Computers and Electronics in Agriculture 155 (2018) 150–156

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Integrating leaf and flower by local discriminant CCA for plant species recognition

Shanwen Zhang[a], Chuanlei Zhang[b,*], Wenzhun Huang[a]

[a] Department of Information Engineering, Xijing University, Xi'an 710123, China
[b] College of Computer Science and Information Engineering, Tianjin University of Science and Technology, Tianjin 300222, China

ARTICLE INFO

Keywords:
Canonical correlation analysis (CCA)
Local discriminant CCA (LDCCA)
Multi-modal plant species recognition
modified LDCCA (MLDCCA)

ABSTRACT

Plant species recognition using a single organ, such as flower and leaf, is not sufficiently reliable, because different species may have very similar flowers or leaves, while the same species may have rather different flowers or leaves. Combining leaves and flowers to recognize plant species can produce positive results. Based on multi-modal learning scheme, an automatic plant species recognition method is proposed by combining leaves and flowers of plant. In the method, a modified local discriminant canonical correlation analysis (MLDCCA) is designed by incorporating the idea of local discriminant embedding (LDE) into canonical correlation analysis (CCA). Firstly, two neighbor graphs are constructed based on the exploration of the manifold that the input data lie on. Then, two projection matrices for dimensionality reduction are obtained by making the within-class neighbor samples most correlated and between-class neighbor samples least correlated, and meanwhile keeping the correlation between leaves and flowers of the same species maximum. Finally, 1-nearest neighbor classifier with geodesic distance is used to recognize the plant species. MLDCCA is a powerful supervised multi-modal dimensional reduction method which can extract the discriminant features from two plant organs, meanwhile preserve the discriminant information and the data structure well. Experimental results on a real leaf and flower image dataset validate the effectiveness of the proposed method.

1. Introduction

Automatic plant species recognition is essential for protecting species diversity, which can help ordinary persons and botanists to recognize the various plant species more rapidly. Plant species can be recognized by their leaf, flower, skin, branch, seed, fruit and entire (Seeland et al., 2017), while leaf based automatic plant species recognition is an important topic in several fields, such as computer science, image processing, algorithm science, botanical science and machine learning, because plant leaf image has a lot robust characteristics and can be easily collected and processed. Nowadays, many efforts have been conducted in extracting various features from leaf for plant species recognition (Liu et al., 2018). Wäldchen and Mäder (2018) systematically reviewed a large number of plant species identification approaches, including 120 peer-reviewed studies published in the last 10 years from 2005 to 2015, and categorized these methods according to the studied plant organ, and the studied features, i.e., shape, texture, color, margin, and vein structure. Furthermore, they compared these methods based on classification accuracy achieved on the publicly available datasets. Zeng et al. (2017) presented a plant species identification method by combining the periodic wavelet descriptor (PWD) of plant leaf with back propagation neural network (BPNN). The experimental results show that the proposed algorithm is effective with a correct identification rate about 90%. Kho et al. (2017) proposed a Ficus species recognition system based on artificial neural network (ANN) and support vector machine (SVM). The results showed that the proposed system is effective to recognize leaf images with an accuracy of 83.3%. Zhang et al. (2017) proposed a two-stage local similarity based classification learning (LSCL) method by combining local mean-based clustering (LMC) and local sparse representation based classification (LWSRC). Zhu et al. (2017) proposed a plant species recognition method based on multi-path sparse coding using SIFT features, which utilizes five paths to model the shape and texture features of plant leaf images by learning the dictionaries with different sizes of hierarchical sparse coding. It is well known that the recognition rates of the above methods rely heavily on the handcraft features extracted from plant leaf images.

Recently, deep learning has led to a series of breakthroughs in many application fields like face recognition, palmprint identification and plant species classification. Deep learning can automatically learn the classification features from the original images. Lee et al. (2017) learned useful leaf features directly from the raw leaf images using

* Corresponding author.

https://doi.org/10.1016/j.compag.2018.10.018
Received 13 July 2018; Received in revised form 7 October 2018; Accepted 12 October 2018
Available online 17 October 2018

convolutional neural networks (CNN), and selected the features based on a deconvolutional network approach, and then reported three unexpected results: (1) different venation orders are the best representative features compared to those of outline shape, (2) multi-level representation of the leaf images are observed by demonstrating the hierarchical transformation of features from lower-level to higher-level abstraction, and (3) the extracted hybrid feature extraction models are able to further improve the discriminative power of plant classification systems. The experiment results validated that deep learning is a promising technology for plant species recognition (Sun et al., 2017b). Ghazi et al. (2017) proposed a plant species recognition method based on three deep CNN (DCNN) architectures, namely GoogLeNet, AlexNet, and VGGNet. Sun et al. (2017a) constructed a leaf image dataset by mobile phone in natural scene. The dataset contains 10,000 images of 100 ornamental plant species in Beijing Forestry University campus. They designed a 26-layer deep learning model consisting of 8 residual building blocks for large-scale plant classification in natural environment. Usually, deep learning model needs a large number of training samples to train its parameters and reduce over-fitting problem. However, it is difficult to collect too many available cucumber diseased leaf images. Also in deep learning methods need high performance computation resource and much time is taken to train the deep learning model.

Therefore, leaf based plant species recognition research is still a challenging topic, because there are large intra-class difference (as shown in Fig. 1A and B) and large inter-class similarity (as shown in Fig. 1C) in a huge number of leaves. From Fig. 1 B and C, it is difficult to discriminate between these confusing leaves. Fig. 1 indicates that the species identification accuracy using a single leaf is still limited with a large number of species.

The appearance of plant flowers is more stable and less variant with such kind of changes. So flowers can cooperate with leaves to enhance plant species identification performance. From Fig. 2, it is found that Ophiopogon japonicus and Chlorophytum have similar leaves but dissimilar flowers, while Rose and Chinese rose have similar flowers but dissimilar leaves. Fig. 2 indicates that some plant species cannot be recognized using a single organ. That is to say, if we take both flower and leaf organs into consideration, the confusable plant species can be correctly identified. It is meaningful to automatically identify plant species with multiple organs of plants because multiple organs can provide sufficient information to distinguish the confusable plant species in some cases.

Goeau et al. (2012) presented an interactive web application system for plant species identification by making use of several different plant organs and views including habit, flowers, fruits, leaves and bark. In the system, all training images were continuously collected during one year through a crowd sourcing application that was set up in the scope of a citizen sciences initiative. Nguyen et al. (2017) proposed a combination of deep learning and hand-designed feature for plant species recognition by fusing the identification results of leaf and flower. Experimental results indicated that deep learning is robust for plant species recognition in natural situations, and the combination of leaf and flower images can improve significantly the recognition rate comparing only

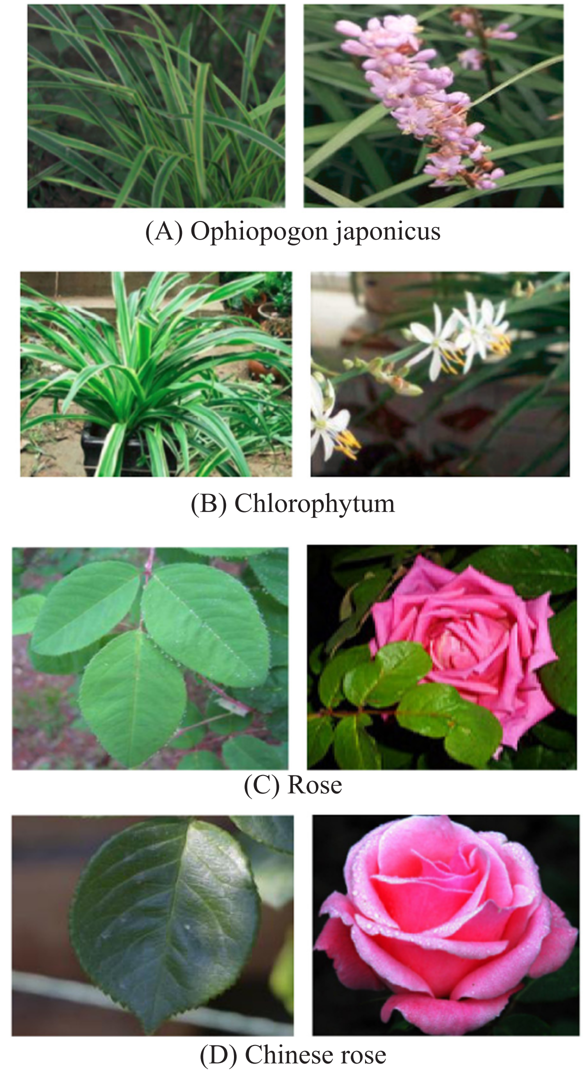

Fig. 2. Leaf and flower examples with similar visual appearance.

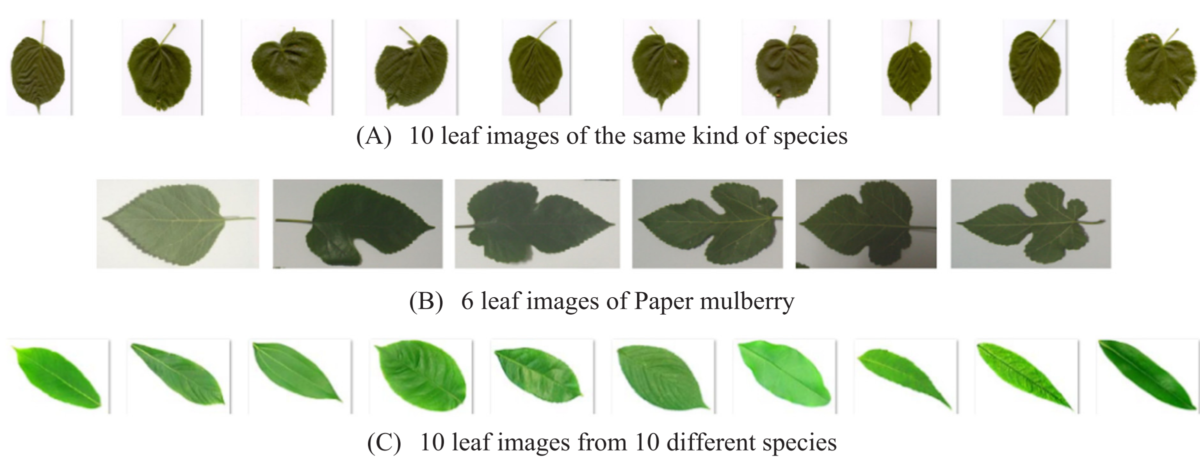

Fig. 1. Leaf examples.

leaf-based plant species recognition. Do et al. (2017) described a fusion technique for species identification from different plant organ images. They firstly extracted confidence scores from each single plant organ using a state-of-the-art DCNN, and then deployed various schemes of the fusion approaches including not only conventional transformation-based approaches (sum rule, max rule, product rule), but also a classification-based approach (SVM, K-nearest neighbor classifier). Finally, they proposed a hybrid fusion model for species recognition. Compared to other feed-forward neural networks, DCNN has achieved numerous significant improvements in recent years. However, it takes too much time, even several days, to train the parameters of DCNN.

In fact, the different plant organ images are nonlinear complex heterogeneous data, and how to combine the multi-organ images together efficiently to recognize species is a key problem in the existing multi-modal species recognize methods. Many supervised manifold learning methods, such as local maximal margin discriminant embedding (MMDE), local discriminant embedding (LDE) and their variants (Huang et al., 2014; Chen et al., 2005), can preserve effectively the discriminative information and the local structure of the data, which have been widely used for image classification. However, they are single-view learning methods, and they do not consider the relation between multi-modal datasets or multiple view feature sets. Canonical correlation analysis (CCA) is a well-known multi-view feature fusion method to extract valuable information from multi-modal datasets (McGarigal et al., 2012). However, it cannot exploit the discriminant information (Hardoon et al., 2014). Several discriminant modified models of CCA have been proposed for nonlinear complex image classification, such as local discriminant CCA (LDCCA) (Huang et al., 2017).

Inspired by LDCCA, MMDE and the improvement of plant species recognition methods, a modified LDCCA (MLDCCA) based plant species recognition method is proposed by combining plant leaf and flower. MLDCCA incorporates the idea of LDE into CCA by maximizing the correlation of neighbor samples from the same class and minimizing the correlation of neighbor samples from different classes, and then two projection matrices are obtained to achieve feature extraction. The contributions of MLDCCA are given as follows:

(1) It utilizes the label information and data structure to combine multiple organs for plant species recognition;
(2) Two neighbor graphs are constructed to measure both the inter-class separability and the intra-class compactness in a local way.
(3) Geodesic distance is introduced to 1-NN classifier.
(4) 50 experiments are conducted on a real leaf&flower dataset to verify the performance of the proposed method.

The rest of this paper is introduced as follows: Section 2 briefly reviews CCA, LDE and geodesic distance measure. Section 3 describes the proposed method in detail. Section 4 applies MLDCCA to plant species recognition. In Section 5 the proposed method is validated on the multi-organ datasets and the experimental results are offered. Section 6 concludes this paper with some future works.

2. Related works

In this Section, canonical correlation analysis (CCA), local discriminant embedding (LDE) (Huang et al., 2014; Chen et al., 2005) and geodesic distance are simply introduced, respectively.

2.1. Canonical correlation analysis

CCA is a multi-modal learning method to find linear relationship between two kinds of multi-view data. Its goal is to seek two projection matrices for dimensionality reduction by maximizing the correlation between two-view data (McGarigal et al., 2012; Hardoon et al., 2014). Suppose a set of pair-wise data $\{(x_i, y_i)\}_{i=1}^n \in R^p \times R^q$ from n samples, where $\{x_i\}_{i=1}^n$ and $\{y_i\}_{i=1}^n$ are two view data with zero mean. Denote $X = [x_1, x_2, ..., x_n] \in R^{p\times n}$ and $Y = [y_1, y_2, ..., y_n] \in R^{q\times n}$, CCA aims to find two projection matrices $w_x \in R^{p\times d}$ and by maximizing the correlation coefficient between X and Y. The objection function of CCA is as follows,

$$\max_{w_x, w_y} \frac{E[w_x^T XY^T w_y]}{\sqrt{E[w_x^T XX^T w_x][w_y^T YY^T w_y]}} = \frac{w_x^T S_{xy} w_y}{\sqrt{w_x^T S_{xx} w_x \cdot w_y^T S_{yy} w_y}} \tag{1}$$

where E represents mathematical expectation, $S_{xx} = XX^T$, $S_{yy} = YY^T$ are the self-covariance matrices of X and Y, respectively, and $S_{xy} = XY^T$ is the cross-covariance matrix of X and Y.

From Eq. (1), we can deduce that the canonical correlation coefficient is affine-invariant to the arbitrary scaling of w_x and w_y. Then, canonical correlation coefficients from the same dataset can be constrained to normalization unit variance. Under these constraints, CCA can be reformulated equivalently as the following optimization problem with constraints:

$$\begin{aligned} &\max_{w_x, w_y} w_x^T S_{xy} w_y \\ s.\ t. \quad & w_x^T S_{xx} w_x = 1,\ w_y^T S_{yy} w_y = 1 \end{aligned} \tag{2}$$

Eq. (2) can be solved by Lagrange multiplier method and we can obtain the following generalized eigen-value decomposition problem (Huang et al., 2017);

$$\begin{bmatrix} & S_{xy} \\ S_{yx} & \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} = \lambda \begin{bmatrix} S_{xx} & \\ & S_{yy} \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} \tag{3}$$

where $S_{xy} = S_{yx}$, λ is the eigen-value.

Under conjugate orthogonal constraints, we can obtain multiple pairs of projection directions $(w_{xi}, w_{yi})_{i=1}^d$, which consist of the first d pairs of eigen-vectors corresponding to the first d largest eigen-values of the generalized eigen-value problems in Eq. (3), where d is the reduction dimensionality.

2.2. Local discriminant embedding (LDE)

Suppose m data points $\{x_i \mid x_i \in R^n\}_{i=1}^m$ and their corresponding class labels $\{c_i \mid c_i \in \{1, 2, ..., C\}\}_{i=1}^m$. Any subset of data points that belong to the same class is assumed to lie on a sub-manifold. For each data point X_i, its k nearest neighbor $N(x_i)$ can be split into two subsets within-class-neighbor $N_b(x_i)$ and between-class-neighbor $N_w(x_i)$, where $N(x_i) = N_b(x_i) \cup N_w(x_i)$. LDE aims to find a projection matrix A. Its optimization function is as follows (Chen et al., 2005);

$$\begin{aligned} &\arg\max_A A^T X(\alpha(D_b - H_b) + (1-\alpha)H_w)X^T A \\ &\quad s.\ t. \quad A^T X D_w X^T A = 1 \end{aligned} \tag{4}$$

where $X = [x_1, x_2, ..., x_n]$, α is a suitable constraint, $0 \leqslant \alpha \leqslant 1$, D_w is a diagonal matrix whose entries are column sum of H_w, i.e. $D_w = \sum_j H_{w,ij}$ and D_b is a diagonal matrix whose entries are column sum of H_b, i.e., $D_b = \sum_j H_{b,ij}$. H_b and H_w denote, respectively, the neighbor matrices defined as:

$$H_{b,ij} = \begin{cases} 1, & \text{if } X_j \in N_b(X_i) \text{ or } X_i \in N_b(X_j) \\ & 0, \text{otherwise} \end{cases} \text{ and } H_{w,ij} = \begin{cases} 1, & \text{if } X_j \in N_w(X_i) \text{ or } X_i \in N_w(X_j) \\ & 0, \text{otherwise} \end{cases}.$$

2.3. Geodesic distance measure

In subspace learning, an image can be represented as a point in the high-dimensional space. The points that are close to each other in the space are considered more similar than the points that are far from each other. A key factor is how to define the similarity between these points for constructing neighbor graphs. The ordinary Euclidean distance measure is not suitable to define this similarity in the plant organ image data space, because the data lying on a nonlinear manifold, the "true distance" between two data points is the geodesic distance, i.e., the distance along the surface of the manifold is rather than the straight-line Euclidean distance.

From the above analysis, the geodesic distance measure is defined as (Bloch, 1995);

$$L(x_i, x_j) = \|x_i - x_j\| + \gamma \cdot l \cdot (1 - \delta(x_i, x_j)) \tag{5}$$

where $l = \max_{ij} \|x_i - x_j\|$ is the data diameter in Euclidean distance, $\gamma \in [0, 1]$ is a tuning parameter to control the amount to which class information should be incorporated, and $\delta(.,.)$ is a character function, i.e., if x_i and x_j belong to same class, $\delta(x_i, x_j) = 1$; otherwise $\delta(x_i, x_j) = 0$, γ is empirically set to 0.01 (Murari et al., 2013).

For the neighbor points, Euclidean distance in the input space can provide a good approximation to geodesic distance, while geodesic distance can be approximated by adding up a sequence of "short hops" between the neighbor points. As for leaf or flower image classification, it is difficult to use Euclidean distance as a similarity measure to efficiently display the difference between the image samples, so the Euclidean distance used as the similarity measure is often replaced by the geodesic distance to obtain the actual distance between the image samples.

3. Modified local discriminant CCA (MLDCCA)

Based on LDCCA, we propose a modified LDCCA (MLDCCA) approach for plant species recognition. Similar to CCA and LDCCA, MLDCCA also aims to find two projection matrices to achieve feature reduction and maintain the locally discriminative information by maximizing the correlation of neighbor samples from the same class and minimizing the correlation of neighbor samples from different classes.

Given n characteristic pairs $\{(x_i, y_i)\}_{i=1}^n \in R^p \times R^q$ from n samples, denote $X = [x_1, x_2, ..., x_n] \in R^{p\times n}$ and $Y = [y_1, y_2, ..., y_n] \in R^{q\times n}$, then MLDCCA includes four stages.

(1) Weighted neighbor graphs. Two weighted neighbor graphs are firstly constructed to represent the locally discriminative information for two-view dataset. Two neighbor weights of X and Y are designed respectively as follows,

$$F_{ij}^x = \begin{cases} \exp(-\|x_i - x_j\|^2/\eta^x))), \text{ if } x_i \text{ and } x_j \\ \quad \text{are neighbors, and } c(x_i) = c(x_j) \\ \quad 0, \qquad \text{otherwise} \end{cases} \tag{6}$$

$$F_{ij}^y = \begin{cases} \exp(-\|y_i - y_j\|^2/\eta^y)\text{ , if } y_i \text{ and } y_j \\ \quad \text{are neighbors, and } c(y_i) = c(y_j) \\ \quad 0, \qquad \text{otherwise} \end{cases} \tag{7}$$

where $\|\cdot\|$ is Euclidean distance, $c(x_i)$ is the label of x_i, and η^x and η^y are two local adjustment parameters, which is often simply set as

$$\eta^x = \frac{1}{n(n-1)} \sum_{i=1,j>i}^{n} \|x_i - x_j\|^2, \ \eta^y = \frac{1}{n(n-1)} \sum_{i=1,j>i}^{n} \|y_i - y_j\|^2 \tag{8}$$

(2) Optimization problem. Based on F_{ij}^x and F_{ij}^y, the optimization problem of MLDCCA is as follows,

$$\max_{w_x, w_y} w_x^T F_{xy} w_y$$
$$s.\ t.\ \ w_x^T F_{xx} w_x = 1,\ w_y^T F_{yy} w_y = 1 \tag{9}$$

where $F_{xx} = XX^T$, $F_{yy} = YY^T$, and

$$F_{xy} = \sum_{i=1}^n X_i Y_i^T + \sum_{i=1}^n \sum_{j=1}^n F_{ij}^x X_i Y_i^T + \sum_{i=1}^n \sum_{j=1}^n F_{ij}^y X_i Y_i^T$$
$$= XY^T + XF^X Y^T + XF^Y Y^T$$
$$= X(I + F^X + F^Y)Y^T$$

where F^X and F^Y are two weighted self-correlation matrices reflecting the similarity between samples, whose elements are F_{ij}^x and F_{ij}^y, respectively, and I is an identity matrix.

Then the Lagrange multiplier method is applied to solve Eq. (9) and the corresponding Lagrange function is

$$L(w_x, w_y, \lambda_1, \lambda_2)$$
$$= w_x^T F_{xy} w_y - \frac{\lambda_1}{2}(w_x^T F_{xx} w_x - 1) - \frac{\lambda_2}{2}(w_y^T F_{yy} w_y - 1) \tag{10}$$

$$\frac{\partial L}{\partial w_x} = F_{xy} w_y - \lambda_1 F_{xx} w_x = 0$$
$$\frac{\partial L}{\partial w_y} = F_{xy} w_x - \lambda_2 F_{yy} w_y = 0 \tag{11}$$

From Eq. (11), we obtain

$$F_{xy} w_y = \lambda_1 F_{xx} w_x, \quad F_{yx} w_x = \lambda_2 F_{yy} w_x$$

Then

$$w_x^T F_{xy} w_y = \lambda_1 w_x^T F_{xx} w_x, \quad w_y^T F_{yx} w_x = \lambda_2 w_y^T F_{yy} w_x$$

Then

$$w_x^T F_{xy} w_y = \lambda_1, \quad w_y^T F_{yx} w_x = \lambda_2$$

Since $F_{xy}^T = F_{yx}$, thus $\lambda_1 = \lambda_2 = \lambda$.

Eq. (11) can be converted to a generalized eigen-value decomposition problem,

$$\begin{bmatrix} & F_{xy} \\ F_{xy} & \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} = \lambda \begin{bmatrix} F_{xx} & \\ & F_{yy} \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} \tag{12}$$

(3) Eigen-value decomposition. After solving Eq. (12), the optimal eigenvector corresponds to the several maximum eigen-value of Eq. (12). Assuming the reduced dimensionality of the two-view feature sets is d, after ordering the eigen-values from large to small, and selecting d ($d \le \min(p, q)$) eigenvectors corresponding the former d eigen-values, two projection matrices can be denoted as follows:

$$W_x = [w_{x1}, w_{x2}, ..., w_{xd}]$$
$$W_y = [w_{y1}, w_{y2}, ..., w_{yd}] \tag{13}$$

(4) Dimensional reduction. After obtaining the project matrices W_x and W_y, any pair (X,Y) can be projected into a low dimensional fusion feature vector $\begin{bmatrix} W_x^T X \\ W_y^T Y \end{bmatrix}$, which is used as features for classification.

From Eqs. (6) and (7), we can conclude that (1) MLDCCA embeds the global and neighbor discriminant information into CCA, because F_{xy} contains the global cross-correlation XY^T and two local weighted cross-correlations XF^XY^T and XF^YY^T; (2) two discriminant neighbor weights between node i and j are allowed to self-tune in terms of the k-neighbors. The discriminant neighbor weights F_{ij}^x and F_{ij}^y reflect the locality around each data point; (3) the smaller the $\|x_i - x_j\|$ (or$\|y_i - y_j\|$), the closer x_i and x_j (or y_i and y_j), and thus the larger F_{ij}^x(orF_{ij}^y); (4) F_{ij}^x and between the ith and jth nodes are allowed to self-tune in terms of the k-neighbor, and they are symmetric and sparse matrices; (5) because LDCCA only utilizes the local neighbor information to compute the correlation matrices without considering the global correlation between all the data pairs, it will be more severely affected by the inaccurate estimation of those similarity matrices than MLDCCA.

4. Plant species recognition based on MLDCCA

Suppose n training image pairs $\{(x_i, y_i)\}_{i=1}^n \in R^p \times R^q$ labeled by

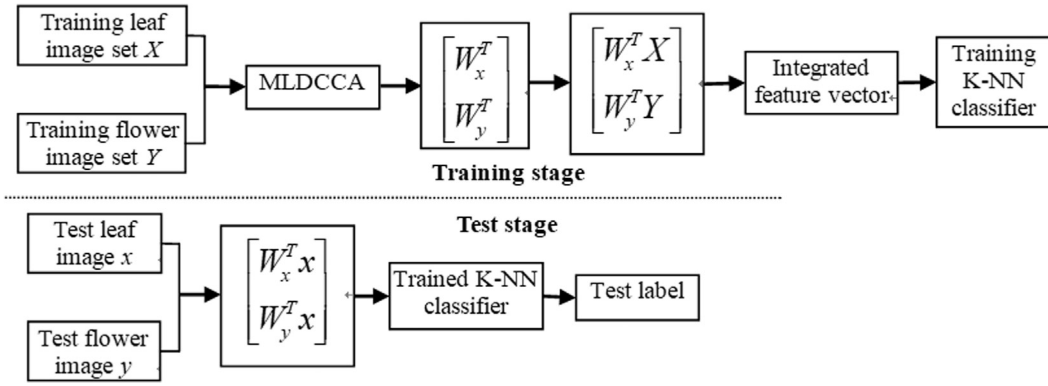

Fig. 3. The workflow of the proposed method.

$\{c_1, c_2, ..., c_n\}$, where x_i and y_i are the ith leaf image and the ith flower image of the same species, respectively, a test image pair (x, y) of plant leaf and flower image of the same species, and the reduced dimensionality d. By MLDCCA, a plant species recognition algorithm is proposed. Its workflow is shown in Fig. 3.

Its steps are described in detail as follows:

(1) Construct two weighted neighbor graphs which weights are defined by (6) and (7), respectively.
(2) Calculate F_{xx}, F_{yy} and F_{xy}.
(3) Do eigen-value decomposition in Eq. (12).
(4) Choose d eigenvectors corresponding to the largest d ($d \le \min(p, q)$) eigen-values, and form the projection matrices W_x and W_y.
(5) By W_x and W_y, project each training image pair (x_i, y_i) and a test image pair (x, y) into low-dimensional fusion feature vectors, denoted as $\begin{bmatrix} W_x^T x_i \\ W_y^T y_i \end{bmatrix}$ and $\begin{bmatrix} W_x^T x \\ W_y^T y \end{bmatrix}$, respectively, $i = 1, 2, ...,n$.
(6) Utilize $\begin{bmatrix} W_x^T x_i \\ W_y^T y_i \end{bmatrix}$$(i = 1, 2, ...,n)$ to train the classifier and then classify $\begin{bmatrix} W_x^T x \\ W_y^T y \end{bmatrix}$ by the trained classifier.

5. Experiments and analysis

This section validates the performance of the proposed MLDCCA for multi-organ plant species recognition, and compares it with the existing methods, such as periodic wavelet descriptor (PWD) (Zeng et al., 2017); hand-designed feature for plant species recognition (HDF) (Nguyen et al., 2017); score-based fusion of multi-organ images to recognize plants (SFMO) (Do et al., 2017) and local discriminant embedding (LDE) (Chen et al., 2005). We also compare the proposed with CCA and LDCCA, because MLDCCA is its modified form. PWD is used to extract the feature from each leaf image for species recognition, and a Back Propagation Neural Network (BPNN) is trained to fulfill the experiment of plant species identification. In HDF, many hand-designed features are extracted from each flower and leaf image and stacked as the vectors for plant species recognition. LDE can tackle some limitations of the global linear discriminant analysis (LDA), which splits the graph Laplacian into two components: within-class adjacency graph and between-class neighbor graph to well characterize the discriminant property of the data. In SFMO, CNN is used to extract feature from each plant organ, and the features are fused by the score-based level fusion strategy. In LDE, CCA, LDCCA and the proposed MLDCCA method, 1-NN classifier is utilized to recognize plant species for its simplicity. All experiments are performed on the platform with 3.0 GHz CPU and 4.0 GB RAM by MATLAB 2011b.

5.1. Dataset

To test and verify the proposed method, we have taken 500 images of plant leaf and flower of 10 plant species. The leaf and flower images were collected in the Zhuque national forest park, Xi'an China. We choose to take images through digital camera, smart phone, and scanner for the picked up leaves and flowers under the same condition to avoid influences by environmental factors such as illumination and orientation, and then construct an image pair dataset of plant leaf and flower. The dataset contains 500 image pairs of leaf and flower from 10 plant species, 50 image pairs per species with color JEPG format of 750 × 1024 pixels. Fig. 4 shows some samples of leaf and flower images.

5.2. Experimental results

In the experiments, each plant organ image (leaf and flower) is aligned, grayscale and resized to 32 × 32 pixels with 256 Gy levels per pixel, and matched as image pairs, where each flower image is firstly cropped with almost background. Each image is concatenated as an image vector, and then normalized with zero mean. To overcome the small-size-sample problem, the image vectors are reduced by PCA with keeping nearly 98% image energy to select the minimum number of principal components as feature dimensionality (Zeng et al., 2017; Zhang et al., 2017). All vectored image pairs are randomly split into training subset and test subset, where the training subset is used to extract two projection matrices for dimensional reduction and train the 1-NN classifier, while the test subset is used to test the performance of the proposed method. Two important parameters, i.e., k (neighbor size) and d (reduction dimensionality) are involved in our tests. A stepwise selection strategy is more feasible to determine the parameters k and d (Huang et al., 2014; Chen et al., 2005). Specifically, k is chosen from 3 to 20 according to the size of the training set and to seeing how k affects final classification accuracy. After fixed k, the optimal d is determined by the maximum recognition rate. The adjustment parameters η^x and η^y in Eqs. (6) and (7) are estimated by Eq. (8). The species recognition is performed using 1-NN classifier with geodesic distance.

A dataset is constructed by all vectored image pairs, and three-fold-cross validation scheme is adopted, namely all vectored image pairs are randomly partitioned into 3 subsets. Each time, one of the 3 subsets is used as the test set and the other two subsets are put together to form a training set. Every image gets to be in a test set exactly once, and gets to be in a training set two times. Three experimental results are obtained on three test subsets, respectively. At the end, the average of the three results is calculated as the experimental result of the three-fold-cross validation experiment. k and d are selected corresponding to the maximum recognition rate. Based on the experiments, we set $k = 10$ and $d = 80$. The three-fold-cross validation experiment is repeated 50 times to reduce the influence of random effects. The final recognition result is

(A) Flowers and leaves of Chinese rose

(B) Flowers and leaves of Rose

(C) Flowers and leaves of Marigold

Fig. 4. Examples of leaf and flower of three species.

computed by averaging the recognition accuracy of these 50 runs. Because all samples in our dataset are positive samples, no negative samples, the recognition rate on each test subset is defined as follows,

$$r = \frac{\text{Number of all samples correctly identified}}{\text{Number of all samples in the test set}} \tag{14}$$

The recognition results on the constructed dataset are shown in Table 1.

To further evaluate how much progress is made by multi-organ input and the 1-NN classifier with two different distance measure, i.e., (1) Euclidean distance simply denoted as ED and (2) geodesic distance shortly denoted as GD, we compare the results of multiple organs with the results of a single organ, i.e., extract the features by LDE from three datasets, respectively, namely the flower image set, leaf image set and Flower + Leaf image set (Chen et al., 2005), and recognize plant species by ED. Here the Flower + Leaf set denotes that the features are extracted from flower images and leaf images of the same species by LDE respectively, and then stacked as vectors for species recognition. PWD is a classical feature extraction method, which can extract multi-resolution feature vector from each leaf image for species recognition. MLDCCA is compared with PWD. We also compare MLDCCA with two widely used multiview learning methods: CCA and LDCCA, and compare the results of MLDCCA by two kinds of classifiers ED and GD. All results are also shown in Table 1.

5.3. Analysis

From Table 1, we find several facts. (1) The proposed method MLDCCA outperforms other methods in any case. The reason may be that MLDCCA fully takes the class and local information and correlation between two different organs into account, and GD can effectively indicate the similarity between two samples. MLDCCA with GD is slightly better than MLDCCA with ED, because GD is suitable for species classification. (2) The results of SFMO and MLDCCA are higher than that of others, which validates that the multi-organ recognition methods can improve the plant species recognition accuracy rate. (3) The results of HDF and LDE on Flower + Leaf image set by combining two organs are not imagine so good, which indicates that simply combining the multi-organ images without considering the relationship between the multi-organ images will reduce the recognition rate, because the stacked features by LDE or HDF from the multi-organ images does not very well describe the plant species. (4) MLDCCA is better than SFMO. The reason

Table 1
The result of classifying 10 species by LDE, PWD, CCA, LDCCA, HDF, SFMO and the proposed method.

	Method									
	LDE			PWD	CCA	LDCCA	HDF	SFMO	MLDCCA	
	Flower	Leaf	Flower + Leaf							
Classifier	ED	ED	ED	BPNN	ED	ED	SVM	SVM	ED	GD
Recognition rate (%)	67.28	75.15	71.82	78.50	81.26	84.53	72.28	91.41	92.73	94.15
Computing time (s)	25	24	27	25	23	26	25	34	20	23

may be CNN in SFMO generally requires a large number of training data to train, while our dataset is not large enough for CNN training. Therefore, the recognition accuracy of SFMO does not benefit from the combination of multi-organ images. Moreover, in experiments, we notice that the performances of SFMO strongly depend on image varieties within each species in the training dataset. The performance of the plant recognition task could be increased when the number of image pairs of each species is large enough. Anyway we can conclude that the combination of two organs, especially with multi-view learning methods and the discriminative information indeed helps to improve the performance of plant species recognition.

From the above analysis, we sum up that the superiority of the proposed method mainly owes to the following three factors, (1) two local neighbor weights and a local correlation are used to effectively express the classification performance of the plant species; (2) geodesic distance measure is exploited to evaluate the similarity in the 1-NN classifier; (3) the multi-organ feature learning methods can extract the discriminant integrated-features, while the single-organ plant recognition method ignores the relationship between the different kinds of organs.

6. Conclusion

Plant recognition based on plant organs is an important task for biological science, ecological science, and agricultural digitization. In this paper, a new method is proposed to automatically classify plant species, which considers images of both leaves and flowers. The method is based on a modified local discriminant canonical correlation analysis (MLDCCA), which is designed by incorporating the idea of local discriminant embedding (LDE) into canonical correlation analysis (CCA). The proposed method fully takes the class and local information and correlation between two different organs into consideration, and geodesic distance can effectively represent the similarity between two samples. Experimental results on real leaf and flower images dataset show that the proposed method outperforms the existing methods. In future, we will explore plant species identification methods based on multiple organs and further analyze their overall performance. Furthermore, we will study deep learning based multi-organ automatic plant recognition system.

Acknowledgments

This work is partially supported by the China National Natural Science Foundation under grant No. 61473237, key research and development plan of Shaanxi No. 2017ZDXM-NY-088 and the key project of Tianjin natural science foundation No. 18JCZDJC32100.

Conflict of interest

No conflicts of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2018.10.018.

References

Bloch, I., 1995. Fuzzy Geodesic Distance in Images. Selected papers from the Workshop on Fuzzy Logic in Artificial Intelligence, Towards Intelligent Systems. Springer-Verlag, pp. 153–166.

Chen, H.T., Chang, H.W., Liu, T.L., 2005. Local discriminant embedding and its variants. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, pp. 846–853.

Do, T.B., Nguyen, H.H., Nguyen, T.T.N., et al., 2017. Plant identification using score-based fusion of multi-organ images. In: IEEE International Conference on Knowledge and Systems Engineering, 2017, https://doi.org/10.1109/KSE.2017.8119457.

Ghazi, M.M., Yanikoglu, B., Aptoula, E., 2017. Plant species recognition using deep neural networks via optimization of transfer learning parameters. Neurocomputing 235, 228–235.

Goeau, H., Bonnet, P., Barbe, J., et al., 2012. Multi-organ plant species recognition. In: ACM International Workshop on Multimedia Analysis for Ecological Data, pp. 41–44.

Hardoon, D.R., Szedmak, S., Shawetaylor, J., 2014. Canonical correlation analysis: an overview with application to learning methods. Neural Comput. 16 (12), 2639–2664.

Huang, P., Tang, Z., Chen, C., et al., 2014. Local maximal margin discriminant embedding for face recognition. J. Visual Commun. Image Represent. 25 (2), 296–305.

Huang, X., Zhang, B., Qiao, H., et al., 2017. Local discriminant canonical correlation analysis for supervised PolSAR image classification. IEEE Geosci. Remote Sensing Lett. 14 (11), 1–5.

Kho, S.J., Manickam, S., Malek, S., et al., 2017. Automated plant species identification using artificial neural network and support vector machine. Front. Life Sci. 10 (1), 98–107.

Lee, S.H., Chan, C.S., Mayo, S.J., et al., 2017. How deep learning extracts and learns leaf features for plant classification. Pattern Recogn. 71, 1–13.

Liu, X., Xu, F., Sun, Y., et al., 2018. Convolutional recurrent neural networks for observation-centered plant identification. J. Elect. Comput. Eng. 2018 1–7.

McGarigal, Kevin, Stafford, Susan, Cushman, Sam, 2012. Canonical Correlation Analysis. Encyclopedia of Statistics in Behavioral Science. John Wiley & Sons Ltd, pp. 263–269.

Murari, A., Boutot, P., Vega, J., et al., 2013. Clustering based on the geodesic distance on Gaussian manifolds for the automatic classification of disruptions. Nucl. Fusion 53 (3) 33006–33014(9).

Nguyen, T.N., Le, T.L., Hai, V., et al., 2017. A combination of deep learning and hand-designed feature for plant species recognition based on leaf and flower images. In: Asian Conference on Intelligent Information and Database Systems, pp. 223–233.

Seeland, M., Rzanny, M., Alaqraa, N., et al., 2017. Plant species classification using flower images-A comparative study of local feature representations. Plos One 12 (2), e0170629.

Sun, Y., Liu, Y., Wang, G., et al., 2017a. Deep learning for plant identification in natural environment. Comput. Intelligence Neurosci. 2017 (4), 7361042.

Sun, Y., Liu, Y., Wang, G., et al., 2017b. Deep learning for plant species recognition in natural environment. Comput. Intelligence Neurosci. 2017 (4), 7361042.

Wäldchen, J., Mäder, P., 2018. Plant species identification using computer vision techniques: a systematic literature review. Arch. Comput. Methods Eng. 25 (2), 1–37.

Zeng, Q., Zhu, T., Zhuang, X., et al., 2017. Using the periodic wavelet descriptor of plant leaf to identify plant species. Multimedia Tools Appl. 76 (17), 1–18.

Zhang, S., Wang, H., Huang, W., 2017. Two-stage plant species recognition by local mean clustering and weighted sparse representation classification. Cluster Comput. 20 (2), 1517–1525.

Zhu, H., Huang, X., Zhang, S., et al., 2017. Plant species recognition via multipath sparse coding. Multimedia Tools Appl. 76 (3), 1–17.