Int. J. Electron. Commun. (AEÜ) 83 (2018) 523–532

Contents lists available at ScienceDirect

Int. J. Electron. Commun. (AEÜ)

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

Regular paper

SAR target configuration recognition via structure preserving dictionary learning

Ming Liu[a,b], Shichao Chen[c,*], Fugang Lu[c], Jun Wang[c]

[a] Key Laboratory of Modern Teaching Technology, Ministry of Education, Xi'an 710119, China
[b] Shaanxi Normal University, School of Computer Science, Xi'an 710119, China
[c] Xi'an Modern Control Technology Research Institute, Xi'an 710065, China

ARTICLE INFO

Keywords:
Synthetic aperture radar (SAR) images
Dictionary learning
Target configuration recognition
Sparse representation

ABSTRACT

Learned dictionaries have been validated to perform better than predefined ones in many application areas. Focusing on synthetic aperture radar (SAR) images, a structure preserving dictionary learning (SPDL) algorithm, which can capture and preserve the local and distant structures of the datasets for SAR target configuration recognition is proposed in this paper. Due to the target aspect angle sensitivity characteristic of SAR images, two structure preserving factors are embedded into the proposed SPDL algorithm. One is constructed to preserve the local structure of the datasets, and the other one is established to preserve the distant structure of the datasets. Both the local and distant structures of the datasets are preserved using the learned dictionary to realize target configuration recognition. Experimental results on the moving and stationary target acquisition and recognition (MSTAR) database demonstrate that the proposed algorithm is capable of handling the situations with limited number of training samples and under noise conditions.

1. Introduction

Synthetic aperture radar (SAR) has been widely used in many application areas due to its overwhelming ability to work day and night under all weather conditions [1–3]. As one of the most important SAR applications, SAR target recognition has attracted increasing interest. Many effective algorithms have been proposed, and the algorithms can be mainly categorized into template-based ones and model-based ones. Template-based algorithms [4,5] suffer from computation burden and misalignment. As a result, model-based ones have become more and more popular due to their great capability of low storage, fast processing and robustness.

Lots of effective model-based algorithms have been proposed. The algorithms based on statistical models [6,7] will fail, when strong statistical relationship does not exist between the training samples and the testing samples. Many subspace based algorithms have been proposed for SAR target recognition [8–10]. However, the performance of these algorithms highly relies on the effectiveness of feature extraction [11]. Physical models of the targets are established for recognition, such as scattering center model [12,13], high range resolution profiles (HRRP) [14,15], etc. Although satisfying performance can be achieved, they suffer from heavy computation complexity. Besides, high-accuracy establishment of the physical models is tough. Geometrical characteristics of the targets like the shadow of the target [16], the length of the target [17] are also exploited for accurate recognition. However, these features are not robust to noise. Mathematical features, such as the nonnegative matrix factorization (NMF) [18], Krawtchouk moments [19], and Zernike moments [20] ignore the relationships among the samples. Nowadays, recognition via neural networks (NN) [21], convolutional neural network (CNN) [11], and deep learning [22] have been popular, however, parameter estimation for these algorithms is so complex that the computation burden is heavy. Moreover, these algorithms usually encounter local minimum and overfitting [23,24].

Sparse representation (SR) has shown its advantages in vast areas [25,26] including SAR target recognition. SR has been introduced to recognize targets in SAR images for the first time by Thiagarajan [27]. SAR target recognition is achieved by a decision fusion method of SR and support vector machine (SVM) [28]. A joint sparse representation based algorithm is proposed which focuses on the situation that multiple views of the same target can be obtained [29]. And a series of monogenic signal based sparse representation algorithms is proposed by Dong et al. [30,31]. In these algorithms, the dictionary used for recognition is comprised by the training samples or feature vectors extracted from the training samples. Many works show that learned dictionaries perform much better than predefined ones [27–31]. For SAR applications, dictionary learning has been validated to be effective

* Corresponding author.
E-mail address: rice0309@163.com (S. Chen).

https://doi.org/10.1016/j.aeue.2017.11.001
Received 30 June 2017; Accepted 1 November 2017

in various fields, such as image classification, ground moving target imaging, compression, and so on [32–34]. Therefore, we aim to realize SAR target configuration recognition with a learned dictionary in this paper.

Structure preserving is of great importance for recognition [9,10], especially for SAR images, which suffer from target aspect angle sensitivity [9,20]. Considering the significance of structure preserving and taking advantage of dictionary learning, a structure preserving dictionary learning (SPDL) algorithm for SAR target configuration recognition is proposed in this paper. Structure information of the datasets is exploited for recognition. Taking the target aspect angle sensitivity of SAR images into account, we fuse two structure preserving factors constructed based on the relationships among the samples into the proposed SPDL algorithm. Intrinsic structure of the datasets can be well preserved. For SAR images, differences among the samples are very explicit with different target aspect angles, although they belong to the same configuration. That is to say, for a given sample (a SAR image of a target), so as to describe it in a more precise way, we want to make the samples that have similar target aspect angles close to it, and we should also apart the samples that have different target aspect angles. Said another way, we would like to capture and preserve the structure of the datasets as complete as possible. The learned dictionary can guarantee the samples that are close to each other in the original space will also be close to each other in the sparse space (by the local structure preserving factor), and the samples that are far from each other in the original space will also be far from each other in the sparse space (by the distant structure preserving factor).

The proposed algorithm is capable of realizing not only target type recognition, but also target configuration recognition. Target configuration recognition can provide more detailed information of the targets than target type recognition, such as the tanks belong to the same type with/without machine guns or fuel barrels, the armored cars belong to the same type with/without splash guards or spotlights. Meanwhile, configuration recognition suffers from more difficulty due to its attempt to judge tiny differences between each other within the same target type [9,10].

The paper is organized as follows. In Section 2, the proposed SPDL algorithm is presented in detail, and the effectiveness of the proposed algorithm is tested on the moving and stationary target acquisition and recognition (MSTAR) database in Section 3. Finally, we conclude the paper in Section 4.

2. Structure preserving dictionary learning (SPDL) algorithm for SAR target configuration recognition

As is validated, preserving the structure of the datasets is of crucial importance for recognition [9,10]. The recognition results under various conditions can be significantly improved by incorporating geometrical structure of the datasets, especially for the SAR target configuration recognition which aims at discriminating tiny differences between targets belong to the same type. This is due to the fact that SAR images suffer from target aspect angle sensitivity [9,20]. Shapes of the targets in SAR images change dramatically with the variation of target aspect angles. The phenomenon is even severe for configurations, and thus resulting in much difficulty in high-precision recognition.

To overcome the problem of target aspect angle sensitivity, we fuse two structure preserving factors into the proposed dictionary learning algorithm to capture and preserve the intrinsic structure of the datasets. Different from the presented works [27–31], in which the dictionaries are predefined, we aim to get better performance under a learned dictionary in this paper. We would like to get a dictionary that can better capture and preserve the structure of the datasets to enhance SAR target configuration recognition performance.

Firstly, we come to see how to learn a dictionary using the training dataset. Define $\mathbf{X}_i = \{\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{iN_i}\}$ as the training dataset formed by the training samples of configuration $i(i = 1,2,...,C)$, where C is the number of target configurations. $\mathbf{x}_{ij}$ represents the $j^{th}(j = 1,2,...,N_i)$ training sample of configuration i, which is a column vector transformed from a SAR image, i.e., $\mathbf{x}_{ij}$ is the j^{th} column of dictionary $\mathbf{X}_i$, and N_i represents the number of the training samples with configuration i. Based on the training datasets, we can solve the following equation to get the sparse vectors [35].

$$\underset{\mathbf{\Phi}_i,\mathbf{A}_i}{\operatorname{argmin}} \|\mathbf{X}_i - \mathbf{\Phi}_i\mathbf{A}_i\|_F^2 + \eta \sum_{j=1}^{N_i} \|\boldsymbol{\alpha}_{ij}\|_1 \tag{1}$$

where $\mathbf{A}_i = \{\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2}, ..., \boldsymbol{\alpha}_{iN_i}\}$ is the sparse vector set, $\boldsymbol{\alpha}_{ij}$ is the sparse vector corresponds to $\mathbf{x}_{ij}$, $\mathbf{\Phi}_i$ is the dictionary to be learned, η is a constant that controls the sparsity, $\|\bullet\|_1$ represents the L1 norm, and $\|\bullet\|_F$ represents the Frobenius norm.

2.1. Construction of the local structure preserving factor

In this paper, we aim to capture and preserve the structure of the datasets to enhance recognition performance. Two structure preserving factors are fused into the dictionary learning model. Firstly, we construct the local structure preserving factor to preserve the local structure of the datasets.

$$\min \sum_{p=1}^{N_i} \sum_{q=1}^{N_i} \|\boldsymbol{\alpha}_{ip} - \boldsymbol{\alpha}_{iq}\|_2^2 \mathbf{S}_{pq}^i \tag{2}$$

where $\boldsymbol{\alpha}_{ip}, \boldsymbol{\alpha}_{iq} \in \mathbf{A}_i$, $\|\bullet\|_2$ represents the L2 norm, $\mathbf{S}^i$ is the similarity matrix, and each entry $\mathbf{S}_{pq}^i$ is given by

$$\mathbf{S}_{pq}^i = \begin{cases} \exp(-\|\mathbf{x}_{ip} - \mathbf{x}_{iq}\|_2 / t_1), & \mathbf{x}_{ip} \in N_k(\mathbf{x}_{iq}) \text{ or } \mathbf{x}_{iq} \in N_k(\mathbf{x}_{ip}) \\ 0, & \textit{otherwise} \end{cases} \tag{3}$$

where t_1 is a constant, $N_k(\mathbf{x}_{ip})$ represents the k nearest neighbors of $\mathbf{x}_{ip}$, whereas $N_k(\mathbf{x}_{iq})$ represents the k nearest neighbors of $\mathbf{x}_{iq}$, $\mathbf{x}_{ip}, \mathbf{x}_{iq} \in \mathbf{X}_i$. From (3), we can tell that if $\mathbf{x}_{ip}$ is one of the k nearest neighbors of $\mathbf{x}_{iq}$ or $\mathbf{x}_{iq}$ is one of the k nearest neighbors of $\mathbf{x}_{ip}$, then a weight calculated by (3) will be imposed on them. In other words, if two samples are close to each other in the original space, then corresponding weight will be imposed on them to reflect their relationships. The closer the samples are, the bigger the corresponding weight will be. A larger weight will lead to a smaller $\|\boldsymbol{\alpha}_{ip} - \boldsymbol{\alpha}_{iq}\|_2^2$ to minimize (2). Local structure of the datasets can be well captured and preserved by the local structure preserving factor.

2.2. Construction of the distinct structure preserving factor

By using (3), we can capture and preserve the local structure of the datasets. However, just preserving the local structure of the datasets might not be enough. We need another factor to further enhance structure preserving. This is due to the fact that, samples that are far from each other in the original space may become close to each other in the sparse space. Hence, we also need to preserve the distant structure of the datasets. That is to say, we should guarantee that samples that are far from each other in the original space still be far from each other in the sparse space. The distant structure preserving factor to achieve this goal is constructed as

$$\max \sum_{p=1}^{N_i} \sum_{q=1}^{N_i} \|\boldsymbol{\alpha}_{ip} - \boldsymbol{\alpha}_{iq}\|_2^2 \mathbf{D}_{pq}^i \tag{4}$$

where $\mathbf{D}^i$ is the difference matrix, and each entry $\mathbf{D}_{pq}^i$ is given by

$$\mathbf{D}_{pq}^i = \begin{cases} \exp(-\|\mathbf{x}_{ip} - \mathbf{x}_{iq}\|_2 / t_2), & \mathbf{x}_{ip} \in F_m(\mathbf{x}_{iq}) \text{ or } \mathbf{x}_{iq} \in F_m(\mathbf{x}_{ip}) \\ 0, & \textit{otherwise} \end{cases} \tag{5}$$

where t_2 is a constant, $F_m(\mathbf{x}_{ip})$ represents the m furthest neighbors of $\mathbf{x}_{ip}$, whereas $F_m(\mathbf{x}_{iq})$ represents the m furthest neighbors of $\mathbf{x}_{iq}$. From (5), we can see that if $\mathbf{x}_{ip}$ is one of the m furthest neighbors of $\mathbf{x}_{iq}$ or $\mathbf{x}_{iq}$ is one of the m furthest neighbors of $\mathbf{x}_{ip}$, then a weight calculated by (5) will be