Towards Autonomous Cotton Yield Monitoring

Howard Brand

Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
In
Mechanical Engineering

Tomonari Furukawa, Chair
Corina Sandu
Steven Thomson

April 29, 2016
Blacksburg, Virginia

Keywords: Yield Variability, Signal Extraction, Yield Estimation, Remote Sensing, Miniature Unmanned Aerial Vehicles, Zone Management, Laplacian of Gaussian

Copyright 2016, Howard Brand

Towards Autonomous Cotton Yield Monitoring

Howard Brand

(ABSTRACT)

One important parameter of interest in remote sensing to date is yield variability. Proper understanding of yield variability provides insight on the geo-positional dependences of field yields and insight on zone management strategies. Estimating cotton yield and observing cotton yield variability has proven to be a challenging problem due to the complex fruiting behavior of cotton from reactions to environmental conditions. Current methods require expensive sensory equipment on large manned aircrafts and satellites. Other systems, such as cotton yield monitors, are often subject to error due to the collection of dust/trash on photo sensors. This study was aimed towards the development of a miniature unmanned aerial system that utilized a first-person view (FPV) color camera for measuring cotton yield variability. Outcomes of the study led to the development of a method for estimating cotton yield variability from images of experimental cotton plot field taken at harvest time in 2014. These plots were treated with nitrogen fertilizer at five different rates to insure variations in cotton yield across the field. The cotton yield estimates were based on the cotton unit coverage (CUC) observed as the cotton boll image signal density. The cotton boll signals were extracted via their diffusion potential in the image intensity space. This was robust to gradients in illumination caused by cloud coverage as well as fruiting positions in the field. These estimates were provided at a much higher spatial resolution (9.0 cm^2) at comparable correlations (R^2=0.74) with current expensive systems. This method could prove useful for the development of low cost automated systems for cotton yield estimation as well as yield estimation systems for other crops.

DEDICATION

To my grandparents whose joy I have gleaned much strength,

To my mother who has been my strong push forward,

To my uncle Reggie, my auntie Debra, and my coubling Lezly for always being a source of inspiration and encouragement.

Thank you for all you have been to me.

ACKNOWLEDGEMENTS

I would like to first express sincere thanks to my committee members, Dr. Tomonari Furukawa, Dr. Corina Sandu, and Dr. Steve Thomson for your support and guidance. I would like to especially thank Dr. Furukawa for your willingness to be my advisor and the opportunities that you made available to me. Dr. Sandu, I would like to thank you for your continued inspiration and support. Dr. Thomson, thank you for taking a chance on me and allowing me the opportunity to do research at the USDA. Thank you for the confidence you placed in me. I have grown a lot as a researcher, and in my awareness of who I want to become professionally.

Special thanks to Dr. Huang for your guidance and your willingness to allow me to conduct research with you. Thank you for all of the encouragement and the guidance you provided, especially with my writing.

Special thanks to Mr. Swafford for your recommendation which was the springboard for this project. You and Miss Betti have been an inspiration for me throughout the years. Thank you doesn't seem to be enough.

Finally, thank you to my mom and my grandparents. You have always been in my corner. I could not have done this without you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

In the field of agriculture there is a shift towards developing robotic technology that allows for the automatic monitoring and analysis of crop conditions. These technologies are oriented towards precision agriculture where ultimately, crop yields are maximized through reliable knowledge of environment-specific crop development. With regard to yield monitoring and prediction systems there are numerous research initiatives aimed at developing parameters, observable by sensory devices, which provide insight on plant growth, development, and potential yield capacity. These parameters include those that influence crop vigor such as disease, lack of water and nutrients, and phenology.

There is also a movement in remote sensing to develop technologies for which crop yield can be measured upon conducting a field investigation. This is desired, in that it is crucial for farmers and agricultural aids to make quick local yield assessments at harvest time. Also important at the end of harvest time is having a proper assessment of within-field yield variability. Proper understanding of yield variability provides insight on the potential yields of crops geographically and insight on zone management strategies and applying necessary treatments. Crop fields are usually too vast for a manual inspection to possibly yield accurate estimations of yield and yield variability. This has been proven to be the case for cotton farmers. Even the most experienced cotton farmers and agricultural aids, more often than not, have incorrect perceptions on cotton yield and yield variability (Rejesus et al., 2013). Current research methods for inferring yield and mapping field variability depend on expensive systems such as a manned aircraft transporting hyperspectral cameras. These methods are often limited to 1-8 m resolution and typically explain approximately 61- 70% of the variability (one field 69%) (Zarco-Tejada, 2005). These methods sometimes involve measurement of physical characteristics of the cotton vegetation before harvesting along with use of spectral vegetation

indices such as the Normalized Difference Vegetation Index (NDVI), Triangular Vegetation Index (TVI), and Normalized Difference Water Index (NDWI). This also involves measurements to be taken during periods of the cotton plant's development. Aside from more expensive methods, other techniques involve more tedious processes such as boll counting. These methods are susceptible to error when not considering other factors such as temperature and seed content. Yield monitoring systems utilized during harvest are limited by spatial resolution even when utilizing differential global positioning system (DGPS) equipped systems. The accuracy of yield monitoring systems is also often negatively affected by dust/trash collect on the photo detector. These effects have been observed to be more severe when measuring across fields with low yields (Vellidis, 2003).

There have been improvements in current crop yield estimation methods not involving vegetative inspection. These methods involve utilizing image systems to observe the crop present during harvest time. This has facilitated the development of many image processing algorithms for the detection of a variety of crop. The two main tasks accomplished in these methods are crop detection and yield estimation. Crop detection and yield estimation presents a number of challenging problems. Current fruit detection methods are developed to accomplish detection under various lighting conditions. This has been accomplished through developing discriminant models using various labeled datasets, employing intensity insensitive color spaces, adaptive methods for certain lighting conditions, or texture filtering based techniques. Current methods are designed for crop for which textures across their surfaces can be observed. For large cotton fields, it is not practical to collect images precise enough to observe and distinguish cotton texture. Current methods are also designed for crop which have more distinct color. This allows the crop to be distinguished from the non-fruit background via an observation on their

chromaticity. Cotton bolls, dead cotton plants, and soil have relatively neutral colors. They have very little chromaticity and more distinguishably characterized by their intensities. The challenges present in yield estimation are related to accounting for multiple detections of individual crop, clustered detections of crop occluding each other, and touching crops. These methods are for larger crop in which the contributions made by individual fruit to the binary signatures of fruit clusters are explicitly observable. Cotton bolls, however, can be surrounded by other cotton bolls. The individual contributions of individual cotton bolls to a clustered binary signature may not be explicitly observed in the binary signature.

The work discussed herein focuses on the development of automatic cotton yield estimation from RGB images taken from a first-person view (FPV) high resolution camera transported autonomously by a cost effective miniature unmanned aerial vehicle. An investigation was conducted to develop a means of detecting cotton bolls, developing a segmentation mask for cotton bolls, and finally deriving a relationship of the mask to cotton lint yield. This has potential as a cost effective solution utilizing affordable and accessible technologies.

2. Literature Review

Crop yield estimation is a large topic within the field of precision agriculture in that there are various types of crops with many different features existing within a variety of environments. There are a wide variety of methods for estimating crop yield and yield variability with a wide variety of sensor configurations. The methodology can be divided into two ideals for crop yield estimation. One ideal is to measure the physical characteristics of a crop and correlate those characteristics to crop yield. The physical characteristics may include height, temperature, leaf vitality, and soil properties. These methods work generally well when the physical traits of interest stipulate the yield production of crops according to a knowledge of crop fruiting behavior. Another ideal is to use a sensor to detect the fruit within the crops and relate the presence of fruit to yield. The techniques can be divided into two stages: fruit signal extraction and yield estimation. The fruit signal extraction involves dividing the measurements of the sensor into a fruit set and a non-fruit set. This division is usually represented in the form of a mask or some labeling convention which spans the space of the measurements. The algorithms developed for signal extraction are designed based on a model of the fruit in the measurement space or within a transformation of the measurement space. The design of the models are based on environmental conditions that may affect observation of the sensor. The masks resulting from signal extraction algorithms serve as inputs to the yield estimation phases. The algorithms developed for yield estimation depend on metrics associated to the yield and vary depending on the crop. Tree crop yield is mainly associated with a manual count while cotton yield is associated with a mass/area density such as kg/ha. For count based methods, it is necessary to handle occlusions in order to obtain more accurate counts. These methods are based on edge detection and shape modelling in order to determine whether an observation is made on more

than one fruit. For other metrics yield estimation is achieved by finding a transformation between the measurement space of sensor and the yield space.

2.1 Current Cotton Yield Estimation Methods

Zarco-Tejada et. al. (2005) conducted a study in which they performed a temporal analysis of the performance on different plant indices in describing the within field variability in cotton crop yield. This temporal study involved hyperspectral images (taken by the Airborne Visible and Near Infrared (AVNIR) hyperspectral sensor) of 1-m resolution to be collected bi-monthly. The yield was estimated with a yield monitor. The indices studied for yield correlation were hyperspectral indexes pertaining to crop growth and canopy structure, chlorophyll concentration, and water content. The variability was composed of the three classes: low, medium, and high yield. Segmentation was performed using the K-means clustering algorithm on the index results using a cluster size of three. Confusion matrices were used to calculate the kappa (κ) coefficient and overall accuracy. One motivation for this research was that the tradition NDVI saturated at larger leaf area index (LAI) values and was effected by shadows and soil background. They found that structural indices related to LAI performed better at early growth stages. They also found that hyperspectral indices related to plant physiological status performed better at later growth stages close to harvest time. RDVI overall accuracy at early stages 61% ($\kappa = 0.39$), drops to 39% ($\kappa = 0.08$) before harvest. MCARI chlorophyll index maintained 51% ($\kappa = 0.22$) overall accuracy.

Domenikiotis et. al. (2004) conducted a study in which the yield of cotton fields could be estimated before harvest time via satellite images. The method involved used of an extension of the Normalized Difference Vegetation Index (NDVI) referred to as the Vegetation Condition

Index (VCI). This index was used to measure impacts of the weather on cotton yield. This was found using a data set over 10-day time intervals from 1982 to 1999. The time periods which yield the results with the highest correlations were identified as critical periods for yield estimation. The researchers recorded achieving a correlation coefficient as high a 0.8. This study was developed for estimating yields for entire fields and was not designed for within field yield variability.

Huang et. al. (2013) conducted a study in which a multispectral camera was used to make observations of the cotton field with differing irrigation and nitrogen treatments. In this study, a model was developed relating the ratio of vegetation index (RVI) with cotton yield. Accounting for soil electrical conductivity (EC) an R2 value of 0.53 was achieved. The multispectral camera sensor was transported across the field on a manned aircraft.

These current methods for cotton yield estimation utilize expensive sensory devices not easily accessible to farmers such as satellite, hyperspectral, and multispectral cameras. Also, these camera systems require installment on other expensive to use equipment such as manned aircrafts. This incurs additional costs making accurate yield variability measurements less of an option for cotton farmers. To this effect, the literature was explored for developing a low cost solution for which an RGB camera could be used to measure cotton yield variability.

2.2 Global Thresholding Methods for Crop Signal Extraction and Yield Estimation

Global thresholding based crop signal extraction involves identifying a space for which the crop exists in a subspace separate from other objects in the sensor's field of view (FOV). This space can be the measurement space of the sensor or a transformation of the measurement

space. Color (or RGB) cameras are popular sensors for which these methods are used as they are relatively inexpensive compared to other camera systems.

Patel et al. (2012) developed an algorithm which could be applied to estimate yields of different tree crops such as apples, pomegranates, oranges, peaches, litchi and plums. These crops were imaged with a color camera under various lighting conditions and varying degrees of occlusion by foliage. These images were retrieved from the internet. The yield estimation procedure began with an extraction of fruit pixels from the images. This involved a transformation of the images from the RGB space to the L*a*b color space. From the L*a*b space, a threshold was then applied to the "a" axis to develop a mask for fruit pixels. A fruit count was developed via a combination of edge detection (done with Sobel operator) and circle fitting in which circles were counted in order to obtain the fruit count. The segmentation accuracy of fruit was 98% and the average yield measurement error was 31.4%. Though the algorithm was purposed for handling occlusions of fruit by foliage, it was not designed to handle occlusions by other fruit. This most likely remained a source of error in terms of fruit count.

Wang et al. (2013) developed a method for estimating the yield of an apple orchard composed of red and green apples. Their measurement system consisted of an RGB stereo rig and a controlled (and snapshot flashing) light source suitable for illuminating apples at night. The sensor array was mounted on an outdoor vehicle which was manually driven. During the data acquisition, both sides of each tree row were scanned. Their method for fruit signal extraction involved the images to be converted from RGB space to HSV space. The red apples pixels were extracted based on ranges selected in the hue and the saturation axes. The red apples had a distinct color from the environment and allowed for HSV threshold for their detection. For extracting green apple, pixels ranges in hue and saturation axes were also implemented. The

saturation axis was instrumental in separating the highly saturated green apples from the lower saturated green foliage. The green apples, however, were much more reflective. The red apples and flash photography caused specular reflections on the apples. The regions of specular reflectance caused the apples to have regions of really low saturation. The regions containing specular reflection were detected by searching for the local maxima and framing a window around. Gradients of scan lines at different orientations within the window were analyzed and compared to a roundness criteria in order to indicate whether the region was a specular reflection or not. The mask for the specular reflections were then combined in order to make a full mask for the green apples. After segmenting the apples, individual apples were identified. The eccentricity of the regions in the apple mask were calculated and a threshold was applied to the eccentricity. This thresholds the objects in the mask for roundness. Regions that usually pass the threshold test are those composed of individual apples which have little occlusion and don't touch other apples. Noise from small round regions which aren't individual apples but round visible parts of occluded apples are removed via comparison of the area of each region with the average area. The minor axes of the elliptical apple regions that have the same normalized second central moments as a remaining round apple regions are calculated. The mean of the minor axis length of all the remaining regions is used as the average diameter. The major axes of the elliptical apple regions with the same normalized second central moments as an apple region were calculated and compared to the average diameter. Those major axis which were greater than twice the average diameter were viewed as touching apples. With the use of a stereo rig and the GPS, the locations of the apples were triangulated and registered according to position. Multiple observations of apples can be merged in this manner improving the count measurement.

In their vision algorithm their yield estimation errors were -3.2% for a red apple block with about 480 trees, and 1.2% for a green apple block with about 670 trees.

This method accounted for touching apples and multiple observations of individual apples in order to improve its count. The work however doesn't account for apples which largely occlude each other. It also utilized a system which has a high power requirement and is rather large. This makes suitable only for crops which can be measured from ground vehicles.

Guijarro et al. (2011) developed a method to identify vegetation, soil, and sky (if any) and to classify plants and soil based on texture. This method was tested on images of barley and corn crops at different growth stages under different illumination conditions. The ground was captured under high quality illumination conditions. The method first involved combining greenness indices: ExG, ExGR, CIVE, and VEG. The indexes were constructed from the individual channels of the RGB images in order to develop new images based on the index relationships. The greenness index images were then combined via a weighted average. The weight depended on the uniformity of the index images, with lower uniformity being given the highest weight. A threshold was then applied to the combined image using the mean of the combined image as the threshold value. Only the most saturated green objects passed. Otsu's method is used on a constructed ExB image to separate blueness from ground gradients to blueness from the sky. An ExR index was also constructed. A threshold on the ExR was accomplished through the use of its mean as the threshold value. In images without sky, soil was detected based on it passing the threshold test on ExR and it failing the threshold test for greenness on the combined greenness image. In images with sky, the detection of the soil also required it to fail the blueness threshold test on the ExB image. Classifications of plants and soil were developed on images with and without sky using the ExB and ExR images. The binary

masks developed from green and red threshold images were used to create plant and soil images. The masked images were then input into a fuzzy clustering algorithm that divided the soils and plants into classifications in the RGB color space. Error in detecting the vegetation in the combined greenness was 8.31% which improved the results of the individual greenness indices. This work presented the benefits of color ration for describing color independently from intensity. It also introduced a framework in which color indices could be combined to improve their individual results. Lastly, they introduced how thresholding color indices can serve as an input to automatic plant classification.

Gong et al. (2013) developed an algorithm to estimate the yield of citrus crops from images taken under various illumination. The signal extraction method involved a threshold of the difference of the red and blue channels and the comparison, red > green, of the red and green channels. A median filter is used for signal extraction to remove noise due to scattering from the sunlight. Noise was removed in masks generated from the signal extraction via a size within region filter. This avoided removing small binary signals pertaining to partially occluded fruit. The quantity estimated by this method is in terms of fruit count. The resulting binary mask is processed using an area filter which identified individual fruits from clusters of fruits. The binary signatures of clustered fruit were then processed using a modified version of the eight point connectedness algorithm. In this modification, a starting point and a transition point was established at the maximum point and the minimum point respectively. The binary signals are then traced in which their orientations are labeled depending on the direction. The points where there were changes from left to right or up to down and vice versa were labelled with 8. The 8's were counted and divided by 4 and the resulting number was the amount of fruit in the cluster. The method achieved a 90% yield accuracy. The images were taken and processed on a cell

phone. The researchers provided a framework for which yield could be estimated with a commonly used mobile electronic device. In this work, the fruit has a very distinctive color from the foliage and is relatively large and exposed when compared to other tree crops.

Teixidó et al. (2012) developed a segmentation algorithm for the detection of paraguayo peaches (which have irregular shapes). The segmentation was accomplished by developing linear models defined on the RGB space, in which pixels could be classified by their distance to the models. These models were developed in order to describe the various colors potentially present in each peach along with all its shades due to illumination variations. Branches and leaves were similarly modeled in this way. A binary image representing the peach samples were then produced. Afterwards, contours were fitted around the binary signal. Each individual signal contours maximum distance between two contour pixels and the major axis length of the ellipse were found. The signal extraction had an overall average error of 11.6%. The error for bright illumination was 19.7%, while the error for low illumination was 8.2%. For occlusions up to 33%, between 34% and 66%, and above 66%, the errors were 8.6%, 12.2%, and 23% respectively.

Søgaard and Olsen (2003) developed a method for determining barley crop rows in a color image by fitting a line over the center of the rows (representing the row center position and row orientation) using linear regression. The accuracy of the line was compared to a position of a reference string positioned in the adjacent inter-row space, parallel to the row. The vegetation signal extraction was accomplished by constructing a grey-level image from the red, green, and blue channels according to this relationship: (2 x green) – red – blue. This difference was designed to amplify green color. The image was then divided into horizontal strips of a set height where elements were assigned to row centers by identifying local centers of mass of the

rows based on their grey levels. A weighted linear regression was then used to plot a line along the rows. The researchers provided a framework for which the grey levels could be studied in terms of their centers mass without a threshold being applied.

Payne et. al. (2013) developed a mango crop yield estimation algorithm which processed color images of the mango trees taken during the daytime. Their method for extracting fruit signal from the background signal was by selecting a range in RGB and YCbCr space and using a texture segmentation approach based on local pixel variations. The resulting blobs making up the final binary mask were counted in order to obtain an estimation of the mango yield. 15 trees were imaged on 4 sides for which the correlation between the manual count and the algorithm estimated count achieved an R2 value of 0.91. For 555 other trees, only imaged on one side, an R2 value of 0.74 was achieved. A bias adjusted root mean square error of prediction of 7.7 was also achieved.

Payne et. al. (2014) extended the previous study on mango crop yield estimate. Less mature mangos were poorly identified when adjusting the filtering settings (YCbCr and RGB) of the original algorithm due to the lower presence of red coloration. The algorithm was changed to be more dependent on texture than color. Their method for fruit signal extraction involved utilizing a controlled diffuse illumination setup designed to sample images during night hours. In the method, a 3x3 variance filter (hessian filter) was applied to the image in order to remove pixels corresponding to largely textured regions such as mango foliage or grass, and in regions of varying low texture such as the sky. The Hessian filter had a smoothing scale of 10 and was applied to the image using an eigenvalue comparison. The largest eigenvalue was able to remove line-like leaves and stems. Next the Cb and Cr layers were averaged, and the trunk was filtered using a blob size filter on the binary image. The Cb and Cr layers were then subtracted,

and the resulting image was thresholded to remove false mango fruit detection. A threshold on the Cb layer removes overexposed, yellow leaves which were identified as a source of error. The binary layers mentioned were combined to form a binary mask for the mango fruit. The particles of the mask were then filtered according to a lower and upper limit on the size of the particles. The remaining particles were then compared to the manual count of the crop yield for the tree. Their method achieved a 78.3% of detection rate with a 10.6% error on a calibration dataset. The resulting correlation between the estimate count and the manual count had an R2 value of 0.63. When they evaluated other datasets, there were mixed results with issues due to variations in foliage characteristics being cited by the researchers.

Onyango, C., & Marchant J. (2003) developed an image segmentation algorithm which combined color information and knowledge of the plant grid in order to inform plant classification (as a crop or weed). In the method, an image was defined as a union of two sets, vegetation and soil. Vegetation was composed of two components, crops and weeds. The algorithm combined class membership likelihoods for crops, weeds, and soil dependent upon a color parameter and a position (pixel coordinate) in order to obtain the probability of an image being a pixel. Given that the color parameter and the position are independent, the probability of a pixel belonging to a crop can be equivalently modeled as the product of the probability that a pixel is a crop conditional on color, and the probability that it is conditional on position. The probability of a pixel being a crop conditional on a color parameter was developed for the image. Applying a close operation and fitting a grid near the centroids of the crops allowed for the calculation of a probability of a pixel being a crop conditional on the position. A Gaussian distribution of a set standard deviation was fit around the grid intersections for the formulation of the conditional probability. The joint probability of a pixel being a crop dependent on color and

position was found by multiplying the two conditional probability functions and applying an open operation on the result to remove noise.

The framework of global thresholding based methods combine thresholds on multiple feature spaces. This allows the crop to be selected based on a multiple dimensional subset existing within a multi-dimensional range of the combined feature space. For current methods, the feature spaces utilized are color and texture spaces. In regard to color spaces, the more sophisticated color spaces used are those that separate chromaticity from intensity. This is utilized in outdoor environments to manage variations in illumination. These methods; however, require that the crop has a distinct chromaticity from its environment. White cotton bolls; however, possess very little chromaticity. Cotton field objects such as dead vegetation and soil also contain very little chromaticity. If the cotton bolls aren't separable based on its value in intensity, then chromaticity can contribute very to uniquely characterizing the cotton. For crop that doesn't exist in unique color ranges texture, filter outputs are also employed to separate the crop from surrounding objects. These filters; however, have been designed around a smoothness constraint in which the smooth crop is separated from surrounding foliage. This is useful when the crop is comparatively much larger than its surrounding texture. Cotton bolls; however, are small and are relatively much closer in size to its surround foliage.

2.3 Learning and Classification Methods For Crop Signal Extraction and Yield Estimation

Classification based crop signal extraction allows crop to be separated from its background in feature space via a hypersurface. Crops may exist in a subspace of the feature space which is has a complex shape, whose boundaries are not appropriately identified via a multi-dimensional range. Classification based methods also provide a framework for which

labeled data can be used to tune the hypersurface, bounding the crop subset according to sensor and field conditions. This is ideal towards developing a yield estimation towards commercial systems which can be used by farmers.

Chaivivatrakul et. al. (2010) developed a method for the detection and 3D reconstruction of pineapples from a video sequence. During the detection process, Harris corners and SIFT keypoint descriptors were extracted and keypoint classification (as pineapple or non-pineapple) was accomplished using support vector machines (SVM). Structure from motion was utilized to obtain a 3D point cloud representing regions of the fruit's surface. A least squares estimation was used to fit an ellipsoid on the resulting point clouds. Two experiments, one indoor and the other outdoor, were conducted in which a 3D reconstruction of the fruit was performed from a set of images. The keypoint classification accuracies of the indoor and outdoor data were 87.79% and 76.81% for base rates 81.42% and 53.83%, respectively. A third experiment was performed in which a 3D reconstruction was developed from indoor data having an average error of 34.96%. In utilizing SVM to classify keypoint features, various kernels and hyperparameter settings were tested in which the radial basis function (RBF) kernel was found to provide the best performance. In utilizing SVM to classify keypoint features, various kernels and hyperparameter settings were tested in which the RBF kernel was found to provide the best performance. They reported issues with scale ambiguity. The researchers provided a framework for combining fruit classification with feature extraction.

Linker et. al. (2012) conducted an experiment estimating the yield of green apple tree crops. Their study involved detecting green apples in an image and estimating the amount of apples present in term of a count. Their method for signal extraction involved using a pixel descriptors based on color and smoothness. These descriptors were assigned to each pixel and

were found by observing the pixel neighborhood. The method also involved using a k-nearest neighboring (KNN) classifier to classify selected points as belonging to apples. This was used to deal with the relatively small amount of apples in the image composition to a wide variety and range of non-apple pixels. The KD-tree stored the count of the samples in each leaf. The probability of pixel belonging to an apple was governed by looking at its corresponding leaf (KD-tree) and observing the number of samples belonging to the leaf. Apples were manually marked in order to develop a calibration set where smoothness was scaled to from 0-255 and leaves (KD-tree) were made based on color and smoothness. The underside of tree leaves (plant) had similar color and were often false positives.

Due to the apples' shape, lighting distribution, and light source angle of projection there existed a reflectance gradient across the apple surface. The places on the apples that were darker were not considered to belong to the pixels. For this reason, the areas of the apples detected as apples were treated as seeds. Assuming that the feature distributions remained the same, but the ratio between the populations (apple and non-apple set) changes, a criteria was developed in order to grow seed regions to edges of the apples. The contours were then divided into segments by estimated arcs that preserve contour endpoints. The neighboring arcs were then combined into potential apple objects. The discrepancy of the combined contours was then quantified by comparing them to an apple model. The model used to approximate the arcs was a circle with radius and center point. Discrepancies were measured in the radial and angular deviations of the arcs to the model. Multiple detections of the same object were accomplished by observing the circle centers and measuring the degree of overlap. The results of the estimated count were compared to a manual count. Under direct illumination the algorithm achieved a detection rate

of 85% a large number of false positives. Under manually diffused light, the algorithm achieved a detection rate of 95% with a false positive rate of less than 5%.

This work provided a means for classifying fruit based on color and smoothness. It also provided a framework for managing illumination gradients across by utilizing detected portions of the fruit as seeds. This allowed the binary signature of the fruit mask to extend to the fruit boundaries despite decreasing gradients in color across the fruit surface.

Kurtulmus et al. (2011) developed a method for detecting green citrus fruits in a color image. The fruit signal extraction involved using three sizes of the scan windows based off of the range of the diameters of the fruit. The windows were shifted by 20 pixels increments in both horizontal and vertical directions. Before performing the window, scanning the background was eliminated via color thresholding. Each window was tested by classifying eigenfruit of intensity component, eigenfruit of saturation component, and circular Gabor texture analysis. A majority voting approach was used to merge results of the classifiers. Two positive votes were enough to determine the sub-window to be fruit. Detection centers were then found from binary images and positive detections were marked with solid circles about the detection centers in the binary images. Multiple detections on the same fruit were merged analyzing blobs that were the result of two circular solids intersecting each other. It then determined new centers by the blobs' major axes converting the multiple detections into one detection. Afterwards, the number of blobs in the binary mask represented the number of fruit in the image. The method achieved a detection rate of 73.5%.

Suzuki et. al. (2008) developed an algorithm which analyzed hyperspectral images in order to segment the crops from the weeds. The algorithm began by thresholding an NDVI image in order to segment plants from soil. The pixel spectrum was then normalized such that

the minimum waveband was zero and the mean was 1. This was to deal with the varying total intensity levels due to illumination and iris lenses of the hyperspectral camera. Afterwards, a set of explanatory variables were generated from wavebands by two different methods RAW and PCA. The RAW method involved a stepwise selection using Wilks' lambda. This was an index for statistical significance that ranged from 0 to 1. In the PCA method, the wavebands were transformed into principle components via principle component analysis. Then in a stepwise manner, explanatory variables were selected from the principle components. These variables were of course orthogonal. In order to discriminate between crops and weeds, linear discriminant analysis (LDA) and a multi-layered neural network (NN) were used. LDA discriminates between categories by finding a linear discriminant function for multi-category pattern classification. The linear discriminant functions were calculated as the ratio of variance between categories and the variance within an individual category. A neural network is a nonlinear model which describes the parallel process of the neurons found in the human brain. Neurons are contained in layers with neurons connected to other neurons in other layers. The back propagation algorithm was used to develop a 3 layer NN model with an input, hidden, and output layer. The back propagation algorithm is a supervised learning technique for the error of the output, and is propagated to the input layer in order to modify the strength of the synapsis connecting the neurons. Four discriminant models were created according to the explanatory variables and discrimination procedure: RAW-LDA, RAW-NN, PCA-LDA, and PCA-NN. Crop and weed pixels were selected and then partitioned into two datasets: modeling and validation. The validation dataset consisted of 30 pixel spectra of a single plant sample from among the other plant samples. The pixel spectra of the remaining plants were used as the modeling dataset which was used to develop a pixel discriminant model. The model was evaluated using the

validation dataset. This process was repeated according to the amount of plant samples. The success rate was calculated as the number of correct predictions over a total number of tests during the model evaluation. Images consisted of crop and one of the 4 weed species. The algorithm had trouble with shaded areas and mistakened shaded crop areas as weeds. The author explained that the model would have to be developed with spectra in shaded areas as well as sunny areas. Pixels at the edges of leaves were also misidentified. The success rate of identifying plants from the soil was 99% while the success rate for identifying the crop from the weeds using the 4 models averaged around 90%.

Hung et. al. (2013) developed an algorithm in which almond fruits were separated from their background in multispectral (RGB-IR) images. This was aimed at providing a yield estimation. The method used an unsupervised learning approach in which unsupervised feature learning was applied to a public dataset of RGB-IR images. This was to address the lack of a feature set developed in the field (researchers noted more clarity is sought). The segmentation procedure involved pixel classification in which neighboring pixels were correlated. This was accomplished under the CRF (conditional random fields) framework. A CRF is a 2D lattice where pixels represent vertices and the relations between pixels represent edges. Every pixel in the image was assigned a label. The optimal labeling of the pixels were found by minimizing an energy functional which is composed of a unary potential, a smoothness term, and a partition term. The unary potential consisted of the likelihood of pixel having a label. The unary potential was measured according to a semi-supervised composed of an unsupervised feature learning step, followed by a supervised feature learning step for a classifier. The unsupervised feature learning step involved a random sample of RGBIR image patch and developing a multi-scale dictionary. Once a multi-scale set was developed, an additional supervised feature learning

approach was applied in order to train a softmax regression classifier. The learned models were then used to perform image segmentation. A global classification accuracy of 88% was achieved. The global accuracy measured the total number of correctly classified pixels of the entire dataset.

Jeon et. al. (2011) utilized an adaptive thresholding method in order to remove the soil background from images of vegetation taking at different illuminations. The algorithm consisted of applying an initial threshold value to an image performing a check of the percentage of the image that passed the threshold test. When the percentage (rate) is over an upper limi,t the threshold value is decreased. When the percentage is under a lower limit the threshold value is increased. The value of increase/decrease is governed by an equation which is dependent on the current rate and rate value dependent on how the current rate relates to the lower/upper limits. Once the soil background was removed, the resulting image was processed through an adaptive neural network (ANN) (trained by 119 images) of weed classifications. Weeds were then classified according to some relationships regarding their expected shape. The segmentation was found to have processing errors of 2.9% and 2.1% for images collected with a stationary and non-stationary camera respectively, while illumination changes took place. The trained ANN initially correctly identified 72.6% of corn plants in field images. They however identified the source of error being related to incomplete images of corn plants at the image edges. After addressing this source of error, they achieved identification rates of 92.5% and 95.1% for two images sets. This study was motivated by plant specific direct applications (PSDA) such as weedicide applications.

Jafari et. al. (2006) developed an algorithm that segmented 7 classes of weeds from an image of weeds and sugar beets. The motivation of this study was to identify weeds for spot spraying. The algorithm was able to identify weeds in images of varying illumination where some plants were illuminated, and some were in shadow. This was accomplished by designing algorithms for shadow and regions and illuminated regions and providing an average luminance test in order to identify them. The results of both the illuminated and shadowed methods were combined using a logical or. The weed segmentation method used in this algorithm was discriminant analysis. Discriminant analysis (DA) is a technique used to build a predictive model of group membership based on observed characteristics of each case. Discriminant functions (DF) are the linear combinations of the standardized independent variables, which yield the biggest mean differences between the groups. Based on the independent variables red (R), green (G), blue (B), linear combinations were developed to separate each weeds from sugar beets. There were two different ways this algorithm was implemented. One way that this algorithm was implemented was to develop 7 DFs for each class of weed, was to develop masks that remove sugar weeds from each weed. The results were combined in order to segment all of the weeds. The other way this algorithm was implemented was to develop on DF, two separate weeds and sugar beets. These were implemented for both shadowed and the illuminated regions of the image, and the results were combined as mentioned earlier. A classification rate of 88.5% was achieved for all weeds except Portulaca in sunlight and 88.1% for weeds except Convolvulus in shadow.

The methods for classification based signal extraction allow crops to be detected under varying, complex conditions such as under varying illuminations, foliage, trunks, and sky. These methods, however, require a labeled measured set from which a hypersurface can be developed.

If this is not provided the subspace containing only crop cannot be identified. These methods also require feature sets for which classifications don't overlap. A unique feature set for which cotton bolls have unique signatures would need to be obtained. This also requires; however, a labeled data for which a ground truth is established for evaluating feature selection.

2.4 Summarizing Conclusions of Literature

Before discussing the conclusions of the literature review, it is important to make mention some aspects of the experimental setup. Firstly, the data set was provided without a field level survey to provide insight on the features observed. Secondly, no information was recorded regarding environmental conditions during flight. Lastly, only one altitude and flight path was selected for the experiment. This does not facilitate a means of attaining a labeled dataset for developing a classifier. Furthermore, features must be identified which are unique to cotton. What is provided in the experimental setup are manually recorded yields. Considering the framework of global thresholding methods, features and corresponding feature extraction techniques can be evaluated based on a resulting correlation with yield. Identifying unique features is crucial to the success of a classification algorithm. Features which do not have overlapping classifications are used in classification based approaches. Thus the study contributes to a fully automatic method for cotton yield estimation.

3. Experimental Setup

Figure 3.1 displays the cotton field of study. The cotton field of study had a total area of 6-ha and was composed of 39 experimental plots. The field is located in a research farm of the United States Department of Agriculture (USDA), Agricultural Research Service (ARS), Crop Production Systems Research Unit at Stoneville, Mississippi (latitude: 33°26'30.86", longitude: -90°53'26.60"). Twenty of the 39 plots were selected for study and are boxed in red in Figure 3.1. These plots were 48.8 m long, 23.2 m wide and were each composed of 24 rows. A randomized complete block design (RCBD) with two additional replicas in each block was used with five nitrogen fertilizer application rates (0, 56, 112, 168, and 224 kg/ha) as treatments. On April 21, 2014, a cotton cultivar (FM1944GLB2) was planted. Nitrogen fertilizer was applied as a urea-ammonium nitrate solution (N-sol, 32%) to each plot at the designated rates. This was accomplished with a side knife drill 57 days after planting (DAP). The cotton was irrigated using a center pivot irrigation system during the growth season. The irrigation cycles were scheduled based on soil moisture content observed by soil moisture sensors. On September 8, 2014, the cotton field was defoliated. The cotton was then machine-harvested with a spindle-type picker on October 1, 2014. The cotton at the central 16 rows of each plot under study were weighed in a load cell equipped buggy. The cotton in the studied plot were collected and ginned at the USDA ARS Cotton Ginning Research Unit at Stoneville, MS. The cotton lint yield for each plot under study was then determined.

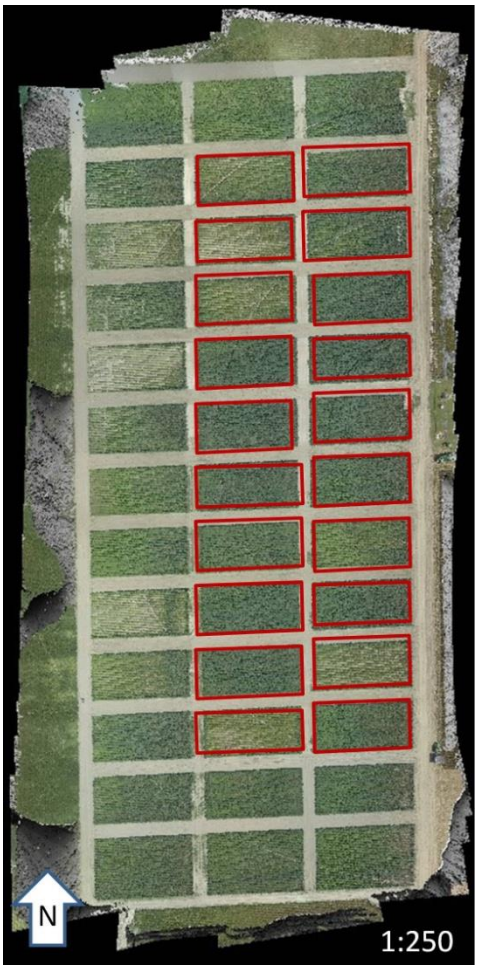


Figure 3.1. Orthographic view of the field RGB point cloud constructed from the images acquired on August 28, 2014. The plots that were study for cotton yield estimation are framed in red.

Color images were taken of the cotton field on August 28, 2014 during the open bolls growth stage and on October 1, 2014 before the harvest operation. The image acquisition was accomplished via a customized RTF X8 octocopter from 3D Robotics. The X8 is a miniature UAV with an 800 gram payload and a flight time of 15 minutes. The X8 is equipped with 2-axis Tarot T-2D brushless gimbal which is designed to stabilize the GoPro HERO camera during flight. A Pixhawk autopilot module is used to control the X8 and allow it to achieve

preprogrammed flight paths. The Pixhawk is equipped with and IMU which allows the gimbal system to make adjustments while the UAV maneuvers through its flight path. The camera used for image acquisition is a 10MP GoPro HERO3+ camera equipped with a 2.97mm f/4.0 non-distortion lens. The lens has a 95^{o} of angle of view (AOV) which allows the camera to achieve a 3 cm $pixel^{-1}$ ground spatial resolution at a flight altitude of 50 m. The total cost of the UAV system is less than $1,500, making it an affordable and economic remote sensing system for farmers.

3.1 Cotton Data Acquisition

In developing a system for data acquisition of the cotton, considerations of the best viewpoint for which the cotton bolls could be fully observed as much as possible was made. The physiology of the full grown cotton plot and its distribution of bolls was a key constraint in the implementation of a robotic system for data acquisition. Ground level views of cotton fields facilitate a lot of occlusion of cotton bolls from the sensor. Because of the fruiting behavior of the cotton plants an aerial view was chosen for observing and measuring cotton bolls. There are many aspects of aerial views of the cotton plants that were considered in the detection of the cotton bolls.

A full grown cotton plant consists of a main stem, fruiting branches, and, when conditions are met, vegetative branches. The main stem produces stem leaves according to a $3/8^{th}$ alternate phyllotaxy. The site on the main stem at which the main stem leaves grow is referred to as the axil. The fruiting branches grow from the same set of axils right above the main stem leaves. At the top of the main stem is a growing point referred to as meristems. This allows the main stem to grow upward and facilitates new points for branches to grow. Fruiting branches are predominantly grown and are composed of cotton bolls and subtending leaves.

Fruiting branches start as single branches with one fruiting bud and a subtending leaf. The subtending leaf provides carbohydrates for the fruiting bud to grow. At the end of fruiting branches are other growing points called axillary meristems. At the axillary meristems, fruiting branches grow new stems with new subtending leaves and fruiting buds. Consequently, the fruiting branches grow outwardly in a "zig-zag" structure with sections of subtending leaf and fruiting bud pairs. The cotton plant grows upwards and outwards in this fashion with the bottom most branches being the longest with more fruiting points and subtending leaves. Vegetative branches grow below the first leaf node on the main stem. These branches only have one meristem and grow straight and upward. New fruiting branches may growth on vegetative branches as well, similar in the manner at which they grow on the main stem with the same phyllotaxy. Vegetative branches occur due to several conditions such as over-fertilization, low population density, and effects from insects and diseases (Ritchie, Bednarz, Jost, & Brown, 2007).

Cotton plants' ability to develop cotton bolls depends on the amount of carbohydrates produced during the photosynthesis from subtending leaves. Because the fruiting branches grow at the same axils as the leaves, they too have a $3/8^{th}$ alternate phyllotaxy. This allows the crop to optimize the amount of sunlight being intercepted by the subtending leaves. The amount of the sunlight the subtending leaves intercept is crucial to the growth and development of the bolls, in that they are a source of carbohydrates for the bolls to grow. When leaves are shaded, the growth of the development of fruiting buds into cotton bolls slows as they utilize other sources of carbohydrates from the plant. Due to the nature of cotton, the organs of the cotton plants are prioritized over the fruiting elements. The less mature fruiting elements are aborted if they create demand on the carbohydrates that overly compete with organs. Competition with

neighboring cotton crops also affect the growth and development of individual plants, as well as their fruiting elements. Neighboring plant branches and leaves can shade leaves (Ponstable et. al. 2010).

Figure 3.2. Patch of cotton field.

In order to harvest the cotton, the crops are sprayed with chemicals so that they undergo defoliation. After defoliation all the cotton branches become dry and adopt a darker brownish color. All of the main stem leaves and subtending leaves fall off of the branches and onto the ground soil. The leaves also have a relatively dark brownish color. Consequently, the ground of

the cotton fields are a mixture of soil and leaf coverage, colors and textures. The leaf coverage depends very much on the amount of leaves each cotton plant had, making the ground landscape have various mixes of the dead leaf and soil concentrations.

In the acquisition of images of the cotton field, aerial views of each plot were taken. With the fruiting branch arrangement being according to a phyllotaxy, the cotton bolls arranged in a manner in which there is little occlusion from a topological view or nadir view. The nadir view of the camera during aerial acquisition was accomplished via a two-axis (roll and pitch) gyroscope. For the experiment, the cotton field was composed of thirty-nine individual cotton plots arranged in a three by thirteen lattice. The UAV used to transport the camera followed a flight path which allowed the camera to capture an image of each plots. The flight path was set at an altitude and speed that allowed the UAV to maintain a low flight duration. Each image where the plots were positioned near the camera optical center was used for signal extraction. This was to minimize distortion in the images of each plot due to the low-distortion lens as possible. The images used were therefore from a nadir perspective with minimal distortion.

The camera view of the field is relatively nadir with the angle between a perpendicular viewpoint and its view vector becoming less significant with altitude. The configuration of the crop allows its bolls to be optimally exposed to the sunlight. It is possible to for individual cotton bolls to have overlapping boundaries at the ground level. There; however, should be some line of sight of some portion of each cotton boll based on phyllotaxy and fruiting behavior due to competition between neighboring plants. The spatial resolution is governed by the altitude and the resolution of the camera sensor. Given the proximity of cotton bolls to each other, the image of individual cotton bolls can be expected to merge together as the altitude of the camera increases. The texture of the background due to leaf and soil concentrations become

smoother as the spatial resolution decreases. The background color is expected to mix as the spatial resolution decreases. Given the altitude required for large field acquisition, the images of the cotton are expected to be a mixture of individual and clustered bolls with soil/leaf mixture backgrounds. Fine details such as cotton bolls edges are also expected to be lost while being slightly mixed with the relatively darker background.

4. Cotton Yield Estimation Phase I

This section will present the first phase in the image analysis of cotton plants and the development of a cotton yield estimation method using image sensor observations at harvest time (Huang et. al., 2016). In this study, assumptions were made regarding the cotton plant image signal leading to the development of cotton yield estimation methods. These assumptions were evaluated based on the results of the corresponding estimation methods relative to the manually recorded cotton yield measurements. First the assumption presented will be an assumption on physiology of crop where the vegetative structure of the crop is observed at harvest time. The results of this method will then be presented along with the motivation for cotton boll observation based methods. A metric for cotton yield estimation will be presented where the cotton field is observed by the camera sensor after defoliation. This metric took cotton signal extraction results as an input. Methods for cotton signal extraction along with governing assumptions will be presented. The yield estimation results of these methods will also be presented along with motivations for the second phase of the investigation.

4.1 Yield Estimation via Height Observation

Images of the cotton field prior to defoliation were taken. During the image acquisition, the camera maintained a nadir view of the field, and thus captured images of each region of the field at multiple viewing angles with each pass. The images were geotagged with coordinates from an onboard GPS module. The coordinates were used to develop a metric three-dimensional reconstruction representing the surface elevation of the field. This reconstruction was in the form of an RGB point cloud and is illustrated in Figure 4.1.

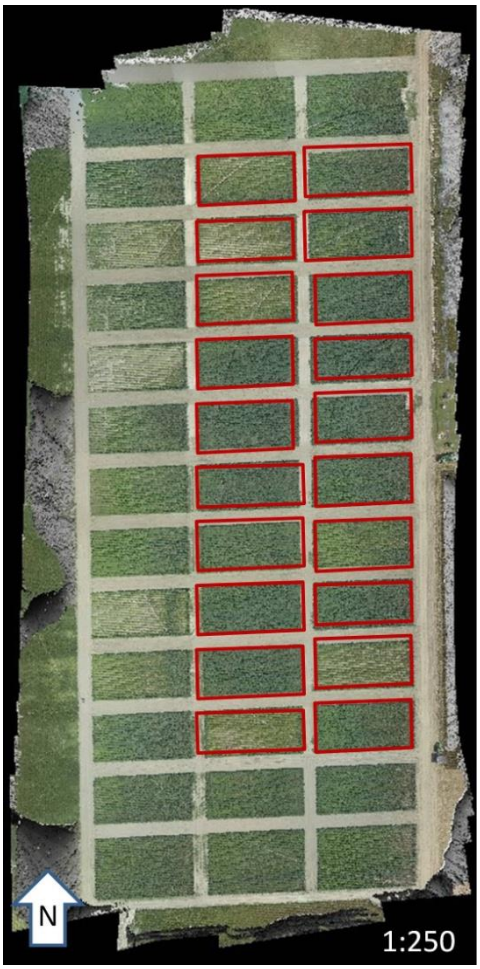


Figure 4.1. Orthographic view of the field RGB point cloud constructed from the images acquired on August 28, 2014. The plots that were study for cotton yield estimation are framed in red.

An assumption can be made that the cotton is proportional to the height the of the cotton canopy. The height may be extracted from the three-dimensional data by identifying the ground terrain and finding the difference of the terrain model from the full elevation model (McDaniel, 2010). The RGB point cloud is a mapping of the elevation model represent in three-dimensional space to RGB values. The terrain elevation model could therefore be estimated by identifying point clouds with RGB values pertaining to soil. The soil points could be extracted via an RGB

threshold operation on the full RGB point cloud, and then used to estimate the terrain elevation model. Let E^f be the full elevation model $E^f: x, y \in X \rightarrow \mathbb{R}$. The RGB point cloud can be modeled as a function $S: \mathcal{V} \rightarrow \mathbb{R}^3$ where $\mathcal{V}$ is the three-dimensional representation E^f according to $x, y, E^f(x, y) \in \mathcal{V} \subset \mathbb{R}^3$. Thus the S is the function that maps the three-dimensional representation of the elevation model $\mathcal{V}$ to an RGB value. The set of terrain elevation points in $\mathcal{V}$ can be represented as a subset $v \subseteq \mathcal{V}$ where $x, y, E^t(x, y) \in v$. It follows that subset v is mapped to soil RGB values by S. The set v can, therefore, be identified through a threshold operation on S for soil points resulting in a masking function $T: v \rightarrow \mathbb{R}^3$ representing the terrain RGB point cloud. This threshold operation is shown in Equation 1.

$$T = \begin{cases} S & S > (t_R, t_G, t_B) \\ \emptyset & otherwise \end{cases} \quad (1)$$

The threshold values t_R, t_G, and t_B on the red, green, and blue channels of the point cloud divide the terrain points from non-terrain points. $\emptyset$ denotes an empty space in which the operation removes the points that don't pass the threshold tests. The resulting terrain points T is illustrated in Figure 4.

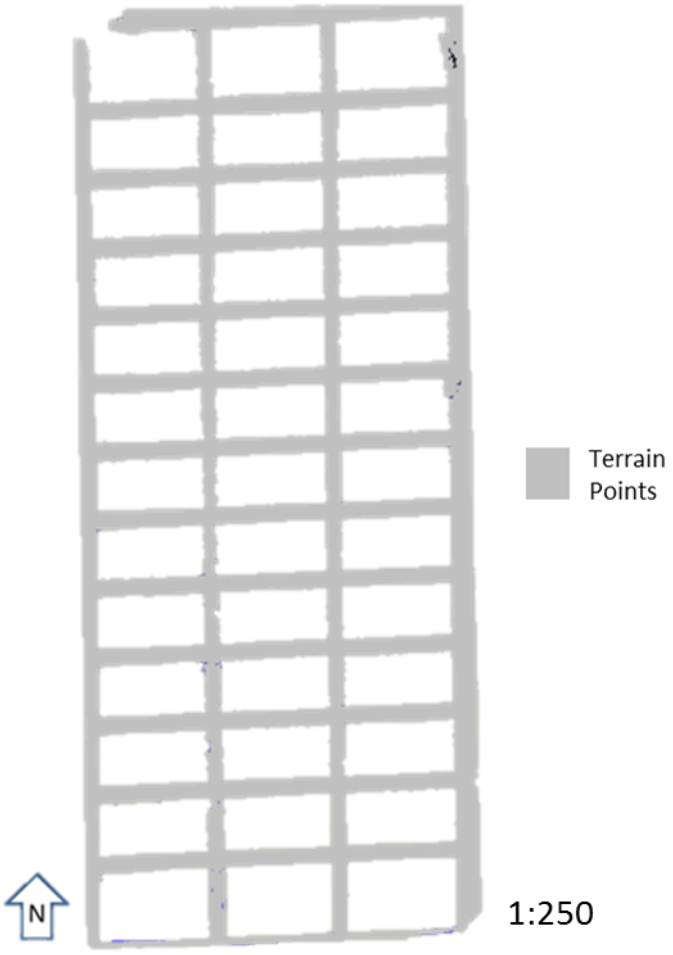


Figure 4.2. Orthographic view of the resulting terrain point cloud T (stray points within are removed).

The terrain elevation subset v was distributed throughout the entire space of $\mathcal{V}$ including the boundaries. An assumption can be made that the overall behavior of the terrain model is captured in v. Therefore, the full terrain elevation model E^T can be estimated through an interpolation of E^t on a prescribed grid space G. An interpolation of E^f on G results in an estimated elevation model E^F on grid space G.

The terrain elevation model E^T was estimated with the use of a linear interpolation of E^t on G. As illustrated in Figure 4, E^t contained significantly large open areas where the crop points occupied. The shape of E^t also exhibited a slightly upward trend towards the crop plot boundaries. Higher order interpolation functions are sensitive to such trends and cause undesired artifacts in an attempt to estimate points in large gaps of missing data. After linear interpolation a Gaussian filter is used to filter Gaussian noise and step artifacts (Eckert, 2008), resulting in the final terrain elevation E^T. The elevation model E^F was estimated via a nearest neighbor interpolation of E^f on G. A nearest neighbor interpolation was sufficient in that it preserved the shape and boundary conditions of each plot in E^f. The height model H was found via a difference taken between E^F and E^T. The height model is illustrated in Figure 5.

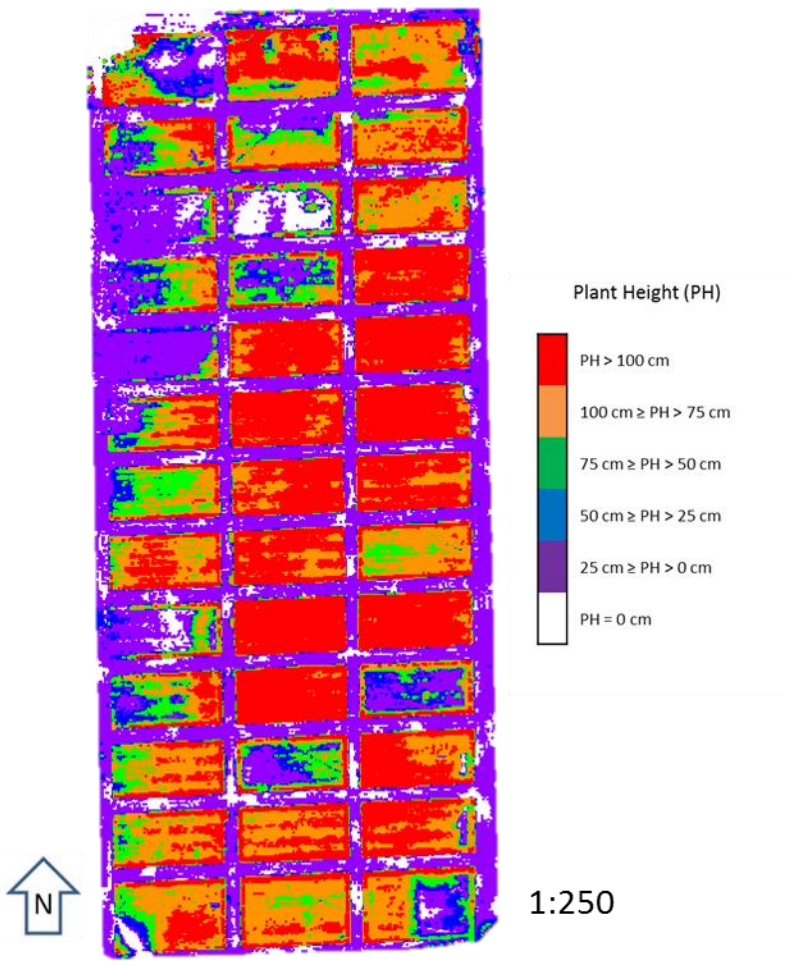


Figure 4.3. Estimated plant canopy height profiles of all plots in the cotton field.

The estimated heights were plotted next to the manually measured heights of fields 103-112 and 203-212. This is illustrated in Figure 4.4 and is observed to have a high linear correlation. This illustrated the potential for estimating crop height via camera sensor.

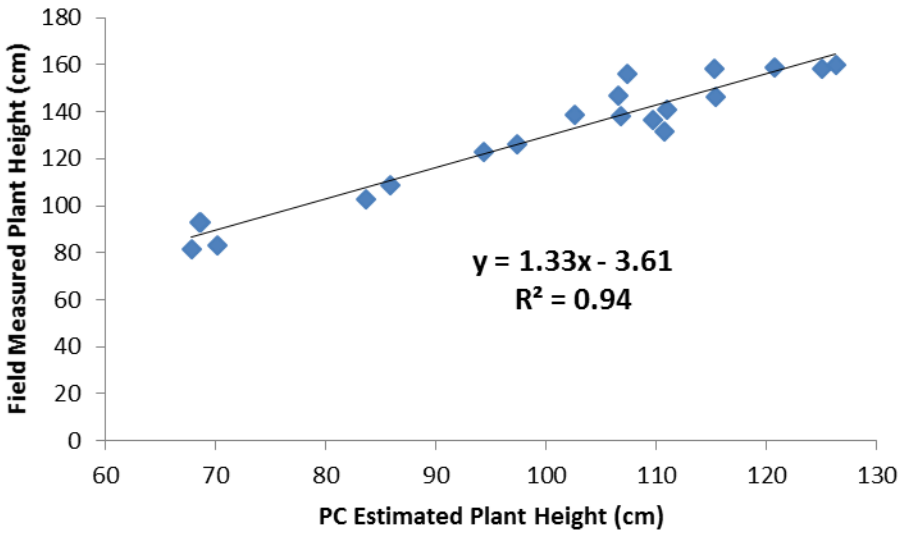
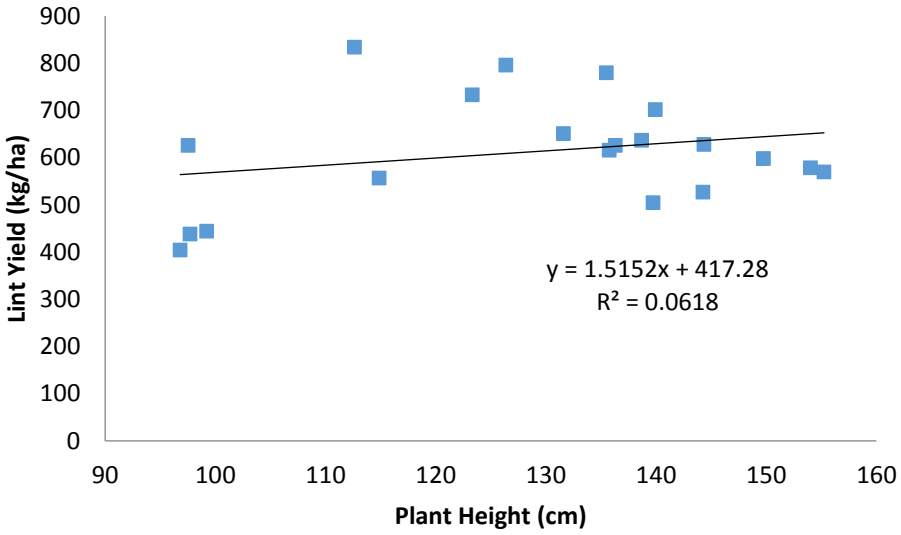
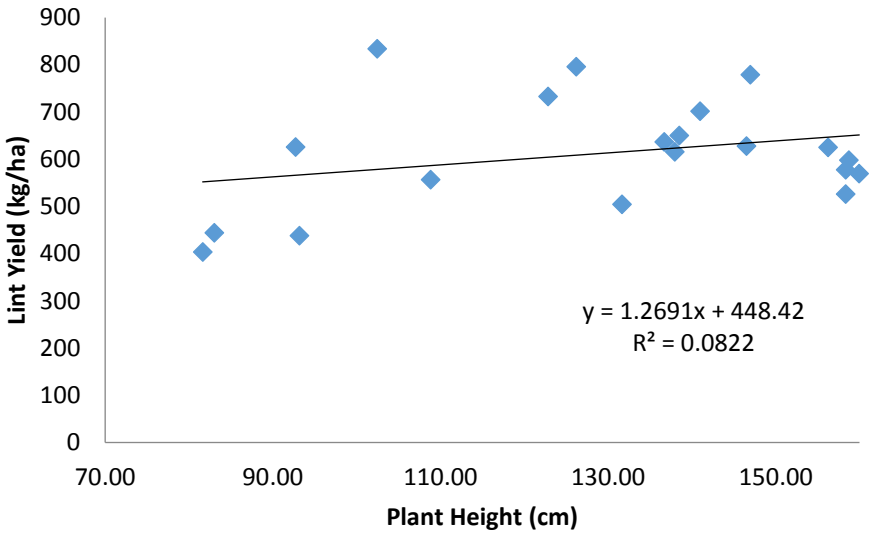


Figure 4.4. Relationship between the PC-estimated plant height and the ruler-measured plant height in field.

The manual heights and the estimated heights were plotted against yield. This is illustrated in Figure 7(a) and 7(b) respectively.



(a)



(b)

Figure 4.5. Relationships between the camera estimated cotton plant (a) and field ruler-measured (b) height and the lint yield.

The correlations of the height, whether collected manually measured or measured using the camera sensor, are low. This is due to the growth and fruiting behavior of the cotton being largely dependent upon the physiological competition and changes in the environment. For instance, cotton can experience a lot of the vegetative growth due to over-fertilization, row spacing. Too much vegetative growth can cause cotton to abort more of their young due to competition of carbohydrates. Also other changes in the environment such as temperature or humidity can cause cotton plants to abort their young as well. The physical features of the plant at the final growth stage alone is insufficient for developing an estimation of the cotton yield (Ritchie et. al, 2007).

4.2 Cotton Coverage Measurement and Field Survey

To estimate the yield, the cotton signals have to be extracted accurately from the images. With the extracted cotton signals, the cotton boll plot coverage can be established, which is defined as a ratio for a cotton unit coverage (CUC) in each plot:

$$CUC = \frac{NoCP}{NoPP} \tag{2}$$

where NoCP is the number of cotton boll pixels in a plot; NoPP is the number of pixels in the plot area. The following sections describe the challenges and solution to accurately extract cotton signals from the images.

A survey was conducted to ascertain the typical features of the defoliated cotton field. In this survey, images of rows of a plot were taken. Figure 4.6 displays an image of a typical portion of the cotton plot. In the image, the cotton rows can be observed to be arranged according to a spacing where irrigation troughs exist in between rows. Plant stems and branches are dead and take on darker brownish color. The plant leaves are also dead and have fallen off of the cotton plants. The dead leaves have a light brownish color are shown to collect mostly underneath the plant and in the troughs. This leaves the bolls exposed while they are arranged according to the plant phyllotaxy. The soil can also be seen to have a relatively grey color. This provides insight on the expected aerial image of the cotton fields. The cotton can be expected to exist in a foreground while the dead branches, dead stems, dead leaves and the soil exist in the background. The dead leaves take up more surface area. Therefore, the background around cotton bolls can be expected to vary based on the amount of vegetative growth of the plant. A heavy leaf based background would be relatively darker in color compared to a heavier soil

based background. The problem being solved in this investigation is to develop a method for extracting the cotton boll foreground pixels in the midst of the non-cotton background pixels.

Figure 4.6. Survey of cotton field: a section of a typical cotton plot with close-up trough and features beneath cotton plant.

4.3 Cotton coverage estimation via RGB observation

The field survey images also reveal that the color of the cotton bolls are relatively bright compared to the background objects. The simplest method for signal extraction to consider would be a simple threshold on the image intensity in the red, green, and blue channels. An RGB image can be modelled as a real-valued function describing radiant flux or intensity in the red, green, and blue spectra emitted from the objects onto the image plane $f: \Omega \to \mathbb{R}^3$ where $\Omega \subset \mathbb{R}^2$.

The image sensor discretizes the image plane Ω into a two-dimensional grid X and quantizes the f into gray-levels. The discrete function of f realized by the image sensor can therefore be denoted as $I(x, y)$, where $x, y \in X \subset \mathbb{R}^2$ denote grid coordinates and the value of I are image gray-levels based on the bit of the quantization. The grey-level depends on the radiant flux density emitted from the light source, the reflectance properties of the objects in the scene, the ISO sensitivity of the camera, the shutter speed, and the angle of view of the camera. It can be assumed that due to the relatively bright, white color of the cotton, its grey-level in I should be relatively larger compared to background.

A threshold operation on I begins by selecting threshold values t_R, t_G, and t_B in the red, green, and blue channels respectively that will divide cotton pixels from pixels from non-cotton pixels. A comparison of each pixel with the threshold values are then performed. Depending on the outcome of the comparison the pixels are labeled 1 or 0. This is illustrated in Equation 3.

$$M(x, y) = \begin{cases} 1 & I(x, y) > (t_R, t_G, t_B) \\ 0 & otherwise \end{cases} \quad (3)$$

The result $M(x, y)$ is an image $M: I \longrightarrow [0,1]$ where pixels pertaining to cotton bolls in $I(x, y)$ are assigned a value of 1 at the corresponding pixel location in $M(x, y)$ while other non-boll pixels are assigned 0. In the case of the cotton signal a threshold operator as described in Equation 2 can be applied to the images using threshold values of high intensity. In performing this operation, the assumption is that the cotton signals as a whole have characteristically higher gray-levels than the other non-cotton signals captured in $I(x, y)$. Figure 4.7 shows the resulting masks of different thresholds along with the original image.

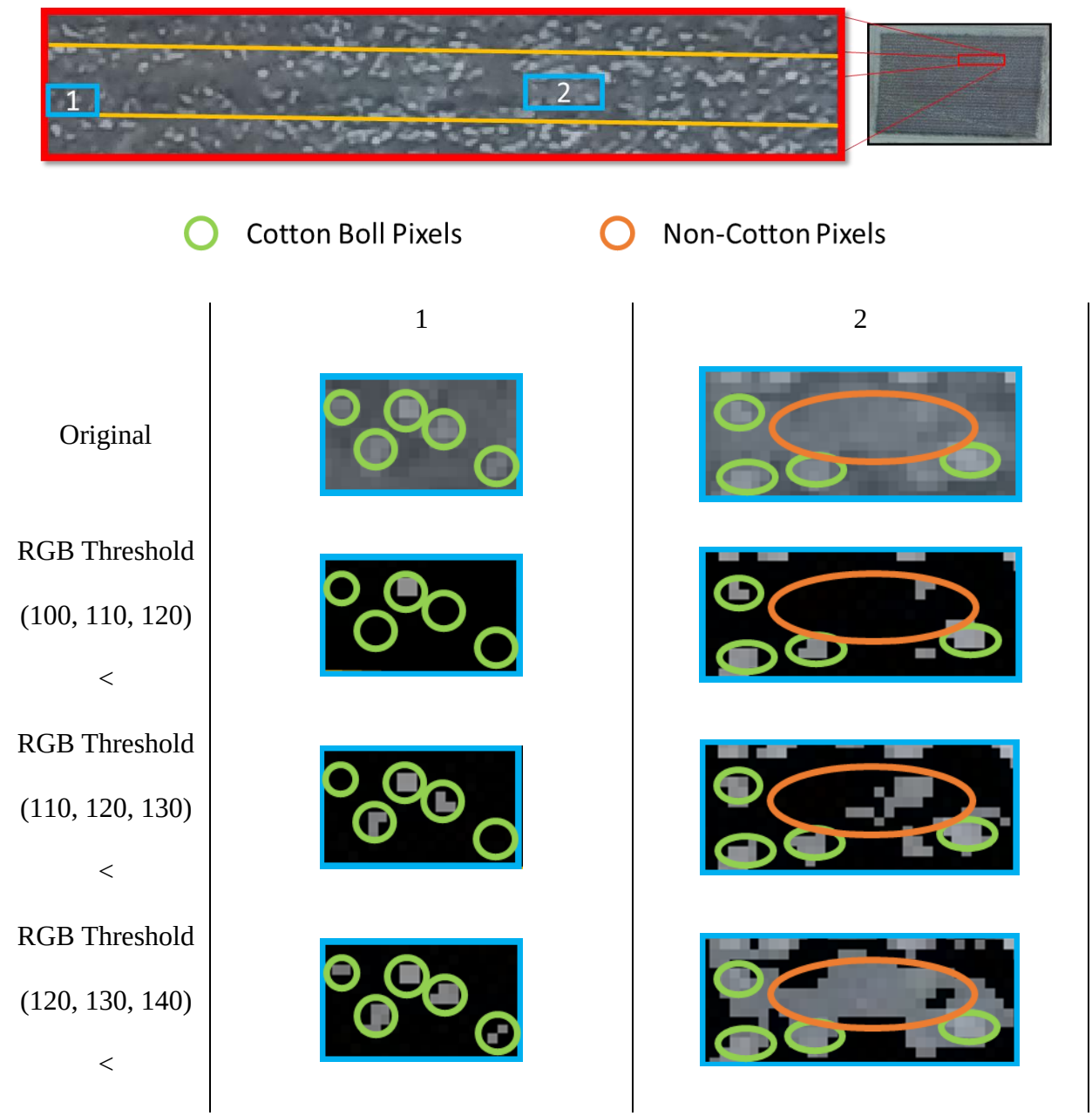


Figure 4.7. RGB thresholded regions using (100, 110, 120), (110, 120, 130), and (120, 130, 140) threshold values.

In the figure above, an image of a plot is shown where a region of interest (boxed in red) is expanded. In the expanded image, the row centers are marked with yellow lines. Two sub-regions of interest are boxed in blue and are numerically labeled 1 and 2. These sub-regions are

expanded with cotton and soil regions identified with green and orange ellipses respectively. The results of the threshold operations performed on the original using threshold value sets, (100, 110, 120), (110, 120, 130), and (120, 130, 140) are shown at the regions 1 and 2 and are compared to origin image. The threshold results are displayed from most conservative to least conservative threshold value sets. From the figure, it is evident that the assumption on characteristic grey-levels of the cotton is violated. The range of the cotton bolls' grey levels intersects that of the soil. A trade-off can be observed between false positives and false negatives. Less conservative threshold values result in masks which label regions of the background as cotton signals. This can be observed in Figure 4.7. As threshold values become less conservative, the more soil pixels are accepted. These false positives exist as large clusters and potentially connect to cotton signals. This is illustrated in sub-region 2 of the (120, 130, and 140) threshold results in the non-cotton region. Utilization of an area filter to remove non-soil cluster results also remove cotton signals. More conservative threshold values result in masks that remove a significant portion of lower grey-level cotton pixels as false negatives, while not labeling background points as false positives. This is illustrated in Figure 4.7 in the threshold results for sub-region 1. The more conservative the threshold value sets are, the more cotton signals are removed from the cotton regions. The range of cotton grey-levels is due to the incident light illuminating the cotton bolls not being consistent across the field. This is due to several factors such as variations in the sunlight caused by clouds, structure of neighboring plants, light source angle. In applying a threshold operator on the cotton field, the chosen threshold value needs to satisfy a trade-off between the cotton signal removed from area filtering and the cotton signal removed as false negatives. Such a trade-off value is different for each image of a part of the field. It is therefore difficult to choose values that produce consistent boll

and non-boll separations. An investigation purposed towards analyzing the effectiveness of the RGB threshold for cotton signal extraction was conducted. The threshold values for the red, green, and blue channels were chosen manually for each image. Figure 4.8 illustrates the yield results where color thresholding at (110, 120, and 130) was used for cotton signal extraction.

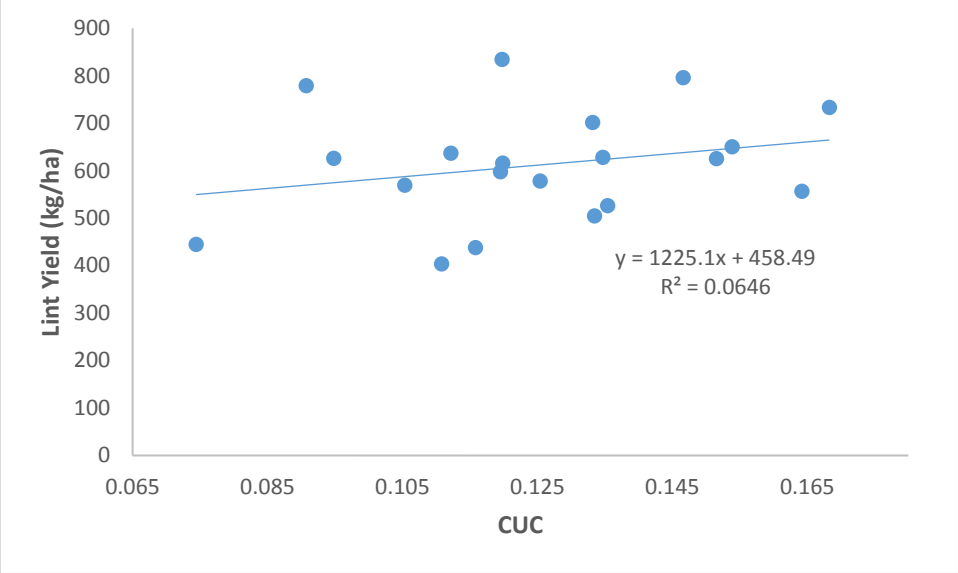
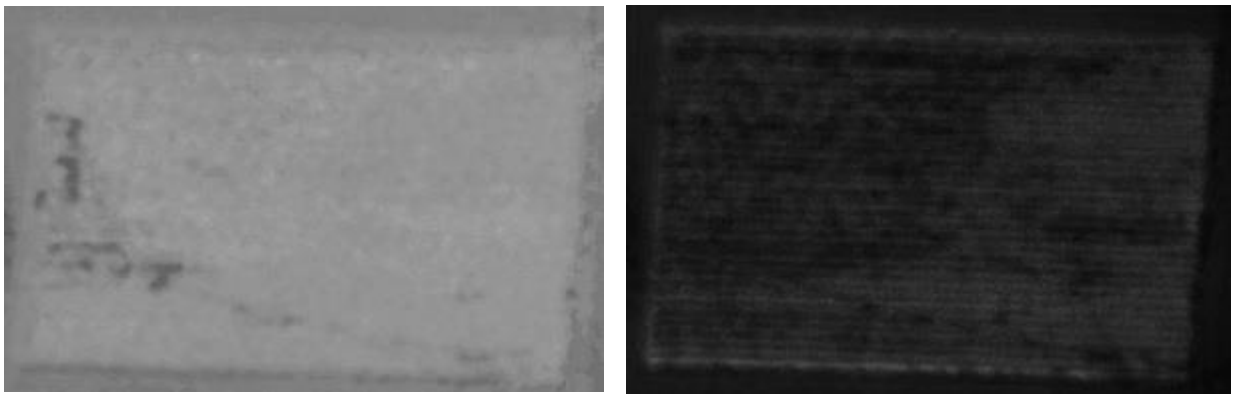


Figure 4.8. The relationships between the estimated CUC, and the measured lint yield using direct pixel intensity thresholding.

From Figure 4.8 it is evident that the assumption of the cotton signal being globally bright than the background is violated. While the cotton is relatively white and have large grey-levels of intensity on all channels of color images, their gray-levels are not characteristically unique enough to overcome global variations in illumination in the presence of bright background regions such as soil. The background histograms, therefore, intersect the cotton signal histograms.

The RGB color space is sensitive to changes in intensity. This is due to the chromaticity and brightness being coupled in the RGB description of color. Selecting a more inclusive range in the RGB space potentially leads to an inclusion of other undesired colors. This is usually managed by employing other color spaces that decouple the chromaticity and the brightness in their description of color. The selection would then be performed based upon the chromaticity, which is less sensitive to changes in illumination. These color spaces include HSV, YCrCb, Lab, etc…. These spaces are commonly used in many fruit and vegetation detection frameworks (Linker et. al. 2012, Patel et. al. 2012, Payne et. al. 2013, Payne et. al. 2014). In the HSV color space chromaticity is described in terms of hue (H) and saturation (S) while intensity or grey-level is measured on a value (V) axis. The cotton field as described by the hue and saturation is illustrated in Figure 4.9.



(a) (b)

Figure 4.9. (a) Hue and (b) Saturation of the original image of the cotton plot as shown on Figure 4.7 (a).

The hue image displayed in Figure 11(a) is very constant. The cotton boll pixels have no unique or distinguishable hue which could be used to separate them from the background pixels. From

Figure 4.9 (b) the overall saturation of the patch is observed tp be low. This indicates that the fields have very neutral colors. Though there is a noticeable variation in the saturation plane, the saturation of the cotton pixels is not unique or characteristic. This illustrates the dependence of the cotton image on intensity and the difficulty of employing intensity independent color descriptions for the cotton field images.

4.4 Cotton Coverage Estimation via Laplacian of Grey Level Observation

Though the cotton as a collective are not observed by the camera as characteristically brighter than the background pixels, the intensity of the cotton boll pixels can be observed to be locally brighter than neighboring background pixels. The cotton pixels cause large discontinuities in the intensity space. This can be observed in the Laplacian of the intensity of the grayscale image $I^g: x, y \in \Omega \to \mathbb{R}$ where only accelerated rates of change intensity space are observed. The discrete Laplacian $\mathcal{L}^D$ of I^g the sum of the partial second derivatives of I^g with respect to x and y and is illustrated in Equation 4.

$$\mathcal{L}^D(x,y) \approx \frac{I(x+\triangle x,y)-2I(x,y)+I(x-\triangle x,y)}{\triangle x^2} + \frac{I(x,y+\triangle y)-2I(x,y)+I(x-\triangle x,y-\triangle y)}{\triangle y^2} \qquad (4)$$

The Laplacian is accomplished through a convolution of image I^g with Laplacian operator $\mathcal{D}^2$. This is illustrated through Equation 5.

$$\mathcal{D}^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{5}$$

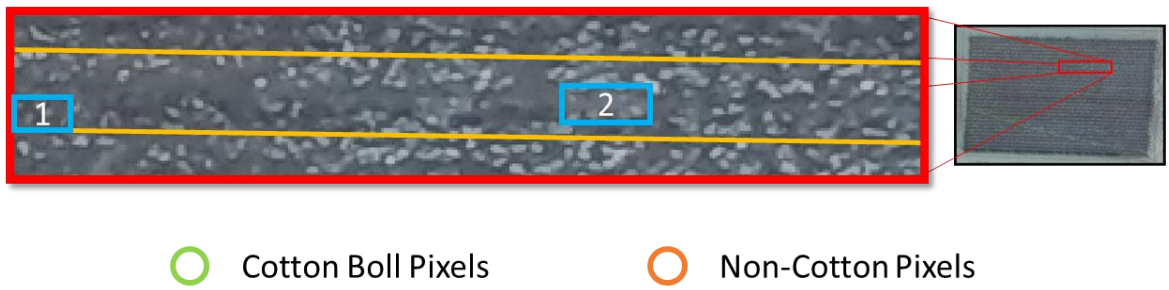
$$\mathcal{L}^D(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I^g(x,y) * \mathcal{D}^2(x-m, y-n)$$

It can be assumed that, because of the cotton pixels relative intensity, the cotton has a characteristic signature in $\mathcal{L}^D$. According to this assumption, a threshold operation on $\mathcal{L}^D$ can be applied with a threshold value t_L to develop a mask $M: \mathcal{L}^D \longrightarrow [0,1]$ for the cotton. The threshold operation is shown in Equation 6.

$$M(x,y) = \begin{cases} 1 & \mathcal{L}^D(x,y) < t_L \\ 0 & otherwise \end{cases} \tag{6}$$

The discontinuities cause by the cotton signal cause local maxima in the intensity. The threshold operator in Equation 5 is designed to local maxima of a certain rate. The Laplacian, being a second derivative operator, is known to be sensitive to noise. A bilateral filter is used to remove noise while preserving the shape and discontinuity relationships of the cotton signals. The Laplacian of the image resulting from the bilateral filter is then obtained. The characteristic signal of cotton is less sensitive to changes in illumination due to inter plant occlusions. The lower intensity cotton boll pixels can be detected without falsely label background pixels as cotton pixels. This is illustrated in Figure 4.10 where the results of the Laplacian threshold operation on sub-regions 1 and 2 of Figure 4.7 are compared to the results of the RGB

thresholding operations on sub-regions 1 and 2 using threshold value sets (100, 110, 120) and (120, 130, 140).



	1	2
Original		
RGB Threshold (100, 110, 120) < RGB Threshold (120, 130, 140) < Laplacian 15 <		

Figure 4.10. RGB thresholded regions using (100, 110, 120) and (120, 130, 140) threshold values and Laplacian Threshold with a threshold value of 15.

In summary, to extract cotton signature pixels from each plot, with the original image a binary mask image was generated to have signatures of cotton bolls. Then, the plot boundary mask was also generated to combine with the cotton boll signature mask to generate the plot cotton boll signature mask. Once this was accomplished a convex hull could be formed around the cotton boll signature pixels in the plot. With the results, the cotton boll signature pixels were counted for estimation of the CUC for the plot. Eventually, with the estimated CUCs for all the plots under study, the cotton yield was estimated. Figure 4.11 shows the work flow for formation of convex hull around the cotton boll signature pixels in a plot.

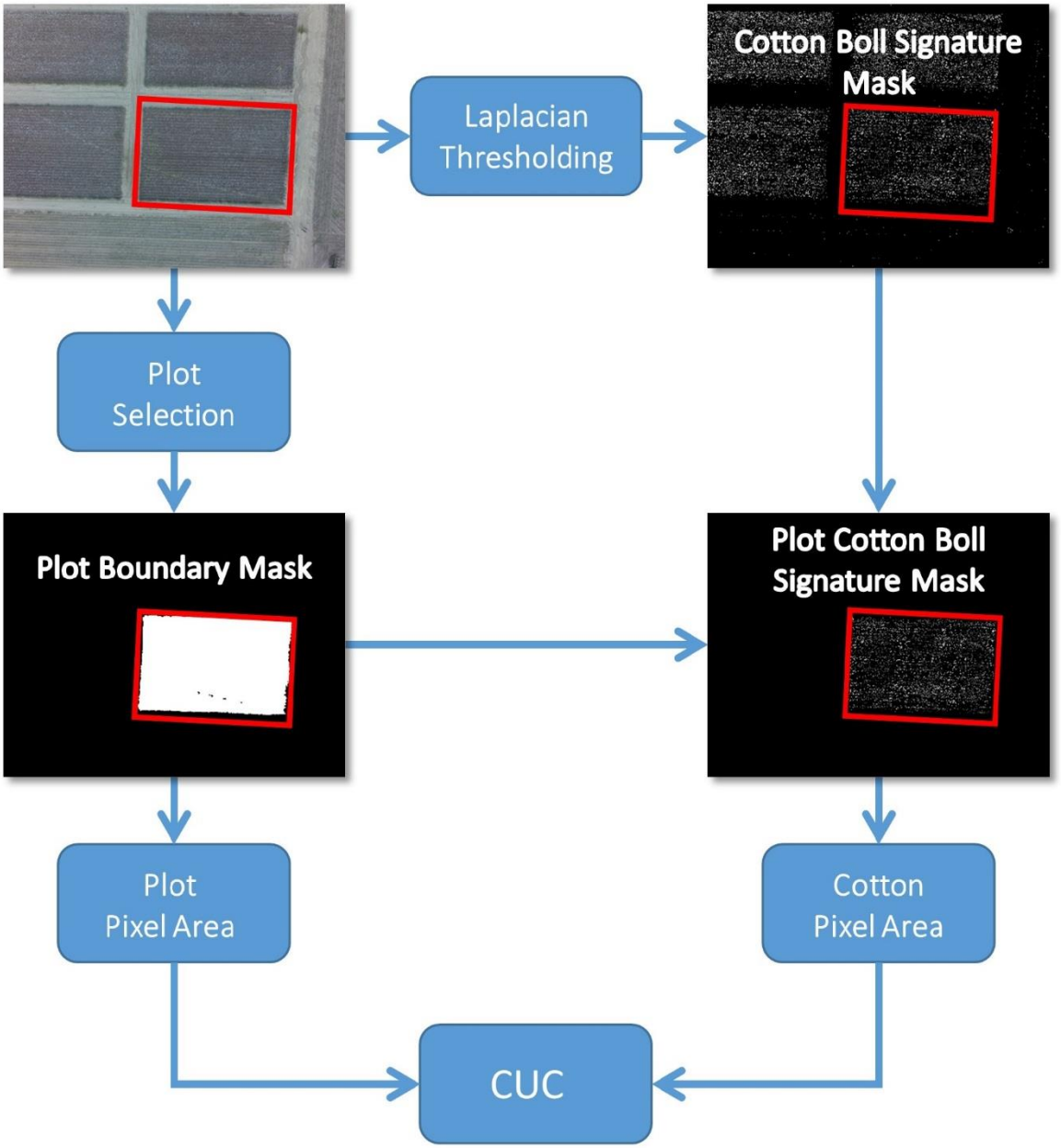


Figure 4.11. Work flow for formation of convex hull around cotton boll signature pixels in a plot.

Figure 4.12 shows the resulting relationships between the estimated cotton plot unit coverage (CUC) based on thresholding the Laplacian of the intensity space and the measured lint yield in kg/ha. A divergence gradient value of 15 was used as the threshold value. From Figure 14, it is

evident that the Laplacian based threshold method provides a significant improvement in observing the cotton signal than the RGB thresholding based method.

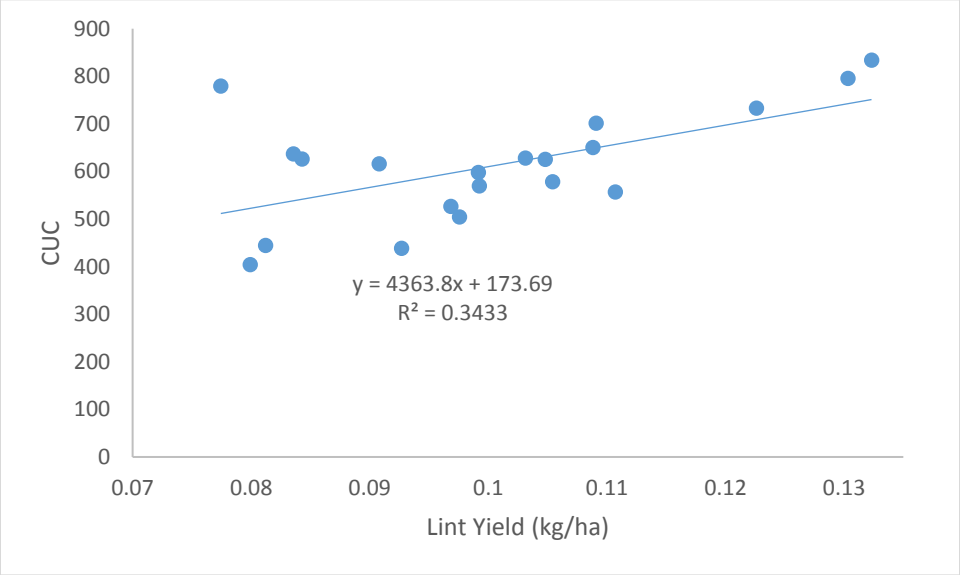


Figure 4.12. The relationships between the estimated CUC, and the measured lint yield using Laplace thresholding.

An inspection of the plot in Figure 4.12 reveals a potential outlier. The studentized residuals of the points with respect to the regression line were taken and plotted alongside the CUC values. This is shown in Figure 4.13.

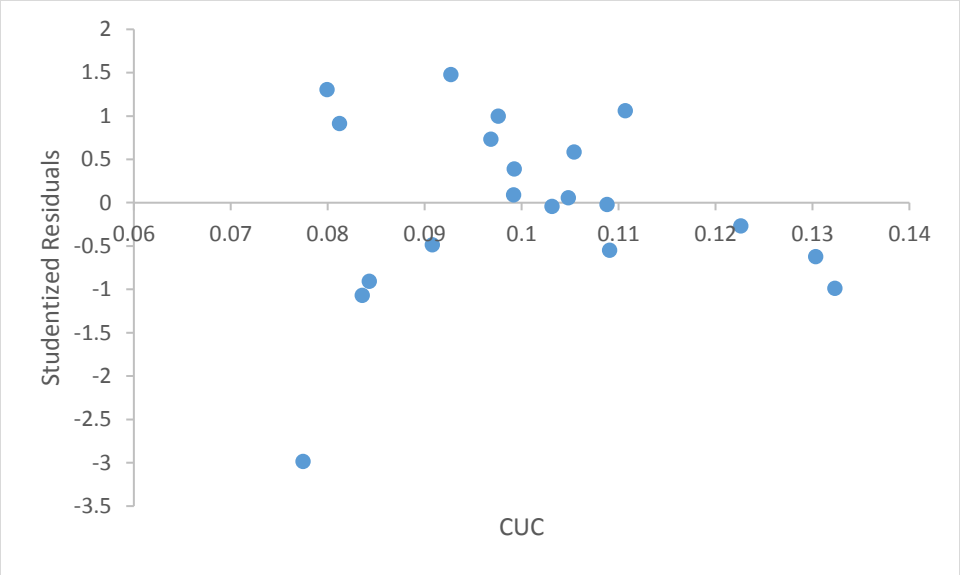


Figure 4.13. Studentized Residuals of Laplacian based cotton unit coverage (CUC)/lint yield regression

A two-tailed t-distribution at a 95% confidence interval indicates that a studentized residual beyond 2.093 is indicative of an outlier. An inspection of Figure 4.13 reveals an outlier at a residual close to 3 (2.98). This point is indicated as an outlier and corresponds with the point 779.2870421, 0.077438 on the plot in Figure 4.12. Considering the high leverage of the outlier suggests a potential for masking other outliers. A robust regression analysis is performed using random sampling consensus (RANSAC). The results of the robust regression is shown in Figure 16.

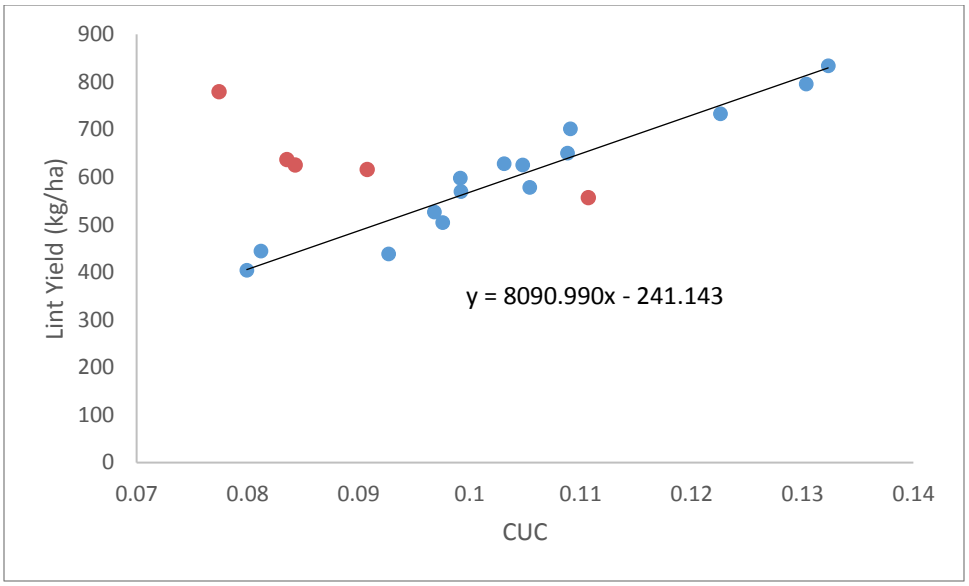


Figure 4.14. Robust regression analysis of lint yield and the Laplacian based cotton unit coverage estimates using random sampling consensus (RANSAC).

The RANSAC algorithm identified five outliers in the data set and developed and fitted regression line to the inliers in the data set. Plotting the residuals of the RANSAC regression reveals the largest residuals occurring as underestimation of the CUC estimates. Plotting absolute residuals alongside the mean intensity levels reveals that the largest residuals occurred at and around the lowest darkest plot. This illustrated in Figure 4.15.

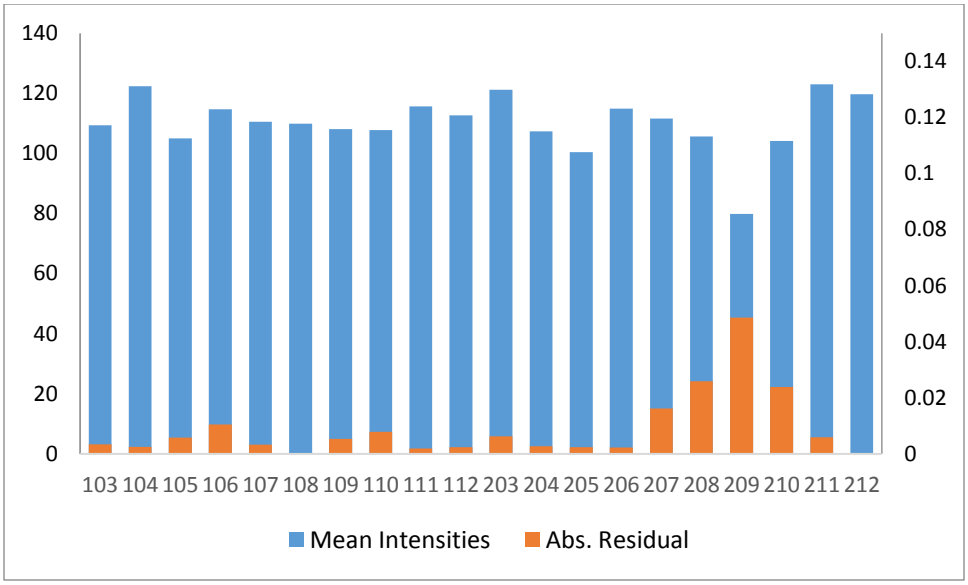


Figure 4.15. Mean intensities and Absolute RANSAC residual for each field plot.

An inspection of the image and mean value reveal that the plot was under shadow which spanned across sections of the surrounding plots. This can be noted in the downward trend around the dark plot. This suggests that the illumination levels are low enough to significantly affect the contrast relationships between the cotton and background signals and therefore the second derivative characteristics.

4.5 Summary

In conclusion, the preliminary investigation revealed that the cotton pixels cause unique characteristic behaviors in the second derivative space of the cotton field images. The cotton was observed to have a very neutral color with very low saturation. The background or non-cotton pixels were mostly composed of soil and dead vegetation pixels. The background pixels were also of neutral color with very low saturation. This was especially the case for flat images caused by cloud coverage impeding the sunlight. This indicated that the appearance of cotton in the image very much relied on the brightness or intensity of the cotton. Chromaticity, therefore,

could not be used to differentiate the cotton from the rest of the scene. It was found that cotton did not have a globally unique characteristic intensity within a given image due to the variation in the incident light available for each cotton boll. Therefore, intensity based methods for detecting the cotton was subject to error, allowing a significant amount of false positives and false negatives to corrupt the detection results. It was also found that cotton plant height does not highly relate to lint yield. This illustrated the difficulty relating structural features to lint yield due to the complexity of the fruiting in cotton plants. Plots imaged under low illumination experienced lower contrasts. This reduced the Laplacian output of the cotton signals making them lower than the global threshold value prescribed for the field. The Laplacian, being an edge detector, was also found to favor the edges of larger cotton boll clusters. The centers of large cotton boll clusters did not contribute to the cotton coverage results.

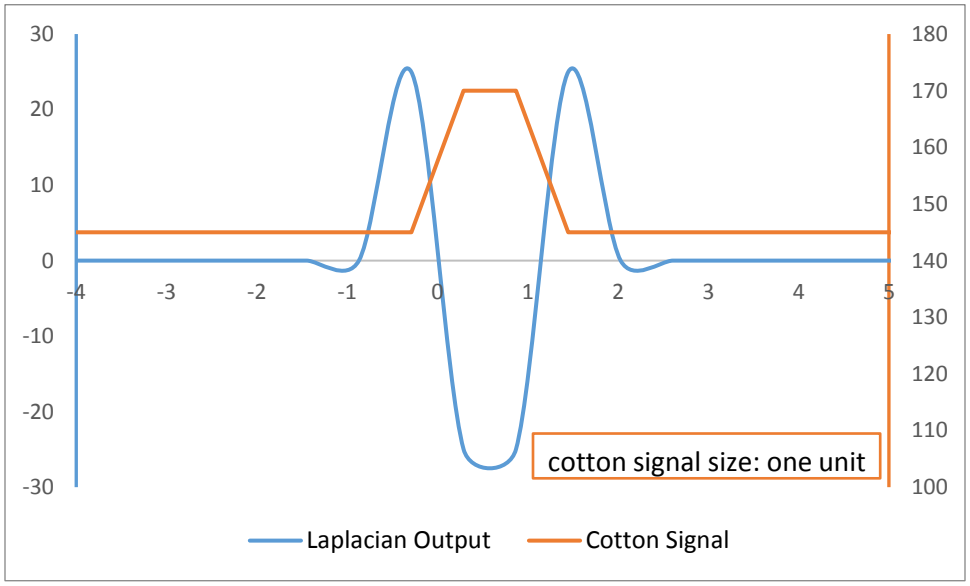
5. Cotton Yield Estimation Phase II

5.1 Laplacian of Gaussian Filter

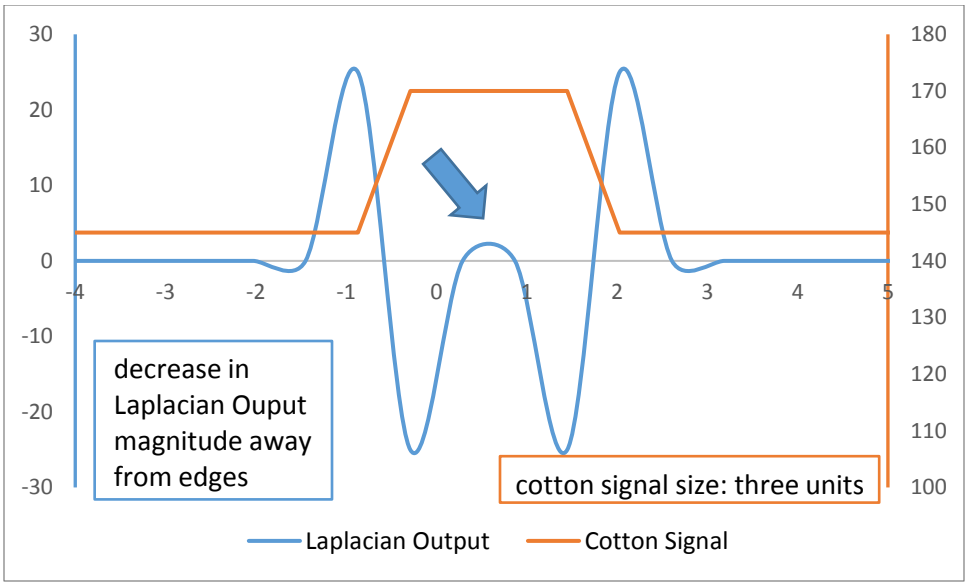
In the Cotton Yield Phase I study, it was found that the cotton signal density could be observed by the camera sensor and could represented by a cotton unit coverage (CUC) metric for a particular field plot. The CUC could be related to the lint yield of the cotton, in that they are both area densities. This metric was found by taking the ratio of the area of the extracted cotton signal by the area of interest in the sensors field of view. This operates on the assumption that areas of interest are observed at viewpoints which have an insignificant level of distortion. For cotton signal extraction an assumption was made that the brightness of the cotton boll signal was characteristically larger than that of the background signal throughout the entire set of field observations. It was determined that this assumption was violated. It was found that the cotton bolls signals existed within a range of intensities even in the presence of constant lighting conditions. The intensity range of the cotton signal intersected with that of the background signal. This was due to the soil content of the background signal and the structural distributions of cotton bolls facilitating different levels of illumination. A characteristic local brightness assumption was posed in order to manage variations illumination due to cotton boll distribution. This stipulated that the cotton signature had a characteristic second derivative signature in the intensity space. The cotton signal was therefore detected by means of a threshold on the Laplacian of intensity space. This assumption was found not to hold in case of a gradient existing in the light source. Other concerns were noted, in that the Laplacian operator detected the edges of the cotton signal due to the step-like behavior of the cotton signal after quantization. In this section an extension of the method shown in Cotton Yield Phase I will be shown where a second derivative operator which measures the second derivative characteristic of entire cotton

boll and cotton boll cluster signals will be described. Afterwards, the assumptions imposed by the operator will be generalized in order to account for variations in the illumination of the cotton field.

Extracting entire cotton boll regions via the image Laplacian was difficult in that cotton boll clusters are quantized under very similar values. Cotton boll clusters induce large plateaus in the intensity space. The discrete Laplacian operator, being a second derivative operator, returns negative values in areas of decreasing slope and zero in areas of constant slope. If the resolution of the discrete Laplacian operation is significantly smaller than the spatial size of the maximal plateau, then the output will be high around the edges of the plateau and be lower towards the center. This is illustrated Figure 5.1 in which a Laplacian operation is applied to simplified one-dimensional, low resolution model of two cotton boll signals of different unit sizes.



(a)

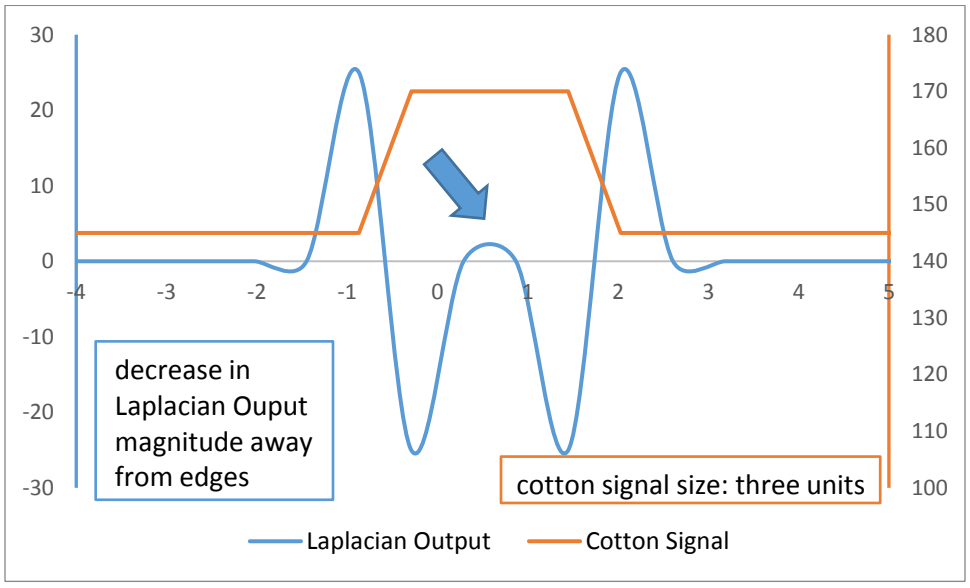


(b)

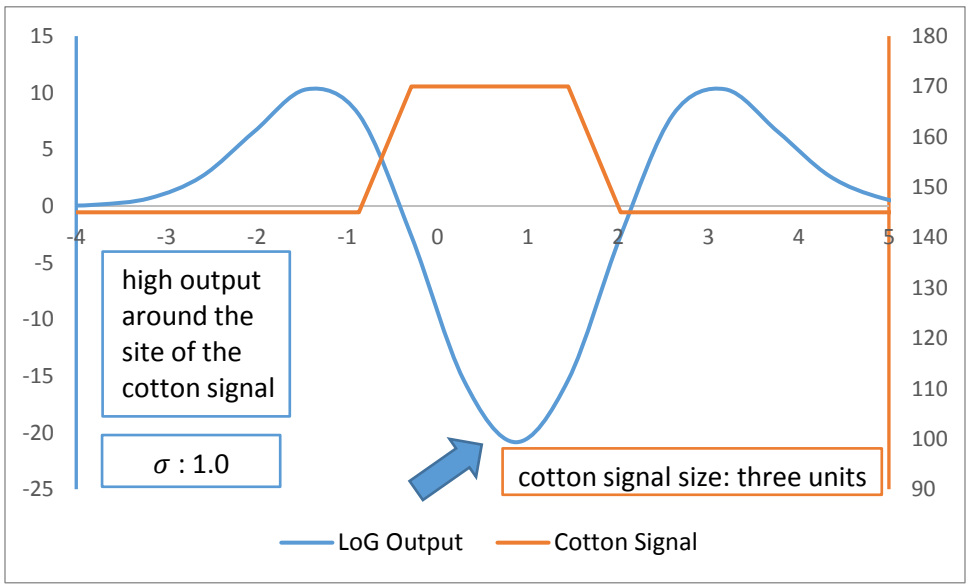
Figure 5.1. The Laplacians of low resolution cotton signals one unit in size (a) and three units in size (b).

The discrete Laplacian operator is conducted on a unit-wise basis. The output of the Laplacian of the cotton signal depended on the number of units describing the signal. In Figure 5.1(a) the Laplacian of the one unit-size cotton signal has a large negative output around site of the cotton signal. As the cotton signal increases in wavelength, as in the three unit-size cotton signal in Figure 5.1(b), the Laplacian of the cotton signal has larger negative outputs towards the edges of the cotton signal. The Laplacian output at the center of the cotton signal, however, low. Cotton signals of a certain size have a low Laplacian outputs towards the center of the signal. The centers are then removed during thresholding. For step like signals, the discrete Laplacian operator favors edges, and the entire region is not described as a maximum. This can be managed by scaling the discrete Laplacian operator to measure the second derivative characteristics over larger unit areas. This is done via the Laplacian of Gaussian operator

[Lathen 2010 and Muntasa et al. 2014]. Figure 5.2(b) shows result of the Laplacian of Gaussian operator applied to the low-resolution, three-unit size cotton signal alongside the Laplacian output in Figure 5.2(a) previously shown in Figure 5.1(b).



(a)



(b)

Figure 5.2. The Laplacian (a) and the Laplacian of Gaussian (LoG) of low resolution cotton signal three units in size (b).

In Figure 5.2(b) the entire step-like cotton signal is detected as a steep global maximum in the cotton field signal and is described as concave up. This is evident in the large negative output of the operator at the site of the cotton boll. It then follows that a threshold operation would allow for the detection of the cotton boll based on its second derivative properties.

The Laplacian of Gaussian operation can be formulated as a Laplacian operation applied to a Gaussian function. This is illustrated in Equation 7 below where σ represents the standard deviation and $x, y \in \mathbb{R}^2$.

$$H(x, y, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

$$G(x, y, \sigma) = H(x, y, \sigma) * I(x, y) \tag{7}$$

$$\Delta G(x, y, \sigma) = \frac{\partial^2}{\partial x^2} G(x, y, \sigma) + \frac{\partial^2}{\partial y^2} G(x, y, \sigma)$$

The plot of the two-dimensional Laplacian of Gaussian operator is displayed in Figure 5.3. A convolution with this operator yields large positive responses to local maxima and large negative responses to local maxima where the extent of region of these extrema are close to a radius of $\sqrt{2}\sigma$. The LoG operator can therefore be designed for local maxima regions of various sizes via the parameter σ.

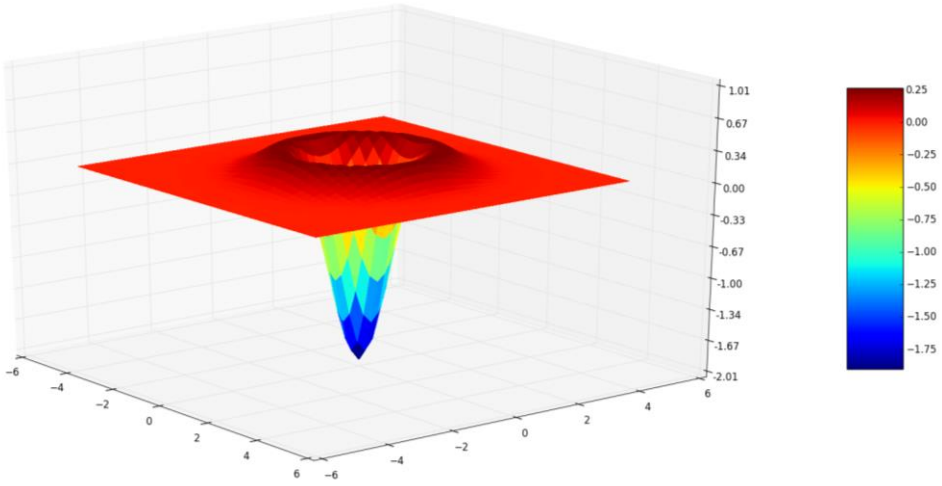


Figure 5.3. Two dimensional Laplacian of Gaussian (LoG) filter output.

Because the LoG is a linear operator, the discrete form of the LoG filter can be applied to the images first by applying a discrete Gaussian filter and taking the discrete Laplacian of the resulting filtered image. This process is shown below in Equation 8 where a single channel image is represented as a discrete function $I: x, y \in \mathbb{R}^2$ and K^G and K^L representing the Gaussian and Laplacian operator kernels respectively. The discrete neighborhood operation involved in the discrete Gaussian and discrete Laplacian are achieved via convolution with their respective kernels.

$$g(x,y) = I(x,y) * K^G$$

$$I^L(x,y) = g(x,y) * K^L \tag{8}$$

The output of Equation 8 is the LoG image I^L. A threshold operation on I^L can be applied with threshold value t_L. Such a threshold operation is shown as Equation 9 below.

$$M(x,y) = \begin{cases} 1 & I^L(x,y) < t_L \\ 0 & otherwise \end{cases} \quad (9)$$

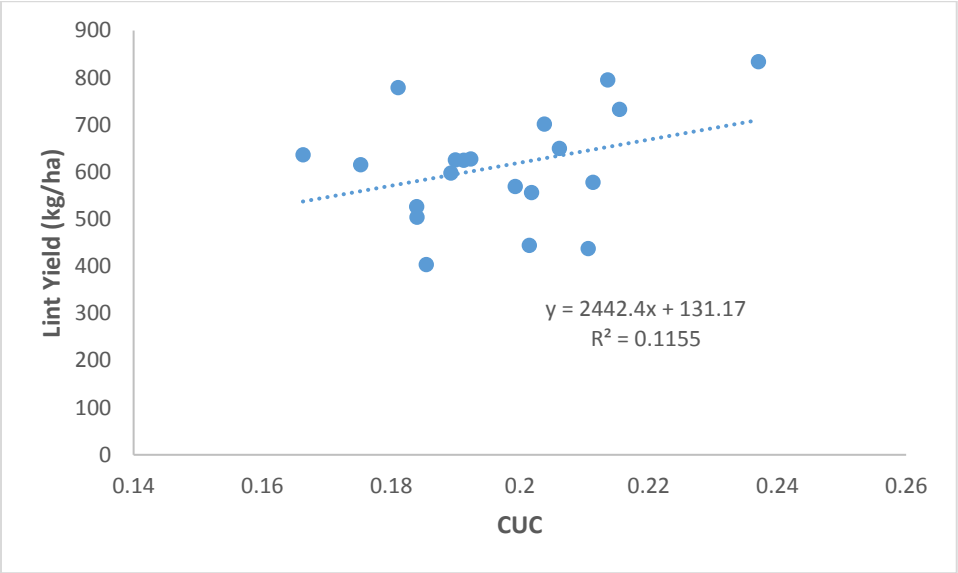
The output of the threshold operation of Equation 9 is a mask $M: \mathrm{I}^L \longrightarrow [0,1]$. The two controllable parameters for this method are therefore the standard deviation σ and the threshold value t_L. The objective, therefore, is to select a value σ and a value t_L such that M realizes most of the cotton signal without accepting a lot of noise.

5.2 Laplacian of Gaussian Experimental Results

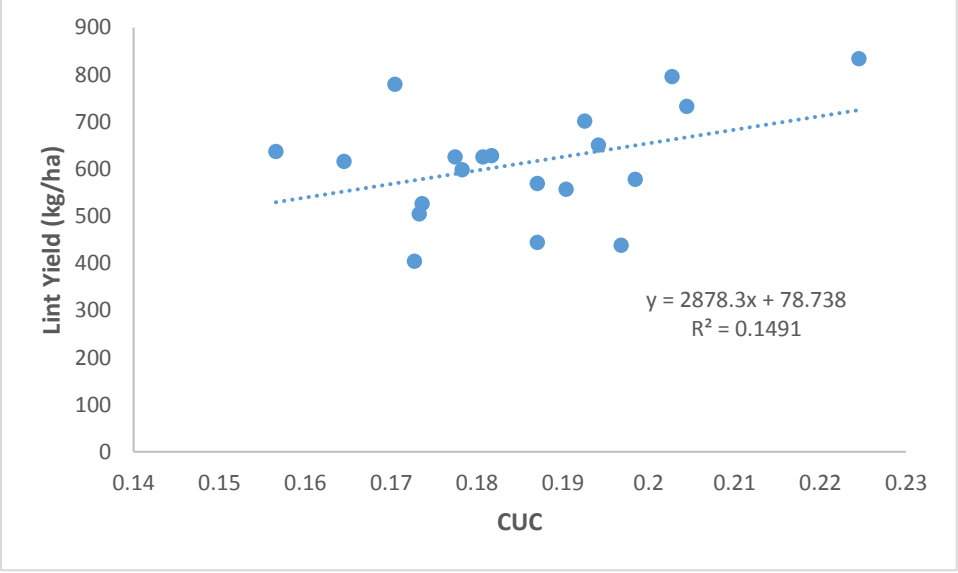
A study for the LoG parameters, σ and t_L was conducted. In this analysis, the CUC measurements were found using the same method field extraction as the Laplacian of grey level observation section. The R^2 were studied during the analysis at different values σ and t_L. The range of values of σ used during the analysis was from 0.5 to 1.5 with increments of 0.5. The range of threshold values t_L was 9-11 with increments of 1. For each value of σ, threshold operations were conducted on the output LoG image according the threshold range and increments. The output masks were then used to obtain CUC measurements which were then linearly correlated to yield at an R^2 value.

The presentation of the parametric study results will begin with a σ value of 0.5. Using the three threshold values of 9, 10, and 11, three masks were generated and utilized in CUC data sets with each threshold. The CUC values were then correlated with yield using a linear

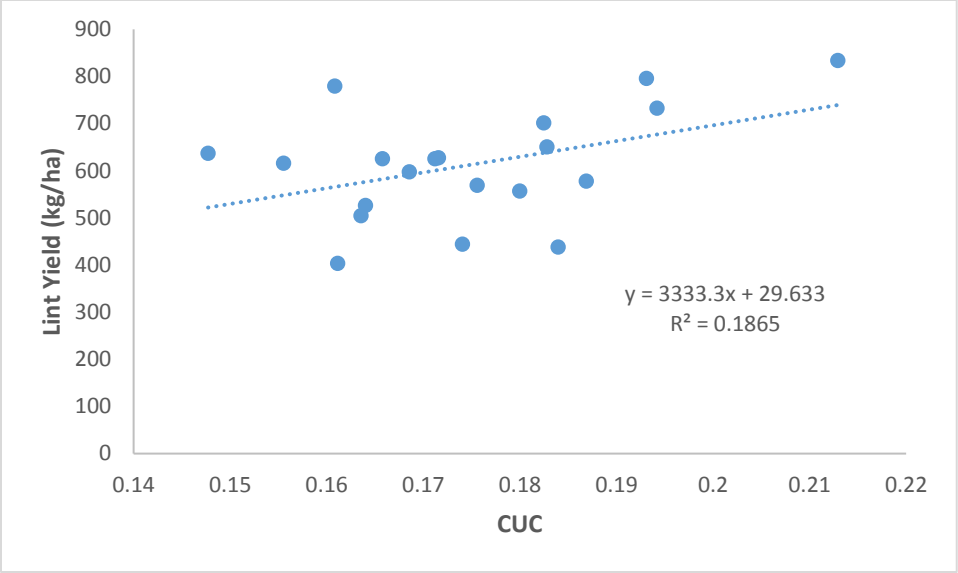
regression. R^2 values for each set could then be obtained. Figure 5.4 displays the three plots for σ of 0.5 along with their t_L and R^2 values.



(a)



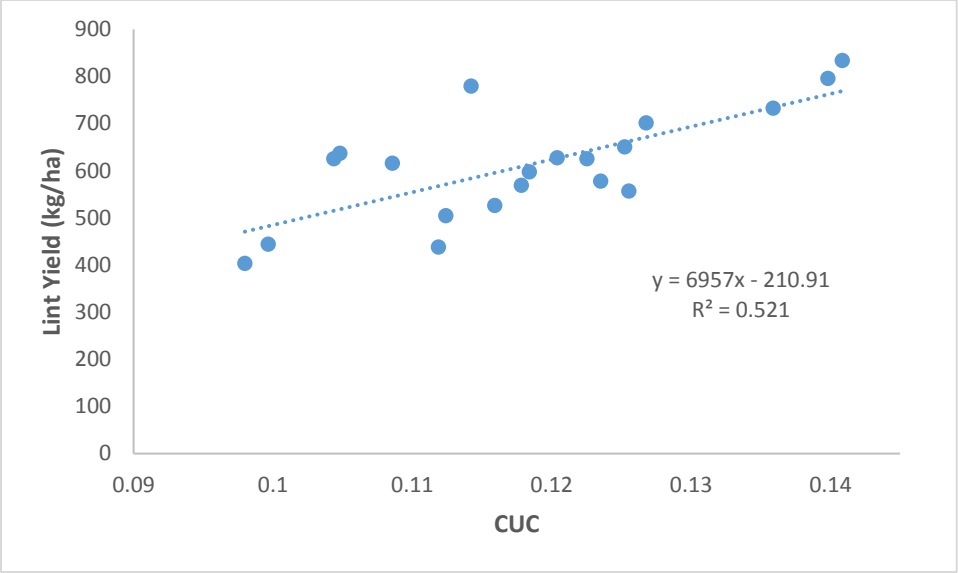
(b)



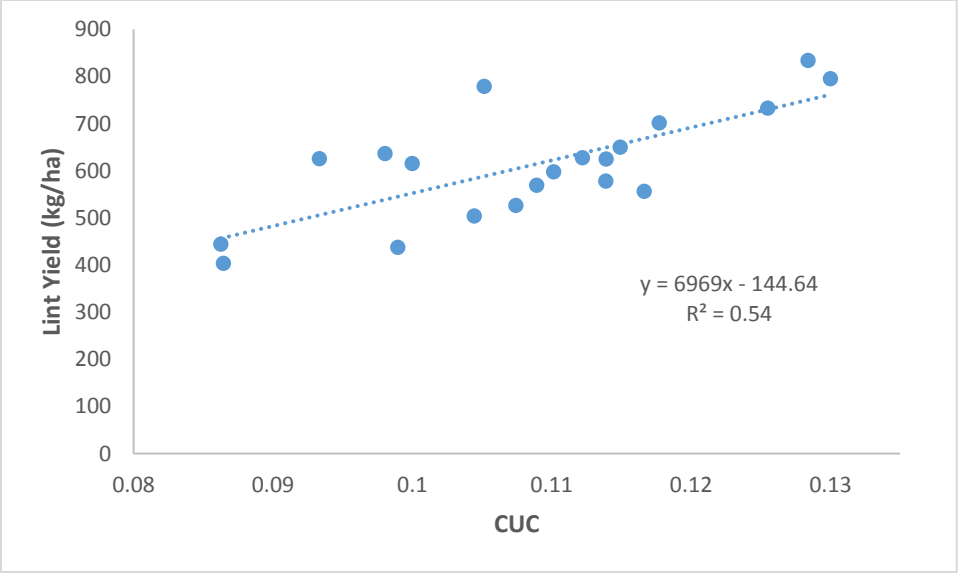
(c)

Figure 5.4. The relationships between the estimated CUC and the measured lint yield. The CUC was obtained using the LoG thresholding method with time step σ = 0.5 for all three plots and threshold values 9 for (a), 10 for (b), and 11 for (c).

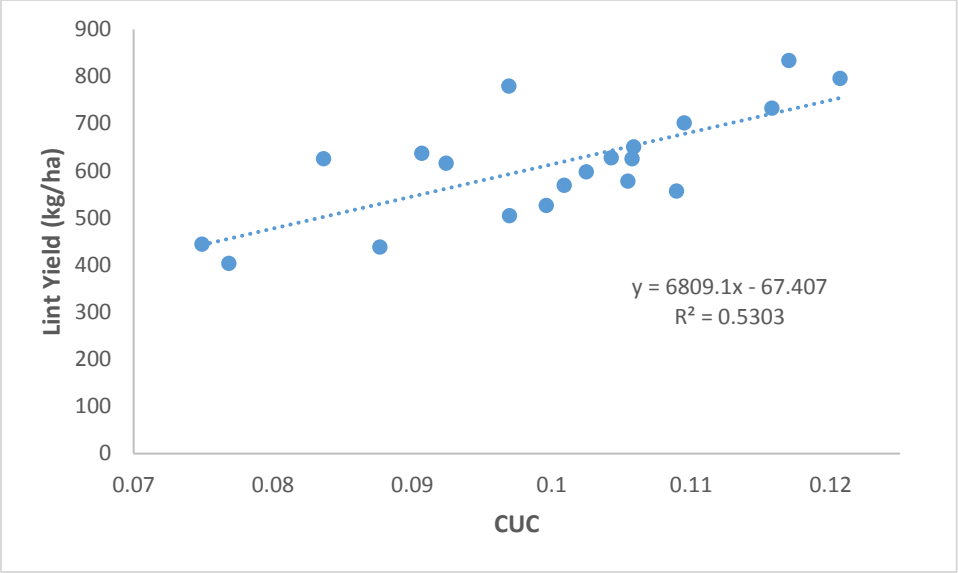
An observation in Figure 5.4 shows that for a LoG filter with a standard deviation 0.5 for cotton pixel extraction, the linear correlation between the CUC estimate and the manual measured lint yield increases as the threshold value increases. The study continued with the utilization of a LoG filter with a σ value of 1.0. Threshold operations were conducted on the filtered output image using the aforementioned threshold values. The masks resulting from these threshold operations were utilized in CUC estimation sets, which were then plotted against the lint yield. These plots, along with their respective linear correlations, are displayed in Figure 22.


900
800
700
600
500
400
300
200
100
0
Lint Yield (kg/ha)
0.09
0.1
0.11
0.12
0.13
0.14
CUC
y = 6957x - 210.91
R² = 0.521

(a)


900
800
700
600
500
400
300
200
100
0
Lint Yield (kg/ha)
0.08
0.09
0.1
0.11
0.12
0.13
CUC
y = 6969x - 144.64
R² = 0.54

(b)

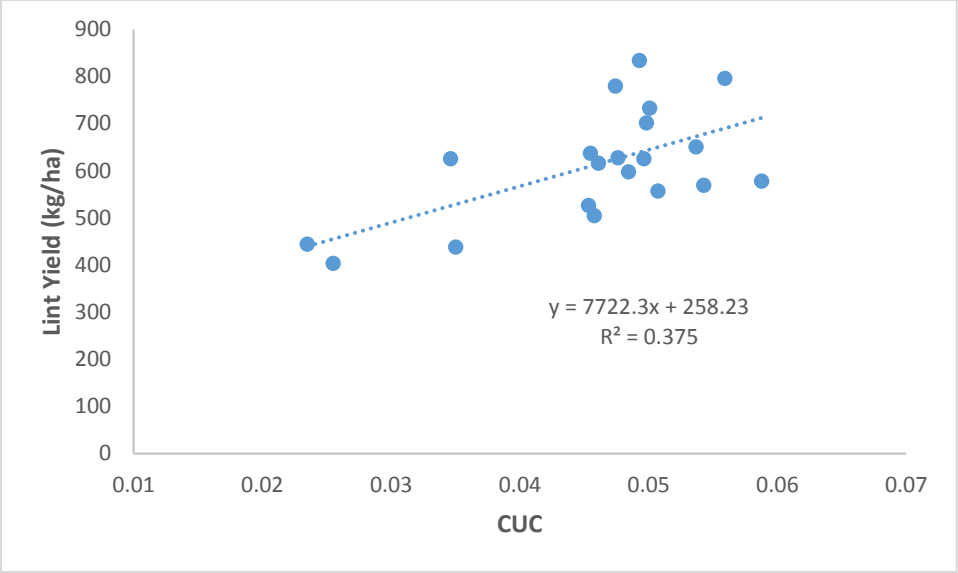


(c)

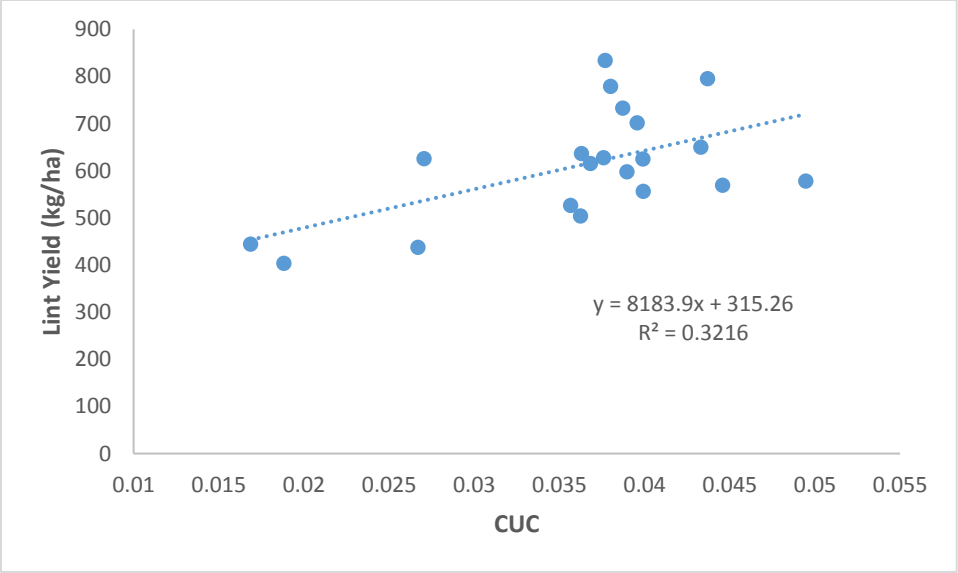
Figure 5.5. The relationships between the estimated CUC and the measured lint yield.

The CUC was obtained using the LoG thresholding method with time step σ = 1.0 for all three plots and threshold values 9 for (a), 10 for (b), and 11 for (c).

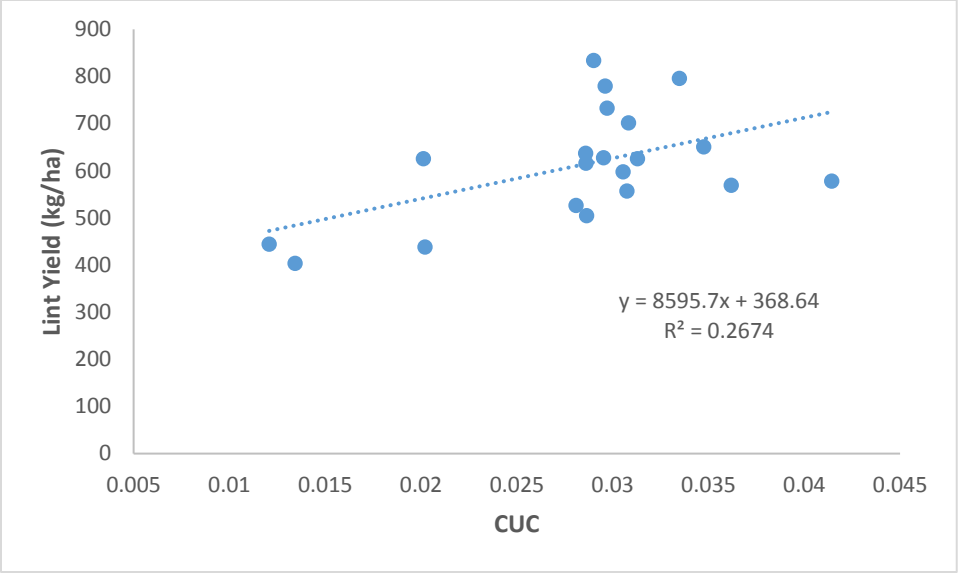
The CUC set which achieves the highest linear correlation is that which is displayed in Figure 22(b). CUC sets developed with threshold values moving away from 10 had smaller correlation values. The last set of CUC estimates were developed using a LoG filter with a σ value of 1.5 and using the aforementioned threshold values. These plots are displayed in Figure 23.


Lint Yield (kg/ha)
900
800
700
600
500
400
300
200
100
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
CUC
y = 7722.3x + 258.23
R² = 0.375

(a)


Lint Yield (kg/ha)
900
800
700
600
500
400
300
200
100
0
0.01
0.015
0.02
0.025
0.03
0.035
0.04
0.045
0.05
0.055
CUC
y = 8183.9x + 315.26
R² = 0.3216

(b)



(c)

Figure 5.6. The relationships between the estimated CUC and the measured lint yield.

The CUC was obtained using the LoG thresholding method with time step σ = 1.5 for all three plots and threshold values 9 for (a), 10 for (b), and 11 for (c).

An observation of the plots in Figure 5.6 reveals an increase in correlation between the CUC set and manually lint yield as threshold value t_L decreases. Different behaviors in correlation values obtained using the threshold values 9, 10, and 11 were observed between the value s of σ 0.5, 1.0, and 1.5. The CUC/lint yield correlation results obtained using the σ and t_L values were tabulated and displayed in Table 1.

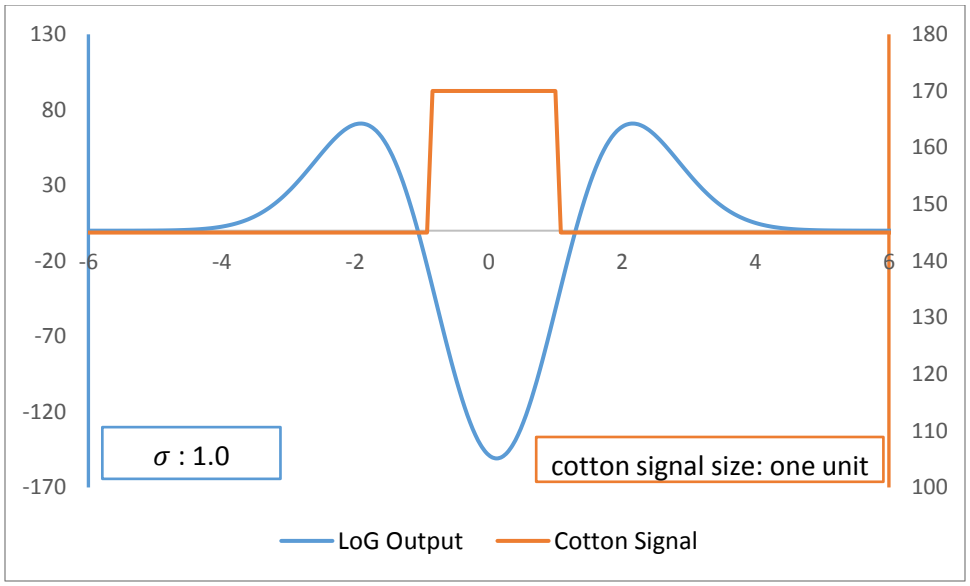
		σ		
		0.5	1	1.5
t_L	9	0.116	0.521	0.375
	10	0.149	0.54	0.322
	11	0.187	0.53	0.267

Table 1. Tabulated response of the Coefficient of Determination (R^2) to σ and t_L.

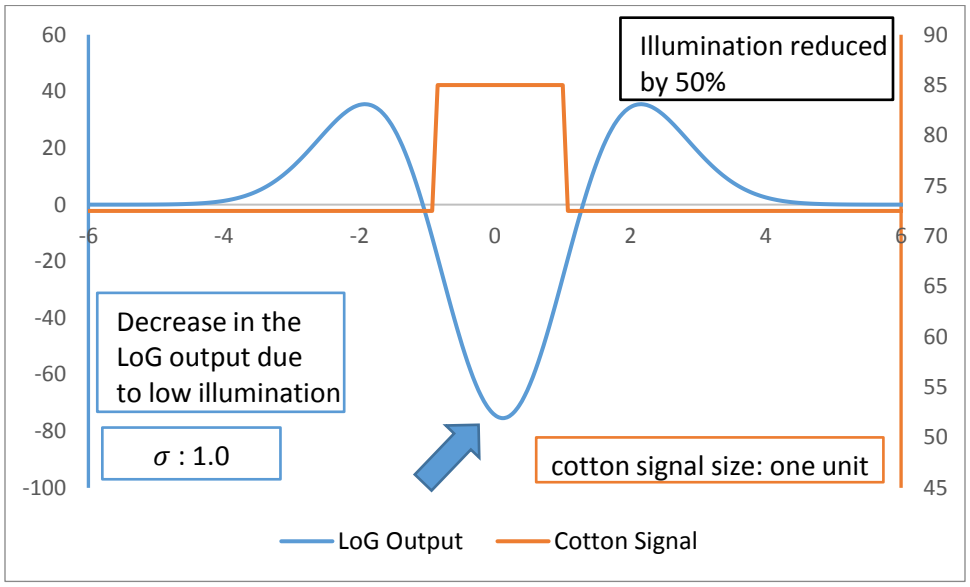
From Table 1, it is evident that the maximum achieved coefficient is 0.54 with σ and t_L values of 1.0 and 10 respectively. This is implemented without any correction being applied to the outliers identified in the cotton unit coverage (CUC) estimation via Laplacian of grey level observation section. This surpasses the R^2 value of 0.3433 from the cotton unit coverage (CUC) estimation via Laplacian of Gaussian grey level observation section. We now turn our attention to managing illumination variations by relating the LoG as a diffusion process.

5.3 Diffusion Response Thresholding for CUC Estimation

As indicated in the phase 1 study, less light may be incident on cotton bolls due to the structural location. This resulted in cotton bolls signals existing in a range of intensities. More severe effects of low incident light was observed in areas where sunlight was occluded by cloud coverage. The second derivative properties of the cotton boll signal changes proportionally to the amount of light illuminating the cotton boll. A decrease in incident light results in a decrease in the output of the Laplacian of Gaussian operator. This is illustrated in Figure 5.7 on the simulated cotton signal.



(a)



(b)

Figure 5.7. Laplacian of Gaussian (LoG) of cotton signal under normal illumination (a) and under 50% of normal illumination (b).

The output of the Laplacian of Gaussian filter is lower when the cotton field receives less incident light. The effect can be noted in the global threshold operation where the filter output is less than the characteristic threshold value for cotton. The cotton would therefore be removed by

the threshold operation. This is an issue in that outdoor images may contain many different gradients in the light source due to cloud coverage. It was therefore necessary to account for the illumination variation in the light source when describing the cotton signals second derivative properties.

The Laplacian of Gaussian of the image can be related to a diffusion process [Image Processing Handbook chapter image enhancement]. In this way, the image grey levels are related to concentrations and blurring between different grey levels, and are synonymous to the diffusions between different concentrations. Equation 10 displays a linear diffusion process with a constant diffusion coefficient D.

$$\frac{dG(x,y,\sigma)}{d\sigma} = D\Delta G(x,y,\sigma) \tag{10}$$

From Equation 10, it is evident that a threshold on the LoG filter result, as implemented in the previous section, is linearly proportional to the threshold on the diffusion rate $\frac{dG(x,y,\sigma)}{d\sigma}$. The assumption made when applying a global threshold, is that the diffusion rate doesn't depend on the concentration. If this assumption is satisfied, then a threshold operation applied to $\frac{dG(x,y,\sigma)}{d\sigma}$ allows regions to be separated based on the diffusion potential due to the reflectance properties of the object. In an image set with constant illumination, this assumption is satisfied in that the surface receives equal energy or incident light, and the relative grey levels are the result of the reflectance properties of the objects. In the case of outdoor images; however, this assumption may be significantly violated.

During aerial image acquisition, it is possible to have cloud coverage impeding the sunlight throughout a crop field of interest. This causes variations in light incident on the landscape. Consequently, the relative grey levels of these images are a function of the incident light variation as well as the material properties. The diffusion rate therefore is also dependent on the light variation as well. This diffusion process; however, can be modeled as an illumination normalized diffusion process where the diffusion coefficient is no longer constant, but is a function of the illumination variation. This is illustrated in Equation 11.

$$\frac{dG(x,y,\sigma)}{d\sigma} = D(x,y,G)\Delta G(x,y,\sigma) \tag{11}$$

To formulate the function D, we observe the simplified model for an image acquired under illumination variation utilized by Hou et. Al. (2010). This model is displayed in Equation 12 with I representing the image function, I_O representing the image without variation, η representing intensity variation, and ξ representing additive.

$$I(x,y) = \eta(x,y)I_O(x,y) + \xi(x,y) \tag{12}$$

Assuming that ξ has a zero-mean and is Gaussian, the application of the Gaussian filter will result in the noise being removed. The application of the Gaussian filter is represented by a function $F^G(\cdot)$. With the Gaussian filter being a linear operator, the application of the Gaussian filter on Equation 12 can be accomplished as shown in Equation 13.

$$F^G(I(x,y),\sigma) = F^G(\eta(x,y),\sigma)F^G(I_O(x,y),\sigma) + F^G(\xi(x,y),\sigma)$$

$$G(x,y,\sigma) = G_\eta(x,y,\sigma)G_O(x,y,\sigma) \tag{13}$$

The functions G_η and G_O denote the Gaussian filtered results of η and I_O. It is a common assumption that η is a slowly varying function. It can also be assumed for outdoor illumination variation on field textures that η varies much more slowly than I_O. A large standard deviation σ_v can be selected to allow the Gaussian filter to remove frequencies above the variation of η. At a value of $\sigma_{v,}$ the function G_O is approximately a constant due to it contributing to much higher frequency features in I. This is illustrated in Equation 14 where G_O is approximated as a constant C_O.

$$G(x,y,\sigma_v) \approx G_\eta(x,y,\sigma_v)C_O$$

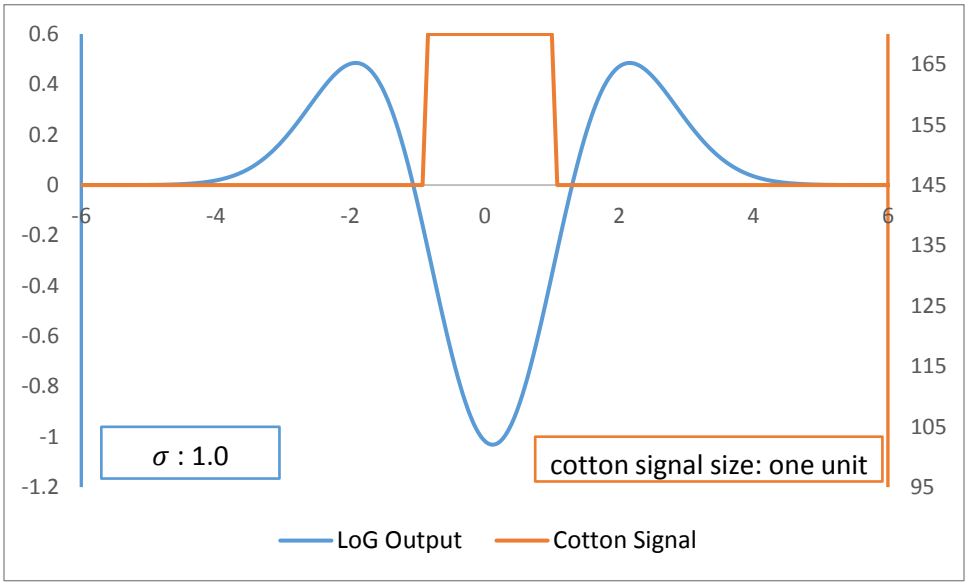
$$G(x,y,\sigma_v) \approx \eta(x,y)C_O \tag{14}$$

Due to the linearity of the Gaussian filter, $G(x,y,\sigma_v)$ can be accomplished with successive Gaussian filtering of $G(x,y,\sigma)$. Given that σ_v selected to retain all the frequency components of η, it can be assumed that $G_\eta(x,y,\sigma_v)$ equals the original signal η. The light variation is directly proportional to contrast and therefore is the diffusion rate. The diffusion coefficient can therefore be modeled as inversely proportional to illumination variation, in order to allow the diffusion rate $\frac{dG(x,y,\sigma)}{d\sigma}$ to characterize the diffusion potential. This is illustrated in Equation 15.

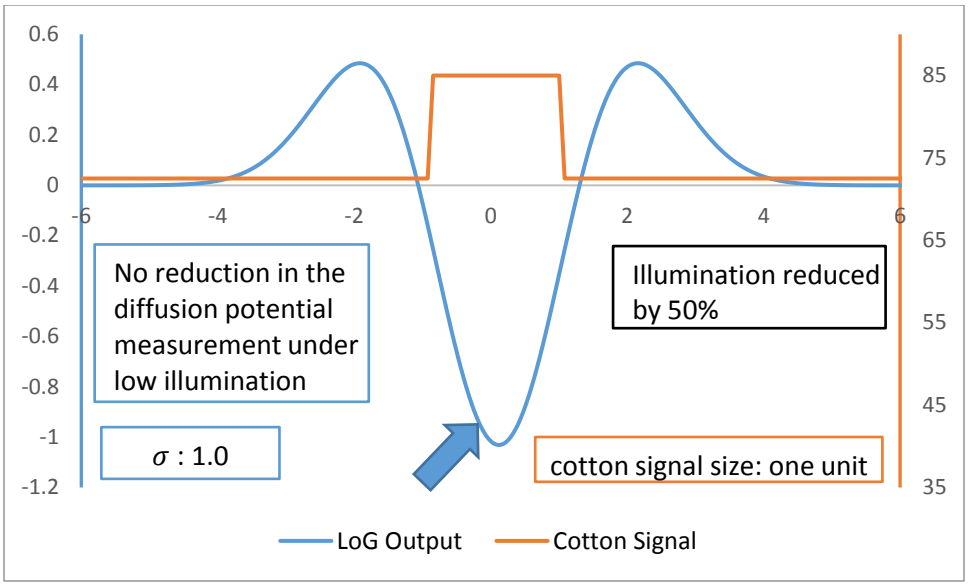
$$D(x, y, G) \propto G(x, y, \sigma_v)^{-1} \tag{15}$$

$$\frac{dG(x, y, \sigma)}{d\sigma} = G(x, y, \sigma_v)^{-1} \Delta G(x, y, \sigma_c)$$

The illumination invariant diffusion potential models the illumination variance due to gradients in the light source. σ_v is chosen to capture low frequency variations in the image [spatial and spatial-frequency analysis in visual optics Westheimer]. Considering the plots in Figure 5.7 they are offset by different low frequency gains caused by different magnitudes in the incident light curve. Normalizing by those gains describes the signal according to its diffusion potential. This illustrated in Figure 25.



(a)



(b)

Figure 5.8. Illumination Variation Normalized Diffusion Potential of cotton signal under normal illumination (a) and under 50% of normal illumination (b).

5.4 Discussion Response Thresholding Experimental Results

For the experiment, LoG was utilized on the image set using the $\sigma = 1.0$ which yielded the best correlation. Afterwards, D was found by filtering the image with a Guassian filter of standard deviation of $\sigma_v = 800$. Afterwards, the diffusion potential $\frac{dG(x,y,\sigma)}{d\sigma,}$ was thresholded at a value of 0.102 in order to generate a cotton mask. The CUC values were estimated from the masks and plotted against the manually obtained lint yield, and is displayed in Figure 5.9 along with the R^2.

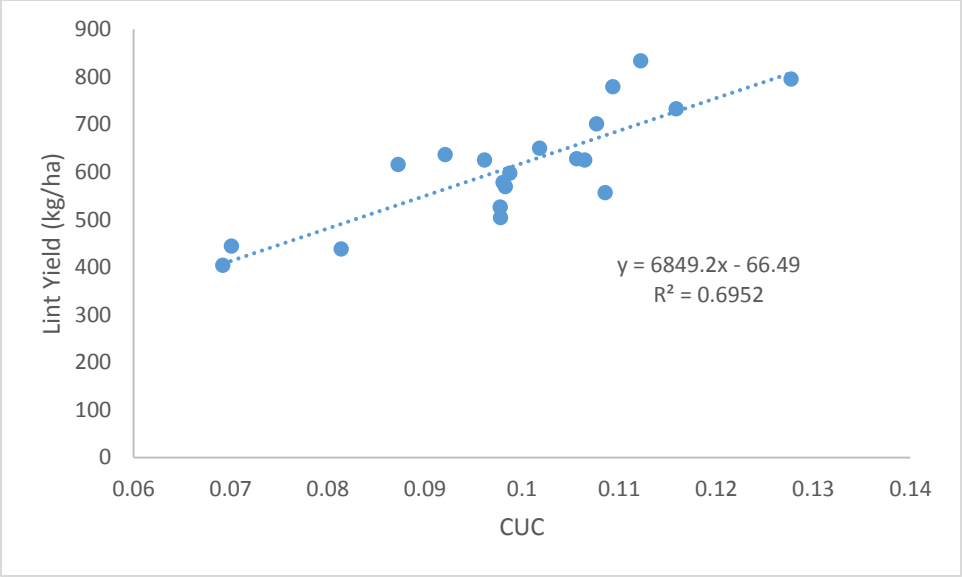


Figure 5.9. The relationships between the estimated CUC and the measured lint yield.

The CUC was obtained using the LoG thresholding method with time step $\sigma = 1.0$ for all three plots and threshold values 0.102.

From Figure 26, it is evident that normalizing the response due illumination variation provided improved CUC values. The correlation of the resulting CUC set is comparable to current expensive state of the art cotton variability systems. The assumption made in implementing this threshold, is that the unmanned aerial vehicle altitudes are constant. This assumption, however; is violated in that the altitudes are not constant. The CUC estimates of the function should also be normalized according to the altitudes. A correction is applied to the CUC using the altitudes and was replotted with the lint yield. The new CUC measurements along with new linear correlation is shown in Figure 27.

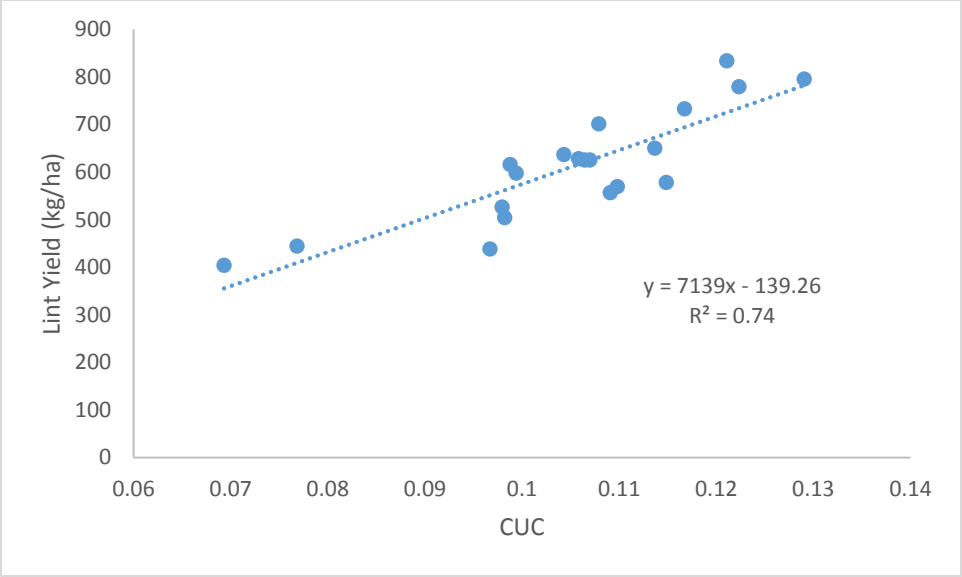


Figure 5.10. The relationships between the altitude corrected CUC and the measured lint yield.

The CUC was obtained using the LoG thresholding method with time step $\sigma = 1.0$ for all three plots and threshold values 0.102

As shown in Figure 27, the correlation increased to 0.74. This is an improvement on the current state of the art methods for cotton yield variation measuring techniques.

6. Summary and Conclusions

In this work, a study was conducted to develop a system that could be utilized to measure cotton yield variability. This system utilized a low cost, high resolution RGB camera for data acquisition mounted on a low cost miniature unmanned aerial vehicle (UAV). The study led to the development of a cotton unit coverage (CUC) metric which could be related to yield. The CUC was introduced as the ratio of cotton pixel area within a region to the region pixel area. A method was introduced to find boundary conditions of regions of study in order to calculate region pixel area. The cotton pixel area was observed through masks separating cotton pixels from the background pixels. Features for developing cotton pixel masks were investigated. Insight was gained on a feature space where the cotton pixel could more uniquely be characterized. It was shown that cotton bolls do not have a unique chromaticity relative to the cotton field, especially when imaged under lower illumination. It was also found that the range of cotton boll intensities may intersect with the range of background intensities. The shortcomings of intensity based methods were then explored. Also explored was the potential for separating cotton signals in the Laplacian of intensity space. This provided motivation for the development of an illumination normalized diffusion model which modeled cotton based on an illumination normalized diffusion potential. The use of the illumination normalized potential for cotton pixel extraction, led to the development of CUCs which explained 74% of the variation in cotton yield at a resolution of 3 cm. This offers improved results when compared to current, expensive state of the art methods (69% of variability at 1 m resolution).

Improvements could be made by accounting for viewpoint distortions on the CUC. Because cotton bolls exist at varying heights off the ground, some small variations in viewpoint could cause some of the row space to not fully be observed. This could potentially cause variations in CUC measurements. The method could also benefit from a scale space implementation in order to manage different cotton boll pixel clusters of different sizes. Finally, in order for this method to be used by farmers, it needs to be automated in which parameters for CUC estimation are automatically identified.

REFERENCES

Alarcon, V. J., & Sassenrath, G. F. (2011). Modeling cotton (Gossypium spp.) leaves and canopy using computer aided geometric design (CAGD). *Ecological Modelling, 222*(12), 1951-1963.

Awrangjeb, M., & Fraser, C. S. (2014). Automatic segmentation of raw LiDAR data for extraction of building roofs. *Remote Sensing, 6*(5), 3716-3751.

Bovik, A. C. (2000). *Handbook of image and video processing*. London;San Diego, Calif;: Academic.

Chaivivatrakul, S., Moonrinta, J., & Dailey, M. N. (2010). *Towards Automated Crop Yield Estimation-Detection and 3D Reconstruction of Pineapples in Video Sequences.* Paper presented at the VISAPP (1).

Constable, G., & Oosterhuis, D. (2010). Temporal dynamics of cotton leaves and canopies *Physiology of Cotton* (pp. 72-79): Springer.

Daugman, J. G. (1988). Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *Acoustics, Speech and Signal Processing, IEEE Transactions on, 36*(7), 1169-1179.

Domenikiotis, C., Spiliotopoulos, M., Tsiros, E., & Dalezios, N. (2004). Early cotton yield assessment by the use of the NOAA/AVHRR derived Vegetation Condition Index (VCI) in Greece. *International Journal of Remote Sensing, 25*(14), 2807-2819.

Eckert, S. (2008). 3d-building height extraction from stereo ikonos data-quantitative and qualitative validation of digital surface models-derivation of building height and building outlines. *JRC Scientific and Technical Reports, JRC43067, EUR, 23255*.

Gamba, P., Dell Acqua, F., & Houshmand, B. (2002). SRTM data characterization in urban areas. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 34*(3/B), 55-58.

Gong, A., Yu, J., He, Y., & Qiu, Z. (2013). Citrus yield estimation based on images processed by an Android mobile phone. *Biosystems engineering, 115*(2), 162-170.

Gonzalez, R. C., & Woods, R. E. (2008). *Digital image processing* (Vol. 3rd). Harlow: Prentice Hall.

Goodman, B., & Monks, C. D. (2003). A Farm Demonstrations Method for Estimating Cotton Yield in the Field for Use by Extension Agents and Specialists. *Journal of Extension, 41*(6).

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P., Burgos-Artizzu, X., & Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture, 75*(1), 75-83.

Hanan, J. (1997). Virtual plants—integrating architectural and physiological models. *Environmental Modelling & Software, 12*(1), 35-42.

Hanan, J. S., & Hearn, A. B. (2003). Linking physiological and architectural models of cotton. *Agricultural Systems, 75*(1), 47-77. doi:http://dx.doi.org/10.1016/S0308-521X(01)00114-7

Hou, Z., & Yau, W.-Y. (2010). *Relative gradients for image lighting correction.* Paper presented at the Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on.

Huang, Y., Brand, H., Sui, R., Thomson, S., Furukawa, T., & Ebelhar, M. (2016). *Cotton Yield Estimation Using Very High-Resolution Digital Images Acquired on a Low-Cost Small Unmanned Aerial Vehicle*. Manuscript submitted for publication.

Huang, Y., & Thomson, S. J. (2015). Remote sensing for cotton farming. *Cotton*(agronmonogr57).

Huang, Y. B., Thomson, S. J., Hoffmann, W. C., Lan, Y. B., & Fritz, B. K. (2013). Development and prospect of unmanned aerial vehicle technologies for agricultural production management. *INTERNATIONAL JOURNAL OF AGRICULTURAL AND BIOLOGICAL ENGINEERING, 6*(3), 1-10. doi:10.3965/j.ijabe.20130603.001

Hung, C., Nieto, J., Taylor, Z., Underwood, J., & Sukkarieh, S. (2013, 3-7 Nov. 2013). *Orchard fruit segmentation using multi-spectral feature learning.* Paper presented at the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems.

Jafari, A., Mohtasebi, S. S., Jahromi, H. E., & Omid, M. (2006). Weed detection in sugar beet fields using machine vision. *Int. J. Agric. Biol, 8*(5), 602-605.

Jähne, B. (2005). *Digital image processing* (Vol. 6th rev. and ext.;1. Aufl.;6th rev. and extend;6th revis and extend;). New York;Berlin;: Springer.

Jeon, H. Y., Tian, L. F., & Zhu, H. P. (2011). Robust Crop and Weed Segmentation under Uncontrolled Outdoor Illumination. *SENSORS, 11*(6), 6270-6283. doi:10.3390/s110606270

Kroyan, A., Lalovic, I., & Farrar, N. R. (2001). *Effects of 95% integral vs. FWHM bandwidth specifications on lithographic imaging.* Paper presented at the 26th Annual International Symposium on Microlithography.

Kurtulmus, F., Lee, W. S., & Vardar, A. (2011). Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture, 78*(2), 140-149. doi:http://dx.doi.org/10.1016/j.compag.2011.07.001

Läthén, G. (2010). *Segmentation Methods for Medical Image Analysis.* Thesis.

Leon, C. T., Shaw, D. R., Cox, M. S., Abshire, M. J., Ward, B., Wardlaw III, M. C., & Watson, C. (2003). Utility of remote sensing in predicting crop and soil characteristics. *Precision Agriculture, 4*(4), 359-384.

Leon, C. T., Shaw, D. R., Cox, M. S., Abshire, M. J., Ward, B., Wardlaw III, M. C., & Watson, C. (2003). Utility of remote sensing in predicting crop and soil characteristics. *Precision Agriculture, 4*(4), 359-384.

Leung, C.-C., Chan, K.-S., Chan, H.-M., & Tsui, W.-K. (2005). A new approach for image enhancement applied to low-contrast–low-illumination IC and document images. *Pattern recognition letters, 26*(6), 769-778.

Linker, R., Cohen, O., & Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agriculture, 81,* 45-57.

Luo, G. (2006). Objective image quality measurement by local spatial-frequency wavelet analysis. *International Journal of Remote Sensing, 27*(22), 5003-5025.

McDaniel, M. W., Nishihata, T., Brooks, C. A., & Iagnemma, K. (2010). *Ground plane identification using LIDAR in forested environments.* Paper presented at the Robotics and Automation (ICRA), 2010 IEEE International Conference on.

Meng, X., Currit, N., & Zhao, K. (2010). Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sensing, 2*(3), 833-860.

Muntasa, A., Siradjuddin, I. A., & Sophan, M. K. (2014). Region Based Segmentation Using Morphological Mathematics And Laplacian Of Gaussian To Automatically Detect Retinal Blood Vessel Image. *International Journal of Advancements in Computing Technology, 6*(6), 93.

Onyango, C. M., & Marchant, J. (2003). Segmentation of row crop plants from weeds using colour and morphology. *Computers and Electronics in Agriculture, 39*(3), 141-155.

Patel, H., Jain, R., & Joshi, M. (2012). Automatic segmentation and yield measurement of fruit using shape analysis. *International Journal of Computer Applications, 45*(7), 19-24.

Payne, A., Walsh, K., Subedi, P., & Jarvis, D. (2014). Estimating mango crop yield using image

analysis using fruit at ‘stone hardening’stage and night time imaging. *Computers and Electronics in Agriculture, 100*, 160-167.

Payne, A. B., Walsh, K. B., Subedi, P., & Jarvis, D. (2013). Estimation of mango crop yield using image analysis–segmentation method. *Computers and Electronics in Agriculture, 91*, 57-64.

Peinecke, N., Wolter, F.-E., & Reuter, M. (2007). Laplace spectra as fingerprints for image recognition. *Computer-Aided Design, 39*(6), 460-476.

Prusinkiewicz, P. (1998). Modeling of spatial structure and development of plants: a review. *Scientia Horticulturae, 74*(1), 113-149.

Pu, Y., Jagtap, J., Pradhan, A., & Alfano, R. R. (2014). Spatial frequency analysis for detecting early stage of cancer in human cervical tissues. *Technology in cancer research & treatment, 13*(5), 421-425.

Rejesus, R. M., Marra, M. C., Roberts, R. K., English, B. C., Larson, J. A., & Paxton, K. W. (2013). Changes in Producers' Perceptions of Within-Field Yield Variability after Adoption of Cotton Yield Monitors. *Journal of Agricultural and Applied Economics, 45*(02), 295-312.

Renton, M., Kaitaniemi, P., & Hanan, J. (2005). Functional–structural plant modelling using a combination of architectural analysis, L-systems and a canonical model of function. *Ecological Modelling, 184*(2), 277-298.

Reuter, M., Biasotti, S., Giorgi, D., Patanè, G., & Spagnuolo, M. (2009). Discrete Laplace–Beltrami operators for shape analysis and segmentation. *Computers & Graphics, 33*(3), 381-390.

Reuter, M., Wolter, F.-E., & Peinecke, N. (2006). Laplace–Beltrami spectra as ‘Shape-DNA’of surfaces and solids. *Computer-Aided Design, 38*(4), 342-366.

Ritchie, G. L., Bednarz, C. W., Jost, P. H., & Brown, S. M. (2007). Cotton growth and development.

Russ, J. C. (2002). Image Enhancement (Processing in the Spatial Domain) *The Image Processing Handbook, Fourth Edition* (Vol. 4th). Boca Raton, FL: CRC Press.

Russ, J. C. (2002). *The image processing handbook* (Vol. 4th). Boca Raton, FL: CRC Press.

Søgaard, H. T., & Olsen, H. J. (2003). Determination of crop rows by image analysis without segmentation. *Computers and Electronics in Agriculture, 38*(2), 141-158.

Suzuki, Y., Okamoto, H., & Kataoka, T. (2008). Image Segmentation between Crop and Weed using Hyperspectral Imaging for Weed Detection in Soybean Field. *Environmental Control in Biology, 46*(3), 163-173. doi:10.2525/ecb.46.163

Swain, K. C., Thomson, S. J., & Jayasuriya, H. P. (2010). Adoption of an unmanned helicopter for low-altitude remote sensing to estimate yield and total biomass of a rice crop. *Transactions of the ASABE, 53*(1), 21-27.

Teixidó Cairol, M., Font Calafell, D., Pallejà Cabrè, T., Tresánchez Ribes, M., Nogués Aymamí, M., & Palacín Roca, J. (2012). Definition of linear color models in the RGB vector color space to detect red peaches in orchard images taken under natural illumination. *Sensors, 2012, vol. 12, núm. 6, p. 7701-7718*.

Tsai, V. J. (2006). A comparative study on shadow compensation of color aerial images in invariant color models. *Geoscience and Remote Sensing, IEEE Transactions on, 44*(6), 1661-1671.

Vellidis, G., Perry, C. D., Rains, G. C., Thomas, D. L., Wells, N., & Kvien, C. K. (2003). Simultaneous assessment of cotton yield monitors. *APPLIED ENGINEERING IN AGRICULTURE, 19*(3), 259-272.

Wang, Q., Nuske, S., Bergerman, M., & Singh, S. (2013). *Automated crop yield estimation for apple orchards.* Paper presented at the Experimental Robotics.

Westheimer, G. (2012). Spatial and spatial-frequency analysis in visual optics. *Ophthalmic and Physiological Optics, 32*(4), 271-281.

Zarco-Tejada, P. J., Ustin, S. L., & Whiting, M. L. (2005). Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. *AGRONOMY JOURNAL, 97*(3), 641-653. doi:10.2134/agronj2003.0257

Zhu, L., Shimamura, H., Tachibana, K., Li, Y., & Gong, P. Building change detection based on object extraction in dense urban areas. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science, Beijing, China, 37*, 905-908.