Combination of Gross Shape Features, Fourier Descriptors and Multiscale Distance Matrix for Leaf Recognition

Cihan Sarı, Ceyhun Burak Akgül, Bülent Sankur
Boğaziçi University, Istanbul, Turkey
{cihan.sari,ceyhun.akgul,bulent.sankur}@boun.edu.tr

***Abstract*—In this study, we have experimented with different image and shape descriptors on the automatic leaf recognition problem. We have studied the effects of gross shape descriptors, Fourier descriptors, multiscale distance descriptors, and the combination of these on the leaf recognition performance using two different datasets. We have achieved 94.62% recognition performance on Flavia, comparable to PNN 90.31% and SVM-BDT 96%. Our performance on SLID dataset, 96.67%, is comparable to MDM-A 93.60% and hierarchical matching of deformable shapes 96.28%.**

***Keywords*—leaf recognition; gross shape features; multiscale distance matrix; Fourier descriptors; SLID; Flavia.**

I. Introduction

It is a very challenging task to determine which plant or family a leaf belongs to for non-professionals. Recognition process can be tiresome and take a long time, especially for a brand new plant and is mostly carried out by professional botanists. Therefore, it is unlikely to use a leaf as alternative medicine or nutrient for non-proffesionals. Leaf recognition is most frequently carried out for pharmacology industry as leaves are commonly used as ingredients. A leaf recognition system would help out this procedure for pharmacology industry and would let non-professionals to have access to the botanic knowledge. Therefore, in last 10 years, leaf recognition has been a popular topic in computer vision.

One of the early works on leaf recognition using shape was by Im et al.[1]. They used a polygon approximation to classify the Acer family variety. Wang et al. continued this work with shape based leaf image retrieval[2]. In their work, they used simple shape features as centroid-contour distance (CCD) curve, eccentricity and angle code histogram (ACH)[2]. They used a large database that contains 1400 color leaf images from 140 Chinese medicinal plants (10 samples from each plant). Wu et al. used basic geometric features; such as diameter, leaf area, and digital morphological features; smooth factor, aspect ratio, etc.[3]. They achieved 90.31% accuracy using probabilistic neural network on Flavia[3] dataset, which contains 1920 color leaf images. Singh et al. proposed machine learning based SVM-BDT techniques for leaf recognition[4]. They reported recognition performance of 96% on Flavia. Felzenszwalb and Schwartz proposed matching shapes based on a hierarchical description of the boundaries[5]. They have achieved up to 96.28% recognition rate on SLID (Swedish Leaf Image Dataset)[6]. Hu et al.[7] proposed a new contour based shape descriptor, Multiscale Distance Matrix (MDM). MDM is invariant to rotation, translation and bilateral symmetry. It can be improved to become invariant to scale by dividing the matrix to average distance (MDM-A) or maximum distance (MDM-M). MDM and MDM-A is described in Section II-C3 in detail. They have tested MDM on Swedish Leaf Image Database (SLID) and ICL Leaf dataset. They achieved experimental performances of 93.60% on SLID and up to 98.20% on a subset of ICL using variants of MDM descriptors.

Main purpose of this study is to combine various feature sets and increase the recognition rate on different leaf datasets. Since we do not know which samples are used as training set and which are used as test set, it is not possible to repeat the previous studies exactly. Therefore, images used for training or testing are marked and saved for repeatability, which are available on request.

On Section II, general recognition flow is explained. Segmentation and alignment methods are given on Section II-B, feature extraction techniques for three feature sets are on section II-C and individual and combined classification of feature sets are detailed on Section II-D. Section III consists of the general properties of datasets and feature sets' individual and combined performances. Finally, conclusion of the study is given in Section IV.

II. Leaf Recognition System

Both general flow of the system and methodology is given on this section.

A. System Flow

Leaf recognition system's workflow is given below in five steps which are explained in detail on their own sections.

- Blue and Saturation channels of color leaf image is used to segment the leaf shape from the background. PCA is used for rotation alignment of the leaf region (Section II-B).
- This region is used to calculate gross shape features, Fourier descriptors and MDM-A descriptors (Section II-C).
- Descriptor and classifier parameters are measured by the feature sets calculated from the training set using cross-validation (Sections II-C and II-D).

- Classifiers' confidence is measured by cross-validation performances for each plant type using fine-tuned parameters (Section II-D).
- New (unseen) instances from the testing set is then recognized by these three classifiers and weighted by the confidence to recognize the plant (Section III).

B. Preprocessing

Some preprocessing is necessary before the feature extraction. Feature sets require of a leaf region, which is computed by two steps, segmentation and alignment.

1) Segmentation: Segmentation depends on the leaf dataset. However, as both Flavia[3] and SLID[6] consist of three channel leaf images with white background, same procedure is used. Leaf images are,

- Converted from RGB color space to HSV color space.
- Blue channel is used to segment general leaf structure, and saturation channel is used to get a sharp separation between the background and the leaf.

$$I(i,j) = \frac{255 - S(i,j) + B(i,j)}{2} \tag{1}$$

where,S(i,j) and B(i,j), are saturation and blue channel's intensity value at coordinates (i,j), respectively. Computed image I (1) is then thresholded as given in (2) for leaf region segmentation.

$$R(i,j) = \left\{ \begin{array}{lll} I(i,j) < 210 & \textit{if} & 1 \\ & \textit{else} & 0 \end{array} \right\} \tag{2}$$

Threshold value 210 is decided experimentally.

2) Alignment: Translation and scale of the leaf can be estimated from the region after the segmentation. Principle Component Analysis (PCA) is used for rotation alignment due to both its speed and its performance against simple shapes as seen in Fig. II-B2.

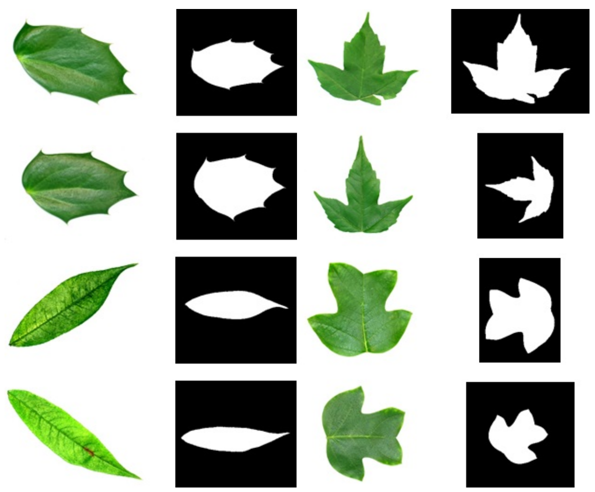

Figure 1. Rotation alignment on simple (left) and complex (right) leaves using PCA

However, there are differences in structure on leaves that belong to the same class, at which complexity of the shapes requires further segmentation before PCA. As seen in Fig. II-B2, PCA does not work on complex shapes as well as simpler ones. This is most likely due to asymmetrical structure of the leaf and torn, ruptured parts of the leaf.

One approach to reduce the complexity would be segmentation of the venation (pattern of the veins). These patterns show the leaf's structure in a more robust way and are much less subject to torn, ruptured parts and has symmetry. However, after some failed attempts to segment the venation repeatedly (Fig. 2), this approach is not taken and only PCA is used for rotation alignment.

Figure 2. Segmentation of venation

C. Feature Extraction

In this study, three different feature sets are used. These are, gross shape features, Fourier descriptors and MDM-A descriptors. Preprocessing required prior to feature extraction can be found in section II-B.

1) Gross Shape Features: Gross shape features consists of rectangularity, aspect ratio, mean hue, eccentricity and convexity. These features are explained as:

- Rectangularity is calculated as the ratio of shape's area to the area of minimum enclosing rectangle. Therefore, it is expected for similar shapes to score a similar rectangularity.
- Aspect ratio is measured from the minimum enclosing rectangle of the leaf region. For rotation invariance, it is always calculated as the ratio of the longer length to the smaller. This feature differs between different classes and is very easy to compute along with rectangularity.
- Mean hue is the only color feature used in this study. It is calculated as the average intensity of the leaf region in hue channel. Hue channel is theoretically invariant to lighting conditions and represents pure color. This feature is used to distinguish two leaf classes which are similar in structure, but different in their pure color.
- Eccentricity for the leaf region is calculated from an ellipse that has the same second moments as the region. It is the ratio of radii of this ellipse.
- Convexity, is calculated from ratio of shape's area to the area of minimum enclosing convex region. Although it is similar to rectangularity in computation, as it encloses the region like a rubber, it describes the local differences better than rectangularity.

2) Fourier Descriptors: Fourier descriptors are calculated from boundary points of the leaf region. Boundary points x, y are sampled into 1000 equidistant points to reduce the effect of the scale. These sample points coordinates are mapped between 0 and 1 to achieve translation and scale invariance.

Afterwards, Fourier transformation from x and y coordinates are calculated separately. Zero frequency is omitted and remaining N magnitude values from x ($|\mathcal{F}\{x\}|$) and y ($|\mathcal{F}\{y\}|$) are used ($2 \times N$ values in total). Due to rapid decrease in magnitude values as number of harmonics increase, natural logarithm of magnitudes are used as features.

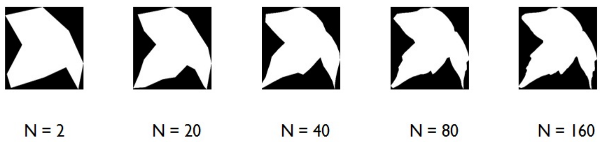

Figure 3. Regions reconstructed using N harmonics

Number of harmonics directly affects the complexity of the feature set. Therefore, increasing it too much causes overlearning the training set, whereas keeping it too low causes trained model to be too simple, and underlearning the data. Results of using different number of harmonics on leaf region can be seen in Fig. 3. Cross-validation is used on training set to determine number of harmonics to use. In Fig. 4, cross-validation error vs. complexity graph can be seen for Flavia dataset. From the results, 180 harmonics are used on Flavia and 160 harmonics are used on SLID dataset.

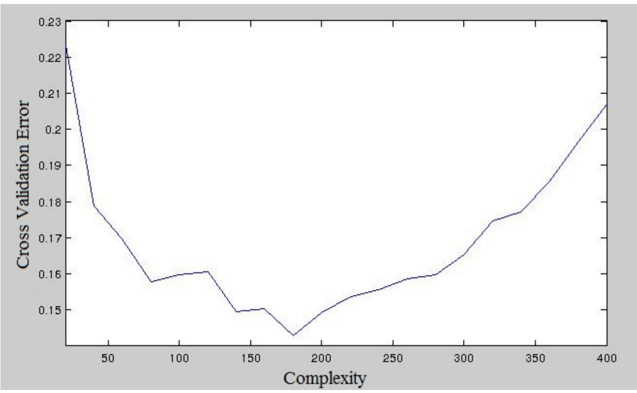

Figure 4. Complexity (number of harmonics) vs cross-validation error for Flavia training set

3) Multiscale Distance Matrix[7]: Multiscale distance matrix (MDM) is computed using boundary points of the region. Feature vector's size is directly proportional to the number of boundary points used, therefore, boundary points are downsampled into 50 equidistant points (Fig. 5).

Figure 5. Region boundary points downsampled to 50 points (left) and the original region with 7553 points (right)

Euclidean distance between these points $(p_1, p_2, ..., p_{50})$ are calculated to create 50×50 distance matrix D. D is symmetrical and all values in its diagonal are zero. From matrix D, MDM is generated by steps given in Fig. 6. These three steps are given below[7].

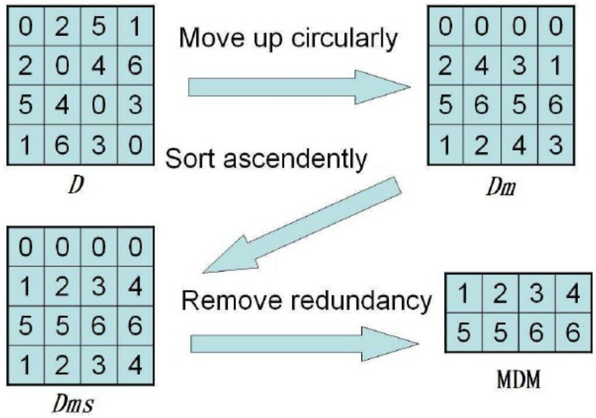

Figure 6. An example illustration of constructing MDM [7]

- For each column of matrix D, it is shifted up circularly so that the first element becomes zeros. By this way, a new matrix D_m is constructed in which the first row has straight zeros.
- For each row of D_m, its elements are sorted ascendently. This generates a matrix D_{ms}. Through this process, D_{ms} becomes invariant to the initial point and arrangement of the points.
- For D_{ms} first and last $[\frac{n-1}{2}]$ redundant rows are omitted to construct a new matrix, which is basic MDM.

It is proposed that MDM is divided by either maximum point of the matrix (MDM-M) or average value (MDM-A) to make the resulting MDM invariant to scale. Both methods are compared and repeatability performances of MDM-A outweighs MDM-M, as measured in the original article[7].

D. Classification

Quadratic discriminant analysis (QDA), support vector machines (svm) and k nearest neighbors (k-NN) are used for classification. These classifiers are used for gross shape features, Fourier descriptors and MDM-A, respectively. Final class is determined by weighing each classifier's performance. The weights are calculated by their cross-validation performance of each class, i.e. each classifier's cross-validation performance measured from training set is stored to determine the description power of each classifier for each class.

III. Experiments

Two leaf datasets are used to measure performance of the individual feature sets and their combined performances. These experiments are

- Individual gross shape features, Fourier descriptors and MDM-A performances on leaf datasets Flavia[3] and SLID[6],

- Combined classification results which is weighted by cross-validation performances of the classifiers.

Leaf datasets Flavia[3] and SLID[6] are split into two groups, training and testing sets randomly. This process is carried out only once and the same sets are used for all process - including the determination of feature extraction and classifier parameters. Therefore, impact of training and testing sets on these parameters and the performance results are fixed. Individual and combined performances are reported separately to measure the impact of the feature sets individually.

A. Datasets

Flavia[3] leaf dataset has leaves from 32 different plants. There are between 50 and 77 leaf images per plant. Total number of samples from all 32 classes is 1907. Each leaf image is a color image (3 channels) with 1600×1200 resolution.

SLID[6] has 15 different plants. There are 75 images for each plant, to a total of 1125 images across the dataset. Image resolution varies between the images and is approximately 1.5 Mpx. Leaves on SLID, just like the Flavia leaf dataset, are color images with 3 channels and white background.

B. Recognition Results

Both the leaf datasets are split into two parts randomly; 90% of the samples are picked as training set and remaining 10% is put into testing set. By doing this split only once as mentioned previously (Section III), three feature set performances are measured individually and together on previously unseen data.

TABLE I
INDIVIDUAL AND COMBINED RECOGNITION PERFORMANCES ON FLAVIA AND SLID DATASETS

Feature set	**Flavia Training**	**Flavia Test**	**SLID Training**	**SLID Test**
Gross shape	82.07%	80.70%	89.65%	86.67%
Fourier	100.0%	91.02%	100.0%	94.17%
MDM-A	88.18%	82.99%	96.22%	92.50%
All	96.18%	**94.62%**	98.51%	**96.67%**

Flavia and SLID datasets' recognition rates using individual feature sets and their combination are given in Table I. Gross shape feature performances are relatively low, but still acceptable, probably due to used classifier. MDM-A provide a performance increase than gross shape features. Fourier descriptors score the best performance amongst all feature sets. This performance differences in feature sets are affected by classification methods and their individual strength.

Combination of all three individual classifiers, it is shown that final recognition error can be substantially decreased (up to 50% on SLID and 40% on Flavia). SLID performances are higher than Flavia in all cases. This is most likely due to high variations in the Flavia dataset.

IV. CONCLUSION

In this study, it is shown that gross shape features can be used to obtain acceptable leaf recognition performances. Fourier descriptors and MDM-A both provide a substantial performance on both SLID and Flavia datasets.

Combining three distinct feature sets and their classifiers using cross-validation performances as weights, these three sets successfully compansates each others' weakness on recognizing a specific class. Each feature set is chosen to represent leaf in a different way so that where one method fails to describe a plant, another may succeed.

We have tested image and shape descriptors on two different datasets. We have used only directly extractable features (no supervisor). We have achieved 94.62% recognition performance on Flavia, comparable to PNN[3] 90.31% and SVM-BDT[4] 96%. Our performance on SLID dataset, 96.67%, is comparable to MDM-A[7] 93.60% and hierarchical matching of deformable shapes[5] 96.28%.

V. FUTURE WORK

In this study, we did not compare classifiers' ability to describe feature sets. Different classification methods could be tested for individual feature sets to minimize errors.

Additional features could be extracted from the leaf, e.g: ACH, CCD. Adding new feature sets could increase the system performance. Especially almost no color-based features are used and some of the leaves in the datasets have similar shapes, but have differences in color.

Classes predicted from individual feature sets are weighed using their cross-validation performances only. Posterior probabilities of the classifiers and class similarities are not taken into account. More intelligent combinations could be used which also respects to posterior probabilities and class similarities.

Another approach could be using decision forest for classification. As it can handle missing information, this is especially important if new color features are introduced, to be able to recognize monochrome leaf images as well as colored ones. In addition, using decision forests, variable importance can be used to select features or feature sets.

REFERENCES

[1] T. K. C. Im, H. Nishida, "Recognizing plant species by leaf shapes-a case study of the acer family," *Proceedings of IEEE International Conference on Pattern recognition*, p. 1171, 1998.
[2] Z. Wang, Z. Chi, and D. Feng, "Shape based leaf image retrieval," in *Vision, Image and Signal Processing, IEE Proceedings-*, vol. 150, no. 1. IET, 2003, pp. 34–43.
[3] S. G. Wu, F. S. Bao, E. Y. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," in *Signal Processing and Information Technology, 2007 IEEE International Symposium on*. IEEE, 2007, pp. 11–16.
[4] K. Singh, I. Gupta, and S. Gupta, "Svm-bdt pnn and fourier moment technique for classification of leaf shape," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 3, no. 4, pp. 67–78, 2010.
[5] P. F. Felzenszwalb and J. D. Schwartz, "Hierarchical matching of deformable shapes," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.
[6] O. Söderkvist, "Computer vision classification of leaves from swedish trees," Ph.D. dissertation, Linköping, 2001.
[7] R. HU, W. JIA, H. LING, and D. HUANG, "Multiscale distance matrix for fast plant leaf recognition," *IEEE transactions on image processing*, vol. 21, no. 11, pp. 4667–4672, 2012.