URSI GASS 2021, Rome, Italy, 28 August - 4 September 2021

SAR Target Configuration Recognition via Multi-Manifold Based Sparse Representation

Ming Liu* (1) and Shichao Chen(2)
(1) School of Computer Science, Shaanxi Normal University, Xi'an 710119, China
(2) No. 203 Research Institute of China Ordnance Industries, Xi'an 710065, China

Abstract

A multi-manifold based sparse representation (MMSR) algorithm which can preserve the local structure of the datasets in the reconstruction space with multiple descriptions is proposed for synthetic aperture radar (SAR) target configuration recognition. Different from the traditional manifold learning-based algorithms that preserve the local structure of the datasets in the low-dimensional space, the proposed MMSR guarantees local structure preserving of the datasets after feature extraction and sparse representation (SR). The proposed MMSR can avoid corrupted relationships of the samples to realize more accurate SAR target configuration recognition. Satisfying results are obtained on the moving and stationary target acquisition and recognition (MSTAR) database.

1 Introduction

As one of the most important application areas of synthetic aperture radar (SAR), target recognition has attracted increasing popularity nowadays. Plenty of new techniques have been adopted to realize SAR target recognition [1-2]. Various advanced algorithms have been employed for SAR target recognition, such as sparse representation (SR) [3] and manifold learning [4].

SR has been successfully used in various fields including pattern recognition. The key point of the SR technique is to solve the sparse representation of the sample as accurate as possible. Recognition result is determined by searching the smallest reconstruction error in the reconstruction space. Manifold learning has been proved to be a powerful technique for local structure preserving, which plays a crucial role for pattern recognition. Fusing the advantages of manifold learning and SR together, a feature extraction algorithm which can preserve the local structure of the datasets in the reconstruction space is proposed to realize high-precision SAR target configuration recognition in this paper. The testing sample is projected onto different manifolds utilizing the prior label information, and SR is implemented in each manifold. Samples that are close to each other in the high-dimensional space will still be close to each other in the reconstruction space by using the proposed algorithm. The one best describes the testing sample will be regarded as the recognition result.

Due to the effective structure preserving of the datasets, the proposed algorithm can realize not only target type recognition, but also target configuration recognition. Configuration recognition [5-6] can provide more detailed information of the samples.

2 SAR Target Configuration Recognition via Multi-Manifold Based Sparse Representation (MMSR)

It has been proved that local structure of the datasets is of crucial importance to recognition [4, 7]. And preserving the local structure of the samples in the reconstruction space is helpful to SR based SAR target recognition [5].

To achieve satisfying performance and robustness, we present a multi-manifold based sparse representation (MMSR) algorithm for SAR target configuration recognition in this paper. In [5], all the targets with different configurations are assumed to lie in the same manifold. However, these targets cannot be fully described by the same manifold. They more likely lie in different manifolds according to the configurations they belong to. In other words, each configuration lies in a separate manifold. To better clarity the point, we take an example to demonstrate the idea in Figure 1.

Here, we assume that there are six configurations, and each color represents an individual configuration in Figure 1. We name the configurations from left to right as configuration 1 to configuration 6. In [5], samples of all configurations are assumed to lie in the same manifold. The relationships of the samples in the Euclidean space are displayed in Figure 1(a). And their relationships in the intrinsic manifold space are displayed in Figure 1(b). From the comparison, we can tell that although configuration 2 and configuration 6 are far from each other actually, they are quite close to each other in the Euclidean space. And if we implement feature extraction and recognition on the basis of the Euclidean space, a wrong relationship will be utilized, and poor recognition results will emerge. We can discover the intrinsic structure of the datasets by using the algorithm proposed in [5]. However, there are still some corrupted relationships in border regions, as shown in the black dot blocks of Figure 1(b). To overcome the problem, we divide the single manifold into separate ones according to the sample labels, as shown in Figure 1(c). In this way, each configuration forms a single manifold to avoid

corrupted relationships. For a testing sample, we will describe it in multiple manifolds, and each manifold will give a description. Among all the descriptions, the one with the same label as the testing sample will perform the best.

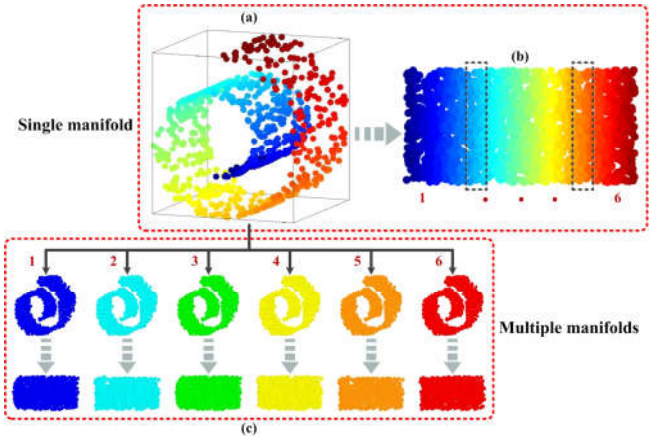

Figure 1. Multi-manifold model.

In MMSR, we construct multiple projection matrices by using the training samples with different labels. That is to say, we will solve the projection matrix $\hat{\mathbf{W}}_i$ by using the training samples of label $i \in (1,2,...,C)$ individually, where C is the number of configurations. In this paper, we try to preserve the local structure of the datasets in the reconstruction space, where recognition is implemented for better recognition results. The objective function of the proposed MMSR is constructed by

$$\hat{\mathbf{W}}_i = \arg\min_{\mathbf{W}_i} \sum_{s,t=1}^{N_i} \left\| \hat{\mathbf{u}}_{is} - \hat{\mathbf{u}}_{it} \right\|_2^2 \mathbf{S}_{ist} \tag{1}$$

where N_i is the number of training samples belong to configuration i, $\hat{\mathbf{u}}_{is}$ and $\hat{\mathbf{u}}_{it}$ are sparse reconstructions of $\mathbf{u}_{is}$ and $\mathbf{u}_{it}$, $\mathbf{u}_{is}, \mathbf{u}_{it} \in \mathbf{U}_i = \left(\mathbf{u}_{i1}, \mathbf{u}_{i2}, ..., \mathbf{u}_{iN_i}\right)$ are the representations of the original training samples of $\mathbf{v}_{is}$ and $\mathbf{v}_{it}$ after feature extraction, $\mathbf{U}_i \in \Re^{d \times N_i}$, $\mathbf{U}_i = \mathbf{W}_i^T \mathbf{V}_i$, $\mathbf{V}_i$ is the original training datasets of configuration i, $\mathbf{v}_{is}, \mathbf{v}_{it} \in \mathbf{V}_i = \left\{\mathbf{v}_{i1}, \mathbf{v}_{i2}, ..., \mathbf{v}_{iN_i}\right\}$, $\mathbf{u}_{is} = \mathbf{W}_i^T \mathbf{v}_{is}$, $\mathbf{u}_{it} = \mathbf{W}_i^T \mathbf{v}_{it}$, and d is the number of the dimensionality after feature extraction. In other words, $\mathbf{W}_i$ is the projection matrix corresponds to configuration i, and $\hat{\mathbf{W}}_i$ is the solution of $\mathbf{W}_i$. $\mathbf{S}$ is a weighted matrix of a graph model who has N_i nodes.

$$\mathbf{S}_{ist} = \begin{cases} \exp\left[-\dfrac{\left\|\mathbf{v}_{is} - \mathbf{v}_{it}\right\|_2}{b}\right], & \mathbf{v}_{is} \in N_k\left(\mathbf{v}_{it}\right) \text{ or } \mathbf{v}_{it} \in N_k\left(\mathbf{v}_{is}\right) \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where b is a constant which needs to be determined in practice, $N_k\left(\mathbf{v}_{is}\right)$ stands for the k nearest neighbors of $\mathbf{v}_{is}$, $N_k\left(\mathbf{v}_{it}\right)$ stands for the k nearest neighbors of $\mathbf{v}_{it}$, $\exp(\bullet)$ stands for the exponential function, and $\|\bullet\|_2$ represents the L2 norm.

Let us take a look at how (1) works to guarantee structure preserving. Taking any two training samples $\mathbf{v}_{is}$ and $\mathbf{v}_{it}$ as an example, if they are close to each other in the high-dimensional space, the corresponding weight calculated by (2) will be large, and minimizing (1) implies that the sparse reconstructions $\hat{\mathbf{u}}_{is}$ and $\hat{\mathbf{u}}_{it}$ will be closer to each other in the reconstruction space. That is to say, we realize structure preserving between the high-dimensional space and the reconstruction space, where recognition is implemented.

Here, we rewrite the objective function as

$$\begin{aligned} J_i &= \frac{1}{2} \sum_{s,t=1}^{N_i} \left\| \hat{\mathbf{u}}_{is} - \hat{\mathbf{u}}_{it} \right\|_2^2 \mathbf{S}_{ist} \\ &= \frac{1}{2} \sum_{s,t=1}^{N_i} \left\| \mathbf{U}_i \boldsymbol{\theta}_{is} - \mathbf{U}_i \boldsymbol{\theta}_{it} \right\|_2^2 \mathbf{S}_{ist} \\ &= \frac{1}{2} \sum_{s,t=1}^{N_i} \left\| \mathbf{W}_i^T \mathbf{V}_i \boldsymbol{\theta}_{is} - \mathbf{W}_i^T \mathbf{V}_i \boldsymbol{\theta}_{it} \right\|_2^2 \mathbf{S}_{ist} \end{aligned} \tag{3}$$

where $\boldsymbol{\theta}_{is}, \boldsymbol{\theta}_{it} \in \boldsymbol{\Phi}_i = \left(\boldsymbol{\theta}_{i1}, \boldsymbol{\theta}_{i2}, ..., \boldsymbol{\theta}_{iN_i}\right)^T$ are sparse representations of $\mathbf{u}_{is}$ and $\mathbf{u}_{it}$, respectively. The s^{th} element of $\boldsymbol{\theta}_{is}$ and the t^{th} element of $\boldsymbol{\theta}_{it}$ are zero.

Defining $\hat{\mathbf{v}}_{is} = \mathbf{V}_i \boldsymbol{\theta}_{is}$, $\hat{\mathbf{v}}_{it} = \mathbf{V}_i \boldsymbol{\theta}_{it}$ and applying some basic mathematical skills, (3) can be expressed as

$$\begin{aligned} J_i &= \frac{1}{2} \sum_{s,t=1}^{N_i} \left\| \mathbf{W}_i^T \hat{\mathbf{v}}_{is} - \mathbf{W}_i^T \hat{\mathbf{v}}_{it} \right\|_2^2 \mathbf{S}_{ist} \\ &= \sum_{s=1}^{N_i} \mathbf{W}_i^T \hat{\mathbf{v}}_{is} \mathbf{H}_{iss} \hat{\mathbf{v}}_{is}^T \mathbf{W}_i - \mathbf{W}_i^T \hat{\mathbf{V}}_i \mathbf{S}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i \\ &= \mathbf{W}_i^T \hat{\mathbf{V}}_i \mathbf{H}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i - \mathbf{W}_i^T \hat{\mathbf{V}}_i \mathbf{S}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i \\ &= \mathbf{W}_i^T \hat{\mathbf{V}}_i \left(\mathbf{H}_i - \mathbf{S}_i\right) \hat{\mathbf{V}}_i^T \mathbf{W}_i \\ &= \mathbf{W}_i^T \hat{\mathbf{V}}_i \mathbf{L}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i \end{aligned} \tag{4}$$

where $\mathbf{H}_{iss} = \sum_t \mathbf{S}_{ist}$, $\hat{\mathbf{V}}_i = \mathbf{V}_i \boldsymbol{\Phi}_i$. $\mathbf{L}_i = \mathbf{H}_i - \mathbf{S}_i$ is a

Laplace matrix, and $\mathbf{H}_i$ is a symmetric diagonal matrix. With the constraint $\mathbf{W}_i^T \hat{\mathbf{V}}_i \mathbf{H}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i = \mathbf{I}$, we have

$$\begin{cases} \arg\min_{\mathbf{W}_i} \mathbf{W}_i^T \hat{\mathbf{V}}_i \mathbf{L}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i \\ s.t. \ \ \mathbf{W}_i^T \hat{\mathbf{V}}_i \mathbf{H}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i = \mathbf{I} \end{cases} \tag{5}$$

where $\mathbf{I}$ is an identity matrix. Hereto, the projection matrix can be obtained by solving the following general eigen-decomposition problem.

$$\hat{\mathbf{V}}_i \mathbf{L}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i = \lambda \hat{\mathbf{V}}_i \mathbf{H}_i \hat{\mathbf{V}}_i^T \mathbf{W}_i \tag{6}$$

The solved $\hat{\mathbf{W}}_i$ consists of the eigen-vectors correspond to the d smallest eigen-values of (6).

Projecting the training samples belong to configuration i by using the solved $\hat{\mathbf{W}}_i$, we can realize local structure preserving of the datasets in the reconstruction space, where SR based recognition is implemented.

After solving all the C projection matrices, we realize multiple feature extractions by the separated projection matrices. Training samples have been projected onto different manifolds according to their labels. The dictionaries are constructed individually according to the labels of the projected training samples, i.e., dictionary i consists of all the projected training samples with label i. What comes next is to implement sparse representation of the testing sample under the dictionary in each manifold. Each sparse representation gives a description of the testing sample. In the end, recognition is realized by identifying which reconstruction error is the smallest. The label corresponds to the smallest reconstruction error will be the final recognition result. The flow diagram of the proposed MMSR algorithm is given in Figure 2.

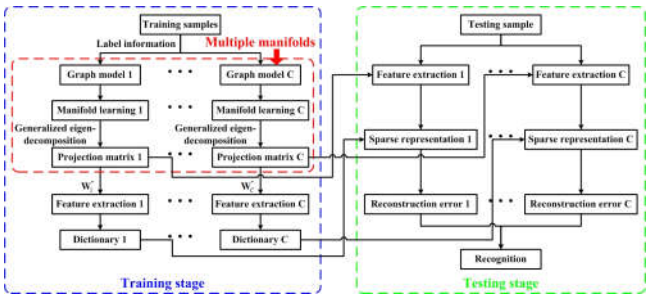

Figure 2. Flow diagram of the proposed MMSR algorithm.

3 Experimental Results and Analysis

The proposed MMSR algorithm is tested on the standard moving and stationary target acquisition and recognition (MSTAR) database [5-6]. The size of the SAR images is 128×128 pixels, and the resolution of the images is 0.3m in both of the range and azimuth directions. Samples taken with the pitching angle 17° and 15° are used as the training samples and testing samples, respectively. Datasets description is given in Table 1. As can be seen from Figure 3, different from the optical images, different targets look quite similar in shape in SAR images. As a result, preserving the local structure of the datasets is of crucial importance to recognition. In the beginning, so as to weaken the negative influences from the background noise, subimages with size 50×50 pixels are extracted. What follows is to normalize the subimages to attenuate the uneven scattering of SAR images.

Table 1. Datasets description.

Training (17°)	BMP2			BTR70	T72		
	9563	9566	C21	C71	132	812	S7
Number of images	233	232	233	233	232	231	228
Testing (15°)	BMP2			BTR70	T72		
	9563	9566	C21	C71	132	812	S7
Number of images	195	196	196	196	196	195	191

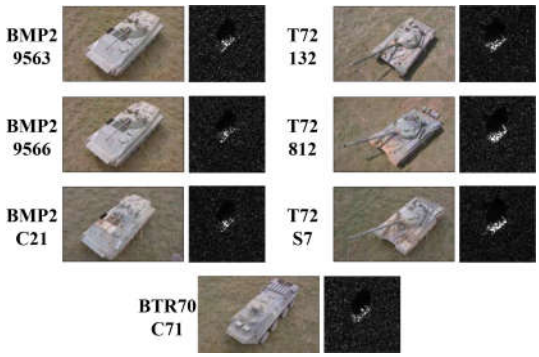

Figure 3. The optical and SAR images of the targets.

To show the effectiveness and advantages of the proposed MMSR algorithm, another three algorithms are taken as competitors with different feature extraction methods (the random projection matrix [8], the LPP algorithm [7], and the single-manifold based sparse representation (SMSR) algorithm [5]). All the four algorithms are based on the SR technique due to its nature discriminative power [8]. Recognition results are displayed in Table 2 and the curves of the probability of recognition error versus dimensionality is demonstrated in Figure 4.

As can be seen from the experimental results, preserving the local structure of the datasets is helpful to recognition. Utilizing LPP to realize feature extraction can achieve better recognition results than the random projection matrix. Actually, random projection matrix simply realizes dimensionality reduction, and more than 90% accurate recognition rate most relies on the discriminative power of SR. No structure information is captured or preserved by using the random matrix. LPP can preserve the local structure of the datasets before and after feature extraction, and therefore resulting in better performance.

Comparing LPP with SMSR, we can further tell the fact that preserving the structure of the datasets in different space shows different performance. LPP realizes structure preserving between the high-dimensional space and the low-dimensional space. And SMSR realizes structure preserving between the high-dimensional space and the reconstruction space.

As for the proposed MMSR, it has an explicit advantage over the competitors. In MMSR, the testing sample has been projected onto various manifolds instead of only one. Besides the advantages of SMSR, we can further get a better description of the testing sample in an established manifold that better fits the testing sample. Corrupted relationships are avoided to contribute to better recognition results.

Table 2. Configuration recognition results under different algorithms.

Algorithms	64	128	256	512	1024
Random matrix (%)	44.54	66.67	78.17	86.74	90.11
LPP (%)	45.79	68.79	81.90	90.62	93.11
SMSR (%)	47.99	72.31	84.69	92.75	95.46
MMSR (%)	48.79	72.89	86.15	93.99	96.34

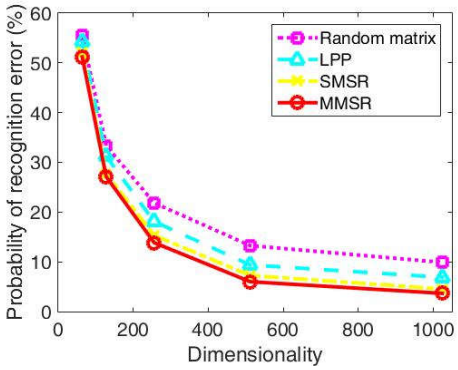

Figure 4. The curves of probability of recognition error versus dimensionality.

4 Conclusion

A multi-manifold based SR algorithm which can guarantee the local structure of the datasets in the reconstruction space is proposed for SAR target configuration recognition. The samples that are close to each other in the high-dimensional space will still be close to each other in the reconstruction space, where recognition is implemented. Important information for recognition is preserved to realize high-precision configuration recognition. The proposed MMSR can avoid corrupted relationships of the datasets. Better discriminating power of the proposed MMSR leads to overwhelming recognition performance.

5 Acknowledgements

This research was funded by the National Natural Science Foundation of China, grant number 61701289, the Young Talent Fund of University Association for Science and Technology in Shaanxi, grant number 20190106, the Fundamental Research Funds for the Central Universities grant number GK202103088, and the Special Support Program for High Level Talents of Shaanxi Province.

6 References

1. L. Chen, X. Jiang, Z. Li, X. Liu, and Z. Zhou, "Feature-Enhanced Speckle Reduction via Low-Rank and Space-Angle Continuity for Circular SAR Target Recognition," *IEEE Transactions on Geoscience and Remote Sensing*, **58**, 11, November 2020, pp. 7734-7752, doi: 10.1109/TGRS.2020.2983420.

2. J. Wang, J. Liu, P. Ren, and C. -X. Qin, "A SAR Target Recognition Based on Guided Reconstruction and Weighted Norm-Constrained Deep Belief Network," *IEEE Access*, **8**, 2020, pp. 181712-181722, doi: 10.1109/ACCESS.2020.3025379.

3. Z. Zhou, Z. Cao, and Y. Pi, "Subdictionary-Based Joint Sparse Representation for SAR Target Recognition Using Multilevel Reconstruction," *IEEE Transactions on Geoscience and Remote Sensing*, **57**, 9, September 2019, pp. 6877-6887, doi: 10.1109/TGRS.2019.2909121.

4. G. Dong, G. Kuang, N. Wang, and W. Wang, "Classification via Sparse Representation of Steerable Wavelet Frames on Grassmann Manifold: Application to Target Recognition in SAR Image," *IEEE Transactions on Image Processing*, **26**, 6, June 2017, pp. 2892-2904, doi: 10.1109/TIP.2017.2692524.

5. M. Liu, S. Chen, J. Wu, F. Lu, X. Wang, and M. Xing, "SAR Target Configuration Recognition via Two-Stage Sparse Structure Representation," *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 4, April 2018, pp. 2220-2232, doi: 10.1109/TGRS.2017.2776600.

6. X. Huang, X. Nie, W. Wu, H. Qiao, and B. Zhang, "SAR Target Configuration Recognition Based on the Biologically Inspired Model," *Neurocomputing*, **234**, April 2017, pp. 185-191, doi: 10.1016/j.neucom.2016.12.054.

7. X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face Recognition Using Laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, 3, March 2005, pp. 328-340, doi: 10.1109/TPAMI.2005.55.

8. J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**, 2, February 2009, pp. 210-227, doi: 10.1109/TPAMI.2008.79.