Texture-Based Leaf Identification

Milan Sulc(✉) and Jiri Matas

Center for Machine Perception, Department of Cybernetics,
Faculty of Electrical Engineering, Czech Technical University in Prague,
Prague, Czech Republic
sulcmila@cmp.felk.cvut.cz

Abstract. A novel approach to visual leaf identification is proposed. A leaf is represented by a pair of local feature histograms, one computed from the leaf interior, the other from the border. The histogrammed local features are an improved version of a recently proposed rotation and scale invariant descriptor based on local binary patterns (LBPs).

Describing the leaf with multi-scale histograms of rotationally invariant features derived from sign- and magnitude-LBP provides a desirable level of invariance. The representation does not use colour.

Using the same parameter settings in all experiments and standard evaluation protocols, the method outperforms the state-of-the-art on all tested leaf sets - the Austrian Federal Forests dataset, the Flavia dataset, the Foliage dataset, the Swedish dataset and the Middle European Woods dataset - achieving excellent recognition rates above 99%.

Preliminary results on images from the north and south regions of France obtained from the LifeCLEF'14 Plant task dataset indicate that the proposed method is also applicable to recognizing the environmental conditions the plant has been exposed to.

1 Introduction

Recognition of plants is a challenging computer vision problem that requires dealing with irregular shapes and textures with high intraclass variability. Interest in methods for visual classification of plants has grown recently [4,5,13,18] as devices equipped with cameras became ubiquitous, making intelligent field guides, education tools and automation in forestry and agriculture practical. Belhumeur et al. [4] discuss, how using such a system in the field a botanist can quickly search entire collections of plant species - a process that previously took hours can now be done in seconds.

Plant recognition has been posed, almost without exceptions [3], as recognition of specific organs such as flowers, bark, fruits or leaves or their combination [1,4,5,7–15,18,19,26,28,32,33]. Leaf recognition has been by far the most popular and a wide range of methods has been reported in the literature [1,4,5,7–15,19,26,28,32,33].

We propose a novel approach to leaf recognition. It achieves excellent recognition rates above 99% on a number of public datasets, outperforming the state-of-the-art. The method uses neither color nor an explicit shape model, focusing

L. Agapito et al. (Eds.): ECCV 2014 Workshops, Part IV, LNCS 8928, pp. 185–200, 2015.
DOI: 10.1007/978-3-319-16220-1_14

on leaf texture, which is represented by a pair of local feature histograms, one computed from the leaf interior, the other from the border. Experimental evaluation of the proposed method shows the importance of both the border and interior textures and that global point-to-point registration to reference models is not needed for precise leaf recognition.

The histogrammed local features are an improved version of the recently proposed rotation and scale invariant descriptor [29] which is based on local binary patterns (LBPs). Describing the leaf with multi-scale histograms of powerful rotationally invariant features derived from sign and magnitude LBPs provides a desirable level of invariance. It avoids the need for registration of the leaf stem, axis and boundary. Compound leaves are handled naturally.

The leaf recognition task is commonly understood as the identification of plant species and several leaf datasets have been collected [5,7,9,19,28,33] containing images of leaves labeled by plant species. While the leaf species is determined by its genotype, its appearnace is influenced by environmental conditions. We provide preliminary results on binary classification of leaves from different locations (south and north regions of France), assuming that the plants are exposed to different environmental conditions while having similar genotypes and show the Ffirst representation is capable of fairly accurate prediction of the collection site location.

The rest of this paper is organized as follows: Section 2 reviews the art in automatic plant identification from images of leaves or combinations of leaves with other images. Section 3 describes the texture recognition method called Ffirst (Fast Features Invariant to Rotation and Scale of Texture). Section 4 explains how the Ffirst descriptor is used for the leaf region. Experiments and results are presented in Section 5. Section 6 concludes the paper.

2 State of the Art

Recognition of leaves usually refers only to recognition of broad leaves, needles are treated separately. Several techniques have been proposed for leaf description, often based on combining features of different character (shape features, colour features, etc.).

The leaf recognition method by Fiel and Sablatnig [5] is based on a Bag of Words model with SIFT descriptors and achieves 93.6% accuracy on a leaf dataset of 5 Austrian tree species. This dataset denoted as AFF is also used in our experiments in Sections 5.3.

Kadir et al. compare several shape methods on plant recognition [8]. Of the compared methods - geometric features, moment invariants, Zernike moments and Polar Fourier Transform - Polar Fourier Transform performed best achieving 64% accuracy on a database of 52 plant species. The dataset has not been published.

Kumar et al. [13] describe Leafsnap, a computer vision system for automatic plant species identification, which has evolved from the earlier plant identification systems by Agarwal et al. [1] and Belhumeur et al. [4]. Compared to the

earlier versions, they introduced a pre-filter on input images, numerous speed-ups and additional post-processing within the segmentation algorithm, the use of a simpler and more efficient curvature-based recognition algorithm instead of Inner Distance Shape Context (IDSC); a larger dataset of images, and a new interactive system for use by non-expert users. Kumar et al. [13] introduce the Leafsnap database of 184 tree species, however at the time of writing this paper it was not publicly available. On this database, 96.8% of queries have a species match within the top 5 results shown to the user with the used method. The resulting electronic field guide, developed at Columbia University, the University of Maryland, and the Smithsonian Institution, is available as a free mobile app for iOS devices. Although the app runs on iPhone and iPad devices, the leaf images are processed on a server, internet connection is thus required for recognition, which might cause problems in natural areas with slow or no data connection. Another limit is the need to take the photos of the leaves on a white background.

Wu et al. [33] proposed a Probabilistic Neural Network for leaf recognition using 12 commonly used Digital Morphological Features (DMFs), derived from 5 basic features (diameter, physiological length, physiological width, leaf area, leaf perimeter). The authors collected a publicly available database of plant leaves called Flavia, containing 1907 images of leaves from 32 species. The average accuracy on the current version of the dataset is 93%[1]. The Flavia dataset is discussed in Section 5.1. In Section 5.3 our results are compared to the best reported by Kadir et al. [7,10] and Lee et al. [14,15], as well as to the results in Novotný and Suk [19], and Karuna et al. [11], who used a different evaluation protocol.

Kadir et al. [9] prepared the Foliage dataset, consisting of 60 classes of leaves, each containing 120 images. Results on the Foliage dataset are compared in Section 5.3. The best reported result by Kadir et al. [10] was achieved by a combination of shape, vein, texture and colour features processed by Principal Component Analysis before classification by a Probabilistic Neural Network.

Söderkvist [28] proposed a visual classification system of leaves and collected the so called Swedish dataset containing scanned images of 15 classes of Swedish trees. Wu et al. [32] introduced a visual descriptor for scene categorization called the spatial Principal component Analysis of Census Transform (spatial PACT), achieving a 97.9% recognition rate on the Swedish dataset. Qi et al. achieve[2] 99.38% accuracy on the Swedish dataset using a texture descriptor called Pairwise Rotation Invariant Co-occurrence Local Binary Patterns (PRI-CoLBP) [26] with SVM classification. In Section 5.3 we provide experimental results on the Swedish dataset.

Novotný and Suk [19] proposed a leaf recognition system, using Fourier descriptors of the leaf contour normalised to translation, rotation, scaling and starting point of the boundary. The authors also collected a new large leaf dataset called Middle European Woods (MEW) containing 153 classes of native or

[1] http://flavia.sourceforge.net
[2] http://qixianbiao.github.io

frequently cultivated trees and shrubs in Central Europe. Their method achieves 84.92% accuracy when the dataset is split into equally sized training and test set. Section 5.3 contains the comparison to our results.

One possible application of leaf description is the identification of a disease. Pydipati et al. [25] proposed a system for citrus disease identification using Color Co-occurrence Method (CCM), achieving accuracies of over 95% for 4 classes (normal leaf samples and samples with a greasy spot, melanose, and scab).

Kim et al. [12] proposed a tree classification method using a combination of leaf, flower and bark photos of the same tree. The description consists of 20 features of wavelet decomposition with 3 levels for a grey and a binary image for description of bark, 32 features of Fourier descriptor for leaves and 72 features in the HS colour space for flowers. The results were obtained on an unpublished dataset consisting of 16 classes. Recognition accuracy of 31%, 75% and 75% is reported for individual leaf, flower and bark classification and 84%, 75% and 100% accuracy for combinations of leaf+flower, leaf+bark and bark+flower. However, in all cases only a single image per class was tested. The statistical significance of such result is questionable and may be prone to overfitting and unreliable.

Pl@ntNet[3] [3] is an interactive plant identification and collaborative information system providing an image sharing and retrieval application for plant identification. It has been developed by scientists from four French research organizations (Cirad, INRA, INRIA and IRD) and the Tela Botanica network. The Pl@ntNet-identify Tree Database provides identification by combining information from images of the habitat, flower, fruit, leaf and bark. The exact algorithms used in the Pl@ntNet-identify web service[4] and their accuracies are not publicly documented.

3 The Ffirst Method

In order to describe the leaf texture independently of the leaf size and orientation in the image, a description invariant to rotation and scale is needed. For applications like intelligent field guides, the recognition method also has to be reasonably fast.

In this section we describe a novel texture description called Ffirst (Fast Features Invariant to Rotation and Scale of Texture), which combines several state-of-the-art approaches to satisfy the given requirements. This method builds on and improves a texture descriptor for bark recognition introduced in [29].

3.1 Completed Local Binary Pattern and Histogram Fourier Features

The Ffirst description is based on the Local Binary Patterns (LBP) [20,22]. The common LBP operator (further denoted as sign-LBP) computes the signs of

[3] http://www.plantnet-project.org/
[4] http://identify.plantnet-project.org/en/

differences between pixels in the 3×3 neighbourhood and the center pixel. LBP have been generalized [21] to arbitrary number of neighbours P on a circle of radius R, using an image function $f(x, y)$ and neighbourhood point coordinates (x_p, y_p):

$$\mathrm{LBP}_{P,R}(x,y) = \sum_{p=0}^{P-1} s(f(x,y) - f(x_p, y_p))2^p, \quad s(z) = \begin{cases} 1 : z \leq 0 \\ 0 : \quad \text{else} \end{cases}. \tag{1}$$

To achieve rotation invariance[5], Ffirst uses the so called LBP Histogram Fourier Features (LBP-HF) introduced by Ahonen et al. [2], which describe the histogram of uniform patterns using coefficients of the discrete Fourier transform. Uniform LBP are patterns with at most 2 spatial transitions (bitwise 0-1 changes). Unlike the simple rotation invariants using $\mathrm{LBP}^{\mathrm{ri}}$ [21,23], which assign all uniform patterns with the same number of 1s into one bin,

$$\mathrm{LBP}^{ri}_{P,R} = \min\left\{\mathrm{ROR}\left(\mathrm{LBP}_{P,R}, i\right) \mid i = 0, 1, .., P-1\right\}, \tag{2}$$

the LBP-HF features preserve the information about relative rotation of the patterns.

Denoting a uniform pattern $U_p^{n,r}$, where n is the number of "1" bits and r denotes the rotation of the pattern, the DFT for given n is expressed as:

$$H(n,u) = \sum_{r=0}^{P-1} h_I(U_p^{n,r}) e^{-i2\pi ur/P}, \tag{3}$$

where the histogram value $h_I(U_p^{n,r})$ denotes the number of occurrences of a given uniform pattern in the image.

The LBP-HF features are equal to the absolute value of the DFT magnitudes (which are not influenced by the phase shift caused by rotation):

$$\text{LBP-HF}(n,u) = |H(n,u)| = \sqrt{H(n,u)\overline{H(n,u)}}. \tag{4}$$

Since h_I are real, $H(n,u) = H(n,P-u)$ for $u = (1, .., P-1)$, and therefore only $\lfloor \frac{P}{2} \rfloor + 1$ of the DFT magnitudes are used for each set of uniform patterns with n "1" bits for $0 < n < P$. Three other bins are added to the resulting representation, namely two for the "1-uniform" patterns (with all bins of the same value) and one for all non-uniform patterns.

The LBP histogram Fourier features can be generalized to any set of uniform patterns. In Ffirst, the LBP-HF-S-M description introduced by Zhao et al. [34] is used, where the histogram Fourier features of both sign- and magnitude-LBP are calculated to build the descriptor. The combination of both sign- and magnitude-LBP called Completed Local Binary Patterns (CLBP) was introduced by Guo and Zhang [6]. The magnitude-LBP checks if the magnitude of the difference of

[5] LBP-HF (as well as LBP^{ri}) are rotation invariant only in the sense of a circular bit-wise shift, e.g. rotation by multiples 22.5° for $\mathrm{LBP}_{16,R}$.

the neighbouring pixel (x_p, y_p) against the central pixel (x, y) exceeds a threshold t_p:

$$\text{LBP-M}_{P,R}(x,y) = \sum_{p=0}^{P-1} s(|f(x,y) - f(x_p,y_p)| - t_p)2^p. \tag{5}$$

We adopted the common practice of choosing the threshold value (for neighbours at p-th bit) as the mean value of all m absolute differences in the whole image:

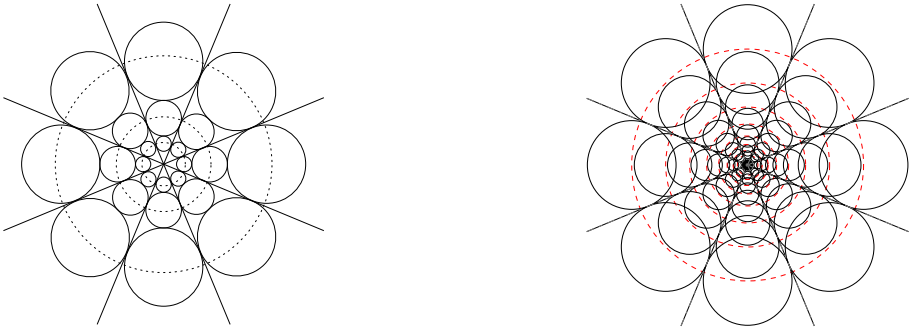
$$t_p = \sum_{i=1}^{m} \frac{|f(x_i,y_i) - f(x_{ip},y_{ip})|}{m}. \tag{6}$$

The LBP-HF-S-M histogram is created by concatenating histograms of LBP-HF-S and LBP-HF-M (computed from uniform sign-LBP and magnitude-LBP).

3.2 Multi-Scale Description and Scale Invariance

A scale space is built by computing LBP-HF-S-M from circular neighbourhoods with exponentially growing radius R. Gaussian filtering is used[6] to overcome noise.

Unlike the MS-LBP approach of Mäenpää and Pietikäinen [17], where the radii of the LBP operators are chosen so that the effective areas of different scales touch each other, Ffirst uses a finer scaling with a $\sqrt{2}$ step between scales radii R_i, i.e. $R_i = R_{i-1}\sqrt{2}$.



(a) Scale space of Mäenpää and Pietikäinen [17]

(b) Scale space from [29] used in Ffirst

Fig. 1: The effective areas of filtered pixel samples in a multi-resolution $\text{LBP}_{8,R}$ operator

[6] The Gaussian filtering is used for a scale i only if $\sigma_i > 0.6$, as filtering with lower σ_i leads to significant loss of information.

This radius change is equivalent to decreasing the image area to one half. The finer sampling uses more evenly spaced information compared to [17], as illustrated in Figures 1a, 1b. The first LBP radius used is $R_1 = 1$, as the LBP with low radii capture important high frequency texture characteristics.

Similarly to [17], the filters are designed so that most of their mass lies within an effective area of radius r_i. We select the effective area diameter, such that the effective areas at the same scale touch each other: $r_i = R_i \sin \frac{\pi}{P}$.

LBP-HF-S-M histograms from c adjacent scales are concatenated into a single descriptor. Invariance to scale changes is increased by creating n_{conc} multi-scale descriptors for one image. See Algorithm 1 for the overview of the texture description method.

Algorithm 1. The Ffirst description method overview

```
R_1 := 1
for all scales i = 1...(n_conc + c − 1) do
  σ_i := R_i sin π/P / 1.96
  if σ_i > 0.6 then
    apply Gaussian filter (with std. dev. σ_i) on the original image
  end if
  extract LBP_{P,R_i}-S and LBP_{P,R_i}-M and build the LBP-HF-S-M descriptor
  for j = 1...n_conc do
    if i ≥ j and i < j + c then
      attach the LBP-HF-S-M to the j-th multi-scale descriptor
    end if
  end for
  R_{i+1} := R_i √2
end for
```

3.3 Support Vector Machine and Feature Maps

In most applications, a Support Vector Machine (SVM) classifier with a suitable non-linear kernel provides higher recognition accuracy at the price of significantly higher time complexity and higher storage demands (dependent on the number of support vectors). An approach for efficient use of additive kernels via explicit feature maps is described by Vedaldi and Zisserman [31] and can be combined with a linear SVM classifier. Using linear SVMs on feature-mapped data improves the recognition accuracy, while preserving linear SVM advantages like fast evaluation and low storage (independent on the number of support vectors), which are both very practical in real time applications. In Ffirst we use the explicit feature map approximation of the χ^2 kernel.

The "One versus All" classification scheme is used for multi-class classification, implementing the Platt's probabilistic output [16,24] to ensure SVM results comparability among classes. The maximal posterior probability estimate over all scales is used to determine the resulting class.

In our experiments we use a Stochastic Dual Coordinate Ascent [27] linear SVM solver implemented in the VLFeat library [30].

4 Describing the Leaf Region

Before a description of a leaf is calculated, it has to be segmented. All datasets used in our experiments contain images of leaves on a white background, thus simple segmentation by thresholding is applicable. The threshold value is set automatically using the Otsu's method. Hole-filling is applied after the thresholding in order to ensure that even lighter spots in the leaf are labeled as foreground. This paper does not address the problem of leaf segmentation on a complicated background.

The Ffirst description is computed on the segmented region $\mathbb{A}$. One option is to describe only such points that have all neighbours at given scale inside $\mathbb{A}$. This description is less dependent on segmentation quality. However describing a correctly segmented border, i.e. points in $\mathbb{A}$ with one or more neighbours outside $\mathbb{A}$, can add additional discriminative information.

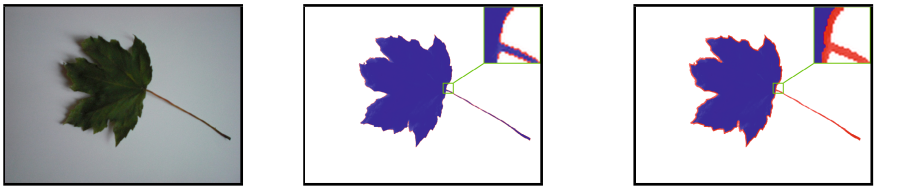
In total there will be 5 variations of the leaf recognition method used in our experiments in Section 5, differing in the processing of the border region:

1. Ffirst^{a} describes all pixels in $\mathbb{A}$. Classification maximizes the posterior probability estimate (i.e. SVM Platt's probabilistic output) over all n_{conc} scales.
2. Ffirst^{i} describes the leaf interior, i.e. pixels in $\mathbb{A}$ with all neighbours in $\mathbb{A}$. Classification maximizes the posterior probability estimate over all n_{conc} scales.
3. Ffirst^{b} describes the leaf border, i.e. pixels in $\mathbb{A}$ with at least one neighbour outside $\mathbb{A}$. Classification maximizes the posterior probability estimate over all n_{conc} scales
4. $\text{Ffirst}^{\text{ib}}_{\sum}$ combines the description from Ffirst^{i} and Ffirst^{b}. Classification maximizes the sum of posterior probability estimates over all n_{conc} scales.
5. $\text{Ffirst}^{\text{ib}}_{\prod}$ combines the description from Ffirst^{i} and Ffirst^{b}. Classification maximizes the product of posterior probability estimates over all n_{conc} scales.

5 Experiments

5.1 Datasets

The following leaf databases are used for results evaluation in Section 5.3, all of them being public with the exception of the Austrian Federal Forest dataset.



(a) Original image (b) Segmentation for R=2.8 (c) Segmentation for R=11.3

Fig. 2: Examples of leaf interior (blue) and border region (red) at different scales

The Austrian Federal Forest (AFF) Datasets were used by Fiel and Sablatnig [5] for recognition of trees based on images of leaves, bark and needles. The datasets are not publicly available, the Computer Vision Lab, TU Vienna, kindly made them available to us for academic purposes, with courtesy by Österreichische Bundesforste/Archiv. In this paper we use the AFF dataset of leaves, which contains 134 photos of leaves (on white background) of the 5 most common Austrian broad leaf trees. The results are compared using the protocol of Fiel and Sablatnig, i.e. using 8 training images per leaf class.

Fig. 3: Examples from the AFF leaf dataset

The Flavia Leaf Dataset contains 1907 images (1600x1200 px) of leaves from 32 plant species on white background, 50 to 77 images per class.

Even though in the original paper by Wu et al. [33] 10 images per class are used for testing and the rest of the images for training, most recent publications use 10 randomly selected test images and 40 randomly selected training images per class, achieving better recognition accuracy even with the lower number of training samples. In the case of the two best result reported by Lee et al. [14,15], the number of training samples is not clearly stated[7]. Some papers divide the set of images for each class into two halves, one being used for training and the other for testing.

Fig. 4: Examples of 4 classes from the Flavia leaf dataset

The Foliage Leaf Dataset [7,9] contains 60 classes of leaves from 58 species. The dataset is divided into a training set with 100 images per class and a test set with 20 images per class.

[7] In [15], the result presented as "95.44% (1820 / 1907)" seems to be tested on all images

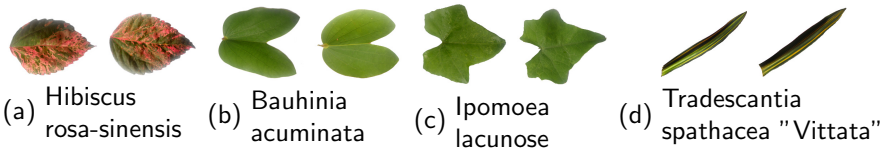


Fig. 5: Examples of 4 classes from the Foliage dataset

The Swedish Leaf Dataset was introduced in Söderkvist's diploma thesis [28] and contains images of leaves scanned using 300 dpi colour scanner. There are 75 images for each of 15 tree classes. The standard evaluation scheme uses 25 images for training and the remaining 50 for testing.

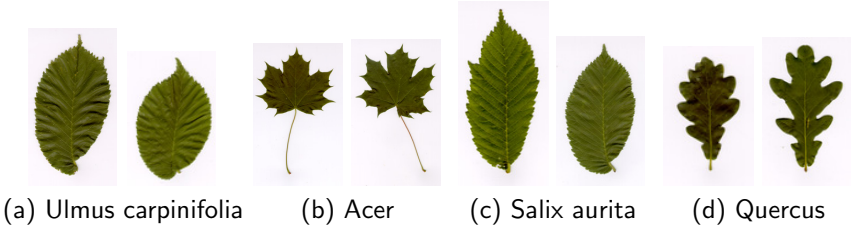


Fig. 6: Examples of 4 classes from the Swedish dataset

The Middle European Woods (MEW) Dataset was recently introduced by Novotný and Suk [19]. It contains 300 dpi scans of leaves belonging to 153 classes (from 151 botanical species) of Central European trees and shrubs. There are 9745 samples in total, at least 50 per class. The experiments are performed using half of the images in each class for training and the other half for testing.

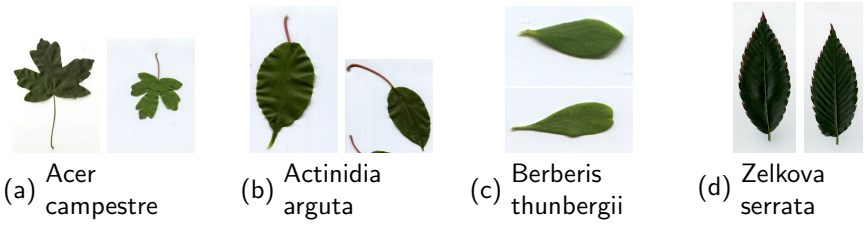


Fig. 7: Examples of 4 classes from the MEW dataset

5.2 Parameters

In all following experiments, we use the same setting of our method: $n_{\text{conc}} = 3$ multi-scale descriptors per image are used, each of them consisting of $c = 6$

scales described using LBP-HF-S-M. The final histogram is kernelized using the approximate χ^2 feature map. In the application, the data are only trained once and the training precision is more important than the training time. Thus we demand high accuracy, setting SVM parameters to: regularization parameter $\lambda = 10^{-7}$, tolerance for the stopping criterion $\epsilon = 10^{-7}$, maximum number of iterations: 10^8. We use the unified setting in order to show the generality of the Ffirst description, although setting the parameters individually for a given dataset might further increase the accuracy.

5.3 Experimental Results

Table 1 shows our classification results on all available datasets, using the standard evaluation schemes. To reduce the effect of random training and test data choice, the presented results are averaged from 10 experiments.

Table 1: Evaluation of Ffirst on available leaf datasets: Austrian Federal Forests, Flavia, Foliage, Swedish, Middle European Woods

	AFF	*Flavia* 10×40	*Flavia* $\frac{1}{2} \times \frac{1}{2}$	*Foliage*	*Swedish*	*MEW*
Num. of classes	5	32	32	60	15	153
Ffirst^{a}	97.8±1.0	98.9±0.6	98.4±0.3	98.6	99.7±0.2	97.7±0.3
Ffirst^{i}	97.6±1.4	98.1±0.8	97.9±0.4	96.7	99.6±0.4	96.9±0.3
Ffirst^{b}	98.9±1.6	98.9±0.4	98.4±0.3	96.1	98.8±0.5	96.0±0.4
Ffirst$^{ib}_{\sum}$	99.8±0.5	99.6±0.3	**99.5±0.2**	98.8	99.8±0.3	98.7±0.1
Ffirst$^{ib}_{\prod}$	**100.0±0.0**	**99.7±0.3**	99.4±0.2	**99.0**	**99.8±0.3**	**99.2±0.1**
Fiel, Sablatnig [5]	93.6	–	–	–	–	–
Novotný, Suk [19]	–	–	91.5	–	–	84.9
Karuna et al. [11]	–	–	96.5	–	–	–
Kadir et al. [10]	–	95.0	–	95.8	–	–
Kadir et al. [7]	–	94.7	–	93.3	–	–
Lee et al. [8] [15]	–	95.4	–	–	–	–
Lee et al. [8] [14]	–	97.2	–	–	–	–
Wu et al. [32]	–	–	–	–	97.9	–
Qi et al. [9] [26]	–	–	–	–	99.4	–

[8] the evaluation schemes in [14,15] are not clearly described, as discussed in Section 5.1

[9] according to the project homepage http://qixianbiao.github.io

5.4 Species Retrieval

In some applications, even results which are not correctly classified may be useful if the correct species is retrieved among the top results. For example, in an intelligent field guide it is enough to show the correct result in a shortlist of possible results, allowing the user to make the final decision.

We conducted a species retrieval experiment, performed using the $\text{Ffirst}^{\text{ib}}_{\prod}$ method on the MEW dataset, the largest available, containing 153 classes. Half of the images were used for training and half for testing. The results are presented in Figure 8.

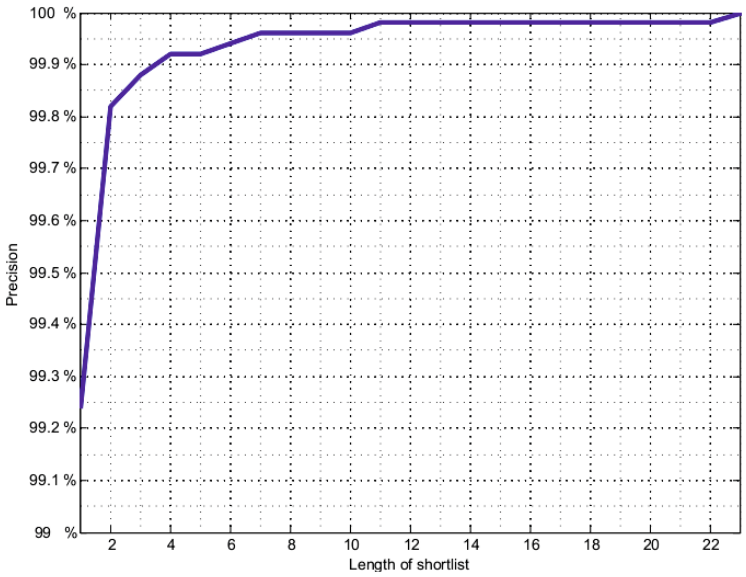


Fig. 8: Retrieval precision for different lengths of shortlist, MEW leaf dataset (153 classes)

5.5 Classifying Leaf Collection Sites

Leaf phenotype, and thus its appearance, is not only determined by the plant species, but also by the influence of the environment. In the experiment, we test whether the Ffirst representation is sufficiently rich to allow determining, besides the leaf species, the location where the leaf was collected.

The experiment was conducted on the publicly available training images from the LifeCLEF'14 Plant identification task[10]. We selected species that have at least 10 specimen collected *from different trees* in the north of France and 10 specimen in the south of France, as illustrated in Figure 9.

[10] http://www.imageclef.org/2014/lifeclef/plant

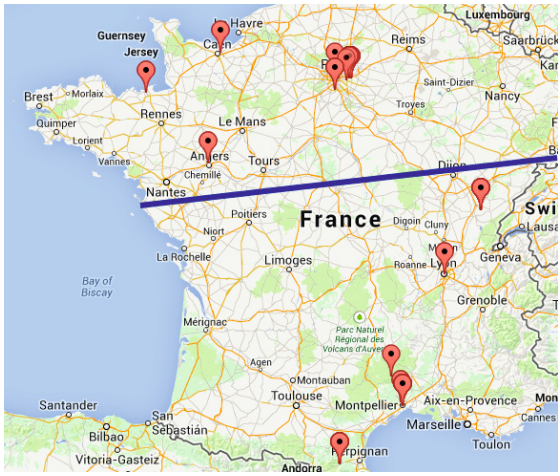


Fig. 9: Two classes of leaf collection sites

Table 2: Enviromental conditions recognition, 10-fold cross validation.

	Betula pendula Roth	*Corylus avellana L.*	*Castanea sativa Mill.*	*Acer campestre L.*
Ffirst[a]	90.0±21.1	90.0±21.1	90.0±21.1	70.0±42.2
Ffirst[i]	85.0±33.8	95.0±15.8	85.0±24.2	70.0±35.0
Ffirst[b]	90.0±31.6	80.0±25.8	80.0±25.8	75.0±35.4
$\text{Ffirst}_{\sum}^{ib}$	90.0±31.6	85.0±24.2	90.0±21.1	85.0±33.8
$\mathbf{Ffirst}_{\prod}^{ib}$	90.0±31.6	85.0±24.2	90.0±21.1	85.0±33.8

The resulting material contained 80 leaf images of 4 species - Betula pendula Roth, Corylus avellana L., Castanea sativa Mill. and Acer campestre L. For each species, a kernelized SVM was trained on the binary North-South classification task. The parameters and the SVM training used the Ffirst method with 10-fold cross validation, exactly as in the species classification. The results presented in Table 2 show that the recognition rate is well above chance, ranging from 85%-90% for the $\text{Ffirst}_{\prod}^{i,b}$. It is important to note that factors beyond phenotype changes due to the environment might facilitate the classification task, e.g. systematic differences in the cameras, dates of acquisition and lighting conditions.

Fig. 10: Examples of misclassified leaves from the Foliage dataset with wrong segmentation (Original image, Segmentation for R=2.8, Segmentation for R=11.3)

6 Conclusions

A new method for leaf classification has been proposed. Its novelties include the use of a pair of histograms representing the texture on the border and in the interior of the leave, the application of Ffirst, the Fast Features Invariant to Rotation and Scale of Texture, and the χ^2 kernel to leaf recognition.

Best results were obtained by the new $\text{Ffirst}^{\text{i,b}}_{\prod}$ method, which combines the classifiers for the leaf border and leaf interior, achieving more than 99% recognition accuracy on all used leaf datasets using the same setting and describing only the gray-scale image texture information. The species retrieval experiment on the largest dataset containing 153 classes shows that the correct result will be displayed among the top 4 results in more than 99.9% of cases.

Even Ffirst^{a}, the simple, less accurate variant not distinguishing the leaf border and interior that is more robust to small errors in leaf segmentation and to leaf border damage, outperforms the state-of-the-art on all tested datasets.

A robust segmentation method might further improve the results. It will also broaden the applicability to leaf picture taken on unconstrained backgrounds. Examples of misclassified leaves with wrong segmentation are shown in Figure 10. Further improvements might be achieved by combining the proposed method with complementary features, such as color or the global leaf shape.

An experiment in plant location classification based on leaf appearance indicates that the Ffirst methods can be used for classification of environmental conditions and, more generally, for the study of phenotype plasticity.

Acknowledgments. Jiri Matas was supported by Czech Science Foundation Project GACR P103/12/G084, Milan Sulc by Czech Technical University project SGS13/142/OHK3/2T/13.

References

1. Agarwal, G., Belhumeur, P., Feiner, S., Jacobs, D., Kress, W.J., Ramamoorthi, R., Bourg, N.A., Dixit, N., Ling, H., Mahajan, D., et al.: First steps toward an electronic field guide for plants. Taxon **55**(3), 597–610 (2006)
2. Ahonen, T., Matas, J., He, C., Pietikäinen, M.: Rotation invariant image description with local binary pattern histogram fourier features. In: Salberg, A.-B., Hardeberg, J.Y., Jenssen, R. (eds.) SCIA 2009. LNCS, vol. 5575, pp. 61–70. Springer, Heidelberg (2009)
3. Barthélémy, D., Boujemaa, N., Mathieu, D., Molino, J.F., Bonnet, P., Enficiaud, R., Mouysset, E., Couteron, P.: The pl@ntnet project: A computational plant identification and collaborative information system. Tech. Rep., XIII World Forestry Congress (2009)
4. Belhumeur, P.N., Chen, D., Feiner, S.K., Jacobs, D.W., Kress, W.J., Ling, H., Lopez, I., Ramamoorthi, R., Sheorey, S., White, S., Zhang, L.: Searching the world's herbaria: a system for visual identification of plant species. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 116–129. Springer, Heidelberg (2008)

5. Fiel, S., Sablatnig, R.: Automated identification of tree species from images of the bark, leaves and needles. In: Proc. of 16th Computer Vision Winter Workshop, Mitterberg, Austria, pp. 1–6 (2011)
6. Guo, Z., Zhang, D.: A completed modeling of local binary pattern operator for texture classification. IEEE Transactions on Image Processing **19**(6), 1657–1663 (2010)
7. Kadir, A., Nugroho, L.E., Santosa, P.I.: Experiments of zernike moments for leaf identification 1 (2012)
8. Kadir, A., Nugroho, L.E., Susanto, A., Santosa, P.I.: A comparative experiment of several shape methods in recognizing plants. International Journal of Computer Science & Information Technology 3(3) (2011)
9. Kadir, A., Nugroho, L.E., Susanto, A., Santosa, P.I.: Neural network application on foliage plant identification. International Journal of Computer Applications 29 (2011)
10. Kadir, A., Nugroho, L.E., Susanto, A., Santosa, P.I.: Performance improvement of leaf identification system using principal component analysis. International Journal of Advanced Science & Technology 44 (2012)
11. Karuna, G., Sujatha, B., Giet, R., Reddy, P.C.: An efficient representation of shape for object recognition and classification using circular shift method
12. Kim, S.J., Kim, B.W., Kim, D.P.: Tree recognition for landscape using by combination of features of its leaf, flower and bark. In: 2011 Proceedings of SICE Annual Conference (SICE) (2011)
13. 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)
14. Lee, K.B., Chung, K.W., Hong, K.S.: An implementation of leaf recognition system (2013)
15. Lee, K.B., Hong, K.S.: Advanced leaf recognition based on leaf contour and centroid for plant classification. In: The 2012 International Conference on Information Science and Technology, pp. 133–135 (2012)
16. Lin, H.T., Lin, C.J., Weng, R.C.: A note on platts probabilistic outputs for support vector machines. Machine Learning 68(3) (2007)
17. Mäenpää, T., Pietikäinen, M.: Multi-scale binary patterns for texture analysis. In: Bigun, J., Gustavsson, T. (eds.) SCIA 2003. LNCS, vol. 2749, pp. 885–892. Springer, Heidelberg (2003)
18. Nilsback, M.E., Zisserman, A.: An Automatic Visual Flora: Segmentation and Classification of Flower Images. Ph.D. thesis, Oxford University (2009)
19. Novotný, P., Suk, T.: Leaf recognition of woody species in central europe. Biosystems Engineering **115**(4), 444–452 (2013)
20. Ojala, T., Pietikainen, M., Harwood, D.: Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In: Proc. IAPR 1994, vol. 1, pp. 582–585 (1994)
21. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. PAMI **24**(7), 971–987 (2002)
22. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. Pattern Recognition **29**(1), 51–59 (1996)

23. Pietikäinen, M., Ojala, T., Xu, Z.: Rotation-invariant texture classification using feature distributions. Pattern Recognition **33**(1), 43–52 (2000)
24. Platt, J.: Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in Large Margin Classifiers 10(3) (1999)
25. Pydipati, R., Burks, T., Lee, W.: Identification of citrus disease using color texture features and discriminant analysis. Computers and Electronics In Agriculture **52**(1), 49–59 (2006)
26. Qi, X., Xiao, R., Guo, J., Zhang, L.: Pairwise rotation invariant co-occurrence local binary pattern. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part VI. LNCS, vol. 7577, pp. 158–171. Springer, Heidelberg (2012)
27. Shalev-Shwartz, S., Zhang, T.: Stochastic dual coordinate ascent methods for regularized loss minimization. arXiv preprint arXiv:1209.1873 (2012)
28. Söderkvist, O.: Computer vision classification of leaves from swedish trees (2001)
29. Sulc, M., Matas, J.: Kernel-mapped histograms of multi-scale lbps for tree bark recognition. In: 2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ), pp. 82–87 (November 2013)
30. Vedaldi, A., Fulkerson, B.: VLFeat: An open and portable library of computer vision algorithms (2008). http://www.vlfeat.org/
31. Vedaldi, A., Zisserman, A.: Efficient additive kernels via explicit feature maps. PAMI 34(3) (2011)
32. Wu, J., Rehg, J.M.: Centrist: A visual descriptor for scene categorization. IEEE Transactions on Pattern Analysis and Machine Intelligence **33**(8), 1489–1501 (2011)
33. 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: 2007 IEEE International Symposium on Signal Processing and Information Technology, pp. 11–16. IEEE (2007)
34. Zhao, G., Ahonen, T., Matas, J., Pietikainen, M.: Rotation-invariant image and video description with local binary pattern features. IEEE Transactions on Image Processing **21**(4), 1465–1477 (2012)