

## Article

# SAR Target Recognition via Local Sparse Representation of Multi-Manifold Regularized Low-Rank Approximation

Meiting Yu \*, Ganggang Dong, Haiyan Fan and Gangyao Kuang

College of Electronic Science, National University of Defense Technology, Changsha 410073, China; dongganggang@nudt.edu.cn (G.D.); hy\_fan@yeah.net (H.F.); kuanggangyao@nudt.edu.cn (G.K.)

\* Correspondence: meitingyu@hotmail.com; Tel.: +86-155-7086-4385

Received: 14 December 2017; Accepted: 27 January 2018; Published: 1 February 2018

**Abstract:** The extraction of a valuable set of features and the design of a discriminative classifier are crucial for target recognition in SAR image. Although various features and classifiers have been proposed over the years, target recognition under extended operating conditions (EOCs) is still a challenging problem, e.g., target with configuration variation, different capture orientations, and articulation. To address these problems, this paper presents a new strategy for target recognition. We first propose a low-dimensional representation model via incorporating multi-manifold regularization term into the low-rank matrix factorization framework. Two rules, pairwise similarity and local linearity, are employed for constructing multiple manifold regularization. By alternately optimizing the matrix factorization and manifold selection, the feature representation model can not only acquire the optimal low-rank approximation of original samples, but also capture the intrinsic manifold structure information. Then, to take full advantage of the local structure property of features and further improve the discriminative ability, local sparse representation is proposed for classification. Finally, extensive experiments on moving and stationary target acquisition and recognition (MSTAR) database demonstrate the effectiveness of the proposed strategy, including target recognition under EOCs, as well as the capability of small training size.

**Keywords:** synthetic aperture radar; target recognition; sparse representation; locality constraint; low-rank approximation; regularized manifold

## 1. Introduction

Synthetic aperture radar (SAR) has been widely applied in civilian and military applications due to its capability of providing all-weather, all-day, high-resolution images. Automatic target recognition (ATR) is an essential topic for SAR image interpretation. Generally, target recognition is divided into two parts: feature extraction and classification. Over the past several decades, various features and classifiers [1–3] have been studied; however, target recognition remains a difficult problem. The major challenge is robustness for recognizing targets under extended operating conditions (EOCs) since targets are classified under EOCs in the presence of variations in configuration, articulation, occlusions and different capture orientations.

Feature extraction is essentially a signal processing procedure for reducing the dimensionality of the target image. Thus, the main objective of feature extraction is to find a low-dimensional representation of a SAR image that could distinguishably represent the target. Different features have been explored to characterize the target signal [4,5], such as scattering centre features [6,7], filter bank features [3,8], pattern structure features [9,10] and statistical features [1,11]. Recently, low-rank matrix factorization (LMF), a particularly useful technique for data representation, has been developed for SAR ATR [12,13]. Specifically, LMF aims to find two matrices  $\mathbf{Z}$  and  $\mathbf{H}$  or more

lower-dimensional matrices whose product provides a good low-dimensional approximation to the original matrix  $\mathbf{X}$  such that  $\mathbf{X} \approx \mathbf{Z}\mathbf{H}$ . From a feature extraction perspective,  $\mathbf{H}$  denotes the low-dimensional representation of original samples  $\mathbf{X}$ , and  $\mathbf{Z}$  represents the mapping between  $\mathbf{X}$  and  $\mathbf{H}$ . Truncated singular value decomposition (TSVD) and non-negative matrix factorization (NMF) are two common LMF techniques. Based on these techniques, some methods have been proposed for characterizing SAR images. Cui et al. [12] proposed  $L_{1/2}$ -norm regularized NMF for sparsely representing SAR images. Dang et al. [13] used incremental NMF with an  $L_p$ -norm constraint to further improve performance. Babaei et al. [14] introduced two NMF variants, i.e., variance-constrained NMF and centre map NMF, to describe SAR images in an interactive system. However, these matrix factorization methods fail to discover the intrinsic geometric structure of samples, which is essential for real-world applications, particularly in the area of feature extraction.

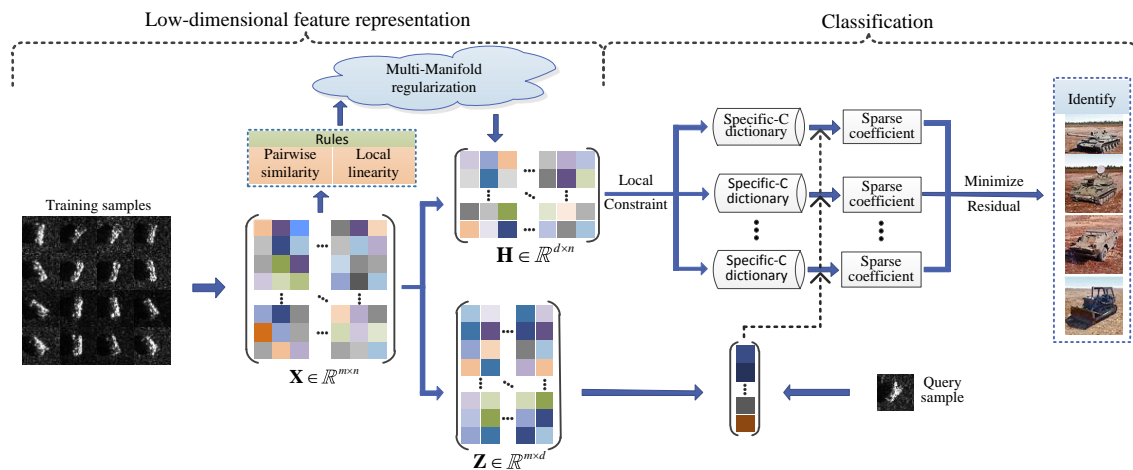
To preserve local geometric structures embedded in the original space, a number of manifold learning approaches have been reported in the literature, such as Laplacian eigenmap (LE) [15], locally linear embedding (LLE) [16], local tangent space alignment (LTSA) [17] and t-distributed stochastic neighbour embedding (t-SNE) [18]. However, these methods are non-parametric techniques, and it is infeasible to directly apply such methods for target recognition. To solve this problem, some works have been devoted to finding parametric extensions of these non-parametric techniques, e.g., locality preserving projection (LPP) [19], neighbourhood preserving embedding (NPE) [20] and parametric supervised t-SNE [10]. It has been shown that the discriminative ability can be enhanced if the intrinsic manifold structure is considered. Motivated by manifold learning methods, manifold regularized LMF techniques have been developed to learn a low-rank approximation that explicitly takes the local geometric structure of samples into account. Cai et al. proposed graph regularized NMF (GNMF) [21], in which a nearest neighbour graph is constructed to model the manifold structures. Zhang et al. [22] demonstrated that the manifold regularized TSVD framework could preserve the nonlinear structure of the sample space more effectively, as the global optimal solution can be reached. However, the selection of the optimal manifold is generally difficult. Moreover, the targets of SAR images consist of multiple variations, such as configuration variation and articulation variation. In particular, the appearance of the same target would vary obviously under different imaging conditions, e.g., the different depression variation. Hence, a single optimal manifold cannot well describe the intrinsic structure of the original space. Although some studies exploited the intrinsic manifold structure by constructing the neighbour graph with different weighting schemes (i.e., 0–1, heat kernel and dot product) [23,24], the representation is relatively weak without considering the property of local neighbours. Consequently, the discriminative ability of the learned features is weak, which further influences the performance of target recognition.

In the classification stage, the classifier makes the final class decision by matching a query sample and the training samples in the specific feature space. The choice of classifier is crucial for achieving a desired classification performance as consensus has evolved that no single feature or classifier is optimal for target recognition. There are many traditional classifiers, such as nearest neighbour (NN) [25], support vector machine (SVM) [26] and adaptive boosting (AdaBoost) [27]. In recent years, the sparse representation-based classification (SRC) [28] technique has achieved some impressive performances on face recognition. Jayaraman et al. [29] reported that the classification of SAR images based on SRC is equivalent to finding the manifold that is closest to the query image. The sparse signal representation can be applied in SAR target classification, as verified in [8,30].

Inspired by the aforementioned works, this paper proposes a new strategy, named local sparse representation of multi-manifold regularized low-rank approximation (MLA-LSR), for target recognition in SAR images. To obtain the low-dimensional representation of SAR images, multi-manifold regularized low-rank approximation (MLA) is proposed by incorporating a multiple manifold regularization term into the TSVD framework. The manifold selection and low-rank approximation are alternately optimized. Specifically, two construction rules, pairwise similarity and local linearity, are utilized for the multi-manifold regularization term. Rather than the traditional

norm graph, the pairwise similarity is described by the joint probabilities, which could capture much of the local structure of the sample space. The local linearity rule considers the local structure property of the samples. Hence, the learned features can well preserve the intrinsic geometric structures and discriminating structures. To predict the identity, local sparse representation (LSR) based on MLA is designed for classification. In contrast to the traditional SRC, LSR constructs the dictionary by local, rather than all, samples in the training set. On the one hand, the locality constraint in LSR takes full advantage of the local structure property of samples obtained by MLA. On the other hand, locality is more essential than sparsity [31]. To optimize the sparse representation, the query sample is represented as a linear combination of class-specific galleries, by which the discriminative ability can be promoted. Finally, the identity of the query sample is reached corresponding to the class with the minimum reconstruction error. A flowchart of the proposed strategy is shown in Figure 1.

The reminder of this paper is organized as follows. In Section 2, a brief review of related works is given, including the LMF and manifold regularized LMF. In Section 3, a multi-manifold regularized low-rank approximation is proposed to learn the low-dimensional representation of SAR images. In Section 4, a local sparse representation for target classification is presented. The effectiveness of the proposed strategy is verified in Section 5. Conclusions are drawn in Section 6.



**Figure 1.** Overview of the target recognition procedure based on the proposed strategy. The proposed strategy consists of two phases: low-dimensional feature representation and classification. The first phase uses a multi-manifold regularized low-rank approximation to obtain the low-dimensional feature representation of training samples  $\mathbf{H}$  and mapping matrix  $\mathbf{Z}$ . The feature of the query sample can be obtained by  $\mathbf{Z}$ . The second phase uses the local sparse representation for classification. The query sample is represented as a linear combination of class-specific dictionaries. According to the representation coefficients, the identity of the query sample is reached by evaluating the class of samples, which could result in the minimal reconstruction residual.

## 2. Review of LMF and Manifold Regularized LMF

### 2.1. Low-Rank Matrix Factorization

Low-rank matrix factorization, which is an important data analysis method, has attracted considerable attention for feature extraction [14], data compression [32] and clustering [33]. NMF and TSVD are two specific LMF techniques. Given a data matrix  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$  with  $n$  samples and  $m$  dimensions, NMF aims to find two non-negative matrices to approximate the original matrix. The decomposition of matrix  $\mathbf{X}$  can be mathematically formulated as

$$\min_{\mathbf{Z} \geq 0, \mathbf{H} \geq 0} \|\mathbf{X} - \mathbf{ZH}\|_F^2 \quad (1)$$

where  $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_d] \in \mathbb{R}^{m \times d}$  and  $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n] \in \mathbb{R}^{d \times n}$  are low-rank matrices, and the Frobenius norm is a well-established proxy to impose similarity between  $\mathbf{X}$  and  $\mathbf{ZH}$ . If we view  $\mathbf{Z}$  as the mapping between the low-dimensional representation  $\mathbf{H}$  and the original samples  $\mathbf{X}$ , then  $\mathbf{h}_i \in \mathbf{H}$  represents the low-dimensional representation of sample  $\mathbf{x}_i$ .

In contrast to NMF, the factorized matrix  $\mathbf{Z}$  is employed with an orthogonality constraint in TSVD. The optimization problem of TSVD is defined as

$$\min_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}_d, \mathbf{H}} \|\mathbf{X} - \mathbf{ZH}\|_F^2 \quad (2)$$

It has been widely applied in data representation because TSVD achieves the best approximation in the Frobenius norm. The two aforementioned optimization problems are generally solved by alternating iterative methods.

## 2.2. Manifold Regularized Low-Rank Matrix Factorization

The techniques NMF and TSVD both fail to preserve the intrinsic geometric structure of data, but it is crucial for actual applications. To overcome this issue, some researchers have incorporated the idea of manifold learning into the LMF framework. Generally, these manifold learning methods are based on the assumption that if two samples are close in the intrinsic manifold of the sample space, then the low-dimensional representations of these two samples are close to each other [15,16]. To preserve the neighbourhood relationships, GNMF is proposed by adding the manifold regularization term to the NMF model. Specifically, a nearest neighbour graph based on heat kernel weight  $\mathbf{W}_{ij}$  is constructed to model the manifold structure. The manifold regularization term can be formulated as

$$\begin{aligned} \frac{1}{2} \sum_{i,j} \|\mathbf{h}_i - \mathbf{h}_j\|^2 \mathbf{W}_{ij} &= \sum_{i=1}^n \mathbf{h}_i^T \mathbf{h}_j \mathbf{D}_{ii} - \sum_{i,j=1}^n \mathbf{h}_i^T \mathbf{h}_j \mathbf{W}_{ij} \\ &= \text{tr}(\mathbf{H} \mathbf{D} \mathbf{H}^T) - \text{tr}(\mathbf{H} \mathbf{W} \mathbf{H}^T) \\ &= \text{tr}(\mathbf{H} \mathbf{L} \mathbf{H}^T) \end{aligned} \quad (3)$$

If  $\mathbf{x}_i \in N_k(\mathbf{x}_j)$  or  $\mathbf{x}_j \in N_k(\mathbf{x}_i)$ , then  $\mathbf{W}_{ij} = \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma})$ ; otherwise,  $\mathbf{W}_{ij} = 0$ , where  $N_k(\mathbf{x}_j)$  denotes the set of  $k$  nearest neighbours of  $\mathbf{x}_j$ ,  $\sigma$  is a constant value,  $\mathbf{D}_{ii} = \sum_j \mathbf{W}_{ij}$ , and  $\mathbf{L} \in \mathbb{R}^{n \times n}$  is the graph Laplacian defined as  $\mathbf{L} = \mathbf{D} - \mathbf{W}$ . GNMF utilizes the graph Laplacian to represent the characteristics of the manifold. The GNMF is formulated as

$$\min_{\mathbf{Z} \geq 0, \mathbf{H} \geq 0} \|\mathbf{X} - \mathbf{ZH}\|_F^2 + \lambda \text{tr}(\mathbf{H} \mathbf{L} \mathbf{H}^T) \quad (4)$$

where  $\lambda$  is a regularization parameter. The optimization problem is solved by an alternating iterative method with a multiplicative rule. Although GNMF has shown good performance on low-dimensional representations of data, it has been verified that the optimal solutions cannot be obtained [22]. Afterwards, a new model, MMF, was proposed to incorporate the manifold regularization term into the TSVD framework. The performance of MMF is superior to that of GNMF. This is because NMF retains the neighbourhood relationship, and the optimal solutions can be reached. The optimization problem of MMF is written as

$$\min_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}_d, \mathbf{H}} \|\mathbf{X} - \mathbf{ZH}\|_F^2 + \lambda \text{tr}(\mathbf{H} \mathbf{L} \mathbf{H}^T) \quad (5)$$

To obtain the optimal solution of Equation (5), a direct algorithm or alternating iterative algorithm can be employed to solve the optimization problem.

### 3. Feature Extraction via Multi-Manifold Regularized Low-Rank Approximation

By incorporating the manifold regularization term into the LMF framework, the manifold structure of the original space is preserved in the low-dimensional representation. However, the selection of the optimal manifold structure is generally difficult. Although some scholars exploit an approximate manifold structure through a union of different weighting schemes (i.e., 0–1, heat kernel and dot product), the representation of the manifold structure is relatively weak without taking the property of local neighbours into account. To address these problems, a new model, named multi-manifold regularized low-rank approximation (MLA), is proposed in this section.

#### 3.1. Multi-Manifold Regularization Term

It is known that SAR images of a given class lie in a manifold, whose dimensions are much lower than the actual dimensions of the image [34]. Without loss of generality, suppose that  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$  are sampled from an underlying manifold  $\mathcal{M}$ .  $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n] \in \mathbb{R}^{d \times n}$  ( $d \ll m$ ) denotes their low-dimensional representations. To obtain the discriminative low-dimensional representations, the intrinsic manifold structure of samples  $\mathbf{X}$  should be preserved by  $\mathbf{H}$ . Recently, Geng et al. developed an approach for the automatic approximation of the intrinsic manifold with an ensemble manifold regularizer (EMR) [35]. It assumes that the intrinsic manifold approximately lies in the convex hull of some pre-given manifold candidates. Therefore, the intrinsic manifold could be approximated by the linear combination of these candidate manifolds. Motivated by EMR, we select the optimal manifold by fusing multiple manifolds.

$$\mathbf{\Psi} = \sum_{i=1}^r \eta_i \mathbf{\Psi}_i; \quad \text{subject to } \sum_{i=1}^r \eta_i = 1, \eta_i \geq 0 \quad (6)$$

where  $\mathbf{\Psi}$  is an affinity matrix that characterizes the manifold of samples and  $\boldsymbol{\eta} = [\eta_1, \dots, \eta_r]$  is a weight vector for basis manifolds. Since the manifold learning problem can be uniformly written as  $\min_{\mathbf{H}} \text{tr}(\mathbf{H}\mathbf{\Psi}\mathbf{H}^T)$ , the intrinsic manifold structure of the sample space can be preserved by optimizing the following objective function:

$$\min_{\mathbf{H}} \text{tr} \sum_{i=1}^r \eta_i (\mathbf{H}\mathbf{\Psi}_i\mathbf{H}^T); \quad \text{subject to } \sum_{i=1}^r \eta_i = 1, \eta_i \geq 0 \quad (7)$$

With the development of manifold learning theory, the affinity matrix  $\mathbf{\Psi}_i$  of the basis manifold can be produced with many manifold learning techniques, including, but not limited to, LE, LLE, and TLSA. Hence, the selection of a suitable basis manifold is important for acquiring the optimal manifold of SAR images. Generally, in the manifold learning techniques, the construction of the affinity matrix typically consists of two rules: pairwise similarity and local linearity. At present, some works [23,24] about the multiple manifold regularizer only consider the construction of basis manifolds through pairwise similarity, which provides a relatively weak representation of the local manifold structure without the local linearity. In the following, we model the basis manifolds taking these two rules into account, which allows the learned low-dimensional representation to preserve the intrinsic and discriminating structure of the sample space.

(1) Based on the pairwise similarity. The manifold learning naturally assumes that if two samples  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are close in the intrinsic geometry of the sample distribution, then the low-dimensional representations of two samples,  $\mathbf{h}_i$  and  $\mathbf{h}_j$ , are also close to each other. This assumption can be employed in terms of similarity. In GNMF and MMF, pairwise similarity measurements based on LE are performed to select  $k$  neighbourhoods for each sample. The heat kernel weight scheme is commonly used for the similarity measurement of SAR images. According to previous works [10], pairwise similarity defined in t-SNE with probability distribution will capture much of the local structure of the sample space in SAR images because the relationships of the samples are calculated by

integrating over all the paths of the neighbourhood graph. The pairwise similarity  $S_{ij}$  is measured by means of Gaussian density and re-normalized

$$S_{ij} = \frac{\exp(\|\mathbf{x}_i - \mathbf{x}_j\|_2^2 / 2\sigma_i)}{\sum_{k \neq l} \exp(-\|\mathbf{x}_k - \mathbf{x}_l\|_2^2 / 2\sigma_i)} \quad (8)$$

where  $S_{ij} = S_{ji}$  and  $S_{ii} = 0$ . The  $\sigma_i$  is the variance parameter of Gaussian that is centered over each sample  $\mathbf{x}_i$ . Since the density of the sample is likely to vary, it is unreasonable to use a single value of  $\sigma_i$  for all the samples. Let  $\mathbf{P}_i$  be the probability distribution over all other given sample  $\mathbf{x}_i$ . To obtain a suitable value of  $\sigma_i$ , a binary search is performed with a perplexity. The perplexity is written as

$$\text{Perp}(\mathbf{x}_i) = 2^{-\sum_j \mathbf{P}_i \log_2 \mathbf{P}_i} \quad (9)$$

where the perplexity can be taken as a smooth measure of the effective number of neighbours [18].

With the defined pairwise similarity matrix  $\mathbf{S} \in \mathbb{R}^{n \times n}$ , the basis manifold regularization term is used for measuring the smoothness of the low-dimensional representation.

$$\frac{1}{2} \sum_{i,j} \|\mathbf{h}_i - \mathbf{h}_j\|_2^2 S_{ij} = \text{tr}(\mathbf{H}(\mathbf{M} - \mathbf{S})\mathbf{H}^T) \quad (10)$$

where  $\mathbf{M}_{ii} = \sum_j S_{ij}$ , and the affinity matrix based on pairwise similarity is defined as  $\mathbf{M} - \mathbf{S}$ .

(2) Based on the local linearity. Although a local manifold structure is artificially designed by exploiting the pairwise similarity, the limitation is obvious, namely, the relatively weak representation of the local manifold without considering the property of local neighbours [16]. Hence, to describe the manifold structure more accurately, the LLE is utilized to construct the basis manifold because the property of local neighbours is fully taken into account in LLE. Specifically, LLE exploits the fact that a nonlinear high-dimensional sample can be linearly reconstructed by its neighbour samples. The linear reconstruction coefficient  $\mathbf{A} \in \mathbb{R}^{n \times n}$  indicates the local manifold structure that should be retained in the low-dimensional space.  $\mathbf{A}$  is obtained via the minimization

$$\arg \min_{\mathbf{A}} \sum_{i=1}^n \|\mathbf{x}_i - \sum_{j \in N_k(\mathbf{x}_i)} \mathbf{A}_{ij} \mathbf{x}_j\|_2^2; \quad \text{subject to} \quad \sum_{j \in N_k(\mathbf{x}_i)} \mathbf{A}_{ij} = 1, \quad \forall i \quad (11)$$

where  $\mathbf{A}_{ij}$  denotes the reconstruction weight between  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , if  $j \in N_k(\mathbf{x}_i)$ ; otherwise,  $\mathbf{A}_{ij} = 0$ .

The smoothness of the low-dimensional representation is measured through the basis manifold regularization term as follows:

$$\sum_{i=1}^n \|\mathbf{h}_i - \sum \mathbf{A}_{ij} \mathbf{h}_j\|_2^2 = \text{tr}(\mathbf{H}(\mathbf{I} - \mathbf{A})(\mathbf{I} - \mathbf{A})^T \mathbf{H}^T) \quad (12)$$

where  $\mathbf{I} \in \mathbb{R}^{n \times n}$  is a unit matrix, and the affinity matrix based on local linearity is defined as  $(\mathbf{I} - \mathbf{A})^T (\mathbf{I} - \mathbf{A})$ .

A large amount of basis manifolds can be obtained by the different parameters for the construction of affinity matrices in Equations (10) and (12). To avoid over-fitting to the single manifold in the manifold regularization term in Equation (7), the multi-manifold regularization term can be rewritten as

$$\min_{\mathbf{H}, \boldsymbol{\eta}} \text{tr} \sum_{i=1}^r \eta_i (\mathbf{H} \boldsymbol{\Psi}_i \mathbf{H}^T) + \lambda \|\boldsymbol{\eta}\|^2; \quad \text{subject to} \quad \sum_{i=1}^r \eta_i = 1, \quad \eta_i \geq 0 \quad (13)$$



### 3.2. Multi-Manifold Regularized Low-Rank Approximation

Compared with the NMF framework, the TSVD framework acquires a smaller approximation error in the Frobenius norm because TSVD relaxes the non-negativity constraint to the orthogonal constraint for the factor  $\mathbf{Z}$ . In our model, a low-rank matrix factorization should preserve the intrinsic manifold structure to the greatest extent possible while the optimal approximation to the sample matrix is also reached. Therefore, we incorporate the multi-manifold regularization term into the TSVD framework, which is formulated as

$$\begin{aligned} \min_{\mathbf{Z}, \mathbf{H}, \boldsymbol{\eta}} \|\mathbf{X} - \mathbf{ZH}\|_F^2 + \lambda_1 \operatorname{tr} \sum_{i=1}^r \eta_i (\mathbf{H} \boldsymbol{\Psi}_i \mathbf{H}^T) + \lambda_2 \|\boldsymbol{\eta}\|^2 \\ \text{subject to } \mathbf{Z}^T \mathbf{Z} = \mathbf{I}_d, \quad \sum_{i=1}^r \eta_i = 1, \quad \eta_i \geq 0 \end{aligned} \quad (14)$$

The objective function, Equation (14), is not convex with respect to  $\mathbf{Z}$ ,  $\mathbf{H}$  and  $\boldsymbol{\eta}$ . Therefore, it is unrealistic to find the optimal solution by the simultaneous optimization of all the variables. To address this problem, the technique of alternating iterations is adopted. The variables are optimized and updated alternately with other variables fixed. In this way, the objective function, Equation (14), can be broken down into two sub-optimization problems as follows.

(1) Update rule for the mapping matrix  $\mathbf{Z}$  and low-dimensional representation matrix  $\mathbf{H}$ . We fix the weight vector  $\boldsymbol{\eta}$  and minimize the cost function with respect to  $(\mathbf{Z}, \mathbf{H})$ . That is, the optimization problem in Equation (14) is rewritten as

$$\begin{aligned} \min_{\mathbf{Z}, \mathbf{H}} \|\mathbf{X} - \mathbf{ZH}\|_F^2 + \lambda_1 \operatorname{tr}(\mathbf{H} \boldsymbol{\Psi} \mathbf{H}^T) \\ \text{subject to } \mathbf{Z}^T \mathbf{Z} = \mathbf{I}_d \end{aligned} \quad (15)$$

This optimization problem shares the same model as MMF. Hence, the global optimal solution of this problem can be calculated directly or iteratively according to MMF. The detailed solving procedure can be found in [22]. We use the iterative method for this problem. The specific iterative steps for updating  $\mathbf{Z}$  and  $\mathbf{H}$  can be given as

$$\begin{aligned} \mathbf{Z} &= \mathbf{U} \mathbf{V}^T \\ \mathbf{H} &= \mathbf{Z}^T \mathbf{X} (\mathbf{I} + \lambda_1 \boldsymbol{\Psi})^{-1} \end{aligned} \quad (16)$$

where  $\mathbf{X} \mathbf{H}^T = \mathbf{U} \mathbf{D} \mathbf{V}^T$  is the SVD of  $\mathbf{X} \mathbf{H}^T$ , and  $\mathbf{D}$  is a diagonal matrix of positive diagonal entries. Note that the updating of  $\mathbf{H}$  can be transformed to solve the linear systems  $\mathbf{H}(\mathbf{I} + \lambda_1 \boldsymbol{\Psi}) = \mathbf{Z}^T \mathbf{X}$  with an iterative method, i.e., preconditioned conjugate gradient method.

(2) Update rule for the weight vector  $\boldsymbol{\eta}$ . By fixing the matrices  $\mathbf{Z}$  and  $\mathbf{H}$ , the optimization problem in Equation (14) can be transformed as follows

$$\begin{aligned} \min_{\boldsymbol{\eta}} \lambda_1 \operatorname{tr} \sum_{i=1}^r \eta_i (\mathbf{H} \boldsymbol{\Psi}_i \mathbf{H}^T) + \lambda_2 \|\boldsymbol{\eta}\|^2 \\ \text{subject to } \sum_{i=1}^r \eta_i = 1, \quad \eta_i \geq 0 \end{aligned} \quad (17)$$

This problem is a constrained quadric optimization problem, which can be solved easily via the Lagrangian multiplier method. When the parameters are renewed, the next round of update is pursued. The iteration is terminated if the convergence condition is met. The procedure of multi-manifold regularized low-rank approximation can be summarized as in Algorithm 1.

**Algorithm 1** Multi-manifold regularized Low-rank Approximation (MLA)

**Input:** training sample matrix  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ ; regularization parameters  $\lambda_1$  and  $\lambda_2$ ;  $r$  candidate basis manifolds; maximum iterations  $T$  and threshold  $\xi$ .

**Output:** mapping matrix  $\mathbf{Z}$  and low-dimensional representation matrix  $\mathbf{H}$ .

**Initialization:**

Initialize  $(\mathbf{Z}, \mathbf{H})$  with TSVD;

Initialize  $\eta = [1/r, 1/r, \dots, 1/r]$ .

**Repeat** for  $t = 1$  to  $T$  do

Update mapping matrix  $\mathbf{Z}$  and low-dimensional representation matrix  $\mathbf{H}$  by Equation (16).

Update  $\eta^t$  according to Equation (17).

Calculate objective function  $O^t$  by Equation (14), if  $|O^t - O^{t-1}| < \xi$ , break.

**End**

**4. Target Classification via Local Sparse Representation of MLA**

Denote  $n_i$  training samples from the  $i$ th class as  $\mathbf{X}_i = [\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,n_i}] \in \mathbb{R}^{m \times n_i}$ . Let  $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_C] \in \mathbb{R}^{d \times n}$  ( $d \ll m$ ) be the low-dimensional feature representation of the training samples  $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n] \in \mathbb{R}^{m \times n}$  obtained by MLA, where  $C$  is the number of the classes,  $d$  is the dimension of the feature vector, and  $n = \sum_{j=1}^C n_j$ . Given a query sample  $\mathbf{y} \in \mathbb{R}^m$ , the low-dimensional feature representation  $\mathbf{T} \in \mathbb{R}^d$  of the query sample is acquired as follows:

$$\mathbf{T} = \mathbf{Z}^T \mathbf{y} \quad (18)$$

where  $\mathbf{Z} \in \mathbb{R}^{m \times d}$  is the projection matrix in MLA.

Recently, sparse-representation-based classification has received considerable attention for target classification in SAR images. It tends to find the sparse representation of the query sample in a dictionary composed of all training samples, and then it assigns the query sample to the class with the minimum reconstruction error. In contrast to the traditional SRC, we propose a novel sparse representation-based method to calculate the representation coefficients by local, rather than all, samples in the training samples. The main reason is that the low-dimensional feature obtained by MLA preserves the local structure property of the sample space, and a query sample can be approximately linearly represented by their local neighbours in the feature space. In addition, locality is more essential than sparsity, as the locality must lead to sparsity but not necessarily vice versa [19]. Specifically, the sparse representation with locality constraint is defined as

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \|\mathbf{Z}^T \mathbf{y} - \mathbf{H} \boldsymbol{\alpha}\|_2^2 + \gamma \|\mathbf{p} \odot \boldsymbol{\alpha}\|_2^2 \\ \text{subject to} \quad & \sum_{j=1}^{n_i} \alpha_{ij} = 1, \quad \forall i \end{aligned} \quad (19)$$

where  $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_C] \in \mathbb{R}^n$ ,  $\alpha_{ij} \in \boldsymbol{\alpha}_i$  are the representation coefficients,  $\gamma$  is a regularization parameter, and  $\odot$  denotes the componentwise multiplication. The constraint  $\sum_{j=1}^{n_i} \alpha_{ij} = 1$  aims to follow the shift-invariant requirements of the reconstruction.  $\mathbf{p} \in \mathbb{R}^n$  is the locality adaptor, which is used to measure the similarities of features between training samples and the query sample. Specifically,

$$\mathbf{p} = [\exp(\|\mathbf{Z}^T \mathbf{y} - \mathbf{h}_{1,1}\|_2) / \delta, \exp(\|\mathbf{Z}^T \mathbf{y} - \mathbf{h}_{1,2}\|_2) / \delta, \dots, \exp(\|\mathbf{Z}^T \mathbf{y} - \mathbf{h}_{c,n_c}\|_2) / \delta]^T \quad (20)$$

The discriminative representation of the query sample can be achieved by Equation (12) when the subspaces spanned by different classes are less correlated. However, in real SAR target recognition scenarios, SAR images of different targets collected under the same conditions show higher similarities than those of the same target with different conditions. Hence, the targets from different classes may



obtain similar features, which will result in the representation of the query sample by training samples from different classes. To avoid the limitation of unsupervised optimization of sparse representation, we introduce the class labels into the model in Equation (19). Specifically, the query sample is sparsely represented by each class independently.

For the  $k$ th target class, suppose that there are  $n_k$  training samples; we compute the low-dimensional representation features  $\mathbf{H}_i = [\mathbf{h}_{k,1}, \mathbf{h}_{k,2}, \dots, \mathbf{h}_{k,n_k}] \in \mathbb{R}^{d \times n_k}$ . It is utilized to learn the representation coefficients

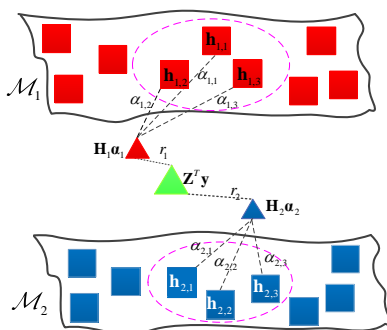
$$\begin{aligned} \min_{\alpha_k} \quad & \|\mathbf{Z}^T \mathbf{y} - \mathbf{H}_i \alpha_k\|_2^2 + \gamma \|\mathbf{p}_k \odot \alpha_k\|_2^2 \\ \text{subject to} \quad & \sum_{j=1}^{n_k} \alpha_{kj} = 1 \end{aligned} \quad (21)$$

where  $\alpha_k = [\alpha_{k1}, \alpha_{k2}, \dots, \alpha_{kn_k}] \in \mathbb{R}^{n_k}$  are the representation coefficients with regards to the  $k$ th target class. The optimal problem can easily be solved via the Lagrangian multiplier method.

According to the optimal representation coefficients  $\alpha$ , the reconstruction error with respect to each class is computed. The inference is reached by evaluating the class with the minimal reconstruction error

$$\text{identity}(\mathbf{y}) = \arg \min_i \|\mathbf{Z}^T \mathbf{y} - \mathbf{H}_i \alpha_i\|, \quad i = 1, 2, \dots, c \quad (22)$$

In fact, the low-dimensional representation features of the  $k$ th target class  $\mathbf{H}_k$  lie in the manifold  $\mathcal{M}_k$ , and the  $\|\mathbf{Z}^T \mathbf{y} - \mathbf{H}_k \alpha_k\|$  can be regarded as the distance between the feature of the query sample  $\mathbf{Z}^T \mathbf{y}$  and the manifold  $\mathcal{M}_k$ . Hence, the decision rule is actually to seek the minimum sample-to-manifold distance. Figure 2 illustrates the basic idea of computing the distance between the sample and manifold.



**Figure 2.** Illustration of LSR by the view of the sample-manifold distance.

The implementation flow of target classification for the proposed strategy MLA-LSR is summarized in Algorithm 2.

**Algorithm 2** Target recognition via MLA-LSR

**Input:** training sample matrix  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$  for  $C$  classes class labels of training samples; query sample  $\mathbf{y}$ .

**Output:** Identification of  $\mathbf{y}$ .

**Step 1:** Learn mapping matrix  $\mathbf{Z}$  and low-dimensional representation matrix  $\mathbf{H}$  for training samples following the detailed description in Algorithm 1.

**Step 2:** Obtain the low-dimensional feature representation of testing samples by Equation (18).

**Step 3:** Solve the sparse representation of query sample over the locality-constrained and specific-class dictionary according to Equation (21).

**Step 4:** Predict the identity by calculating the class that could produce the minimal reconstruction error by Equation (22).

## 5. Experiments and Discussion

In this section, we perform experiments on the publicly available MSTAR database, a gallery collected with X-band and  $1 \times 1$  foot resolution. The target images cover the full  $0^\circ \sim 359^\circ$  aspect angles at different depression angles  $\{15^\circ, 17^\circ, 30^\circ, 45^\circ\}$ . The target images are approximately  $128 \times 128$  pixels. To exclude the redundant background, the target images are cropped  $64 \times 64$  pixels from the centre of the images. To validate the efficacy of the proposed strategy, we first carry out a series of fundamental experiments, comparing to various feature models and some popular classification models. Then, we demonstrate the effectiveness of the proposed strategy to configuration variation, depression and articulation variation, small sample size, and recognition with ten-class targets. In our experiments, we empirically set the regularization parameters  $\lambda_1 = 10$  and  $\lambda_2 = 5$  in MLA and  $\gamma = 0.1$  in LSR. The typical parameter values of the basis manifolds are between 30 and 50 for perplexity in the pairwise similarity rule and between 6 and 12 for  $k$  in the local linearity rule.

### 5.1. Fundamental Verifications

This paper develops a new strategy for target recognition in SAR images. To obtain the intrinsic low-dimensional representation of a SAR image, the multiple manifold regularization term is incorporated into the TSVD framework. Then, the local sparse representation is employed to make the decision, by which the recognition performance can be further improved. To verify the proposed strategy, we first perform two sets of experiments for different low-dimensional feature representation models and different classification models. Four military vehicles, BMP2, BTR60, T72 and T62, are used for conducting the experiments. The type and number of samples in different depression views are listed in Table 1. The items in parentheses represent different configurations with structural modifications, which will be explained in detail in Section 5.2.1.

**Table 1.** The type and number of samples in different depression views.

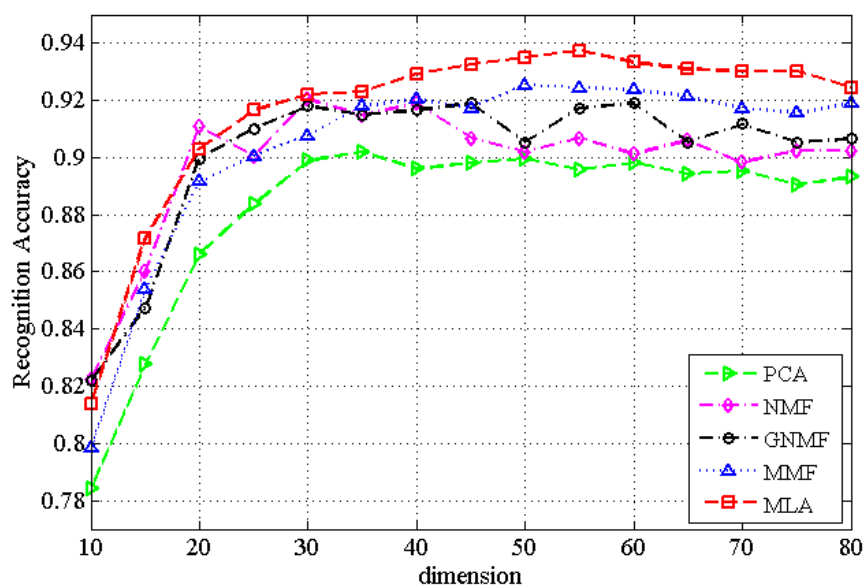
Type	Depr.	BMP2	BTR60	T72	T62	Total
Training	$17^\circ$	233(SN_9563)	256	233(SN_132)	299	1020
Testing	$15^\circ$	196(SN_9566) 196(SN_c21)	195	195(SN_812) 195(SN_s7)	273	1246

#### 5.1.1. The Comparison of Feature Representation Model

To assess the performance of the proposed feature representation model, the recognition performance of the proposed method MLA is compared with PCA [36], NMF [37], GNMF [21] and MMF [22]. We set the regularization parameters in GNMF and MMF as  $\lambda = 10$  and use the classical k-nearest-neighbour graph with  $k = 12$ . In previous reports [12,21,22], these baseline methods are always combined with a linear SVM classifier for classification. To ensure a fair comparison,

the recognition performances for different representation models, in conjunction with a linear SVM classifier (with a cost function parameter  $c = 6$ ), are shown in Figure 3.

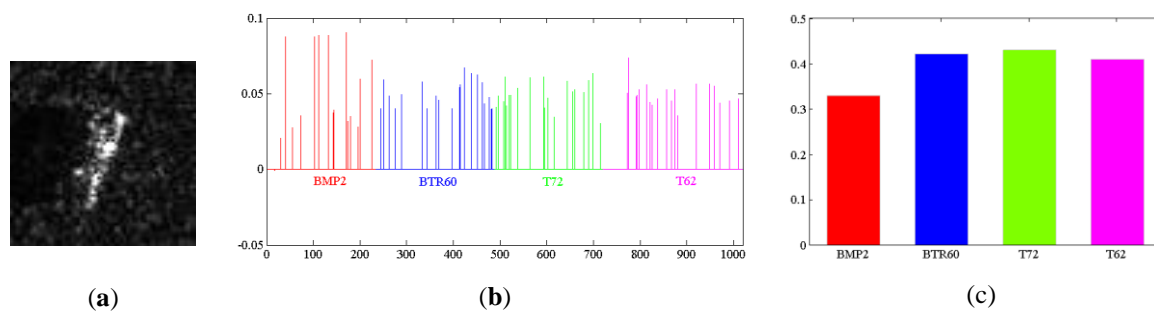
Figure 3 shows the recognition performances versus the change in the feature space dimension. When the dimension of the feature space exceeds 30D, the best performance of MLA is superior to the best performance of the baseline methods at each individual feature dimension. Specifically, the recognition performances of PCA and NMF are inferior to those of all the other methods. The recognition performances of GNMF and MMF are significantly better since they preserve the geometric structure of the sample space in the low-dimensional representation. By contrast, the proposed method MLA does a much better job in preserving the intrinsic geometric and discriminating structure of the sample space via multiple manifold regularization. Moreover, the performances of NMF and GNMF strongly depend on a good choice of the optimal dimension. In contrast, the other three methods retain relatively stable results above a certain feature dimension since the best low-rank approximation  $\mathbf{H}$  for PCA, MMF and MLA can be obtained via using the orthogonality constraint for mapping matrix  $\mathbf{Z}$ . The above experimental results confirm the effectiveness of the proposed low-dimensional representation model for SAR images.



**Figure 3.** The recognition rates of different feature representation models across the range of the dimension of feature space.

### 5.1.2. The Comparison of Classification Model

To improve the classification performance, this paper employs the local sparse representation to make a decision. Since the manifold of given classes that is closest to the query sample could be found by the sparse representation [29], SRC has been applied to separate features of different targets. In contrast to SRC, in this paper, a local sparse representation model, LSR, is proposed to limit the feasible set of representations. In this way, the sparsity can be harnessed, and the local structure of features can be fully utilized. Meanwhile, a class-specific model is introduced to supervise the optimization of the sparse representation, which can improve the discriminative ability. The class of the query sample is assigned to the object class which results in the minimum reconstruction errors. A practical example for LSR can be found in Figure 4. To validate the classification model LSR, a set of comparative studies based on MLA with 100D are employed, where SVM and SRC are compared with LSR. The experimental results are presented in Table 2.



**Figure 4.** A practical example for LSR: (a) a test sample is randomly chosen from BMP2; (b) the representation coefficients for the test sample are plotted together with the four classes—in fact, the representation coefficients with each class are obtained independently; and (c) the residuals of the test sample with respect to the four classes.

**Table 2.** The comparison of classification models based on the feature representation model MLA.

Method	SVM	SRC	LSR
MLA	0.9350	0.9390	0.9582

As shown in Figure 4, the representation coefficients are sparse, and the smallest residual corresponds to the correct class (BMP2). In Table 2, we can see that the proposed LSR model achieves the best performance. Specifically, the recognition accuracy of LSR is 0.9582, which is 1.72 percent better than SRC and 2.32 percent better than SVM. This result reveals that LSR is much more effective than the other two classification models for separating features of different target classes.

### 5.1.3. Influence of the Related Factors on Performance

(1) The validation of multi-manifold regularization term. In the proposed low-dimensional representation model, MLA, the multiple manifold regularization term is used for preserving the intrinsic manifold structure. To verify the effectiveness of the multi-manifold regularization term in MLA, we perform an experiment and compare the recognition performance with different single-manifold regularization terms. The recognition results are shown in Table 3.

**Table 3.** The recognition accuracies obtained using different regularizations.

Regularization	Single Manifold <sub>1</sub>	Single Manifold <sub>2</sub>	Multi-Manifold
Accuracy	0.9398	0.9486	0.9582

In Table 3, single manifold<sub>1</sub> and single manifold<sub>2</sub> denote that only the pairwise similarity rule and local linear rule are utilized, respectively, whereas both rules are used for multi-manifold. As shown in Table 3, the low-dimensional representation model with the multi-manifold regularization term achieves better recognition accuracy than those of single-manifold regularization terms. The experimental results reveal that each single-manifold regularization could describe some underlying manifold structure. Specifically, the single manifold<sub>1</sub> preserves local neighbor structure by exploiting approximately pairwise similarities. The single manifold<sub>2</sub> represents the local manifold structure by exploiting the local symmetries of linear reconstruction between each sample and its neighbors. However, the performance can be improved by aggregating the manifold structures via multiple manifold regularization.

(2) Setting of the regularized parameters. In the proposed MLA-LSR strategy, there are two essential parameters, i.e., the multiple manifold regularization parameter  $\lambda_1$  and local constraint regularization parameter  $\gamma$ , which are both empirically set in the experiments. To obtain the

approximate optimal values for  $\lambda_1$  and  $\gamma$ , we evaluate the influence of  $\lambda_1$  and  $\gamma$  for target recognition. The recognition performance is shown at different values of parameters in Figure 5.

In the experiments, we choose the values of  $\lambda_1$  and  $\gamma$  in the sets of  $\{0.1, 0.3, 1, 3, 10, 30, 100\}$  and  $\{0.01, 0.05, 0.1, 0.5, 1\}$ , respectively. As shown, better recognition performances can be obtained when the value of  $\lambda_1$  ranges from 1 to 100. Either small or large  $\lambda_1$  can degrade the recognition performance. This result is because a small value for  $\lambda_1$  cannot well reflect the manifold structure in the low-dimensional feature representation, whereas a large value for  $\lambda_1$  is likely to cause a potential overfitting. For the parameter  $\gamma$ , the performance is affected by the selection of the parameter  $\lambda_1$  to some extent. For example, if  $\lambda_1$  ranges from 1 to 30, the recognition performance is robust to the change of  $\gamma$  from 0.01 to 1. By amount of the experimental analysis, the recognition performance of the proposed method is relatively robust to changes in the  $\lambda_1$  and  $\gamma$  parameters, when the the value of  $\lambda_1$  is between 1 and 50 and the value of  $\gamma$  is between 0.05 to 0.5. Hence, we empirically set the same regularization parameter  $\lambda_1 = 10$  and  $\gamma = 0.1$  in the all of our experiments.

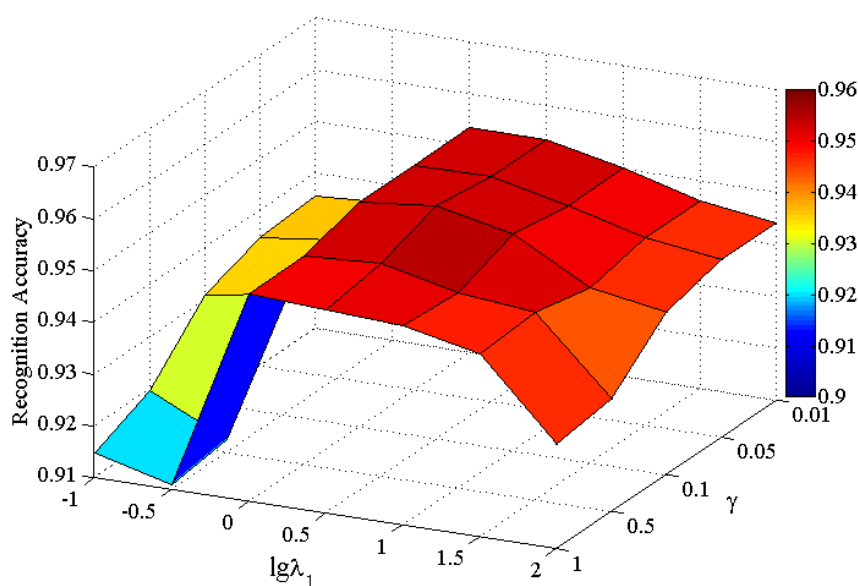


Figure 5. The recognition accuracy across the range of the parameters.

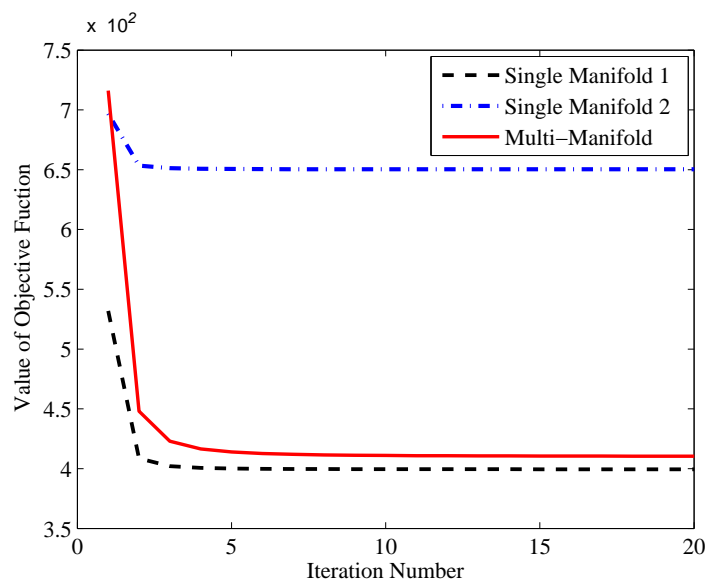
#### 5.1.4. Convergence Analysis

The technique of alternating iterations is used for minimizing the objective function of MLA. Thus, it is necessary to analyse its convergence. To evaluate the convergence of MLA, the experiment is performed with four military vehicles, BMP2, BTR60, T72 and T62. The type and number of samples can be found in Table 1. Figure 6 shows the convergence curves of different manifold regularization. It displays the changes of the value of the objective function with the number of iterations. As shown, in general, the objective functions of all the methods can rapidly converge within 20 iterations. In particular, the methods with single-manifold regularized term have faster convergence speed than MLA.

#### 5.2. Recognition Performance Evaluation

In this section, the proposed strategy, MLA-LSR, is compared with several methods in terms of recognition performance under different extended operation conditions. We compared it not only against SRC [28], LSR, GNMF [21], MMF [22] and MLA-SR, but also against the newly developed methods, i.e., joint sparse representation of monogenic signal (TJSR) [38], sparse representation on Grassmann manifold ( $\text{KSR}^M$ ) [39], and all-convolutional network (A-ConvNets) [40]. Specifically, SRC and LSR are employed with the PCA feature. MLA-SR stands for MLA features based on SRC.

To perform a fair comparison, GNMF and MMF use the LSR for classification. Additionally, the same dimension of 80D is selected in the following experiments.



**Figure 6.** The convergence curve of different regularization. The single manifold<sub>1</sub> and the single manifold<sub>2</sub> denote the methods with pairwise similarity rule and local linearity rule, respectively. The multi-manifold represents the proposed MLA method.

### 5.2.1. Performance on Configuration Variation

Due to the physical difference and structure modifications for targets, the configuration variation is not inevitable for target recognition in SAR images. To validate the efficacy of the proposed strategy for configuration variation, four class targets are chosen from BMP2, BTR60, T72 and T62, as listed in Table 1. BMP2 and BTR60 are armoured personnel carriers, while T72 and T62 are main battle tanks. Example images are shown in Figure 7. In addition, BMP2 and T72 consist of multiple variants. Only the standards (SN 9563 for BMP2 and SN 132 for T72) are available for training, and the other variants are used for testing.

Figure 8 presents a detailed comparison of MLA-LSR with a variety of methods. The MLA-SR and MLA-LSR methods outperform all the baseline SRC, LSR, GNMF and MMF methods. The results indicate that the low-dimensional representation of SAR images obtained by MLA could be more effective than those obtained by the single-manifold learning methods. In addition, the recognition accuracy for MLA-LSR is 0.9582, which is 1.72% better than that of MLA-SR. This result demonstrates that LSR produces more believable inference than SRC. MLA-LSR is also shown to have better performance than the latest TJSR, KSR<sup>M</sup> and A-ConvNets methods. From the confusion matrices, we can see that BMP2 and T72 are more likely to be misclassified, as different types of targets are employed in the training and testing. However, MLS-LSR still maintains a higher recognition accuracy than those of the other methods.



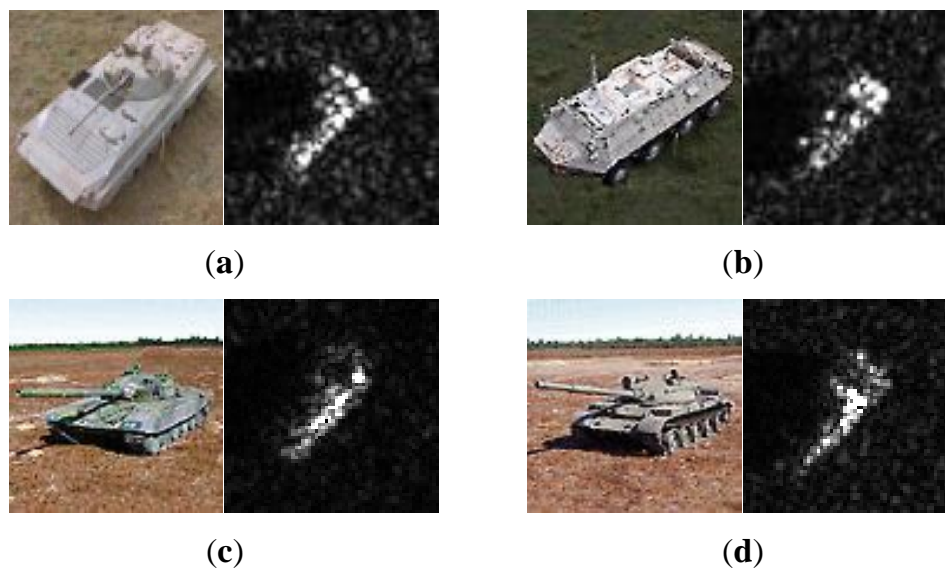


Figure 7. Illustration of four-class target: (a) BMP2; (b) BTR60; (c) T72; and (d) T62.

0.8925	SRC				0.9246	LSR			
	BMP2	BTR60	T72	T62		BMP2	BTR60	T72	T62
BMP2	<b>0.9362</b>	0.0230	0.0128	0.0281	BMP2	<b>0.9439</b>	0.0026	0.0459	0.0077
BTR60	0.0256	<b>0.9692</b>	0.0051	0	BTR60	0.0154	<b>0.9692</b>	0	0.0154
T72	0.0337	0.0285	<b>0.7435</b>	0.1943	T72	0.0233	0.0052	<b>0.8342</b>	0.1373
T62	0.0110	0	0.0037	<b>0.9853</b>	T62	0.0073	0	0	<b>0.9927</b>
0.9117	TJSR				0.9270	KSR <sup>44</sup>			
	BMP2	BTR60	T72	T62		BMP2	BTR60	T72	T62
BMP2	<b>0.9170</b>	0.0230	0.0510	0.0153	BMP2	<b>0.9311</b>	0.0281	0.0357	0.0051
BTR60	0.0256	<b>0.9385</b>	0.0103	0.0103	BTR60	0	<b>0.9897</b>	0	0.0103
T72	0.0181	0	<b>0.8349</b>	0.1269	T72	0.0570	0.0026	<b>0.8420</b>	0.0984
T62	0	0	0.0037	<b>0.9927</b>	T62	0	0.0037	0	<b>0.9963</b>
0.9278	GNMF				0.9342	MMF			
	BMP2	BTR60	T72	T62		BMP2	BTR60	T72	T62
BMP2	<b>0.9617</b>	0	0.0255	0.0128	BMP2	<b>0.9260</b>	0.0102	0.0587	0.0051
BTR60	0.0205	<b>0.9641</b>	0.0103	0.0051	BTR60	0.0205	<b>0.9641</b>	0.0051	0.0103
T72	0.0155	0	<b>0.8290</b>	0.1554	T72	0.0207	0	<b>0.8834</b>	0
T62	0.0073	0	0	<b>0.9927</b>	T62	0.0037	0	0	<b>0.9963</b>
0.9436	A-ConvNets				0.9390	MLA-SR			
	BMP2	BTR60	T72	T62		BMP2	BTR60	T72	T62
BMP2	<b>0.9517</b>	0.0368	0.0025	0.0089	BMP2	<b>0.9796</b>	0.0077	0.0077	0.0051
BTR60	0.0051	<b>0.9898</b>	0	0.0051	BTR60	0.0205	<b>0.9590</b>	0.0051	0.0154
T72	0.0054	0.0027	<b>0.8261</b>	0.1658	T72	0.0078	0.0026	<b>0.8497</b>	0.1399
T62	0.0037	0.0037	0.0073	<b>0.9854</b>	T62	0.0073	0	0	<b>0.9927</b>
0.9582	MLA-LSR								
	BMP2	BTR60	T72	T62					
BMP2	<b>0.9847</b>	0	0.0102	0.0051					
BTR60	0.0205	<b>0.9641</b>	0.0051	0.0103					
T72	0.0078	0.0052	<b>0.9041</b>	0.0829					
T62	0.0073	0	0	<b>0.9927</b>					

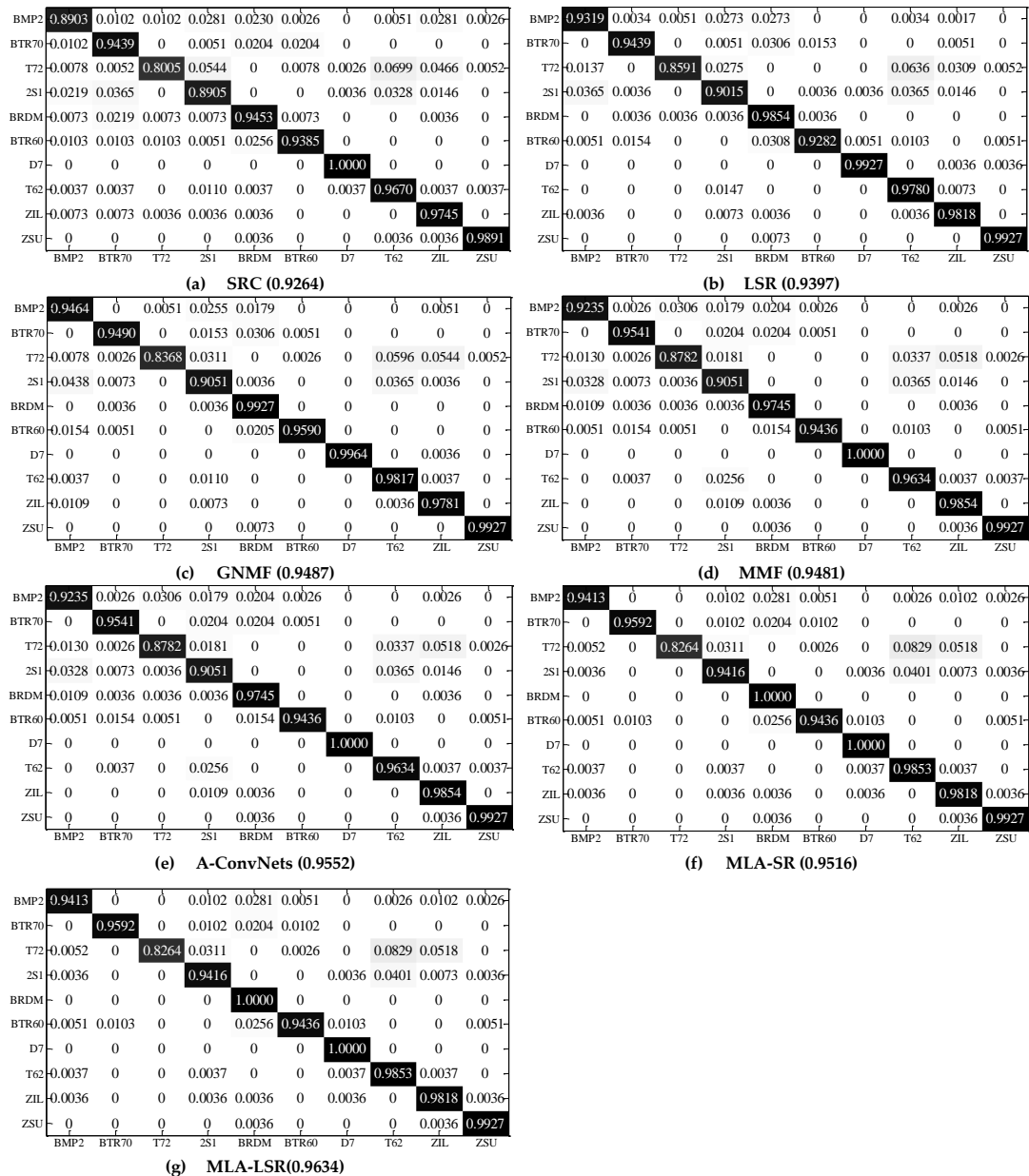
Figure 8. Confusion matrices by different methods. For each method, the average recognition accuracy is given on the top left corner of matrix.

### 5.2.2. Performance on Ten-class

Target recognition performed on all of the ten-class targets in the MSTAR database is a more challenging problem. Details of the numbers and types of the ten-class targets are listed in Table 4. Similarly, for BMP2 and T72 with multiple variants, the standards are employed for training, while the others are utilized for testing. To analyse the misclassification of targets, confusion matrices of the studied methods are shown in Figure 9.

Table 4. The number of images for ten-class targets.

Depr.	BMP2	BTR70	T72	2S1	BRDM	BTR60	D7	T62	ZIL	ZSU
17°	233(SN_9563)	233	233(SN_132)	299	298	256	299	299	299	299
15°	196(SN_9566) 196(SN_c21)	196	195(SN_812) 191(SN_s7)	274	274	195	273	273	274	274



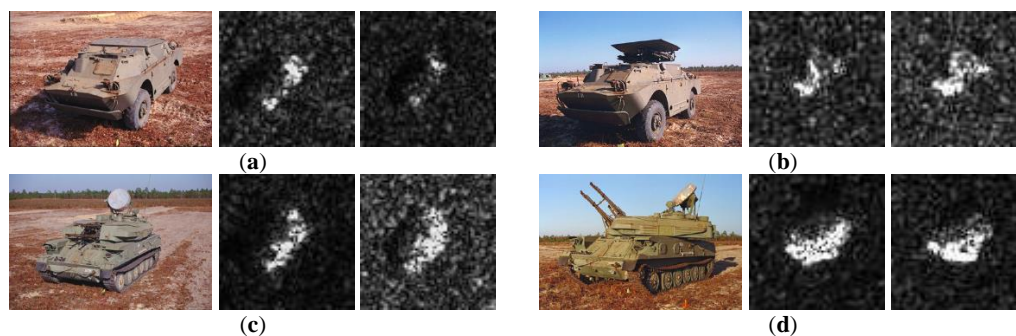
**Figure 9.** The confusion matrices of SRC, LSR, GNMF, MMF, MLA-SR, A-ConvNets, and MLA-LSR. For each matrix, the entry denotes the percentage of classifying the target in the row as the target in column. The average accuracies are given in parentheses.

The results presented in Figure 8 show that the MLA-LSR method outperforms all the other methods in terms of the recognition accuracy on the ten-class targets. Specifically, the proposed MLA-LSR strategy achieves a high recognition accuracy of 0.9634. It outperforms SRC, LSR, GNMF, MMF, A-ConvNets and MLA-SR by a margin of 3.7, 2.37, 1.47, 1.53, 0.82 and 1.18 percent, respectively.

From these matrices in Figure 8, most targets without variants could be correctly identified in MLA-LSR, which is attributed to the combination of MLA and LSR to improve the discriminative ability. In addition, similar to the results shown in Section 5.2.1, the performance of MLA-LSR is also consistently better than that of the other methods for the targets with variants, particularly for BMP2.

### 5.2.3. Performance on Larger Depression and Articulation Variation

Images collected under different depression angles would change drastically. Hence, the feature extraction and classification methods should be robust to changes in depression angles. In this subsection, the performance of the proposed strategy, MLA-LSR, is evaluated under different depression angles. Three vehicle targets, 2S1, BRDM2 and ZSU234, are used in the experiments. Images captured at a  $17^\circ$  depression angle are used for training, whereas the ones acquired at  $30^\circ$  and  $45^\circ$  depression angles are used for testing. In particular, BRDM2 and ZSU234 have articulated variants in the testing, such as opening or closing hatch and straight or rotational turret. The examples with depression and articulated variants are shown in Figure 10. The number and depression information of these targets are listed in Table 5. The entry in parentheses means the number of targets with articulated variants.

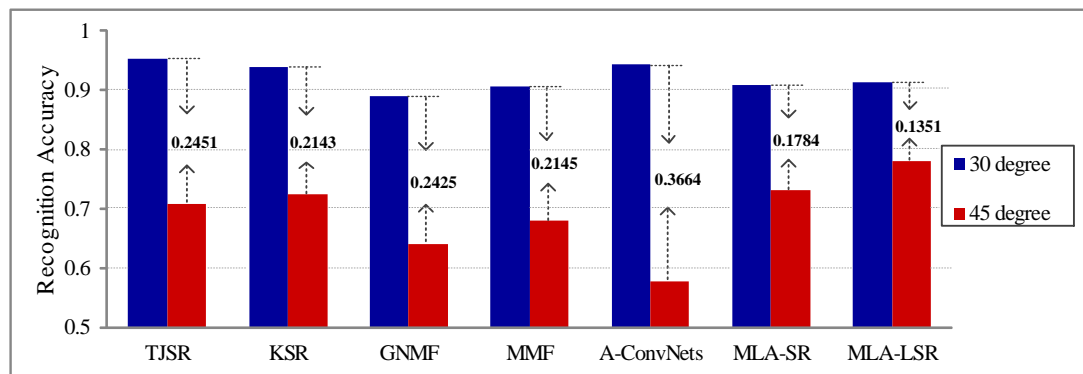


**Figure 10.** Examples of BRDM2 and ZSU234 with depression and articulated variants: (a,b) display SAR images of BRDM2 collected at  $30^\circ$  and  $45^\circ$  with hatch variation; and (c,d) display SAR images collected at  $30^\circ$  and  $45^\circ$  with gun turret variation.

**Table 5.** The number of images for large depression and articulation variation.

Type	Depression	2S1	BRDM2	ZSU234	Total
Training	$17^\circ$	299	298	299	896
Testing	$30^\circ$	288	287 (133)	288 (118)	1114
	$45^\circ$	303	303 (120)	303 (119)	1148

Table 6 shows the results of our experiments with MLA-LSR, as well as the confusion matrices. The recognition accuracies of MLA-LSR compared with TJSR,  $KSR^M$ , GNMF, MMF and MLA-SR are given in Figure 11. When the depression angle for testing is  $30^\circ$ , two targets, BRDM2 and ZSU234, can easily be misclassified due to the articulation variation. When the depression angle has a drastic changes of  $28^\circ$  from  $17^\circ$  for training to  $45^\circ$  for testing, the recognition accuracies of the three targets are slightly degraded due to the abrupt change in the signatures of the same targets from different depression angles. Regardless, the recognition accuracy of the proposed MLA-LSR strategy still reaches 0.7778, compared to 0.7073 for TJSR, 0.7238 for  $KSR^M$ , 0.6408 for GNMF, 0.6803 for MMF, 0.5763 for A-ConvNets and 0.73 for MLA-SR. In addition, the decrease of recognition accuracy for MLA-LSR is 0.1351, which is 11%, 9.92% 10.74%, 7.94%, 23.13% and 4.33% better than its six competitors. The experimental results prove that the proposed strategy is much more robust towards variations in depression and articulation than the competitors, particularly for large depression variations.



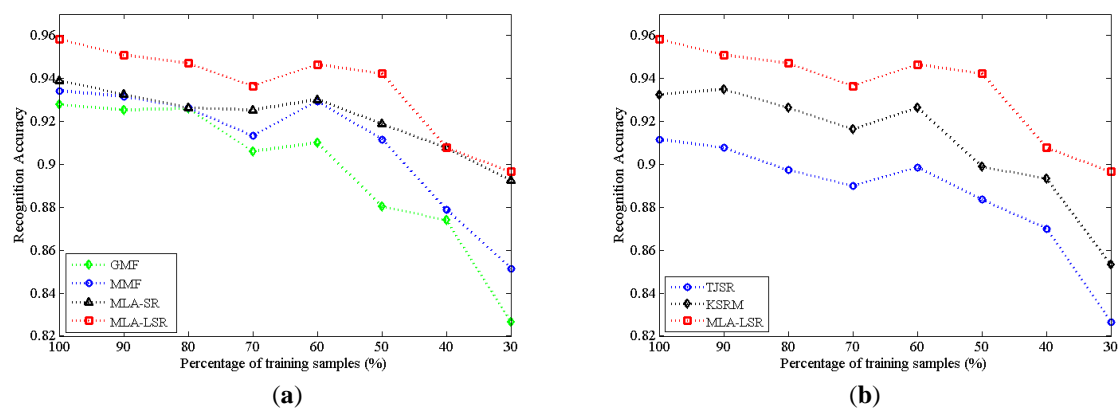
**Figure 11.** The recognition rates of various methods. The number on the right of the bar represents the decrease in recognition performance across the range of depression variation from 30° to 45°.

**Table 6.** Recognition accuracies of MLA-LSR on the depression and articulation variation database.

Target	30° Depression			45° Depression		
	2S1	BRDM2	ZSU234	2S1	BRDM2	ZSU234
2S1	0.9826	0.0069	0.0104	0.7492	0.1716	0.0792
BRDM2	0.0762	0.8333	0.0905	0.0804	0.6478	0.2719
ZSU234	0.0296	0.0246	0.9458	0.0308	0.0403	0.9289
Average	0.9129			0.7778		

#### 5.2.4. Performance on Small Sample Size

An important consideration in target recognition is the robustness to limited training samples. Similar to the experiment set in Table 1, four-class targets are employed, but the number of training samples is gradually decreased. Specifically, the sizes of the training samples are chosen to be 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% of the size of the original training samples. Figure 12 plots the recognition performances of MLA-LSR and its competitors as a function of the size of training samples. The performances of all methods degrade when the number of training samples decreases. However, the performance of the proposed MLA-LSR strategy is still superior to those of the other methods under different training sizes. The results demonstrate that MLA-LSR is capable of performing well on limited training samples.



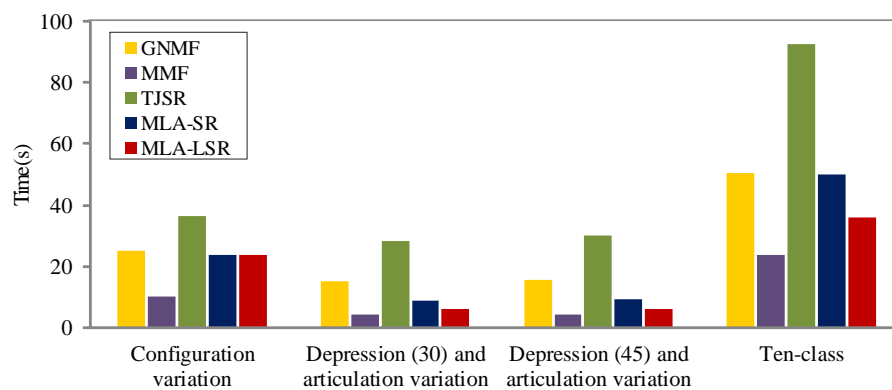
**Figure 12.** The recognition accuracy across the size of training.

### 5.3. Computational Complexity

The computational cost of the proposed method consists of the calculation of MLA and LSR. For MLA, the primary computational load is the alternative iteration of  $\mathbf{H}$ ,  $\mathbf{Z}$  and  $\eta$ . The updating rule in Equation (16) for  $\mathbf{H}$  involves inverting a matrix  $\mathbf{I} + \lambda_1 \Psi$ . Actually, it is equivalent to solving  $d$  linear systems. Thus, the preconditioned conjugate gradient (PCG) [41] can address the linear system efficiently, which needs  $4n + n^2$  in each iteration. Combining with the cost of  $dn^2$  for calculating  $d$  right-hand-side vectors, the computational complexity for updating  $\mathbf{H}$  is approximated as  $p(4n + n^2)d + dn^2$ , where  $p$  is the iterative times, typically  $p < 10$ . For updating rule in Equation (16) for  $\mathbf{Z}$ , the time cost is  $2md^2 + 20d^3 + (m + d)d^2$ , specifically,  $2md^2$  for  $\mathbf{X}\mathbf{H}^T$ ,  $20d^3$  for SVD [42] and  $(m + d)d^2$  for forming  $\mathbf{Z}$ . The remaining cost of MLA in each iteration is  $r^3 + nd$  for updating  $\eta$  in Equation (17). In addition, MLA requires  $O(rmn^2)$  to construct the affinity matrices of multiple manifolds. Hence, the overall computational complexity of MLA is approximately  $O(t((p + 1)dn^2 + md^2) + rmn^2)$ , where  $t$  is the number of iterations. For LSR, the computational complexities for obtaining the low-dimensional feature representation of a query sample in Equation (18) and the sparse representation for each class in Equation (21) are  $md$  and  $O(n_c^2)$ . Since we need to calculate  $c$  local sparse representation systems, the total workload of LSR is approximately  $O(md + cn_c^2)$ . Therefore, the cost of the proposed MLA-LSR method is approximated as

$$O(t((p + 1)dn^2 + md^2) + rmn^2) + O(md + cn_c^2) \quad (23)$$

Figure 13 shows the runtime of MLA-LSR and its four competitors under different experimental conditions, i.e., target recognition under configuration variation, depression (both  $30^\circ$  and  $45^\circ$ ) and articulation variation, as well as ten-class targets. From the results in Figure 13, some conclusion can be drawn. (1) The construction of optimal manifold based on multiple manifolds would moderately add computational load of MLA-LSR. For example, the runtime of MLA-LSR is higher than that of MMF, as MMF only utilizes a single-manifold. However, the MLA-LSR iteration converges fast. Therefore, the computation cost of the proposed method is still acceptable under the realistic applications. (2) The locality constraint would improve the computation efficiency of sparse representation. For all the experiments in Figure 13, the runtime of MLA-LSR is lower than that of MLA-SR. The results verify that the computation cost of LSR is low. (3) The computation efficiency of MLA-LSR is much higher than that of TJSR. According to the order of experiments in Figure 13, the runtime of MLS-LSR is 12.44 s, 22.05 s, 22.94 s and 56.62 s lower than TJSR, respectively.



**Figure 13.** Comparison of the runtime of MLA-LSR with GNMf, MMF, TJSR and MLA-SR under different experimental conditions.

## 6. Conclusions

In this paper, a new strategy, MLA-LSR, is proposed for SAR target recognition. In the proposed method, MLA refers to obtaining a low-dimensional representation of SAR images. LSR refers to

using a local sparse representation for classification. In the proposed MLA model, a multi-manifold regularization is incorporated into the TSVD framework. In this way, MLA can not only acquire the optimal low-rank approximation of the original samples but also capture the intrinsic manifold structure. In the proposed LSR classification method, a locality constraint and class-specific model are used to form the dictionary, by which the query sample is sparsely represented. The inference is reached corresponding to the class with the minimum reconstruction error. Compared to the traditional sparse-representation-based classification method, LSR improves the discriminative ability. Extensive experiments are performed on the MSTAR database to evaluate the performance of MLA-LSR. From these experimental results, some conclusions can be drawn. (1) The multi-manifold regularization term employed in the low-dimensional representation is more effective than a single-manifold term for capturing the intrinsic manifold structure of samples. (2) The locality constraint in the sparse representation model is more appropriate to discriminate the features obtained by MLA than the sparsity constraint. (3) The proposed strategy is a contender for establishing a firm foothold on territory from which stronger invariances for configuration variation, depression and articulation variation can be built. (4) The proposed strategy achieves a significant improvement compared to competitors with small training sizes.

**Acknowledgments:** The authors would like to thank the handling Associate Editors and the anonymous Reviewers for their constructive comments and suggestions for this paper. Meiting Yu would like to thank Haipeng Wang for sharing the network parameters of A-ConvNets. This work was supported by the National Natural Science Foundation of China under Grant No. 61701508 and No. 61601481.

**Author Contributions:** Meiting Yu proposed the original idea and performed the experiments; Ganggang Dong, Haiyan Fan and Gangyao Kuang reviewed the idea and contributed to the discussion of the results; Ganggang Dong and Haiyan Fan provided much analysis; Meiting Yu wrote the paper.

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

## References

1. Song, S.L.; Xu, B.; Yang, J. SAR target recognition via supervised discriminative dictionary learning and sparse representation of the SAR-HOG feature. *Remote Sens.* **2016**, *8*, 683.
2. Umamahesh, S.; Vishal, M. SAR automatic target recognition using discriminative graphical models. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 591–606.
3. Dong, G.; Kuang, G.; Wang, N.; Wang, W. Classification via sparse representation of steerable wavelet frames on grassmann manifold: Application to target recognition in SAR image. *IEEE Trans. Image Proc.* **2017**, *26*, 2892–2904.
4. Tang, Y.Y.; Yuan, H.; Li, L. Manifold-Based Sparse Representation for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7606–7618.
5. Zhang, L.; Zhang, Q.; Du, B.; Huang, X.; Tang, Y.Y.; Tao, D. Simultaneous Spectral-Spatial Feature Selection and Extraction for Hyperspectral Images. *IEEE Trans. Cybern.* **2018**, *48*, 16–28.
6. Ding, B.Y.; Wen, G.J.; Huang, X.H.; Ma, C.H.; Yang, X.L. Data augmentation by multilevel reconstruction using attributed scattering center for SAR target recognition. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 979–983.
7. Zhou, J.X.; Shi, Z.G.; Cheng, X.; Fu, Q. Automatic target recognition of SAR images based on global scattering center model. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3713–3729.
8. Dong, G.; Wang, N.; Kuang, G. Sparse representation of monogenic signal: With application to target recognition in SAR images. *IEEE Signal Proc. Lett.* **2014**, *21*, 952–956.
9. Huang, X.Y.; Qiao, H.; Zhang, B. SAR target configuration recognition using tensor global and local discriminant embedding. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 222–226.
10. Yu, M.T.; Zhao, L.J.; Zhao, S.Q.; Xiong, B.L.; Kuang, G.Y. SAR target recognition using parametric supervised t-stochastic neighbor embedding. *Remote Sens. Lett.* **2017**, *8*, 849–858.
11. Clemente, C.; Pallotta, L.; Proudler, I.K.; De Maio, A.; Soraghan, J.J.; Farina, A. Pseudo-Zernike Based Multi-Pass Automatic Target Recognition from Multi-Channel SAR. *IET Radar Sonar Navig.* **2015**, *9*, 457–466.
12. Cui, Z.; Cao, Z.; Yang, J.; Feng, J.; Ren, H. Target recognition in synthetic aperture radar images via non-negative matrix factorisation. *IET Radar Sonar Navig.* **2015**, *9*, 1376–1385.



13. Dang, S.; Cui, Z.; Cao, Z.; Liu, Y.; Min, R. SAR target recognition via incremental nonnegative matrix factorization with  $L_p$  sparse constraint. In Proceedings of the 2017 IEEE Radar Conference, Seattle, WA, USA, 8–12 May 2017; pp. 530–534.
14. Babaei, M.; Yu, X.; Rigoll, G.; Datcu, M. Immersive Interactive SAR Image Representation Using Non-negative Matrix Factorization. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2844–2853.
15. Belkin, M.; Niyogi, P. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering. 2002, pp. 585–591. Available online: [papers.nips.cc/paper/1961-laplacian-eigenmaps-and-spectral-techniques-for-embedding-and-clustering.pdf](http://papers.nips.cc/paper/1961-laplacian-eigenmaps-and-spectral-techniques-for-embedding-and-clustering.pdf) (accessed on 29 January 2018).
16. Roweis, S.T.; Saul, L.K. Nonlinear dimensionality reduction by locally linear embedding. *Science* **2000**, *290*, 2323–2326.
17. Zha, H.; Zhang, Z. Spectral properties of the alignment matrices in manifold learning. *SIAM Rev.* **2009**, *51*, 545–566.
18. Der Maaten, L.V.; Hinton, G.E. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
19. He, X.; Niyogi, P. Locality Preserving Projections. 2003, pp. 186–197. Available online: [papers.nips.cc/paper/2359-locality-preserving-projections.pdf](http://papers.nips.cc/paper/2359-locality-preserving-projections.pdf) (accessed on 29 January 2018).
20. He, X.; Cai, D.; Yan, S.; Zhang, H.J. Neighborhood preserving embedding. In Proceedings of the 10th IEEE International Conference on Computer Vision, Beijing, China, 17–21 October 2005; pp. 1208–1213.
21. Cai, D.; He, X.; Han, J.; Huang, T.S. Graph regularized nonnegative matrix factorization for data representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 1548–1560.
22. Zhang, Z.; Zhao, K. Low-rank matrix approximation with manifold regularization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1717–1729.
23. Wang, J.J.; Bensmail, H.; Gao, X. Multiple graph regularized nonnegative matrix factorization. *Pattern Recognit.* **2013**, *46*, 2840–2847.
24. Trigeorgis, G.; Bousmalis, K.; Zafeiriou, S.; Schuller, B. A deep matrix factorization method for learning attribute representations. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 417–429.
25. Cover, T.M.; Hart, P.E. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27.
26. Tison, C.; Pourthie, N.; Souyris, J.C. Target recognition in SAR images with support vector machines. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 23–28.
27. Sun, Y.; Liu, Z.; Todorovic, S.; Li, J. Adaptive boosting for SAR automatic target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 112–125.
28. Wright, J.; Yang, A.Y.; Ganesh, A.; Sastry, S.; Ma, Y. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 210–227.
29. Thiagarajan, J.J.; Ramamurthy, K.N.; Knee, P.; Spanias, A.; Berisha, V. Sparse representations for automatic target classification in SAR images. In Proceedings of the 2010 4th International Symposium on Communications, Control Signal Processing (ISCCSP), Limassol, Cyprus, 3–5 March 2010; pp. 1–4.
30. Ding, B.; Wen, G. Exploiting multi-view SAR image for robust target recognition. *Remote Sens.* **2017**, *9*, 1150.
31. Wang, J.; Yang, J.; Yu, K.; Lv, F.; Huang, T.S.; Gong, Y. Locality-constrained linear coding for image classification. In Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 13–18 June 2010; pp. 3360–3367.
32. Gong, L.; Mu, T.; Goulermas, J.Y. Evolutionary Nonnegative Matrix Factorization for Data Compression. In *Intelligent Computing Theories and Methodologies, Proceedings of the 11th International Conference on Intelligent Computing (ICIC 2015), Fuzhou, China, 20–23 August 2015*; Huang, D.S., Bevilacqua, V., Premaratne, P., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Part I, pp. 23–33.
33. Gillis, N.; Kuang, D.; Park, H. Hierarchical Clustering of Hyperspectral Images Using Rank-Two Nonnegative Matrix Factorization. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2066–2078.
34. Berisha, V.; Nitesh, N.S. Sparse manifold learning with application to SAR image classification. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Honolulu, HI, USA, 15–20 April 2007; pp. 349–354.
35. Geng, B.; Tao, D.; Xu, C.; Yang, L.; Hua, X. Ensemble manifold regularization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 1227–1233.

36. Mishra, A.K.; Motaung, T. Application of linear and nonlinear PCA to SAR ATR. In Proceedings of the 2015 25th International Conference on Radioelektronika, Pardubice, Czech Republic, 21–22 April 2015; pp. 1089–1092.
37. Cao, Z.J.; Min, R.; Pi, Y.M.; Xu, Z.W. The feasibility analysis of applying NMF in SAR target recognition. In Proceedings of the IEEE International Conference on Digital Signal Processing (DSP), Singapore, 21–24 July 2015; pp. 721–725.
38. Dong, G.; Kuang, G.; Wang, N.; Zhao, L.; Lu, J. SAR target recognition via joint sparse representation of monogenic signal. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3316–3328.
39. Dong, G.; Kuang, G. SAR target recognition via sparse representation of monogenic signal on grassmann manifolds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1308–1319.
40. Chen, S.; Wang, H.; Xu, F.; Jin, Y. Target Classification Using the Deep Convolutional Networks for SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4806–4817.
41. Hestenes, H.R.; Stiefel, E. Methods of conjugate gradients for solving linear analysis. *J. Res. Nat. Bur. Stand.* **1952**, *49*, 81–85.
42. Demmel, J.; Kahan, W. Accurate singular values of bidiagonal matrices. *SIAM J. Sci. Stat. Comput.* **1990**, *11*, 873–912.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).