MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

3—30
A Hierarchical Method of Recognizing Plant Species by Leaf Shapes

Cholhong Im *
Department of Information Science, Faculty of Science, The University of Tokyo

Hirobumi Nishida †
Software Research Center, Ricoh Co., Ltd.

Tosiyasu L. Kunii ‡
Professor, Computational Science Research Center, Hosei University
Senior Partner, MONOLITH CO.,Ltd.,
Advisor in Science and Technology, Fukushima Prefecture Government

Abstract

In order to recognize plant species from their shapes, shapes of contours of leaves can be used. However, contours of leaves are complicated and have variations. In order to deal with the complexities and variations of contours of leaves, a hierarchical representation of leaves is presented in this paper. First, global structures and shapes of leaves are represented by polygons whose vertices are critical points of curvature of smoothed contours of leaves. Second, detailed shapes of mid-most pieces are represented by linear approximation of the contours of pieces. Last, teeth on the contour of leaves are represented by triangles. Such triangles are detected from contours that are not smoothed enough to remove teeth. Features at each hierarchy are represented statistically. Such a hierarchical representation is more powerful than previous methods that represent only global structures of leaves.

1 Introduction

In computer vision, a lot of methods for recognizing objects have been presented and have been applied to several kinds of objects. However there are few methods related to recognizing plant species. This is mainly due to the complexities and variations of their shapes even in the same species.

We can generally classify plants according to their shapes and structures of flowers or genitals[6]. However they have very complex structures and need anatomical analysis for the classification. On the other hand, shapes of leaves are planar and we can use contours of leaves as features to discriminate species because of their varieties of shapes.

Leaves consist of a number of pieces and there are teeth along their contours. In addition to such complexity of contours, they have variations even in the same species.

In order to deal with the variations and complexities, we adopt a hierarchical and statistical method of representation and recognition to reflect structural properties and detailed shapes of leaves. We can consider triangular pieces that protrude all around as the basic structure of leaves. Besides, there are varieties of shapes of pieces to discriminate plant species. We can use the properties of a set of teeth along the contour of piece as features to discriminate species. We use such a hierarchical structure of leaves to recognize plant species. At each hierarchy, we calculate statistics of features. We define a similarity measure as probability based on the statistics. The recognition of plant species is done based on the similarity measure.

In section 2, we address previous works and their problems. We describes the representation method of structures and detailed shapes of leaves in the first part of section 3. In section 3.4 we define a similarity measure. In section 3.5, a detection method of teeth is presented. In section 4, an overall recognition method is presented.

2 Previous Works

Previous studies [5, 1] of recognizing plant species use contours of leaves as a feature to discriminate species also. However they emphasized the representation of structures of complex contours and they use only global structures of leaves for matching descriptions of any two plants. Abbasi [1] uses positions of peaks of curves in the upper part of scale space images to compare any two contours of leaves. Though this can represent global structures, global shapes and more detailed shapes cannot be represented enough. Tsukioka [5] uses critical points of curvature of contours. Contours are represented by the distance between critical points and value of curvature at a characteristic scale. However because of local variations of values of curvature, matching based on this representation does not work well.

Analysis of teeth on the contour is done using wavelet local extrema [2]. They analyze teeth by

*7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: chim@is.s.u-tokyo.ac.jp
†1-1-17 Koishikawa, Bunkyo-ku, Tokyo 112-0002, Japan. E-mail: hn@src.ricoh.co.jp
‡1-25-21-602 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: tosi@kunii.com

Figure 1: Polygonal Approximation

Figure 2: Polygonal Approximation of A Mid-Most Piece of A Leaf

classifying vectors of critical values that are obtained by following critical points of curvature filtered by wavelet transformation and detecting a regular sequence of critical points. However shapes of teeth of leaves are not so regular and this method does not work well in such cases.

3 A Representation of Leaves

In this paper, a hierarchical representation and recognition method of plant species from contours of leaves is presented. Such a method is needed because different methods of detections are needed at each hierarchy. There are three hierarchies. The first hierarchy is a representation of global structure and shapes of leaves. The second is a representation of local detailed shape. The third is a representation of teeth. The first and second hierarchies are described in detail in a paper[3].

3.1 A Representation of Global Structures and Shapes

Because approximate shapes of pieces can be considered as triangles, a piece can be represented by an apex and two pits beside it. Apexes of leaves correspond to the critical points of curvature of the contour with positive value and pits correspond to the critical points with negative value. Therefore the structure is represented by a polygon whose vertices are critical points (a sequence of alternate pits and apexes as in Fig1) of curvature of the smoothed contour [4].

3.2 A Representation of Local Detailed Shapes

We represent only a mid-most piece of leaves in detail because the shape is more stable than the others. We use a linear approximation composed of four line segments that minimize the difference of arc length and the sum of length of line segments(Fig.2) to approximate curves of pieces.

3.3 A Representation of Features of a Contour

We represent the features obtained by the methods of previous sections by the tuple of arc length between two end points of the line segment and angles between two consecutive line segments.

3.4 A Similarity Measure

We define the similarity measure based on statistical properties of shapes. We calculate statistics for each features from the training set for each species. Using the statistics, we define the similarity measure for species as probability that the observed sample belongs to the species, assuming that the distribution of each feature is normal and independent of each other.

3.5 Detection of Teeth

Because we can consider teeth of leaves are triangular, a tooth can be represented by an apex and two pits beside it. In order to detect teeth on contours of leaves, we smooth contours slightly. However, slightly smoothed contours have many critical points caused by discreteness of contours and noises. Therefore we should select the points that constitute teeth from the critical points on the contours.

3.5.1 Conditions for a Tooth

We assume following conditions of a tooth to be satisfied(Fig. 3).

1. The line connecting two pits is included inside the contour.
2. The height and width of triangle exceed some threshold.
3. The ratio of height to the length of the base exceeds some threshold.
4. The angles formed at pits are lower than some threshold.

The condition 1 can be checked by drawing a line between two pits on the binerized image of a leaf. The second condition is necessary because there are a lot of triangles generated by noises and discreteness of contours. The condition 3 avoids detecting false teeth. The condition 4 is used to check whether the critical points with negative values are pits.

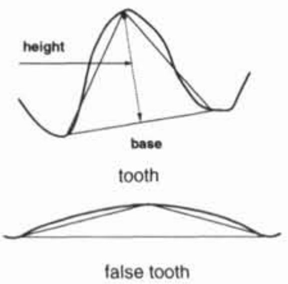

Figure 3: A Tooth of a Leaf

Figure 4: A Tip of a Tooth

3.5.2 A Procedure for Detecting Teeth

Teeth are detected in an order of the size which is represented by arc length.

For each size, triangles are detected along the contour of a leaf according to the conditions described in the previous section. However at the stage of detection of small teeth, it is possible to detect tips of teeth(Fig. 4) because of noises or discreteness of contours. In order to avoid such problems, a larger triangle detected at the following stages replaces the smaller one. The larger triangle should include only one smaller triangle. Results of detection are presented(Fig. 5, 6). In Fig6, tips of pieces are detected as teeth. Such tips can be discriminated by the angle formed by the base line of a tip and those of adjacent teeth.

3.5.3 A Representation of Teeth

Features of teeth on the contours of leaves are represented by the number of teeth, size of teeth and approximate shapes of teeth. Size of teeth is represented by the length of base and height of the triangle that approximates a tooth. Such quantities are normalized by dividing them by the perimeter of the contour of the leaf. An approximate shape of a tooth is represented by the ratio of height to the length of the base and angle formed by the base and the line that connects apex of teeth and the center of the base. All the features except the number of teeth are represented statistically. We represent species statistically by considering statistics of teeth of each leaf as features. A similarity measures is defined in the same manner as the similarity measure of structures and shapes of leaves.

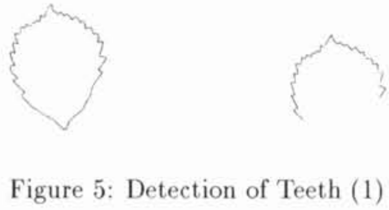

Figure 5: Detection of Teeth (1)

Figure 6: Detection of Teeth (2)

4 Recognition Method

Recognition of species is done hierarchically also. Before recognition, we calculate statistics of each feature. In the first stage of recognition, teeth of the leaf are detected. Next, the global structure of the leaf and the parameter for similarity measure are calculated. The leaf is classified according to the number of apexes and then based on the values of the similarity measure, the leaf is classified further. If the species of the leaf is not determined in this stage, the approximation of the mid-most piece is calculated and the similarity measure is calculated. Based on the values of the similarity measure, the leaf is classified. If the species of the leaf is not determined in this stage, features of teeth that are detected in the first stage are used to classify plant species.

5 Experimental Result

Classification rates of each leaf are summarized in table 1, 2,4. 1st is the stage of recognizing global structures. 2nd is the stage of recognizing mid-most pieces. 3rd is the stage of recognizing teeth. Some leaves of the 1st row in table 1 are recognized at the 3rd stage. Because the leaves of the species have large teeth and the other species have small teeth or do not have teeth, they can be recognized at this stage. Some leaves of the last row in table 2 are recognized at the 3rd stage. They have similar values of similarity measure with the species of the 3rd row of the table 2. Because the leaves of the last row have no teeth, they can be recognized at the stage. When the leaf have an apex, it needs the 3rd stage only. Leaves of the 1st row of table 4 have no teeth. Leaves of the 2nd row in table 4 have smaller teeth than the leaves of the 3rd row. Average of parameters of teeth are summarized in table 3.

	1st	2nd	3rd	sum
	0.083	0.583	0.083	0.749
	0.417	0.417	0.000	0.834
	1.0	0.0	0.000	1.0
	0.909	0.000	0.000	0.909
	0.429	0.571	0.000	1.000
	0.847	0.153	0.000	1.000

Table 1: Classification Rate for More than 5 Apexes

6 Discussion

We solve the problem of previous methods that two leaves are compared only by their structural properties. Because there are a lot of plant species that are structurally similar, it is not effective to compare leaves only by their structure. The angle formed by connecting apexes and pits are more stable than the value of curvature at the apexes and pits and the sequence of them can represent global structures and shapes of leaves. Detailed shapes can be stably represented by linear approximation compared with the critical points.

The detection method of teeth presented in this paper can detect approximate shapes of teeth. Our method assumes approximate shapes of teeth as triangles and detects such triangles from contours. Therefore teeth are detected even if there are not exact regular sequences of teeth.

The hierarchical method we presented can deal with the detailed shapes of contours that cannot be appropriately dealt with by the previous methods.

	1st	2nd	3rd	sum
	0.846	0.077	0.000	0.923
	0.917	0.000	0.000	0.917
	0.625	0.125	0.000	0.750
	0.625	0.125	0.125	0.875

Table 2: Classification Rate for 3 Apexes

	height	base	ratio	angle	number
(b)	0.0033	0.0127	0.0002	0.6947	25
(c)	0.0124	0.0203	0.0003	1.0573	19

Table 3: Averages of Parameters of Teeth

	Teeth	3rd
(a) 		1.000
(b) 		0.917
(c) 		0.917

Table 4: Classification Rate for 1 Apex

Acknowledgments

All the images of leaves used in this paper are offered by Herbarium of the University of Tokyo.

References

[1] S. Abbasi, F. Mokhtarian, and J. Kittler. Reliable classification of chrysanthemum leaves through curvature scale space. In *Proc. International Conference on Scale-Space Theory in Computer Vision*, pages 284–295, Utrecht, The Netherlands, 1997.

[2] E. Hussein, Y. Nakamura, and Y. Ohta. Analysis of detailed patterns of contour shapes using wavelet local extrema. In *Proc. of ICPR*, pages 335–339, 1996.

[3] C. Im, H. Nishida, and T.L. Kunii. Recognizing plant species by leaf shapes - a case study of the acer family. In *Proc. of 14th ICPR*, pages 1171–1173, Brisbane, Australia, 1998.

[4] F. Mokhtarian and A. Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Trans. on PAMI*, PAMI-8(1):34–43, 1986.

[5] T. Tsukioka. Coding boundary shapes of leaves based on multiresolution. In *Master thesis University of Tokyo, Dept. of Information Science*, 1994.

[6] T. E. Weier, C. R. Stocking, M. G. Barbour, and T. L. Rost. *Botany: an introduction to plant biology 6th ed.* John Wiley & Sons, 1982.