A shape-based approach for leaf classification using multiscale triangular representation
Sofiène Mouine, Itheri Yahiaoui, Anne Verroust-Blondet

To cite this version:
Sofiène Mouine, Itheri Yahiaoui, Anne Verroust-Blondet. A shape-based approach for leaf classification using multiscale triangular representation. ICMR ’13 - 3rd ACM International Conference on Multimedia Retrieval, Apr 2013, Dallas, United States. hal-00818115

HAL Id: hal-00818115
https://inria.hal.science/hal-00818115
Submitted on 26 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A shape-based approach for leaf classification using multiscale triangular representation.

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

Itheri Yahiaoui
Inria Paris-Rocquencourt
78153 Le Chesnay, France
itheri.yahiaoui@inria.fr

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

ABSTRACT
In this paper we introduce a new multiscale shape-based approach for leaf image retrieval. The leaf is represented by local descriptors associated with margin sample points. Within this local description, we study four multiscale triangle representations: the well known triangle area representation (TAR), the triangle side lengths representation (TSL) and two new representations that we denote triangle oriented angles (TOA) and triangle side lengths and angle representation (TSLA). Unlike existing TAR approaches, where a global matching is performed, the similarity measure is based on a locality sensitive hashing of local descriptors. The proposed approach is invariant under translation, rotation and scale and robust under partial occlusion. Evaluations made on four public leaf datasets show that our shape-based approach achieves a high retrieval accuracy w.r.t. state-of-art methods.

keywords: shape descriptor, leaf image retrieval, plant identification, multiscale triangle representation, local description

1. INTRODUCTION
The large number of existing plant species in the world makes human identification of them tedious and time consuming, particularly for non-expert stakeholders such as land managers, foresters, agronomists, amateur gardeners, etc. Hence, an automatic plant identification tool should speed up the plant species identification task. This identification tool may be useful even for experienced botanists. Plant identification is based on the observation of its organs, i.e. buds, leaves, fruits, stems, etc. An interesting review of existing approaches developed for plant species identification can be found in [11]. A large amount of information about the taxonomic identity of a plant is contained in its leaves. This is due to the fact that leaves are present on the plants for at least several months, which is not generally the case for other organs such as fruits or flowers. Therefore, most plant identification tools based on Content-Based Image Retrieval techniques [20, 27, 5, 8, 7, 3, 9, 10] work on leaf image databases. Leaves can be characterized by their shape, color and texture. Leaf color may vary with the seasons and geographical locations. In addition, different plant species can have almost the same color leaves. Thus, color is not sufficiently discriminant to be used alone in a plant identification task.

In this paper, we focus on the shape of the leaf and on shape-based approaches for leaf recognition. To describe the shape of a leaf, one can develop a specific approach or adapt a generic shape retrieval method to the particular case of leaves. Specific approaches [13, 8] are based on the botanical characterization of leaf shapes. They extract morphological characters such as: Aspect Ratio, Rectangularity, Convex Area Ratio, Convex Perimeter Ratio, Sphericity, Circularity, Eccentricity and Form Factor. Shape feature extraction techniques (cf. [26] for a survey) can be subdivided into two families: global approaches, where the shape is represented by one feature descriptor, and local approaches, where a set of local descriptors are computed at some interesting points of the shape. When global features are extracted, a global measure is used to compute the similarity of the shapes. The Curvature Scale Space approach [28] has been tested on leaves in [28, 8]. Fourier-based descriptors have also been used [41, 32, 30]. Multiscale approaches [20, 28, 2, 12, 23] have been introduced to enrich the shape description and render it more robust to noise and contour deformations. Local approaches compute local features of landmark points of the object. Landmark points can be boundary points [6, 25] or salient points [29] of the shape. Then, a feature-to-feature matching is performed to retrieve the most similar pairs of points of two different shapes. A 2D histogram derived from the shape context [6] computing inner distances and angles between sample points of the leaf margin has been proposed in [25, 5]. Local approaches obtained good results on the Swedish leaf dataset [25] and on the ImageCLEF2011 plant identification task [29]. We want to benefit here from the advantages of both the multiscale approaches and the local ones. For this purpose,
we associate a multiscale local description of the shape based on triangles to a sample of the shape contour points. Before describing our approach, let us study the possible triangle representations.

Several authors have proposed representations based on triangles built from feature points \cite{18, 19, 36}, Tao an Grosky \cite{36} describe the shape by a Delaunay triangulation and estimate the density of triangles discrete angles by a global histogram. In \cite{18, 19}, the authors use the angles given by the medians of the triangles, joining the labelled feature points to encode their spatial relationships.

Multiscale schemes based on triangles have been introduced to describe the contour of a shape \cite{21, 34, 33, 14, 2}. All of them represent the triangles by their areas at each scale. Shen et al. \cite{21} showed that the triangle area representation (TAR) is affine-invariant and proposed a fast error minimization algorithm for computing correspondence matching. El Rube et al. \cite{14} suggested computing TAR at multiscale wavelet levels (MTAR) to reduce the noise effect on the shape boundary. More recently, Alajlan et al. \cite{2, 1} made the triangle normalization locally for each scale and used a dynamic space warping matching to compute the optimal correspondence between two shapes.

Although TAR is affine-invariant and robust to noise and deformation, it has a high computational cost since all the boundary points are used. Moreover, TAR has two major limitations:

- The area is not informative about the type of the triangle (isosceles, equilateral, etc.) considered, which may be crucial for a local description of the boundary.
- The area is not accurate enough to represent the shape of a triangle. Figure 1 shows three triangles that are equal in area but which have different shapes.

In fact, the triangles in Figure 1 are not similar triangles. They do not fulfill any of the following three properties:

(i) All three pairs of corresponding side lengths are in the same proportion.
(ii) All three pairs of corresponding angles are the same.
(iii) Two pairs of side lengths have the same proportion and the included angle is equal.

To our knowledge, the Triangle Area Representation has not been really compared to other triangle representations (using side lengths, angles, etc.). We want here to take into account the similarity property in our triangle representation. Thus, in Section 2, we present three other triangle descriptions (TSL, TSLA and TOA), based on the triangles side lengths, their angles, or both, and integrate these representations in a multiscale based approach. Experiments are made on four public leaf datasets and presented in Section 3.

![Figure 1: Three triangles having the same area](image)

**2. MULTISCALE TRIANGLE REPRESENTATION**

The shape boundary is represented by a sequence of \( N \) sample points \( p_1, ..., p_N \) uniformly distributed over the contour and numbered in a clockwise order. Then, each contour point \( p_i \) is represented by \( N_i \) triangles computed at different scales (see Figure 2). \( N_i \) is then the number of triangles and the number of scales. Unlike other multiscale triangular representations, we introduce \( d(k) \) the distance between the triangle points at scale \( k \), expressed in the number of boundary points, with \( 1 \leq k \leq N_s \) and \( d \) being an increasing function such that \( d(N_s) \geq N/2 \). In addition, to describe \( p_i \), we do not systematically use all the remaining boundary points. We select only two sets of \( N_i \) points on both sides of \( p_i \). The choice of \( N_i \) depends on whether we are seeking to capture local or global information. The distance \( d(k) \) may be either uniform or logarithmic (cf. Figure 2).

In what follows, each boundary point \( p_i \) is associated with \( N_i \) triangles \( T_i^1, ..., T_i^{N_i} \), \( T_k \) being the triangle defined by the contour points \( p_{i-d(k)}, p_i \) and \( p_{i+d(k)} \), \( 1 \leq k \leq N_s \). Four triangle representations associated to each \( p_i \) are introduced:

- TAR\((p_i) = (\text{TAR}(T_i^1), ..., \text{TAR}(T_i^{N_i})), \)
- TSL\((p_i) = (\text{TSL}(T_i^1), ..., \text{TSL}(T_i^{N_i})), \)
- TSLA\((p_i) = (\text{TSLA}(T_i^1), ..., \text{TSLA}(T_i^{N_i}))) \) and
- TOA\((p_i) = (\text{TOA}(T_i^1), ..., \text{TOA}(T_i^{N_i}))) .

The shape is then described by \( N \) feature vectors \( T(p_i); 1 \leq i \leq N \), \( T \) being either TAR, TSL, TSLA or TOA triangle representation. We will show that all these triangle representations are invariant to translation and rotation of the shape. By normalizing the description locally, we also obtain a scale invariant description of the shape.

**2.1 Triangle area representation(TAR)**

Here, for each triangle \( T \), TAR\((T) = A(T) \), where \( A(T) \) is the signed area of \( T \). TAR is affine-invariant, robust to noise and provides information about local concavities or convexities at a given boundary point as the signed area is computed. Let \((x_i-d(k), y_i-d(k)), (x_i, y_i)\) and \((x_i+d(k), y_i+d(k))\) be the respective coordinates of the points \( p_{i-d(k)}, p_i \) and

![Figure 2: Multiscale triangular representation. (a) \( N \) boundary points of the leaf. Here \( N = 22 \). (b) \( N \) boundary points are selected on each side of \( p_i \) and \( p_i \) is represented by two triangles \( (N_i = 2 \text{ with } d(1) = 2 \text{ and } d(2) = 4) \) (c) \( p_i \) is represented by four triangles and logarithmic distance between triangle points \( (N_i = 4 \text{ with } d(1) = 4, d(2) = 2, d(3) = 4 \text{ and } d(4) = 8) \)
The signed area of the triangle formed by this triplet is given by:

\[ A(T^k_i) = \frac{1}{2} \begin{vmatrix} x_{i-d(k)} & y_{i-d(k)} & 1 \\ x_i & y_i & 1 \\ x_{i+d(k)} & y_{i+d(k)} & 1 \end{vmatrix} \]

A normalization is made locally with respect to the maximum area in each scale as in [2]. However, the TAR used in this paper is quite different from the original TAR.

- We consider here a subset of points on the contour, unlike the original TAR where all boundary points are used.
- We define the number of scales as a parameter \( N_s \). In [2], \( \frac{N}{2} - 1 \) scales are systematically used where \( N \) is the number of points on the contour.
- The matching process is different. A dynamic space wrapping is used in [2] to compare global signatures of the shapes at each scale. Here, the feature associated to a contour point takes into account all the selected scales and then a similarity measure based on a locality sensitive hashing is used to find similar points.

### 2.2 Triangle side lengths representation (TSL)

TSL uses only the side lengths to represent a triangle. Let \( L_{1k}, L_{2k} \) and \( L_{3k} \) be the three sides lengths sorted in ascending order \( (L_{1k} \leq L_{2k} \leq L_{3k}) \) of triangle \( T^k_i \) formed by the points \( p_{i-d(k)}, p_i \) and \( p_{i+d(k)} \), \( k \in \{1, ..., N_s\} \) of the shape contour. Let \( M_k = L_{1k}/L_{3k} \) and \( N_k = L_{2k}/L_{3k} \). Then TSL(\( T^k_i \)) = \( (M_k, N_k) \).

The three side lengths of \( T^k_i \) are proportional to \( M_k, N_k \) and 1; this is also the case for any triangle similar to \( T^k_i \). Thus similar triangles have an equal TSL representation and TSL is invariant under scale, translation, rotation and reflection around the contour points.

Figure 3 shows an example of four triangles having the same TSL representation. T2 is a result of a rotation of T1 around \( p_i \) while T3 is the mirror image of T1 w.r.t. a horizontal line. Note that the vertex angle \( p_i \) of \( T_4 \) is different from the vertex angle \( p_i \) of the other triangles. However, the TSL representation is the same since the triangle side lengths are sorted.

### 2.3 Triangle represented by two side lengths and an angle (TSLA)

Let \( \theta \) be the absolute value of the vertex angle at point \( p_i \) of triangle \( T^k_i \). TSLA representation of \( T^k_i \) is the triplet \( (M_k, N_k, \theta) \), where \( (M_k, N_k) = \text{TSL}(T^k_i) \). TSLA is more accurate than TSL. For example, the similar triangles \( T_1 \) and \( T_4 \) in Figure 3 have different TSLA representations because the respective angles at \( p_i \) are distinct. On the other hand, TSL(\( T_1 \)) and TSL(\( T_3 \)) are equal. Like TSL, the TSLA representation is invariant under scale and reflection around the contour points.

### 2.4 Triangle represented by two oriented angles (TOA)

TOA uses only angle values to represent a triangle. Let \( \phi_k = \angle p_{i-d(k)}p_ip_{i+d(k)} \) and \( \psi_k = \angle p_{i-d(k)}p_{i+d(k)}p_{i-d(k)} \) two successive oriented angles of triangle \( T^k_i \). Then TOA(\( T^k_i \)) = \( (\phi_k, \psi_k) \).

The angle orientation provides information about local concavities and convexities (cf. Figure 4). In fact, an obtuse angle means convex, an acute angle means concave. TOA is not invariant under reflection around the contour point: here, only similar triangles having equal angles at \( p_i \) will have equal TOA values.

### 2.5 Matching Method

TSL, TSLA and TOA represent local descriptions of contour points. In fact, a feature vector \( F_i \) is associated to each contour point \( p_i \), \( i \in \{1, ..., N\} \). The size of the signature using TSL, TSLA and TOA depends on the number of scales \( N_s \).

Size([TSL(\( p_i \))]) = Size([TOA(\( p_i \))]) = 2 \times N_s.

Size([TSLA(\( p_i \))]) = 3 \times N_s.

Size([TAR(\( p_i \))]) = N_s.

When a small number of scales is used, we obtain a compact representation of each contour point. The matching process is a feature-to-feature comparison. It is the same for all the triangle descriptors presented above. The features matching is done by an approximate similarity search technique based on a Locality Sensitive Hashing (LSH) method [31]. We use the Multi Probe Locality Sensitive Hashing technique [22] 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 \( k \)-nearest neighbors \( (k \leq n \times n) \) of each feature \( F_i \) are searched for 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.

### 3. Experimental Results on Leaves

Our descriptors have been tested on four leaf datasets: the Swedish leaf dataset [35], the Flavia dataset [38], the ImageCLEF dataset in 2011 [16] and in 2012 [17]. In all the
experiments, a leaf image contains a single leaf on an uncluttered background. A preprocessing step is required to isolate the leaf area. First, we apply the Otsu threshold method to remove the background and keep only the mask corresponding to the leaf. A closed contour is then extracted from the leaf mask. Note that the input of all the representations described above is a sequence of $N$ boundary points regardless of other leaf features like texture, color and venation.

### 3.1 The Swedish leaf dataset

The Swedish leaf dataset [35] contains 1125 images of leaves uniformly distributed in 15 species. Figure 6 shows sample leaves. The Swedish leaf dataset is very challenging because of its high inter-species similarity. One can notice the similarity of shapes of the first, third, and ninth species. To compare our approach with existing ones, we adopted the evaluation protocol used in [35, 25, 37, 4]. We randomly split each class into two sets: a training set containing 25 images and a testing set consisting of the remaining 50 images; we computed the classification rate given by the nearest neighbor (1-NN).

Table 1 shows the classification rate of the different representations compared to shape-based methods found in the literature. The number of boundary points is the same for all the proposed descriptors ($N = 400$). We are positioned first using the TSLA representation with 6 scales ($N_s = 6$) and a logarithmic periodicity between scales ($d(k) = 2^k$). TSLA outperforms all state-of-art methods while the TSL representation obtains the third best score with 95.73% using 7 scales ($N_s = 7$) and a logarithmic distance between triangles at different scales. The parameters that give the best results for the TOA and the TAR descriptors are respectively $N_s = 64$, $d(k) = 2 \times k$ and $N_s = 4$, $d(k) = 2^k$.

To compare the triangular representations, we use the same parameters that give the best classification rate (96.53%) with TSLA and we apply them on TSL, TOA and TAR. The performance of the descriptors for each class is shown in Figure 5. Examining the results per class on the Swedish leaf dataset, we notice that:

<table>
<thead>
<tr>
<th>Method</th>
<th>Classification rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSLA</td>
<td>96.53%</td>
</tr>
<tr>
<td>TSL</td>
<td>95.73%</td>
</tr>
<tr>
<td>TOA</td>
<td>95.20%</td>
</tr>
<tr>
<td>TAR</td>
<td>90.40%</td>
</tr>
<tr>
<td>Shape Tree [15]</td>
<td>96.28%</td>
</tr>
<tr>
<td>SPTC+DP</td>
<td>95.33%</td>
</tr>
<tr>
<td>MDS+SC+DP</td>
<td>95.33%</td>
</tr>
<tr>
<td>IDSC + DP</td>
<td>94.13%</td>
</tr>
<tr>
<td>IDSC + learned distance [4]</td>
<td>93.80%</td>
</tr>
<tr>
<td>sFACT (on contour) [37]</td>
<td>90.77%</td>
</tr>
<tr>
<td>Fourier [1]</td>
<td>89.60%</td>
</tr>
<tr>
<td>SC + DP</td>
<td>88.12%</td>
</tr>
<tr>
<td>Söderkvist [35]</td>
<td>82.40%</td>
</tr>
</tbody>
</table>

1 methods tested in [25]
3.2 The Flavia dataset

The Flavia dataset is composed of 1907 scans of leaves belonging to 32 species (see Figure 8). Several methods were tested in [24] on Flavia. To compare our approach with these methods, we used the same evaluation metrics as those in [24]: the Mean Average Precision (MAP) and the recall/precision curves. The precision $P$ and the recall $R$ values are given by:

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

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

The MAP value 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 $q$:

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

$P(k)$ 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 results are reported in Table 2. The TSL, TOA, and TSLA descriptors significantly outperform other methods. In this experiment, we used the same parameters for all the representations: 400 boundary points and 10 scales ($N = 400, N_s = 10$). Despite the high similarity of the shapes of different species, our approach shows a high capability to discriminate between species.

The recall/precision curves in Figure 9 show the performance of our descriptors on the Flavia dataset. We observe that TSL, TOA, TSLA perform in a similar way. The TAR representation gives the lowest performance. The recall/precision curves also prove that the angular information (TSLA, TOA) enhances the retrieval performance.

3.3 Comparison with ImageCLEF2011 results

Let us now introduce the context of the plant identification task of ImageCLEF 2011[16]. The ImageCLEF2011 dataset contains three categories of images:
- scans of leaves acquired using a flat-bed scanner
- scan-like leaf images acquired using a digital camera
- free natural photos

For each category, the images are divided into two sets: a training set and a test set. The goal of the task is to find the correct tree species of each test image. The identification score is quite different from the classic measures 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

<table>
<thead>
<tr>
<th>Methods in [24]</th>
<th>D2</th>
<th>MSDM</th>
<th>GEDT</th>
<th>REM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP</td>
<td>42.82</td>
<td>47.91</td>
<td>48.01</td>
<td>57.21</td>
</tr>
<tr>
<td>Ours</td>
<td>TAR</td>
<td>TSL</td>
<td>TOA</td>
<td>TSLA</td>
</tr>
<tr>
<td>MAP</td>
<td>50.81</td>
<td>65.94</td>
<td>68.37</td>
<td>69.93</td>
</tr>
</tbody>
</table>

Table 2: Mean Average Precision on the Flavia dataset

![Figure 9: Recall/Precision curves on the Flavia dataset](image-url)
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:

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

$U$: number of users (who have at least one image in the test data).
$P_u$: number of individual plants observed by the $u$th user.
$N_{u,p}$: number of pictures taken of 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 of the $p$th plant observed by the $u$th user.

We focus on scans and scan-like images. The first category contains 2349 images for training and 721 test images. For the scan-like category, 717 images are used for training and 180 images for testing. Table 4 shows the identification scores of our descriptors compared to other submitted runs of ImageCLEF2011. Our identification scores are higher than all the scores of the other methods on scans as well as scan-like images. If we average the score between the two categories, the TOA representation is slightly better.

Note also that the identification score using TOA is nearly the same on scans and scan-like images, although the noise that may exist in scan-like images (shadows, cluttered background, etc.) This demonstrates that a shape-based approach is suitable for a plant identification task.

### 3.5 Robustness to partial occlusion

In this experiment, we evaluate the performance of the proposed shape representation under partial occlusion. Partial occlusion may be due to uneven lighting conditions or overlapping objects. We picked five leaf images from the Flavia dataset and we applied three different types of occlusion: lobe occlusion, half leaf occlusion and multi occlusions. This eliminates from 20% up to 50% of the contour points. (cf. Figure 11). Taking into account that some boundary points are more sensitive to occlusion than others, we simultaneously applied multi occlusions of different parts of the leaf for the image $I_5$ in Figure 11.

In fact, the robustness of the descriptor depends on the choice of the number of triangles $N_a$. Here, the descriptor parameters are the same as those used previously on the Flavia dataset. Five retrieval tests were carried out using the occluded leaf images as queries. The results are presented in Table 5. We compare the obtained results to the scores of classification when original images are used as queries (100% for all the descriptors with $knn = 15$).

### 3.4 Comparison with ImageCLEF2012 dataset

The formula used to rank the runs in the ImageCLEF2012 plant identification task [17] is nearly the same as in 2011 (see [17] for details). The scan dataset contains 4870 images for training and 1760 test images. The scan-like category contains 1819 images for training and 907 images for testing. We obtain the second best results for all the triangle descriptors on the scan images and the three descriptors TSL, TSLA and TOA outperform the ImageCLEF2012 runs for the scan-like images. Note that the method that achieved the best score on scans (0.58) used a set of 27 features describing the shape, the texture and the color of the leaf [40]. In our case, we used only one shape descriptor.

<table>
<thead>
<tr>
<th>run_id</th>
<th>Scans</th>
<th>Scan-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFSC_USP-run2</td>
<td>0.562</td>
<td>0.402</td>
</tr>
<tr>
<td>inria_media_plantnet-run1</td>
<td>0.685</td>
<td>0.464</td>
</tr>
<tr>
<td>IFSC_USP-run1</td>
<td>0.411</td>
<td>0.430</td>
</tr>
<tr>
<td>LIRIS-run3</td>
<td>0.546</td>
<td>0.513</td>
</tr>
<tr>
<td>LIRIS-run1</td>
<td>0.539</td>
<td>0.543</td>
</tr>
<tr>
<td>Sabanci-okan-run1</td>
<td>0.682</td>
<td>0.476</td>
</tr>
<tr>
<td>LIRIS-run2</td>
<td>0.530</td>
<td>0.508</td>
</tr>
<tr>
<td>LIRIS-run4</td>
<td>0.537</td>
<td>0.538</td>
</tr>
<tr>
<td>inria_media_plantnet-run2</td>
<td>0.477</td>
<td>0.554</td>
</tr>
<tr>
<td>IFSC_USP-run3</td>
<td>0.356</td>
<td>0.187</td>
</tr>
<tr>
<td>DFF+GP [39]</td>
<td>0.778</td>
<td>0.725</td>
</tr>
<tr>
<td>TSL</td>
<td>0.802</td>
<td>0.757</td>
</tr>
<tr>
<td>TOA</td>
<td>0.794</td>
<td>0.780</td>
</tr>
<tr>
<td>TSLA</td>
<td>0.796</td>
<td>0.779</td>
</tr>
<tr>
<td>TAR</td>
<td>0.721</td>
<td>0.636</td>
</tr>
</tbody>
</table>

Table 3: Normalized classification scores of the scan and scan-like images on the ImageCLEF2011 dataset using the evaluation metric of [16]

<table>
<thead>
<tr>
<th>Top 3 Scores</th>
<th>Scans</th>
<th>Scan-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAR</td>
<td>0.52</td>
<td>0.61</td>
</tr>
<tr>
<td>TOA</td>
<td>0.54</td>
<td>0.63</td>
</tr>
<tr>
<td>TSLA</td>
<td>0.53</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Table 4: Normalized classification scores of the scan and scan-like images using the evaluation metric of [17] (ImageCLEF2012)

The best score on scans (0.58) used a set of 27 features describing the shape, the texture and the color of the leaf [40]. In our case, we used only one shape descriptor.

Table 5: Classification rates of occluded images.

<table>
<thead>
<tr>
<th>run_id</th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$I_3$</th>
<th>$I_4$</th>
<th>$I_5$</th>
<th>Average</th>
<th>Lost</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSL</td>
<td>100</td>
<td>93.3</td>
<td>100</td>
<td>93.3</td>
<td>94.4</td>
<td>94.4</td>
<td>5.6</td>
</tr>
<tr>
<td>TSLA</td>
<td>100</td>
<td>93.3</td>
<td>100</td>
<td>93.3</td>
<td>90.6</td>
<td>90.6</td>
<td>9.4</td>
</tr>
<tr>
<td>TOA</td>
<td>100</td>
<td>93.3</td>
<td>100</td>
<td>90.6</td>
<td>96</td>
<td>96</td>
<td>4</td>
</tr>
<tr>
<td>TAR</td>
<td>93.3</td>
<td>86.6</td>
<td>80</td>
<td>93.3</td>
<td>86.6</td>
<td>86.6</td>
<td>13.4</td>
</tr>
</tbody>
</table>

Figure 10: Partial occlusions on Flavia leaf images. First row: original images. Second row: occluded leaves.
The percentage of lost information is less than 10% for TSL, TSLA and TOA. The TOA descriptor obtained the best average score of classification (96%) over the five occluded images. The TSL description is the best in terms of robustness against multi occlusions applied on the image $I_5$. Retrieval queries using $I_5$ can be seen in Figure 10.

4. CONCLUSION

In this paper, we have presented a multiscale shape-based approach for leaf classification. We have compared four triangle representations based either on area (TAR) or side lengths and angles. We have introduced two triangle representations: TSLA and TOA. The experiments carried out on four leaf datasets show that using angles and side lengths is more appropriate than the area for triangle description. However, to compare our descriptors to the original TAR, a matching method based on dynamic programming should be developed. Moreover, the angular information provides a more precise description when it is jointly used with the triangle side lengths. Our approach is invariant to translation, rotation and scale. We have also shown that the different local descriptors are robust under partial occlusion. In our future work, we want to include other leaf features such as venation 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


A. Joly and O. Buisson. A posteriori multi-probe locality

N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs,
W. J. Kress, I. C. Lopez, and J. V. B. Soares. Leafsnap: A
computer vision system for automatic plant species
identification. In 12th European Conference on Computer

H. Laga, S. Kurtek, A. Srivastava, M. Golzarian, and
S. Miklavcic. A riemannian elastic metric for shape-based
plant leaf classification. In Digital Image Computing:

H. Ling and D. Jacobs. Shape classification using the
inner-distance. IEEE Transactions on Pattern Analysis

Y. Minggijiang, K. Kidiiyo, and R. Joseph. A survey of shape
feature extraction techniques. Pattern Recognition

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

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

S. Mouine, I. Yahiaoui, and A. Verroust-Blondet. Advanced
shape context for plant species identification using leaf
image retrieval. In 2nd ACM International Conference on

Plant species identification using elliptic Fourier leaf shape
analysis. Computers and Electronics in Agriculture,

L. Paullevé, H. Jégoù, and L. Amsaleg. Locality sensitive
hashing: A comparison of hash function types and
querying mechanisms. Pattern Recognition Letters,

E. Persson and K.-S. Fu. Shape discrimination using
fourier descriptors. IEEE Transactions on Pattern Analysis
and Machine Intelligence, 8:388–397, 1986.

D. Shen, H. H. Ip, and E. K. Teoh. Affine invariant
detection of perceptually parallel 3d planar curves. Pattern

D. Shen, W. Wong, and H. Ip. Affine-invariant image

O. J. Ööderkvist. Computer vision classification of leaves
from swedish trees. Master’s thesis, Linköping University,

Y. Tao and W. I. Grosky. Image indexing and retrieval
using object-based point feature maps. J. Vis. Lang.

J. Wu and J. Rehg. Where am i: Place instance and
category recognition using spatial pact. In Computer
Conference on, pages 1 –8, june 2008.

S. Wu, F. Bao, E. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L.
Xiang. A leaf recognition algorithm for plant classification
using probabilistic neural network. In Signal Processing
and Information Technology, 2007 IEEE International
Symposium on, pages 11 –16, dec. 2007.

I. Yahiaoui, O. Mzoughi, and N. Boujemaa. Leaf shape
descriptor for tree species identification. In IEEE
International Conference on Multimedia and Expo, ICME

C. Büyük, E. Atpoulou, and C. Tirkas. Sabanci-akan
system at images2012: Combining features and
classifiers for plant identification. In CLEF (Online
Working Notes/Labs/Workshop), 2012.

C. T. Zahn and R. Z. Roskies. Fourier descriptors for plane
1972.