Normalized Cross Correlation Template Matching for Oil Palm Tree Counting from UAV image

Sharifah Nurul Husna Syed Hanapi\textsuperscript{1,a}, S A A Shukor\textsuperscript{1,b}, Jalal Johari\textsuperscript{2}

\textsuperscript{1}Faculty of Electrical Engineering Technology, Kampus Alam Pauh Putra, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia
\textsuperscript{2}Jalal Johari Consultants Sdn. Bhd., UG-13A, LEXA Galleria, No. 45, Jalan 34/26, Wangsa Maju, 53300 Kuala Lumpur, Malaysia

Email: \textsuperscript{a}nurulhusna@unimap.edu.my, \textsuperscript{b}shazmin@unimap.edu.my

Abstract. Tree crown detection and counting from remote sensing data such as images from Unmanned Aerial Vehicle (UAV) shows significant role in this modern era for vegetation monitoring. Since the data processing would depend on raw data available and for this case the RGB data, thus a suitable method such as template matching is presented. Normalized cross correlation is widely used as an effective measure in similarity between template image and the source or testing images. This paper focuses on six (6) steps involved in the overall process which are: (1) image acquisition, (2) template optimisation, (3) normalized cross correlation, (4) sliding window, (5) matched image and counting, and (6) accuracy assessment. Normalized cross correlation and sliding window techniques proposed for this work resulted in 80% to 89% F-measure values. This result indicates that UAV image data with appropriate image processing method/s have the potential to provide vital information for oil palm tree counting. This would be beneficial in plantation management to estimate yield and productivity. However, there are still rooms for improvement to achieve better results.

1. Introduction

Tree counting studies are not only important in forestry but also in plantation sector especially in oil palm plantation. This study is very important for plantation managers or farmers in order to manage and estimate their oil palm production. In large plantation area, the cost of fertilizers, pesticide and labours, among other things, are high. Thus, by getting the best approximate number of trees in the plantation area, all those costs can be minimized and reduced. Other than that, this study is also important for yield estimation. By knowing the number of trees, the number of fruits produced can be estimated by multiplying the number of trees with the average number of fruits per tree.

In terms of remote sensing data, satellite imagery with multispectral or hyperspectral band would give more advantages as there are more information can be processed compared to normal UAV image data which only provide RGB image. Multispectral data has more spectrum input which is applicable for many more classifier input instead of only using the red, green and blue spectrums as input. This gives more context to the input images especially as the infrared and near-infrared channels has shown strong correlation to the presence of vegetation [1]. Thus, for UAV image which only provide RGB data, the task of tree detection and counting would be more challenging, as methods like vegetation indices of normalized difference vegetation index (NDVI) and soil adjusted vegetation index (SAVI)
cannot be conducted. Thus, a more suitable method such as template matching is proposed that would give comparable results as multispectral or hyperspectral images algorithm.

For oil palm tree detection and counting, there are various conditions that will affect the process such as the density of grass and other weeds. Problem also arises for the overlapping leaf. This situation usually occurred in the mature oil palm tree plantations where the gaps between trees are narrower [2].

Template matching is a high-level image processing method which has been used over the last few decades and is still being used until now due to its advantages. Template matching is the process of comparing the template image to the testing image in order to find the best matched location [3]. This method usually applies for identifying characters, numbers, animals, pedestrian and other objects. In manufacturing, template matching is very useful in parts detection as part of the quality control process [4]. There are many techniques that can be used in template matching such as sum of absolute difference, sum of squared difference, cross correlation and normalized cross correlation. Normalized cross correlation is said to be more sensitive to revolution in addition to scale changes. That is why this technique is more reliable compared to other conventional cross correlation technique [3].

This paper would present the experimental process of tree detection and counting algorithm using standardized normalized cross correlation technique of template matching. The rest of this paper is prepared as follows: in Section 2, a few related studies are reviewed while in Section 3 the methodology is presented. The experimental results and discussion of the proposed method is discussed in Section 4 and finally, Section 5 concludes the finding of the overall process.

2. Related Studies
Since the 1980’s, there were few studies that use the concept of synthetic model for template matching algorithm. For example, [5] did a template matching algorithm involving radiometric and geometric technique to develop a synthetic model for different tree crown shape and size. While in [6], tree crown ‘circular model’ was developed using convex grey-level curves and edge contours at multiple scales for estimating each tree crown area.

The use of template matching and segmentation of RGB satellite imagery was done by [7] for counting purpose. The technique used was generating a masking function and scan the template from top to bottom for tree detection process. Other study used object’s boundary as a criterion for template matching algorithm. However, the boundary technique could be misleading due to image occlusion and distortion [8].

A circular autocorrelation of polar model shape matrix (CAPS) was proposed in [9] as shown in Figure 1. The detection was done by generating shape feature represented by CAPS curve vector. The labeled generic shape features extracted from the training set was trained using SVM and the best decision boundary between positive and negative training set was determined. Another template matching technique which used nearest neighbour (NN) with efficient sliding window computation was presented in [10]. This algorithm can deal with changes in appearance, illumination, non-rigid transformations, viewpoint, occlusion and illumination. However, the method is relatively slow due to the computational process of NN and sliding window.

One of the latest studies related to UAV image data was done by [11] which applied entropy, variance of images and NDVI for counting apple in an orchard. Trees were detected using sliding window with specified NDVI, entropy and variance threshold values which give an accuracy of higher than 93%. However, the method still could not solve the problem of detecting small trees and trees that are planted too close with each other.
3. Methodology

The general block diagram that describes the operation of the proposed template matching operation is shown in Figure 2.

![Block diagram of the overall process](image)

Figure 2. Block diagram of the overall process

3.1 Data Acquisition and Image Pre-processing

Images were taken with Quadcopter DJI Phantom 4 as shown in Figure 3 which is equipped with 14.4 Megapixel camera. This small lightweight UAV can take images at 25 minutes maximum flight time and store them in microSD. The UAV was flown over the Kuala Ketil oil palm plantation in Kedah for data collection which covers about 5 hectares plantation area. Different data sets were taken at different site within the study area with flight altitude of 250 m height.

![Quadcopter DJI Phantom 4](image)

Figure 3. Quadcopter DJI Phantom 4
3.2 Template optimization

The acquired images were then processed for the detection of tree crown. First, the best template from the images is chosen and undergone through the conventional image processing step which are image filtering, image binarization, morphological operation and feature extraction. These processes were done in order to extract the features of oil palm tree crown. Figure 4 below shows the template optimization processes.

![Image Processing for Template Image](image)

**Figure 4.** Image processing for the template image: (a) original chosen image, (b) noise removal and filtering, (c) contrast enhancement, (d) binarized image, (e) erosion, (f) feature extraction

3.3 Normalized Cross Correlation

Generally, normalized cross correlation follows the following formula:

$$p(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}] [f(x - u, y - v) - \bar{f}]}{\left(\sum_{x,y} [f(x, y) - \bar{f}]^2 \sum_{x,y} [f(x - u, y - v) - \bar{f}]^2\right)^{0.5}}$$  \hspace{1cm} (1)

where,

- $f$ is the image.
- $\bar{t}$ is the mean of the template
- $\bar{f}_{u,v}$ is the mean of $f(x,y)$ in the region under the template.

The formula indicates the calculation of getting local sums by precomputing running sums which will be used to normalize the cross correlation to get correlation coefficients. The process of finding the best image matched was done by comparing the template image with the testing image which are both in grayscale image as shown in Figure 5. The maximum value where the template image has correspondence (pixel by pixel) value to the testing image, located at $(x,y)$ shows the similarity matched between these two images [3].
3.4 Sliding window

Sliding window technique is the process of determining object of interest by sliding a fixed size rectangular region across an image. The size of the sliding window should be optimized to detect enough number of feature points in the window. Too small window size would underestimate the detection of object while too large window size would affect the concentration of the image [4]. Thus, in this experiment, the chosen size is based on the size of object of interest (tree crown) which is 50x50 window size. Then, the thresholding values of normalized cross correlation obtained is further used for the sliding process.

4. Results and Discussion

The algorithm was successfully tested to 4 (A, B, C and D) testing image at different location of Kuala Ketil’s plantation area. First of all, the RGB image of template and testing image were resized to 90x 90 pixel and 500x500 pixel respectively. The images were then converted to grayscale image for the normalized cross correlation process.

The correlation image was represented with bright pixel while the dark pixels showed low correlation values. The point of the correlation obtained for the corresponding testing image was plotted using surface plot. This surface plot is used in determining the threshold values for the sliding window process. From these 4 images, the values of thresholding range between 0.22 – 0.26. Then, the sliding window process took place where the matched images were automatically counted in the algorithm. The correlation value which is greater than the threshold value selected were converted into vector point representation on testing image for counting purposes. The results in Figure 6 and Figure 7 show the counted tree crown were marked with rectangular line which is the same size as the window size.

---

**Figure 5.** Grayscale of template and testing image
For the validation process of the oil palm tree counting, the result obtained was compared to the ground truth. The accuracy assessment is based on the following formula and it has been summarised in Table 1.

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{2}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{3}
\]

\[
\text{F-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}
\]

where TP (true positive) is correctly detected tree crown, FP (false positive) is incorrectly detected as tree and FN (false negative) is tree crown that is not detected [12].
Table 1. Accuracy percentage for the counted tree crown

| Dataset | Number of detected tree crown | True value | FP | FN | Precision | Recall | F-measure |
|---------|-------------------------------|------------|----|----|-----------|--------|-----------|
| A       | 63                            | 63         | 10 | 12 | 0.84      | 0.82   | 0.83      |
| B       | 56                            | 68         | 4  | 12 | 0.93      | 0.81   | 0.87      |
| C       | 70                            | 57         | 12 | 2  | 0.83      | 0.97   | 0.89      |
| D       | 59                            | 61         | 13 | 10 | 0.78      | 0.82   | 0.80      |

The overall performance shows that the F-measure is between 80% to 89% accuracy. The FP value is due to the overlapping tree crown and grass or bushes which are detected as tree. For the FN value, it is due to the small oil palm tree which are smaller in size compared to the other old trees. The darker image and brighter image also affect the detection where FN values are more prone to darker image while FP values are prone to brighter image.

5. Conclusion

The research on tree detection and counting has contribute to the enhancement of plantation and inventory management. The use of remote sensing such as UAV also give great impact to the agricultural sector especially in oil palm plantation. This research work demonstrates one of the efficient ways of counting trees for the benefit of farmers and managers as compared to the traditional way of hard labour. It can be said that the results obtained is acceptable and still can be improved with the integration of template matching and machine learning algorithm. The accuracy also can be improved by using feature extracted values of the template image during the sliding window process.

References

[1] Srestasathiern P, Rakwatin P 2014 Oil palm tree detection with high resolution multi-spectral satellite imagery Remote Sensing 6(10) 9749-74
[2] Shafri HZ, Hamdan N, Saripan MI 2011 Semi-automatic detection and counting of oil palm trees from high spatial resolution airborne imagery International journal of remote sensing 32(8) 2095-115.
[3] Swaroop P, Sharma N 2016 An overview of various template matching methodologies in image processing International Journal of Computer Applications 153(10) 8-14
[4] Tsai DM, Lin CT 2003 Fast normalized cross correlation for defect detection Pattern Recognition Letters 24(15) 2625-31
[5] Pollock R 1996 The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown image model (Doctoral dissertation, University of British Columbia)
[6] Brandtberg T, Walter F. 1998 Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis Machine Vision and Applications 11(2) 64-73
[7] Vibha L, Shenoy PD, Venugopal KR, Patnaik LM 2009 Robust technique for segmentation and counting of trees from remotely sensed data IEEE International Advance Computing Conference 2009 pp 1437-1442
[8] Aljarrah IA, Ghorab AS, Khater IM 2012 Object recognition system using template matching based on signature and principal component analysis International Journal of Digital Information and Wireless Communications (IJDWIC) 2(2) 156-63
[9] Manandhar A, Hoegner L, Stilla U 2016 Palm tree detection using circular autocorrelation of polar shape matrix ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences 3 465-72
[10] Talker L, Moses Y, Shimshoni I 2018 Efficient sliding window computation for NN-based template matching In Proceedings of the European Conference on Computer Vision (ECCV) 2018 pp 404-418
[11] Haddadi A, Leblon B, Patterson G 2020 Detecting and Counting Orchard Trees on Unmanned Aerial Vehicle (UAV)-Based Images Using Entropy and NDVI Features The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 2020 43 1211-5

[12] Kalantar B, Idrees MO, Mansor S, Halin AA 2017 Smart counting–oil palm tree inventory with UAV Coordinates 13(5) 17-22.