Shape-based image retrieval in botanical collections

Itheri Yahiaoui[1], Nicolas Hervé[1] and Nozha Boujemaa[1],

[1] Projet IMEDIA, INRIA Rocquencourt, Domaine de Voluceau-Rocquencourt-B.P. 105.
78153 LE CHESNAY Cedex, France
{itheri.yahiaoui, nicolas.herve, nozha.boujemaa}@inria.fr

Abstract. Apart from the computer vision community, an always increasing number of scientific domains show a great interest for image analysis techniques. This interest is often guided by practical needs. As examples, we can cite all the medical imagery systems, the satellites images treatment and botanical databases. A common point of these applications is the large image collections that are generated and therefore require some automatic tools to help the scientists. These tools should allow clear structuration of the visual information and provide fast and accurate retrieval process. In the framework of the plant genes expression study we designed a content-based image retrieval (CBIR) system to assist botanists in their work. We propose a new contour-based shape descriptor that satisfies the constraints of this application (accuracy and real-time search). It is called Directional Fragment Histogram (DFH). This new descriptor has been evaluated and compared to several shape descriptors.

Keywords: Content-based image retrieval, region-of-interest based visual query, shape descriptor, botanical image processing.

1 Introduction

The content-based image retrieval (CBIR) systems have proven to be very useful in many fields to browse and search very huge image databases. Botanists are usually brought to use large collections of plants images. They need automatic tools to assist them in their work. Some very recent initiative in botany can be found in [1][12]. In a recent work Reference [1], we were interested in issues that are specific to the study of the function of genes in plants. These genetic modifications generate visual changes in the visual appearance of the plants. The biologists must find which phenotypes are visually similar; indeed, visual resemblance between phenotypes reflects similarities in the roles of the genes whose expression was blocked when obtaining these phenotypes.

At first time, we use a number of descriptors that were already implemented in our generic CBIR system IKONA [3]. Our existing shape descriptor based on the Hough transform captures the global structural information inside the plant. However, we need to describe more precisely the external shape of the plants to improve the plant characterization.

Different contour based descriptors were proposed in the literature. The well-known *shape contexts* and *curvature scale space* (CSS) descriptors record good performances but use complex and heavy methods to extract descriptors and to compute distances between them. In the case of a large database, with real-time performances issues, as in our application, this could be a bottleneck. As a consequence, we propose a new shape descriptor based on the external outline of a region that addresses these constraints. It is called Directional Fragment Histogram (DFH).

This paper focuses on a new extension of this shape descriptor and on its evaluation against state of the art descriptors. It is organized as follows. In the next section, we overview the related works on shape description. Then we expose in section 3 our proposal for a new contour-based shape descriptor. Section 4 presents the evaluation of this descriptor.

2 Related works

The shape is a very important aspect in the recognition of objects and therefore plays a major role in CBIR systems. It has been widely studied in the past years and many shape descriptors have been proposed. They differ in their ability to be invariant to translation, rotation and scale, and in their computational performances. One generally distinguishes two kinds of shape descriptors: region-based descriptors and contour-based descriptors. For the latter, one can cite the Fourier descriptor [11], the chain code histogram [5], the edge orientation histogram [6], the shape context descriptor [2] or the curvature scale space descriptor [8]. The chain code histogram (CCH) is a simple histogram of the 8 directions of the freeman code computed on the contour of a shape. The contour shape descriptor chosen in the MPEG-7 standard is an extension of the CSS descriptor [7]. It uses a representation of a contour in the curvature scale space and computes a feature vector of the corresponding region by using the maxima of curvature zero-crossing in this space. These maxima, called peaks, correspond to the main inflection points of the contour between its convex and concave portions. The contour is analyzed several times, at different scales after a smoothing process. This brings scale invariance to this descriptor and a good robustness to noise in the contour. In order to compute a similarity measure between two descriptors, a matching algorithm is used. It is designed to ensure a rotation invariance of the descriptor. Additionally, the feature vector includes the overall eccentricity and circularity of the contour.

3. Shape descriptor

We propose a new approach to describe the shape of a region, inspired by an idea related to the color descriptor in [9]. This new generic shape descriptor is computed using the outline of a region. In our case, the region of interest is the plant. We first need to separate each plant from the background of the image. Once this plant mask is extracted, we are able to compute the outline that represents the external shape of the plant.

3.1 Directional Fragment Histograms

Our new shape descriptor, called Directional Fragment Histogram (DFH), is computed using the outline of the region. We consider that this contour is composed of a succession of elementary components. An elementary component can either be a pixel or the elementary segment between two neighbour pixels. We associate to each elementary component of the contour a direction information which can take **N** different values d_0, d_1,...,d_{N-1}. A segment of size **s** is composed of exactly **s** elementary components.

For example, given a contour **C** composed of **m** elementary components:

C = ec_1 ec_2 ec_3 ec_4 ec_5 ec_6 ec_7 ... ec_i ... ec_m

The two first segments of size 4 are:

$\mathbf{Seg_4 1}$ = ec_1 ec_2 ec_3 ec_4 $\mathbf{Seg_4 2}$ = ec_2 ec_3 ec_4 ec_5

The total number of all possible segments is equal to **m** in the case of a closed contour, and is equal to **m-s+1** for open contours.

For each possible segment of the external contour, we identify groups of elementary components having the same direction (orientation) in this segment. Such groups are called directional fragments. The DFH codes two kinds of information. At a local level, it codes the relative length of these groups of elementary components within a given segment. This relative length is expressed as a percentage of the segment length. We choose to quantify the percentage axis into **J** percentiles. At a global level, the DFH codes the elementary components frequency distribution. The length of the segment defines the scale **s** of the DFH.

A DFH at scale **s** is a two-dimensional array of values that contains **N** x **J** bins. We illustrate in Figure 1. the extraction procedure of the fragment histogram supposing that we have 8 different directions and 4 different fraction ranges.

	0 25	25 50	50 75	75 100
d0	+1			
d1				
d2		+1		
d3				
d4				
d5				
d6				
d7			+1	

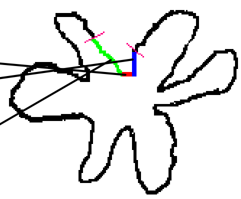

Figure 1. Extraction of the directional fragment histogram

The direction information associated to elementary components can be absolute with respect to a fixed reference. It can also be relative by deriving the absolute direction information. The use of relative direction information allows this new shape descriptor to be rotation invariant by focusing only on directions changes. As a counterpart, we are, of course, unable to distinguish the global directions. The DFH descriptor is in both cases invariant to translation but not to scale. To decrease the sensitivity of this descriptor to scale, we set a segment length to be proportional to the total length of the global outline of the considered shape. The fragment histogram DFH can also be normalized by the number of all the possible segments. To achieve real-time search, we use a simple L1 distance to compare the descriptors.

3.2 Freeman-DFH

A basic version of this new general framework for shape descriptor is to describe the contour of each region by a freeman code using 8 directions. It is a chain code representing a contour by a sequence of orientations of unit-sized line segments. In order to approximate rotation invariance, we compute a chain code derivative using modulo 8 differences code. This new chain code is another numbered sequence that represents relative directions of region boundary elements, measured as multiples of counter-clockwise 45° direction changes. DFH is then computed on this contour representation.

3.3 Gradient-DFH

A contour representation by a freeman chain code (absolute or relative versions) that uses only 8 directions might discretize too roughly the direction space. Thus, the contour coding is too sensitive to low variations of direction. We investigate a more elaborate version of DFH that consists in using a gradient **N**-directional chain code instead of the 8-connexity freeman chain code. In order to associate this local directional information to each extracted pixel belonging to the external contour, we compute the corresponding gradient angle on the original image. Then we project these angles to a quantified directional space. This quantification is based on a pre-defined parameter equal to **N**. The outline will then be represented by a string composed of directions resulting from this quantified space. Here again, the directional chain code can be derived to gain rotation invariance using modulo **N** differences.

4. Evaluation

The Freeman-DFH descriptor has been integrated in the IKONA system and tested on the genetically modified plants database. Our perception of these results is that they are good in general. This has been confirmed by the biologists during a demonstration session.

In order to automatically evaluate the performance of this new shape descriptor we need a larger database, with clear ground truth information (which is not the case with the genetically modified plants database). The shape should be the most pertinent feature to discriminate different classes. We use a Swedish trees leaves database collected by the Swedish Museum of Natural History for a project with Linköping University [10]. It contains 15 different Swedish trees species. Each species is represented by 75 images. All these images represent isolated leaves. We note that although they are from different species, some leaves look very similar to a non specialist user.

Figure 2. Representative leaves of the 15 different species

4.1 Leaves description

Usually the color descriptors are used in CBIR systems because they characterize well visual similarity. As known, the color of most leaves is green during the spring season. This green hue may vary due to climatic conditions and sun exposure. In addition this color changes to a variety of hot colors in autumn. This happens for almost all species of trees. When the leaves fall from the trees and decompose, they become dry and brown. This visual aspect modifications affect both the color and the texture of the leaves, as shown in the next figure (Figure 7).

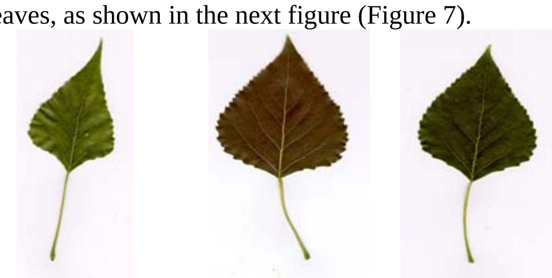

Figure 3. Different colors for the same species leaves

Thus, even is this features vary from one species to an other, we suppose they won't be discriminant enough. This shows that this database is particularly well adapted to be discriminated by shape descriptors.

4.2 External contour extraction and coding

As for the INRA database we need the outlines of the leaves to compute our shape descriptor. In this database we have a high contrast between the leaves and the white background. Thus the separation, and consequently the contour pixels, are easily obtained. We distinguish the external contour from all the possible contours (due to small gaps or overlapping in composite leaves) by keeping the longest one.

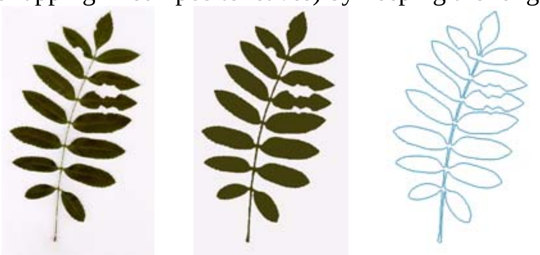

Figure 4. External contour extraction

4.3 Comparison of DFH versions

We compared the DFH descriptor computed respectively on the absolute and on the relative directional chains codes. Figure 5. represents a comparison of precision curves of both freeman and gradient DFH descriptors computed on absolute chain codes. We included also the combined version (relative + absolute) of gradient DFH.

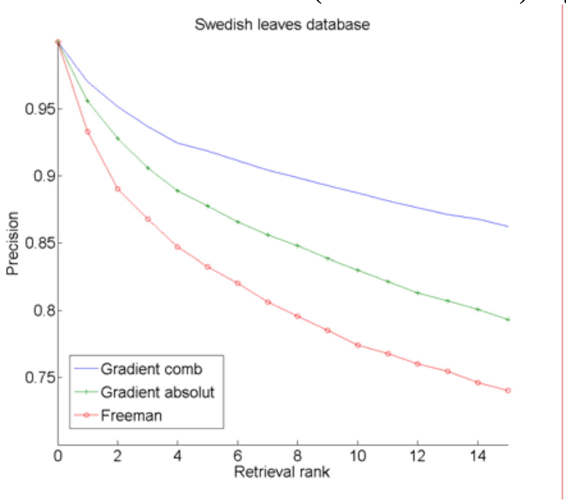

Figure 5. Comparison of DFH versions

The gradient DFH improves the performances of the freeman DFH. This is due to a finer quantization of the angle space that is made available with the use of a gradient directional space instead of the classical freeman chain code. For the gradient DFH version, we notice that the combination of the relative and absolute chain codes provides the best results. In this database, the rotation variance is low. This small variation is corrected by the introduction of the relative directional information in the computation of the DFH descriptor.

4.4 Comparison with other shape descriptors

In order to measure the contributions of Gradient-DFH descriptor, we evaluate it jointly with some existing shape descriptors on the Swedish database and compare the relative performances. These shape descriptors are:

- IKONA shape descriptor (Hough transform [4])
- Edge Orientation Histogram (EOH) [6] that we adapted to leaves outlines
- MPEG-7 CSS descriptor
- CCH descriptor

For our descriptor, these results are obtained using 40 quantified directions, 6 quantified percentiles. The segment lengths are fixed to 1.1% of the contour size. These parameters where chosen empirically. We are then using 240 bins histograms. The EOH descriptor was computed on the original gradient space without quantification. For CSS, we use the original MPEG-7 implementation.

These results demonstrate both the good performance and the contribution of the Gradient-DFH shape descriptor to the existing descriptors. The results are slightly better than those obtained with CSS. The Hough transform records average performances because it captures structural information of leaf images. Both CCH and EOH give lower results because they focus only on the global shape without taking into account local distribution of directions.

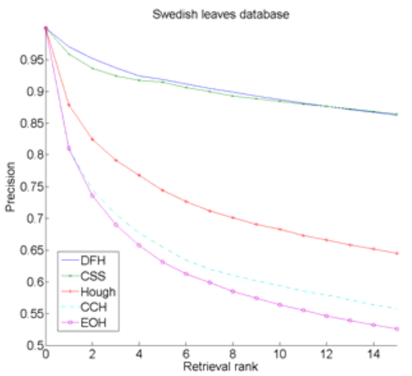

Figure 6. Precision curves

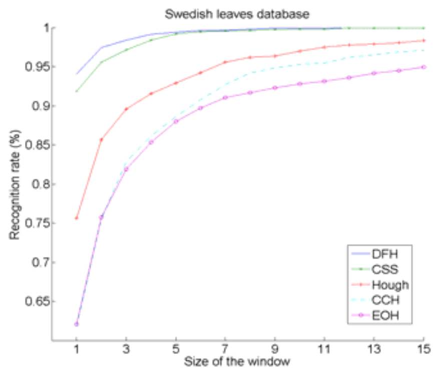

Figure 7. ROC curves

One of the main reasons for introducing this new shape descriptor was the computation performances. We compare the time needed to compute all the possible similarity measures between the 1125 images of the database (about 1.2 million measures computed). On a P4-2.8GHz computer, it takes 0.5 second for the DFH descriptor and 123 seconds for the CSS descriptor. The gain is obvious. In other hand, the CSS descriptor is coded on only 116 bits (average on the entire database), but this is not a critical point for our application.

5. Conclusion

In order to build an automatic tool that will help botanists to study the impact of genetic modification on the visual appearance of plants, we propose a new generic shape descriptor. We evaluate the performance of this descriptor using a ground truth leaves database and compare it to known shape descriptors. The results show the effectiveness of this approach for similarity based retrieval.

6. References

[1] P. Belhumeur and al., An Electronic Field Guide: Plant Exploration and Discovery in the 21st Century. http://www.cfar.umd.edu/~gaaga/leaf/leaf.html

[2] S. Belongie, J. Malik and J. Puzicha. *Shape matching and object recognition using shape contexts*. IEEE transactions. Pattern analysis and machine intelligence. 2002. vol 24 no 24.

[3] N. Boujemaa, J. Fauqueur, M. Ferecatu, F. Fleuret, V. Gouet, B. LeSaux and H. Sahbi, *IKONA: interactive specific and generic image retrieval*, International workshop on Multimedia Content-Based Indexing and Retrieval (MMCBIR'2001), Rocquencourt, France, 2001

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

[5] J. Iivarinen and A. Visa. *Shape recognition of irregular objects*. 1996. Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling, Proc. SPIE.

[6] A. K. Jain and A. Vailaya, *Image retrieval using color and shape*, Pattern Recognition, 29(8):1233--1244, August 1996.

[7] B.S. Manjunath, Philippe Salembier and Thomas Sikora, *Introduction to MPEG-7*, Wiley, 2002, ISBN 0-471-48678-7

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

[9] R.J. Qian, P.L.J. van Beek and M.I. Sezan, *Image retrieval using blob histograms*, in IEEE Proc, Int'l. Conf. On Multimedia and Expo, New York City,July 2000.

[10] O. Söderkvist, *Computer vision classification of leaves from swedish trees*, Master thesis, 2001

[11] C.T. Zahn and R.Z. Roskies. *Fourier descriptors for plane closed curves*. IEEE Transactions on computers, march 1972

[12] J. Zou and G. Nagy, Evaluation of Model-Based Interactive Flower Recognition. (ICPR'04), Cambridge, United Kingdom, 2004.

[13] I. Yahiaoui and N. Boujemaa, Content-based image retrieval in botanical collections for gene expression studies, (ICIP'05), Genova, Italy, September 2005