Neurocomputing 205 (2016) 341–348

Contents lists available at ScienceDirect

Neurocomputing

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

Cucumber disease recognition based on Global-Local Singular value decomposition

Shanwen Zhang, Zhen Wang

Department of Electronics and Information Engineering, XiJing University, Xi'an 710123, China

ARTICLE INFO

Article history:
Received 29 December 2015
Received in revised form
21 February 2016
Accepted 19 April 2016
Available online 10 May 2016

Keywords:
Cucumber disease leaf
Cucumber disease recognition
Singular value decomposition (SVD)
Global-Local SVD (GL-SVD)

ABSTRACT

Plant leaf based plant disease recognition is becoming an important research topic in the pattern recognition and image processing. Many conventional methods are not effective for plant disease recognition because of the complexity of the disease leaf images. Singular value decomposition (SVD) has shown its advantage over other classical algorithms in terms of clustering accuracy and robustness to the data representation and feature extraction. Nonetheless, there is still considerable room to improve the practical application performance of SVD. To improve the recognition rate of cucumber disease, based on the Gglobal-Local SVD, a recognition method of cucumber disease is proposed. First, the spot image is segmented by the watershed algorithm from each cucumber disease leaf image. Second, each spot image is divided into a few blocks, and then the combining features of Global-Local singular values are extracted and ordered from each block by SVD. Third, the key-point vectors are constructed and their dimensionalities are adjusted to equal to each other. Finally, a SVM classifier is adopted to recognize the class of the unknown disease leaf image. Comparing with the existing cucumber disease recognition methods, the proposed one can extract the key-point features from the spot image whose dimension is significantly lower than that of the original space. Also, it can effectively avoid the key information loss in the usual stochastic feature extraction and selection process. The approach is tested on three kinds of cucumber disease leaf images. In the experiments, the original singular values are replaced by the nonnegative Global-Local Logarithmic singular values. The experimental results show that the proposed method is effective and feasible for the cucumber disease recognition, and has the highest recognition rate and more practical value.

1. Introduction

Cucumber is an important favorite vegetable in the world. But in the growth process of cucumber, due to the bad ecological environment, non-biological factors and the impact of various factors, cucumber diseases will seriously affect the yield and quality of cucumber and cause economic losses to farmers. Timely and accurate identification of the cucumber diseases is helpful to achieve scientific and rational use of pesticides, improve the cucumber quality and protect the environment. The cucumber diseases can be identified and judged by the eyes as soon as the leaves appear. But this identification method depends on the knowledge, experience level and subjective consciousness of agricultural technical personnel [1,2]. Cucumber disease detection from the unhealthy symptoms that appear on the cucumber leaves is an important research topic [3]. By now, many image processing methods have been used by many researchers for cucumber disease leaf image detection and analysis, such as Support Vector Machine (SVM), neural network, Color Histogram, Edge Histogram (EH), Adaptive Thresholding (AT), Edge Direction Histogram Texture (EDHT), Contrast Stretching Technique (CST), HSV Color Space (HSV-CS) and Scale Invariant Feature Transform (SIFT), etc., and many detection methods of the cucumber leaf diseases have been proposed by the computer image processing technology and pattern recognition methods [4–7]. By extracting the characteristics of color, shape and texture of infected cucumber leaf lesion, many methods have been presented for agricultural producers to timely and accurately recognize the disease type and damage degree of cucumber. In order to deal with the high dimensional characteristics of cucumber leaf image, Gao et al. [8] proposed a method of approximate nearest neighbor detection algorithm based on a Kd-tree to overcome the inefficiency of the high dimensional vector detection. In the method, subspace data structure and data variance characteristic vector of cucumber disease image is calculated and a number of pixel values closer to the adjacent area are detected to analyze connecting regional, so a nearest neighbor

E-mail address: wjdw716@163.com (S. Zhang).

http://dx.doi.org/10.1016/j.neucom.2016.04.034
0925-2312/

chain is constituted which can detect the diseases of cucumber leaf image. Dong et al. [9] mainly studied the leaf diseases of cucumber downy mildew, powdery mildew and anthracnose by image processing and recognition technology. The procedures of filtering noise, segmenting the leaf spot disease of cucumber leaf, extracting the feature parameters of color, shape and texture from the lesion, are implemented. Based on the nearest neighbor method, the cucumber diseases are identified. The experimental result showed that the disease recognition accuracy rates of this method are more than 96%. Tian et al. [10] proposed a method of recognizing cucumber leaf diseases based on computer image processing and SVM. In the method, the features of color, shape and texture are extracted from the color spot image of cucumber, and SVM is used to recognize the cucumber diseases. The recognition results by using the shape and texture features is better than that of only using the shape feature. Guo et al. [11] presented an automatic recognition method for cucumber diseases. In the method, the color and texture features are extracted from the representative lesions of disease images, and Bayes classifier is applied to the cucumber disease recognition of powdery mildew, anthracnose, botrytis and downy mildew. Si et al. [12] studied the cucumber downy mildew by adopting image preprocessing, feature extraction, parameter optimization and pattern recognition method. In the method, the image samples are analyzed by the gray statistics, then color and geometry characteristics are extracted, and digitalized, and normalized. The genetic algorithm (GA) is used to optimize these parameter values after a large number of samples of training, a group of more accurate characteristic parameter values are obtained.

From the above analysis, it is cleared that, extracting and selecting the optimal features of color, shape and texture from the disease leaf images is very important in detecting cucumber diseases. Now, many researchers face the issue of long computational time with difficulty to extract a lot of features from the blurred disease leaf images. With so many selected features but so little consensus about which feature are better or worse, many practitioners lack guidelines to decide which features can be used. In pattern recognition community, there has been enormous effort and emphasis on finding the “optimal” features. Many methods were proposed to obscure the other important factors that can clarify the role of the feature selection in the overall pattern recognition process. Because of the variability and complexity of disease leaf image, many existing feature extraction based disease recognition methods have not remarkable effective. Singular value decomposition (SVD) is a most widely-used multivariate statistical technique [13–15]. Its purpose is to reduce the number of freedom degrees in a complex system which is to be modeled. SVD is applied extensively to the study of linear inverse problems, and is useful in the analysis of regularization methods. It is widely used in statistics where it is related to principal component analysis (PCA) and to correspondence analysis (CA), in signal processing (SP) and pattern recognition (PR). The singular values (or eigenvalues) by SVD represent the intrinsic properties of an image and have better stability, and are valid features in object recognition such as face recognition [16]. Many SVD based recognition methods were proposed by using the singular values as the feature vectors [17]. Although SVD is a really old and classical method, there is still considerable room to improve the practical application performance of SVD. Only in 2015 and 2016, many newly SVD based face recognition or other pattern recognition methods have been proposed [18–22]. It was known from these methods that SVD as feature extraction improved the performance of classification accuracy and this is proven since recognition rate accuracy is higher with SVD as feature extraction as compared to other feature extraction based recognition methods. The smaller the singular values are, the less energy along the corresponding eigenvector there is. Therefore, the smallest eigenvalues are often considered to be due to noise. Motivated by the recent progress and application of SVD, we propose a novel plant classification algorithm based on Global-Local SVD (GL-SVD). The singular values as the classifying features are extracted by SVD from leaf images, and then the recognition task is implemented by an optimal classifier. The highlights of the paper are (1) the global and local singular values of leaf images are applied to plant disease recognition; (2) we skillfully adjust the dimensionalities of two global and local singular vectors by constituting a classifying vector; (3) it is find that nonnegative Global-Local Logarithmic singular values have certain discriminant ability; (4) the key-point vector is constituted as the input of SVM classifier.

The rest of this paper is organized as follows. Section 2 introduces SVD. Section 3 proposes a plant classification method based on GL-SVD. Experimental results and analyses are given in Section 4. Section 5 concludes this paper and points out our future work on improving our proposed algorithm.

2. Related works

Formally, SVD of an $m \times n$ real matrix M is a following factorization

$$M = U\Delta V^{\mathrm{T}} \tag{1}$$

where U is an $m \times m$ real unitary matrix, Δ is an $m \times n$ rectangular diagonal matrix with non-negative real numbers on the diagonal, and V^{T} is an $n \times n$ real unitary matrix, especially, Δ is an m non-square matrix with zero entries everywhere except on the leading diagonal with elements arranged in descending order of magnitude.

The SVD of a matrix M is typically computed by a two-step procedure. Firstly, the matrix is reduced to a bidiagonal matrix [23]. Secondly, SVD of the bidiagonal matrix is computed by an iterative method. Suppose the diagonal entries $\Delta_{i,i}$ of Δ are the singular values of M. A common convention is to list the singular values in descending order. In this case, the matrix Δ is uniquely determined by M. Note that the matrices U and V are not unique. The m columns of U and the n columns of V are called the left-singular vectors and right-singular vectors of M, respectively. Then, Eq. (1) indicates that: (1) An $m \times n$ matrix M has at least one and at most $p = \min(m, n)$ distinct singular values; (2) It is always possible to find a unitary basis for R^m consisting of left-singular vectors of M; (3) It is always possible to find a unitary basis for R^n consisting of right-singular vectors of M; (4) Qualitatively U represents a vector basis for the most relevant information while $\Delta_{i,i}$ represents the variability in the information.

The matrix Δ is computed by the following procedure:

(1) M^T and $M^T M$ are computed.
(2) The eigenvalues of $M^T M$ are determined and sorted in descending order, in the absolute sense. The nonnegative square roots of these are the singular values of M.
(3) M is constructed by placing singular values in descending order along its diagonal.

It is known that the Rank of a Matrix is the number of nonzero singular values. Since Δ is a diagonal matrix, its non-diagonal elements are equal to zero. However, one would need to know $\boldsymbol{U}$ first, which we have not defined yet. Either way, the alternate expression for Δ is obtained by post- multiplying by $\boldsymbol{V}$ and pre-multiplying by M^T.