Knowledge-Based Systems 65 (2014) 50–59

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A novel particle swarm optimization trained support vector machine for automatic sense-through-foliage target recognition system

Shijun Zhai *, Ting Jiang

Key Laboratory of Universal Wireless Communication, Ministry of Education, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

ARTICLE INFO

Article history:
Received 30 May 2013
Received in revised form 21 January 2014
Accepted 4 April 2014
Available online 19 April 2014

Keywords:
Target recognition
Support vector machine
Adaptive chaos particle swarm optimization
Sparse representation
Principal component analysis

ABSTRACT

Detection and recognition of targets that are embedded in foliage is of interest to both military and civilian research. Due to multipath propagation effects of rough surfaces, scattering from trees and ground tend to overwhelm the weak backscattering of targets, which makes it more difficult for sense through foliage target detection and recognition. In this paper, a novel intelligent recognition model based on support vector machine (SVM) and novel particle swarm optimization for sense through foliage targets recognition is proposed. SVM is a powerful novel tool for solving the recognition problem with small sampling, nonlinearity and high dimension. A new adaptive chaos particle swarm optimization (ACPSO) is developed in this study to determine the optimal parameters for SVM with the highest accuracy and generalization ability. Moreover, the measured real target echo signals are processed using sparse representation. Principal component analysis (PCA) is performed to extract the features of targets. Then, a hybrid feature selection is used to remove the redundant and irrelevant information of the features. The computational results on different real measurement datasets validate the effectiveness of the proposed approach.

1. Introduction

The efficient and accurate detection and recognition of target that are embedded in foliage provides a broad range of applications, such as detecting intrusion for security systems. Currently, the detection and recognition of targets, such as human and weapons that are hidden in foliage is still remains a challenging issue due to the presence of large amount of unwanted echo echoes. Scattering from the tree trunk and movement of the branches make foliage as a time varying and rich scattering environment. Therefore, it is difficult to detect and recognize the target in such environments.

In recent years, there have been many efforts undertaken to investigate foliage penetration and sense through foliage target detection. Sheen et al. [1] measured one-way transmission properties of foliage using a bistatic and coherent wide-band system. Fleischman et al. [2] made measurements of two-way foliage attenuation by synthetic aperture radar (SAR) and discussed probability dependency for frequency, polarization and depression angle. Based on Ultra-wide band (UWB) radar, Kapoor et al. [3] proposed an alpha-stable model while Liang et al. [4–6] presented a log–logistic model for foliage clutter. These studies have showed that foliage contains many spikes and angular fluctuations, and it's difficult to achieve effective and accurate target detection and recognition in forest. Moreover, Liang et al. [7–10] investigated sense through foliage target detection using signal processing and UWB radar sensor network system. But these systems were not able to provide investigations on different target recognition. Nowadays, various sense-through-foliage target detection and recognition systems based on machine learning have been proposed [11,12], canonical and particle swarm optimization optimized SVM are applied as classifiers. However, the recognition performance of these classifiers is greatly dependent upon the selection of SVM parameters and the method has a low convergence rate and sometimes is easy to trap into local optimum.

Although target echo signal contains useful information for the type of target, it is difficult for us to make accurate judgments due to the huge amount of data. With the development of the artificial intelligence, neural network as an intelligent technology has been widely applied to the automatic target recognition system [13,14]. However, neural network is based on the traditional empirical risk minimization which may suffer from several drawbacks, such as

* Corresponding author. Address: P.O. Box 96, Beijing University of Posts and Telecommunications, No. 10 Xi Tu Cheng Road, Beijing 100876, China.

E-mail address: zhaishijun@bupt.edu.cn (S. Zhai).

http://dx.doi.org/10.1016/j.knosys.2014.04.005
0950-7051/© 2014 Elsevier B.V. All rights reserved.

over-fitting and slow convergence velocity and easily relapsing into local extremum. Therefore, neural network is difficult to establish a global optimal model and thus lack the generalization capability. Nowadays, support vector machine (SVM) [15] and extreme learning machines (ELM) [16] have attracted a lot of research interests and yielded many real-world applications due to their good generalization performance. These two methods are both frequently used intelligent algorithms, and they may be used interchangeably in many practical cases. But, SVM does well in the case of small sample size and high dimension, it is frequently used in case of there is not enough samples obtained, such as target recognition and human action recognition.

Support vector machine, developed by Vapnik, is a novel machine learning method based on statistical learning theory. Based on the idea of structural risk minimization principle, SVM minimizes the empirical risk of the training samples, can overcomes the problems of over-fitting, failing local minimum and low-convergence rate existing in neural network and does well in the case of small sample size and high dimension. In the SVM, the crucial operation is that a kernel function is used to map the input data into a higher dimensional feature space so that the classification problem becomes linearly separable. There are many kinds of existent SVM kernels such as linear kernel, the polynomial kernel, and the radial basis function kernel. Since radial basis function kernel has excellent non-linear classification ability, radial basis function kernel is frequently used in the SVM. Nowadays, SVM has been successfully applied to numerous nonlinear classification and pattern recognition problems such as target detection and recognition [17,18], electrocardiogram signal classification [19], and fault diagnosis [20]. However, in those applications, the performance of SVM is greatly dependent upon the selection of SVM parameters, but the appropriate SVM parameters is very difficult to select due to the lack of the corresponding theoretical basis. In order to solve the effect for classification accuracy of SVM due to the selected improper parameters of SVM, particle swarm optimization and chaos particle swarm optimization have been applied to select the proper parameters of SVM [19,20], but the particle swarm optimization method is easy to trap into local optimum and the chaos particle swarm optimization method has slow convergence. An new adaptive chaos particle swarm optimization (ACPSO) based on chaos queues and adaptive weight and acceleration coefficients which improve the convergence speed and the abilities of searching for the global optima is adopted to select the parameters of SVM.

On the other hand, for target recognition, feature extraction and type recognition are two crucial steps. As we know, the radar target echo signals contain information about the targets. However, the target echo signals are very complicated with the features of non-linear and non-stationary processes in foliage environment. In this study, to extract target features from target echo signals, target echo signals are processed with sparse representation (SR) [12,21,22]. A set of overcomplete basis [23] can represent the essential information in a signal using a very small number of non-zero elements. Useful information can be revealed in the resulting sparse representations while redundant information can be filtered. Consequently, SR is a suitable method for target feature extraction. Sparse coefficients are used to describe target features. The dimension of the sparse coefficients is reduced using principal component analysis (PCA) [24]. The extracted sparse feature set contains irrelevant or redundant features as well as significant features. However, irrelevant and redundant features can spoil the performance of the SVM classifier and increase the computation burden. To further improve the recognition accuracy and reduce the dimension of features, a hybrid feature selection is used to further reduce the dimension of features; the significant features are selected as input of classifiers to automatically recognize the types of target.

According to the above analysis, the proposed SVM model is applied to sense-through-foliage target recognition in this paper, in which ACPSO is used to optimize the parameters of the SVM. The remainder of this paper is organized as follows: Section 2 provides a brief introduction to the theory of SVM. In Section 3, parameters selection of SVM with ACPSO is presented. Section 4 gives the proposed target recognition methodology. The experimental results are reported in Section 5. Finally, the conclusion and future work are presented in Section 6.

2. Support vector machine

SVM is a learning machine for two-class classification problems. The purpose of SVM classification is to find optimal separating hyperplane by maximizing the margin between the separating hyperplane and the data. Let $\{x_l, y_l\}_{l=1}^{L}$ be the training sample set, where x_l is the input vector, $y_l \in \{-1, +1\}$ is the class labels. SVM try to find an optimal hyperplane $f(x) = 0$ that separates the given data when two classes are linearly separable.

$$f(x) = w \cdot x + b = 0 \tag{1}$$

where w denotes the weight vector, and b denotes the bias term. w and b are used to define the position of separating hyperplane. The optimal hyperplane can be found by solving the following constrained optimization problem:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\|w\|^2 \\ &\text{Subject to} \quad y_l\langle w \cdot x_l + b\rangle \geqslant 1, \quad l = 1, 2, \ldots, L \end{aligned} \tag{2}$$

In the case where the two classes can be separated, the SVM determines the separating hyperplane which maximizes the margin between the two classes. Typically, most practical problems involve classes which are not separable. In this case, the SVM classifier is obtained by solving the following optimization problem:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{l=1}^{L}\xi_l \\ &\text{Subject to} \quad \begin{cases} y_l\langle w \cdot x_l + b\rangle \geqslant 1 - \xi_l, & l = 1, 2, \ldots, L \\ \xi_l \geqslant 0, & l = 1, 2, \ldots, L \end{cases} \end{aligned} \tag{3}$$

where the ξ_l are slack variables that allow the SVM to tolerate misclassifications and C is the penalization parameter, which used to control the trade-off between minimizing training errors and model complexity. By the Lagrangian multipliers algorithm, the above-mentioned optimization problem is transformed into the dual quadratic optimization problem, that is:

$$\begin{aligned} &\text{Maximize} \quad \sum_{l=1}^{L}\alpha_l - \frac{1}{2}\sum_{l,m=1}^{L}\alpha_l\alpha_m y_l y_m (x_l \cdot x_m) \\ &\text{Subject to} \quad \sum_{l=1}^{L}\alpha_l y_l = 0, C \geqslant \alpha_l \geqslant 0, \quad l = 1, 2, \ldots, L \end{aligned} \tag{4}$$

The linear decision function is created by solving the dual optimization problem, which can be written as

$$f(x) = sign\left(\sum_{l=1}^{L}\alpha_l y_l (x_l \cdot x) + b\right) \tag{5}$$

SVM can also be used in non-linear classification by using kernel function $K(x_l, x_m)$. The value of $K(x_l, x_m)$ equals to $\varphi(x_l) \cdot \varphi(x_m)$, where $\varphi(\cdot)$ is the transformation function, which makes the input data into high-dimensional feature space. Then, the non-linear decision function is described as below: