Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE 5 (2018) 234–245

journal homepage: www.elsevier.com/locate/inpa

Feature extraction of overlapping hevea leaves: A comparative study

Sule T. Anjomshoae *, *Mohd Shafry Bin Mohd Rahim*

Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Malaysia

ARTICLE INFO

Article history:
Received 15 December 2017
Received in revised form
31 January 2018
Accepted 2 February 2018
Available online 12 February 2018

Keywords:
Rubber tree leaves
Feature extraction
SIFT
Harris
FAST
Hough transformation
Boundary extraction

ABSTRACT

Automation of rubber tree clone classification has inspired research into new methods of leaf feature extraction. In current practice, rubber clone inspectors has been using several leaf features to identify clone types. One of the unique features of rubber tree leaf is palmate leaflets. This characteristic generates different leaflet positions, where the leaves are overlapping or separated. In this research, we propose keypoint extraction and line detection methods to extract shape and axil (angle between petioles) features of leaflet positions. The results of keypoint extraction methods, namely, SIFT, Harris, and FAST, were compared and discussed for shape feature extraction. Next, Hough transformation and boundary-tracing methods were compared to identify the suitable axil detection method. The evaluation result demonstrates the proper keypoint extraction method for shape context and the clear advantages of Hough Transformation in accuracy of angle detection.

1. Introduction

The rubber tree, scientifically known as Hevea Brasiliensis, is prevalent in Thailand, Malaysia, Indonesia, the southern part of India, and Sri Lanka [1]. These countries have maintained a successful clone development and breeding program for decades. Today, the rubber tree has been the primary source of natural rubber for worldwide use. This tree has high export growth as one of the most profitable agri-industrial ventures.

Asia's well-known rubber tree boards are Rubber Research Institute of Malaysia (RRIM), Rubber Research Institute of India (RRII), and Rubber Research Institute of Sri Lanka (RRISL). These boards have the primary role in recommending and distributing clones to the cultivators. One of the important factors that affect the properties of raw rubber is the clonal origin of the rubber tree [2]. The clone inspection and verification is critical because it must guarantee that recommended rubber clones produce the maximum yield in the future. One way to identify rubber clones is based on the physical characteristics of leaves. This method requires experts with adequate experience. Therefore, the automation of this process for clone classification is the subject of new research.

Currently, plant classification and recognition methods are implemented on plant components, such as flowers, leaves, and bark [3]. Because the reproductive organs such as flowers, only available in a particular season, the plant classification systems based on leaves are more widespread. Leaf-based plant classification methods generally use the characteristics

* Corresponding author.
E-mail addresses: tsule3@live.utm.my (S.T. Anjomshoae), shafry@utm.my (M.S.B.M. Rahim).
Peer review under responsibility of China Agricultural University.
https://doi.org/10.1016/j.inpa.2018.02.001

of leaves such as shape, color, and texture [4]. However, existing plant classification methods (e.g. WAPSI: web application for plant species identification) are unable to classify rubber tree clones due to leaves having similar physical features. Rubber tree leaves have other features that differentiate clones from one another. Leaf tip, leaf base, form of the leaf, and leaf margin are among the attributes that might be used for the clone classification system. Another feature of the rubber tree is that three compound leaflets radiate from one mutual leaf base (palmate leaflets). These leaflets exist in three positions: overlapping, touching, or separated. Overlapping and non-overlapping leaves can be identified based on shape and angle features.

In this paper, we introduce the features of overlapping and non-overlapping leaves of the rubber tree and present a framework to extract these features. SIFT (Scale-invariant feature transform) is proposed to extract shape features from rubber tree leaf images. However, we present a comparative study of the SIFT, Harris, and FAST (Features from accelerated segment test) methods for keypoint extraction. Then, the angle feature extraction is carried out to discern the overlapping and non-overlapping leaves. The proper angle extraction method is identified based on comparing two methods: Hough transform and boundary-tracing. Hough transform is an effective tool for detecting geometrical features, while boundary-tracing is an alternative method for line computation and angle extraction. Both methods are discussed in the following sections.

The paper is organized as follows. The next section presents an overview on keypoint extraction methods and line detection. Section 3 introduces the features of rubber tree leaf to recognize different clones. In Section 4, overlapping and non-overlapping rubber tree leaf features are introduced. Section 5 explains the feature extraction of hevea leaf for leaflet positions. Section 6 presents the experimental results and the performance evaluations. In Section 7, we discuss the results of the comparative analysis and Section 8 concludes the paper.

2. Related work

Plants are commonly recognized based on leaf shape and texture. Shape-based classification methods are considered as contour-based and region-based classification [5,6]. Several researches were focused on the contour-based methods [7–9]. Lee et al. [10] presented Fast Fourier transform (FFT) using frequency domain data. They used FFT to extract twenty features by calculating the distance between the centroid and all points on the leaf contours. Caballero et al. [11] used the contour-based descriptors on their web application for plant species identification. The contour descriptors are implemented for the assessment of similarity between two images. Zheng and Wang [12] developed leaf shape extraction method based on visual consistency. Although, this method can describe shapes of several plant leaves, it is unable to recognize the shape features for clone within a particular type. Bhardwaj et al. [13] extracted four shape features: area convexity, volume fraction, moment invariant, and inverse difference moment. These features enhanced the classification result for the structurally complex images. Wu et al. [14] have studied Probabilistic Neural network (PNN) to classify 32 types of plants based on their shapes, while Kadir et al. [15] integrated vein and color features with shape and texture features in their PNN method.

Many feature extraction methods have considered different plant leaf features, rather than clone leaf features. Available classification methods are not suitable for overlapping rubber tree leaf recognition because they focus on features of different plant types. Furthermore, these studies paid little attention to compound leaflet identification. This research focuses on SIFT keypoint extraction and Hough angle extraction methods for overlapping and non-overlapping rubber tree leaf identification. There have been several works on keypoint extraction and line detection. A few related papers are reviewed here.

SIFT keypoints are widely used in computer vision applications that require fast and efficient feature matching, such as object detection, feature description, and object tracking [16–19]. Pan and Lyu [20] presented a method to detect duplication of a particular region in the same image based on SIFT features. Mehrotra et al. [21] suggested SIFT keypoints to extract local features from a noise independent iris image to overcome the effect of partial occlusion due to eyelids. Tao et al. [22] proposed a method that described an airport by a set of SIFT keypoints and identified the airport by keypoint matching algorithm. This algorithm used the clustering information from matched keypoints after locating the corresponding keypoints. Zahedi et al. [23] presented a recognition method using SIFT features for the Farsi and Arabic font. The method is robust to varying size, scale, and rotation of the fonts.

Hough transformation is an effective tool for detecting predefined shapes. This method is commonly being used for image processing, pattern recognition, and computer vision [24]. One of the early implementations proposed by Stylianidis and Patias [25] was the detection of straight lines for digital close range photogrammetry. These features were helpful for further photogrammetric work, such as sensor calibration, and image orientation. Later, Weiss [26] introduced Hough transformation in real-time symbol detection. Hough transforms produced results of high accuracy and performance

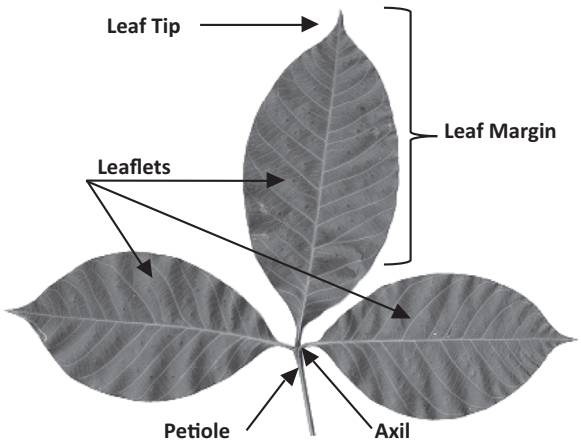

Fig. 1 – External features of rubber tree leaf.

Table 1 – Clones with different features. PB 350 (Prang Besar), RRIM 3001 and RRIM 2025 (Rubber Research Institute of Malaysia).

Clones	PB 350	RRIM 3001	RRIM 2025
Shape	Round	Elliptic	Obovate
Color	Fade	Shiny	Clear
Leaf base	Obtuse	Attenuate	Cuneate
Position	Overlap	Separate	Separate
Leaf tip	Cuspidate	Aristate	Acuminate

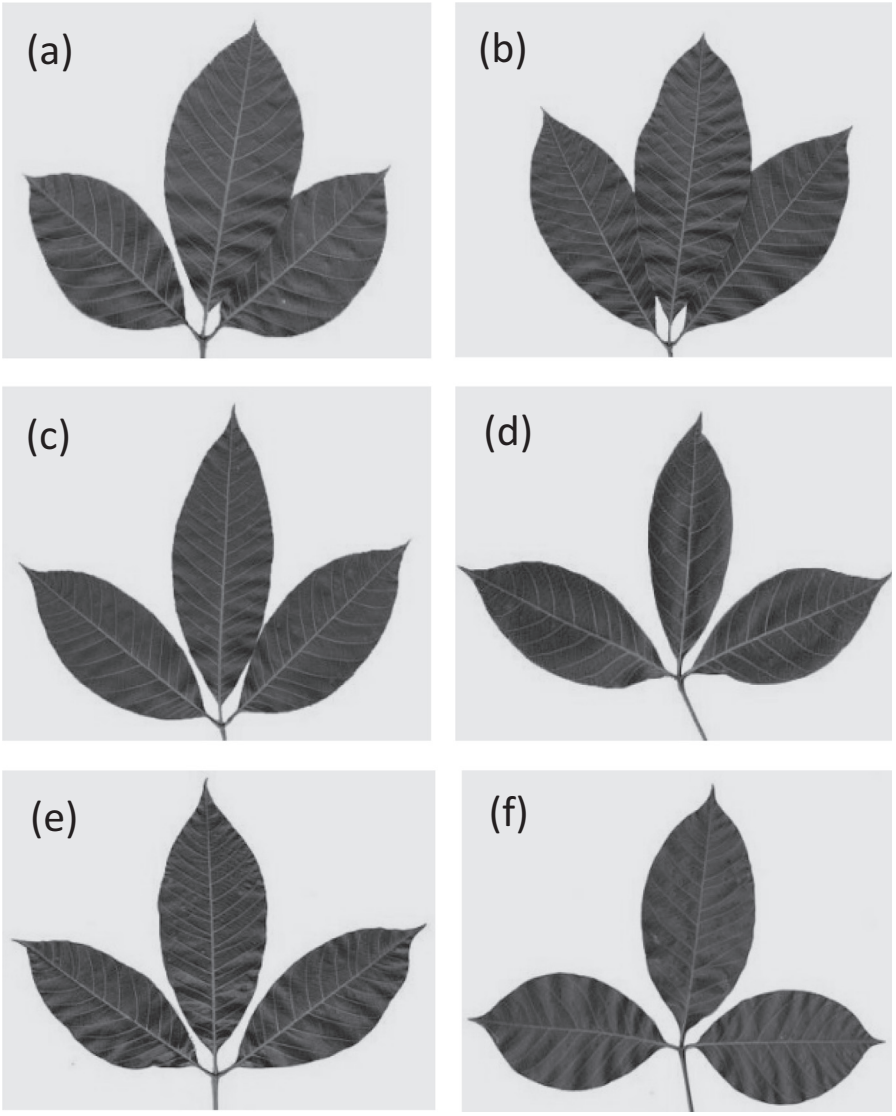

Fig. 2 – Rubber tree leaf samples with different shapes and leaf positions (a) obovate-overlapping leaflets, (b) rounded-overlapping leaflets, (c) elliptic-touching leaflets, (d) elliptic-separated leaflets, (e) elliptic-separated leaflets, (f) rounded-separated leaflets.

in real-time video streams. Similar work was presented by Chakraborty [27]; he extracted large-scale events in video streams using Hough transformation. A maximum margin algorithm was implemented to examine the weights of the visual vocabularies. The method is applied directly to the extracted features to avoid redundant comprehensive examination, which is inapplicable for the activity detection problem. Hough transformation has also been used on

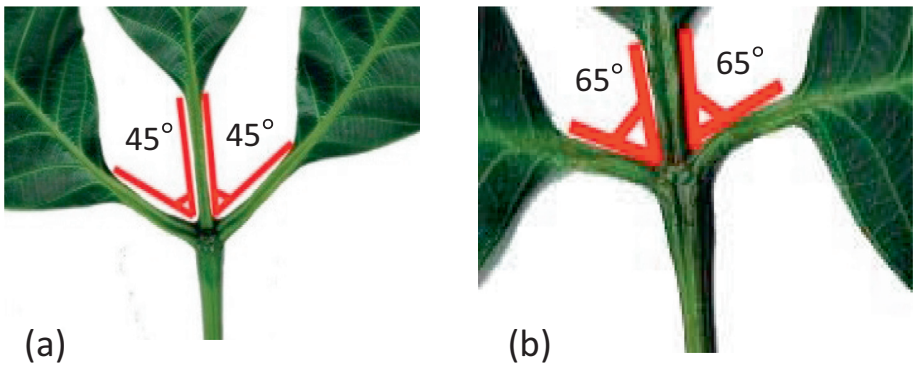

Fig. 3 – Petiole degrees of rubber tree leaflets for clone types (a) RRIM 2001 and (b) PB350.

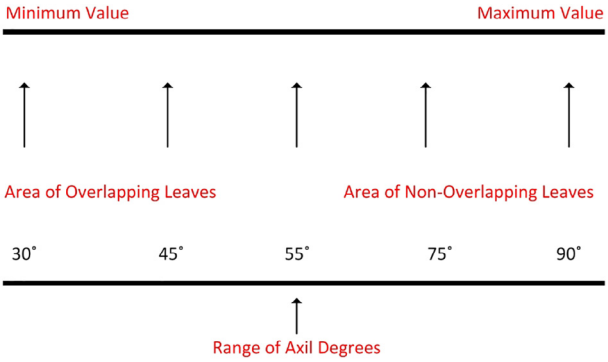

Fig. 4 – Range of axil degrees obtained from ground truth data set.

graphics-based shape extraction for complex building descriptions by Cui et al. [28]. The proposed algorithm worked on the roofs of flats constructed using right angles. Although Hough transformation method has been applied in many different cases, angle extraction for axil computation has not been explored. This paper demonstrates angle extraction of rubber tree leaf petioles using Hough transformation.

3. General features of rubber tree leaf to identify clone types

A typical rubber tree has composite leaves that are broken up into three separate leaflets. Leaf colors are either light green or dark green; retains green leaves throughout the year. However, the different color areas might be caused by diseases or lack of nutrients [29]. Fig. 1 demonstrates a number of the external features of the rubber tree leaf along with their botanical terms.

There are more specific features that differentiate rubber clones from one another. Some of the leaves demonstrate oval shape with a semi-glossy surface, while others represent an elliptical shape with dark shining green. Table 1 shows rubber tree clones with their distinguishing leaf features. Rubber tree clone names are designated by letters indicating its place of origin and a serial number as shown in the Table 1.

3.1. *Dataset*

The data set used in this research is based on our previous work, which was collected at the Rubber Research Institute of Malaysia (RRIM), Kuala Lumpur [30]. The recommended trees for automated clone classification were an average two years old, which were mature enough for clone inspection. The samples contain a total of 250 leaves from 25 different rubber trees that were classified into five classes manually based on their clone name.

4. Overlap and non-overlap leaf features

4.1. *Shape*

Rubber tree leaf shape appears in three different forms, which are elliptic, obovate, and round. Fig. 2 shows the different leaf forms with different leaflet positions. The middle leaf might be occluded by one adjacent leaf or by both adjacent leaves, while in some cases, leaves are separated from each other.

4.2. *Angle*

Although leaf shape identifies the overlapping and non-overlapping leaves, the other way to identify the leaflet position is the angle between petioles. In botanical terminology, angle between petioles corresponds to the axil. The axil degrees of overlapping and non-overlapping rubber tree leaves assist in differentiating the clone origin. Overlapping leaves makes an acute angle while non-overlapping leaves have the right or obtuse angle. Therefore, the ranges of axil degrees were explored to identify overlapping and non-overlapping leaves. Fig. 3 demonstrates different clones with different angles.

Before assessing the angle feature extraction, the ranges of axil degrees of overlapping and non-overlapping leaves are identified. This study demonstrated that the angle degrees of overlapping leaves range from 30° to 55° while the angle degrees of non-overlapping leaves range from 55° to 90° such as shown in Fig. 4.

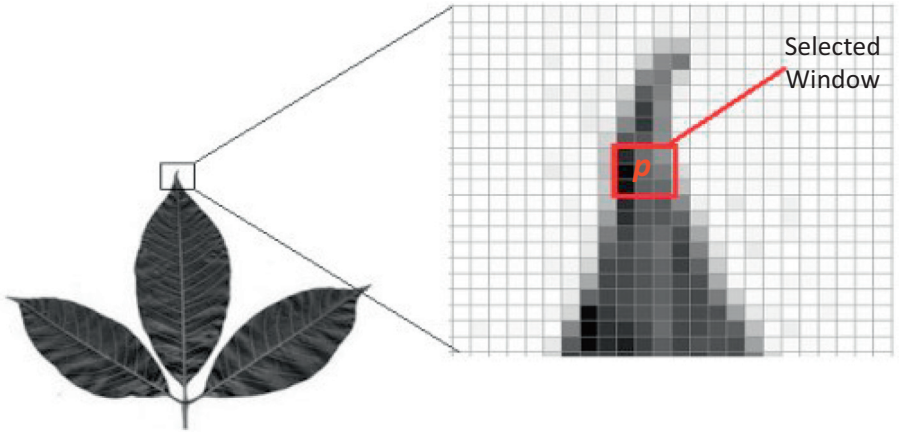

Fig. 5 – Segmented patch demonstrates Harris corner detection. The pixel at p is the center of a candidate corner. Around the red square, pixels are brighter than *p*.

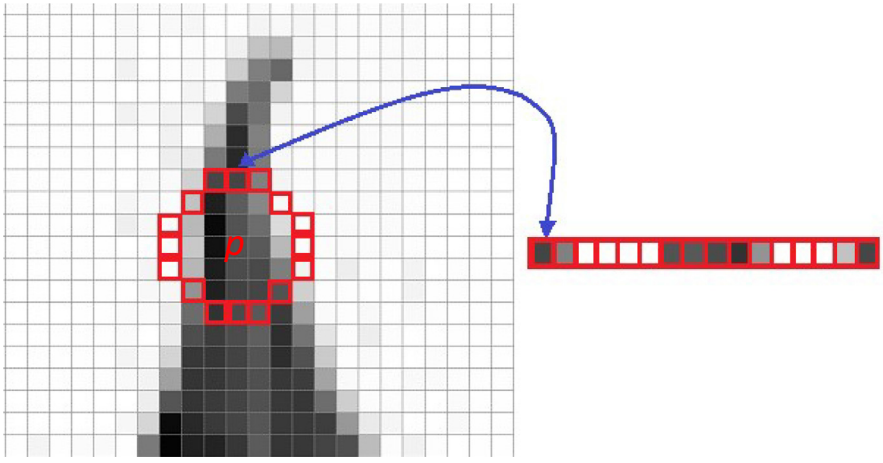

Fig. 6 – The 16 values surrounding pixel p stored in vector form.

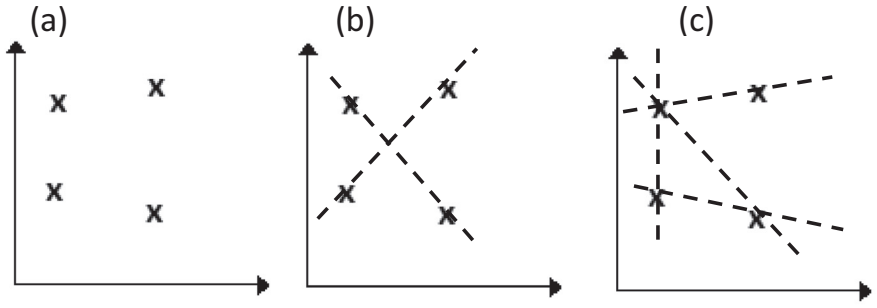

Fig. 7 – (a) White pixels, (b) possible straight lines, (c) other possibilities of straight lines.

5. Feature extraction of hevea leaf based on position

Now that we discussed the shape and angle features of rubber tree leaves, next a comparative study of keypoint feature extraction and angle extraction methods will be presented. Here, we proposed keypoint extraction method to extract shape feature of the rubber tree leaves. This approach facilitates extracting further features such as corner, edge, and blob. The key features of overlapping and non-overlapping leaf assist in identifying similar shapes through comparison, using the nearest neighbor algorithm. This process can be implemented by constructing a trained template consisting of various leaves with different positions. Next, keypoints in the input leaf image are compared with keypoints of the template image to examine the position of leaflets accordingly. It

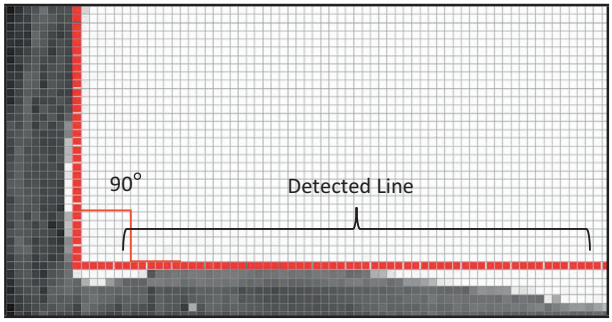

Fig. 8 – Petiole line detection for angle computation.

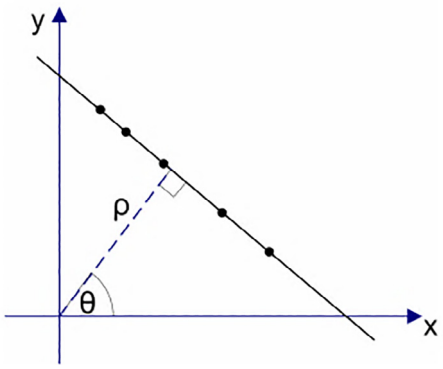

Fig. 10 – Rho (ρ) and theta (θ) representation of a straight line.

is necessary to use a powerful descriptor that is stable to image translations, rotations, and scale changes, to detect critical features throughout the matching process. However, this paper is dedicated to extracting only these critical features. In this research, SIFT is proposed to extract shape features from rubber tree leaf images. However, comparative study of SIFT, Harris, and FAST keypoint extraction methods is conducted to obtain a convincing result. The second part of the feature extraction of rubber tree leaf is the axil detection through comparison of two prior line detection methods: Hough transformation and Boundary-tracing. Both methods were implemented in the region where the petioles of rubber tree leaves create the axil. In the following subsections, we describe these methods starting with keypoint feature extraction methods.

5.1. *SIFT algorithm*

Lowe [31] proposed SIFT to address feature matching challenges that arise due to scaling, rotation, and transformation. Scale invariant technique implies that if the variables are multiplied by a common factor, features of the object remain unvarying. This paper uses SIFT keypoint detection to find the main features of the rubber tree leaves such as occluded region of the leaflets. SIFT algorithm is defined through four main phases, as follows:

Finding Scale-space extreme: This phase involves examinations on all scales and image locations that are implemented by a difference-of-Gaussian (DoG) function. The results of these findings are also considered as potential interest points.

Keypoint localization: Once a potential keypoint has been spotted by comparison of the neighboring pixels, the next stage is to implement a complete fit to the nearby data for location, scale, and ratio of principal curvatures. The 3D quadratic function was used to identify the current location of the keypoint candidates by expanding the scale space function. Hereafter, the process attempts to remove some insignificant candidates from the list of keypoints. Poorly localized keypoints or candidates with low contrast on the edges are eliminated at this stage.

Orientation assignment: Remaining points are assigned to a fitting position based on the image gradient directions. The gradient orientation of sample points around the keypoints is formed from the gradient histogram. The gradient histogram covers the 360° radius orientations. The maximum point in the histogram is identified and any other local points with 80% of the peak value are used to generate a keypoint with that orientation.

Keypoint descriptor: The image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

5.2. *Harris corner detection algorithms*

Harris detection is one of the widely implemented feature extraction algorithms and is primarily used for corner

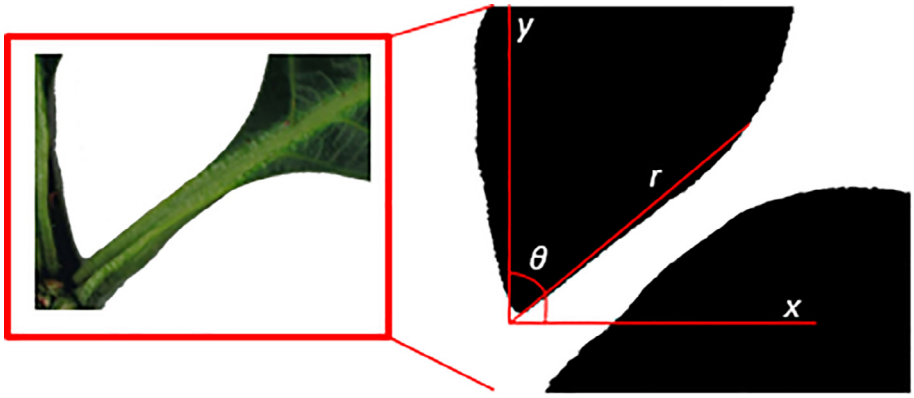

Fig. 9 – Threshold operation on petioles.

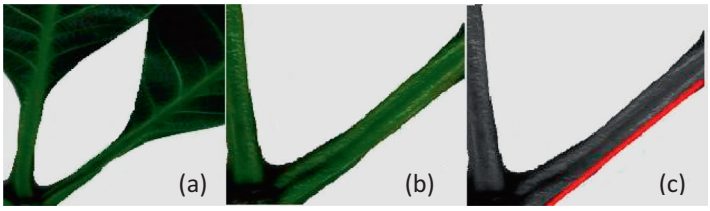

Fig. 11 – (a) Petioles, (b) region of interest, (c) detected line.

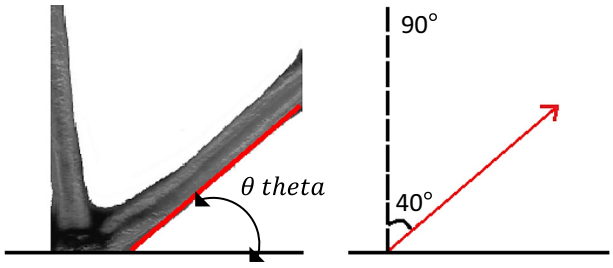

Fig. 12 – Detected line and the calculation of the angle degrees.

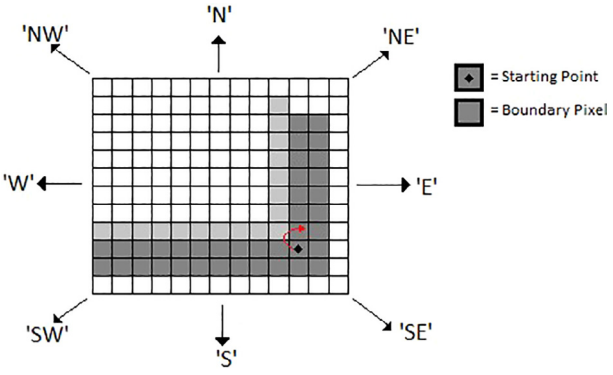

Fig. 13 – Initial search direction for the next pixel.

detection. The central principle of this method is that the intensity changes mostly at the corner in multiple directions. This can be expressed by examining the variations in the intensity of movement in a selected window (see Fig. 5) [32].

This method is inspired by the self-correlation function in signal processing. The Harris corner detector calculates each pixel gradient and if the absolute gradient values in two directions are both greater than a threshold, then the pixel is accepted as a corner.

5.3. *FAST algorithm*

FAST (features from accelerated segment test) is a corner detection method that has been used in many computer vision tasks such as object tracking and mapping [33]. The advantages of FAST corner detection is that the algorithm is computationally inexpensive. The FAST algorithm searches for the corner within a circle of 16 pixels if the point p is a potential corner. Pixels in this circle are assigned a number from 1 to 16 in a clockwise direction. Fig. 6 demonstrates the 16 values surrounding pixel p that are stored in vector form.

This detection method uses the pixel intensities from the 16 pixel circle as a feature vector. These features can then be classified as positive and negative at the pixel indexes depending on whether their intensities are greater or lesser than the center. This classification is convenient because positive features would not be compared to negative ones.

5.4. *Angle extraction using Hough transform*

The next operation is the detecting the petioles as a straight lines using Hough Transformation. The principle of Hough transforms for straight lines is that if there is one white pixel in a binary image, many straight lines can go over from it (see Fig. 7). All the lines can go over from each other in the same image as well [24].

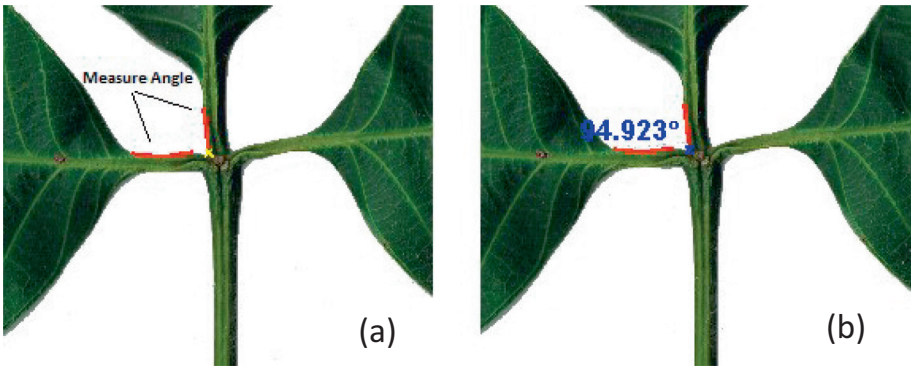

Fig. 14 – (a) Boundary-tracing in region of interest, (b) result of angle computation using dot product.

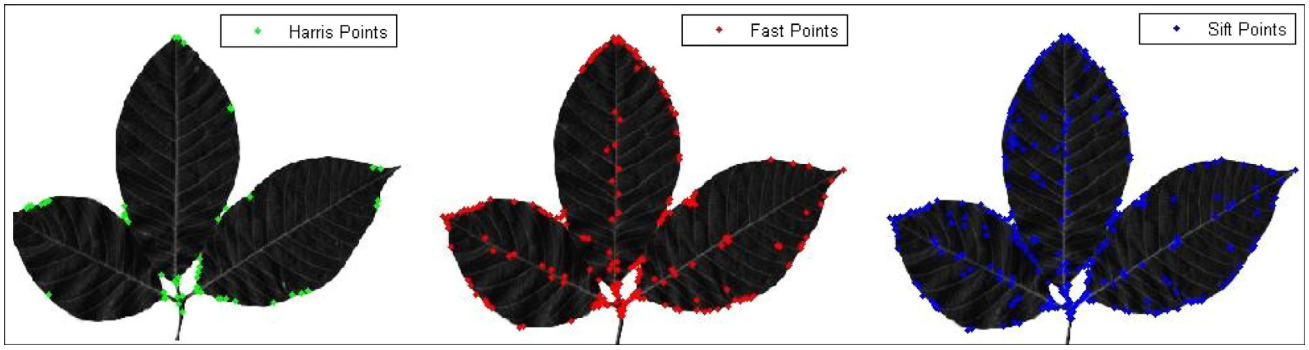

Fig. 15 – Comparison of keypoints detection results for Harris, FAST and SIFT methods.

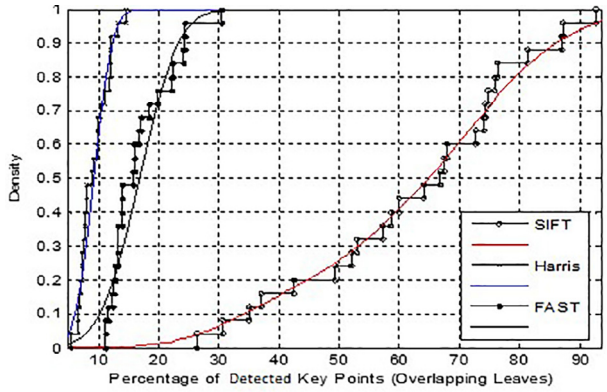

Fig. 16 – Detected keypoint results for overlapping leaves.

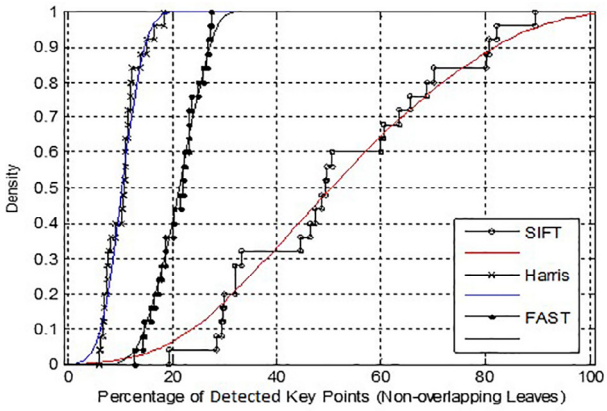

Fig. 17 – Detected keypoint results for non-overlapping leaves.

The region of interest for the angle extraction was set in the image where the petioles intersect as presented in Fig. 8. Leaf petioles might be found in partially curve forms. To compute an accurate angle, the petiole line must be simplified and detected as a straight line starting from an axil point to the leaf base, rather than tracing the petiole edges.

Therefore, the shape-based image threshold has been implemented to minimize the variances [34]. It is defined as the weighted sum of variances of background and foreground:

$$\sigma_{\omega}^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2 \quad (1)$$

Where w_i is the probabilities of the two classes, threshold t where these two classes spread at its minimum, and σ_i^2 is the variances of these two classes. Fig. 9 shows the threshold operation for axil computation.

Straight lines can be detected by setting the mathematical expression as:

$$y = m * x + b \quad (2)$$

where m is the incline and b is where the line intercepts the y-axis. These properties, m and b, can be used to describe the straight line as a single point (m, and b) in the parameter-space extended by two parameters, m and b. However, the problem is that the presentation of the line (m, and b) as a point in the parameter space goes to infinity and the parameter space becomes infinitely large. Therefore, it is required to redesign the expression of the line with some parameters that have limitations. This enhancement is carried out by replacing slope and an intercept with a distance and angle parameters.

The distance ρ (*rho*) is the distance from the origin to the line along a vertical vector and the angle θ (*theta*) is the angle between the x-axis and the ρ vector (Fig. 10). It can be written as:

$$y = \frac{\cos(\theta)}{\sin(\theta)} * x + \frac{\rho}{\sin(\theta)} \quad (3)$$

This expression of ρ can be reorganized as:

$$\rho = x * \cos(\theta) + y * \sin(\theta) \quad (4)$$

The values of ρ and θ are limited by $\theta \in [0, 180]$ in degrees and $\theta \in [0, \pi]$ in radians and $\rho \in [-D, D]$ where D is the diagonal of the image. Thus, the line can be converted into a single point in the parameter space with the parameters θ and ρ. This space is also called Hough space.

The results of the Hough transform is saved in a matrix named as accumulator. One dimension of this matrix is the *theta* value, which is an angle and the other dimension is *rho* value, which is a distance. Each element has a number of points and pixels that lie on the line with the parameters (*rho*, *theta*). Therefore, the element with the highest value shows the line that is most represented in the input image. Fig. 11 displays the result on rubber tree leaf petiole.

Each element of the vector determines the *theta* value for the corresponding column. The acceptable range of theta values is $-90° \leqslant \theta \leqslant 90°$, and the default is −90:89; therefore, the result will be obtained with $90° - \theta$ (Fig. 12).

Table 2 – Result comparisons of axil degrees for Hough Transform (HT), Boundary Tracing (BT) with Ground Truth (GT) and differences from actual value for overlapping and non-overlapping rubber tree leaf images.

Overlapping leaves axil degrees						Non-overlapping leaves axil degrees					
Image ID	GT ± 1	HT	Dif	BT	Dif	Image ID	GT ± 1	HT	Dif	BT	Dif
Oimg1	43°	42°	1°	45°	2°	Nimg1	61°	61°	0°	66°	5°
Oimg2	55°	55°	0°	59°	4°	Nimg2	60°	61°	1°	62°	2°
Oimg3	52°	52°	0°	54°	2°	Nimg3	67°	65°	2°	53°	14°
Oimg4	49°	48°	1°	47°	2°	Nimg4	70°	72°	2°	67°	3°
Oimg5	47°	47°	0°	40°	7°	Nimg5	79°	78°	1°	83°	4°
Oimg6	41°	42°	1°	43°	2°	Nimg6	64°	63°	1°	74°	10°
Oimg7	45°	45°	0°	50°	5°	Nimg7	66°	65°	1°	66°	0°
Oimg8	50°	51°	1°	35°	15°	Nimg8	66°	65°	1°	72°	6°
Oimg9	44°	43°	1°	45°	1°	Nimg 9	53°	51°	2°	65°	12°
Oimg10	45°	45°	0°	34°	11°	Nimg10	82°	80°	2°	90°	8°
Oimg11	48°	48°	0°	40°	8°	Nimg11	62°	62°	0°	58°	4°
Oimg12	45°	44°	1°	51°	6°	Nimg12	61°	60°	1°	72°	12°
Oimg13	50°	51°	1°	41°	9°	Nimg13	57°	57°	0°	63°	6°
Oimg14	41°	42°	1°	47°	9°	Nimg14	58°	58°	0°	66°	8°
Oimg14	52°	50°	2°	49°	3°	Nimg14	59°	57°	2°	66°	7°
Oimg16	50°	50°	0°	43°	7°	Nimg16	67°	67°	0°	56°	11°
Oimg17	46°	47°	1°	37°	9°	Nimg17	88°	89°	1°	78°	10°
Oimg18	54°	54°	0°	42°	12°	Nimg18	70°	72°	2°	79°	9°
Oimg19	47°	48°	1°	43°	4°	Nimg19	58°	58°	0°	84°	26°
Oimg20	46°	46°	0°	48°	2°	Nimg20	69°	68°	1°	67°	2°
Oimg21	42°	41°	1°	59°	17°	Nimg21	85°	86°	1°	76°	9°
Oimg22	41°	41°	0°	47°	6°	Nimg22	85°	85°	0°	70°	15°
Oimg23	45°	43°	2°	35°	10°	Nimg23	64°	64°	0°	83°	19°
Oimg24	46°	47°	1°	58°	12°	Nimg24	77°	77°	0°	78°	1°
Oimg25	51°	52°	1°	47°	4°	Nimg25	65°	64°	1°	78°	14°

Table 3 – Descriptive statistics and t-test results for ground truth, Hough transformation and boundary tracing for overlapping and non-overlapping rubber tree leaf images.

Outcome	t-test			95% CI for mean difference			
	M	SD	R		r	t	df
Overlapping							
Ground Truth1	47	4.02	14				
Ground Truth2±	47	4.12	15				
Ground Truth3±	46.9	4.22	15				
Hough	46.96	4.148	14	−0.346, 0.426	0.97	0.832	24
Boundary	45.56	7.095	25	−1.830, 4.710	0.06	0.372	24
Non-overlapping							
Ground Truth1	67.7	9.7	35				
Ground Truth2±	68.4	9.7	35				
Ground Truth3±	67.6	9.56	35				
Hough	67.4	10.03	39	−0.152, 0.792	0.99	0.175	24
Boundary	70.88	9.34	37	−7.330, 1.010	0.44	0.131	24

5.5. *Angle extraction using Boundary-tracing*

The Boundary-tracing algorithm is the computation of the row and column coordinates of all the pixels on the border of an object. It is initialized by specifying a location of an initial point on the object. It can be defined as nonzero pixels belong to an object and pixels with the value 0 constitute the background [35] (example of boundary tracing can be seen in Fig. 13).

It is required to specify the row and column coordinates of the starting point, and the direction of the first step. Initial search direction of the starting point to the north ('N') is demonstrated in Fig. 12. To obtain the edge of the horizontal petiole, the adjacent pixels are shifted and inspected in the column until the object pixel occurs or until arrival at a background pixel. This procedure will be repeated for the other petiole, but this time, tracing should be vertical.

Boundary-tracing algorithm extracts x, y locations of the boundary points throughout the line. It is important to extract as many points belonging to the edges as possible to obtain more accurate point of intersection and angle calculations.

The number of points should be determined experimentally. Because the initial point of the horizontal line was found by tracing from north to south, it is the safest to set the initial point starting from the north. Fig. 14 demonstrates the result of boundary tracing angle detection.

6. Experimental results

The result of the experiment is discussed under two subsections as keypoint extraction methods and angle detection methods.

6.1. *Comparison of SIFT, Harris and FAST algorithms*

The experiment employs total of fifty overlapping and non-overlapping rubber tree leaves, which consist of five different clones: P350, RRIM 2001, RRIM 2002, RRIM 2025, and RRIM 3001. Fig. 15 demonstrates the comparison of keypoints obtained from SIFT, Harris, and FAST methods. The output of SIFT shows that extracted features usually lie on high-contrast regions of the image, such as edges, vein, and occluded regions of the leaflets. As seen by this figure, the results of Harris and FAST algorithms in terms of number of detected features are limited.

The output of the experiment indicates that using the scale invariant keypoint extraction algorithm is able to detect blobs and ridges features that are critical for matching process. Although Harris and FAST methods are often used for feature extraction, the number of detected keypoints are insufficient for a reliable matching process. An evaluation of the number of detected keypoints has been performed with regard to the leaflet position identification. The statistic results of experiments are shown in Figs. 16 and 17 respectively. This figure shows the percentage of detected keypoints and the density of detected features at the same location.

6.2. *Comparison of Hough and Boundary-trace algorithms*

In this section, we present the comparison of two different angle extraction methods, Hough transformation (HT) and boundary-tracing (BT) methods. Initially, ground truth (GT±) of angle degrees is identified manually using a graphics editing software. However, the angle degrees might be subject to variability since they are selected by a user. Therefore, three distinct ground truth set are specified by different individuals to consider measures of the inter subject variability. Table 2 displays the comparison between one of the selected ground truth results, Hough transformation, and the boundary-tracing results. As revealed in Table 2, Hough Transformation axil degrees remain in the correct range for both overlapping and non-overlapping leaves, while boundary-tracing angle degrees deviated from the range.

To analyze the correlation between variables, a paired samples *t*-test was performed. Some of the frequently used measures of inter subject variability: mean (M), standard deviation (SD), and the range (R) is compared to analyze the variations. Ground truth dataset overlapping ($M = 47$, $SD = 4.02$, $R = 14$), non-overlapping ($M = 67.7$, $SD = 9.7$, $R = 35$), with Hough transformation overlapping ($M = 46.96$, $SD = 4.14$, $R = 14$), non-overlapping ($M = 67.40$, $SD = 10.03$, $R = 39$) and boundary tracing overlapping ($M = 45.56$, $SD = 7.09$, $R = 25$), non-overlapping ($M = 70.88$, $SD = 9.347$, $R = 37$) was compared respectively. There was a significant difference between ground truth and boundary tracing results for the overlapping leaves at the level of significance ($t(24) = 0.909$, $p < .372$) compare to Hough Transformation in the condition of $t(24) = 0.214$, $p < .832$. Same as overlapping, non-overlapping leaf results of boundary tracing method shows statistically higher difference ($t(24) = -1.564$, $p < .131$), while Hough Transformation appears to be closer ($t(24) = 1.398$, $p < .175$). These results show that Hough and boundary tracing methods are different in their performance and the boundary tracing angle values is significantly varied from the ground truth. The summary of the analysis is shown in Table 3.

The probability distribution fitting in the accuracy of the angle detection is also evaluated (see Figs. 18 and 19). This

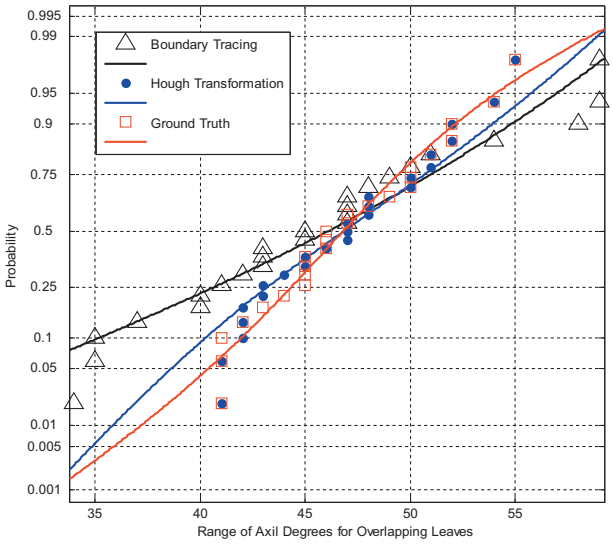

Fig. 18 – Probability distributions analysis for overlapping leaf angle degrees.

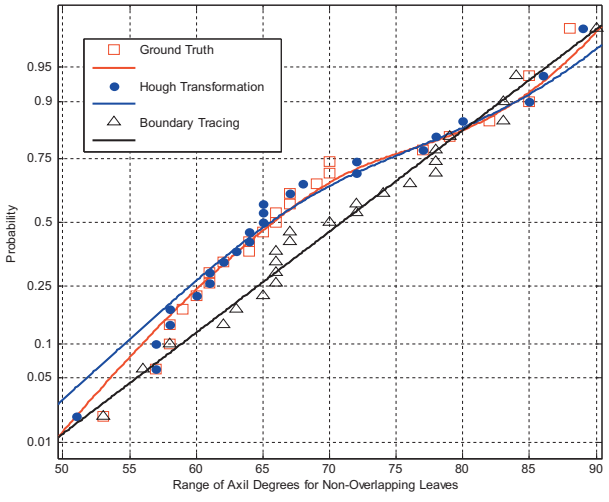

Fig. 19 – Probability distributions analysis for non-overlapping angle degrees.

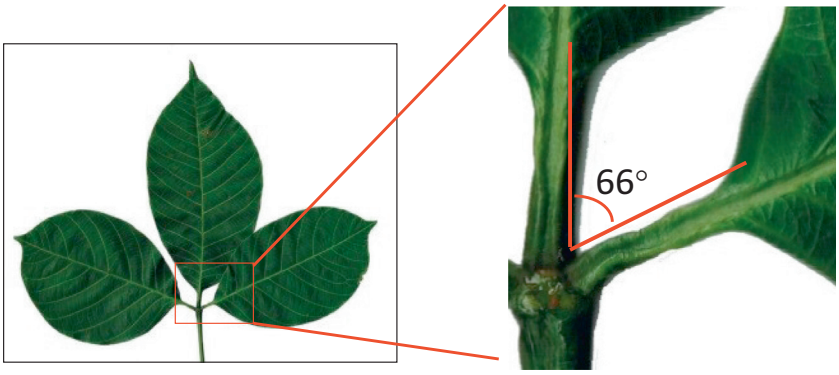

Fig. 20 – Exceptional case of overlapping leaflet with axil degree out of its range.

experiment carried out to identify if the data set is approximately fitted to right range in terms of leaflet position identification. Departures from the red line (ground truth) indicate departures from normality. Also note that the overlap range was 30–55° and the angle degrees of non-overlapping leaves range from 55° to 90°. The figures show that the result of Hough transformation is relatively fitter to the ground truth dataset than the boundary-tracing algorithm. The result of boundary-tracing angle extraction for the overlapping leaves shows values up to 59°, which is an excessively high number for the overlapping leaves. Non-overlapping leaves axil degrees also deviate from the ground truth dataset. The reason for this is that the boundary-tracing algorithm traces the curved shape of the petioles. Therefore, it affects the intersection point and the position of the lines, consequently the computation of the angle degree.

In certain cases, the axil degree of occluded leaflet is above 55° as shown in Fig. 20. The leaf base feature must be considered, such as obtuse leaflet tend to be larger, thus it makes the leaflets overlapping. Identification of these outliers requires further study possibly using the leaf base feature.

7. Discussions

This paper proposed a framework for feature extraction of rubber tree leaves to provide insights into the overlapping leaf recognition. A comparative study of SIFT, Harris and FAST methods for keypoint extraction, Hough Transformation and boundary tracing method for angle computation is carried out to measure their performance. Shape feature is extracted using the difference-of-Gaussian function, which lie throughout high-contrast regions. The results demonstrated that the SIFT descriptor outperforms other methods.

Angle feature is a dynamic attribute of the framework, yet it plays a key role in the leaflet position identification. The challenge of angle computation is that petioles might be found in curvy forms. Therefore, smooth line detection is crucial for the precise angle computation. To compute the angle accurately, the petiole line is simplified using threshold operation. Then, Hough Transform is used to detect petioles as straight lines from the axil point to the leaf base.

8. Conclusions

This paper investigated the performance of keypoint extraction and line detection methods for overlapping and non-overlapping rubber tree leaf features. The results indicate that the keypoint extraction method to obtain leaf shape information helps to extract more features, including edge, corner, blob, and ridge.

In this paper, we used axil feature of rubber tree leaves for leaflet position identification. The result suggests that the values of touching leaflets degrees can be considered under non-overlapping leaves with respect to the angle range. The proposed framework presents the advantage of using Hough transformation for accurate rubber tree angle computation. The method helps to detect the petiole as a straight line; it disregards the partially curved shape of the petiole. Therefore, it allows to determine the angle between petioles accurately. In conclusion, the results demonstrate that the SIFT and Hough transformation methods can successfully extract shape and axil features of overlapping rubber tree leaflets.

Conflict of interest

The authors declare that there is no conflicts of interest.

REFERENCES

[1] Shigematsu A, Mizoue N, Kajisa T, Yoshida S. Importance of rubberwood in wood export of Malaysia and Thailand. New Forest 2011;41(2):179–89.
[2] Ong EL. Characterization of new latex-timber clones of natural rubber. J Appl Polym Sci 2000;78(8):1517–20.
[3] Wei OC. Digital image recognition system for rubber clones produced in Malaysia. In: Irc 2012 international rubber conference; 2012.
[4] Du JX, Wang XF, Zhang GJ. Leaf shape based plant species recognition. Appl Math Comput 2007;185(2):883–93.
[5] Anjomshoae ST, Rahim MS, Javanmardi A. Hevea leaf boundary identification based on morphological transformation and edge detection. Pattern Recognit Image Anal 2015;25(2):291–4.

[6] Hu R, Jia W, Ling H, Huang D. Multiscale distance matrix for fast plant leaf recognition. IEEE Trans Image Process 2012;21 (11):4667–72.
[7] Feng Y, Zhang S. Supervised locally linear embedding for plant leaf image feature extraction. In: International conference on intelligent computing. Springer, Berlin, Heidelberg, 2009 September 16. p. 1–7.
[8] Kumar N, Belhumeur PN, Biswas A, Jacobs DW, Kress WJ, Lopez IC, Soares JV. Leafsnap: a computer vision system for automatic plant species identification. In: Computer vision–ECCV 2012, Springer, Berlin, Heidelberg, 2012, p. 502–16.
[9] Wang Z, Chi Z, Feng D, Wang Q. Leaf image retrieval with shape features. In: International conference on advances in visual information systems 2000 November 2. Springer, Berlin, Heidelberg; 2000.
[10] Lee KB, Chung KW, Hong KS. An implementation of leaf recognition system based on leaf contour and centroid for plant classification. In: Ubiquitous information technologies and applications 2013. Springer, Dordrecht; 2013.
[11] Caballero C, Aranda MC. WAPSI: Web Application for Plant Species Identification using fuzzy image retrieval. In: International conference on information processing and management of uncertainty in knowledge-based systems. Springer, Berlin, Heidelberg, 2012 July 9, p. 250–9.
[12] Zheng XD, Wang XJ. Feature extraction of plant leaf based on visual consistency. In: International symposium on Computer Network and Multimedia Technology. CNMT 2009, IEEE, 2009 January 18. p. 1–4.
[13] Bhardwaj A, Kaur M, Kumar A. Recognition of plants by leaf image using moment invariant and texture analysis. Int J Innovat Appl Stud 2013;3(1):237–48.
[14] Wu SG, Bao FS, Xu EY, Wang YX, Chang YF, Xiang QL. A leaf recognition algorithm for plant classification using probabilistic neural network. In: 2007 IEEE international symposium on signal processing and information technology, 2007 December 15.
[15] Kadir A, Nugroho LE, Susanto A, Santosa PI. Leaf classification using shape, color, and texture features. arXiv preprint arXiv:1401.4447, 2013 November 20.
[16] Sirmacek B, Unsalan C. Urban-area and building detection using SIFT keypoints and graph theory. IEEE Trans Geosci Remote Sens 2009;47(4):1156–67.
[17] Castle RO, Murray DW. Keyframe-based recognition and localization during video-rate parallel tracking and mapping. Image Vis Comput 2011;29(8):524–32.
[18] Piccinini P, Prati A, Cucchiara R. Real-time object detection and localization with SIFT-based clustering. Image Vis Comput 2012;30(8):573–87.
[19] Geng C, Jiang X. Face recognition using sift features. In: 2009 16th IEEE International Conference on Image Processing (ICIP), IEEE; 2009. p. 3313–6.
[20] Pan X, Lyu S. Detecting image region duplication using SIFT features. In: 2010 IEEE international conference on acoustics speech and signal processing (ICASSP), 2010 March 14. IEEE; p. 1706–9.
[21] Mehrotra H, Majhi B, Gupta P. Robust iris indexing scheme using geometric hashing of SIFT keypoints. J Network Computer Applicat 2010;33(3):300–13.
[22] Tao C, Tan Y, Cai H, Tian J. Airport detection from large IKONOS images using clustered SIFT keypoints and region information. IEEE Geosci Remote Sens Lett 2011;8(1):128–32.
[23] Zahedi M, Eslami S. Farsi/Arabic optical font recognition using SIFT features. Procedia Comput Sci 2011;3:1055–9.
[24] Duda RO, Hart PE. Use of the Hough transformation to detect lines and curves in pictures. Commun ACM 1972;15(1):11–5.
[25] Stylianidis E, Patias P. Using Hough transform in Line extraction. Int Arch Photogram Remote Sens 33:119–24.
[26] Weiss JM. Real-time feature detection using the Hough transform. In: CAINE 2008. p. 168–73.
[27] Chakraborty B, Gonzalez J, Roca FX. Large scale continuous visual event recognition using max-margin Hough transformation framework. Comput Vis Image Underst 2013;117(10):1356–68.
[28] Cui S, Yan Q, Reinartz P. Complex building description and extraction based on Hough transformation and cycle detection. Remote Sens Lett 2012;3(2):151–9.
[29] Martins MB, Zieri R. Leaf anatomy of rubber-tree clones. Scientia Agricola 2003;60(4):709–13.
[30] Anjomshoae ST, Rahim MS. Enhancement of template-based method for overlapping rubber tree leaf identification. Comput Electron Agric 2016;1(122):176–84.
[31] Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vision 2004 Nov 1;60(2):91–110.
[32] Dawood M, Cappelle C, El Najjar ME, Khalil M, Pomorski D. Harris, SIFT and SURF features comparison for vehicle localization based on virtual 3D model and camera. In: 2012 3rd International conference on Image Processing Theory, Tools and Applications (IPTA), 2012 Oct 15. IEEE; p. 307–12.
[33] Rosten E, Porter R, Drummond T. Faster and better: a machine learning approach to corner detection. IEEE Trans Pattern Anal Mach Intell 2010;32(1):105–19.
[34] Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybernet 1979;9(1):62–6.
[35] Gonzalez Rafael C, Woods Richard E, Eddins Steven L. Digital image processing using MATLAB. USA: Editorial Pearson-Prentice Hall; 2004.