ELSEVIER
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Technology 6 (2012) 171 – 179
Procedia Technology

2nd International Conference on Communication, Computing & Security [ICCCS-2012]

Plant Identification Based on Fractal Refinement Technique (FRT).

Antony Jobin[a,*], Madhu S. Nair[b], Rao Tatavarti[c]

[a,]Department of Computer Science, Rajagiri College of Social Sciences, Kalamassery, Kochi 683104, Kerala, India*
e-mail: antony.jobinkj@gmail.com
[b]Department of Computer Science, University of Kerala, Kariavattom, Thiruvananthapuram 695581, Kerala, India
e-mail: madhu_s_nair2001@yahoo.com
[c] Department of Civil Engineering, Gayatri Vidya Parishad College of Engineering, Madhurawada, Visakhapatnam 530048, Andhra Pradesh, India
e-mail: rtatavarti@gmail.com

Abstract

We propose here a new algorithm for plant classification and identification based on fractal dimension. It is a simple and efficient technique for identifying plants using three levels of fractal refinement on leaf images. Contour, Contour-Nervure and Nervure fractal dimensions are computed and are used in the first, second and third level of refinement respectively. A 50 set species with each set containing 10 samples are used for training the algorithm. The performance of the algorithm was examined with a test set of 500 leaves arbitrarily selected from different groups of species. The fault acceptance rate (FAR), the fault rejection rate (FRR) and the classification accuracy of the algorithm were analyzed experimentally and demonstrated that the proposed method has an accuracy rate of 84%.

Keywords: Fractal Refinement Technique;FRT; Fractal Dimension; Identification of Plants

* Antony Jobin. Tel.: +919995431287.
E-mail address: antony.jobinkj@gmail.com

doi:10.1016/j.protcy.2012.10.021

1. Introduction

Conventional identification and classification of plants evolved over hundreds of years and involves manual comparison of certain characteristics and then assigning a particular plant to a known taxonomic group and ultimately arriving at a species, the result of which depends on the method adopted out of the several methods available and is also time consuming. Traditionally, flowers and stems were the main parts in plants used for the identification purpose [1, 2]. Shape and arrangement of corolla, calyx, receptacle, pistil, stamen, ovary, ovule, etc., were the main parts of flowers used for this purpose. All these manual tasks are time consuming and require the help of experts, but could give very accurate results. If this process could be automated, would make identification task easier. Increasing the accuracy rate could also make immense changes in biological science in terms of both time and money. Against this back drop researchers seriously began addressing the problem of taxonomic data processing in the 1970s using computer based information networks [3], with the belief that a significant improvement can be expected if the plant identification can be carried out by a computer, automatically or semi-automatically, assisted by image processing and computer vision techniques. By using a computer-aided plant identification system, non-professionals can also take part in this process. Presently, there are several automatic-identification theories and procedures are available like “Shape based leaf image retrieval” [4], CAPSI [5], etc. Among them the comparative studies shown that the fractal based classification got an improved accuracy rate from 60% to 74% as that of the “Shape based leaf image retrieval” which specifically depends on the centroid-contour distance (CCD), eccentricity, angle code histogram(ACH) which made [4] fractal dimension based identification of plants quite popular and accurate. The fractal dimension based plant identification theory is based on the fact that leaves of different plants have different shape, the shapes of different plants’ leaves varies, leaves of the same species have a certain similarity, the difference of leaves of different species is obvious[6, 7]. Hence the contour shape and nervure part of leaves would be different for different leaves, which mean that their combination shows a uniqueness or distinguishable nature among different species. Since they are irregular in shape, fractal dimension could be used for the purpose. Hence, measuring the fractal dimension of leaves can quantitatively describe and classify even morphologically complex plants [8-10].Fractals are an important branch of mathematics. Most physical systems of nature and many artifacts are not regular geometric shapes of the standard geometry derived from Euclid. Fractal geometry offers almost unlimited ways of describing, measuring and predicting this natural phenomenon. Two of the most important properties of fractals are self-similarity and non-integer dimension [11]. Fractal geometry has permeated many areas of science, such as astrophysics, biological sciences, computer graphics, etc [12]. Fractal dimensions can be used in a more general sense referring to any of the dimensions commonly used to characterize fractals [13]. Methods of calculating fractal dimension are categorized into 4. Among them we use pixel covering method [14] which belongs to the first group, method based on simple pixel borders. In the pixel covering method, initially the image of leaf is converted into multi-valued gray-scale image which is then transformed into bi-valued gray scale image. A sub-image is then extracted from it which contains completely the feature part of the leaves. This, sub-image is then split into boxes of size $k=2, 4, 6 \ldots m$ pixels. Then the number of boxes $N(2), N(4), N(6) \ldots, N(m)$ that contain at least one pixel with feature part in it, is counted. If, ‘$€^*$’ denotes the size of a pixel then,

$€ = N(k)\, €^*$, represents the side length of a single box. Then the fractal dimension could be found as,

$N(k)=1/(€)^D$, taking logarithm on both sides of the equation (1) we get,
(1)

$$log(N(k) = -D\, log(€) \quad (2)$$

$$i.e.,\ D = -log(N(k)) / log(€) \quad (3)$$

where *'D'* is the fractal dimension. The fractal theory could hence be applied for the purpose of classification. There are many theories available about the current context. Among them the important one is the comparative study of the Box-counting method and the Bouligand–Minkowski method of finding the fractal dimension for the purpose of plant classification [15]. The paper has shown that it could classify the plant species with a precision rate of 74% (that is with a rate of 26% of miss-classification). It has also proved that the traditional method of computer aided classification based on centroid-contour distance has only an accuracy rate of 60% [4], which signifies importance of the fractal characteristic of leaves for the purpose of plant classification. But the process had only used the contour and the venation of the leaves for the purpose. The focus on the nervure characteristic of leaves, which could be considered as a unique identity of each species leaves, are not taken into consideration for the classification process.In this paper we propose a Fractal Refinement algorithm for plant classification and identification based on fractal dimension, with the similar number of datasets as that of [4], with the focus given to the nervure part of leaf which could classify the plant species accurately and thereby to increase the efficiency and correctness of identification of plants. Three levels of extraction with two levels in nervure part are considered in newly suggested classification process, which had ultimately brought up the accuracy rate from 74% to 84%. The Classification process could be divided into two phases: 1) Specimen collection and recording phase - to collect and record the representative fractal value of a new specimen in the file and, 2) Refining and Identification phase – to identify the sample by refining the collection of recorded species with that of the property of the sample to get a high probabilistic matching species from the recorded species. The fractal refinement technique consists of three levels of refinement - refinement based on contour structure of leaves, refinement based on contour-nervure structure of leaves and lastly the refinement based on nervure part of leaves.

2. Fractal Refinement Technique for Identification of Leaves:

The technique is completely based on the fractal dimension based refinement of recorded species to obtain a good probabilistic matching species for the input sample. The process consists of two phases:

2.1 Specimen Collection and Recording Phase

This phase is again divided into two parts. In the first part, *i.e*., specimen collection, *'n'* samples of each species were collected (for a disrepair-mature leaf, and for good clarity camera; $3 \leq n \leq 5$, else $n > 5$, for computing correct fractal dimension for a species). Mature leaves without any disrepair are selected to get the representative fractal dimension. Samples collected are then placed in front of a white light source, so that digital camera with high clarity could capture nervure details of each and every corner of the leaves. Captured leaf images are then transferred to the storage space in the computer. Brightness and contrast of the images are appropriately adjusted and the background noise, if present, is removed from the images. Resizing of image into 800 pixels wide and 600 pixels high would neither lose details of the image nor impact the speed of image processing due to its large size.

In the second part of this phase *i.e.*, recording of fractal values of the collected species, the *'n'* processed images of the *'n'* samples of the collected species is then converted into *'n'* gray scale images. The resulting multi-valued gray scale images are then converted into bi-valued gray scale images; each pixel representing white or black. Next, obtain the contour part of thus formed images i.e., the outer edge of the images and calculate the fractal dimensions $Fc_1, Fc_2, \dots Fc_n$ of the contour part of these *'n'* samples of the collected species. Finding the mean and variance of these *'n'* contour-fractal values would define the range in which the real fractal value of the contour part of the collected species would be lying. The mean and the standard deviation of the computed fractal dimensions is computed as follows,

$$\mu = (1/n)\ (Fc_1+Fc_2+\dots+Fc_n) \quad (4)$$

$$\sigma = ((1/n)\ [\ (Fc_1-\mu)^2 + (Fc_2-\mu)^2 + \dots + (Fc_n-\mu)^2)^{1/2} \quad (5)$$

Hence, the contour-fractal dimension of the collected species would be in the range $(\mu-\sigma,\ \mu+\sigma)$. Any sample with contour-fractal value coming in this range has got a chance that it belongs to this species. Record the computed μ and σ value of each of the species in a file f_1.

Next step is to extract the contour-nervure part from the bi-valued gray-scale images of these 'n' samples. After acquiring the contour-nervure part, the fractal dimensions $Fcn_1, Fcn_2..Fcn_n$ are calculated for each of the species. On getting *'n'* contour-nervure fractal values of *'n'* samples, find the mean and standard deviation of these *'n'* fractal values to get a representative fractal dimension range $(\mu_{cn}-\sigma_{cn},\ \mu_{cn}+\sigma_{cn})$. Record the mean μ_{cn} and standard deviation σ_{cn} of each of the species in a file f_2. Any sample whose contour-fractal dimension comes in the range $(\mu-\sigma,\ \mu+\sigma)$ and contour-nervure fractal dimension comes in the range $(\mu_{cn}-\sigma_{cn},\ \mu_{cn}+\sigma_{cn})$ has got a greater chance of being in that group of species.

Next step is to extract the nervure part of species from the *'n'* bi-valued gray-scale images. On getting the *'n'* nervure part of the *'n'* samples of the same species, we could then compute their fractal dimension. Getting the *'n'* fractal dimension of the *'n'* nervure samples $Fn_1, Fn_2 \dots Fn_n$, we then compute the mean and standard deviation to get their representative fractal dimension range $(\mu_n-\sigma_n,\ \mu_n+\sigma_n)$. Record the computed mean μ_n and standard deviation σ_n of each of the species in a file f_3. Hence any sample whose contour value comes in the range $(\mu-\sigma,\ \mu+\sigma)$, contour-nervure fractal dimension comes in the range $(\mu_{cn}-\sigma_{cn},\ \mu_{cn}+\sigma_{cn})$ and finally the nervure fractal dimension comes in the range $(\mu_n-\sigma_n,\ \mu_n+\sigma_n)$ would have a high probability of being in that group of species. For each new species which could be included in the identification process should have an entry in all the three of the above discussed files.

2.2 Refining and Identification Phase

In this phase, we try to identify the species to which an unknown sample belongs. A disrepair-mature leaf of the unknown species is taken and kept in front of a white light source for capturing the image with a high resolution camera. The image of this unknown sample is pre-processed by passing through all of the stages discussed in previous phase like increasing contrast, brightness, eliminating the background noise included in the image and resizing of the image into 800*600 resolution (to speed up the processing without any loss of its structure). After pre-processing, the sample image is then converted into multi valued gray-scale image. This gray-scale image is then converted into bi-valued gray-scale image. From the bi-valued gray-scale image, extract the contour part and find its fractal dimension *Fc,* then extract the contour-nervure part and find its fractal dimension *Fcn* and lastly obtain the nervure part and find its fractal dimension *Fn*. These fractal dimension values are then used for three-level refinement.

2.2.1 First Level Fractal Refinement: Check contour fractal dimension of the sample, *'Fc'* with each entry in the file *f1* that stores the contour fractal dimensions of the recorded species. Check whether there exists any entry/entries whose contour fractal dimension range includes the sample's contour dimension. If exist, they are the first level possibility species.i.e., First-level Possibility Vector = $\{ i \mid \mu_i - \sigma_i <= Fc <= \mu_i + \sigma_i, \ \forall\ i \in i^{th}$ *Records in file1* $\}$

Here, we are refining the recorded species in the file *f1* containing contour dimension, based on the matching of dimensional value of the testing sample. And only this first level possibility vector is moved onto the second level refinement.

2.2.2 Second Level Fractal Refinement: In the second level refinement, only those species which are in first level possibility vector would only be refined and goes to the third level of refinement. In the second level of refinement, the contour-nervure fractal dimension is being used for the purpose of refinement. In this level, the contour-nervure dimension of the testing sample *'Fcn'* would be matched with the dimensional values in file *f2*, corresponding to those species which are in First-level Possibility Vector. The condition for second level possibility vector is given below:

Second-level Possibility Vector = $\{ i \mid \mu_{cni} - \sigma_{cni} <= Fcn <= \mu_{cni} + \sigma_{cni}, \ \forall\ i \in$ *first level possibility vector* $\}$ (6)

Those species, whose contour-nervure dimensional range contains the contour-nervure dimensional value of testing sample, are taken into the Second-Level-Possibility Vector.

2.2.3 Third Level Fractal Refinement: In the third level, refinement is done to the second level possibility vector. In this level, the fractal dimension of the nervure part *'Fn'* of the testing sample is compared with the records of nervure dimensions of recorded species contained in the file *f3*. There we find out the matching species whose nervure dimensional range contains the nervure dimension of the testing sample. This will lead to the formation of Third-Level Possibility Vector.

Third-level Possibility Vector = $\{ i \mid \mu_{ni} - \sigma_{ni} <= Fn <= \mu_{ni} + \sigma_{ni}, \ \forall\ i \in$ *second level possibility vector* $\}$ (7)

Here the species in the Second level possibility vector would only be considered and from them the accurate species would be chosen. The Fractal Refinement technique discussed above has been diagrammatically represented in Fig 1.

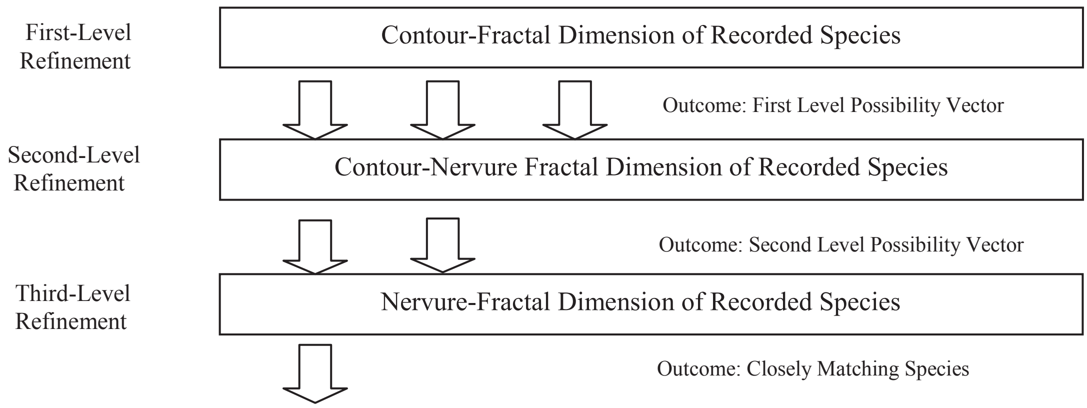


Fig.1: Fractal Refinement Technique (FRT).

3. Experimental Results:

This section presents the results of the experiments conducted to demonstrate the performance of the proposed algorithm. The method has been implemented in the MATLAB on an Intel Core2 Duo 2.93 GHz with 1024 X 768 resolutions. In the experiment, 50 known species were considered; each with 10 samples. The species selected for training the algorithm are shown in Fig 2. The fractal dimension of contour, contour-nervure and nervure thus found were then saved in files separately. The Fractal values computed for arbitrarily selected four species are shown in Table 1.

To study the Fault-Acceptance Rate (FAR), Fault-Rejection Rate (FRR) and the degree of accuracy of the proposed FRT algorithm, 500 new testing samples with wide variety of species, were taken randomly from the botanical gardens of Rajagiri College of Social Sciences, Kochi and St.Albert's College, Kochi.

Fig.2: Species of leaves that were considered for the study.

This sample set was then divided into five groups, namely Test group1, Test group2, Test group3, Test group4 and Test group5. Each group containing 100 test samples and each of the samples are pre-processed (as in Fig 3), converted into multi-valued grey-scale image (as in Fig 4). From them contour (as in fig 5), contour-nervure (as in Fig 6) and nervure sub images are extracted (as in Fig 7) for finding their respective fractal dimensional values.

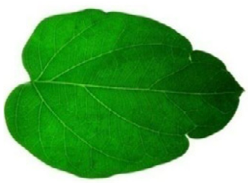


Fig.3:Pre-processed image of leaf

Fig.4:Multi-valued Gray-Scale Image of leaf

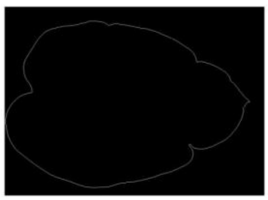


Fig.5: Contour Image of leaf

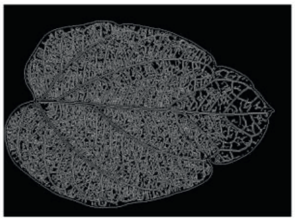
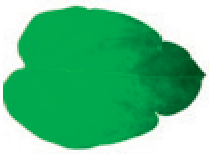
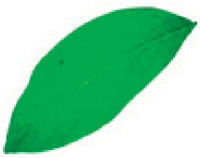


Fig.6: Contour-Nervure Image of leaf

Fig.7: Nervure Image of leaf

On finding their contour, contour-nervure and nervure fractal dimension, the second phase of the proposed algorithm was applied to the dimensions of the sample. After three levels of fractal refinement the result thus obtained is shown in Table 2. From the observations it is clear that the Fault Acceptance Rate of the algorithm
is 4.2% , Fault Rejection Rate of the algorithm is 11.8% and rate of miss-classification is 16% and hence the accuracy of the proposed FRT algorithm is 84%.

Table 1: Fractal value of 4 collected species, to be recorded in the files.

Serial no.	Species	Contour Fractal Dimension	Contour-Nervure Fractal Dimension	Nervure Fractal Dimension
1		1.0682±0.004657	1.5831±0.01484	1.6243±0.015725
2		1.0853±0.01911	1.5595±0.030144	1.5017±0.077398
3		1.0503±0.010755	1.6593±0.023307	1.6643±0.016014
4		1.0661±0.004561	1.6558±0.015846	1.6733± 0.01064

4. Conclusions

The proposed algorithm is based on three levels of refinement, *i.e.* refinement based on contour, contour-nervure and nervure fractal dimensions. The effectiveness of algorithm lies in the fact that the resultant identified species is found after three levels of refinement. In these three levels the test sample values are cross checked with that of dimensional values stored in the files. The result that is finally obtained should satisfy the conditions at all the three levels with the fractal dimension values of the test sample and this will reduce the chance of getting an erroneous result. The drawback of the proposed method is that it completely depends on the clarity of the image, *i.e.* better the clarity and the quality of the image the better would be the result obtained. In order to overcome this drawback, novel techniques independent of the clarity of the image could be added to this algorithm and could be implemented as future work. However, with the advances in electronic technologies high definition cameras have become common and cheap and therefore the proposed algorithm can be effectively used for plant identification.

Test Group:	No. samples in the test group	No. of correctly Classified samples	No. of Miss-classified samples			% of Accuracy of the Proposed FRT Algorithm
			FAR	FRR	Miss	
Test Group 1	100	80	5	15	20	80%
Test Group 2	100	90	4	6	10	90%
Test Group 3	100	80	6	14	20	80%
Test Group 4	100	80	4	16	20	80%
Test Group 5	100	90	2	8	10	90%
Total	**500**	**420**	**21**	**59**	**80**	**84%**

Table 2: Experimental Result Analysis

FAR=Fault-Acceptance Rate, FRR=Fault-Rejection Rate and miss=Total number miss- classified.

References

E. D. Dibble and K. J. Killgore, “Measurement of Plant Architecture in Aquatic Plants,” Journal of Freshwater Ecology,1996, vol. 11, pp.
311–318.

H. Molenaar, D. Barthélémy, P.de. Reffye, A. Meinesz, and I. Mialet, “Modelling architecture and growth patterns of Posidonia oceanica,” Aquatic Botany, 2000, vol. 66, pp. 85–99.

Larry E. Morse, “Computer-Assisted Storage and Retrieval of the Data of Taxonomy and Systematics”,Taxon 1974, Vol. 23, No. 1, pp.
29-43.

Wang, Z. , Chi, Z. ; Feng, D. , “Shape based leaf image retrieval” Vision, Image and Signal Processing, IEE Proceedings -Feb 2003,
150 , Issue: 1 , pp34 - 43.

Ji-Xiang Du, De-Shuang Huang, Xiao-Feng Wang, Xiao Gu, “Computer-Aided Plant Species Identification (CAPSI) Based on Leaf shape Matching Technique” Transactions of the Institute of Measurement and Control 28, 3 (2006) pp. 275-284.

Wei Jiang, Cui-cui Ji, Hua Zhu "Fractal Study on Plant Classification and Identification" International Workshop on Chaos-Fractals Theories and Applications, 2009. IWCFTA '09, 6-8 Nov. 2009,pp. 434 – 438.

T. K. Liu, X. X. Zou, T. T. Zhao, H. B. Huang, J. G. Wang, and L. Ma, "Fractal Characteristics of Leaves of Osmanthus fragrans Lour," Journal of Wuhan Botanical Research, 2005, vol. 23(2), pp. 199–202.

J. D. Corbit and D. J. Carbary, "Fractal Dimension as quantitative measure of complexity in plant development," Proceedings of the Royal Society of London B, 1995, vol. 262, pp. 1–6.

E. C. Marie and C. D. John, "Reverse Engineering of Biological Complexity," Science, 2002, vol. 295, pp. 1664–1669.

A. Christoph, O. Charles, and C. C. Travis, "Special Feature: Evolution of biological complexity," PNAS, 2000, vol. 97, pp.4463–4468.

Benoit B. Mandelbrot, The Fractal Geometry of Nature, W.H.Freeman & Co, 1 edition (16 July 2010).

Mandelbrot, B.B. The Fractal Geometry Of Nature. San Francisco 1982.

Eric W. Weisstein. Fractal dimension. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/FractalDimension.html.

R. D. Peng, H. P. Xie, and Y. Ju, "Computation Method of Fractal Dimension for 2-D Digital Image," Journal of China University of Mining & Technology, 2004, vol. 33(1), pp. 19–24.

Odemir Martinez Bruno , Rodrigo de Oliveira Plotze, Mauricio Falvo, Ma´rio de Castro "Fractal dimension applied to plant identification" Information Sciences, 15 June 2008, Information Sciences,Volume 178, Issue 12, Pages 2722–2733.