the SVM-C are fused obeying Bayesian rule to make the decision. The Moving and Stationary Target Acquisition and Recognition (MSTAR) SAR image database is used to test the performance of the proposed method. The experimental results show that the FSR-C can predict testing SAR images with considerable recognition accuracy and high real-time ability, and the decision fusion recognition method can improve the recognition accuracy and still be fast.

1. Introduction

Taking advantage of acquiring image in inclement weather or during night as well as day, Synthetic Aperture Radar (SAR) image is widely applied in civilian and military. SAR image target recognition is the most important step in SAR image interpretation and analysis [1].

Target recognition generally consists of three processes, i.e., preprocessing, feature extraction and classification [2]. The image feature is extracted, and it directly affects the probability of correct classification (PCC) [3]. There are many methods for SAR image target feature extraction, such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), Linear Discriminant Analysis (LDA) and Discrete Wavelet Transform (DWT) [4–6]. These methods have different characteristics, and fit different tasks. A variety of target classification algorithms have been used in SAR image, such as Support Vector Machine (SVM) [7,8], Adaboost [9], Neural Network (NN) [10], Gaussian Mixture Models (GMM) [11,12] and so on.

Wu Tao et al. [8] studied SAR image target recognition with SVM classifier. They divided the training samples into several groups according to the SAR image aspect angles and train SVM classifier for each group. In the test, they first measured the aspect angle through segmentation, and then extracted PCA feature. At last, the result is predicted in the corresponding SVM classifier. The method got high recognition accuracy. However, in this method, the PCC would decrease with the increasing of the aspect angle interval, and target segmentation is very time-consuming and hard to achieve high accuracy.

Sparse Representation (SR) is a new method for target recognition [13,14]. Wright et al. [13] proposed the SR method for face recognition, and Thiagarajan et al. [14] modified the approach for classifying targets in SAR image. In Ref. [13] or [14], the dictionary is composed of the training vectors that are normalized, and the sparse representation of the testing data is computed with the dictionary. The sparse representation is a locally linear approximation with respect to the corresponding class. The testing image is reconstructed from the sparse representation coefficient and the dictionary, and is used to classify. The Orthogonal Matching Pursuit (OMP) is used in Ref. [14] instead of l_1-norm in Ref. [13] to solve the sparse representation of the testing image. The SR method with OMP solution (SR-OMP) proposed by Thiagarajan et al. is still time-consuming because of the iteration in OMP and the reconstruction process of the testing image.

To improve the target recognition performance, many fusion methods are proposed [15–18]. Fig. 1 displays three types of decision fusion methods that are often used for SAR image recognition. Fig. 1 (a) is a multi-angle decision fusion method

* Corresponding author. Tel.: +86 73 188 822 866.
E-mail addresses: shutao_li@yahoo.com.cn, shutao_li@hnu.edu.cn (S. Li).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.01.033

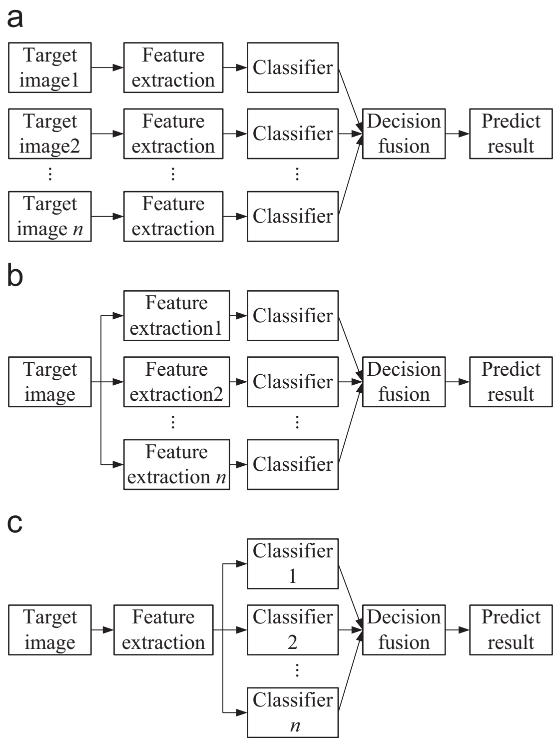


Fig. 1. Three types of decision fusion methods for SAR image. (a) multi-angle decision fusion, (b) multi-feature decision fusion and (c) multi-classifier decision fusion.

which needs several SAR images of one target with different aspect angles to predict correctly. Fig. 1 (b) is a multi-feature decision fusion method, and it extracts various features of the target image, such as PCA, ICA, and LDA. Fig. 1 (c) is a multi-classifier decision fusion method which improves the recognition accuracy by using multi-classifier instead of multi-angle testing images or multi-feature. The multi-angle decision fusion method will bring higher PCC than multi-feature and multi-classifier decision methods, but it is implemented with difficulty in real-time system, because there is no time to acquire more than one testing SAR image.

In many cases, the real-time ability is demanded for the recognition system, especially in the application of military [19–21]. A real-time recognition system aims to effectively and rapidly identify the target. Compared to the single-strategy recognition methods, the fusion methods can improve the recognition accuracy. To meet the requirement of real-time ability, the methods taken part in the fusion must save computational cost and be fast.

Because SAR image is formed based on point scatter distribution on the target surface which is different from the optical image, it is sensitive to the aspect angle, hence, the images of the same target taken at different aspect angles show great differences [6]. In SR classifier, training images are set as atoms, and they composed a dictionary. The usage of the dictionary benefits the classification, especially for the recognition system with large scale of training set. The dictionary makes the training images ordered, and it is easy to increase the training set which can be carry out through increasing atoms in the dictionary. SR classifier with the dictionary can predict testing image fast. SVM is a classical method for target recognition, and it can achieve high recognition accuracy with short predicting time. In SR classifier, training images are transformed to vectors and set to be atoms as the image features, and the testing image is classified through matching every atoms. In SVM classifier (SVM-C), the PCA features of training images are extracted and used to find a separating hyperplane, and the testing image is classified by the distance to the hyperplane. SR and SVM are two fast and effective recognition methods, and they predict targets from different principles. Considering the recognition accuracy and real-time ability at the same time, the decision fusion method of SR and SVM (DFSS) is proposed in this paper. In Fig. 1, we list three types of decision fusion methods that are often used for SAR image recognition. The DFSS method is the synthesis of multi-feature decision fusion method and multi-classifier decision fusion method.

Weak classifiers are usually adopted in online training fusion frame to reduce the computation cost [22,23]. In this paper, our aim is to obtain high recognition PCC in short time with no increasing input information. We select two well-performed strong classifiers of SR and SVM in fusion frame to further improve the recognition PCC. Training stage in SVM and dictionary constructing stage in SR are time consuming, however, the two stages are offline implemented, and they do not effect the predicting time.

First, we propose the fast SR classifier (FSR-C) with Matching Pursuit (MP) solution. In FSR-C, to save predicting time, the SAR image pixels are directly used as the features. The dictionary is composed of the training images vectors, and only one nonzero element in the SR of the testing image is computed via MP. The testing image is classified through the location of nonzero element in the SR coefficient. Via the PCA feature of the image target, another classify result through SVM-C can be obtained. Finally, we compute the confidence of every class of targets through the prediction results of FSR-C and SVM-C based on Bayesian rule, and through comparison, the class with the maximal confidence is selected as the decision result.

We propose FSR-C in Section 2, and introduce the SVM-C in Section 3. The proposed decision fusion method of SR and SVM is present in Section 4. Performance comparison and analysis of several recognition methods for SAR image are provided in Section 5. Finally, a concluding remark is given in Section 6.

2. Fast target recognition based on SR

In SR, a signal y can be represented as [22]

$$y = \mathrm{D}\alpha + \varepsilon \tag{1}$$

where D is the dictionary, α is the coefficient which is sparse, and ε is the residual vector. In this Section, we propose a fast target recognition algorithm based on SR.

2.1. Related work

The dictionary D is composed of atoms d_m ($1 \le m \le M$ where M is the total number of atoms). Assuming there are J classes of targets, all the training images are transformed into vectors, and ranked in classes. All the training images are included in the dictionary D. Fig. 2 displays the dictionary composition.

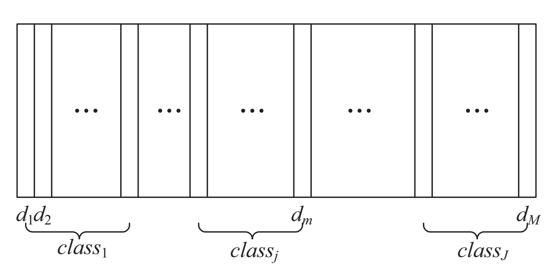


Fig. 2. Dictionary composition.

The Eq. (1) is a linear system, and can be represented as

$$y = d_1\alpha_1 + d_2\alpha_2 \cdots + d_M\alpha_M + \varepsilon \quad (2)$$

Setting the testing image as y, and the testing image is decomposed into a linear expansion of atoms (training images) selected from the dictionary. The coefficient α is sparse. Wright et al. used the l_1-norm to solve the coefficient α in Ref. [13] for face recognition, that is

$$\alpha = \text{argmin}||\alpha||_1 \text{subject to} ||y - \text{D}\alpha||_2 \le \varepsilon \quad (3)$$

Thiagarajan et al. used the OMP method to solve α, and fixed the number of nonzero elements as 5 or 9 in Ref. [14] for SAR image target recognition. Keeping the elements correlative with the class j and setting the others as zero in the coefficient α, a reconstructed coefficient $\hat{\alpha}(j)(1 \le j \le J)$ is obtained. The reconstructed testing image $\hat{y}(j)$ is got through $\hat{y}(j) = \text{D}\hat{\alpha}(j)(1 \le j \le J)$. In Refs. [13] and [14], they all classify the testing image through the error of y and $\hat{y}(j)$. The testing image belongs to the class in which the error is the smallest.

2.2. Fast SR classifier

We solved the problem (1) by the method of MP [24], and stop it after the first iteration. The testing image can be approximately represented with one atom, that is

$$y = \langle y, d_{m^*} \rangle d_{m^*} + \varepsilon^* \; 1 \le m^* \le M \quad (4)$$

To minimize the residual vector ε^*, $d_m \in \text{D}$ must be chosen to ensure $|\langle y, d_m \rangle|(1 \le m \le M)$ is the maximum. We find out the atom d_{m^*} in the training set witch is the most matching with the testing image.

In the coefficient $\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_M \end{bmatrix}^T$, there is one element α_{m^*} is nonzero, and it is

$$\alpha_{m^*} = \langle y, d_{m^*} \rangle (1 \le m^* \le M) \quad (5)$$

The testing image is predicted to the class which the atom d_{m^*} is in. Algorithm 1 below summarizes our proposed FSR-C.

Algorithm 1. Fast classification based on SR

1: Input: The training dictionary D
   The testing SAR image y
2: Transform the testing image y to a vector;
   Solve $y = \text{D}\alpha + \varepsilon$ based on MP to find out the most matching atom d_{m^*} and α_{m^*};
3: Classify the testing image via the location of α_{m^*} in the SR coefficient α;
4: Output: Result of the classification

In solving $y = \text{D}\alpha + \varepsilon$, we use MP method and just iterate one time to seek one most matching atom. This save more computational cost than solving l_1-norm in Ref. [13], and is one fifth or ninth of that in Ref. [14] which need five or nine iteration. In classification, we do not need to reconstruct the testing image and compute the error. So our algorithm is much faster than l_1-norm and SR-OMP.

3. SAR target recognition based on SVM

SVM is a popular and useful technique for data classification, and in the SVM-C, the PCA feature of image target is often applied [25,26]. The PCA feature is valid to target recognition and could be extracted fast [18].

Before using the SVM-C, we first train it. The PCA features of some samples $x_k(x_k \in \mathbb{R}^w, k = 1,2\cdots K)$ are set to be training vectors, and they are mapped into a higher dimensional space through function $\phi(x)$. Then SVM training is to find a linear separating hyperplane, as

$$g(x) = wx + b = 0 \; (x \in \mathbb{R}^w) \quad (6)$$

where w is a weight vector which has the same dimension as the feature space and b is the bias. The hyperplane should have the maximal margin between the separating hyperplane and training samples in this higher dimensional space.

The training set of instance-label pairs are $(x_k, y_k)(k = 1,2,\cdots K)$, where $y_k \in \{1,-1\}$. The SVM-C requires the solution of the following optimization problem:

$$\min_{w,b,\eta} \frac{1}{2} w^T w + C \sum_{k=1}^{K} \eta_k \text{ subject to } y_k(w^T\phi(x_k) + b) \ge 1 - \eta_k \; \eta_k \ge 0 \quad (7)$$

where C is the penalty parameter of the error term.

RBF kernel function is used in mapping training vectors into a higher dimensional space, as

$$K(x_k, x_g) = \phi(x_k)^T \phi(x_g) = \exp(-\gamma||x_k - x_g||^2) \; \gamma > 0 \quad (8)$$

where γ is the kernel parameter.

The methods of cross-validation and grid-search are used to identify the two parameters C and γ [27]. The training set is divided into v subsets with the same size. One subset is tested sequentially using the classifier trained on the remaining $v-1$ subsets. Thus, each instance of the whole training set is predicted once so the accuracy is the percentage of data which are correctly classified. In grid-search, pairs of (C,γ) are tried and the one with the best cross-validation accuracy is picked. To save consuming time, a coarse grid is used first. After a better region is found out on the grid, a finer grid search on that region can be conducted. The best (C,γ) is found out, and the classifier is trained again on the whole training set to generate the final SVM-C.

The FCA feature of the testing image is extracted, and it can help the trained SVM-C predict accurately. What we need to solve is a multi-class problem, but the SVM is suite for binary problem. Here one-against-one scheme and voting rule for decision making are used in SVM prediction for SAR image [8].

4. Decision fusion strategy based on Bayesian

Recognition accuracy and real-time ability are both important for a target recognition system. We make decision fusion strategy of FSR-C and SVM-C based on Bayesian rule, and it can improve the recognition accuracy. The decision fusion scheme is illustrated in Fig. 3.

Two prediction results of r_1 and r_2 for a testing image are respectively obtained from FSR-C and SVM-C. We assume that recognition accuracy of the two classifiers are $P_{\text{FSR-C}}$ and $P_{\text{SVM-C}}$. There are J classes of target, and it is possible that the testing image belongs to the random class. The confidence of that the target which is predicted to the class T_q by the FSR-C method belongs to the class $T_n(1 \le n \le J, 1 \le q \le J)$ can be computed through the Bayesian rule, as follows:

$$P_{FSR-C}(T_n|T_q) = \frac{P(T_n)P_{FSR-C}(T_q|T_n)}{\sum_{j=1}^{J} P_{FSR-C}(T_q|T_j)P(T_j)} (1 \le n \le J, 1 \le q \le J) \quad (9)$$

where $P(T_j)$ is the possibility of that the testing image belongs to the class j, and $P_{FSR-C}(T_q|T_n)$ is the possibility of that the testing image in class n is predicted to the class q by the method of FSR-C.

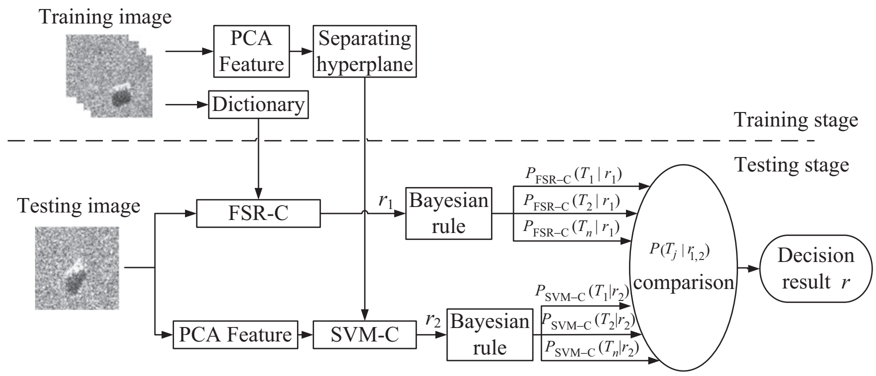


Fig. 3. Decision fusion scheme of FSR-C and SVM-C. Training stage of SVM and dictionary constructing stage of SR are offline implemented.

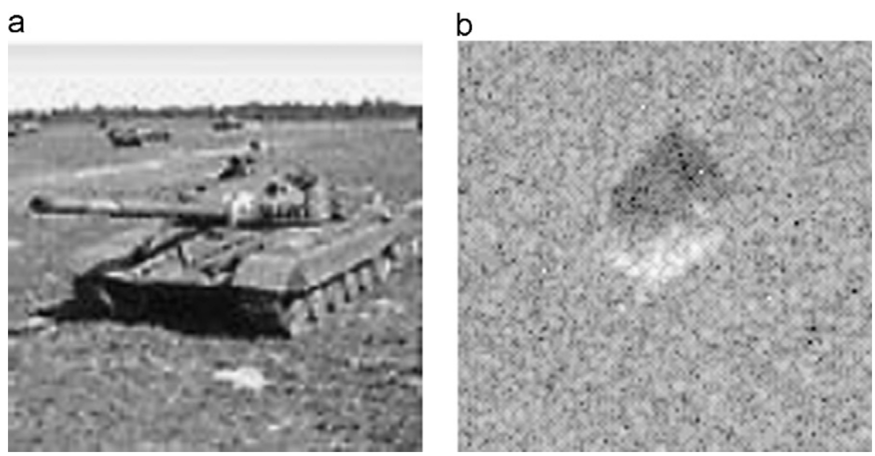


Fig. 4. T72 target pictures from optical camera and SAR. (a) Optical image and (b) SAR image.

The confidence $P_{SVM\text{-}C}(T_n|T_q)$of that by the SVM-C method also can be obtained.

The decision fusion is made based on the Bayesian rule. The confidence of each class for the testing image may belonging to are obtained respectively through the FSR-C method and the SVM method, and the confidence of the same class obtained respectively through the two methods are add together. The class with the maximal confidence is selected as the decision result. The decision fusion based on Bayesian rule can be formulated as follows:

$$r = T_{j^*}, j^* = \max_{j^* = 1,2\cdots J}(P(T_{j^*}|T_{r_{1,2}})) \tag{10}$$

where

$P(T_j|T_{r_{1,2}}) = P_{FSR\text{-}C}(T_j|T_{r_1}) + P_{SVM\text{-}C}(T_j|T_{r_2})$.

In the FSR-C, the training image and the testing image are directly transformed into vectors without feature extraction, and in the SVM-C, the PCA feature of the image is extracted. Our decision fusion method is the synthesis of the multi-feature and multi-classifier methods. Both of the FSR-C and the SVM-C can be implemented rapidly. So our decision fusion strategy of FSR-C and SVM-C based on Bayesian rule can predict the testing images with high recognition accuracy and save predicting time.

5. Experiments

In this section, Moving and Stationary Target Acquisition and Recognition (MSTAR) SAR image set is used in the experiments. First the parameters of the SVM-C are selected, and then the PCC of FSR-C, SVM-C and DFSS methods is tested. Finally, we compare our method with some others, and the comparison result and analysis is given.

5.1. Data preparation

The MSTAR data set was collected in September of 1995 at the Redstone Arsenal, Huntsville, AL by the Sandia National Laboratory (SNL) SAR sensor platform. SNL used an X-band SAR sensor in one foot resolution spotlight mode. The data set includes ten classes of military target: BMP2 (tank), BTR70 (armored car), T72 (tank), 2S1 (cannon), BTR_60 (armored car), BRDM_2 (truck), D7 (bulldozer), T62 (tank), ZIL_131 (truck) and ZSU_23/4 (cannon). SAR image and optical image are formed in different ways, and SAR image is sensitive to the target aspect angle. Fig. 4 is the T72 target images from optical camera and SAR, and Fig. 5 is SAR images with different aspect angles of BMP2, BTR70 and T72.

There are many different aspect versions of each target at each depression angle. We use the images at 17° depress angle as training set, and the images at 15° depress angle as testing set. Table 1 lists the classes and sample numbers of the training and testing set in the database. The size of image is 128×128. We cut the chip to 80×80 to delete the edge of the image, and this can improve the recognition accuracy and reduce much computational cost. For further reduce the predicting time in FSR-C, we downsample the chip to the size of 40×40.

5.2. PCC test for DFSS

First, we use three classes of target, BMP2, BTR70 and T72 in the MSTAR database for experiments. In the testing set, BMP2 contains three variants (Sn_9563, Sn_9566 and Sn_c21), and T72

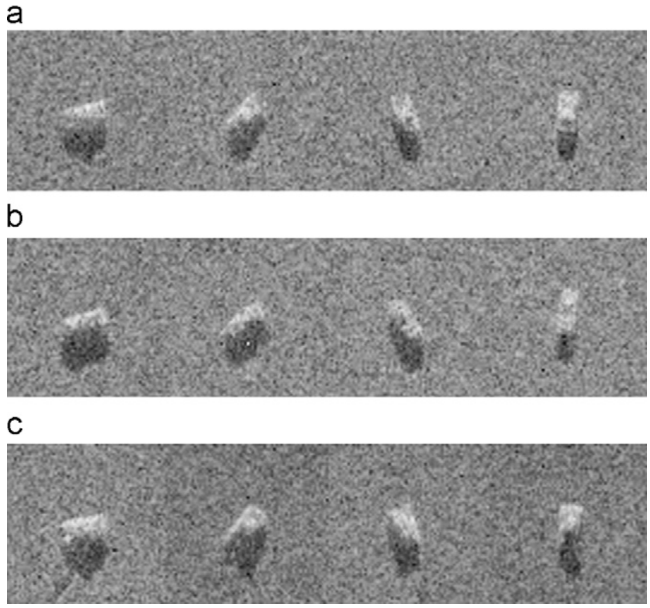


Fig. 5. SAR images with different aspect angles in classes of (a) BMP2, (b) BTR70 and (c) T72.

Table 1
Classes and sample number of training and testing set.

class	Training Set	Testing Set
	233 (Sn_9563)	195 (Sn_9563)
BMP2	232 (Sn_9566)	196 (Sn_9566)
	233 (Sn_c21)	196 (Sn_c21)
BTR70	233 (Sn_c71)	196 (Sn_c71)
	232 (Sn_132)	196 (Sn_132)
T72	231 (Sn_812)	195 (Sn_812)
	228 (Sn_s7)	191 (Sn_s7)
2S1	299	274
BTR_60	256	195
BRDM_2	298	274
D7	299	274
T62	299	273
ZIL_131	299	274
ZSU_23/4	299	274

contains three variants (Sn_132, Sn_812 and Sn_s7). In most cases, we can not obtain the variant of the training image which is same with the testing image we need to predict. We select three variants of Sn_9563, Sn_c71, and Sn_132 from the three classes of BMP2, BTR70 and T72 as the training set in our experiments. There are 698 images in the training set, and 1365 images in the testing set.

In our experiments, the PCC is used to measure the recognition performance.

$$PCC = \frac{n_c}{n_t} \times 100\% \tag{11}$$

where n_c is the sample number of correct classification and n_t is the total sample number.

The PCA features of the SAR image is extracted as the input of the SVM-C. In PCA features extraction, the variance-covariance matrix of the training set is decomposed with the Singular Value Decomposition (SVD) method, and the number of PCA features corresponds to the number of biggest eigenvalues that are selected in the decomposition. We select the number of PCA features to ensure that the SVM-C have the highest recognition accuracy. Fig. 6 displays the relationship of PCC and the number of PCA features. If the number of PCA features is more than 60, the PCC would not increase as the increasing of the number of PCA features. We fix the number of PCA features as 60 in the later experiment.

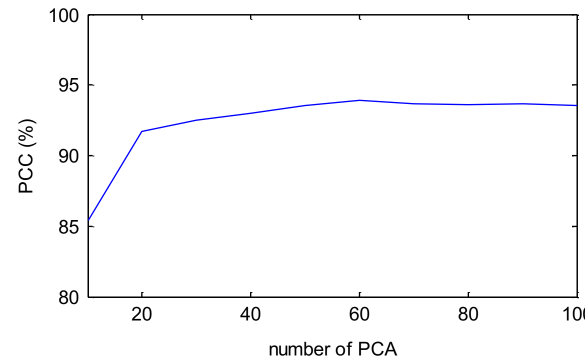


Fig. 6. Relationship of PCC and the number of PCA features.

Table 2
Confusion matrix for the testing set in three classes by FSR-C and SVM-C.

Method	Testing image	Prediction result		
		BMP2	BTR70	T72
	BMP2	548	17	22
FSR-C	BTR70	0	196	0
	T72	80	11	491
	BMP2	529	15	43
SVM	BTR70	0	195	1
	T72	15	9	558

In the training of SVM-C, there are two parameters C and γ that affect the recognition performance and need to be selected with the methods of cross-validation and grid-search. Because the one-against-one scheme and voting rule are used in SVM-C prediction, three groups of (C,γ) need to be identified for the three classes, and they are $(C,\gamma)_{BMP2-BTR70}$, $(C,\gamma)_{BMP2-T72}$ *and* $(C,\gamma)_{BTR70-T72}$ corresponding to three SVM-Cs to be trained. Sometimes obtaining the best recognition accuracy as high as 100% in the training progress may bring the overfitting problem. The cross-validation procedure can prevent the overfitting problem well [25], and we select the three best groups of (C,γ) as follows

$$\begin{aligned} (PCC,C,\gamma)_{BMP2-BTR70} &= (98.92\%,2,0.0313) \\ (PCC,C,\gamma)_{BMP2-T72} &= (98.27\%,0.3536,0.125) \\ (PCC,C,\gamma)_{BTR70-T72} &= (98.70\%,0.5,0.3536) \end{aligned} \tag{12}$$

In the Eq. (12), the PCC is got for the training set. Through training, we obtain three hyperplanes that have the maximal margin between the separating hyperplane and training samples.

We test the recognition performance of the FSR-C and SVM-C with the testing set, and Table 2 lists the confusion matrix for each class by FSR-C and SVM-C. Based on Table 2, the confidence of the three classes of targets for the two methods is computed through Eq. (9), and they are listed in Table 3. The $P(T_n)$ can be obtained via the concrete application conditions. In this paper, we set the $P(T_n)$ for every classes via the testing set.

$$\begin{cases} P(T_{BMP2}) = \frac{587}{1365} = 0.43 \\ P(T_{BTR70}) = \frac{196}{1365} = 0.14 \\ P(T_{T72}) = \frac{582}{1365} = 0.43 \end{cases} \tag{13}$$

Table 3 is consistent with Table 2. For example, the confidence of the class of BTR70 is 0 if the prediction result of the testing image is the class of BMP2 by the FSR-C method. That is to say, it is impossible that the testing image belongs to the class of BTR70.

Table 3
Confidence of three classes of target for FSR-C and SVM-C (%).

Method	Prediction result	Confidence of each class		
		BMP2	BTR70	T72
FSR-C	BMP2	87.26	0	12.74
	BTR70	7.59	87.50	4.91
	T72	4.29	0	95.71
SVM-C	BMP2	96.53	0	3.47
	BTR70	6.85	89.04	4.11
	T72	7.14	0.17	92.69

Table 4
Recognition accuracy of the proposed method for three classes of targets (%).

		Method		
		FSR-C	SVM-C	DFSS
Testing classes	BMP2	93.35	90.11	91.65
	BTR70	100	99.48	99.48
	T72	84.36	95.87	96.04
Over-all PCC		90.47	93.92	94.65
Average PCC		92.57	95.15	95.72
std		7.84	4.72	3.92

This is because that the FSR-C method can predict the testing image in the class of BTR70 correctly and will not predict it to the other class, and this is showed in Table 2. If the testing image is predicted to the class of BMP2, the confidence of class of BMP2 is 87.26% from the FSR-C method, and this is lower than 96.53% which is got from the SVM-C method, though FSR-C can get higher PCC than SVM-C for the class of BMP2. The FSR-C classes more other targets to the class of BMP2 than SVM-C, and so the confidence of the class of BMP2 from FSR-C is decreased.

The FSR-C and SVM-C are fused based on the Bayesian rule as described in the Section 4. We test the proposed fusion method, and the result is listed in Table 4. The recognition performance is measured through over-all PCC, average PCC and standard deviation (std). Over-all PCC is got through the number of targets predicted correctly divided by the number of targets for all the classes. Average PCC and std are the average and standard deviation of every class PCC. The FSR-C and SVM-C methods can predict nearly every testing image for the class of BTR70, because there is just one variant in the testing set which is included in the training set. However, for the classes of BMP2 and T72, the PCC of the methods are not as high as that for the class of BTR70. The average PCC is obtained as 92.57% from FSR-C, and 95.15% from SVM-C. In Table 4, we find that the FSR-C can predict the testing images in the class of BMP2 with high PCC, but with low PCC for the class of T72. In the contrary, from the SVM-C method we can get high PCC for the class of T72, but low PCC for the class of BMP2.

DFSS method has shown better performance in terms of the over-all PCC, average PCC and std of the classification accuracy than FSR-C and SVM-C. The average PCC of DFSS is improved to as high as 95.72%, and this is higher than that of the FSR-C method by 3.15%, and that of the SVM-C method by 0.57%. The std of DFSS method is lower than FSR-C and SVM-C, and this indicates that DFSS method is more stable. The DFSS method does not achieve the highest PCC from FSR-C or SVM-C for the class of BMP2. This is because the FSR-C method gives its prediction result lower confidence for the class of BMP2, though it can predict this class of target with higher PCC than the SVM-C method. Which class the testing image belongs to is unknown before the prediction, and so the decision of which classifier to select can not be made. The prediction results are obtained respectively through the two methods, the confidence of the prediction results are obtained through the Bayesian rule, and the result with the maximal confidence is selected as the decision. Though the DFSS method does not achieve the highest PCC for every class, it improves the average PCC much, and this is the advantage of the Bayesian rule.

To further test the recognition accuracy of our method, ten classes of targets in the MSTAR database are used in experiments. More training samples are added to the dictionary as atoms in FSR-C method. In SVM-C method, $\frac{10\times(10-1)}{2}=45$ SVM classifiers are trained because of one-against-one scheme and voting rule adopted for multi-class problem. One testing image will receive $J-1=9$ votes at most. To further improve the PCC, We assign weight to the SVM-C result before computing the confidence, and the weight value is $w=num_{vote}/(J-1)$. Table 5 lists the test result. The FSR-C method gets the average PCC as high as 95.16%. DFSS method achieves the highest performance, and it improves the average PCC to 97.33% from 96.45% and reduces the std to 2.65% from 3.55%.

5.3. Comparison with other methods

For the three classes of targets, BMP2, BTR70 and T72, we test the prediction accuracy and time of FSR-C and DFSS, and compare them with SR-l_1 method [13], SR-OMP method [14], EnSR-C method and EnSVM-C method [28,29]. SR-l_1 is the SR method with l_1-norm solution, and it is proposed for face recognition. We use this method for SAR image target recognition. The 40×40 image is so large for SR-l_1 that the method takes much time to predict one target, so we downsample the image to 6×6 in SR-l_1 method to reduce the computational cost. SR-OMP is the SR method with OMP solution. In SR-OMP, the number of nonzero elements in the sparse represent of the testing image is fixed as 5. In SVM-C and EnSVM-C, the PCA feature will be extracted, and so the 80×80 image is used. In SR-OMP and FSR-C, we use the 40×40 image. For comparison, we design the EnSR-C method and the EnSVM-C method. The EnSR-C method is the decision fusion

Table 5
Recognition accuracy of the proposed method for ten classes of targets (%).

		Method		
		FSR-C	SVM-C	DFSS
Testing classes	BMP2	88.41	88.92	94.89
	BTR70	99.48	98.97	91.31
	T72	77.66	97.07	97.08
	2S1	97.44	91.60	99.48
	BTR_60	99.48	96.41	98.46
	BRDM_2	91.60	97.08	99.63
	D7	99.63	99.27	96.33
	T62	98.60	96.33	97.25
	ZIL_131	99.63	99.27	99.27
	ZSU_23/4	99.63	99.63	99.63
Over-all PCC		92.60	95.71	96.62
Average PCC		95.16	96.45	97.33
Std		7.29	3.55	2.65

Table 6
Comparison of prediction accuracy and time for three classes of targets.

Method	SR-l_1	SR-OMP	FSR-C	SVM-C	EnSR-C	EnSVM-C	DFSS
Over-all PCC (%)	78.09	91.43	90.47	93.92	92.53	93.19	94.65
Average PCC (%)	82.07	93.20	92.57	95.15	94.06	94.24	95.72
Std	16.57	5.46	7.84	4.72	4.73	3.93	3.92
Time (ms)	269.2	29.3	3.9	6.5	32.1	12.9	11.6

Table 7
Comparison of prediction accuracy and time for ten classes of targets.

Method	SR-l_1	SR-OMP	FSR-C	SVM-C	EnSR-C	EnSVM-C	DFSS
Over-all PCC (%)	87.26	94.59	92.60	95.71	94.84	95.56	96.62
Average PCC (%)	90.76	96.43	95.16	96.45	96.62	96.23	97.33
Std	12.34	4.90	7.29	3.55	4.75	3.74	2.65
Time (ms)	1294.1	194.8	13.5	59.6	202.0	132.9	75.8
Increased time(ms)	1024.9	165.6	9.6	53.1	169.9	120.0	63.8

of SR-OMP and FSR-C based on Bayesian rule, and the EnSVM-C is the decision fusion of SVM-C with polynomial kernel and SVM-C with Gaussian kernel based on Bayesian rule.

The experiments are performed on the personal computer with a 2.6 GHz AMD CPU and 2 G bytes of memory. This computer runs on Windows XP with MATLAB 7.10 installed. Table 6 lists the recognition performance of the methods. The time is average of the testing images, and it is for predicting one vector which is transformed from the testing image, not including the processes of composing the dictionary for SR or training for SVM.

From Table 6, it can be seen that SR-l_1 method takes the most time and obtains the lowest PCC, and so it is not suitable for SAR image recognition. The FSR-C method take the shortest time as 3.9 ms which is one seventh time of SR-OMP for predicting one SAR testing image, and its PCC is just little lower than the SR-OMP method. The SVM-C method gets the average PCC as 95.15% with 6.5 ms for predicting one SAR testing image. Both FSR-C and SVM-C can be implemented fast.

FSR-C and SVM-C are fused based on the simple Bayesian rule, and the confidence of the classes of target for the two methods is computed before prediction. So the predicting time of DFSS is almost the sum of FSR-C and SVM-C. The DFSS method is much faster than SR-OMP, and it can predict with higher PCC. EnSVM-C method and EnSR-C method achieve higher PCC than FSR-C method, but lower PCC than DFSS method, and they take much longer time than FSR-C and DFSS methods. SVM-C with polynomial kernel and SVM-C with Gaussian kernel predict targets using the same image feature. SVM-C with polynomial kernel is not suitable for high-dimensional input data, so EnSVM-C method does not help to improve the recognition PCC.

For 10 classes of targets, the recognition performance of our method is compared with that of other methods, and Table 7 lists the test result. Because of the increasing of classes of targets, the prediction time of all the methods increase. FSR-C method is the fastest with average prediction time of 13.5 ms, and its increased time is much lower than that of other methods. The increased time of DFSS is nearly the sum of FSR-C and SVM-C. The DFSS method is still much faster than EnSVM-C method and EnSR-C method, and it achieves the highest PCC.

In single methods that take part in the comparison, FSR-C is much faster than SR-l_1 method, SR-OMP method and SVM-C method, and its PCC is just a little lower than SR-OMP method and SVM-C method. In fusion methods, DFSS method takes advantage of multi-feature decision fusion and multi-classifier decision fusion simultaneously. DFSS obtains the highest PCC as the most stable method and it is still much faster than EnSVM-C method and EnSR-C method, even than the single method of SR-l_1 and SR-OMP. With the increasing of target classes, the increased predicting time of FSR-C method and DFSS method are much lower than that of other methods.

6. Conclusions

The decision fusion method of SR and SVM based on Bayesian rule is proposed in this paper, and the MSTAR SAR radar image database is used to test the recognition performance. We first proposed the FSR-C method with MP solution. The FSR-C can predict SAR images with high PCC and high real-time ability. PCA feature is used in the SVM-C, and the number of PCA features is selected to make the SVM have the highest PCC. FSR-C and SVM-C are two strong classifiers. Both of the two methods achieve high recognition performance, but they classify testing images through different principles. We fuse the two strong classifiers based on Bayesian rule to obtain higher recognition accuracy. The FSR-C is faster than SVM-C, but its recognition accuracy is lower than that of SVM-C. The three methods of FSR-C, SVM-C, and DFSS can be used separately, and they are selected based on the demand of recognition system to predicting time and accuracy.

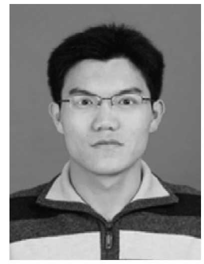
Acknowledgment

This paper is supported by the National Natural Science Foundation of China (No. 61172161) and the Fundamental Research Funds for the Central Universities, Hunan University.

References

[1] A.K. Mishra, M. Bernard, Automatic target recognition using multipolar bistatic synthetic aperture radar images, IEEE Trans. Aerosp. Electron. Syst. 46 (4) (2010) 1906–1920.
[2] N. Sandirasegaram, R. Englist, Comparative analysis of feature extraction (2D FFT and wavelet) and classification (L_p metric distances, MLP NN, and HNeT) algorithms for SAR imagery, Proc. SPIE 5808 (2005) 314–325.
[3] M. Cheng, B. Fang, C.M. Pun, Y.Y. Tang, Kernel-view based discriminant approach for embedded feature extraction in high-dimensional space, Neurocomputing 74 (9) (2011) 1478–1484.
[4] K. Delac, M. Grgic, S. Grgic, Independent comparative study of PCA, ICA, and LDA on the FERET data set, Int. J. Imaging Syst. Technol. 15 (5) (2006) 252–260.
[5] S.K. Goumas, M.E. Zervakis, G.S. Stavrakakis, Classification of washing machines vibration signals using discrete wavelet analysis for feature extraction, IEEE Trans. Instrum Meas. 51 (3) (2002) 497–508.
[6] M. Ceccarelli, A. Petrosino, Multi-feature adaptive classifiers for SAR image segmentation, Neurocomputing 14 (4) (1997) 345–363.
[7] Q. Zhao, J.C. Principe, Support vector machines for SAR automatic target recognition, IEEE Trans. Aerosp. Electron. Syst. 37 (2) (2001) 643–654.
[8] T. Wu, X. Chen, X.W. Ruang, L. Niu, Study on SAR target recognition based on support vector machine, In: Proceedinga of the 2nd Asian-Pacific Confrence on Synthetic Aperture Radar, Xi'an, China, October 2009, pp. 856–859.
[9] Y. Sun, Z. Liu, S. Todorovic, J. Li, Adaptive boosting for SAR automatic target recognition, IEEE Trans. Aerosp. Electron. Syst. 43 (1) (2007) 112–125.
[10] H. Pan, L.Z. Xia, Efficient object recognition using boundary representation and wavelet neural network, IEEE Trans. Neural Networks 19 (12) (2008) 2132–2149.
[11] J.A. O'Sullivan, M.D. DeVore, V. Kedia, SAR ATR performance using a conditionally Gaussian model, IEEE Trans. Aerosp. Electron. Syst. 37 (1) (2001) 91–108.
[12] M.D. DeVore, J.A. O'Sullivan, Performance complexity study of several approaches to automatic target recognition from SAR images, IEEE Trans. Aerosp. Electron. Syst. 38 (2) (2002) 632–648.
[13] J. Wright, A.Y. Yang, A. Ganesh, et al., Robust face recognition via sparse representation, IEEE Trans. Pattern Anal Mach. Intell. 31 (2) (2009) 210–227.
[14] J.J. Thiagarajan, K.N. Ramamurthy, P. Knee, et al., Sparse representations for automatic target classification in SAR images, In: Proceedings of the 4th International Symposium on Communications, Control and Signal Processing, Limassol, Cyprus, March 2010, pp. 1–4.
[15] A. Sinha, H.M. Chen, D.G. Danu, et al., Estimation and decision fusion: A survey, Neurocomputing 71 (13–15) (2008) 2650–2656.
[16] M.Z. Brown, Analysis of multiple-view Bayesian classification for SAR ATR, Proc. SPIE 5095 (2003) 265–274.

[17] A. Aziz, A soft-decision fusion approach for multiple-sensor distributed binary detection systems, IEEE Trans. Aerosp. Electron. Syst. 47 (3) (2011) 2208–2216.
[18] G. Simone, A. Farina, F.C. Morabito, et al., Image fusion techniques for remote sensing applications, Inf. Fusion 3 (1) (2002) 3–15.
[19] H. Cevikalp, H.S. Yavuz, M.A. Cayet, A. Barkana, Two-dimensional subspace classifiers for face recognition, Neurocomputing 72 (4–6) (2009) 1111–1120.
[20] T. Balz, S. Becker, N. Haala, M. Kada, Using real-time SAR simulation to assist pattern recognition applications in urban areas, Pattern Recognition Image Anal. 18 (3) (2008) 412–416.
[21] M. Greenspan, L. Pham, N. Tardella, Development and evaluation of a real time SAR ATR system, In: IEEE Proceedings Radar Conference, Dallas, Texas, May, 1998, pp. 38–43.
[22] J. Chuanyi, M. Sheng, Combinations of weak classifiers, IEEE Trans. Neural Networks 8 (1) (1997) 32–42.
[23] T. Parag, F. Porikli, A. Elgammal, Boosting adaptive linear weak classifiers for online learning and tracking, In: IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, June 2008, pp. 1–8.
[24] S.G. Mallat, Z.F. Zhang, Matching pursuit with time-frequency dictionaries, IEEE Trans. Signal Process 41 (12) (1993) 3397–3415.
[25] X.Y. Wang, Q.Y. Wang, H.Y. Yang, J. Bu, Color image segmentation using automatic pixel classification with support vector machine, Neurocomputing 74 (18) (2011) 3898–3911.
[26] X.Z. Wang, S.X. Lu, J.H. Zhai, Fast fuzzy multicategory SVM based on support vector domain description, Int. J. Pattern Recognition Artif. Intell.s 22 (1) (2008) 109–120.
[27] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 1–27.
[28] Y.S. Dong, K.S. Han. Boosting S.V.M. ,Classifiers by ensemble, In: ACM Special Interest Tracks and Posters on World Wide Web. Chiba, Japan, May 2005, pp. 1072–1073.
[29] D.F. Llorca, F. Vilarino, Z. Zhou, G. Lacey, A multi-class SVM classifier ensemble for automatic hand washing quality assessment, In: BMVC Proceedings on British Machine Vision Conference, Warwick, UK, September 2007, pp. 213–223.



Haicang Liu received the B.S. and M.S. degrees from the College of Electrical and Information Engineering, from Hunan University, Changsha, China, in 2006 and 2011 respectively. He is currently working towards the Ph.D. degree in image processing in Hunan University, Changsha, China. His research interests are in SAR image processing, pattern recognition, target detecting and tracking.

Shutao Li received the B.S., M.S., and Ph.D. degrees in electrical engineering from Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively.

In 2001, he joined the College of Electrical and Information Engineering, Hunan University. From May 2001 to October 2001, he was a Research Associate with the Department of Computer Science, Hong Kong University of Science and Technology, Kowloon, Hong Kong. From November 2002 to November 2003, he was a Postdoctoral Fellow with the Royal Holloway College, University of London, U.K., working with Prof. J.-S. Taylor. From April 2005 to June 2005, he was a Visiting Professor with the Department of Computer Science, Hong Kong University of Science and Technology. He is currently a Full Professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 150 refereed papers. His professional interests are information fusion, pattern recognition, bioinformatics, and image processing. Dr. Li served as a member of the Neural Networks Technical Committee from 2007 to 2008. He has won two Second-Grade National Awards at the Science and Technology Progress of China in 2004 and 2006.