Kernel linear representation: application to target recognition in synthetic aperture radar images

Ganggang Dong
Na Wang
Gangyao Kuang
Yinfa Zhang

Kernel linear representation: application to target recognition in synthetic aperture radar images

Ganggang Dong,[a,][*] Na Wang,[a] Gangyao Kuang,[a] and Yinfa Zhang[b]

[a]National University of Defense Technology, College of Electronics Science and Engineering, Street YaWaChi 137#, Changsha 410073, China
[b]School of Xi'an Communication, Xi'an 710106, China

Abstract. A method for target classification in synthetic aperture radar (SAR) images is proposed. The samples are first mapped into a high-dimensional feature space in which samples from the same class are assumed to span a linear subspace. Then, any new sample can be uniquely represented by the training samples within given constraint. The conventional methods suggest searching the sparest representations with ℓ_1-norm (or ℓ_0) minimization constraint. However, these methods are computationally expensive due to optimizing nondifferential objective function. To improve the performance while reducing the computational consumption, a simple yet effective classification scheme called kernel linear representation (KLR) is presented. Different from the previous works, KLR limits the feasible set of representations with a much weaker constraint, ℓ_2-norm minimization. Since, KLR can be solved in closed form there is no need to perform the ℓ_1-minimization, and hence the calculation burden has been lessened. Meanwhile, the classification accuracy has been improved due to the relaxation of the constraint. Extensive experiments on a real SAR dataset demonstrate that the proposed method outperforms the kernel sparse models as well as the previous works performed on SAR target recognition.
© The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JRS.8.083613]

1 Introduction

Synthetic aperture radar (SAR) has been widely used in many fields, such as environmental monitoring, surveillance, and reconnaissance, due to its ability to work 24-hour a day and its robustness with inclement weather conditions. Automatic target recognition (ATR) is a fundamental topic of SAR image interpretation. It has been studied extensively in the past two decades, yet it is still an open problem. Particularly, it is challenging to perform target recognition in the extended operating conditions,[1–3] in which a single operational parameter is significantly different between the images used for training and those used for testing. A typical SAR ATR system identifies the unknown through three sequential stages:[4–6] detection,[7] discrimination,[8] and classification.[9] First, targets as well as various clutter false alarms (e.g., buildings, trees, streetlights, etc.) are detected. Then, natural and manmade clutter false alarms are rejected in the discrimination stage, followed by a classifier to give the identity. Only the final procedure, classification, is studied in this article.

Sparse and redundant signal representations have recently drawn much interest in computer vision, signal and image processing,[10] due to the fact that signals and images of interest can be sparse or compressible with respect to a given dictionary. In Ref. 11, Huang and Aviyente propose the sparse representation (SR) over a redundant basis set for signal classification. The sparse coefficients are obtained by optimizing an objective function that includes the measurement of the reconstruction error and sparsity level. In Ref. 12, Wright et al. generalize the SR

*Address all correspondence to: Ganggang Dong, E-mail: dongganggang@nudt.edu.cn

technique for robust face recognition. By assuming that a linear subspace can be spanned using the samples belonging to the same class, any new sample can be recovered by a linear combination of the training samples from all classes. To search for the most parsimonious representation, the constraint, ℓ_1-norm minimization, is imposed on the encoding coefficients. Similarly, an SR technique has been introduced to the ATR framework. In Ref. 13, Chen et al. propose a sparsity-based algorithm for automatic target detection in hyperspectral imagery. The test pixel is linearly reconstructed by the training pixels with a sparsity constraint, and the characteristics of coefficients on reconstruction are used to make a decision. To circumvent pose estimation and the multiple preprocessing procedures, Ref. 14 performs target classification in SAR images with the SR method. They explain the advantage of SR from the perspective of manifolds' learning. In Ref. 9, Patel et al. present a block sparsity-based classification method, in which the inherent block structure of encoding coefficients has been exploited with a group sparsity constraint. Although great performances have been reported in these works, they may be not effective if the dataset is not linearly separable in the original space.

To cover the shortage of the linear SR techniques, some works transform the samples into a new feature space induced by a nonlinear mapping. However, it is infeasible to solve the presented problem directly due to the implicit nonlinear mapping. In Ref. 15, Gao et al. relax the nonconvex problem via alternately computing sparse coefficients and a learning redundant dictionary. In Ref. 16, Yin et al. propose another approach to search the most compact representations. Both the test and the training samples in feature space are projected into a novel space by left multiplying a linear operator. In Ref. 17, Zhang et al. circumvent the explicit computation in feature space with a dimension reduction strategy. However, these methods are computationally expensive due to optimizing a nondifferential objective function. Moreover, in the feature space, there is actually no need to impose the strong sparsity constraint on the representations. A much weaker constraint, ℓ_2-minimization, can play the same role but with less computational consumption.[18,19]

To improve the performance while alleviating the computational cost, a new classification method named the kernel linear representation (KLR) is presented in this article. All the samples are first mapped into an implicit feature space by a nonlinear mapping. The classification is then implemented with respect to the data in the feature space. In the feature space, it is assumed that the samples belonging to the same class approximately span a linear subspace, thus any new sample can be represented by a linear combination of all the training samples. To uniquely recover the test sample, the conventional methods impose the strong sparsity constraint on the encoding coefficients. The idea in this article is to limit the feasible set of the representation with a much weaker constraint, ℓ_2-norm minimization. Due to the convexity and differentiability, the presented problem can be solved in closed form. Thus, the complicated procedure to optimize the strong sparsity constraint problem has been circumvented. The unknown is identified according to the characteristics of representations on reconstruction. Compared with the forerunners' works,[15–17] the proposed method performs much faster due to the analytic solution, and the accuracy has been improved because of the relaxation of the constraints.

Although the proposed method is of broad interest to object recognition in general, the studies and experimental results in this article are confined to high-resolution SAR target recognition (i.e., MSTAR dataset). The proposed method does not rely on any preprocessing procedure, such as pose estimation, noise reduction, and binary-value. It is robust to small variations in configuration, pose, and depression angles.

2 Methodology

2.1 *Sparse Representation*

SR aims to succinctly recover the signal over a given dictionary. It is based on the assumption that a linear subspace can be spanned by samples belonging to the same class. Given sufficient training samples of the i'th class, $\mathbf{X}_i = [\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \ldots, \mathbf{x}_{i,n_i}] \in \mathbb{R}^{m \times n_i}$, where m is the data dimension stacked as a column. Any new sample $\mathbf{y} \in \mathbb{R}^m$ from the same class will approximately lie in

the linear span of the training samples associated with the i'th class, $\mathbf{y} = \mathbf{x}_{i,1}\alpha_{i,1} + \mathbf{x}_{i,2}\alpha_{i,2} + \ldots + \mathbf{x}_{i,n_i}\alpha_{i,n_i}$, where $\alpha_i = [\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n_i}]^T \in \mathbb{R}^{n_i}$ is the coefficients. Since the membership of the new sample is initially unknown, a dictionary has been defined by concatenating the n training samples of all k distinct classes, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k] \in \mathbb{R}^{m \times n}$, where $n = \sum_{i=1}^{k} n_i$. Then, the linear representation of $\mathbf{y}$ can be rewritten in terms of all training samples

$$\mathbf{y} = \mathbf{X}_1\alpha_1 + \mathbf{X}_2\alpha_2 + \cdots + \mathbf{X}_k\alpha_k = \mathbf{X}\alpha, \tag{1}$$

where $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_k]^T \in \mathbb{R}^n$ is the representation vector whose entries are zeros except those associated with the i'th class in theory. Usually, we consider that the underdetermined systems ($m < n$), and its solution is not unique. The popular methods are to search the sparsest solution by ℓ_0-norm minimization[12]

$$\min_{\alpha} \|\alpha\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{X}\alpha\|_2 \leq \varepsilon, \tag{2}$$

where $\| \cdot \|_0 : \mathbb{R}^n \mapsto \mathbb{R}$ counts the nonzero entries and ε is the error tolerance. In Eq. (2), the objective function is to measure the sparsity level, whereas the constrained term is to control the reconstruction error (i.e., fidelity). Due to the nonconvex and nondifferential properties, solving Eq. (2) is nondeterministic polynomial (NP)-hard. Thanks to the development of compressed sensing theories, the solution of the ℓ_0-norm minimization problem is equal to the one of the ℓ_1-norm minimization if the solution is sparse enough[20]

$$\min_{\alpha} \|\alpha\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{X}\alpha\|_2 \leq \varepsilon. \tag{3}$$

At present, many algorithms have been presented to solve Eq. (3), as comprehensively reviewed in Ref. 21. Then, the test sample is classified as the class whose training samples can generate the minimum residual

$$\min_{i=1,\ldots k} \{\|\mathbf{y} - \mathbf{X}_i\hat{\alpha}_i\|_2^2\}. \tag{4}$$

2.2 Kernel Sparse Representation

In the previous works,[9,12,14] although great performance has been reported, they may be not effective if the dataset is not linearly separable originally. To improve the performance, it is natural to cast the data into a new feature space in which the data separability between different classes has been enhanced. Suppose the feature space (denote by $\mathfrak{I}$) is induced by the nonlinear mapping, $\phi: \mathbb{R}^m \mapsto \mathfrak{I}$, and the kernel function, $\kappa: \mathbb{R}^m \times \mathbb{R}^m \mapsto \mathbb{R}$, is defined as the inner product in the feature space

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle, \tag{5}$$

where $\langle \cdot, \cdot \rangle$ is the inner product. Commonly used kernel functions include a polynomial kernel, a sigmoidal kernel, and a Gaussian radial basis function (RBF).

In the feature space, it is assumed that the samples belonging to the same class approximately span a linear subspace. Thus, any new sample $\phi(\mathbf{y}) \in \mathfrak{I}$ can be recovered over all the training samples in the feature space

$$\phi(\mathbf{y}) = \phi(\mathbf{X})\alpha, \tag{6}$$

where $\phi(\mathbf{X}) = [\phi(\mathbf{X}_1), \phi(\mathbf{X}_2), \ldots, \phi(\mathbf{X}_k)]$; $\phi(\mathbf{X}_i) = [\phi(\mathbf{x}_{i,1}), \phi(\mathbf{x}_{i,2}), \ldots, \phi(\mathbf{x}_{i,n_i})]$, $\phi(\mathbf{x}_{i,j}) \in \mathfrak{I}$; $\alpha \in \mathbb{R}^n$ is the representation vector. To create the unique solution of Eq. (6), the previous works impose a strong sparsity constraint (i.e., ℓ_1-norm minimization) on the representations

$$\min_{\alpha} \|\alpha\|_1 \quad \text{s.t.} \quad \|\phi(\mathbf{y}) - \phi(\mathbf{X})\alpha\|_2 \leq \varepsilon.$$

2.3 *Kernel Linear Representation*

The previous works[15–17] compute the sparsest representation with an ℓ_1-norm minimization constraint. The idea here is to limit the feasible set of the representations with a much weaker constraint, ℓ_2-norm minimization

$$\min_{\alpha}\|\alpha\|_2 \quad \text{s.t.} \ \|\phi(\mathbf{y}) - \phi(\mathbf{X})\alpha\|_2 \le \varepsilon. \tag{8}$$

Which is also equal to the unconstrained problem

$$\min_{\alpha}\{f(\alpha) = \|\phi(\mathbf{y}) - \phi(\mathbf{X})\alpha\|_2 + \lambda\|\alpha\|_2\}. \tag{9}$$

Obviously, Eq. (9) can be solved in closed form due to the convex and differential objective functions. By unfolding the fidelity term, it can be rewritten as

$$\min_{\alpha}\left\{ \begin{array}{l} f(\alpha) = [\phi(\mathbf{y}) - \phi(\mathbf{X})\alpha]^T[\phi(\mathbf{y}) - \phi(\mathbf{X})\alpha] + \lambda\|\alpha\|_2 \\ = \phi(\mathbf{y})^T\phi(\mathbf{y}) + \alpha^T\phi(\mathbf{X})^T\phi(\mathbf{X})\alpha - 2\phi(\mathbf{y})^T\phi(\mathbf{X})\alpha + \lambda\|\alpha\|_2 \\ = 1 + \alpha^T\Phi\alpha - 2\Phi_{\mathbf{y}}\alpha + \lambda\alpha^T\alpha \end{array} \right\}, \tag{10}$$

where $\Phi_{\mathbf{y}} = [\kappa(\mathbf{y}, \mathbf{x}_1), \kappa(\mathbf{y}, \mathbf{x}_2), \dots, \kappa(\mathbf{y}, \mathbf{x}_n)]^T$, and

$$\Phi = \begin{pmatrix} \kappa(\mathbf{x}_1, \mathbf{x}_1) & \dots & \kappa(\mathbf{x}_1, \mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ \kappa(\mathbf{x}_n, \mathbf{x}_1) & \cdots & \kappa(\mathbf{x}_n, \mathbf{x}_n) \end{pmatrix}$$

is the kernel Gram matrix. It is apparent that the objective function of Eq. (10) is tractable since it only refers to the matrix operation of finite dimension, $\Phi_{\mathbf{y}} \in \mathbb{R}^n$ and $\Phi \in \mathbb{R}^{n \times n}$, rather than dealing with a possibly infinitely dimensional dictionary, $\phi(\mathbf{X})$. An important hint of this formulation is that the computation of $\Phi_{\mathbf{y}}$ and Φ only requires the dot products. It is, therefore, feasible to solve Eq. (10) with Mercer kernel tricks[22] regardless of the nonlinear mapping ϕ. With the kernel trick [Eq. (5)], searching the coefficients over the dictionary, $\phi(\mathbf{X})$, is then converted to compute the representations in terms of the kernel Gram matrix Φ. So, the explicit computation in the feature space has been circumvented. Following the mathematic rules, we conduct the partial differential of the representation

$$\frac{\partial f(\alpha)}{\partial \alpha} = 2\Phi\alpha - 2\Phi_{\mathbf{y}} + 2\lambda\alpha. \tag{11}$$

By setting the derivative to be zero, $[\partial f(\alpha)/\partial \alpha] = 0$, it is easy to obtain the analytic solution

$$\hat{\alpha} = (\Phi + \lambda\mathbf{I})^{-1}\Phi_{\mathbf{y}}, \tag{12}$$

where $\mathbf{I}$ is the identity matrix whose dimension is the same as kernel Gram matrix.

Similarly, the decision is made by the characteristics of the representations on reconstruction

$$\min_{i=1,\dots,k}\{\|\Phi_{\mathbf{y}} - \Phi\delta_i(\hat{\alpha})\|_2^2\}, \tag{13}$$

where $\delta_i(\cdot): \mathbb{R}^n \mapsto \mathbb{R}^n$ is the mapping to pick out the coefficients associated with the i'th class.

2.4 *Validation*

The ability to determine whether the input sample is valid or not is crucial for the classifier to work in real-world situations. A typical target recognition system, for example, should reject the civil vehicles, buildings, and trees. In the proposed framework, the validation judgment is in terms of the reconstruction error. That is, the algorithm accepts or rejects a test sample based on how small the minimum residual is. Given a solution $\hat{\alpha}$ found by Eq. (12), the residual
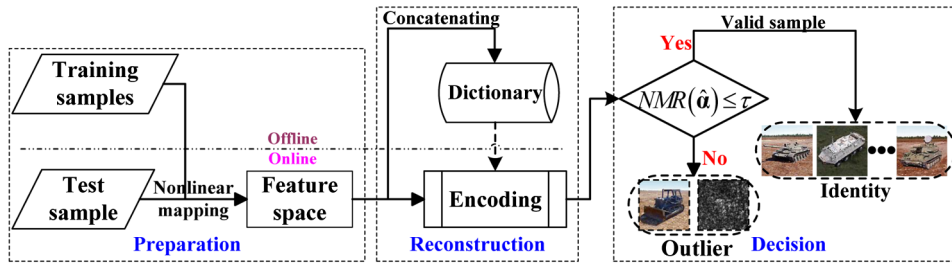
Fig. 1 The block diagram of the proposed algorithm.

vector can be built as $\mathbf{e} = [e_1, e_2, \ldots, e_k] \in \mathbb{R}^k$, where $e_i = \|\Phi_{\mathbf{y}} - \Phi\delta_i(\hat{\alpha})\|_2^2$, $i = 1, \ldots, k$ is the residual associated with the i'th class. To measure the quality of the test sample, an index named the normalized minimum residual (NMR) has been defined as

$$\text{NMR}(\hat{\alpha}) = \min_{i=1,\ldots,k} \left\{ \frac{\|\Phi_{\mathbf{y}} - \Phi\delta_i(\hat{\alpha})\|_2^2}{\sum_{j=1}^{C} \|\Phi_{\mathbf{y}} - \Phi\delta_j(\hat{\alpha})\|_2^2} \right\}. \tag{14}$$

For any coefficient vector $\hat{\alpha}$, the smaller the $\text{NMR}(\hat{\alpha})$ is, the better the quality of the test sample is, and hence belief that the test sample is a valid one is higher, vice versa. Given a threshold $\tau \in (0, 1)$, a test sample is accepted as valid if $\text{NMR}(\hat{\alpha}) \leq \tau$, and otherwise it is rejected as the outlier. Only the test sample that passes the criterion can be assigned a class label.

The framework of the proposed method has been pictorially summarized in Fig. 1, where it has been divided into three sequential stages. The first stage is devoted to data preparation. All the samples are mapped into the feature space with a nonlinear mapping. In the second stage, the training samples are used to formulate the redundant dictionary to encode an input test sample as a linear combination of them via the ℓ_2-norm minimization. The last stage contributes to the decision. The identity is assigned to the test sample if it passes the rejection criterion [Eq. (14)], otherwise it is determined to be an outlier.

3 Experimental Results and Discussions

This section demonstrates the performance of the proposed method on MSTAR database, a collection done using a 10-GHz spotlight SAR sensor with a one-foot resolution. For each target, images are captured at different depression angles over a full 0 to 359 deg range of aspect view. To ensure the accuracy and efficiency, a wide variety of experiments are performed, including configuration variations, pose and depression angle variations, outlier rejection, etc. In all the experiments, the center 80×80 pixels patch is used as the input. The cropped images are first subsampled by a factor of ρ, and the subsampled images are then mapped into the feature space. The subsampling factor is chosen from a given interval $\rho = \{1/400, 1/256, 1/100, 1/64\}$, which corresponds to sizes 4×4, 5×5, 8×8, and 10×10 pixels. Gaussian RBF, $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|_2^2)$, is employed as the kernel function, and the width parameter, γ, is assigned as 200, 50, 10, and 5 for the subsampled image of 4×4, 5×5, 8×8, and 10×10 pixels, respectively. The baseline methods include linear SR method[12] and kernel sparse techniques, i.e., Gao et al.'s method[15] [Kernel sparse representation (KSR)], Yin et al.'s method[16] [Kernel sparse representation projection, (KSRP)], and Zhang et al.'s method[17] [Kernel sparse representation with dimension reduction (KSRDR)], linear support vector machine (SVM) and kernel SVM.

3.1 Parameter Setting

In the proposed method, there is a free parameter to be set, the regularizer λ. It is used to balance the fidelity and "sparsity." First, it circumvents the difficulty when the kernel Gram matrix, Φ, is not invertible, so the solution is more stable. Second, it introduces a particular "sparsity" to the representations, and the sparsity is much weaker than the ℓ_1-norm minimization. To determine λ, five targets, BMP2, BTR70, T72, T62, and BTR60, are employed for several groups of

Table 1 The number of aspect view images available for different targets in parameter setting, where the training samples are in bold.

Angle	BMP2	T72	BTR70	T62	BTR60	SUM
17°	**233**	**232**	**233**	**299**	**256**	**1253**
15°	587	582	196	273	195	1833

experiments. The number of aspect view images available for different targets is shown in Table 1, and the recognition rates with different parameters are drawn in Fig. 2.

As can be seen from Fig. 2, when λ decreased from 0.08 to 0.006, the accuracy first improves and the subsequently degrades. The peak value happens at $\lambda = 0.02$ or 0.03. Though different rates have been obtained, the accuracy varies very slightly. Thus, we can come to the conclusion that the proposed method is insensitive to the regularizer. It is, however, necessary for stable matrix inversion. Hereafter, the parameter will be fixed as $\lambda = 0.02$.

3.2 Configuration Invariance

This subsection examines the invariance of different algorithms under different configurations. Three basic targets, T72, BMP2, and BTR70, have been utilized. For T72 and BMP2, there are three variants with different serial numbers and small structural modifications, SN_132, SN_812, SN_s7, SN_9563, SN_9566, and SN_c21. Only the standard configurations, SN_132 (T72), SN_9563 (BMP2), and SN_c71 (BTR70), at a 17-deg depression angle are used for training, whereas all configurations at a 15-deg depression angle are used for testing, as detailed in Table 2. The performance obtained by different algorithms is listed in Table 3, where the former item gives the recognition rates, and the latter item shows the run-times in seconds.

As can be seen from Fig. 3, kernel-based methods, KSR, KSRP, KSRDR, and KLR, significantly outperform the linear one, SR. The average improvements of 12.18%, 14.65%, 13.86%, and 19.65% have been obtained by KSR, KSRP, KSRDR, and KLR against SR. This is because the dataset is not linearly separable in the original space, and it can be remedied by mapping the data into a new feature space. The proposed method achieves the best recognition rates, i.e., 0.7728, 0.8593, 0.9406, and 0.9575, and the average improvements of 7.48%, 5.0%, 5.79%, 8.04%, and 4.71% have been obtained against KSR, KSRP, KSRDR, SVM, and KSVM. The good performance can be attributed to two characteristics. First, the high-order structure of data has been exploited with the nonlinear mapping. Second, the discriminative ability of the encoding coefficients has been enhanced due to the relaxation of constraints.
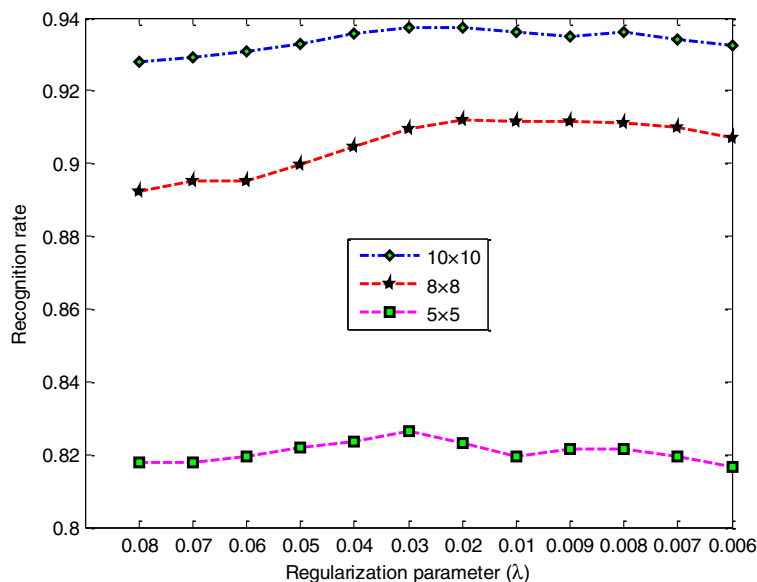

Fig. 2 The recognition rates of kernel linear representation with different parameters.

Table 2 The number of aspect views available for targets in configuration invariance.

	BMP2			T72			BTR70	
Angle	[SN_9563]	SN_9566	SN_c21	[SN_132]	SN_812	SN_s7	[SN_c71]	SUM
17°	**233**	–	–	**232**	–	–	**233**	**698**
15°	195	196	196	196	195	191	196	1365

Table 3 Performance of various methods in configuration invariance.

Dim.	SR	SVM	KSVM	KSR	KSRP	KSRDR	KLR
4 × 4	0.5692/20	0.7018/0.5	0.6901/0.5	0.7413/35	0.7582/58	0.7626/28	**0.7728**/4.5
5 × 5	0.6373/22	0.8095/0.5	0.7926/0.5	0.8131/56	0.8087/107	0.8007/28	**0.8593**/4.5
8 × 8	0.7399/32	0.8241/0.7	0.9164/0.7	0.8219/108	0.8717/155	0.8659/28	**0.9406**/4.5
10 × 10	0.7978/36	0.8732/0.8	0.9428/0.8	0.8549/155	0.8915/165	0.8695/28	**0.9575**/4.5

The computational cost, measured by run-times, is pictorially shown in Fig. 4. It can be observed that the proposed method performs much faster than the linear SR and kernel SR methods, and slower than SVM and KSVM. For the low-dimensional classification (16-D), KLR achieves the highest recognition rate, 0.7728, with a runtime of 4.5 s, and it is 4.4, 7.7, 12.8, and 6.2 times faster than SR, KSR, KSRP, and DSRDR. For the high-dimensional classification (100-D), KLR again obtains the highest recognition rate, 0.9575, with a run-time of 4.5 s, and it is 8, 34, 36, and 6.2 times faster than SR, KSR, KSRP, and DSRDR. This is because the complicated procedure to ℓ_1-norm minimization has been circumvented.

3.3 Comparison with Conventional Approaches to Synthetic Aperture Radar Target Recognition

The past two decades have witnessed great developments in SAR target recognition, and a wide variety of techniques are presented, such as template matching methods,[1] subspace [principal
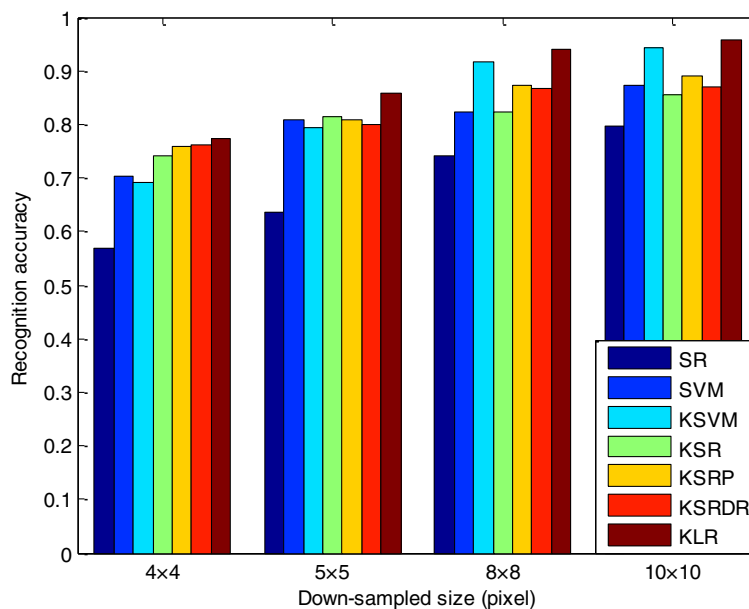

Fig. 3 The recognition rates of different methods in configuration invariance experiment.
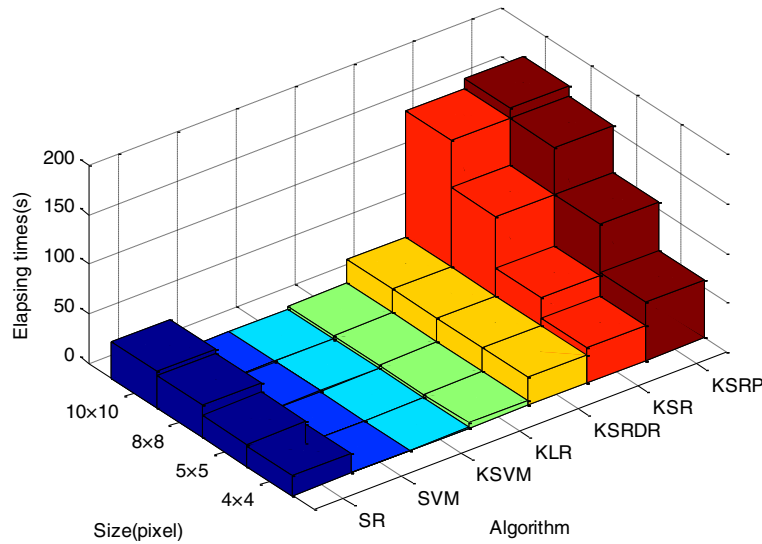
Fig. 4 The runtimes of different methods in configuration invariance.

component analysis (PCA), linear discriminant analysis (LDA), etc.] methods,[23,24] transform [discrete cosine transform (DCT), etc.],[25] Fourier domain techniques,[26] etc. This subsection compares the proposed method with the previous works performed on SAR target recognition. Following the prototype of the above experiment, BMP2, BTR70, and T72 have been utilized. The classification accuracies obtained using different algorithms are shown in Table 4, where TM denotes the template matching method with the templates at 10-deg increments and a mask individualizing the targets.

As can be observed from Table 4, the template matching method achieves a similar performance to subspace (PCA) and transform (DCT) based methods (0.9040 versus 0.9076, 0.9055). All of them outperform another supervised classification technique, LDA. This is because there are only three different classes in the training dataset, which results in very lower dimensional projected data (i.e., 2-D) with LDA. The Fourier domain algorithm outperforms the TM-, PCA-, DCT-, and LDA-based methods. The proposed method, KLR achieves the highest recognition rate, 0.9575, and the improvements of 5.35%, 4.99%, 7.5%, 5.2%, and 4.03% have been obtained compared with TM, PCA, LDA, DCT, and FT. The experimental results are mainly due to the advantage of nonlinear mapping, which maps the data into a feature space where the data separability between different classes has been enhanced. Furthermore, the classification with collaborative representation of all training samples also contributes to the good performance, i.e., KLR recovers the test sample with training samples of all the classes, while the others reconstruct the test sample using several relevant ones.

3.4 *Outlier Rejection*

In the "open world" classification issues, the classifier should be capable of rejecting the unknown objects which do not belong to the training set classes (i.e., outlier). Thus, this subsection considers the rejection of confusers and clutters. Three military vehicles, BMP2 (SN_9563), BTR70 (SN_c21), and T72 (SN_132), are used as the standard targets, whereas D7 (a bulldozer) is specified as the confuser to be rejected. Samples of the standard target and the confuser are illustrated in Fig. 5. Following the previous works,[3,26,27] the 1159 "target-like" clutter chips are generated from

Table 4 Comparison with several conventional approaches to synthetic aperture radar target recognition.

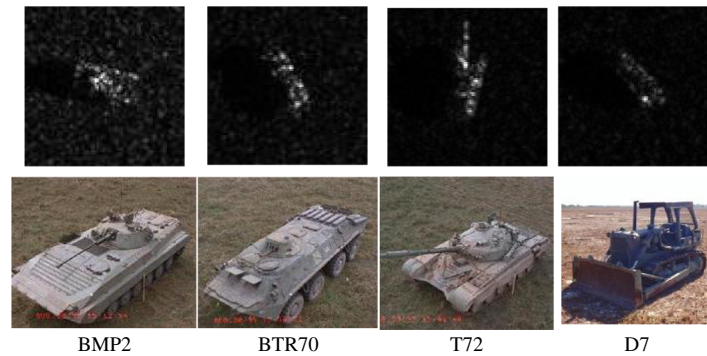|	TM	PCA	LDA	DCT	FT	KLR#1	KLR#2
Accuracy	0.9040	0.9076	0.8825	0.9055	0.9172	0.9406	**0.9575**

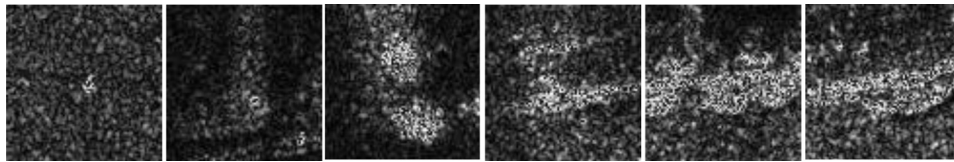**Fig. 5** Samples of the standard target and confuser.

Fig. 6 Samples of the clutter chips.

the 100 scene images of MSTAR clutters, and 200 ones of them are used as the clutters to be rejected. A small portion of the clutter chips are demonstrated in Fig. 6. The total number of chip images available for standard targets and the outliers are listed in Table 5. By tuning the threshold τ through a range of values in $(0, 1)$, a receiver operating characteristic (ROC) curve, which plots the false reject rate against false accept rate for all possible thresholds, can be created. The ROC curve visually reflects the classifier's ability to determine whether a given test sample is in the training dataset. The results of confuser rejection and clutter rejection experiments have been drawn in Figs. 7 and 8, where the performance of four different algorithms, KLR, KSR, KSRP, and KSRDR, have been evaluated.

As can be seen from Figs. 7 and 8, it is more difficult to determine the confuser (D7) than the clutter. This is because the confuser D7 is a manmade object. It produces similar scattering centers to three standard targets, as easily observed from Fig. 5. The clutters, however, are mainly composed of buildings, trees, roads, streetlights, etc. Their scattering centers are much more random and dispersive than typical tactical ground targets, as demonstrated in Fig. 6. Moreover, whether in the confuser rejection experiment or in the clutter rejection experiment, the proposed method significantly outperforms all the reference algorithms, KSR, KSRP, and KSRDR.

3.5 *10-Object Recognition*

To further evaluate the performance, a group of more challenging experiments have been subsequently performed. Several different kinds of targets, main battle tank (T72 and T62), armored personnel carrier (BTR70 and BTR60), bulldozer (D7), truck (ZIL131), antiaircraft gun (ZSU_23/4), armored truck (BMP2 and BRDM_2), and howitzer (2S1), are employed, as

Table 5 The number of chip images available for the rejection experiment.

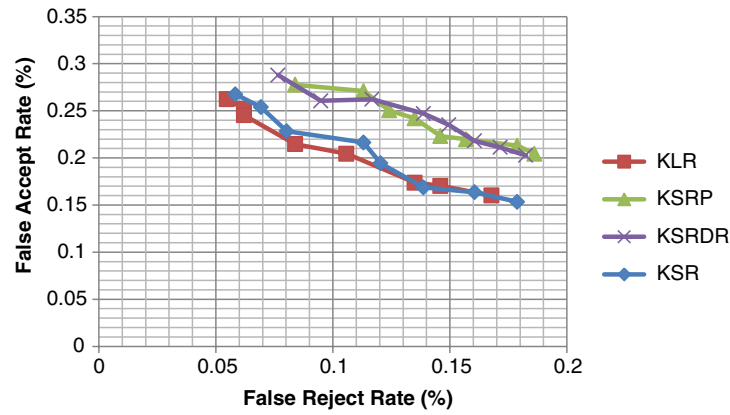|	BMP2	BTR70	T72	D7	Clutter
Training	**233**	**233**	**232**	–	–
Testing	195	196	195	274	200

Fig. 7 Receiver operating characteristic (ROC) curves for confuser rejection.
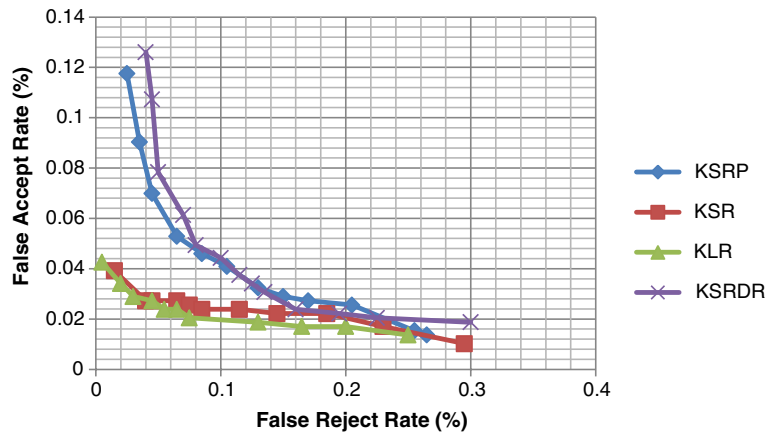

Fig. 8 ROC curves for clutter rejection.

Table 6 The number of aspect view images available for different targets in 10-object recognition.

Angle	BMP2	T72	BTR70	BTR60	2S1	BRDM2	D7	T62	ZIL131	ZSU23/4	SUM
17°	**233**	**232**	**233**	**256**	**299**	**298**	**299**	**299**	**299**	**299**	**2747**
15°	587	582	196	195	274	274	274	273	274	274	3203

Table 7 The results of 10-object recognition experiment.

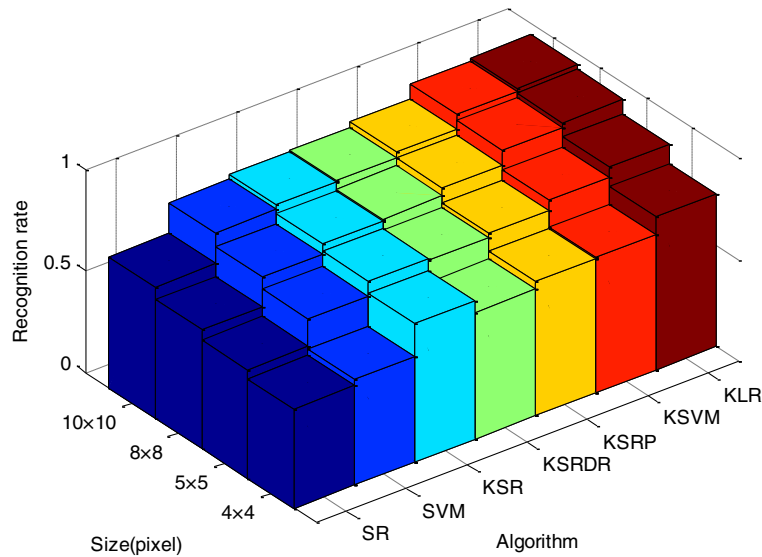| Dim.	SR	SVM	KSVM	KSR	KSRP	KSRDR	KLR
4 × 4	0.4889	0.5176	0.6646	0.6781	0.6584	0.6219	**0.7499**
5 × 5	0.5332	0.6665	0.8005	0.7399	0.7519	0.7322	**0.8476**
8 × 8	0.5816	0.7258	0.8979	0.7764	0.8239	0.7848	**0.9191**
10 × 10	0.6453	0.7986	0.9244	0.8204	0.8534	0.8261	**0.9428**

Fig. 9 The recognition rates of different methods in 10-object recognition experiment.

detailed in Table 6. The recognition rates of different algorithms are listed in Table 7, with the corresponding diagram shown in Fig. 9.

The results conform to the above experiments. Higher accuracies are obtained in high-dimensional space (100-D and 64-D) than in low-dimension space (25-D and 16-D). Compared to the linear models (i.e., SR and SVM), better performance has been obtained using kernel models (KSVM, KSR, KSRP, KSRDR, and KLR). This is due to the nonlinear mapping which improves the data separability. As for KLR, the recognition rates of 0.7499, 0.8476, 0.9191, and 0.9428 have been obtained, which outperforms all the baseline methods. Two sides contribute to the experimental results. The improvement in accuracy against KSVM results from the advantage of collaborative representation, whereas the elevation in accuracy than KSR, KSRP, and KSRDR is due to the relaxation of constraints.

4 Conclusion

Recent years have witnessed a considerable resurgence of interest in sparse signal representation. This popularity comes from the fact that signals in most problems can be well recovered over a small set of basis vectors. However, these methods are computationally expensive due to optimizing the nondifferential objective function. To improve the performance while boosting the speed, a new classification algorithm named KLR is presented in this article, and it is applied to target recognition in high-resolution SAR images. The classification is implemented with respect to the data in a feature space induced by a nonlinear mapping. Since the mapping is of implicit form, it is infeasible to solve the presented problem directly. To produce the unique solution, the conventional methods impose strong sparsity constraint on the representation. Our idea in this article is to limit the feasible set of the representation with a much weaker constraint, ℓ_2-norm minimization. Thus, the complicated procedure to optimizing sparsity constraint problem has been converted to a simple least-square fitting. Due to the convexity and differentiability, the new problem can be solved in closed form. Therefore, the computational cost has been significantly lessened, and the classification accuracy has been simultaneously improved. Extensive experiments, including configuration invariance, depression angle invariance, outlier rejection, and 10-object recognition, have been carried out on the MSTAR dataset. The experimental results demonstrate that the proposed method performs much faster than various reference algorithms. Moreover, it achieves much higher recognition rates than the kernel sparse models, as well as the conventional approaches to SAR target recognition. The improvement in accuracy against the previous works performed on SAR target recognition is due to the advantage of

collaborative representation, whereas the elevation in speed over kernel sparse models results from the relaxation of constraints.

The proposed method refers to matrix inverse for computing the encoding coefficients. Thus, it is difficult to deal with a large-scale classification task due to the bottleneck of realizing a high-dimensional matrix inverse. Future attention will be paid to covering the shortage with dictionary learning skills. Another intriguing question for future work is whether the presented framework can be useful for other stages of ATR, for example, automatic target detection and discrimination.

References

1. T. Ross et al., "Standard SAR ATR evaluation experiments using the MSTAR public release data set," *Proc. SPIE* **3370**, 554–565 (1998).
2. J. C. Mossing and T. D. Ross, "An evaluation of SAT ATR algorithm performance sensitivity to MSTAR extended operating conditions," *Proc. SPIE* **3370**, 554–565 (1998).
3. R. Singh and B. Kumar, "Performance of the extended maximum average correlation height filter and the polynomial distance classifier correlation filter for multiclass SAR detection and classification," *Proc. SPIE* **4727**, 265–279 (2002).
4. L. M. Novak, G. J. Owirka, and C. M. Netishen, "Performace of a high-resolution polarimetric SAR automatic target recognition system," *Lincoln Lab. J.* **6**(1), 11–23 (1993).
5. G. J. Owirka, S. M. Verbout, and L. M. Novak, "Template-based SAR ATR performance using different image enhancement techniques," *Proc. SPIE* **3721**, 302–319 (1999).
6. J. A. O'Sullivan et al., "SAR ATR performance using a conditionally Gaussian model," *IEEE Trans. Aerosp. Electron. Syst.* **37**(1), 91–108 (2001).
7. M. D. Bisceglie and C. Galdi, "CFAR detection of extended objects in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.* **43**(4), 833–843 (2005).
8. J.-I. Park, S.-H. Park, and K.-T. Kim, "New discrimination features for SAR automatic target recognition," *IEEE Geosci. Remote Sens. Lett.* **10**(3), 476–180 (2013).
9. V. M. Patel, N. Nasrabadi, and R. Chellappa, "Sparsity-motivated automatic target recognition," *Appl. Opt.* **50**(10), 1425–1433 (2011).
10. B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: a strategy employed by V1," *Visual Res.* **37**(23), 3311–3325 (1997).
11. K. Huang and S. Aviyente, "Sparse representation for signal classification," in *Proc. Advanced Neural Information Processing Systems (NIPS)*, Vancouver, Canada (2006).
12. J. Wright et al., "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–228 (2009).
13. Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Top. Signal Process.* **5**(3), 629–640 (2011).
14. J. Thiagarajan et al., "Sparse representation for automatic target classification in SAR images," in *Proc. 4th Int. Symposium on Communications, Control and Signal Processing*, pp. 1–4, IEEE (2010).
15. S. Gao, I. W.-H. Tsang, and L.-T. Chia, "Kernel sparse representation for image classification and face recognition," in *Proc. 11th Eur. Conf. Computer Vision (ECCV'10) Part IV*, pp. 1–14, Springer (2010).
16. J. Yin et al., "Kernel sparse representation based classification," *Neurocomputing* **77**, 120–128 (2012).
17. L. Zhang et al., "Kernel sparse representation-based classifier," *IEEE Trans. Signal Process.* **60**(4), 1684–1695 (2012).
18. L. Zhang, M. Yang, and X. C. Feng, "Sparse representation or collaborative representation: which helps face recognition?," in *Proc. Int. Conf. Computer Vision (ICCV'11)*, Barcelona, Spain (2011).
19. Q. Shi et al., "Is face recognition really a compressive sensing problem?," in *Proc. Int. Conf. Computer Vision (ICCV'11)*, pp. 553–560, IEEE (2011).
20. E. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?," *IEEE Trans. Inf. Theory* **52**(12), 5406–5425 (2006).
21. J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE* **98**(6), 948–958 (2010).

22. S. Lyu, "Mercer kernels for object recognition with local features," in *Proc. Int. Conf. Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 2, pp. 223–229, IEEE (2005).
23. Q. Zhao et al., "Synthetic aperture radar automatic target recognition with three strategies of learning and representation," *Opt. Eng.* **39**(5), 1230–1244 (2000).
24. Y. Chen et al., "Experimental feature-based SAR ATR performance evaluation under different operational conditions," *Proc. SPIE* **6968**, 69680F (2008).
25. J. Cui, J. Gudnason, and M. Brookes, "Radar shadow and superresolution features for automatic recognition of MSTAR targets," in *Proc. Int. Radar Conf.*, pp. 534–539, IEEE (2005).
26. R. Patnaik and D. Casasent, "MINACE filter classification algorithms for ATR using MSTAR data," *Proc. SPIE* **5807**, 100–111 (2005).
27. A. R. Wise, D. Fitzgerald, and T. D. Ross, "The adaptive SAR ATR problem set," *Proc. SPIE* **5427**, 366–375 (2004).

Ganggang Dong received his BEng degree in UAV application engineering from Artillery Academy, Hefei, China, in 2004 and his MAEng degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2012. He is currently working toward his PhD degree in the National University of Defense Technology. His research interests include the applications of compressed sensing and sparse representations.

Biographies for other authors are not available.