Advanced shape context for plant species identification using leaf image retrieval

Sofiene Mouine
INRIA Paris-Rocquencourt
78153 Le Chesnay, France
sofiene.mouine@inria.fr

Itheri Yahiaoui
INRIA Paris-Rocquencourt
78153 Le Chesnay, France
CReSTIC Université de Reims, FRANCE
itheri.yahiaoui@inria.fr

Anne Verroust-Blondet
INRIA Paris-Rocquencourt
78153 Le Chesnay, France
anne.verroust@inria.fr

ABSTRACT

This paper presents a novel method for leaf species identification combining local and shape-based features. Our approach extends the shape context model in two ways. First of all, two different sets of points are distinguished when computing the shape contexts: the *voting set*, i.e. the points used to describe the coarse arrangement of the shape and *the computing set* containing the points where the shape contexts are computed. This representation is enriched by introducing local features computed in the neighborhood of the computing points. Experiments show the effectiveness of our approach.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms

Theory

Keywords

Image retrieval, plant species identification, shape context

1. INTRODUCTION

Identifying a plant can be a tricky task even for experienced botanists, considering the huge number of species existing in the world. This task is of great importance for a number of professionals such as land managers, foresters, agronomists, etc., and can be also useful for amateur gardeners. Plant identification is generally based on the observation of the morphological characteristics of the plant (such as general character, structures of stems, roots and leaves, embryology and flowers) followed by the consultation of a guide or a known database. An important amount of information about the taxonomic identity of a plant is contained in its leaves. Moreover, leaves are present on the plants for several months in a year, whereas flowers and fruits may remain only several weeks. This is why most plant identification tools based on Content-Based Image Retrieval techniques work on leaf image databases [17, 29, 20, 24, 31, 10, 4, 7, 23, 25, 6, 3, 8, 9, 13]. A leaf can be characterized by its color, its texture, and its shape. The color of a leaf may vary with the seasons and climatic conditions. In addition, as most plants have similar colors, this feature is not discriminant enough for the species recognition. Thus, generally, only shape and texture information are taken into account in similarity based leaf image retrieval schemes.
Several techniques have been introduced to solve the problem of automatic leaf identification.
Existing methods generally use a shape-based approach. This is not the case for [13] where shape and texture descriptors on oriented patches centered around Harris points are computed and a large scale matching method [16] performs the leaf identification. No prior segmentation is made and Harris points are not necessarily located on the leaf margin. This approach is generic and works well on scans of leaves [14] but may be fastidious for images with a cluttered background.
Shape-based approaches mainly work on the overall shape or on the contour of the leaves.
A first group of methods extracts morphological plant characters commonly used in botany. Du et al. [10] compute eight features, Aspect Ratio, Rectangularity, Convex Area Ratio, Convex Perimeter Ratio, Sphericity, Circularity, Eccentricity and Form Factor, from the boundary of the leaves. Morphological features are also retained and used in the identification process in the parameterized segmentation representation of leaves proposed by Cerruti et al. [9]. Eccentricity is used in the two-stage approach of Wang et al. [29] and of Caballero and Aranda [7] to reduce the search space. Other shape feature extraction techniques [19] have been adapted or introduced to solve the plant retrieval problem. Neto et al. [24] used elliptic Fourier functions on the leaf shape. Statistical methods based on the method of Complex Networks [3, 8] on the contour of the leaf or extracting fractal dimension [6] on the contour and the venation of the leaf have been proposed. Yanikoglu et al. [32] obtain good results on leaf scans with a combination of texture, shape and color descriptors among which some are based on algorithms of mathematical morphology. Different types of 2D histograms using geometric features, such as curvature, lengths, relative orientation, distances, etc., computed on the boundary points of the leaf have been presented in [31,

ICMR '12, June 5-8, Hong Kong, China

18, 7]. In [31], a 2D directional fragment histogram computes directions and relative lengths on a succession of elementary fragments on the contour. Another 2D histogram derived from the shape context [5] computing inner distances and angles between sample points of the leaf margin is proposed in [18, 4]. Curvature of the leaf contour is used in the Curvature Scale Space representation [21, 1] and in Caballero and Aranda 's approach [7]. Note that the shape features computed from the leaf margin can be enriched with venation features [25, 23]. Most of these shape-based approaches are adapted to the leaf identification problem, but their effectiveness may depend on the quality of the contour obtained by the segmentation process.
In this paper, we combine local and shape-based features to obtain an efficient and effective leaf identification method. For this purpose, we select appropriate salient points of the leaf and model local information and spatial relations by a shape context based approach [5].
This paper is structured as follows. A family of shape context based approaches is introduced in section 2. They work on different sets of selected points of the leaf. Local features are then introduced in our model to enrich the image description. Experimental results are presented in section 3. The last section concludes the paper and presents our future work.

2. ADVANCED SHAPE CONTEXT

Shape context technique [5] has proven its efficiency for shape retrieval, even for leaf images, with the inner shape context [18, 4]. To describe properly the boundary of a shape and obtain good retrieval results, a dense sampling of the contour points is necessary. Then a large number of histograms are computed and compared, making the overall technique expensive. To solve this problem, Xie et al. [30] have introduced the skeletal shape context, which uses a medial axis transform to produce an optimal sampling of the shape contour with a smaller number of points.
Shape context retains only global shape information of an image region. To enrich this description Amores et al. [2] have extended the shape context approach for object class-recognition, introducing color and edge information in the histogram. They first perform region segmentation and use the region boundaries as contours of the image. Their "generalized correlograms", encoding both local and spatial information, are computed on a sampled set of contour points. In our case, spatial and local information (see Section 2.2) are computed separately for each salient point and combined through concatenation. Moreover, we use here another strategy to encode the spatial information in the images.
In our shape context based approach, we want to reduce the computational cost while preserving or increasing the shape matching precision. We think that introducing two different sets of points that play different roles in the shape context scheme and choosing them appropriately will help us to achieve this goal. Thus we distinguish the *voting points*, which is the set of points used to build the shape context histograms from the *computing points* on where the shape contexts are computed.
In fact, the computing points correspond to characteristic or salient points of the object and, for efficiency purposes, the cardinality of this set has to be low. The voting points are other points belonging to the object and must add information on the shape when involved in the computation of shape context on the computing points. Here, as these points are used only once for each histogram, we can use a representative number of voting points. The notion of saliency depends on the application and on the type of the considered dataset. For example, for a polygonal shape, the computing points can be the extrema of boundary curvature points and the voting points the boundary points.
Let us now present our approach in more details.

2.1 Advanced shape context

Given a set of n points $\mathcal{V}$ and a point p of $\mathbb{R}^2$, the *advanced shape context* of $\mathcal{V}$ on p is a discrete representation of the set of n vectors defined by the pairs of points (p, q) with $q \in \mathcal{V}$. It is represented by a coarse histogram $aSC(p, \mathcal{V})$ where each pair of points (p, q), represented by a radius r and an angle θ, contributes to the bin k using the log-polar quantization introduced in [5] and used more recently in [2, 18, 28].

$$aSC(p, \mathcal{V})_k = \#\{q \in \mathcal{V} : q - p \in bin_p(k)\}$$

In our implementation, the radius is quantized into 5 bins and the angle into 12 bins.
In the rest of this paper, the set $\mathcal{V}$ is denoted the *voting set* of points and the set $\mathcal{C}$ of points p of $\mathbb{R}^2$, where the advanced shape context $aSC(p, \mathcal{S})$ is computed, is called the *computing set*. The sets $\mathcal{C}$ and $\mathcal{V}$ are not necessarily distinct.
Let us now return to our application.
We suppose in this paper that a leaf image consists in a leaf picture with a nearly white uniform background. A leaf is not only characterized by its margin: its venation network may be significant. Some plant species have a high intra-variability of the leaf shape, and leaves from different species may have globally similar shapes as shown in Figure 1. Thus, in this case, the venation network information may

Figure 1: Top row: Overall shape similarity between different species (*Pittosporum tobira, Arbutus unedo, Rhamnus alaternus*). Bottom row: Intra-variability of the species *Ficus carica*

be useful for an identification task and the shape context based approaches that consider only the leaf margin [18, 4,

30] may be insufficient.
In fact, according to botanists, characteristics (salient) points of a leaf are extrema of curvature on the margin, when leaves are lobed or toothed, and junction points, when the venation network can be extracted from the image.
In our case:
- The leaf margin is a closed contour calculated by using the Otsu thresholding method since the image background is homogeneous.
- The salient points are approximated using the generic Harris detector, which is known to detect edge corners efficiently.

In order to test our approach on images of leaves, three scenarios are proposed by varying the computing set $\mathcal{C}$ and the voting set $\mathcal{V}$ of points in the image.

SCO Spatial relations between margin points.
Here the computing set $\mathcal{C}$ and the voting set $\mathcal{V}$ are identical. They involve the margin points, i.e. n points extracted from the margin by a uniform quantization (as illustrated on the first leaf of Figure 2).

$$\mathcal{C} = \mathcal{V} = \{\text{margin points}\}$$

This scenario corresponds to the shape context proposed by Belongie et al. [5]. Note that the venation network is not introduced here.

SC1 Spatial relations between salient points.
As in the previous case, $\mathcal{C}$ and $\mathcal{V}$ represent the same set. They both contain n salient points computed with the Harris corner detector (the cross points on the second leaf of Figure 2).

$$\mathcal{C} = \mathcal{V} = \{\text{salient points}\}$$

This scenario is similar to the logo retrieval approach [27]. The goal here is to determine whether the spatial relationships between salient points on the leaf area can characterize leaves of a given species and to evaluate the influence of the number of considered salient points on the plant identification performance.

SC2 Spatial relations between salient and margin points
Here we want to measure the spatial relationships between the salient points described in the context defined by the leaf margin (respectively the cross points and the circles of the rightmost leaf of Figure 2). The voting set of points $\mathcal{V}$ is composed of all the margin points. The Harris points form the computing set $\mathcal{C}$.

$$\mathcal{C} \neq \mathcal{V},\ \mathcal{C} = \{\text{salient points}\} \text{ and } \mathcal{V} = \{\text{margin points}\}$$

As mentioned above, the salient points may lay inside the leaf or may belong to the leaf margin. Our aim is to study the correlation between the venation network and the margin of the leaves belonging to the same species.

2.2 Local properties

The advanced shape context captures a spatial configuration of points without taking into account local properties of the image around the set $\mathcal{C}$ of computing points. Thus, to enrich the description, a set of local features computed on the neighborhood of each point of $\mathcal{C}$ is introduced. As the color is not a discriminant feature for leaves, we focus on texture and shape. To describe the texture and the shape, three local features are extracted from the grey-level of an image patch located around each computing point:
- A 16 dimensional Hough histogram, hough_4_4, based on the Hough transform, is used to represent simple shapes in an image [11]. The histogram computes the tangential flow of pixels along the edges in the image projected to the position vector of each pixel.
- A Fourier histogram, fourier_8_32 [11], based on the Fourier transform, which contains information about texture and scale: two histograms are computed in the complex plane from the Fourier transform. They represent two types of distributions of the energy: the first one is computed with a circular partition, the second uses a wedge partition. Both have an equal importance in the final signature, which gives a description of the energy in the image at several scales as well as a description of the local small scale behavior in a number of predefined directions.
- A 8 dimensional classical Edge Orientation Histogram, which is known to be suitable for non-uniform textures.
These three features have given promising results when associated with Harris points on scans of leaves in [14].
In the following, the combination of the Hough, the Fourier and the edge orientation histogram is denoted by *std*.

Figure 2: From left to right: points used in scenario SC0, SC1 and SC2. The small circles represent the sample points on the leaf margin. The cross points represent the salient points computed with Harris detector.

2.3 Matching Method

The matching process is the same for all the scenarios.
Let I be an image and let n be the number of points of I in the computing set $\mathcal{C}_I$. I is represented by n feature vectors $F_1, F_2, ..., F_n$ where F_i is the feature vector associated to the i^{th} point of $\mathcal{C}$. F_i is of dimension k with:

$$k = nbins_r \times nbins_\theta \text{ or } k = nbins_r \times nbins_\theta + n_{std}$$

where $nbins_r$ and $nbins_\theta$ are respectively the number of bins for quantified log-polar distances and the number of bins for quantified angles and n_{std} is the size of the std feature vector presented in Section 2.2.
The features matching, is done by an approximate similarity search technique based on a Locality Sensitive Hashing (LSH) method [26]. We use here the Multi Probe Locality Sensitive Hashing technique proposed by Joly and Buisson [15] and the distance L_2 to compute the similarity between two feature vectors. The principle of this algorithm is to project all the features in an L dimensional space and to use hash functions to reduce the search and the cost time. At query time, the features $F_1, F_2, ..., F_n$ of the query image are mapped onto the hash tables and the knn nearest neighbors

of each feature F_i are searched in the buckets associated to F_i. These n lists of candidate feature matches are used as input for a voting system to rank images according to the number of matched features.
One can notice here that if scenario SC0 has the same histograms as the shape context methods [5, 22], the overall approach is not identical, as our matching method is different. It is the same for the SC1 scenario and the logo retrieval approach of [27] as, in their case, the LSH approximate search is performed on shapeme histograms.

3. EXPERIMENTAL RESULTS

All the approaches presented above have been tested on the Pl@ntLeaves dataset that was used for the plant identification task organized within ImageCLEF 2011 [12]. The Pl@ntLeaves dataset contains three categories of images: scans of leaves acquired using a flat-bed scanner, scan-like leaves acquired using a digital camera and free natural photos. We first tested our approaches on the training subset of scans of leaves, which contains 2349 images. Then all the scan and scan-like images of the ImageCLEF2011 (train +test images) have been used to evaluate the previous scenarios (except scenario SC1 whose performance was not convincing). We also compared our results with the scores of identification of ImageCLEF2011 obtained on two categories of images: scans and scan-like leaf images. Figure 3 gives a view of the different species present in the database.

Figure 3: A selection of scan-like leaves (first two rows) and scan leaves of the Pl@ntNet test dataset (Only one leaf per species is kept).

To evaluate the three scenarios SC0, SC1 and SC2, we use the precision P and recall R measures defined respectively by

$$P = \frac{\#\text{relevant images}}{\#\text{retrieved images}} \quad \text{and}$$

$$R = \frac{\#\text{retrieved relevant images}}{\#\text{relevant images}}$$

and the Mean Average Precision (MAP). It is measured on a set of queries Q and is defined as follows:

$$MAP = \frac{\sum_{q \in Q} AP(q)}{|Q|}$$

where the average precision score $AP(q)$ is computed for each query q:

$$AP(q) = \frac{\sum_{k=1}^{n} (P(k) \text{ x } f(k))}{\#\text{retrieved relevant images for q}}$$

$P(k)$ is the precision at cut-off k in the list of retrieved images and $f(k)$ is equal to 1 when the image at rank k is relevant and 0 otherwise. The MAP value is correlated with the precision value P. It gives a more general result taking into account all the possible queries of the database.

3.1 Results on the training scan dataset

Scenarios SC0, SC1, SC2 have been tested with computing sets of points of cardinality 50 and 400. Precision Recall curves are shown in Figure 4. The std curve corresponds to the retrieval using only Harris points associated with local features around them. Enriched SC2 is presented by the curve SC2+std.

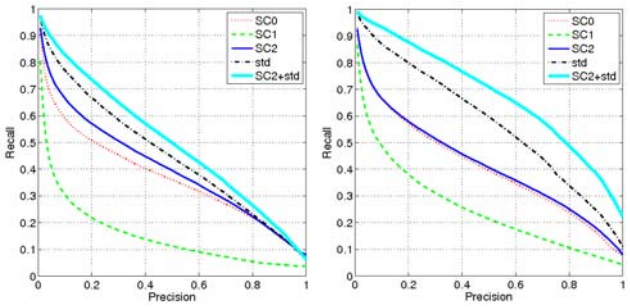


Figure 4: Recall/Precision curves on the training scan dataset using 50 points (Left) and 400 points (right).

Approach	SC0	SC1	SC2	std	SC2+std
MAP (50 points)	0.36	0.14	0.40	0.45	0,50
MAP (400 points)	0.40	0.25	0.42	0.57	0,68

Table 1: MAP values for the different scenarios on the training scan dataset

Examining these precision-recall curves and the MAP values of Table 1, we can notice that:

- All the scenarios give better results with 400 points compared to 50 points.
- The information provided by the Scenario SC1 is less meaningful compared with other scenarios. This is due to the fact

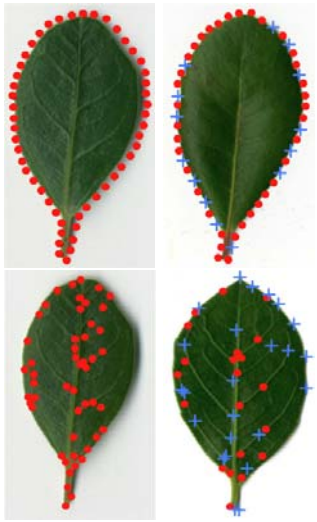
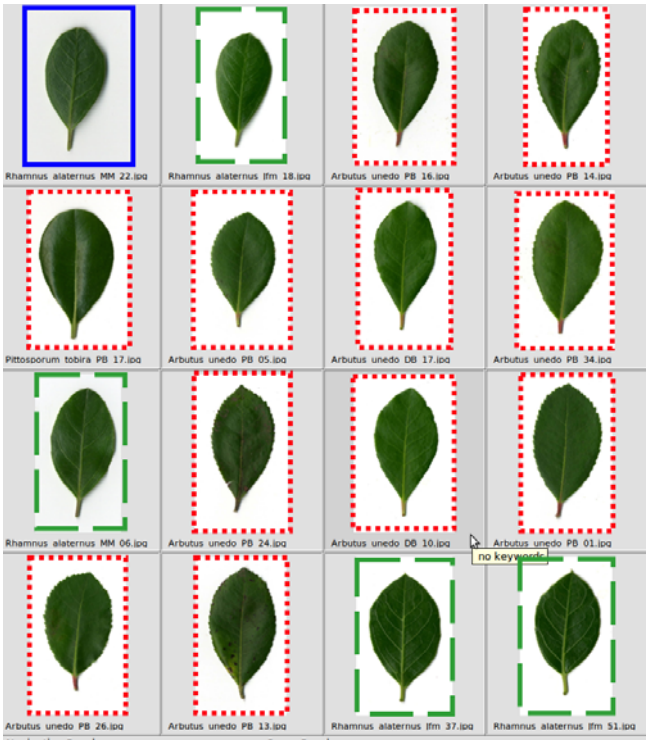


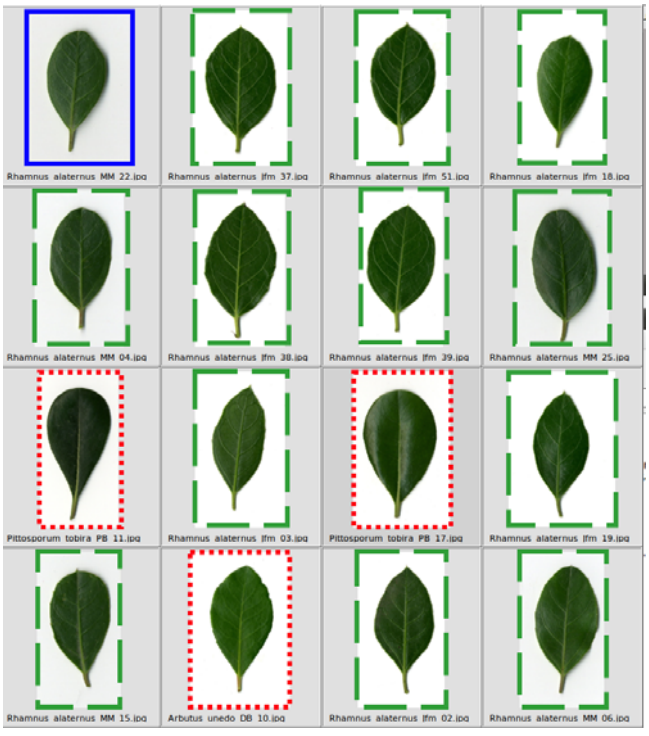
Figure 5: Top row: the query image used in Figure 6 and a response example using SC0, red points on the right image are the points that matched with contour points of the query image. Bottom row: the query image used in Figure 6 and a response example using SC2, red points on the right image are the points that matched with Harris points of the query image

that, when only 50 Harris points are computed, they provide a very coarse representation of the margin and the venation network.
- Scenario SC2 obtains the best R/P curve and MAP value compared to SC1 and SC0. However, the std method that computes local features around Harris points, is more effective than SC2. Moreover, the results are improved when local features are associated to Harris points in scenario SC2, which corresponds to the enriched SC2 (SC2+std).
- The enriched SC2 (SC2+std) outperforms all the other scenarios. It has the best Recall/Precision curves and the best MAP value. When 400 computing points are used, the MAP value of SC2+std is 0.68. This proves that the joint use of spatial information and local information increases the identification rate on the training Scan dataset.
- SC2 obtains good results with a small number of computing points. When SC2 is used with only 50 points, it gives similar performances to SC0 with 400 computing points. This enables to speed up the running time of the matching step without a loss of performance. In fact, in this case, most of significant points, from a botanic point of view, are detected and the spatial information computed with SC2 is relevant. When 400 points are used, SC2 results increase slightly: when all the characteristic points of the leaf have been found, the remaining Harris points will not increase the accuracy of the spatial information. However, SC2+std has a better score thanks to the local texture and shape information added around the computing points.

Let us compare the retrieval results obtained using scenarios SC0 and SC2 on a Rhamnus alaternus leaf in Figure 6.
With SC0, only four images are relevant among the first 15 returned images which gives a precision $P = 27\%$ for $knn = 15$. In fact, the overall shape of all the returned



(a)



(b)

Figure 6: Two retrieval responses with the same query image using SC0 for (a) and SC2 for (b). Query image Q is framed by a solid line, relevant retrieved images i.e leaves from the same species of Q are framed by a dashed line and false positive images are framed by a dotted line

	# of images	# of individual plants	# of users
Scan train	2349	151	17
test	721	55	13
Scan-like train	717	51	2
test	180	13	1

Table 2: Statistics of the composition of the training and test data of the Pl@ntLeaves dataset

leaves is very similar. Consequently, additional information about the leaf taxonomy is needed to be robust to shape similarity between the species.

This was our first motivation to build SC2 by separating two sets of points. Figure 6(b) shows the retrieval response using SC2 without including local features. The precision for the first 15 returned images is $P = 80\%$ which is much higher than the results of SC0. This can be explained by the fact that SC2 includes, with the use of salient points, informative characters within the leaf area such as venation, texture, etc. Note that the information on the contour is not lost in this schema since the voting set contains points of the leaf margin. Moreover, salient points computed with a Harris detector may be located on the boundary, in particular, for the toothed leaves contour.

In Figure 5, we show two retrieval results with the same query image of Figure 6. In the case of SC0, an important number of points of the returned image matched with the contour points of the query image. However, the returned image belongs to another species. In fact, the entire configuration of the shape is very similar but the venation network is different.

Thus, the spatial representation provided by SC0 may not be appropriate in this particular case. On the other hand, matched points computed with SC2 are either on the nervation, on the contour or inside the leaf. The main advantage of SC2 is that the contour is not considered as an exclusive source of information. Here, the retrieved image and the query image belong to the same species.

3.2 Comparison with ImageCLEF2011 results

Let us now introduce the context of the organized identification task ImageCLEF 2011[12]. The goal of the task was to find the correct tree species of each test image. The identification score is quite different from the classic measures presented above such as the MAP value and recall-precision curves. Two assumptions guided the identification score S definition:

- The leaves from the same tree may be more similar than leaves from different trees (the classification rate on each individual plant is averaged).
- Photos taken by the same person will have nearly the same acquisition protocol (S measures the mean of the average classification rate per user).

Then, S is defined as follows in ImageCLEF 2011:

$$S = \frac{1}{U} \sum_{u=1}^{U} \frac{1}{P_u} \sum_{p=1}^{P_u} \frac{1}{N_{u,p}} \sum_{p=1}^{N_{u,p}} s_{u,p,n}$$

U: number of users (who have at least one image in the test data).

species	Test ScL	Test Sc	Train Sc
Acer campestre	9	22	24
Acer monspessulanum		22	45
Acer negundo		21	17
Acer platanoides		10	12
Aesculus hippocastanum		4	25
Albizia julibrissin			45
Alnus glutinosa			8
Arbutus unedo	1	41	34
Betula pendula		3	76
Carpinus Betulus			33
Castanea sativa		45	24
Celtis australis		24	39
Cercis siliquastrum	11	20	49
Corylus avellana		20	55
Cotinus coggygria	29		64
Crataegus azarolus			38
Crataegus monogyna		54	23
Diospyros kaki	7		43
Eriobotrya japonica		21	14
Fagus sylvatica			28
Ficus carica	19		43
Fraxinus angustifolia			74
Fraxinus ornus			33
Ginkgo biloba		15	34
Gleditsia triacanthos			32
Ilex aquifolium	26		40
Juglans nigra	22		16
Juglans regia		4	30
Laburnum anagyroides			30
Laurus nobilis	16	6	36
Ligustrum vulgare			26
Magnolia grandiflora			23
Malus sylvestris			17
Nerium oleander	8		88
Olea europaea		25	125
Paliurus spina-christi		7	65
Phillyrea angustifolia			15
Pistacia lentiscus		36	41
Pistacia terebinthus			47
Pittosporum tobira			67
Platanus x hispanica	2		49
Prunus mahaleb			49
Prunus serotina			24
Prunus spinosa			40
Punica granatum			38
Quercus coccifera		33	16
Quercus ilex	29	18	123
Quercus petraea			14
Quercus pubescens		18	24
Rhamnus alaternus		52	54
Rhamnus cathartica			27
Robinia pseudoacacia		26	32
Salix caprea			26
Sambucus nigra			9
Sophora japonica			32
Sorbus domestica			21
Syringa vulgaris		20	51
Tilia cordata			23
Viburnum lantana			17
Viburnum tinus		94	47
Vitex agnus-castus		2	55

Table 3: Species appearing either on the test scan or on the scan-like dataset versus the train scan dataset on Pl@ntLeaves database of ImageCLEF 2011.

P_u: number of individual plants observed by the u^{th} user.
$N_{u,p}$: number of pictures taken from the p^{th} plant observed by the u^{th} user.
$s_{u,p,n}$: classification score (1 or 0) for the n^{th} picture taken from the p^{th} plant observed by the u^{th} user.

We focus on scans and scan-like images. Table 2 and Table 3 describe the number of images and their distribution between the different datasets.

The evaluation metric S is used to compare our approach with the others. Only the top 10 scores of ImageCLEF2011 are presented in Table 4. More details about the methods and the complete list of scores can be found in [12]. We present here the score of SC0 and SC2 using 50 computing points.
Both SC0 and SC2 outperform all the other approaches on the scan-like dataset. We obtain the best identification score using SC0 (S=0.706). SC2 is also better than the top ten scores of ImageCLEF 2011 with a score S=0.677. In fact, Harris detector is known to be more robust to illumination changes than other interest points detectors. But in scan-like images, some Harris points can be located on light reflection points, which are not characteristic points of the leaf shape. That is why the spatial description provided by SC2 is less accurate than SC0.
We also obtain good results on the scan images. SC2 performs better than SC0 on the scans and they are respectively positioned in the third and the fourth place with respect to ImageCLEF 2011 runs.
In scans and scan-like dataset, SC2+std, which had the best result on the training scan dataset, loses its first position.

run_id	Scans	Scan-like
IFSC_USP_run2	**0.562**	0.402
inria_imedia_plantnet_run1	**0.685**	0.464
IFSC_USP_run1	0.411	0.430
LIRIS_run3	0.546	0.513
LIRIS_run1	0.539	**0.543**
Sabanci-okan-run1	**0.682**	0.476
LIRIS_run2	0.530	0.508
LIRIS_run4	0.537	**0.538**
inria_imedia_plantnet_run2	0.477	**0.554**
IFSC_USP_run3	0.356	0.187
SC0	**0.654**	**0.706**
SC2	**0.676**	**0.677**
SC2+std	**0.650**	**0.590**

Table 4: Normalized classification scores of the scan and scan-like images on the Pl@ntLeaves dataset using the evaluation metric of [12]. Top 6 results per image type are highlighted in bold.

Approach	SC0	SC2	std	SC2+std
MAP (50 points)	0.32	0.36	0.32	0,44

Table 5: MAP values for the whole Pl@ntLeaves dataset (scan and scan-like, train and test images)

We have computed the MAP values for the whole Pl@ntLeaves dataset. These results are presented in Table 5.
Now, it is clear that our method performs better on MAP: SC2 is better than SC0 and SC2+std obtains the best MAP value. That confirms the results obtained in the training scan dataset. The ranking difference between MAP and the ImageCLEF scores may be due to the normalizations made on users and individual plants while computing the score S. In fact, by examining the statistics of the composition of the Pl@ntLeaves dataset in Table 2, we notice that the number of users is much smaller than the number of images (17 users for 1349 images for the traning scan dataset) and that the number of images is also about ten times the number of individual plants. Then the difference between MAP and S scores can be explained by the fact that the combination of local and spatial relationships between Harris points is more efficient to retrieve almost identical images (taken by the same user or images of leaves belonging to the same individual plant) than similar ones.

4. CONCLUSION

In this paper, we have presented a new approach extending the shape context method for plant species identification on leaf images. To compute histograms, two sets of points have been introduced: the computing set and the voting set. Several scenarios based on different computing and voting sets have been proposed and tested on scans and scan-like leaf images. Scenarios SC0 and SC1 work respectively only on leaf margin points and only on salient points. Scenario SC2 uses both contour points and salient points of the leaf in the spatial representation. Experiments on the training scan dataset of ImageCLEF 2011 have shown that combining the contour and the local texture and shape information is suitable for leaves description. In some cases, as in Figure 5, the leaf margin information is not sufficient for the species identification, which strengthens our choice of adding salient points in the spatial description. This is one of the main contribution of our advanced shape context method compared to other shape context approaches [18, 4, 30].
We have shown the effectiveness of the enriched SC2 on scans of leaves: SC2+std obtains the best Recall/Precision curves and the best MAP value. Furthermore, our approach has been compared favorably to ImageCLEF2011 runs.
The notion of saliency depends on the application. In this framework, we used a generic corner detector (Harris) to compute salient points on the images. Work in progress extends this approach in two ways:
- We are studying specific detectors for leaf salient points based on botanical expertise
- We are investing new scenarios using venation network and margin points.

Acknowledgements

This research has been conducted with the support of the Agropolis Foundation through the Pl@ntNet project. We would like also to thank Vera Bakić for her help to accomplish this work and Richard James for revising the English of this paper.

5. REFERENCES

[1] S. Abbasi, F. Mokhtarian, and J. Kittler. Reliable classification of chrysanthemum leaves through curvature scale space. In *Scale-Space Theory in Computer Vision*, volume 1252, pages 284–295. LNCS, 1997.

[2] J. Amores, N. Sebe, and P. Radeva. Context-based object-class recognition and retrieval by generalized correlograms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(10):1818 –1833, Oct. 2007.

[3] A. R. Backes, D. Casanova, and O. M. Bruno. A complex network-based approach for boundary shape analysis. *Pattern Recognition*, 42(1):54–67, 2009.

[4] P. Belhumeur, D. Chen, S. Feiner, D. Jacobs, W. Kress, H. Ling, I. Lopez, R. Ramamoorthi, S. Sheorey, S. White, and L. Zhang. Searching the world's herbaria: A system for visual identification of plant species. In *European Conference on Computer Vision (ECCV)*, pages 116–129, 2008.

[5] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509 –522, apr 2002.

[6] O. M. Bruno, R. de Oliveira Plotze, M. Falvo, and M. de Castro. Fractal dimension applied to plant identification. *Information Sciences*, 178(12):2722 – 2733, 2008.

[7] C. Caballero and M. C. Aranda. Plant species identification using leaf image retrieval. In *ACM International Conference on Image and Video Retrieval (CIVR)*, pages 327–334, 2010.

[8] D. Casanova, J. B. Florindo, and O. M. Bruno. IFSC/USP at ImageCLEF 2011: Plant identication task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.

[9] G. Cerutti, L. Tougne, A. Vacavant, and D. Coquin. A parametric active polygon for leaf segmentation and shape estimation. In *International Symposium on Visual Computing (ISVC)*, pages 202–213, 2011.

[10] J.-X. Du, X.-F. Wang, and G.-J. Zhang. Leaf shape based plant species recognition. *Applied Mathematics and Computation*, 185(2):883 – 893, 2007.

[11] M. Ferecatu. *Image retrieval with active relevance feedback using both visual and keyword-based descriptors*. PhD thesis, University of Versailles Saint-Quentin-en-Yvelines, 2005.

[12] H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthelemy, J.-F. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The CLEF 2011 plant images classification task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.

[13] H. Goëau, A. Joly, S. Selmi, P. Bonnet, E. Mouysset, L. Joyeux, J.-F. Molino, P. Birnbaum, D. Bathelemy, and N. Boujemaa. Visual-based plant species identification from crowdsourced data. In *19th ACM international conference on Multimedia*, pages 813–814, 2011.

[14] H. Goëau, A. Joly, I. Yahiaoui, P. Bonnet, and E. Mouysset. Participation of INRIA& Pl@ntNet to ImageCLEF 2011 plant images classification task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.

[15] A. Joly and O. Buisson. A posteriori multi-probe locality sensitive hashing. In *16th ACM international conference on Multimedia*, pages 209–218, 2008.

[16] A. Joly and O. Buisson. Random maximum margin hashing. In *CVPR*, pages 873–880, 2011.

[17] T. Kunii, C. Im, and H. Nishida. Recognizing plant species by leaf shapes: A case study of the acer family. In *14th International Conference on Pattern Recognition (ICPR)*, pages Vol II: 1171–1173, 1998.

[18] H. Ling and D. Jacobs. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):286 –299, Feb. 2007.

[19] Y. Mingqiang, K. Kidiyo, and R. Joseph. A survey of shape feature extraction techniques. *Pattern Recognition Techniques, Technology and Applications*, 2008. ISBN: 978-953-7619-24-4, InTech.

[20] F. Mokhtarian and S. Abbasi. Matching shapes with self-intersections: application to leaf classification. *IEEE Transactions on Image Processing*, 13(5):653 –661, May 2004.

[21] F. Mokhtarian, S. Abbasi, and J. Kittler. Robust and efficient shape indexing through curvature scale space. In *British Machine Vision Conference (BMVC)*, 1996.

[22] G. Mori, S. J. Belongie, and J. Malik. Efficient shape matching using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(11):1832–1837, 2005.

[23] Y. Nam, E. Hwang, and D. Kim. A similarity-based leaf image retrieval scheme: Joining shape and venation features. *Computer Vision and Image Understanding*, 110(2):245 – 259, 2008.

[24] J. C. Neto, G. E. Meyer, D. D. Jones, and A. K. Samal. Plant species identification using elliptic Fourier leaf shape analysis. *Computers and Electronics in Agriculture*, 50(2):121 – 134, 2006.

[25] J. Park, E. Hwang, and Y. Nam. Utilizing venation features for efficient leaf image retrieval. *Journal of Systems and Software*, 81(1):71 – 82, 2008.

[26] L. Paulevé, H. Jégou, and L. Amsaleg. Locality sensitive hashing: A comparison of hash function types and querying mechanisms. *Pattern Recognition Letters*, 31(11):1348–1358, 2010.

[27] M. Rusiñol and J. Lladós. Efficient logo retrieval through hashing shape context descriptors. In *9th IAPR International Workshop on Document Analysis Systems (DAS)*, pages 215–222, New York, NY, USA, 2010. ACM.

[28] E. Shechtman and M. Irani. Matching local self-similarities across images and videos. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1 –8, june 2007.

[29] Z. Wang, Z. Chi, and D. Feng. Shape based leaf image retrieval. *IEE Proceedings on Vision, Image and Signal Processing*, 150(1):34 – 43, Feb. 2003.

[30] J. Xie, P.-A. Heng, and M. Shah. Shape matching and modeling using skeletal context. *Pattern Recogn.*, 41:1756–1767, May 2008.

[31] I. Yahiaoui, N. Hervé, and N. Boujemaa. Shape-based image retrieval in botanical collections. In *7th Pacific Rim Conference on Multimedia (PCM)*, volume 4261, pages 357–364, 2006.

[32] B. A. Yanikoglu, E. Aptoula, and C. Tirkaz. Sabanci-Okan system at ImageClef 2011: Plant identification task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.