DRONE BASED IMAGE PROCESSING FOR PRECISION AGRICULTURE

by

MUHAMMAD ARIF SYAFIQ BIN MD SHARIF

Thesis submitted in fulfilment of the requirements for the Bachelor degree of Engineering (Honours) (Aerospace Engineering)

June 2019

ENDORSEMENT

I, Muhammad Arif Syafiq Bin Md Sharif hereby declare that all corrections and comments made by the supervisor and examiner have been taken consideration and rectified accordingly.

(Signature of Student)

Date:

(Signature of Supervisor)

Name: DR. MOHAMMAD NISHAT AKHTAR

Date:

(Signature of Examiner)

Name: ASSOCIATE PROFESSOR DR. ELMI BIN ABU BAKAR

Date:

DECLARATION

This thesis is the result of my own investigation, except where otherwise stated and has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any other degree.

(Signature of Student)

Date:

ACKNOWLEDGMENT

I am using this opportunity to express my deepest gratitude to my thesis advisor and final year project supervisor, Dr Nishat Akhtar. With his patience guidance throughout the year has led me to the completion of this final year project. His vast knowledge in the field of image processing and thought-provoking suggestions have guided me towards success completion of this project. His advices and comment are valuable and useful which help in my thesis writing. He was patient in monitoring me in order to ensure that I am on the right track along the way of completion of this project.

Thank you to the examiner, Dr Elmi Abu Bakar, for the involvement in my final year project and thesis assessment.

Moreover, I would like to show my gratitude to the Dr Ho Hann Woei for his help regarding the final year project. Without him I would not be able to complete my final year project as he lend and help me to fly the drone to take the images for the project. I would like to thank the final year project coordinator Ir Dr Nurulasikin Binti Mohd Suhadis, Dr Aslina Anjang Binti Ab Rahman, Ir Dr Parvathy Rajendran, and other lectures who have given the talk and guideline for the final year project and also thesis writing.

Finally yet importantly, I would like to thank both my beloved parents, MD Sharif Embi and Siti Hajar Hashim who have always support me throughout the completion of the project. Their love, steady encouragement and support have helped me keep a positive attitude in completing this project. I would like to dedicate this thesis to my parents and my family.

DRONE BASED IMAGE PROCESSING FOR PRECISION AGRICULTURE

ABSTRACT

In today's world, with an advent of technological advancements, the use of automated monitoring in agriculture is gaining increase in demand. In the agricultural field, yield loss occurs primarily due to widespread disease. Most of the disease is detected and identified when the disease progresses to a severe stage. A specially equipped UAV can perform several important tasks in agriculture, including monitoring the agriculture land and perform disease detection for several plants at an early stage. Currently, disease traits in agriculture are visually assessed, which can be time-consuming, less accurate and more subjective. Hence, in this project, image processing is used for the detection of plant disease. Detection of plant disease using automated image processing method is beneficial as it can reduce huge work of monitoring in big farms comprising of numerous crops. Moreover, in order to monitor big farms it is a viable option to use unmanned aerial vehicle on specific drone (UAV) to take the snap shots of various diseased plants from multiple angles. This study proposes a parallel image segmentation algorithm in order to detect the diseased leaf in Coconut, Palm, Banana, Dwarf Palmetto and Sapodilla plants acquire using Parrot PF728000 Anafi Drone with 4K HDR Camera. At first, the parallel K-means clustering algorithm was applied on the acquired image to segregate various components acquired using UAV. Post K-means clustering, the diseased portions of the plants were assessed using Hue-Saturation-Value (HSV) based image segmentation algorithm. Moreover, a comparison for image segmentation was also done on non-K-means clustered image and K-means clustered image for which a difference of

$18394mm^2$, $80931mm^2$, $43361mm^2$, $16293mm^2$ and $77542mm^2$ corresponding to Coconut, Banana, Dwarf Palmetto, Palm and Sapodilla was obtained sequentially. The outcome of this project reflects that high-throughput phenotyping techniques will potentially improve the throughput and objectivity of detecting healthier plants and other crops, and will subsequently contribute to the development of new cultivars in breeding programs and yield estimation in precision agriculture.

DRON BERASASKAN PEMPROSESAN IMEJ UNTUK KETEPATAN PERTANIAN

ABSTRAK

Dalam dunia hari ini, dengan kepesatan kemajuan teknologi, penggunaan pemantauan secara automatik dalam bidang pertanian semakin meningkat dalam permintaan. Bidang pertanian, kehilangan hasil berlaku terutamanya disebabkan oleh penyakit yang merebak. Kebanyakan penyakit dikesan dan dikenalpasti apabila penyakit itu berlanjutan ke tahap yang serius. UAV yang dilengkapi kamera boleh melakukan beberapa tugas penting dalam bidang pertanian, termasuk memantau keadaan tanah pertanian dan melakukan pengesanan penyakit tumbuhan pada peringkat awal. Pada masa ini, ciri-ciri penyakit dalam pertanian dinilai secara visual, yang boleh memakan masa, kurang tepat dan lebih subjektif. Oleh itu dalam projek ini, pemprosesan imej dijalankan untuk pengesanan penyakit tumbuhan. Pengesanan penyakit tumbuhan menggunakan kaedah pemprosesan imej automatik bermanfaat kerana ia dapat mengurangkan kerja pemantauan besar di ladang-ladang besar yang terdiri daripada banyak tanaman. Lebih-lebih lagi, untuk memantau ladang-ladang besar adalah satu pilihan yang boleh digunakan untuk menggunakan kenderaan udara tanpa pemandu (UAV) untuk mengambil tangkapan imej dari pelbagai tanaman berpenyakit dari pelbagai sudut. Kajian ini mencadangkan satu algoritma segmentasi imej selari untuk mengesan daun berpenyakit dalam Tanaman Kelapa, Kelapa sawit, Pisang, ‘Dwarf Palmetto’ dan ‘Sapodilla’ memperoleh menggunakan dron Parrot PF728000 Anafi dengan 4K HDR kamera. Pada mulanya, algoritma K-means yang selari digunakan pada imej yang diambil untuk memisahkan

pelbagai komponen yang diperoleh menggunakan UAV. Selepas K-means berkelompok, bahagian-bahagian yang berpenyakit tumbuhan telah dinilai menggunakan algoritma segmentasi imej berasaskan Hue-Saturation-Value (HSV). Selain itu, perbandingan untuk pemisahan imej juga dilakukan pada imej bukan berkelompok dan imej berkelompok K-means yang mana perbezaannya adalah 18394mm^2, 80931mm^2, 43361mm^2, 16293mm^2 dan 77542mm^2 sepadan dengan Kelapa, Pisang, ‘Dwarf Palmetto’, Kelapa Sawit dan ‘Sapodilla’ diperoleh secara berurutan. Hasil daripada projek ini penggunaan teknik ‘phenotyping’ tinggi akan berpotensi meningkatkan keupayaan serta mampu mengesan tumbuh-tumbuhan yang lebih sihat dan tanaman lain, dan akhirnya akan menyumbang kepada pembangunan kultivar baru dalam program pembiakan dan estimasi hasil pertanian ketepatan.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

UAV	:	Unmanned Aerial Vehicle
HDR	:	High Dynamic Range
RGB	:	Red-Green-Blue
HSV	:	Hue Saturation Value
HSI	:	Hue Saturation Intensity
SVM	:	Support Vector Machine
ANN	:	Artificial Neural Network
2D	:	2-Dimensional
GUI	:	Graphical User Interface
OpenCV	:	Open Source Computer Vision
CMOS	:	Complementary Metal-Oxide Semiconductor
ASPH	:	Aspherical Lens
JPG/JPEG	:	Joint Photographic Experts Group
Ms	:	Milliseconds
GSD	:	Ground Sampling Distance
Mbps	:	Megabits per second
YCbCr	:	Luminance, Chromaticity blue and Chromaticity red
NDVI	:	Normalized Difference Vegetation Index

LIST OF SYMBOLS

P_i	:	Probability function
f_i	:	Number of pixels having grey level ‘i’
N	:	Number of pixels in the images
$w(t)$	:	Class probability
$\hat{X}$	:	Mean centred image
$N(x)$	:	Current neighbouring pixel of point x
δ	:	Difference between pixels
$l(x)$	:	Pixel value of point x
i	:	Index of the cluster
t_p	:	Threshold point
C	:	Cluster
a	:	Amplitude
Q	:	Quality factor
t	:	Image without noise
s	:	Filtered image
$\sigma_{s,t}$	:	Covariance between 2 images
σ_s^2	:	Variance of filtered image
σ_t^2	:	Variance of source image without noise
$d(x)$	:	Distance from center of cluster
$\phi(x)$	:	Input leaf
$K(x,y)$	:	Kernel function

CHAPTER 1

INTRODUCTION

1.1 Research background

Unmanned Aerial Vehicle (UAV), popularly referred to as Drone, is an aircraft operated remotely by a human operator or an on board computer autonomously. UAV remote sensing is a new means of remote sensing with high image resolution, fast, flexible, low cost, etc. (Wang et al., 2013). Satellite imagery and controlled aircraft data could compensate for high costs, high weather impacts, long cycles of work, and so on which is why we replace with the UAV as the cost will be reduce and can reach the work area that satellite and controlled aircraft cannot reach.

UAV has been widely used in numerous areas such as agriculture, forestry, soil resources, etc. It plays an important role in the acquisition of information on agricultural production and the monitoring of agricultural conditions, in particular in agriculture. UAV is widely used in rice lodging monitoring, monitoring of pests and diseases, monitoring crop growth and diagnosing crop nutrients, which has achieved an ideal application effect (Jia, Su, Shen, Yuan, & Xu, 2016).

Agriculture is the backbone of human sustenance on this world. A decade or two back, it was associated solely with the production of basic food crops. Agricultural productivity is one of the most important primary sector on which economy highly depends and turn out to be significantly more than simply a means to feed ever-growing populations. Nowadays with growing population we need the productivity of the agriculture to be increased a lot to meet the demands increasing population means that there has to be an increased focus to the primary sector.

Plants plays an important role in providing the energy to the human being and essential source to the issue of global warming. Agriculture is the foundation of economy in India and about 50% of the population is involved in farming activities directly or indirectly as mentioned by (Barbedo, 2013). It is imperative where in over 70% population relies upon horticulture in India (Varshney & Dalal, 2016). That means it feeds great number of people.

Therefore, plant diseases have turned into dilemma because it may cause significant reduction in quantity and quality of the agricultural products. This will cause a serious problem on the countries whose economics are primarily dependent on the agriculture. So that, detection of diseases in plant is important in order to prevent negative impact. Hence, detection of diseases from the initial stages is crucial to the production of agriculture.

The major diseases of plants are viral, fungus and bacterial disease like *Alternaria*, *Antharacnose*, bacterial spot, canker, etc.,. These are further classified in Figure 1.1 below. The environmental changes will cause the viral disease, fungus disease is due to essence of fungus in leaf and bacterial disease is causes by presence of germs in leaf or plants as mentioned by (Prakash, Saraswathy, Ramalakshmi, Mangaleswari, & Kaviya, 2017).

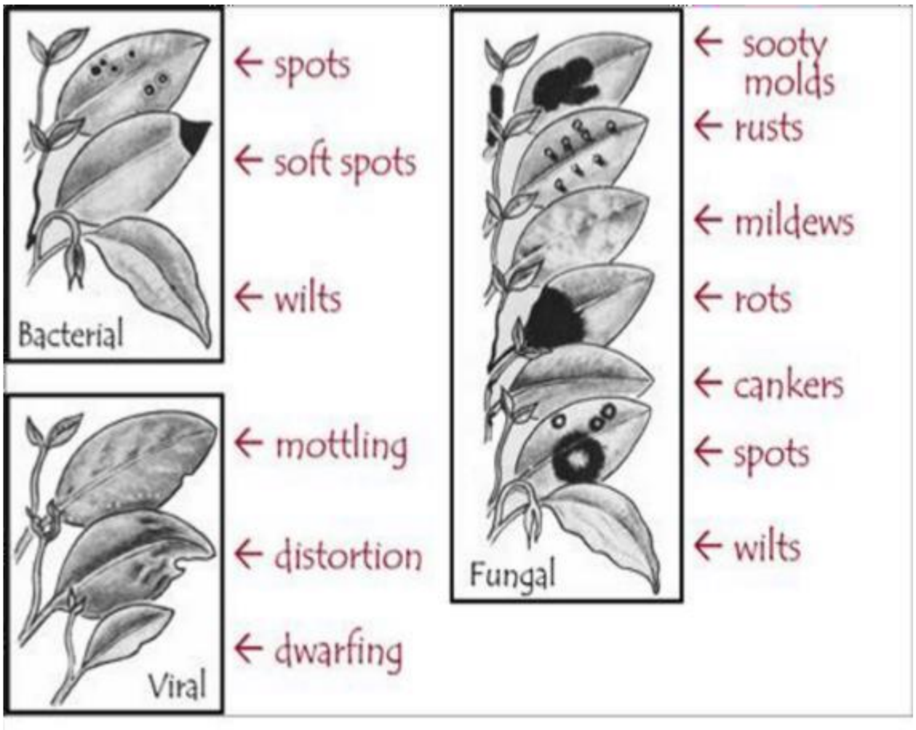

Figure 1.1 Type of diseases (Mainkar, Ghorpade, & Adawadkar, 2015)

For example an infection named little leaf infection is a risky disease found in pine trees in United States. The influenced tree has a stunted growth and dies within 6 years. Its effect is found in Alabama, Georgia parts of Southern US. In such situations early detection could have been save the tree (Dhygude & Kumbhar, 2013).

In plants, disease can be found in different parts such as leaves, fruits and stems. The plant leaf for the detection of disease is viewed as which demonstrates the disease symptoms. This research provides the prologue to image processing technique used for leaf disease detection.

1.2 Problem Statement

So far the existing method for diseased leaf detection is carried out by experts using naked eye observation after which the further identification and analysis is done. In order to carry out this process, a large team of experts as well as continuous monitoring of plant is required, which costs very high when it is implemented in large farms.

At the same time, in many developing and under developing countries, farmers do not have adequate facilities to contact experts for diseased leaf detection and analysis as a result of which the consulting experts even cost high as well as time consuming too.

For such scenario, the suggested automated technique proves to be beneficial in monitoring large fields of crops using UAV. Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. This also supports machine vision to provide image based automatic process control, inspection, and robot guidance.

Plant disease identification by visual way is more laborious task and at the same time, less accurate and can be done only in limited areas. Whereas if automatic detection technique is used, it will minimize the time and effort and will increase the accuracy. Generalized diseases in many plants comprises of brown and yellow spots, early and late scorch, and others are fungal, viral and bacterial diseases.

For input data disease which will be acquired using Parrot PF728000 Anafi Drone with 4K HDR Camera, samples of plant leave like Palm with leaf blight and lesions, Dwarf Palmetto (palm species) with leaf blight and scorching, Sapodilla leaf with bacterial and fungal disease, Banana leaf with early scorch disease and Coconut leaf with leaf scorching. The initial infections on lower leaves develop distinct brown to blackish lesions or spots with yellow halos around them.

1.3 Research objective

A research study designed to implement and assess the performance of the proposed technique involves the following research objectives:

I. To design parallel K-means algorithm for various tree leaf sample acquired using drone.

II. To implement Hue-Saturation-Value (HSV) image segmentation on acquired sample post K-means.

1.4 Research scope

This work are to recognize the sicknesses that conventionally occur in the plants leaf by the properties of the leaves like shading and its tendency. We proposes a framework which can give progressively precise outcomes identified with the distinguishing proof of malady.

Here, the multiple image acquired from the UAV will be applied with K-means clustering algorithm to segregate various components acquired using UAV. Next, pre-processing the images to resize and remove the noise. Furthermore, convert the RGB colorspace to HSV colorspace format by using HSV threshold algorithm. HSV threshold is mapped to assess the regions of interest.

Various image segmentation algorithms for diseased leaf detection were not automated and have not been implemented in parallel manner over multiple compute cores. In order to overcome the limitations, this work proposes a new parallel automated image segmentation algorithm for diseased leaf detection for processing bulk input images taken from UAV simultaneously.

1.5 Thesis outline

This thesis contains of five major chapters that includes an introduction, literature review, methodology, results and discussions and lastly conclusions and recommendations.

Chapter 1 introduces the main idea of the project which is briefly explained and some information regarding the research background, problem statement, research objective, research scope and limitations, and thesis outline.

Chapter 2 describes about literature review. The nomenclature of the plant leaf disease is explained briefly and where it usually occurred on the leaves. Also, what are the causes effect the diseases and type of the diseases that may occur. In this chapter will explained about how to detect the diseases in plant leaf. Previous studies regarding the automatic detection of the diseases were presented too. This gives a whole idea on how to improve and proposes a new technique from existing technique.

In Chapter 3, the proposed method and technique used for disease detection in plant leaf are presented. Briefly explained about installing the software, Image acquisition, K-means clustering, image pre-processing and automated image segmentation for this projects.

Chapter 4 explained about the results obtaining from the experiment. The used of K-means clustering algorithm and image segmentation on HSV threshold. The clustering images and analysis of segmentation are presented in the figure and the computed data are present in the table. Then, discussion including proposed algorithm limitations will be carried out in this chapter

Chapter 5 conclude all the findings and recommendations for future works.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

In plants, diseases may be discovered in numerous parts such as fruits, stems and leaves. Diseases in plant should be detected and identified sooner to avoid loss and plants from dying. The technique of digital image processing is improved as many intellectuals contribute more ideas in the corresponding field. Digital image processing involves the use of computer algorithms to process images on digital images.

Image segmentation is a process of splitting the image into several parts. In agriculture, the image of plants will usually be segmented for disease detection purposes. Thus, researchers have been implemented various techniques of image segmentation for detection of diseases occurring in plant. Hence a low cost, faster and accurate system should be developed. Literature review is done to know the state of art for leaf diseases segmentation methods so that improvements can be made on the current methods.

2.2 Plant leaf disease symptoms

Some of the common occurring disease in plants leaf are bacterial disease leaf spots, leaf scorch and leaf blight with tattering. Bacterial disease leaf due to Pathogen-caused leaf spot diseases, particularly those of stone fruit trees and such vegetables as tomato, pepper and lettuce are of two types, those caused by bacteria and those caused by fungus (Chaube, 2017).

Leaf spotting of either kind is generally similar in appearance and effect. Infected plants have brown or black water-soaked spots on the foliage, sometimes with a yellow halo, usually uniform in size. Leaf scorch (also called leaf burn, leaf wilt, and sun scorch) results in browning of plant tissues, leaf margins and tips, and yellowing or darkening of veins which may lead to eventual wilting and abscission of the leaf (Servin et al., 2015).

Blight is a rapid and complete chlorosis which results in death of plant tissues such as leaves, branches, twigs, or floral organs. Accordingly, many diseases that primarily exhibit this symptom are called blights (Chaube, 2017; Martinelli et al., 2015).

As example, Figure 2.1 shows the Banana leaf having spot diseases as mentioned earlier and Figure 2.2 is the Banana leaf having leaf scorch.

Figure 2.1: Banana leaf spot

Figure 2.2: Banana leaf scorch

Similar to other crops, oil palm also gets infected which shown in Figure 2.3 by numerous pathogens i.e., fungi which ultimately results in reduction of yield (Nurdiansyah, Clough, Wiegand, & Tscharntke, 2016; Suwannarach, Kumla, & Lumyong, 2015). Palm oil producing countries have reported numerous diseases for which the most common one is *Ganoderma*. The basal stem rot or *Ganoderma* butt rot disease results in leaf blight and is caused by *Ganodermaboninense,* is the most severe oil palm disease in Malaysia and Indonesia (Izzuddin et al., 2018). Aluminium deficiencies in oil palms are acute in Sumatra (Ahmad et al., 2018; Olafisoye, Oguntibeju, & Osibote, 2017). Smith et al.(Smith, Evers, Yule, & Gan, 2018) recorded over forty ailments of oil palms. They identified that effects cause the ailments from the various growth phases: Germination (with brownish germ ailments) and toddler stage growth. It is worth to be noted that seed decay and brown germ are acute in Malaysia (Olafisoye et al., 2017; Tong, 2017) which is likely due to potassium, phosphorus, boron and calcium deficiency (Oviasogie et al., 2018; Woittiez, van Wijk, Slingerland, van Noordwijk, & Giller, 2017).

Figure 2.3: Palm leaf blight with tattering

2.3 Related work

A disease detection system has the ability to not only detect early symptoms of defective plants, but can also avoid the disease from spreading. The application of image processing on numerous fields has emerge as extensively vital since past few years. In the beginning, the implementation of image processing changed into restricted to specially scientific and terrestrial images. Progressively, people commenced to use this on agricultural purpose inclusive of plant recognition, sickness identity and control and so on. Many different methods of image processing and pattern recognition have been implemented for identification of ailments happening on plant leaves by the researchers.

2.3.1 Image processing

In (Sariputra & Shirolkar, 2016), Sariputra and Shirolkar addressed the fundamental steps to be taken to categorize a disease by processing the image of the impacted leaf. They presented the method that includes image acquisition, pre-processing of images, segmentation of useful components, extraction of features and statistical analysis using a method of spatial gray-level dependence.

Dheeb Al Bashish & et al. (2010) formulated image processing techniques that consists of two steps. First step, captured images are segmented using the K-means clustering methods and secondly segmented images are transmitted through a neural network pre-trained for the classification of disease that affect a plant leaves. The test result mentioned by (Al Bashish, Braik, & Bani-Ahmad, 2010) shows that the neural network classifier based on statistical classification supports accurate and automatic detection of leaf diseases with an accuracy of around 93 percent.

Singh and Misra described that essentially there are four steps in image segmentation of a diseased leaf, out of which, first one is used to convert the input leaf image in Red-Green-Blue (RGB) colorspace to Hue-Saturation-Intensity (HSI) colorspace. In the second step, a value for thresholding was set and green pixels were removed using the masking technique. For the third step, by using the value of threshold defined in second step, diseased portion of the leaf was identified and extracted. Lastly in the fourth step, the segmentation is done (Singh & Misra, 2017)

Satish Madhogaria & et al. (2011) has explained about implemented an automatic pixel-based classification method for detecting unhealthy regions in leaf images (Madhogaria, Schikora, Koch, & Cremers, 2011). The methods mentioned consists of three main steps. The first step, image segmentation is done to split the images into foreground and background. Next, Support vector machine (SVM) is used to predict the class of each pixel belonging to the foreground. Lastly, further refinement via neighbourhood-check to pass over all falsely-categorised pixels from second step.

One such technique uses the thresholding and back propagation network. Input is an image of the grape leaf on which thresholding is performed to mask green pixels. Using K-means clustering segmented portion of the disease is obtained. ANN is then used for classification of the disease as mentioned by (Sannakki, Rajpurohit, Nargund, & Kulkarni, 2013).

Sahoo et el. (2018) used automated dead zone detection in 2D leaf's image. As per their work, segmentation of lesion-based region was carried out by defining a particular threshold value and in their final step, several diseases were categorized by computing the quotient of leaf area and lesion area. It is worth to be noted that the method utilized by them (Sahoo, Panda, Barik, & Panda, 2018) is fast and accurate for the computation of leaf disease severity.

Patil et al.(2011) demonstrated the triangle threshold and simple threshold methods (Patil & Bodhe, 2011). As per their work, these threshold methods were used to lesion region area and segment the leaf area respectively. In final step, categorization of disease is finished by computing the remainder of leaf area and lesion area. According to the research done, the given method is quick and exact for calculating leaf disease severity and leaf area calculation is finished by using threshold segmentation.

Zulkifli Bin Husin et al (2012) captured the chilli plant leaf photograph and processed to determine the health repute of the chilli plant. Their methods is ensuring that the chemical compounds should apply to the diseased chilli plant most effective. They used the MATLAB for the feature extraction and image recognition. Next, Pre-processing is done using Fourier filtering, edge detection and morphological operations in this paper. Computer vision extends the image processing paradigm for classification of objects. Input image is taken with digital camera and LABVIEW software tool to build the GUI (Husin, Shakaff, Aziz, & Farook, 2012).

Naikwadi et al. (2013) used histogram matching to detect diseased leaf in plants. For the diseased leaf the histogram matching is completed on the basis of edge detection technique and color feature(Naikwadi & Amoda, 2013).

Arya et al.(2018) (Arya, Anjali, & Unni, 2018) used image processing and genetic algorithm with Arduino to detect diseased leaf in plants. For the diseased leaf, the histogram matching was done detecting the edge.

Bai et al. (2017) (Bai, Li, Fu, Lv, & Zhang, 2017) and Ma et al.(2018) (Ma et al., 2018) came up with a leaf spot detection algorithm implemented using neighbourhood grayscale information. In their paper, process of ailment spot detection was done by comparing the impact of HSI, CIELAB, and YCbCr colorspace. The next step is applying Gaussian blur on the image for soothing the image. In their final step, Otsu's threshold method applied on colour component and calculation of threshold was done to find the diseased spot on the leaf. However, there was some noise because of background which was shown in their experimental result, camera flash and vein. CIELAB colour model is used to get rid of that noise.

From the paper published by (Phadikar, Sil, & Das, 2012), they established an automated classification system based on morphological changes caused by brown spot and rice plant leaf blast diseases. Using the Bayes and SVM Classifier, radial distribution of the hue from the centre to the boundary of the spot images was used as a feature to classify the diseases.

In (Arivazhagan, Shebiah, Ananthi, & Varthini, 2013), the authors discussed leaf disease detection using image processing and neural network where they used Colour Co-occurrence Methodology to extract the features for texture analysis. They used multi-class SVM (Support Vector Machine) which is a binary classifier for disease classification purposes. They used a winner-takes-all strategy in this classifier, in which the classifier with the highest output function assigns the class.

The basic image properties dealt with image segmentation are its dissimilarity and similarity. Sharp changes in the intensity of image causes dissimilarity whereas similarity corresponds to the process of combining and matching the pixels with the neighbouring one based on its gray level pixel value match (Demšar, Harris, Brunsdon, Fotheringham, & McLoone, 2013; Senthilkumaran & Vaithegi, 2016).

Some of the widely recognized techniques to implement image segmentation are; Otsu's threshold method for automated image segmentation, region growing and region merging technique, edge detection method, watershed transformation and histogram thresholding based algorithms (Firdousi & Parveen, 2014).

2.3.2 Unmanned Aerial Vehicles (UAVs) and Precision Agriculture

The use of Unmanned Aerial Vehicles (UAVs) is viewed an efficient field data collection method. The use of UAVs in agriculture is rapidly becoming widespread, while the implementation of aerospace engineering and sensor technology is reducing costs. Drone technology provides high-tech remodelling to the agricultural industry, with planning and strategy based on data collection and processing in real time.

UAVs collect images at high visual resolutions allowing a centimetre comparison of crop differences. Thermal images can be acquired from the UAV depending on the type of camera used as mentioned in their paper (Wright, Rasmussen, Ramsey, Baker, & Ellsworth, 2004). They also provide instant visual data on large crop areas, helping farm managers make quick decisions.

UAVs can also send and receive live videos to the receiving station on the ground from their flight operation. The ground station provides a user interface that incorporates flight planning, flight control and/or image acquisition (Laliberte & Rango, 2011; Xiang & Tian, 2011).

High-resolution satellite imagery costs and availability often limit their applications in precision agriculture thus UAV could be a cheaper and more practical substitute for high-resolution remote sensed data for satellite and general aviation aircraft (Zhang & Kovacs, 2012).

Drones equipped with cameras can view every centimetre of a piece of land from several angles in research areas to create sustainable agriculture. Drones can create a digital field map, detect crop health issues, find missing animals (Mogili & Deepak, 2018).

UAVs are currently also used in different business and industries. These include the use for crop dusting or precise farming of unmanned helicopters in particular (Ipate, Voicu, & Dinu, 2015). Patel et al.(2013) designed an innovative quadcopter for an infrared camera survey of the farm to show the difference between infected or diseased crops and mature crops (Patel, Patel, Faldu, & Dave, 2013).

In general, UAV-based remote sensing can record images with GSD values 5cm - 20cm from flying height of about 160m - 400m. In the case for precise agriculture, the

GSD values less than 15cm is sufficient to be able to see clearly the individual tree or some object boundary (Figure 2.4) (Rokhmana, 2015).

Figure 2.4: Visual image captured by Parrot PF728000 Anafi Drone

2.4 Segmentation model for diseased leaf detection

The aim of image segmentation is to cluster the entire pixels into specified salient image regions, i.e., regions corresponding to individual objects, surfaces or natural part of objects. Segmentation is an essential process of object recognition, image compression, image database look-up and occlusion boundary estimation within stereo or motion system.

The researchers these days are dealing with the problem of over segmentation of images which ultimately leads to inaccurate results and therefore, leaves a room for enhancing this problem (Danisman, Bilasco, Martinet, & Djeraba, 2013; Jolliffe & Cadima, 2016).

Thresholding is considered an important technique for image segmentation which has got potential to identify and extract the target portion of an image from its actual background on the principal of distribution of gray-levels in an image object.

According to Otsu's method, an image is considered to be a two-dimensional grayscale intensity function which contains N pixels including gray levels ranging from 1 to L (Hancock, Baddeley, & Smith, 1992).

As per Otsu's analysis, the number of pixels having gray level 'i' is denoted by 'f_i'. Therefore, the probability function (P_i) of gray level 'i' in an image with N pixels could be written as Eqs. (2.1) (Tremeau & Borel, 1997).

$$\mathrm{P_i} = {}^{\mathrm{f_i}}/_{\mathrm{N}} \tag{2.1}$$

For the analysis of bi-level thresholding of an image, the pixels could be divided into two classes C_1 and C_2 respectively. C_1 consists of first tier of gray level $(1 \ldots\ldots, t)$ and C_2 consists of second tier of gray level (t+1……...,L). Therefore, the gray level probability distribution for the two classes could be written as Eqs. (2.2) and Eqs. (2.3) (Yin & Wu, 2017).

$$\mathrm{C_1} = \mathrm{P_1}/\omega_1(\mathrm{t}) \ldots \ldots \mathrm{P_t}/\omega_1(\mathrm{t}) \tag{2.2}$$

$$C_2 = \mathrm{P_{t+1}}/\omega_2(\mathrm{t}), \mathrm{P_{t+2}}/\omega_2\,(\mathrm{t}), \ldots \mathrm{P_L}/\omega_2(\mathrm{t}) \tag{2.3}$$

Where $\omega_1(t) = \sum_{i=1}^{t} P_i$ and $\omega_2(t) = \sum_{i=t+1}^{L} P_i$

Above grey level probability distribution method could also be applied for M number of classes assuming that there are $M-1$ thresholds, $[t_1, t_2 \ldots\ldots, t_{M-1}]$ which divide the original image into M classes: C_1 for $[1\ldots.., t_1]$, C_2 for $[t_1+1\ldots\ldots, t_2]\ldots.., C_i$ for $[t_{i-1}+1.., t_i]$ and C_m for $[t_{M-1}+1\ldots\ldots\ldots, L]$ (Yin & Wu, 2017). Eqs. (2.4) represents a column vector.

$$x = \begin{matrix} x_1 \\ \vdots \\ x_\mathrm{n} \end{matrix} \tag{2.4}$$

If the entered values in Eqs. (2.4) are random pixel variables with a precise mean, then the segmented matrix $[seg(X_i, X_j)]$ value Σ is given by Eqs. (2.5) (Prasetyo, Adityo, Suciati, & Fatichah, 2017).

$$\sum\nolimits_{\mathrm{ij}} = \mathrm{seg}(\mathrm{X_i}, \mathrm{X_j}) = \mathrm{val}\left[(\mathrm{X_i} - \sigma_\mathrm{i})(\mathrm{X_j} - \sigma_\mathrm{j})\right] \tag{2.5}$$

Where $\sigma_i = val(X_i)$ and $\sigma_j = val(X_j)$ are the assumed value of the i_{th} and the j_{th} entry in the vector X.

Now let us assume there are n such images to be segmented and if a single image is denoted by vector x, , then the sample computed segmentation could be given by the formula in Eqs. (2.6) (Prasetyo et al., 2017).

$$Seg = \frac{1}{n}\sum_{i}^{n}(x_i - \bar{x})(x_i - \bar{x})^T = \frac{1}{n}\hat{X}\hat{X}^T \tag{2.6}$$

Where i= index for the set of n images, $\bar{x}$ = average of n image pixels

Eqs. (2.6) could also be rewritten in matrix form using $\hat{X}$ to denote the mean centred images $(x_i - \bar{x})$ Eqs. (2.7).

$$\begin{pmatrix} \vdots & & \vdots \\ \hat{x}_i & \ldots & \hat{x}_n \\ \vdots & & \vdots \end{pmatrix} * \begin{pmatrix} \cdots & \hat{x}_1 & \cdots \\ \cdots & \vdots & \ldots \\ \ldots & \hat{x}_n & \cdots \end{pmatrix} \tag{2.7}$$

Let us divide the image patches into v number of pixels based on their similarity. On similarity basis, let us categorize the set of pixels into different clusters i.e., $C_1, C_2, \ldots . C_y$.

Now let us define the set group of every unsigned pixel which at least borders one of the clusters as defined in Eqs. (2.8) (Gu et al., 2015).

$$S = [x \notin \bigcup_{i=1}^{v} \mathrm{C_i} \bigwedge \exists \mathrm{k} : \mathrm{N}(x) \bigcap \mathrm{C_k} \neq \emptyset] \tag{2.8}$$

Here, *x* is the pixel to be assigned, where N(*x*) denotes the current neighbouring pixel of point x which is a part of cluster C_k. As per Eqs. (2.8), x does not lie the cluster C_i and k belongs to pixel x such that *N(x)* is a part of cluster C_k (Cluster with *k* pixels).

Now let us denote δ as the difference of measure between the pixels as defined Eqs. (2.9) (Gu et al., 2015).

$$\delta(x, C_i) = \left| l(x) - mean_{y \in C_i}[l(y)] \right| \tag{2.9}$$

Where $l(x)$ denotes the pixel value of point x and i denotes the index of the cluster such that $N(x)$ intersect $C_i l(y)$ denotes the pixel value of point y. Now to select whether $q \in S$ and cluster C_j where $j \in [1, n]$ such that:

$$\delta(q, C_j) = \min_{x \in s, k \in (1,n)} \{(x, C_k)\} \tag{2.10}$$

Where S is defined in Eqs. (2.8)

Now if $\delta(q, C_j)$ is lesser than the predefined threshold point t_p set by the programmer, the pixel is assigned to cluster C_j, else it must be assigned to another most considerable cluster C such that:

$$C = arg \min_{C_k} \{\delta(Z, C_k)\} \tag{2.11}$$

Now if $\delta(q, C_n) < t_p$, then the pixel is allocated to C_n. If neither of the condition is satisfied, then the formation of new cluster C_{n+1} takes place.

After the pixel has been allocated to the cluster, the mean pixel value of the cluster must be updated. According to Gedraite and Hadad (Gedraite & Hadad, 2011), the function which is used to generate the kernel is a Gaussian function comprising of 2 dimensions and could be defined using Eqs. (2.12):

$$f(q,r) = a.\,e^{-\left\{\frac{(q-q_0)^2}{2\delta q^2}+\frac{(r-r_0)^2}{2\delta r^2}\right\}} \tag{2.12}$$

Where q and r are the vectors, a is the amplitude, (q_0, r_0) is the centre, δq and δr is the standard deviation in q and r direction.

Filter is defined using the variance of the Gaussian distribution. This parameter drastically affects the filtering results. The quality factor (Q) function defined for segmented image using Gaussian blur is given by the Eqs. (2.13) (Gedraite & Hadad, 2011):

$$Q(s,t) = \frac{\sigma_{s.t}}{\sigma_s{}^2.\sigma_t{}^2}\,.\,2\,.\frac{\bar{s}.\bar{t}}{s^2+t^2}\,.\frac{2\,.\sigma_s{}^2.\ \sigma_t{}^2}{\sigma_s{}^2+\ \sigma_t{}^2} \tag{2.13}$$

Where, t is the image without noise and s is the filtered image,$\sigma_{s.t}$ is the covariance between two images, $\sigma_s{}^2$ is the variance of filtered image and $\sigma_t{}^2$ is the variance of source image without noise. $\bar{s}$ and $\bar{t}$ are the mean of images s and t. This quality factor determines the covariance between two images, the distribution in the contrast and distortion in luminance.

2.5 Conclusions

Over the years, multiple image segmentation algorithms have been used to analyse the images. Nowadays, wide range of algorithm is being used to carry out the process of image segmentation such to extract the region of interest area, which is an essential feature that reflects important information about the image surface.

Thus, this project proposed to design parallel K-means algorithm for Coconut, Banana, Dwarf Palmetto, Palm, and Sapodilla images acquired using drone. Then, we will be implement Hue-Saturation-Value (HSV) image segmentation on the acquired sample post K-means.

Since the proposed research is dealing with image data, therefore, it is appropriate to use OpenCV library. Furthermore, OpenCV has capability to exploit high degree of parallelism due to its available rich set of libraries. This scenario makes the condition more favourable for parallel image processing in an efficient manner.

CHAPTER 3

METHODOLOGY

3.1 Overview

This chapter describes the methodology followed to achieve the research objectives. This project broken down into two phases. Phase one consists of data collection, data sampling and setup the frameworks while phase two consists of image segmentation and image analysis. This chapter also presents the mathematical model used and developed for the image segmentation of the leaf plants. Our goal is to find a simple way to segment the regions of interest by using K-means clustering algorithm and HSV threshold algorithm.

3.2 Overall process

The project flows shown in Figure 3.1. Drone will take the images of the proposed plant which is Coconut, Palm, Banana, Dwarf Palmetto, and Sapodilla. The input image will be in RGB colorspace. Firstly, the images will be standardized and resized to 500x500 pixels resolution.

K-means clustering applied to the standardized images. By applying the K-means clustering algorithm we get to segregate the colours of the images based on a set of features into K number of classes. The clustering image will be used for image segmentation later.

Post K-means clustering, the bilateral filter applied to the images. Hence, the Red-Green-Blue (RGB) colorspace will be converted to Hue-saturation-Value (HSV) colorspace and apply threshold to generate contours of the images. Then, threshold image consists of non-zero pixel value will be counted.

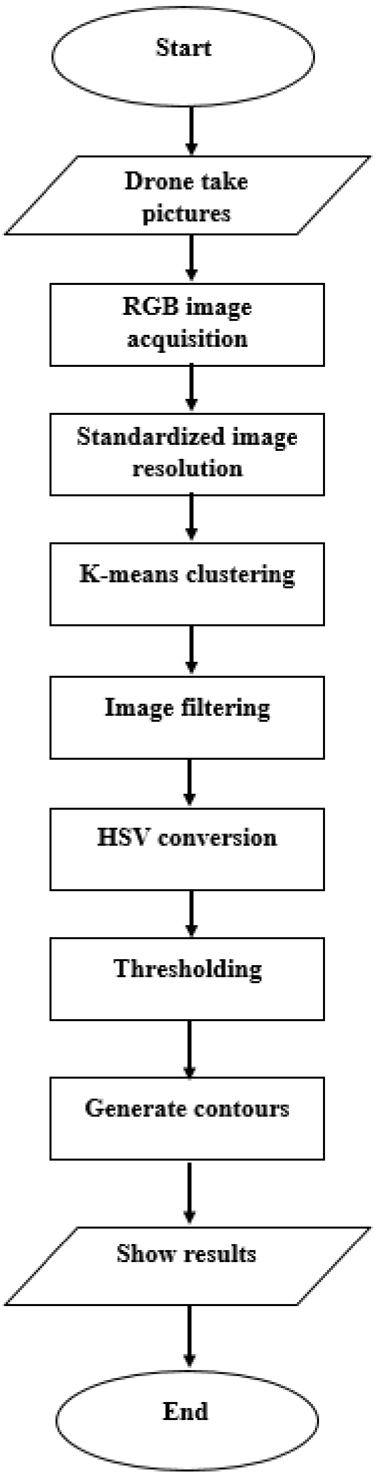

Figure 3.1: Overall process

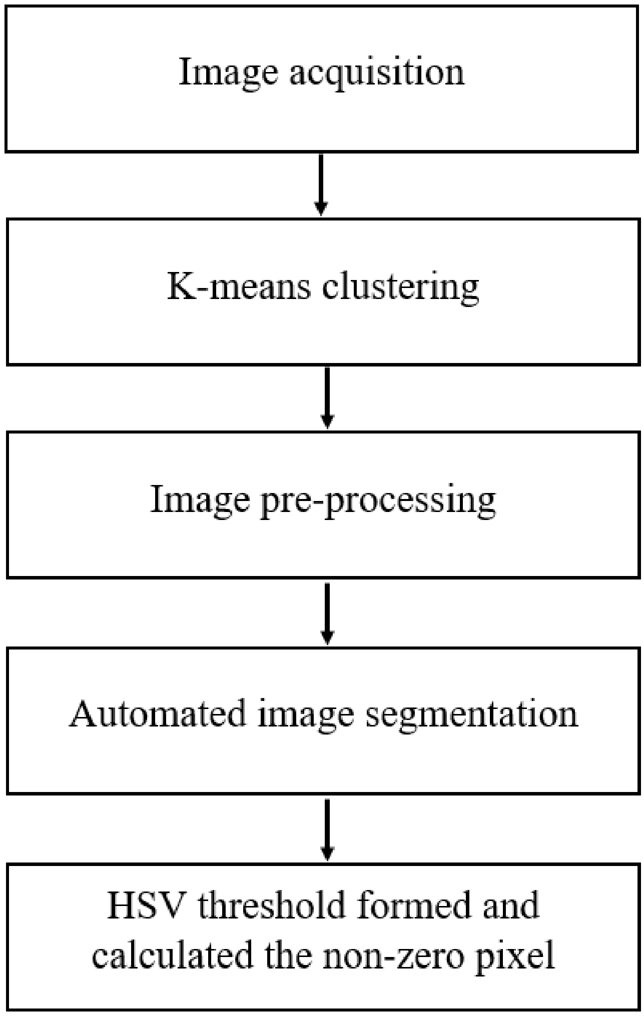

Figure 3.2: Basic steps for plant disease detection

3.3 Install OpenCV 2.4.9 software

We will be using Open Source Computer Vision library (OpenCV) for image pre-processing. OpenCV is library of programming functions mainly aimed at real-time computer vision (C++ interface). The library has over 2,500 optimized algorithms, including a comprehensive set of classical and state-of - the-art computer vision and machine learning algorithms. OpenCV has more than 47,000 people in the user community and is estimated to have over 18 million downloads. The library is widely