Applied Soft Computing 67 (2018) 164–171

Contents lists available at ScienceDirect

Applied Soft Computing

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

Combining sparse representation and singular value decomposition for plant recognition

Shanwen Zhang[a], Chuanlei Zhang[b,*], Zhen Wang[a], Weiwei Kong[a]

[a] Department of Information Engineering, Xijing University, Xi'an, 710123, China
[b] School of Computer Science and Information Engineering, Tianjin University of Science and Technology, 1038 Da Gu Nan Lu, Tianjin, 300222, China

ARTICLE INFO

Article history:
Received 28 August 2016
Received in revised form 4 January 2018
Accepted 25 February 2018
Available online 1 March 2018

Keywords:
Plant leaf image
Plant species recognition
Sparse representation (SR)
Singular value decomposition (SVD)

ABSTRACT

Plant recognition is one of important research areas of pattern recognition. As plant leaves are extremely irregular, complex and diverse, many existing plant classification and recognition methods cannot meet the requirements of the automatic plant recognition system. A plant recognition approach is proposed by combining singular value decomposition (SVD) and sparse representation (SR) in this paper. The difference from the traditional plant classification methods is that, instead of establishing a classification model by extracting the classification features, the proposed method directly reduces the image dimensionality and recognizes the test samples based on the sparse coefficients, and uses the class-specific dictionary learning for sparse modeling to reduce the recognition time. The proposed method is verified on two plant leaf datasets and is compared with other four existing plant recognition methods The overall recognition accuracy of the proposed approach for the 6 kinds of plant leaves is over 96%, which is the best classification rate. The experimental results show the feasibility and effectiveness of the proposed method.

1. Introduction

Plants play an important role and carry significant information in the development of human society [1]. With the environment deterioration, the urgent situation is that many plant species are on the risk of extinction. There are about 22% to 47% plant species of all known plants endangered, and there are about 100,000 to 150,000 plants that probably will die out in next few years. To protect environment and ecological balance, it is the first priority to protect plants and recognize plant species. Plant recognition is a critical problem especially for biologists, chemists and environmentalists [2,3]. It can be performed by human experts manually, but it is time-consuming and low-efficiency. So, automatic plant recognition is an important research area and many plant recognition methods have been proposed. Compared with the traditional methods such as cell and molecule biology methods, the leaf based plant recognition is a better choice. With a camera, smart phone and even Internet of things, it is easy, low-cost and convenient to collect and digitalize plant leaves, and extract features automatically via the image processing methods for the plant classification and recognition [4,5]. Du et al. [6] proposed a method to describe the characteristics of plant leaf images based on the outline fractal dimension and venation fractal dimension. Du et al. [7] also introduced a shape recognition method based on radial basis probabilistic neural network which is trained by orthogonal least square algorithm (OLSA) and optimized by recursive OLSA. The plant recognition is performed through the modified Fourier descriptors of the leaf shape. Jyotismita et al. [8] proposed a method of characterizing and recognizing plant leaves by combining the texture and shape features, where the texture features are modeled by Gabor filter and gray level co-occurrence matrix (GLCM), and shape features are captured using a set of curvelet transform coefficients together with invariant moments. Bama et al. [1] presented a content based image retrieval system by combining color and texture features. The texture and color features are extracted by wavelet transforms and color histogram, which are robust to scaling and translation of the leaf images. In fact, the combination of the shape features, color features, texture features and other features of the leaf images is very useful in leaf recognition [9,10]. It is easy to extract more than 100 kinds of features from each leaf image [9–12], but how to select the effective features from so many kinds of features still keeps an open problem.

Genetic Algorithm (GA) and Neighborhood Rough Set (NRT) are used to select minimal feature subset with high classification accuracy to improve the classification performance of the classification model, and reduce the redundant or irrelevant features [13]. Many

* Corresponding author.
E-mail address: 97313114@tust.edu.cn (C. Zhang).

https://doi.org/10.1016/j.asoc.2018.02.052
1568-4946/

dimensionality reduction methods, such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Locality Preserving Projections (LPP) and Locally Linear Embedding (LLE), are often applied to reduce the dimensionality of the extracted features. Valliammal et al. [14] described an optimal approach for feature subset selection to classify the leaf images based on GA and Kernel PCA (KPCA). Recent researches show that the high dimensional leaf images lie on a low-dimensional nonlinear manifold, and some nonlinear manifold learning methods have been proposed for the plant recognition [15–17]. Although several effective plant recognition methods have been proposed, the classical methods and manifold learning methods have some limitations as follows:

(1) Leaf image itself and the extracted features are sensitive to the illumination, orientation and scaling of leaf image, and a preprocessing step prior to the feature extraction is often applied to making correction for the random factors of translation, rotation and scaling.
(2) Each selected feature is regarded as being equally important in the classification task, but it is not.
(3) In manifold learning based recognition methods, the projections obtained from directly solving the eigen-equation are the nonzero linear combiner of all the image vectors, and each variable of the patterns (image vectors) is regarded as being equally important also.

Actually, the contribution of different leaf images and different features is not equal for the classification task. Some features might even degrade the recognition performance. In recent years, several sparse representation based classification (SRC) methods have been widely applied to face recognition, human movement recognition and tumor classification [18–21]. In SRC, the main idea is that each test sample can only be represented as a sparse linear combination of the training samples. In theory, a test sample can be well represented by a small number of the training samples from the same class. In fact, there is much noise in the real dataset, which always degrades the classification accuracy, and causes the discriminative features to be unobvious. Thus, the test samples cannot be represented effectively by some training samples from the different classes [22]. Wagner et al. [23] proposed a conceptually simple face recognition system that achieves a high degree of robustness and stability to illumination variation, image misalignment and partial occlusion. Zuo et al. [24] applied SRC to generic image classification with background clutter and scale, translation, and rotation variations within the same image class. Although SR is effective for data classification, it is difficult to implement SRC because of the high dimensionality of the original data. For example, in implementing classical SRC, a 128×128 image in a data recognition system has 16384 dimensions, which requires many training samples to ensure the effectiveness of SRC. A well-known approach to deal with this problem is two-step strategy, i.e., the data dimensionality is firstly pre-reduced using PCA, and then SRC is implemented. Many researchers have proved that singular value features by Singular Value Decomposition (SVD) are ideal feature description with four characteristics [25]: stability, transpose invariance, rotation invariance, shift invariance, image transformation invariance. SVD is applied widely in data compression and pattern recognition [26]. SVD based feature extraction methods not only weaken the impact of light and expression, but also reduce the dimensions and the computational complexity, and meanwhile retain most of the effective characteristics of plant leaf images for the following plant species recognition [27]. Motivated by the recent progress and success in SR and SVD, we propose a novel plant classification algorithm by combining SVD and SR. The singular value features are firstly extracted from the original leaf image, and then SR is used to represent the test samples, and the plant species is recognized by the SR coefficients. As for the classification tasks, the proposed method is different from the classical plant classification methods. Instead of extracting the classifying features of color, shape and texture from each leaf image, the singular value features can be directly extracted from the leaf images and the plant species are recognized according to the SR coefficients.

The rest of this paper is organized as follows. Section 2 describes SRC method. Section 3 introduces SVD. Section 4 proposes a plant recognition method by combining SVD and SR. Experimental results and analysis are given in Section 5. Section 6 concludes this paper and points out the future work to improve the performance of the proposed method.

2. Sparse representation based classification (SRC)

SRC has shown its effectiveness in face recognition experiment [23,24]. In order to ensure all samples to be sparsely decomposed by the same dictionary, many predefined and adaptive algorithms, such as wavelets, curvlets, contourlets, steerable wavelets filters, short-time Fourier transforms, etc., have been adopted to design the over-complete dictionary [28]. SRC is introduced as following.

Suppose n training samples $X = [x_1, x_2, \ldots, x_n]$ belong to k object classes, where $x_i \in R^m$, generally $m \ll n$, and then the whole data set can be re-expressed as $W = [W_1, W_2, \ldots, W_k] \in R^{m \times n}$, where $W_i = \left[v_{i,1}, v_{i,2}, \ldots, v_{i,n_i}\right]$, $v_{i,j}$ is the j^{th} sample of the i^{th} class, n_i is the sample number of the i^{th} class. Based on the theory of SR, a new test sample $y \in R^m$ of the i^{th} class is ideally laid in the linear span of the training samples W_i, denoted as follows,

$$y = \alpha_{i,1} v_{i,1} + \alpha_{i,2} v_{i,2} + \cdots + \alpha_{i,n_i} v_{i,n_i} \tag{1}$$

where $\alpha_{i,j} \in R$ is the sparse representation coefficients of $y, j = 1, 2, \ldots, n_i$.

Eq. (1) can be also rewritten as

$$y = Wx \tag{2}$$

where $x = [0, \ldots, 0, \alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n_i}, 0, \ldots, 0] \in R^n$.

To obtain a sparse x, SRC solves the following ℓ_1-minimization problem,

$$\begin{aligned} x' &= arg\min \|x\|_1 \\ \text{s.t.} \quad & Wx = y \end{aligned} \tag{3}$$

where x' is a coefficient vector, and $\|x\|_1 = \sum_i |x_i|$ denotes the l_1-norm.

Due to the existence of noise in y, Eq. (3) can not hold exactly in practical problem. Then a generalized version of Eq. (3) can be reconstructed as follows

$$J(x, \varepsilon) = \min_x \{\|Wx - y\|_2 + \varepsilon \|x\|_1\} \tag{4}$$

where $\varepsilon > 0$ is an adjusting parameter which can be determined experimentally, W is the dictionary of coding atoms, x is the coding vector of y over W, and $\|x\|_2 = \sqrt{\sum_i |x_i|^2}$ denotes the l_2-norm.

Eq. (4) aims to find the best matching coefficients by minimizing the sum of squared errors, which can be solved by the truncated Newton interior-point method [30,31]. The parameter ε can balance the degree of noise and ensure that $(W^T W)^{-1}$ is valid. Eqs. (2)–(4) mean that the test samples can be approximately represented by a linear combination of all training samples. The solution of Eq. (4) can be expressed as follows

$$x' = (W^T W + \varepsilon I)^{-1} y \tag{5}$$

where I is an identity matrix.

Least-angle regression (LARS) is an algorithm for fitting linear regression models to high-dimensional data [21], which can pro-