Landmark Analysis Of Leaf Shape Using Polygonal Approximation

Zakhi Firmansyah¹, Yeni Herdiyeni¹, Bib Paruhum Silalahi², Stephane Douady³
¹Department of Computer Science, Faculty of Mathematics and Natural Sciences. Bogor Agricultural University, West Java Indonesia
²Department of Mathematic, Faculty of Mathematics and Natural Sciences. Bogor Agricultural University, West Java Indonesia
³Laboratoire: Matiere et System Complex, Université Paris Diderot, France

E-mail: zakhi99@gmail.com, yeni.herdiyeni@ipb.ac.id, bibparuhum@gmail.com

Abstract. This research proposes a method to extract landmark of leaf shape using static threshold of polygonal approximation. Leaf shape analysis has played a central role in many problems in vision and perception. Landmark-based shape analysis is the core of geometric morphometric and has been used as a quantitative tool in evolutionary and developmental biology. In this research, the polygonal approximation is used to select the best points that can represent the leaf shape variability. We used a static threshold as the control parameter of fitting a series of line segment over a digital curve of leaf shape. This research focuses on seven leaf shape, i.e., eliptic, obovate, ovate, oblong and special. Experimental results show static polygonal approximation shows can be used to find the important points of leaf shape.

1. Introduction
Botanist observe various features in plants using taxonomic key to identify the species of a plant. One feature that is observed to determine the species of plants are leaf shape. Each class has a unique shape. In addition, the general class of leaf shape is also determined by the ratio of length to width leaf and the widest position of the leaf. Previous research related to plant identification such as Gwo et al. (2013) use key point along leaf boundary, centroid dan fuzzy histogram to representation leaf shape. There are so many works related shape identification using landmark analysis. Takemura et al. (2002) introduces a new shape analysis approach using the well known wavelet transform and exploring fish shape representation by landmarks. In this paper, a new method using static polygonal approximation to analyze landmark of leaf shape is proposed.

1.1. Landmark
Landmark is a biologically definable point on an organism, that can be sensibly compared between related organisms. It typically requires domain-specific expert knowledge to choose a suitable set of landmarks.
Landmark-based shape analysis has been the core of morphometrics geometry and has been used as a quantitative tool in evolutionary and developmental biology (Klingenberg, 2010). However, it exhibits three main fundamental limitations (Laga 2014):

First, it assumes that the landmarks are already in correspondence across shapes. Shapes of plant leaves, however, exhibit large non-rigid variations, even within a single species. This makes finding the correspondences across shapes even more challenging.

Second, landmark-based methods are not invariant to all the shape-preserving transformations. They can be made, in a preprocessing step, invariant to translation, scale, and rotation of the shapes. However, they are not invariant to parameterization of the shapes.

Finally, landmark-based methods assume that the shapes to be analyzed are points in a high-dimensional Euclidean space and are distributed following a Gaussian model. This limits the type of statistical analysis that can be performed to Gaussian distribution only.

1.2. Leaf Shape

Leaf shape is one of the most important features in describing the plant. Humans can easily identify the different types of leaves and classify them into different species based on information found on the leaves. Based on the shape, the leaves are grouped into 16 classes (Harlow dan Harrar, 1969). There are classes to distinguish the width and length ratio and their unique shape. Fig. 1 shows the 16 classes of shape i.e. there are: (1) Accicular, (2) Scalelike, (3) Linear, (4) Oblong, (5) Lanceolate, (6) Oblanceolate, (7) Ovate, (8) Obovate, (9) Elliptical, (10) Oval, (11) Orbicular, (12) Reniform, (13) Cordate, (14) Deltoide, (15) Rhomboid, (16) Spatulate. Each class has a characteristic shape of leaves typical.

![Fig. 1. Leaf Shapes Of Harlow and Harrar.](image-url)
1.3. Laminar shape

The simplest way to describe the overall shape of the lamina is to locate the axis or, in some cases, the zone of greatest width that lies perpendicular to the axis of greatest length or long axis (Wing et al. 1999):

a. **Elliptic**
   The widest part of the leaf is on an axis in the middle of the long axis of the leaf.

![Elliptic](image1.png)

Fig. 2. Elliptic

b. **Obovate**
   Widest part of the leaf is on an axis in the apical 2/5 of the leaf

![Obovate](image2.png)

Fig. 3. Obovate

c. **Ovate**
   Widest part of the leaf is on an axis in the basal 2/5 of the leaf

![Ovate](image3.png)

Fig. 4. Ovate

d. **Oblong**
   Widest part of the leaf is a zone in the middle 1/3 of the long axis where the opposite margins are roughly parallel.

![Oblong](image4.png)

Fig. 5. Oblong
2. Methods

2.1. Polygonal Approximation

Polygonal Approximation is a simplification of the method of representation form the curve (Prasad et al. 2012). For example, the set \(\{P_1, P_2, P_3, ..., P_n\}\) as shown in Figure 6, is a series of dots representing a sequential digital curve, then given treatment so that the set of points is reduced to \(\{Q_1, Q_2, Q_3, ..., Q_m\}\) with \(m < n\).

![Figure 6. Polygonal Approximation](image)

There are two main ways to reduce the polygonal approximation point (Grigore O et al. 2003):

- **Min – \(\varepsilon\)**: is the approach to determine the number of points so that predictions can represent essentially the shape of the curve as shown in Figure 7.

![Figure 7. Min - \(\varepsilon\) Polygonal Approximation](image)

- **Min - \#:** is an approach that determines the margin of error which is a deviation limits in representing the final results as shown in Figure 8.

![Figure 8. Min - \# Polygonal Approximation](image)

Prasad et al. (2012) optimize the polygonal approximation of digital curves that had previously been done by Lowe (1987), Ramer (1972), Douglas & Peucker (1973) or L-RDP, digital curve \(e = \{P_1, P_2, ..., P_n\}\), with \(P_i\) is i the number of points the edge of the pixel in the digital curve e. Line through a pair of pixels \(P_a (x_a, y_a)\) and \(P_b (x_b, y_b)\) is given by Equation 1.

\[
x(y_a-y_b)+y(x_a-x_b)+y_a x_b - y_b x_a = 0
\]

(1)

and deviation \(d_i\) of the pixels \(P_i (x_j, y_j) \in e\) of the line passing through the pair \(\{P_1, P_n\}\) is given by Equation 2.
\[ d_i = |x_i(y_1 - y_N) + y_i(x_N - x_1) + y_N x_1 - x_N| \]  

(2)

From the level difference \( d_i \) can be seen how far the deviations that occur, as shown in Figure 9.

![Figure 9. The degree of difference \( d_i \)](image)

Based on the above picture, \( d_i \) is the degree of difference of \( P_2 \) against the pair of points \( P_1 \) and \( P_4 \), and \( d_i \) have the biggest difference level. By this means that the greater the degree of similarity between the \( d_i \) the digital curve and the results are becoming increasingly different polygonal approximation.

Thus, pixels with a maximum deviation can be found, denoted by \( P_{\text{max}} \), then given the pair \( \{P_1, P_{\text{max}}\} \) and \( \{P_{\text{max}}, P_N\} \), so we get two new pixels of the curve \( E \) using Equation 1 and 2 above. This process is repeated until certain conditions are met by all segments of the line.

The conditions used by L-RDP [Douglas & Peucker (1973), Lowe (1987), Ramer (1972)] is that for each line segment, the maximum deviation of pixels contained in the corresponding edge segment is less than a certain tolerance value as Equation 3 below.

\[ \max(d_i) < d_{\text{tol}} \]  

(3)

With \( d_{\text{tol}} \) is the threshold or margin selected. As shown in Figure 10.

![Figure 10. Illustration \( d_{\text{tol}} \) on Polygonal Approximation](image)
3. Results and discussions

In this paper used image dataset from Machine Learning Repository UCI [21]. The Leaves were collected in the Royal Botanic Gardens, Kew, UK. All there are 100 species of leaf and every species there are 16 image data of a binary image. In this paper have been classified some image data according to Benson leaf shape. There are five leaf shape that can be analyzed is Eliptic, obovate, ovate, oblong and special

3.1. Static Threshold Polygonal Approximation in Leaf Shape

We know that edge shape of leaf consists of successive points. That meant the points can choose the best points which could indicate landmarks of leaf shape (Fig. 11). Static value $d_{tot}$ 0.5, 1.0, 1.5, and 2.0 Polygonal Approximation is used in this research

![Eliptic and Obovate Leaf Shapes](image)

Figure 11. Sample Leaf shape with static value $d_{tot} 2.0$

3.2. Eliptic Shape Analysis

Eliptic is the shape of the leaves that have the widest part at the center of the leaf. Elliptical leaves have the broadest width in the middle and then taper off at the ends As we see in the Fig. 12 a and b, the elliptic shape is determine in point 4, 5, 6 which have more point in the middle of the leaf for threshold 2.0.

In point 1, 2, 3 respectively be formed as the base of the leaf shape and 7, 8 be formed as the apex of the leaf shape which have less broader than the middle.

![Eliptic Shape Analysis](image)

Figure 12. Eliptic Shape Analysis

3.3. Obovate Shape Analysis

Obovate is the shape of the leaves that have the widest part of the leaf is on an axis in the apical 2/5 of the leaf. As we see in the Fig. 13 a and b is two species from obovate shape, the obovate shape is determine in point 7, 8, 9, 10, 11 which have more point in the apex of the leaf for threshold 1.0.

In point 1, 2, 3 respectively be formed as the base of the leaf shape and 4, 5, 6 be formed as the middle of the leaf shape which have less broader than the apex
3.4. Ovate Shape Analysis

Ovate is the shape of the leaves that have widest part of the leaf is on an axis in the basal 2/5 of the leaf. As we see in the Fig. 14 a and b is two species from ovate shape. the ovate shape is determine in point 1, 2, 3, 4 which have more point in the basal of the leaf for threshold 2.0.

In point 5, 6 and 7 respectively be formed as the middle of the leaf shape and 8, 9 be formed as the apex of the leaf shape which have less broader than the middle.

3.5. Oblong Shape Analysis

Oblong is the shape of the leaves that have the widest part of the leaf is a zone in the middle 1/3 of the long axis where the opposite margins are roughly parallel. As we see in the Fig. 15 a and b, the oblong shape is determine in point 3, 4 and 5, 6 which have more point in the middle of the leaf and roughly parallel for threshold 1.5.

In point 1, 2 respectively be formed as the apex of the leaf shape and 7, 8 be formed as the base of the leaf shape which have less broader than the middle 1/3 of the long axis.

4. Conclusions

This study presents shape representation of the leaf based on landmark using polygonal approximation method. Based on the experimental results, polygonal approximation method can represent landmarks of four leaf shapes such as elliptic, obovate, ovate, and oblong. To get the best landmark, necessary to determine the value of static threshold manually, it is not effective so that required dynamic threshold.

5. References

[1] Dilip K. Prasad, Chai Quek, Maylor K. H. Leung, and Siu-Yeung Cho. 2011. “A parameter independent line fitting method,” 1st Asian Conference on Pattern Recognition (ACPR 2011), Beijing, China, 28-30 Nov;
[2] Dilip K. Prasad and Maylor K. H. Leung. 2012. "Polygonal Representation of Digital Curves,” Digital Image Processing, Stefan G. Stanciu (Ed.), InTech.
[3] Douglas, D. H., & Peucker, T. K. 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: The International Journal for Geographic Information and Geovisualization, 10(2), 112-122.
[4] Grigore O et al. 2003. On the Implementation of Polygonal Approximation Algorithms. Department of Information and Computing Sciences, Utrecht University. technical report UU-CS-2003-005.

[5] Gwo., et al. 2013, Plant Identification Through Images: Using Feature Extraction of Key Points on Leaf Contours, Botanical Society of America

[6] Harlow WM, Harrar ES. 1969. Textbook of Dendrology. Mc.Graw-Hill: United States of America

[7] Joao Camargo Neto, George E. Meyer, David D Jones, Ashok K. Samal. 2006. Plant species identification using Elliptic Fourier leaf shape analysis.

[8] Laga H, Kurtek S, Srivastava A, Miklavcic S J. 2014. Landmark-Free Statistical Analysis of The Shape of Plant Leaves. Journal of Theoretical Biology. Volume 363, 21 December 2014, Pages 41–52.

[9] Lowe, D. G. 1987. Three-dimensional object recognition from single two-dimensional images. Artificial Intelligence, 31(3), 355-395.

[10] Lukasiewicz , T. U. Straccia . 2009 . Description logic programs under probabilistic uncertainty and fuzzy vagueness. International Journal of Approximate Reasoning 50 : 837 – 853 .

[11] Ramer, U. 1972. An iterative procedure for the polygonal approximation of plane curves. Computer Graphics and Image Processing, I(3), 244-256.

[12] Takemura et al. 2002. Shape Analysis and Classification using Landmarks: Polygonal Wavelet Transform

[13] Wing et al. 1999 Manual of Leaf Architecture - morphological description and categorization of dicotyledonous and net-veined monocotyledonous angiosperms by Leaf Architecture Working Group. 65p.