Neurocomputing 196 (2016) 125–132

Contents lists available at ScienceDirect

Neurocomputing

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

SAR image target recognition via Complementary Spatial Pyramid Coding

Shaona Wang*, Licheng Jiao, Shuyuan Yang, Hongying Liu

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi 710071, China

ARTICLE INFO

Article history:
Received 7 February 2015
Received in revised form
15 December 2015
Accepted 3 February 2016
Available online 24 March 2016

Keywords:
Synthetic aperture radar (SAR)
Automatic target recognition
Complementary Spatial Pyramid Coding
Sparse coding

ABSTRACT

Many works have been recently presented to extract efficient features for automatic target recognition of synthetic aperture radar (SAR) images. However, they are limited in the discriminative ability of similar targets and robustness to the remarkable speckle noises and background clutters existed in images. In this paper, we propose a Complementary Spatial Pyramid Coding (CSPC) approach in the framework of Spatial Pyramid Matching (SPM). Both the coding coefficients and coding residuals are explored to develop more discriminative and robust features for representing SAR images. Multiple codebooks are first built from some training example images, where each codebook is formulated by local features of a certain class of samples. Then multiple sparse coding models are developed to derive features of a target under these codebooks. Additionally, these coding residuals are further sparsely encoded in the same way to that of local features. Finally, the encoded local features and the residual features are pooled according to spatial pyramid respectively, then concatenated to form the complementary features for the subsequent classification. The experiments on Moving and Stationary Target Acquisition and Recognition (MSTAR) public database verify the superior performance of the proposed method to some related approaches.

1. Introduction

Automatic target recognition (ATR) systems are used to identify one or a group of target objects in a scene. These ATR systems are to detect and classify targets using various image and signal processing techniques. Due to the all-weather ability and robustness towards environmental condition of synthetic aperture radar (SAR), researchers have drawn much attention on the automatic target recognition (ATR) based on SAR images in recent years, particularly in civilian and military fields [1,2]. However, the target in SAR images are sensitive to the variation in the pose and the speckle noise, as shown in Fig. 1. How to detect and recognize the specified targets in the SAR images still remains to be studied and explored.

The available SAR ATR approaches can be mainly divided into two groups: template based methods [3] and model based methods [4,5]. Template based methods need to generate a large number of templates, taken from the images of target in the dataset according to different poses. Candidate targets are classified based on the correlation with the templates. However, the computational burden of the template-based methods is huge and hard to be fast implemented. Besides, the matching is difficult due to the clutter background as well as the variations in pose and depression [6]. Moreover, SAR images are very sensitive to different articulations or configurations of the targets, which may lead to images with large differences. To alleviate these problems, many models based methods [7,8] have been proposed to solve SAR ATR problem. The key step of these methods is the effective feature extraction, such as Principal Component Analysis (PCA) [8] and Linear Discriminant Analysis (LDA) [9] and so on. PCA and LDA are very popular feature extraction techniques which can overcome the change of background and the interference of speckle noise. However, they can only extract the linear features. In order to make up this shortage, there are some nonlinear feature extraction methods including spectral feature analysis (SFA) [10] and gaussian process latent variable model (GPLVM) [11] etc. In [10], a nonlinear feature extraction algorithm based on the multiparameter spectral clustering has been proposed for SAR automatic target recognition. Compared with the PCA, the algorithm produced promising results in terms of recognition accuracy.

Recently, many successful midlevel feature extraction algorithms have been proposed [12–15] to connecting low-level local features and high-level concepts. First, low-level local feature descriptors such as SIFT [16] and HOG [17] are extracted to capture distinctive details of the images. However, the local descriptors are

* Corresponding author.
E-mail address: shaonaw@163.com (S. Wang).

http://dx.doi.org/10.1016/j.neucom.2016.02.059
0925-2312/

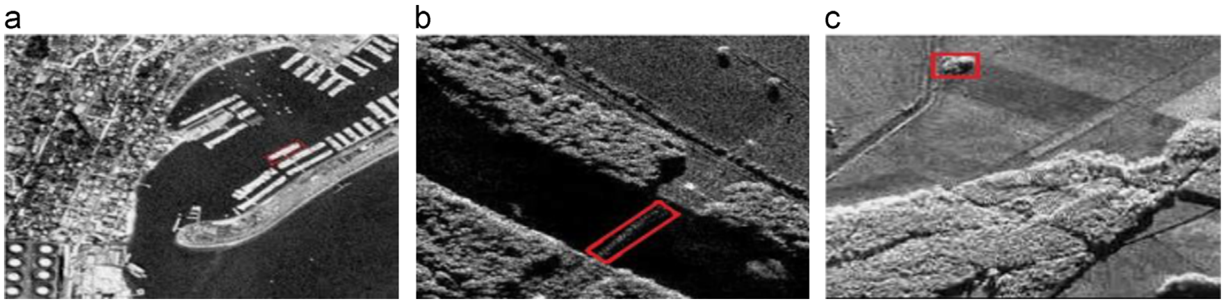

Fig. 1. Illustration of targets in the ATR task. The targets are marked in the red rectangle frames. (a) Vessel. (b) Bridge. (c) Tank. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rarely directly fed into image classifiers due to the computational complexity and their sensitiveness to noise. A common strategy is to integrate the local descriptors into a midlevel feature to represent the images. One of the most popular algorithms is the spatial pyramid matching (SPM) model for image classification. In this method, local feature descriptors are encoded using a learned codebook to compute a rough geometric correspondence between local features in the image. However, there still exists drawback in the framework that the Vector Quantization (VQ) which conducts hard assignment is used in the coding process. This results in loss of meaningful information and limits us to learn the discriminative classification representation.

Therefore, various schemes have been proposed to overcome the drawback. Yang et al. [12] proposed the ScSPM method where sparse coding (SC) was used to accurately represent the input features instead of vector quantization (VQ). In [18] and [19], the spatial weighting methods were proposed, which could suppress the influence of background clutter. To capture the discriminative information lost in traditional coding methods, Wang et al. [20] proposed the linear distance coding (LDC) method where the robust image-to-class distance was employed. Benefiting from alleviating the discriminative information lost in the quantization process, LDC has achieved competitive classification performance on multiple datasets. In particular, Knee et al. [15] extended the spatial pyramid method by applying the sparse representation to describe the local features in the framework of SPM for the SAR images target classification.

However, the representation capability still restricts the SAR image target recognition. The one reason may be that the characteristic of the synthetic aperture radar (SAR) images are sensitive to the target aspect angles. The sensitivity of SAR images to the aspect angles results in diversity of the images in the same class. The local feature coding using the codebook generated from the whole training data will bring too much noise. Combining with the characteristics of SAR images, we construct multiple codebooks, where each codebook is formulated by local features of a certain class. Then, the local features are encoded using these codebooks respectively to force the local features to be close to their corresponding classes, while to be far away from other classes [21].

The other reason could be the unavoidable loss of the discriminative information in the coding process, which limits the representation capability. Inspired by the success of LDC method and the sparse representation for automatic target classification in SAR images, the residuals could capture the discriminative information lost in the sparse coding process. In order to improve the discriminative power of the learned local features, we apply the multiple features sparse coding to represent the images [22]. In addition to the local feature, the coding residual is encoded for capturing the lost discriminative information. The residual coding plays a complementary role to the original local feature coding. The recognition power of combination of them would outperform each individual of them.

Here we provide an analysis on the importance of the residuals in the sparse coding process. Because the reconstruction residuals can be used to classify the image to the class achieving minimum residual in sparse representation [23,24], it indicates that the residual could capture the discriminative information. To verify the discrimination of the residuals, we apply sparse coding to the SAR images. Fig. 2 shows the results with example images per class. Fig. 2(a) represents the targets in the center of the original images. Fig. 2(b) is the reconstructed images using the sparse coding coefficients. Fig. 2(c) is the obtained sparse coefficients. Fig. 2(d) is the reconstructed images using the residuals. Fig. 2(e) is the obtained residual coefficients corresponding to each class. It is obvious that the residuals can identify the category of the example images when the coding coefficients can validate the target [25]. This demonstrates that the residuals are discriminative to represent the targets to some extent.

In this paper, we propose a Complementary Spatial Pyramid Coding (CSPC) approach which takes the advantage of the robustness of the spatial pyramid sparse coding model and meanwhile captures the discrimination of the residual in the coding process. The key point is the two step sparse coding which regards residuals obtained in the sparse coding of the local features as a kind of discriminative features. These residuals are further encoded into sparse codes to capture the salient information of the image. Then, we directly concatenate the resulting image representations of local features and residual features, and yield remarkable improvement of the performance as expected. Since the characteristic of the synthetic aperture radar (SAR) image's sensitivity to the target aspect angles, codebooks are constructed by using the samples of each class in order to avoid bringing noise and to better capture the detail information of the SAR images in the quantize/encode step.

The main contributions of this work can be summarized as follows:

1) Because the synthetic aperture radar (SAR) images are sensitive to the target aspect angles, the variation of the intra-class may be larger than the inter-class. To reduce the interference of noise and get more discriminative local image representation, we construct a codebook for each class which utilizes the prior label information in the step of sparse coding.
2) We propose a Complementary Spatial Pyramid Coding (CSPC) approach, which conducts the linear coding and max-pooling on the residual vectors to collect the lost meaningful information of images. Compared with the SPM methods, such process can avoid the desired information missing in the step of local feature coding.

The rest of this paper is organized as follows. Section 2, the residual feature descriptors and the combination with the original