Comparison of Leaf Recognition by Moments and Fourier Descriptors

Tomáš Suk[1], Jan Flusser[1], and Petr Novotný[2]

[1] Institute of Information Theory and Automation of the ASCR,
Department of Image Processing,
Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic
{suk,flusser}@utia.cas.cz

[2] Charles University in Prague, Faculty of Education,
Biology & Environmental Studies Department
M.D. Rettigové 4, 116 39 Praha 1, Czech Republic
petr.novotny@pedf.cuni.cz

Abstract. We test various features for recognition of leaves of wooden species. We compare Fourier descriptors, Zernike moments, Legendre moments and Chebyshev moments. All the features are computed from the leaf boundary only. Experimental evaluation on real data indicates that Fourier descriptors slightly outperform the other tested features.

Keywords: leaf recognition, Zernike moments, Fourier descriptors, Legendre polynomials, Chebyshev polynomials.

1 Introduction

Recognition of plant species by their leaves is an important task in botany. Its automation is at the same time a challenging problem which can be resolved by visual pattern recognition methods. The plant leaves have high intraclass variability and sometimes the leaves of different plants are very similar, which makes this task difficult even for botanists.

Various approaches can be found in the literature. Kumar et al. [1] use a histogram of curvatures. The curvature is computed as a part of a disk with center on a leaf boundary covered by the leaf. The disks of several radii are used. Chen et al. [2] use another type of curvature. Kadir et al. [3] use polar Fourier transformation supplemented by a few color and vein features.

Nanni et al. [4] use the combination of inner distance shape context, shape context and height functions. Wu et al. [5] use simple geometric features as diameter, length, width, area, perimeter, smooth factor, aspect ratio, form factor, rectangularity, narrow factor, convex area ratio, ratio of diameter to perimeter, ratio of perimeter to length plus width and four vein features. The features are evaluated by principal component analysis and neural network. Söderkvist [6] uses similar features as a supplement to geometric moments with support vector machine as a classifier. In [7], Zernike moments are used.

R. Wilson et al. (Eds.): CAIP 2013, Part I, LNCS 8047, pp. 221–228, 2013.

Since the most discriminative information is carried by the leaf boundary (see Fig. 2c), all above-cited papers employ boundary-based features. We decided to objectively compare the most popular ones – Fourier descriptors, Zernike moments, Legendre moments, Chebyshev moments, and a direct use of the boundary coordinates – on a large database of tree leaves.

2 Data Set

In the experiments, we used our own data set named Middle European Woody Plants (MEW 2012 – Fig. 1, [8]). It contains all native and frequently cultivated trees and shrubs of the Central Europe Region. It has 151 botanical species (153 recognizable classes), at least 50 samples per species and a total of 9745 samples (leaves). In the case of compound leaves (Fig. 2b), we considered the individual leaflets separately.

Fig. 1. Samples of our data set (different scale – MEW 2012 scans cleaned for this printed presentation): 1st row – *Acer pseudoplatanus*, *Ailanthus altissima* (leaflet of pinnately compound leaf), *Berberis vulgaris*, *Catalpa bignonioides*, *Cornus alba*, 2nd row – *Deutzia scabra*, *Fraxinus excelsior* (leaflet of pinnately compound leaf), *Juglans regia*, *Maclura pomifera* (male), *Morus alba*, 3rd row – *Populus tremula*, *Quercus petraea*, *Salix caprea*, *Tilia cordata* and *Vaccinium vitis-idaea*.

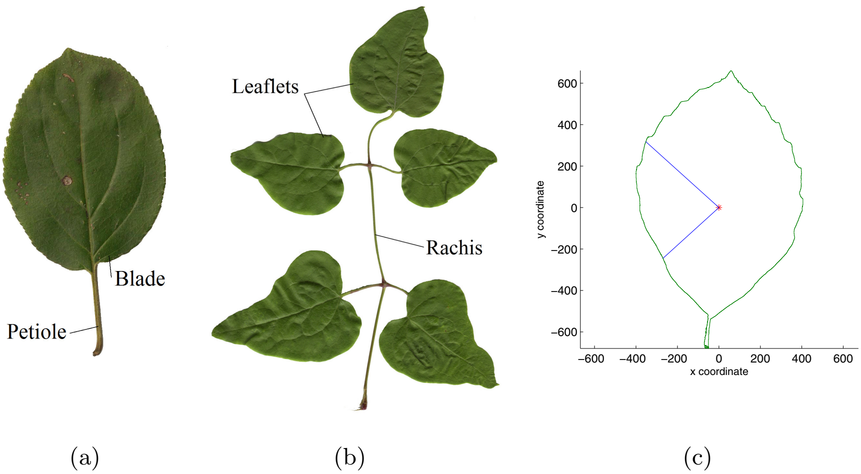

Fig. 2. (a) A simple leaf (*Rhamnus cathartica*). (b) pinnately compound leaf (*Clematis vitalba*). (c) the boundary of the leaf (*Fagus sylvatica*).

3 Method

3.1 Preprocessing

The preprocessing consists namely of the leaf segmentation and boundary detection. The scanned green leaves on white background are first segmented by simple thresholding. The leaves are converted from color to graylevel scale and then we compute the Otsu's threshold [9]. The contours in the binary image are then traced. Only the longest outer boundary of the image is used, the other boundaries (if any) and holes are ignored.

Then we compute the following features: Cartesian coordinates of the boundary points (CB), polar coordinates of the boundary (PB), Fourier descriptors (FD), Zernike moments computed of the boundary image (ZMB), Legendre moments (LM), Chebyshev moments of the first kind (CM1) and that of the second kind (CM2).

All the features need to be normalized to translation and rotation. The normalization to the translation is provided by a subtraction of the centroid coordinates m_{10}/m_{00} and m_{01}/m_{00}, where m_{pq} is a geometric moment. The rotation normalization in the case of the direct coordinates, Legendre and Chebyshev moments is provided so the principal axis coincides with the x-axis and the complex moment c_{21} would have non-negative real part

$$\begin{aligned} &\theta = \frac{1}{2}\arctan\left(\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right) \\ &\text{if } (\mu_{30}+\mu_{12})\cos(\theta) - (\mu_{21}+\mu_{03})\sin(\theta) < 0 \text{ then } \theta := \theta + \pi, \end{aligned} \tag{1}$$

where μ_{pq} is a central geometric moment. All the boundary coordinates are multiplied by a rotation matrix corresponding to the angle $-\theta$. The starting point of the coordinate sequence is that one with the minimum x-coordinate. If there is more such points, that one which minimizes y-coordinate is chosen.

3.2 Direct Coordinates

The simplest method is the direct use of the boundary coordinates as the features. To normalize the features with respect to scaling, we resampled the boundaries of all leaves to the constant number of samples n_d. Nearest neighbor interpolation was found slightly better than the linear interpolation for this purpose. Then we used the Cartesian coordinates $a_{2j-1} = x_j/n_o$ and $a_{2j} = y_j/n_o$ as the features, so we have $2n_d$ features together. n_o is the original number of the boundary points, x_j and y_j are the resampled coordinates, $j = 1, 2, \dots n_d$.

As an alternative, we tried to use the polar coordinates $a_j = \sqrt{x_j^2 + y_j^2}/n_o$ and $\varphi_j = \arctan(y_j/x_j)$, where the angle is used with a lowered weight.

3.3 Fourier Descriptors

The Fourier Descriptors [10] are defined as Fourier transformation of the boundary

$$F(u) = \sum_{k=1}^{n_o} (x_k + iy_k) e^{-2\pi iku/n_o}, \tag{2}$$

where x_k and y_k are the original boundary point coordinates, n_o is their number, u is the relative frequency (harmonic). We use the descriptors in the range $u = -n_h, -n_h + 1, \dots n_h$, where n_h is an empiric value common for all leaves, $F(-u) = F(n_o - u)$. After the translation normalization $F(0) = 0$ and it is not further considered.

The scaling invariance can be easily provided by normalization of the amplitudes by the squared boundary length. Another problem is that the magnitude of the amplitude falls quickly with the frequency and we need the appropriate weight of the features in the classifier, therefore we use the normalization $a_u = 10(|u|+1)|F(u)|/n_o^2$. The phase must be normalized to the rotation of coordinates and to the change of the starting point: $\varphi_u = \text{angle}(F(u)) - \vartheta - u\rho$, where $\vartheta = (\text{angle}(F(1)) + \text{angle}(F(-1)))/2$ and $\rho = (\text{angle}(F(1)) - \text{angle}(F(-1)))/2$.

3.4 Zernike Moments

The Zernike moments are frequently-used visual features, see e.g. [11], defined as

$$A_{n\ell} = \frac{n+1}{\pi} \sum_{k=1}^{n_o} R_{n\ell}(r_k) e^{-i\ell\varphi_k}, \tag{3}$$

where r_k and φ_k are the polar coordinates of the boundary and the radial function $R_{n\ell}(x)$ is a polynomial of the nth degree

$$R_{n\ell}(x) = \sum_{s=0}^{(n-|\ell|)/2} \frac{(-1)^s (n-s)!}{s!((n+|\ell|)/2 - s)!((n-|\ell|)/2 - s)!} x^{n-2s}. \tag{4}$$

The parameter n is called *order* and ℓ is called *repetition*. Since ZM's were designed for 2D images, we treat the leaf boundary (which is actually 1D information) as a 2D binary image.

This explicit formula becomes numerically unstable for high orders, therefore three recurrence formulas were developed. They are known as Prata method, Kintner method and Chong method, we used the Kintner method [12]. The scaling invariance is provided by a suitable mapping of the image onto the unit disk. The points in the distance κn_o from the centroid are mapped onto the boundary of the unit disk, where κ is a constant found by optimization of the discriminability on the given dataset. The value $\kappa = 0.3$ was determined for MEW2012. The parts of the leaf mapped outside the unit disk are not included into the computation. The moment amplitudes are also normalized both to a sampling density and to a contrast: $a_{n\ell} = |A_{n\ell}|/A_{00}$, the phases are normalized to the rotation as $\varphi_{n\ell} = \text{angle}(A_{n\ell}) - \ell \cdot \text{angle}(A_{31})$.

3.5 Legendre and Chebyshev Moments

The one-dimensional moments can be computed by

$$P_n = \sum_{k=1}^{n_o} (x_k + iy_k) K_n \left(2\frac{k-1}{n_o - 1} - 1 \right), \tag{5}$$

where x_k, y_k are the boundary coordinates normalized to rotation and starting point by (1). $K_n(x)$ is a Legendre or Chebyshev polynomial. They can be computed by the recurrence formula

$$K_0(x) = 1,\ K_1(x) = \alpha_0 x,\ K_n(x) = \alpha_1 x K_{n-1}(x) - \alpha_2 K_{n-2}(x), \tag{6}$$

where $\alpha_0 = 2$ for the Chebyshev polynomials of the second kind otherwise $\alpha_0 = 1$. $\alpha_1 = 2 - \frac{1}{n}$ and $\alpha_2 = 1 - \frac{1}{n}$ for the Legendre polynomials, while $\alpha_1 = 2$ and $\alpha_2 = 1$ for the Chebyshev polynomials.

The amplitude features are used as $a_n = |P_n|/n_o^2$ and the phase features as $\varphi_n = \text{angle}(P_n)$. There is the coefficient $1/n_o^2$ because of the scaling normalization.

3.6 Leaf Size

The leaf size has big intraclass variability – the largest leaf is approximately twice as large as the smallest one. Regardless, the size bears some interesting information, we must use it only with a suitable weight (see the choice of w_s in the next section). When comparing the sizes of two leaves, we must compensate for the resolution of the images if they are different. Then we find diameters $d_m^{(a)}$, $d_m^{(b)}$ of both leaves and define the distance between the leaves as

$$\delta_s(a, b) = 1 - e^{-\frac{\left(d_m^{(a)} - d_m^{(b)}\right)^2}{2 d_m^{(a)} d_m^{(b)}}}. \tag{7}$$

4 Classifier

We use a simple nearest neighbor classifier with optimized weights of individual features. While we can use just L_2 norm for comparison of the amplitude features, the phase features are angles in principle and we have to use special distance

$$\delta_\varphi(\alpha, \beta) = \min(|\alpha - \beta|, 2\pi - |\alpha - \beta|). \tag{8}$$

The distance of two leaves in the feature space is than evaluated

$$d_f(\ell, q) = w_s \delta_s(d_m^{(q)}, d_m^{(\ell)}) + \left(\sum_{k \in S_A} (a_k^{(\ell)} - a_k^{(\ell)})^2 \right)^{\frac{1}{2}} + \\ + w_f \sum_{k \in S_P} w_c(k) \delta_\varphi(\varphi_k^{(\ell)}, \varphi_k^{(q)}), \tag{9}$$

where S_A is the set of all indices, for which a_k is an amplitude feature. Similarly, S_P is the set of all indices, for that φ_k is a phase feature. The weight w_f is constant for a given type of features, while $w_c(k)$ depends on the order of the feature. We use $w_c(k) = 1/|u_k|$ for FD and $w_c(k) = 1/n_k$ for all the moments, where u_k is the current harmonic and n_k is the current moment order. In the case of CB and PB, $w_c(k)$ has no meaning. The parameters and weights of all features were optimized for MEW2012.

In the training phase, the features of all leaves in the data set are computed. In the classification phase, the features of the query leaf are computed, they are labeled by index (q) in Eq. (9), while the features labeled (ℓ) are successively whole data set features. We only consider one nearest neighbor from each species. Where the information whether the leaf is simple or compound is available, only the corresponding species are considered.

5 Results

In the experiments, we divided randomly the leaves of each species in the data set into two halves. One of them was used as a training set and the other half was tested against it. The results are in Tab. 1. The Fourier descriptors slightly outperform the other tested features. The reason of their superiority to moments in this task lies in numerical properties of the features. Since the leaves are similar

Table 1. The success rates (f – boundary features only, s – the leaf size, c – information whether the leaf is simple or compound)

test	CB	PB	FD	ZMB	LM	CM1	CM2
f	64.55%	63.16%	79.88%	69.03%	66.69%	69.13%	74.98%
f & c	67.42%	66.67%	81.84%	72.31%	69.01%	71.92%	77.47%
f & s	74.01%	73.12%	85.43%	78.10%	75.04%	77.38%	77.19%
f & s & c	76.47%	76.16%	86.86%	80.70%	77.31%	80.14%	79.69%

to one another, we need to use high-order features to distinguish them. However, when calculating the high-order moments, floating-point overflow and/or underflow may occur for the orders higher than 60 (even for orthogonal moments calculated by recurrent relations), which leads to a loss of precision. Fourier descriptors are not so prone to overflow/underflow. Although they may also suffer with numerical errors when calculating high-frequency coefficients, the influence of these errors appears to be less significant. Another reason could lie in the shape of the basis functions, which in case of Fourier descriptors can better characterize the shape of most leaves. The direct use of the boundary coordinates, without computing any sophisticated features, produces slightly worse results than both Fourier descriptors and moments. It is also interesting that the leaf size is more important than the information, whether the leaf is simple or compound.

Finally, we compared the performance of the automatic method with the performance of humans. We asked 12 students of computer science to classify the leaves visually. The experiment setup was such that they could see the query leaf and could simultaneously browse the database and compare the query with the training leaves. Unlike the algorithm, they worked with full color images, not with the boundaries only. Each test person classified 30 leaves. The mean success rate was 63% which is far less than the success rate of the algorithm regardless of the particular features used. Hence, the public web-version of our method [13] could be a good leaf recognition tool for non-specialists, which provides them with better performance and higher speed than their sight.

6 Conclusion

We have tested several types of features in a specific task - recognition of wooden species based on their leaves. We concluded that Fourier descriptors are the most appropriate features which can, when combined with the leaf size, achieve the recognition rate above 85%. A crucial factor influencing the success rate is of course the quality of the input image.

In this study, the leaves were scanned in the laboratory. The system is not primarily designed to work with photographs of the leaves taken directly on the tree. In such a case, the background segmentation and elimination of the perspective would have to be incorporated. We encourage the readers to take their own pictures and to try our public web-based application [13].

Acknowledgement. This work was supported by the grant No. P103/11/1552 of the Czech Science Foundation.

References

1. Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C., Soares, J.V.B.: Leafsnap: A computer vision system for automatic plant species identification. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part II. LNCS, vol. 7573, pp. 502–516. Springer, Heidelberg (2012)

2. Chen, Y., Lin, P., He, Y.: Velocity representation method for description of contour shape and the classification of weed leaf images. Biosystems Engineering 109(3), 186–195 (2011)
3. Kadir, A., Nugroho, L.E., Susanto, A., Santosa, P.I.: Foliage plant retrieval using polar fourier transform, color moments and vein features. Signal & Image Processing: An International Journal 2(3), 1–13 (2011)
4. Nanni, L., Brahnam, S., Lumini, A.: Local phase quantization descriptor for improving shape retrieval/classification. Pattern Recognition Letters 33(16), 2254–2260 (2012)
5. Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.X., Chang, Y.F., Xiang, Q.L.: A leaf recognition algorithm for plant classification using probabilistic neural network. In: 7th International Symposium on Signal Processing and Information Technology ISSPIT 2007, p. 6. IEEE (2007)
6. Söderkvist, O.J.O.: Computer vision classiffcation of leaves from swedish trees. Master's thesis, Linköping University (September 2001)
7. Kadir, A., Nugroho, L.E., Susanto, A., Santosa, P.I.: Experiments of Zernike moments for leaf identification. Journal of Theoretical and Applied Information Technology 41(1), 113–124 (2012)
8. MEW2012: Download middle european woods (2012), `http://zoi.utia.cas.cz/node/662`
9. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9(1), 62–66 (1979)
10. Lin, C.C., Chellapa, R.: Classification of partial 2-D shapes using Fourier descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence 9(5), 686–690 (1987)
11. Flusser, J., Suk, T., Zitová, B.: Moments and Moment Invariants in Pattern Recognition. Wiley, Chichester (2009)
12. Kintner, E.C.: On the mathematical properties of Zernike polynomials. Journal of Modern Optics 23(8), 679–680 (1976)
13. MEWProjectSite: Recognition of woods by shape of the leaf (2012), `http://leaves.utia.cas.cz/index?lang=en`