

Multiscale Fractal Dimension Modelling on Leaf Venation Topology Pattern of Indonesian Medicinal Plants

Aziz Rahmad¹, Yeni Herdiyeni¹, Agus Buono¹, Stephane Douady²

1. *Department of Computer Science, Faculty of Mathematics and Natural Sciences, Bogor Agricultural University, West Java Indonesia*

2. *Laboratoire: Matière et Système Complexe, Batiment Condorcet 10 rue Alice Domon et Leonie Duquet, Université Paris Diderot CC 7056 75205 PARIS Cedex*

Email: aziz.rahmad12p@apps.ipb.ac.id, yeniherdiyeni@ipb.ac.id, pudesha@gmail.com, douady@lps.ens.fr.

Abstract— This research proposed a new model to differentiate leaf venation topology patterns using Multiscale Fractal Dimension. Identification of medicinal plants is important considering wide range of biodiversity in Indonesia and significant role of medicinal plants in Indonesia. Plants identification can be performed with shape analysis using plant leaf venation as a feature. Multiscale Fractal Dimension is a shape analysis method that analyze shapes through its complexity. In this research three Indonesian medicinal plants species has their leaf venation topologies modelled with Multiscale Fractal Dimension. The result shows that while the difference is not remarkably clear, there are irregularities that can be made more evident with multiscale analysis. Future works can include Multiscale Fractal Dimension as one technique to identify plants.

Index Terms— medicinal plant, leaf venation topology, multiscale fractal dimension

I. INTRODUCTION

INDONESIA has a wide range of biodiversity. To date, there are approximately 38.000 recognized plant species in Indonesia [1], of which there are more than 2.000 herbal plants among them [2] and of which almost 80% of them hasn't been cultivated and have to be retrieved directly from wilderness [3]. Herbal medicines have a significant role in Indonesia as it is widely used by people to cure diseases and maintain health because of its inexpensiveness and little to none negative side effects [4]. The significance of herbal medicine combined with the fact that the majority of it is still out there in the wilderness makes identification and classification of herbal plants is paramount.

One of the most common plant organ used for identification is leaf, because it has simple shape, always available, and could be picked without harming the plant. However most information on the leaf like morphological, physiological and genetic properties are

contained in leaf venation [5]. Leaf venation topology is one of the most easily identifiable properties of leaf venation. There are two main variation of leaf venation topology, pinnate with one primary vein and palmate with more than one primary vein. Palmate topology is divided into five more variations. They are parallelodromous, campylodromous, acrodromous, actinodromous, and palinactinodromous [6]. These variations can be distinguished from one another by their distinct shapes.

There are many methods to analyze the shape of an object, one of them is through its complexity [7]. Complexity of a shape is related to the irregularity pattern presented by the shape. One way to estimate shape complexity is by calculating its fractal dimension [8]. There are several definitions of fractal dimension. Bouligand-Minkowski dimension is one of the most widely used definition for shape analysis because it has the most accurate and consistent result [7].

As with all real world objects, leaf venation is not a fractal, albeit it has self-similarity on limited scale [11]. In this case, fractal dimension could be estimated. The simplest estimation method commonly used is linear regression which produce a numeric dimension value. Another more complex method is by applying derivative to track the change of irregularities on different scales, thus produce more detailed information. This method is known as multiscale analysis [12].

This paper propose the modelling of leaf venations with Bouligand-Minkowski fractal dimension using multiscale analysis. Result of this experiment is presented with plot analysis.

II. RESEARCH METHOD

A. Proposed Techniques

The research methodology can be seen in Figure 1. The method consist of data acquisition, preprocessing, fractal modelling and analysis.

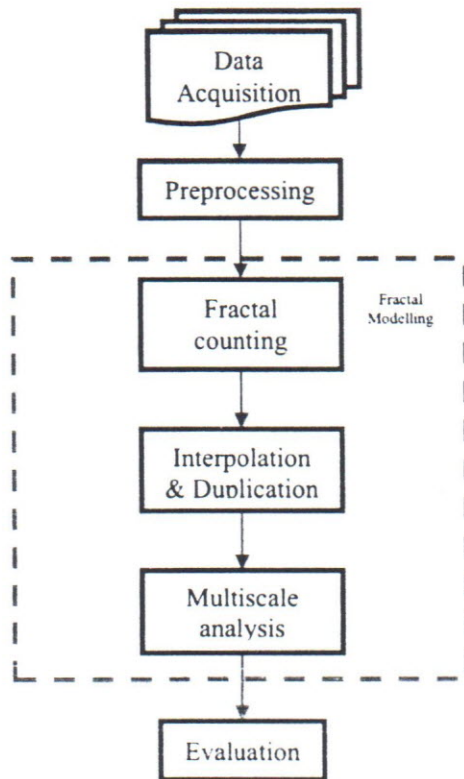


Fig. 1. Proposed Techniques

B. Experimental Material

In this experiment, nine leaves from three different species were used. The three species were *Jatropha curcas* Lin., *Smilax china*, and *Ficus deloidea* L. which represents different venation topology, which are actinodromous, acrodromous, and palinactinodromous, respectively. The images used in this experiment is acquired from Herdiyeni et al. [13]. The images were captured in Biopharmacy Garden of IPB and in the greenhouse of Medicinal Plants of Tropical Forest Ex-situ Conservation Center, Forestry Faculty IPB.

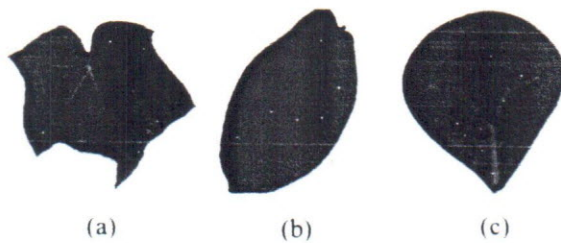


Fig. 2. Leaf species used in the experiment (a) *Jatropha curcas* Lin., (b) *Smilax china*, and (c) *Ficus deloidea* L.

C. Preprocessing

Every digital images of the leaves was uniformly resized to ensure that the size factor will not interfere with the result. The color of each images was also converted to greyscale since the color of the leaf is irrelevant in this experiment. Afterwards, the region of interest in each images, the leaf venation topology, were segmented.

The segmentation method in this experiment

involves edge detection, Gaussian blur, and thresh holding. Edge detection marked every edge in the image by detecting contrast changes of greyscale intensity in the image. Edge detection was applied twice, (i) primary veins segment and (ii) the outline, which will be useful in noise removal later. Gaussian blur was applied to thicken the region of interest and erase the small edges and noises. This results in only primary veins edges remains. Thresh holding is done by converting the image to black and white, thus sharpen the veins already blurred by Gaussian.

After the segmentation it is necessary to apply additional noise filtering. It is done by removing chunks of small regions, leaving only the primary veins in the image. Another filtering is done by using the outline edge from edge detection earlier. This is done to remove the outline from the primary veins. Figure 2 illustrates preprocessing steps used in this experiment.

Because different cameras were used to capture the leaves, it is difficult to find universal parameter for segmentation process. Consequently, the segmentation parameters were entered manually to achieve uniform result.



Fig. 3. Preprocessing steps

D. Fractal Dimension Modelling

Fractal dimension is a measure of how fragmented a fractal object is [7]. It could identify how complex a fractal is by comparing the changes of irregularity in a shape as the scale changes. Fractal dimension is also characterize self-similarity of the shape. An object is said to be self-similar if it is approximately similar to its parts. Self-similarity is an important feature of fractal, as fractals have self-similarity in every scale.

There are many definitions of fractal dimension, one of which is Bouligand-Minkowski dimension. The formula to model the fractal dimension according to Bouligand-Minkowski can be defined as

$$DB = 2 - \lim_{r \rightarrow 0} \frac{\log A(r)}{\log(r)} \quad (1)$$

where $A(r)$ can be defined as the number of counted element and r is the size of the counting window [14]. There are two major steps in fractal dimension modelling, counting and estimation. The counting step is done to produce a curve that describes different measurements on different scales. The estimation step is done to get the actual fractal dimension from the curve by estimating the limit. This estimation part is important because self-similarity of real world objects are only on limited scale.

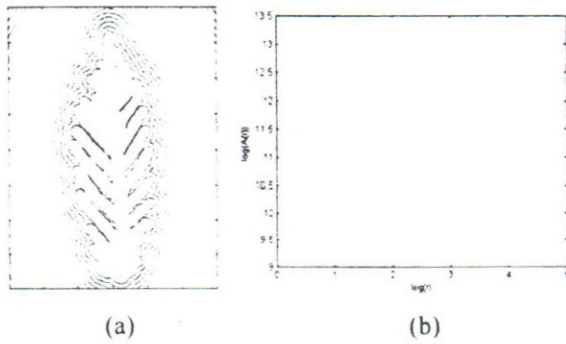


Fig. 4. Illustration of fractal counting with dilation (a) dilated image (b) plot of dilation distance and dilation area

One of the techniques to count Bouligand-Minkowski dimension is the dilation method. This method computes the influence area by calculating the area of a shape ($A(r)$) after being dilated by a disc of a certain radius (r). Since this method, particularly the dilating process, is considerably time-consuming, another more efficient method is used in this experiment. Euclidean Distance Transform (EDT) can calculate the Euclidean distance between each foreground pixels to the region of interest with good performance [15]. From the distance map the area of dilated shape ($A(r)$) can be calculated by summing up all the pixels whose distance is r or less from the region of interest. The result of this step can be presented as a log-log curve of $r-A(r)$ or dilation curve. For this experiment, the distance were calculated to $r = 100$. Figure 3(a) shows an illustration on dilation method performed on an image and 3(b) shows the resulting plot of shape area and its radius.

The curve produced from the counting method is an empirical curve because it is consisted of data points. To get the actual fractal dimension, the estimation of the curve is necessary. The simplest way to estimate the curve is by using linear regression. This method calculate the curve by drawing a straight line that approximate the curve. This straight line can then be calculated for its gradient to obtain the fractal dimension. As the gradient is a numeric value, so does the fractal dimension. While this method is simple, the result is often cannot describe the complex nature of object shape. The similar shape may have the exact same fractal dimension value, thus the small difference of the shape may not be noticed.

Better information on image shape can be obtained by calculating first derivative of the log-log curve. The result will be a curve that binds fractal dimension changes to the dilation radius changes. The curve is called Multiscale Fractal Dimension (MFD). It was defined as

$$MFD = 2 - \frac{du(t)}{dt} \quad (2)$$

where $u(t)$ represent $\log(A(r))$ and t represents $\log(r)$.

Unlike dilation curve that tracks the change of measurement based on size, the multiscale curve tracks

the rate of change of the measurement. This means multiscale curve has richer information than dilation curve. Dilation curve is always linear because $A(r)$ will always increase as r increases. Multiscale curve, on the other hand, depends on how much $A(r)$ increase as r increases. This means that it is possible for multiscale curve to have local maxima and minima values. These values can represents the whole curve into a much smaller number and still retain important information of the curve such as the curve shape [9].

One of the main problem of multiscale curve is that there are many redundant information contained within. These redundant information could slow the computational process and blurred the differences among fractal dimensions of different images. This problem can be solved by applying descriptor to the curve. There are several descriptors that could be used to remove redundant information from multiscale curve, one of which is known for its short computing time is Fourier descriptor. Fourier descriptor could also produce data that is invariant to rotation, translation, and scale, thus it is suitable for pattern recognition [11].

The problem with derivation is that it has tendency to enlarge high-frequency noise. This could affect the result greatly because sometimes the shape under analysis has unpredictable noises. This problem can be solved by applying low-pass filter to the curve. One of the commonly used low-pass filter is Gaussian filter. Thus, the Fourier descriptor of the derivative is defined as

$$\frac{du(t)}{dt} = F^{-1}\{F\{u(t)\} \{g_{\sigma}(t)\}(j2\pi f)\} \quad (3)$$

where f as frequency, j as imaginary number $\sqrt{-1}$ and $g_{\sigma}(t)$ is Gaussian filter applied to $\log(r)$ and defined as

$$g_{\sigma}(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-t^2}{2\sigma^2}\right) \quad (4)$$

In the equation above σ represents standard deviation that could be adjusted. In this experiment, standard deviation $\sigma = \{1 \dots 10\}$ were used.

It is necessary to provide a curve with good sampling and uniform interval. Unfortunately the dilation curve has inconsistent interval in its nature. There are two ways to approach this. The first one is to remove the initial points with very low sampling. The second one is by using linear interpolation to fill the space between every two points of the sampled curve by its average point.

The discontinuity of the dilation curve can produce either overshoot or undershoot near its Fourier approximation. This phenomenon is known as Gibbs phenomenon [10]. The solution to this problem is by applying duplication and reflection to the curve, to make the curve seems continuous. The Fourier approximation can later be cropped according to the original curve limits.

III. RESULTS AND ANALYSIS

A. Preprocessing

Preprocessing in this experiment were performed using Matlab and GIMP. Additional noise that unable to be removed in Matlab were removed using GIMP. Figure 5(a) shows some leaf images from each species and Figure 5(b) shows the same images after being preprocessed.

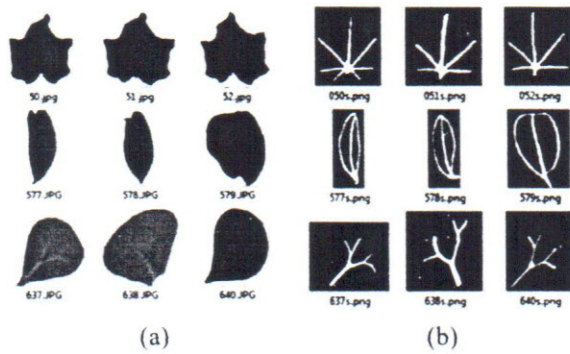


Fig. 5. (a) Original leaf images (b) preprocessed venation images

B. Fractal Counting

Measurement of fractal dimension in this experiment was performed using EDT method. Figure 6 shows the log-log curve of the fractal counting result. The figure shows that while dilation curves of *Jathropa curcas Lin* (red) and *Ficus deloidea L.* (green) has formed separated clusters, *Smilax china* (blue) dilation curves are in mixed positions. This is mainly caused by lack of uniqueness possessed by *Smilax china* leaf venations. Unlike *Jathropa curcas Lin* and *Ficus deloidea L.* leaf venations which possess distinctive branches of primary veins, the leaf venations of *Smilax china* has only three curved primary veins. These veins evidently did not provide *Smilax china* with enough features to separate it from other, more complex venations. The blue curves in the plot are seen mixed mostly with green curves. This is caused by the similarity between the two venation topologies.

It is worth mentioning that the curve itself describes increases in the area of the objects, not the shape of the objects itself. It is from the different way the area changes that the difference of shape can be deduced. On smaller dilation radius the area of each leaf venation are similar. This is because the scale of each leaf venation was deliberately uniformed in preprocessing step. As dilation radius became wider the difference between areas of shapes became more evident. This is illustrated by gradient of the curve. Higher gradient means bigger area increase rates. Therefore dilation curves do not associated with shapes directly, instead it describe the uniqueness of shapes through area increase rates.

C. Fractal Estimation

Dilation curves that has been interpolated and duplicated can then have its fractal dimension estimated by multiscale analysis. The result of multiscale analysis to each curve is presented in Figure

7. It is clearly shown that unlike dilation curves, MFD curves have local maxima and minima. These local extrema can represents the curve without omitting important features of the curve.

From the trends of MFD curves at a glance it looks similar. It should be noted that red and green curves are clustered like their counterparts in dilation curves while the blue curves are not. However unlike the condition in dilation plot where the blue curves are mixed with the green curves, in MFD plot the blue curves are mostly above the green curves. This shows that while *Smilax china* and *Ficus deloidea L.* leaf venation topologies are similar in shape but there are some different irregularities between them that is more evident after multiscale analysis.

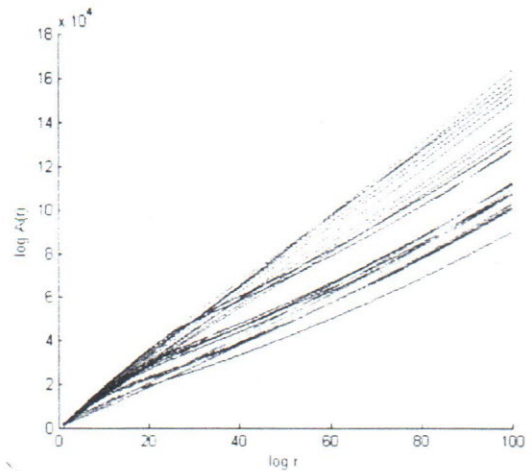


Fig. 6. Log-log dilation curves

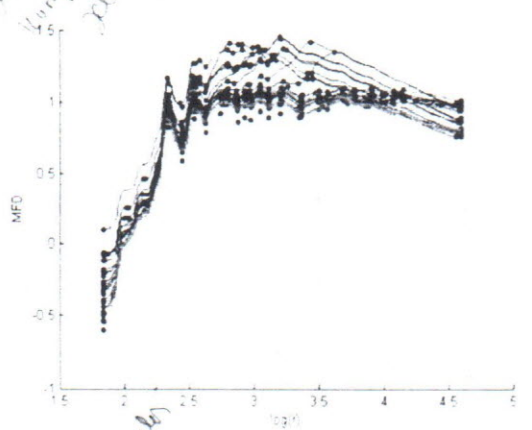


Fig. 7. MFD curves

IV. CONCLUSION

This paper presents an experiment in modelling leaf venation using Bouligand-Minkowski multiscale fractal dimension. The fractal counting is performed using Euclidean distance transform and the estimation is performed using multiscale analysis. Fourier descriptor have been performed on multiscale fractal dimension curve. The result from both Euclidean distance transform and multiscale analysis is presented through plots. While it is not remarkably clear, there are some irregularities that could be detected with dilation

but more evident after multiscale analysis.

REFERENCES

- [1] Bappenas. 2003. *Indonesia Biodiversity Strategy and Action Plan 2003-2020*. Jakarta: Bappenas.
- [2] Zuhud, EAM. 2009. *Potensi Hutan Tropika Indonesia sebagai Penyanga Bahan untuk Kesehatan Bangsa*. *Jurnal Bahan Alam Indonesia*, 6(6).
- [3] Masyhud. 2010. *Lokakarya Nasional Tumbuhan Obat Indonesia 2010*. <http://www.dephut.go.id> [11 July 2014].
- [4] Gogtay NJ, Bhatt HA, Dalvi SS, Kshirsagar NA. 2002. The use and safety of non-allopathic Indian medicine. *Drug Saf*. 25:1005-1019.
- [5] Sack L, Scoffoni C. 2013. Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future. *New Phytologist*.
- [6] Daly DC, Hickey LJ, Johnson KR, Mitchell JD, Wilf P, Wing SL. 2009. *Manual of leaf architecture*. CABI.
- [7] Backes AR, Bruno OM. 2008. Fractal and Multi-Scale Fractal Dimension analysis: a comparative study of Bouligand-Minkowski method. *INFOCOMP Journal of Computer Science*, 7(2): 74-83.
- [8] Mandelbrot BB. 2000. *The Fractal Geometry of Nature*. New York: W.H. Freeman and Company.
- [9] Bruno OM, de Oliveira Plotze R, Falvo M, de Castro M. 2008. Fractal dimension applied to plant identification. *Information Sciences*, 178(12): 2722-2733.
- [10] Schroeder M. 1991. *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*. New York: WH. Freeman
- [11] Florindo JB, Backes AR, de Castro M, Bruno OM. 2012. A comparative study on multiscale fractal dimension descriptors. *Pattern Recognition Letters*, 33(6): 798-806.
- [12] Emerson CW, Siu-Ngan Lam N, Ouattrochi DA. 1999. Multi-scale fractal analysis of image texture and patterns. *Photogrammetric Engineering and Remote Sensing*, 65: 51-62.
- [13] Herdiyeni Y, Nurfadhilah E, Zuhud EA, Damayanti EK, Arai K, Okumura H. 2013. A Computer Aided System for Tropical Leaf Medicinal Plant Identification. *International Journal on Advanced Science, Engineering and Information Technology*, 3(1): 23-27.
- [14] Tricot C. 1995. *Curves and fractal dimension*. Springer.
- [15] Bruno OM, Costa LDF. 2004. A parallel implementation of exact Euclidean distance transform based on exact dilations. *Microprocessors and Microsystems*, 28(3): 107-113.