AUTOMATED CO-REGISTRATION OF MULTITEMPORAL SERIES OF MULTISPECTRAL UAV IMAGES FOR CROP MONITORING

Polycarp Ochieng' Mc'Okeyo

Enschede, The Netherlands, March 2018

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the award of a Master of Science degree in Geo-Information Science and Earth Observation.

Specialization: Geoinformatics

SUPERVISORS

Dr. Claudio Persello

Dr. Francesco Nex

ADVISORS

Dr. Ir. Anton Vrieling

Mr. Phillip Jende, MSc.

THESIS ASSESSMENT BOARD

Prof. Dr. Ir. Alfred Stein
(Chairperson)

Dr. Ir. Thomas Groen
(External Examiner, ITC- Department of Natural Resources)

DEDICATION

I dedicate this thesis to my mother Wifrida Aoko Pinywuoro and father Samwel Ochieng' Mak'Okeyo. Although our time together on earth was short lived, your contributions to my life have been with me in every part of my academic journey. Dad, I hope I made you proud!

Profoundly, I devote this academic achievement to my lovely daughter Billiah Fridah Monyangi. You have been my source of inspiration, and the reason to keep on forging ahead relentlessly. The eighteen months away from you has yielded something worth dedicating to you. May I also be your inspiration when your time to do the same comes.

DISCLAIMER

ACKNOWLEDGEMENTS

First and foremost, my continued appreciated dialogue goes to the ITC fraternity: in a special way, I wholeheartedly thank my supervisors Dr. Francesco Nex, Dr. Claudio Persello and Dr. Anton Vrieling for believing in me and giving me unparalleled guidance throughout the research period. My sincere gratitude goes out to Watse Siderius for assisting in image acquisition as the drone pilot, and Phillip Jende for his insights as my academic advisor and inspiration to be at the top of things during my research.

I wish to pour out my love and gratitude to Nyabuto Emily Barongo; you have been my best friend and confidant during this research period. Thank you for pushing me to achieve my set targets, thank you for instilling the much-needed positive energy to override the challenges that befell me. Your role in my success cannot go unhailed. Thank you!

To my larger family in Kenya, I have no words to express how grateful I am. Uncle Moses Owuor Pinywuoro, you have been an instrumental person in my life; I celebrate you for being a focal point of inspiration to me and the family at large; thank you for believing in me. Aunty Bibian Achieng' Onyango, thank you for always being there for me, and thank you for loving me unconditionally; your good will and wishes for me have been will me throughout my academic journey. In a special way, I wish to thank Maurice Samambo and Wilbroda Buyeshe for unceasingly praying for me and channelling their love to me miles away. I thank you for being awesome parents, may the Almighty God Bless you abundantly. Dr. Henry Onderi Nyabuto, thank you for being the yardstick and posing the challenge to me. Thank you for seeing the potential in me.

My utmost appreciation goes to my employer, my friend, my father-figure, Kennedy Kubasu; thy wisdom has guided and inspired me this far; you have been the latent energy behind my accomplishments since our paths crossed. Your stewardship is unmatched. Thank you for supporting me and shaping my career. I look forward to venture in uncharted geospatial horizons with you.

To my friends, Roni Rezaul and Dinah Ogara, you are the best. I salute you for your direct and indirect support. You were literally my support system throughout this journey, I appreciate you both and I will forever remain indebted to you for the support you gave me.

Lastly, my sincere appreciation goes to the Netherland Fellowship Programme for making it possible for me to achieve and live my dream. The scholarship has opened an array of new opportunities, in addition to putting me on the world platform in matters Geoinformation processing and Earth Observation. I have gained a lot during my stay in the Netherlands.

ABSTRACT

The technologies and applications of aerial photography have exponentially evolved in the past two decades. The advent of purpose specific micro sensors and unmanned aerial vehicles (UAVs) have gone a long way in revolutionizing the applications of aerial images from crop monitoring to hazard assessment. The use of UAVs has become widespread due to its operational flexibility, ultra-high spatial resolution acquisition, and inexpensiveness. However, the application of multitemporal series of multispectral UAV imagery still suffers a setback of significant misregistration errors, and therefore becoming a concern for applications such as precision agriculture. In addition, the micro sensors have different spectral properties thus an integration and/or comparative analysis is inevitable. Researchers have widely studied the application of UAV imagery; topical issues range from spectral and spatial properties to image registration and multi-sensor integration. In crop monitoring, accurate co-registration of images acquired within and in different epochs remains to be exhaustively researched on; spectral band-to-band alignment is fundamental to precise crop sensing.

Although the application of UAV-based sensors is widespread, direct image georeferencing and co-registration is done using GCP; this is usually costly and time consuming. This research proposes a novel approach for automatic co-registration of multitemporal UAV imagery using intensity-based keypoints. The approach is based on multispectral orthophoto alignment. However, photogrammetric products such as Digital Surface Models (DSM) present an interesting challenge to automatically co-register a multitemporal series of such products including the several spectral channels, from three to twelve, captured by the UAV-based micro sensors. This research makes an effort to investigate inherent intra-epoch and inter-epoch co-registration errors in multispectral imagery, and the spatial data quality of photogrammetric products of two UAV-based cameras (Parrot Sequoia and Micro MCA Tetracam).

A successful image registration involves four major stages, feature detection, feature description or descriptor extraction, feature matching, and geometric transformation. Estimation of geometric transformation from matched point pairs yields a 2D transformation matrix, in this case a similarity transformation matrix, which maps the inlying matched pairs. The inliers are selected randomly iteratively until sufficient pairs fitting the transformation model is reached. In this study, existing keypoints detection algorithms were assessed to identify the most viable algorithm to be implemented for co-registration of multi-spectral and multitemporal UAV images. Intensity-based registration algorithms (SURF, BRISK, MSER and KAZE) were tested and optimally parameterized. A detailed comparison on the performance of these algorithms was done and an informed decision was reached to pursue further experiments with only SURF and KAZE.

The co-registration error analysis shows that optimally parametrized SURF and KAZE algorithms can obtain co-registration accuracies of 0.1 and 0.3 pixels for intra-epoch and inter-epoch images respectively. To obtain better intra-epoch co-registration accuracy, collective band processing is advised. On the other hand, the quality of the DSM per band is evaluated and the results show that the red band is best fit for DSM extraction. In addition, the spatial variability of the UAV-based spectral features and that of Sentinel 2B satellite imagery are compared; the results show that regardless of the differences in spectral bandwidths, and spatial resolution, they are highly correlated. A positive correlation of 0.93 and 0.77 was obtained for the maize field and non-vegetated area respectively. However, the UAVs take pride in the spatial resolution advantage to reveal intrinsic intra-farm variability. On the other hand, a comparison of spectral response of vegetation using Parrot Sequoia and Micro MCA Tetracam show that although both cameras are able to correctly sense active and declining photosynthetic activity in crops, spectral and radiometric calibration is key to achieving optimal response for the Micro MCA camera. In light of the results obtained in this study, descriptor-based methods are fit for co-registration of multispectral imagery for crop monitoring. In addition, UAV based multispectral cameras have different specifications and thus differ in the quality of their respective photogrammetric outputs; using the same system for monitoring purposes is advised.

Keywords: Co-registration, Multitemporal, Multi-spectral, Unmanned Aerial Vehicles, Feature detection, Feature description, Feature matching, Geometric Transformation, Crop Monitoring, Orthophoto, Digital Surface Model, Ultra-high resolution.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

Crop cultivation is arguably the predominant form of agriculture, and as such, is important to feed the world's population. In developing countries, for every four people, three live in the rural periphery, and their livelihood is directly or indirectly dependent on agriculture (World Bank, 2008). The ever-increasing need to optimize yield while sustainably exploiting finite resources remains a challenge to farmers. In this light, it is prudent to have an efficient decision support system for farm management. Crop cultivation faces unique problems not common to other economic activities, from climate or weather related problems to soil erosion and degradation, thereby warranting the need to use specialized technologies like Remote Sensing (RS) in crop management (Atzberger, 2013). According to Mulla (2013), Remote Sensing technologies have been applied in monitoring crops for close to three decades now. Over the years, a large number of sensors and RS techniques have been used to better map and monitor crop growth, including the development of low-altitude aerial systems like drones.

Contemporary monitoring of crops is rooted in three fundamental principles: optimizing yield, economic efficiency, and curbing environmental impacts (Allahyari, Mohammadzadeh, & Nastis, 2016). In the effort to undertake crop production by adhering to these principles, Precision Agriculture (PA) has advanced as a feasible system by implementing geospatial technologies to observe, measure and match agricultural production to the potential of specific areas (Elarab, Ticlavilca, Torres-Rua, Maslova, & McKee, 2015). In so doing, PA brings together spatial data from an array of sources, from UAV photography to terrestrial laser scanning. FAO (2013), points out that access to such technologies and improved agro-ecological practices have shaped crop production. Allahyari et al. (2016) assert that to maximize crop production sustainably, embracing and implementing PA as a farm management system is key.

1.2. Research Justification and Problem Statement

The full range of satellite imagery from coarse, medium, very high resolution (VHR), and synthetic-aperture radar, has been used for crop monitoring for a while. For example, Inoue et. al. (2014) assess the biophysical variables in paddy rice using high-resolution satellite SAR imagery. The capabilities and applications of SAR interferometry for PA are discussed by Lopez-Sanchez and David Ballester-Berman (2009). The availability of VHR multispectral imagery has made it possible to detect crop stress and diseases, determine soil nitrogen content, extract vegetation indices like NDVI, and ultimately estimate crop yield (Lee et al. 2010). Mulla (2013), reviews the contribution of satellite imagery to PA and notes that the analysis of multitemporal images at varying crop growth stages has enabled scientists to identify heterogeneous patches within similar farms. Potential crop-stressed zones are identified with ease and a management decision is made. Lobell (2013) illustrates that VHR imagery like RapidEye, (5m spatial resolution) presents more spatial information and an opportunity to map crops and estimate yield with a comparatively higher accuracy than medium and low-resolution imagery like Landsat and MODIS respectively. However, with VHR data like Ikonos, Quickbird, Spot-5 and Worldview that attain up to between 0.5m to 2.5m spatial resolution, image analysis becomes computationally demanding, and a balance between the spectral and spatial characteristics has to be struck (Dalla-Mura et al. 2009). The major setbacks for VHR satellite imagery include the high costs, the complexity involved in tasking (i.e. to capture a scene when needed), the cloudy weather constraint, and relatively long revisit period (Issei et al. 2010; Gevaert et al. 2015).

Recently, the application of the drone technology in crop monitoring is rife. Nex & Remondino (2014) review the use of unmanned aerial vehicles for 3D mapping applications and highlights Agriculture as a domain that consumes digital surface models (DSM) and orthoimages to extract useful information on crop status. Mazur (2016) explains that the simulation of multitemporal UAV images shows the exact growth of crops and reveals inadequacies in production. In addition, UAV imagery surpasses satellite imagery in spatial resolution. ultra-high Resolution multispectral and multitemporal imagery is undoubtedly an enabler of Precision Agriculture to obtain actionable crop properties (Elarab et al. 2015).

UAVs are embraced across domains because they are flexible low-altitude RS platforms; due to this fact, the spatial and temporal resolution is ultra-high, and they are not affected by cloud occlusion. UAV-based sensors can achieve ground sampling distances (GSD) of 5cm, 3cm or less depending on the flight parameters and the aim of the acquisition (Nex & Remondino, 2014). This is still ten times higher the spatial resolution of the best VHR satellite imagery. In addition, UAVs provide an inexpensive alternative to satellites and other platforms for

aerial image acquisition; they increasingly offer tools and inspire innovations that seal the gap between terrestrial and aerial (high-altitude) platforms (Nex et al. 2015). Conversely, UAVs face a couple of drawbacks: regulatory constraints on the application of drones and licensing of drone pilots vary from country to country; limited areal coverage due to the battery endurance per flight; the instability of lightweight platforms; atmospheric elements such as strong winds and rain affect drone operations; the payload limit; image co-registration complexities, and difficulties in radiometric and geometric corrections (Remondino et al. 2011; Nex et al. 2014; Freeman et al. 2015;Yang et al. 2017).

Accurate image co-registration is vital for reliable change detection assessment and accurate comparative analysis of crop phenology (Fytsilis et al., 2016; Tilly et al., 2014). Several models and algorithms that automate the co-registration process have been proposed. However, multispectral cameras with several lenses still suffer misregistration setbacks as demonstrated in related works of Jhan et al. (2017); Jhan et al. (2016); and Rey et al. (2013). This is partly due to the fact that the technology space is dynamic and new camera sensors with different specifications and more abilities are continuously being engineered. Sub-pixel accuracy co-registration of multispectral bands is desirable for accurate extraction of vegetation indices. On the other hand, co-registration of multitemporal series is vital for reliable spatiotemporal analysis of crop's spectral properties. Misclassification of crop growth per pixel, vegetation index extraction errors, interpolation errors in values between available observations, and difficulties in predicting variations in harvest index are some of the inherent errors due to misregistration (Lobell, 2013). Figure 1.1 illustrates the misregistration problem.

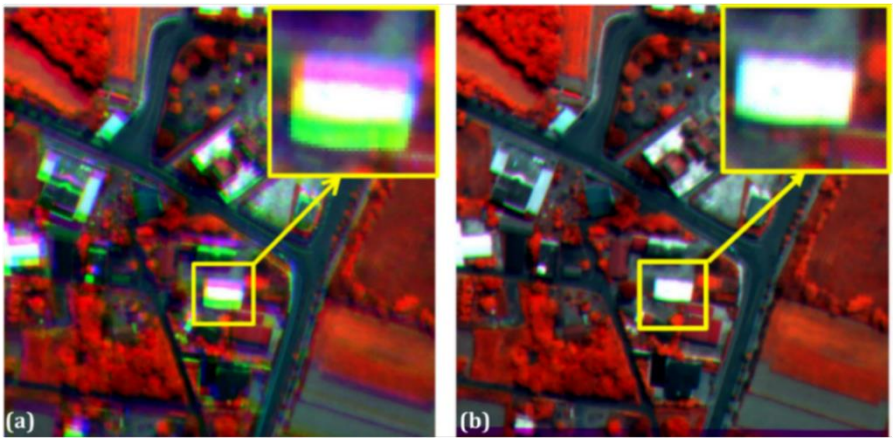


Figure 1. 1. Misregistration effect of the false colour image. (a) Original imagery and (b) Correctly registered. SOURCE: (Jyun Ping Jhan et al., 2016)

With these error sources in mind, defining a method to accurately register or co-register UAV-based images ultimately translates to accurate modelling and quantification of yield, which is key to closing the yield gap. Precise crop monitoring and by extension spatial yield quantification could help to increase farm productivity by informing management decisions for under-performing areas of the field.(Mueller et al. 2013).

1.3. Research Identification

1.3.1. Research Pivot

Although an array of methodologies and aerial sensors have been used for crop monitoring, gaps still exist. This research used ultrahigh resolution multispectral UAV imagery to explore intra-farm spectral and spatial variability of maize crop. In addition, the spatial data quality of two multispectral sensors were investigated. As opposed to the RGB cameras, the multispectral cameras used in this research (Micro MCA Tetracam and Parrot Sequoia) have spectral channels within the red edge and NIR regions of the electromagnetic spectrum which are important for sensing photosynthetic activity in vegetation. In this research, the spectral variability of Parrot Sequoia and Micro MCA Tetracam is explored; further, the spatial variability of spectral features between the UAV platforms and Sentinel 2B is assessed.

This study aims at accurate co-registration of multitemporal UAV images by comparing existing algorithms, and adopting or modifying the best-performing one; extraction and assessment of digital surface models (DSM), orthorectification, extraction of biophysical properties of the crop like Normalized Difference Vegetation Index (NDVI), comparative analysis of spectral variability between the two UAV-based cameras, and assessment of the spatial variability of spectral features between the UAV and satellite platforms such as Sentinel 2B.

1.4. Research Objectives and Questions

The main objective of this study is to investigate intra-epoch and inter-epoch misregistration of multispectral UAV imagery, and to explore the potentials of unmanned aerial systems for crop monitoring. The specific objectives and corresponding research questions are as follows:

Objective 1: To review the application of UAVs in crop monitoring and evaluate the contemporary methods of UAV image co-registration.

The research questions are:

a) What is the role of UAV imagery in crop monitoring and what milestones have been achieved?
b) What is the added value of additional spectral channels for crop monitoring?
c) What are the state-of-the-art methods used for UAV image co-registration?

Objective 2: To establish the best band combinations for photogrammetric processing and image analysis.

a) What is the effect of different band combinations on the accuracy of band-to-band registration?
b) Which is the best band combination for image co-registration within an epoch?
c) Which is the best band for DSM extraction?

Objective 3: To accurately co-register multitemporal series of multispectral imagery and DSMs, and to assess the co-registration errors.

a) What are the inherent co-registration errors? How can these errors be minimized?
b) Which is the best band combination for image co-registration between epochs?
c) Which registration method best fits this research problem?

Objective 4: To evaluate the similarity between spectral signatures from crops as extracted from Micro-MCA Tetracam, Parrot Sequoia imagery and Sentinel 2B.

a) What is the spatial variability of the spectral features extracted from UAVs and Sentinel 2B?
b) What is the variability in spectral response of vegetation between Sequoia and Micro MCA Tetracam?

1.5. Novelty and Innovation Aim

Although the application of UAV-based sensors is widespread in crop monitoring, direct image georeferencing and co-registration is done using GCP; this is usually costly and time consuming. This research proposes a novel approach for automatic co-registration of multitemporal UAV imagery using intensity-based keypoints. In addition, this research aims at investigating the feasibility and applicability of the Micro-MCA Tetracam in crop monitoring. This entails accurate co-registration of the multitemporal series of multispectral images, and extraction of spectral features of maize crop from the UAV-acquired images. The originality of this research is based on the co-registration methodology of the multispectral images and the comparative analysis of spectral variability between two UAVs and Sentinel 2B. Assessment of the spatial variability of spectral features between the aforementioned sensors is also a novel idea.

1.6. Thesis Structure

This study is organised into seven major chapters addressing various aspects of the research. Chapter one includes the introduction, background, justification, motivation, problem statement, research objectives and questions, and the novelty of this study. Chapter two contextualizes this study in the crop monitoring domain and underpins it to the theoretical background of RS using UAVs. An in-depth review on methods of image co-registration is done in this section. Chapter three basically describes the equipment, data and software platforms used in this study. The methodology is covered in chapter four. Results and analysis are presented in chapter five and the discussion in chapter six. Lastly, the conclusions drawn from the research and recommendations are presented in chapter 7.

2. LITERATURE REVIEW

This chapter underpins the pillar concepts of this research on a theoretical background. The concept of UAV-based crop monitoring is described in section 2.1. A description of the spectral response of vegetation and the crop indices feasible to be extracted are described in section 2.2. The building blocks of photogrammetry and image co-registration are discussed comprehensively in sections 2.3 to 2.5.

2.1. The application of UAVs and Multispectral sensors in Crop Monitoring

Remote sensing technologies have gone a long way in hatching a range of in-situ earth observation applications. Satellite based sensors continue to make advancements to enhance spatial and temporal resolutions. However, sophisticated technologies, operational rigidity, and high economic demand plague these advancements. Although satellite based sensors have the areal extent advantage, they still suffer from comparatively coarse spatial and temporal resolution; cloud cover; and timing constraints to capture real-time events due to fixed acquisition timings (Matese et al. 2015). On the other hand, the aircraft surveys are more flexible and have increased spatial-temporal resolution, but are quite costly, operationally demanding and still affected by weather elements (Greene, 2012). For certain applications, these drawbacks are undesirable. Precision Agriculture (PA), employs techniques and equipment able to sense intra-farm heterogeneity driven by intrinsic spatial elements such as diseases, soil nutrients and water content among others. The UAVs have demonstrated to be the suitable remote sensing platform for this application (Matese et al. 2015; Nex et al. 2014; Song, 2016).

2.1.1. Unmanned Aerial Vehicles

The UAVs are flexible to operate and provide an inexpensive method to obtain the critical elements of precise crop monitoring, which include ultra-high spatial and temporal resolutions. The most common UAVs used for crop monitoring include helicopters, fixed-wings, multi-rotors, blimps and flying wings (Sankaran et al. 2015). The UAVs are equipped with single or multiple mini-cameras, and Global Navigation Satellite Systems for flight and image geolocation. Table 1 shows some of the UAVs used for crop monitoring.

Specification	**Description**			
	Multi-rotor	**Helicopter**	**Quadcopter**	**Fixed Wing**
Pics				
Model	Matrice 600 Pro	Vapor 55	Phantom 4	ebee
Manufacturer	DJI Technologies	Pulse Aerospace	DJI Technologies	Sensefly
Power/Motors	6	2	4	1
Weight (Kgs)	10	20	1.4	0.73
Payload (Kgs)	5.5	5	0.5	0.15
Endurance (mins)	38	60	28	45
Reference	(DJI, 2016a)	(Aerospace, 2017)	(DJI, 2016c)	(EBee, 2015)

Table 1 Examples of UAVs used for crop monitoring.

The helicopter drones can take off and land vertically and hover; their payload is comparatively large and can support larger sensors. Nevertheless, operational complexities, limited hovering ability, and relatively high maintenance costs constrain widespread application of helicopter drones (Swain et al. 2010;Chapman et al. 2014). The flying speed of fixed wing drone is high and has a longer flight time. However, the lack of free hover ability and the image blur attributed to high speeds and altitudes are some of its setbacks (Link, Senner, & Claupein, 2013). Multirotor UAVs are low cost; they have the ability to hover; low take-off and landing requirements; and are most frequently used for crop monitoring. However, multi-rotor UAVs suffer from relatively short flight time, lower payload and sensitivity to weather (Peña et al. 2013; Näsi et al. 2015).

2.1.2. UAV-based Multispectral Sensors

Multispectral sensors capture images within and beyond the visible range of the electromagnetic spectrum. The UAV-based sensors are small in size and light weight to enable mounting on most UAV platforms. They are relatively inexpensive, have fast frame imaging and as such have been widely applied for crop monitoring (Yang et al. 2017). The use of multispectral sensors allows for ultrahigh resolution aerial imagery. The spectral channels capture the absorption and reflectance of radiant energy by crops; this information is then used to characterize crop health status, detect invasive crop species, and predict crop yield (Garcia-Ruiz et al., 2013; López-Granados et al. 2016; Song, 2016).

In comparison to RGB cameras, multispectral cameras have more spectral channels thus more spectral information is captured by these sensors. The multispectral sensors are widely used in precision agriculture applications today; they reveal crop properties such as chlorophyll and water content, nitrogen concentration, leaf area index, photosynthetically active radiation among other biophysical properties (Swain et al. 2010; Elarab et al. 2015; Zhu et al. 2013). Table 2 shows some of the multi-spectral UAV-based sensors used for crop monitoring.

Specification	**Description**			
Pics				
Model	Micasense Red edge	Mapir	Parrot Sequoia	Micro MCA
Manufacturer	Micasense, Seattle	Peau Productions	Parrot, France	Tetracam Inc.
Bands	5	4	4 plus RGB	6
Focal length (mm)	5.5	3.97	3.98	9.8
Image Size	1280 x 960	4032 x 3024	1280 x 960	1280 x 1024
Weight (g)	175	65	135	530
Shutter	Global	Global	Global	Rolling
Reference	(Micasense, 2016)	(Geosense, 2017)	(Parrot Sequoia, 2016)	(Jhan et al., 2016)

Table 2 Examples of the multi-spectral UAV-based sensors used for crop monitoring.

2.2. Vegetation Spectral Response

Vegetation is uniquely distinguishable in the NIR region of the electromagnetic spectrum due to its spectral property. It has a low reflectance in the blue and red regions of the electromagnetic spectrum; this is attributed to chlorophyll for photosynthesis. Nevertheless, it peaks at the green region which increases the green colour of vegetation. The reflectance is however higher in the NIR than in the visible range due to cellular structure of leaves (Jensen, 2007; Govaerts & Verhulst, 2010).

The dominant factors affecting the leaf reflectance include the leaf pigments, cell structure, and the water content. "Light falling on a leaf can be reflected, absorbed or transmitted. Absorption in the visible and infrared regions of the spectrum is primarily driven by stretching and bending of covalent bonds between oxygen, carbon, hydrogen and nitrogen present in plant biochemical components like sugar, lignin, cellulose and proteins. In addition, pigments responsible for leaf color also constitute principal absorbing molecules. Because of the central function of these pigments in photosynthesis, chlorophyll content is generally regarded as a good indicator of plant physiological health. Many nutrient deficiencies result in a decrease in chlorophyll content, a concomitant increase in reflectance in the visible range (400 – 700 nm) a decrease in reflectance in the infrared (700-1100 nm) ranges and blue shift in the red edge inflection point. Visually, chlorotic changes are perceived as yellowing of leaves" (Crop Copter, 2017). Figure 2.1 and 2.2 shows the vegetation spectral response in the spectrum.

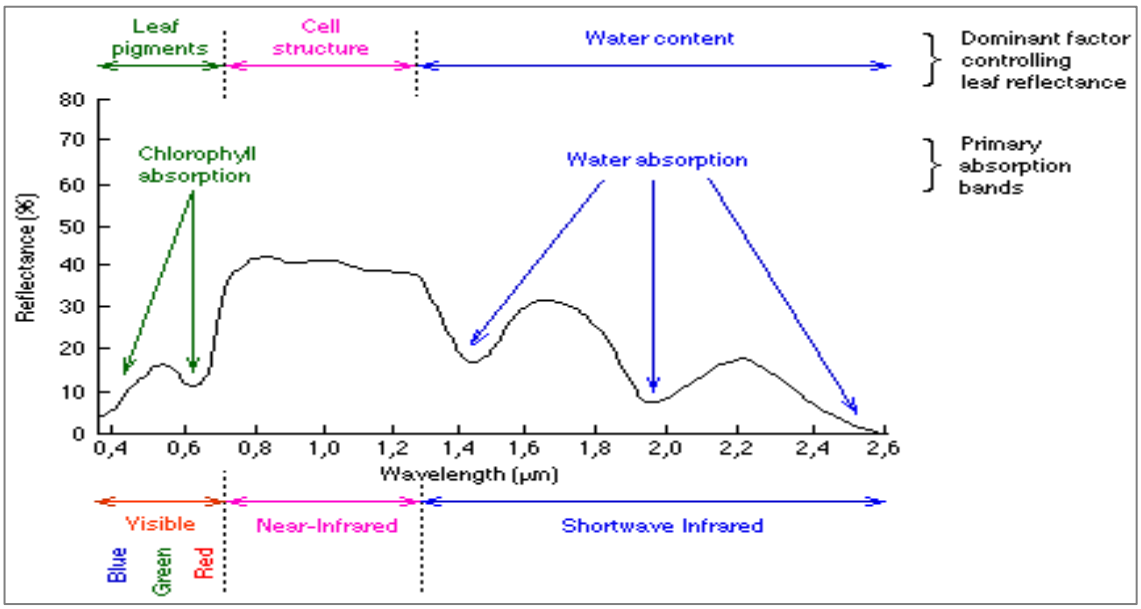


Figure 2.1 Vegetation Reflectance Curve. Source: (Crop Copter, 2017)

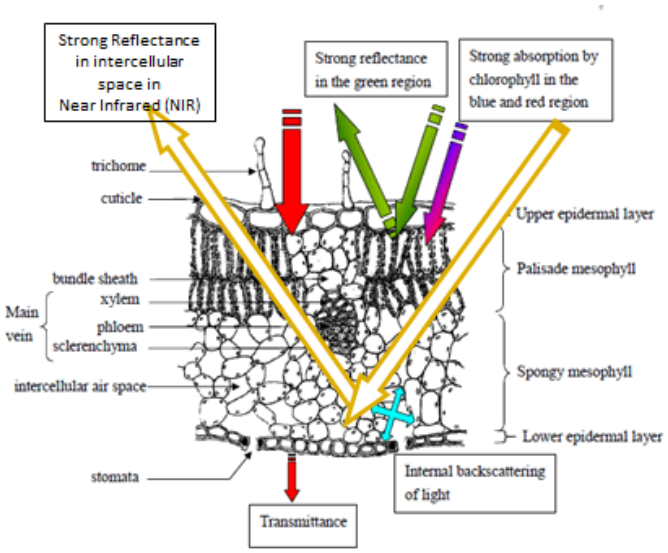


Figure 2.2 Leaf Structure in relation to interaction mode with light. Source: (Crop Copter, 2017)

2.3. Automatic Image Co-registration

"Image co-registration is the process of geometrically aligning two or more images to integrate or fuse corresponding pixels that represent the same objects" (Harris Geospatial, 2017). This is done by modelling the geometric relationship between two overlapping images using transformation models such as affine, projective and polynomial among others. Image registration has four main components; feature detection, feature description, feature matching and geometric transformation (Zitová & Flusser, 2003; Madhuri, 2014).

2.3.1. Feature detection

Feature detection refers to identification of twin points (also known as tie points or keypoints) between successive overlapping images. Information about keypoints in two images is an integral part of the co-registration process. Desirable keypoints must be devoid of noise, blurs, illumination variances and geometric differences; this facilitates retrieval of corresponding points from multiple images of the same scene by different sensors.

The four main types of feature detectors are discussed as follows; emphasis is put on the detectors used in this research.

a) Edge detectors

A significant local change in intensity of an image is referred to as an edge. Edge detectors are therefore mathematical model constructs that detect abrupt intensity discontinuities or sharp changes in brightness in an image (Shrivakshan & Chandrasekar, 2012). Steps in edge detection include filtering, enhancement and finally detection. Since gradient computation based on intensity is prone to noise, filtering is done to improve the performance of the edge detector algorithm. However, noise reduction through filtering is counter-useful since it reduces the edge strength, therefore a tradeoff has to be made (Davies, 2005). The major edge detectors used in image processing include:

i) **Canny detector:** A multi-stage algorithm developed by John F. Canny in 1986. It is the first derivative of a Gaussian and closely approximates the operator that optimizes the product of signal-to-noise ratio (Davies, 2005b).
ii) **Sobel detector**: Developed by Sobel Feldman in 1990. It performs a spatial gradient measurement and highlights regions of sharp spatial frequency as edges. It uses two filtering kernels at each point in the image; one checks the vertical edge for maximum response, and the other checks the horizontal edge (Bin & Samiei, 2012).
iii) **Prewitt detector**: It is similar to the Sobel detector apart from the constant of the partial derivatives. In Prewitt, the constant is equal to 1 while in Sobel the constant is 2, therefore having different convolution masks.
iv) **Roberts detector**: This is a first order operator which uses the approximation between two adjacent pixels of the diagonal direction of the gradient to detect the edges.

Generally, edge detectors are not suitable for image registration because edges are not distinct nor localized.

b) Corner detectors

The intersection of lines or edges can be defined as a corner. It is characterized by large variations in image gradient in two directions. Corners are vital for image registration because they are localized and can be used to orient objects and provide their dimensions (Davies, 2005a). The major corner detectors used in image processing include:

i) **Harris detector:** Developed by Harris and Stephens in 1988. The Harris corner detector finds the intensity differences by calculating the gradient of each pixel in all directions. If the absolute gradient value in both directions is great, that point is regarded as a corner (Chen et al. 2009). The Harris Corner detector algorithm makes use of eigenvalues of the matrix that results from expanding the second term. If the first eigen value is higher than the second eigen value (or vice versa), then an edge is detected. If both eigenvalues are small, then a flat region of uniform intensity is detected. Lastly, if both eigenvalues are large and approximately equal to each other, then a corner is detected.

ii) **Förstner detector:** Developed by Förstner & Gülch in 1987. The main aim was to provide a fast operator for detection and localization of discrete points, corner and centers of circular features within an image for photogrammetric applications. One major advantage is that the Förstner detector has the ability to detect features with a subpixel accuracy making it a reliable tie point detector.

iii) **FAST (Features from Accelerated Segment Test)** algorithm was developed and presented in a paper by Rosten & Drummond (2006). The detector selects a pixel, p and defines a circular region around this pixel with a radius equal to three pixels. Intensity values of a subset of pixels, n within this circular region are compared to the intensity value of p plus or minus a threshold value, t. Pixel p is considered a corner if all the surrounding n pixels are brighter than $I_p + t$ or darker than $I_p - t$.

Other detectors such as SUSAN (Smallest Univalue Segment Assimilating Nucleus), Moravec, Wang and Brady, Trajkovic and Hedley corner detectors among others exist (Hassaballah, Abdelmgeid, & Alshazly, 2016). Despite being able to detect localized features, corner detectors are not invariant to scale changes of an image thus not optimally suitable for aerial image registration.

c) Region detectors

Regions or blobs, are areas of different intensity in an image. Comparatively, these regions significantly differ in brightness to neighbouring regions. These regions are invariant to scale and this property makes them the best candidates for detecting corresponding points between two images of the same or different scales (Mikolajczyk et al. 2005). These regions do not have a fixed shape but adapt based on the underlying image intensity values. This property poses the question, are these regions really invariant? Mikolajczyk et al. 2005, explain that these regions are more covariant than invariant; they adapt covariantly with the transformation. However, it is important to note that the feature descriptors derived from these regions are typically invariant.

There are two types of blob detectors; the differential and watershed detectors. The former are based on differential expressions and the latter on local extrema of the image intensity (Kong, Akakin, & Sarma, 2013). The Laplacian of Gaussian and the difference of Gaussians are among the common region detectors used in image processing.

i) **Laplacian of Gaussian (LoG):** The Laplacian is a differential operator given by the divergence of the gradient of a function on Euclidean space. On the other hand, the Gaussian function is a rapidly decaying isotropic function that modifies the input by convolution. This combination forms the operational structure of the LoG (Gonzalez & Woods, 2004). The LoG first smoothens the input image using a Gaussian function. The filtering is at different scales and is defined by the standard deviation to reduce noise. Given the Gaussian function in equation 1.

$$G(x, y; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \qquad (1)$$

The Gaussian scale-space representation $L\ (x, y; \sigma)$ of an image $f\ (x, y)$ would therefore be

$$L\ (x, y; \sigma) = f\ (x, y) * G\ (x, y; \sigma) \qquad (2)$$

Where * is the convolution operator. The Laplacian operator (∇^2) would therefore be a summation of the differentiation of the image elements.

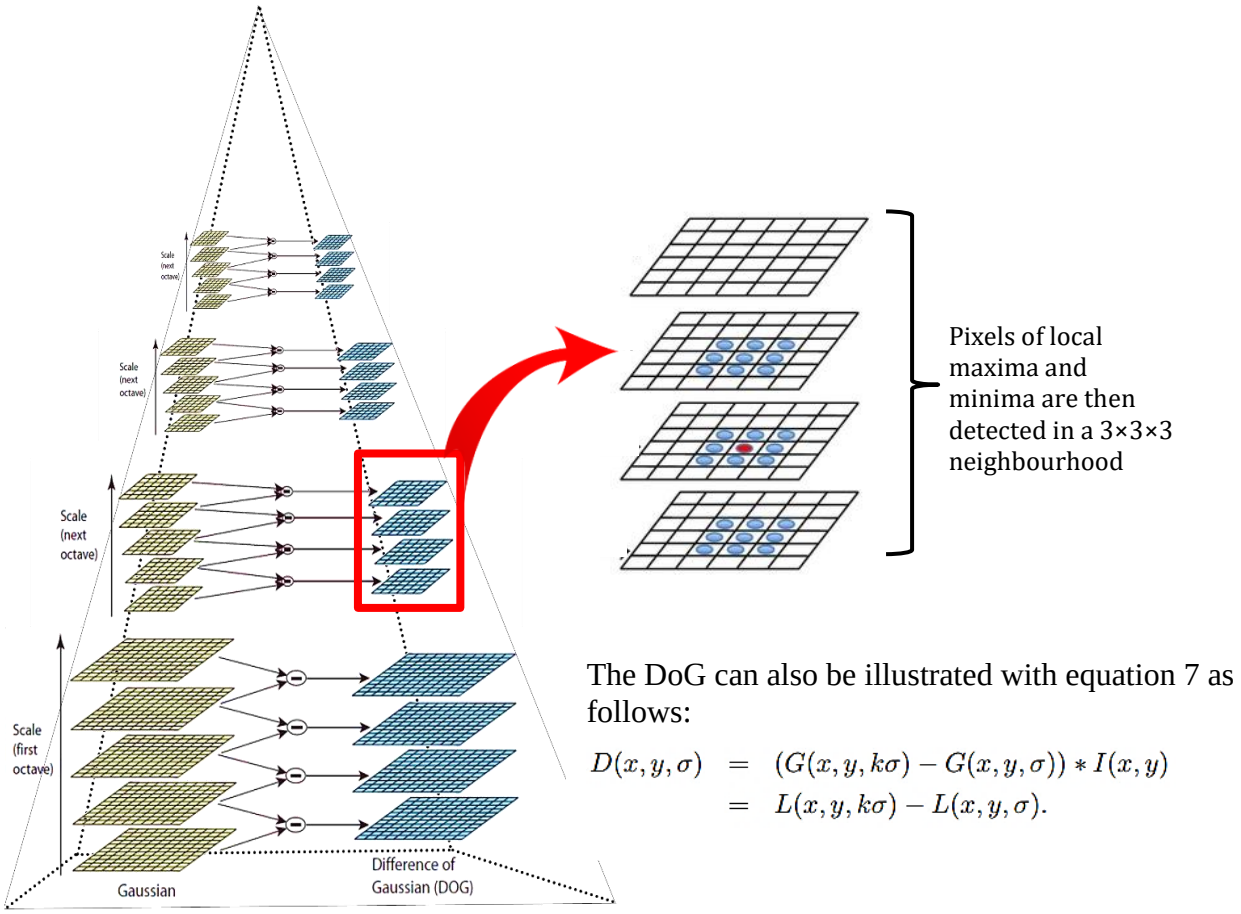
$$\nabla^2 = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \qquad (3)$$

The LoG representation is therefore

$$\nabla^2 \mathrm{G}(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \qquad (4)$$

As the scale of the standard deviation (σ) increases, blob-like features begin to appear and converge at a given scale. The size of the blobs is directly proportional to the standard deviation. Then Laplacian operator yields strong positive responses for regions of dark intensity and strong negative responses for blobs of light intensity (Kong et al. 2013).

ii) **Difference of Gaussians (DoG):** Uses two Gaussian functions of different space constants to approximate the LoG. Due to this structural difference, and the ability to detect stable blobs, the DoG is more efficient than the LoG (Huertas & Medioni, 1986; Lowe, 1999). The DoG operator subtracts filtered images per octave in successive levels of scale. An octave refers to a set of ordered images in a given Gaussian scale space. The layering of each subsequent level of scale forms a pyramid. The bottom-most image in this layering is the original image and the upper ones are scaled down images. The scaling down of images is done by a fixed factor. The pixels of local extrema are detected as keypoints. Figure 2.3 illustrates the DoG.



The DoG can also be illustrated with equation 7 as follows:

$$\begin{aligned} D(x,y,\sigma) &= (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) \\ &= L(x,y,k\sigma) - L(x,y,\sigma). \end{aligned}$$

Figure 2.3 Diagram illustrating the Difference of Gaussians (Modified from Lowe,1999)

The LoG and DoG have inspired a number of image registration algorithms including Lowe's **Scale Invariant Feature Transform (SIFT)** and the **Speeded Up Robust Features (SURF)** among others.

The **SIFT** algorithm, developed in 2004 by David G. Lowe, implements the DoG to detect keypoints and compute the respective descriptors. First and foremost, a space-scale filtering is done. The result is the LoG for the image with various standard deviation (σ) values. The LoG detects blobs of various sizes due to change in the σ. The standard deviation is therefore the scaling parameter. But LoG is cost thus **SIFT** implements the DoG. Once the DoG is found, the local extreme is searched over the scale-space by comparing one pixel with eight (8) neighbours as well as nine (9) in previous scales. The other stages of the **SIFT** algorithm include keypoints localization, orientation assignment, and keypoints description.

Despite implementing DoG, the **SIFT** algorithm remains to be computationally costly. To overcome this setback, Bay et al. 2006, developed the **SURF** algorithm which implements the Determinant of Hessian (DoH) instead of the Difference of Gaussians (DoG) to detect blobs. The **SURF** algorithm relies on the Hessian matrix $H(p, \sigma)$, for both scale and location. Where the determinant is maximum or minimum keypoints are detected. The H matrix is denoted as follows:

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix}$$

Where L_{xx}, L_{yy} and L_{xy} are second-order derivatives of intensity with respect to the x direction, y direction and both x and y directions respectively. **SURF** thus approximates the LoG with a box filter, which is easily computed with the help of integral images and can be done in parallel for varying scales. See Figure 2.4.

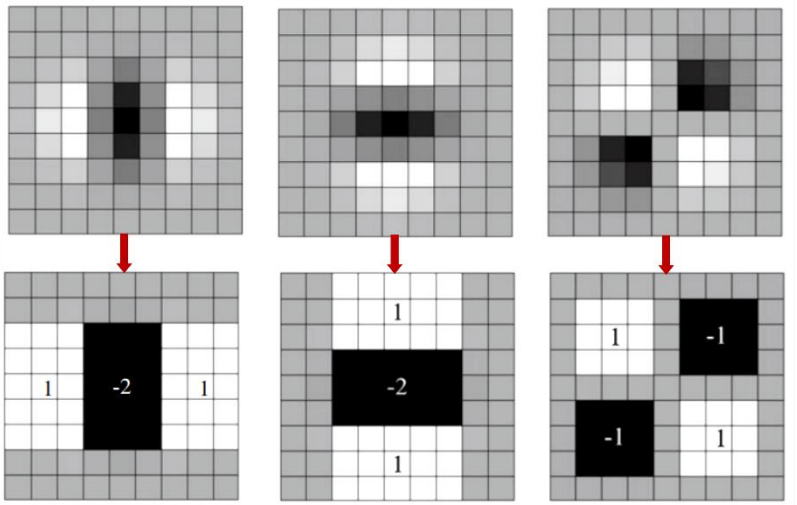


Figure 2.4 Box filters used in SURF to approximate LoG. From left to right are Lxx, Lyy, and Lxy. (Fan et al. 2015)

2.3.2. Feature description

Feature description refers to the encoding of more information about the detected keypoints; that includes but is not limited to image gradients or intensity of neighbouring pixels in relation to the detected blobs. This information is used as a "fingerprint" to distinguish detected features from each other (Lowe, 2004). To enable successful matching between corresponding images, the descriptor needs to be autonomous and invariant to illumination, scale, and orientation.

There are two main types of feature descriptors, and they include:

a) Float descriptors

These descriptors use image intensity gradients to encode for autonomous features. The image gradients of pixels around the center pixel (detected feature point) are computed in form of float digits. One of the eight possible orientations is assigned and weighted. The encoding is thereafter stored in a vector with the same dimensions as the descriptor.

The **SIFT** algorithm by David Lowe, 2004 is an example that implements the float descriptor. Notably, it is the foundation from which other descriptors have been developed. The **SIFT** descriptor is constructed by dividing the local region of 16 by 16 sample points. Histograms of gradient orientations are computed in 4x4 grids, and concatenated together to a unit vector as the descriptor (Fan et al. 2015). The orientations of the image gradient of each of these pixels are determined and simplified to eight possible values. These values are resolved for all pixels within a 4 by 4 array resulting in a descriptor with eight possible orientations stored in a 4 by 4 array. At the end, the SIFT descriptor vector has 128 dimensions thus making it computationally expensive and time consuming (Onyango Fredrick, 2017). See Figure 2.5.

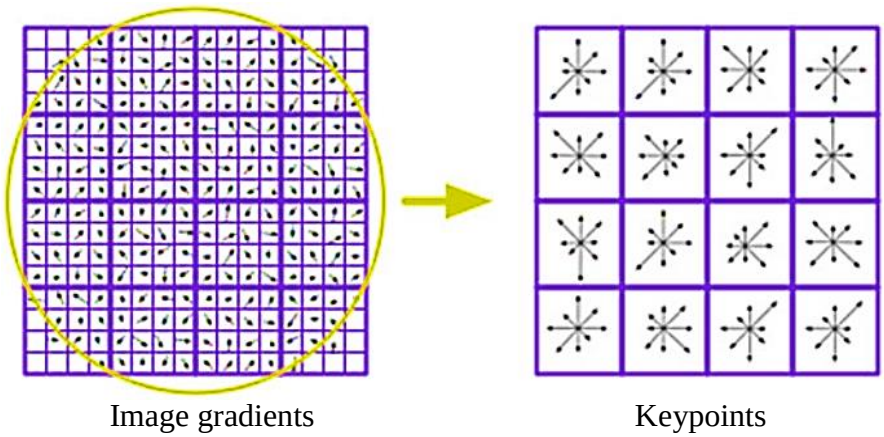


Figure 2.5 Diagram illustrating the construction of the SIFT descriptor (Adel, Elmogy, & Elbakry, 2014)

The **SURF** algorithm also implements a float descriptor. It is faster in descriptor computation than **SIFT** because it adopts wavelet response to assign orientation in both vertical and horizontal directions. A neighbourhood of 20 by 20 is taken around the Keypoints. For feature description, it divides this neighbourhood into 4x4 sub-regions of which horizontal and vertical wavelet responses are taken and a vector, $v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$, is formed. **SURF** computes a 64-dimension feature vector which is half the dimension of the SIFT feature vector. Gaussian weights are applied to it and plotted in a space as illustrated in Figure 2.6. To estimate the dominant orientation, all responses within a sliding orientation window at an angle of 60° are summed up. Remarkably, the use of integral images makes it quite easy to find wavelet response at any scale (Bay et al. 2006).

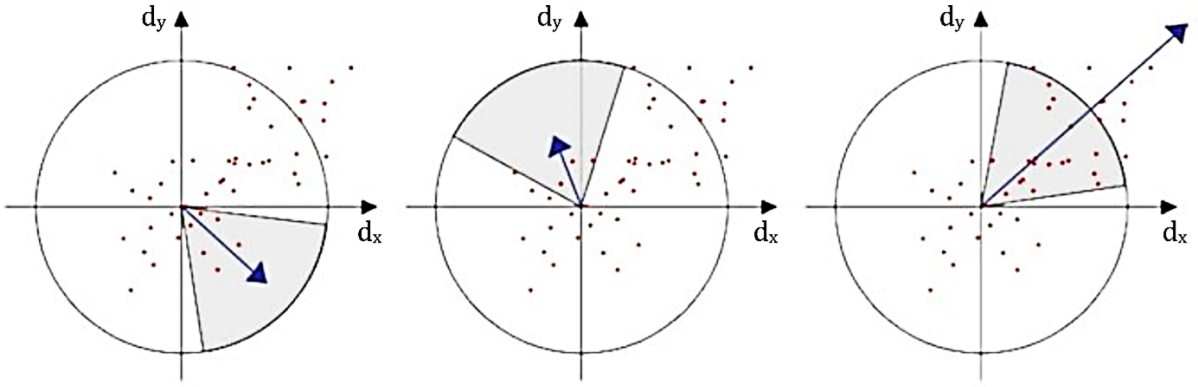


Figure 2.6 Diagram illustrating orientation estimation of the SURF descriptor (Yuan, 2017)

Another state-of-the-art registration algorithm that implements a float descriptor and is worth mentioning is the KAZE algorithm. It is a multi-scale 2d feature detector and descriptor in nonlinear scale spaces. Since algorithms such as SIFT and SURF are unable to preserve object boundaries due to smoothening at all scales, the KAZE solves this by detecting features in nonlinear scale spaces (Fernández et al. 2012), but is computationally expensive.

b) Binary descriptors

The steady increase in the need to process data quickly and efficiently has invoked the need to develop computationally inexpensive descriptors. To this end, the binary descriptors provide a workable alternative to the widely used float descriptors since they offer similar keypoints description performance at a reduced computational cost. The binary descriptors encode features as zeros or ones, and store this information in a vector form where each digit represents the results of an intensity comparison of a pixel-pair. Figure 2.7 illustrates a binary descriptor.

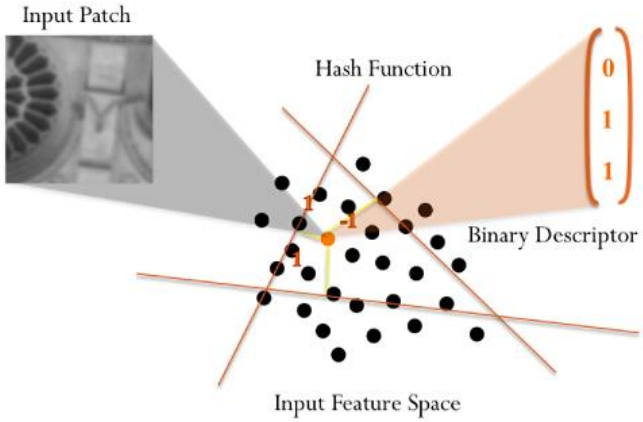


Figure 2.7 Diagram illustrating how binary descriptors are encoded (CVLAB, 2015)

The structure of the binary descriptors encompasses a sampling pattern and sampling pairs. The pattern defines the neighbourhood around the Keypoint for description while the sampling pairs qualifies pixel-pairs to use for intensity comparison (Levi & Hassner, 2015).

The **Binary Robust Elementary Features (BRIEF)** by Calonder, Lepetit, Strecha, and Fua, 2010 is arguably the 'mother' of binary descriptors. It describes a Keypoint by comparing intensities of random pixel-pairs around the detected feature. The comparison yields binary strings which can be matched quickly and easily using an XOR operation. Since **BRIEF** only compares intensities, it is much faster to extract than the float descriptors that compute histograms and gradients (Levi & Hassner, 2015). As much as the BRIEF algorithm is suitable for real-time applications is suffers a setback; it is not as reliable and robust since it has minimal tolerance to in-plane rotation and scale change (Leutenegger, Chli, & Siegwart, 2011).

The **Binary Robust Invariant Scalable Keypoints (BRISK)** is another good example of a binary descriptor; it was developed by Leutenegger, Chli, & Siegwart in the year 2011. It uses evenly spread sampling points on a set of scaled concentric circles whose sizes are directly proportional to the standard deviation (σ) of the Gaussian filter applied to each point. Orientation is computed from the sampled pixel pairs as shown in equation 7. Thereafter, all the computed local gradients are summed up for all long pairs (a pair of sampling points that are beyond a set minimum threshold) and the short pairs (a sampling pair less than a maximum threshold) are rotated by this orientation angle thus making the descriptor invariant to rotation.

$$g(p_i, p_j) = (p_i, p_j).\frac{I(p_j, \sigma_j) - I(p_i, \sigma_i)}{\|p_j - p_i\|^2} \quad (7)$$

Where *g(pi , pj)* is the local gradient between a sampling pixel-pair *(pi , pj)* and *I* is the smoothed intensity derived after applying a Gaussian filter.

The **BRISK** descriptor is constructed by computing comparisons between a pair of short pixel -pairs as shown in equation 8.

$$b = \begin{cases} 1, & I(p_j^{\alpha}, \sigma_j) > I(p_i^{\alpha}, \sigma_i) \\ 0, & otherwise \end{cases} \quad (8)$$

Where $p_j^{\propto}, p_i^{\propto}$ are short pixel-pairs whose intensities are compared.

A value of one (1) is assigned when the first point in a pair has an intensity larger than the second point. When the vice versa is true, a value of 0 is assigned.

2.3.3. Feature matching

Feature matching is the process of finding corresponding features between a pair of overlapping images. This is achieved by comparing the descriptor values of paired keypoints using the descriptor distance (Strecha et al. 2011). The descriptor distance is a measure of similarity and not a metric distance. Various methods are used to compute descriptor distances such as L1 Norm, L2 Norm and Hamming distances. Further, the type of descriptors being matched dictates which similarity measure to use. For instance, float descriptors are compared using L1 and L2 Norm distances whereas binary descriptors are compared using Hamming distances.

Two feature vectors match when the distance between them is less than the threshold set by the *"MatchThreshold"* parameter. The matching method can either be exhaustive or approximate. The exhaustive method matches features "A" to all the features in "B" by computing the pair-wise distance between feature vectors in "A" and "B". On the other hand, the approximate method matches features "A" to the nearest in features "B" using an efficient approximate nearest neighbor search.

2.3.4. Outlier Removal

Outliers refer to observations or measurements that acutely deviate from other observations. In image registration, outliers are point pairs that do not contribute in estimation of the transformation matrix. Discarding of outliers is done by setting a metric threshold or margin within which to sample the inliers. A well-known method for outlier removal is the **Random Sample Consensus (RANSAC).** The Random Sample Consensus has been instrumental in computer vision and photogrammetry domains. RANSAC has offered solutions to the

Keypoints correspondence problem, and accurate estimation of the transformation matrix between successive cameras. The RANSAC, developed by Martin Fischler and Robert Bolles in 1981, fits a model to a sample dataset; it picks a random sample of paired keypoints based on a set threshold and estimates the transformation matrix between the two images pairs using the inliers only. The main idea is to find the optimal partition of points in the inlier set and estimate the model. RANSAC is a non-deterministic approach; its results are only to a certain probability, with this probability increasing as more iterations are allowed (Fischler & Bolles, 1981). The RANSAC has a three-stage standard approach: that is, (i) sampling the number of data points required to fit the model, (ii) computing the model parameters using the sampled data points, and (iii) scoring by the fraction of inliers within a preset threshold of the model. These three stages are done iteratively until the best model with a high confidence level is found. Figure 2.8 illustrates how RANSAC really works.

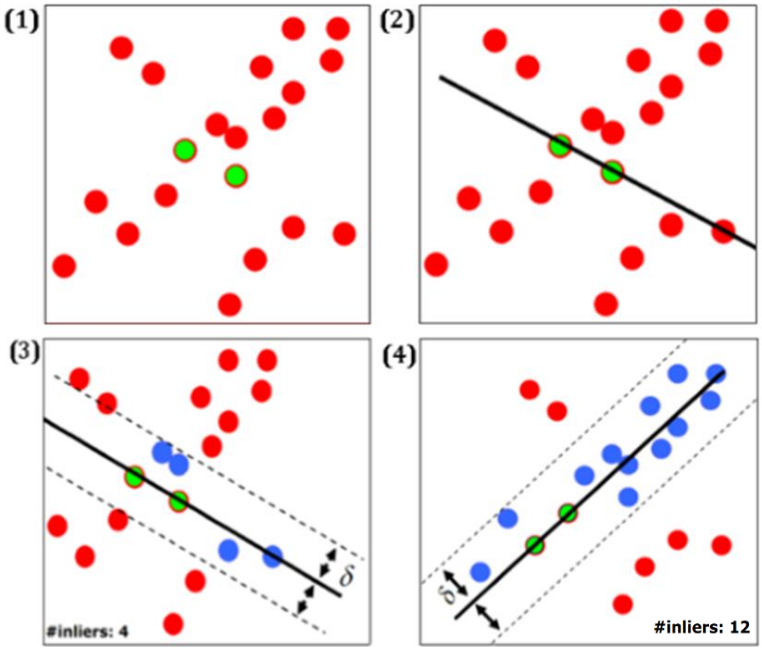


Figure 2.8 RANSAC standard approach: (1) Sample Selection (2) Model Parameters estimation (3) Inlier selection within set threshold (4) Optimal result after iteration (Xiaoyu Yuan, 2017)

2.4. Related works

The use of multitemporal series of multispectral UAV imagery to monitor crops has attracted agricultural practitioners and researchers in equal measure; most research works have demonstrated that increased spatial resolution has improved the ability to extract more precise crop details. López-Granados et al. (2016), use multispectral UAV imagery in early detection of grass weed within grass crop, and Yemane-Tumlisan (2017), extracts vegetation indices, models crop heights, and estimates biomass using multitemporal series of UAV-based RGB imagery. An evaluation of the application of UAV imagery for intra-field nitrogen monitoring and yield estimation by (Song, 2016), reveals that a correlation of the vegetation indices and biomass are higher for UAV imagery than of Rapid-Eye sensor. Similarly, Haghighattalab et al. (2016) carry out phenotyping of wheat using UAV-based multispectral imagery; high correlation between VIs and ground truth measurements was observed.

Co-registration of multispectral imagery using proprietary software has been demonstrated to still have significant misregistration errors (Jhan et al., 2016; Rey et al., 2013); this has attracted researchers to propose methodologies and develop algorithms to resolve this problem. Kelcey & Lucieer (2012) suggest procedures to calibrate the six bands Mini MCA camera including radiometric correction, noise reduction and affine transformation for simultaneous image registration and correction of lens distortion. The scale invariant feature transform (SIFT) has been utilized for extraction of feature points and matching for the band-to-band registration of the Mini MCA images (Turner et al. 2014). Aband-to-band methodology to register and orthorectify the Mini MCA12 imagery has been developed by Jhan, Rau, and Huang (2016). They propose the use of interior orientation parameters (IOPs) of each camera, and the relative orientation parameters (ROIs)

between the master camera and the slaves to find conjugate features between the image pairs. Closely related to this, Suárez et al. (2009) carry out aerial triangulation and orthorectification on a single band and estimate the exterior orientation parameters (EOPs) of the five other bands using the relative orientation between the master camera and the other five. In addition, Mesas-Carrascosa et al. (2016) assess optimal UAV flight parameters for generating accurate orthomosaics of the Micro-MCA Tetracam images; and propose increasing the forward and lateral overlaps from 60% and 30% to 70% and 40% respectively. Misregistration due to parallax issues for multi-camera sensors like the Red edge Micasense and Parrot Sequoia have been investigated; a robust and adaptive band-to-band image transform is proposed for multi-lens co-registration (Jhan et al., 2017).

Monitoring applications have attracted research on co-registration of multitemporal UAV imagery. Aicardi, Nex, Gerke, & Lingua, 2016, propose an image-based approach for co-registration of multitemporal imagery; the results attained are comparable to GCP-based strategies. Wei et al. 2017, proposed a four-steps method for multi-temporal image co-registration for monitoring of dynamic agricultural terraces. In an effort to enhance crop monitoring, Turner et al., 2014b developed a methodology to co-register UAV-acquired visible (RGB), multispectral and thermal images; the matching algorithm was based on the Scale Invariant Feature Transform (SIFT). Phase correlation method has been used to determine overlapping areas in images, thus estimation of rotation and translation between reference image and slaves (Li Tang et al., 2016).

3. EQUIPMENT, DATA & SOFTWARE

3.1. Equipment

This research seeks to accurately co-register and assess the data quality acquired by the Micro-MCA Tetracam camera mounted on the Matrice 600 UAV, and the Parrot Sequoia camera mounted on the Phantom 4 UAV.

3.1.1. UAVs - Phantom 4 and Matrice 600 Pro

The Phantom 4 UAV platform is a quadcopter with an advanced stereo vision positioning system giving it accurate hovering even without GNNS support. It has a magnesium skeleton which reduces weight while keeping stiffness at a maximum to minimize vibration. It weighs 1.4kgs, has a payload limit of 500g, and a battery and power management system that ensures flight endurance of up to 28 minutes. Phantom 4 is equipped with an obstacle sensing system to avoid collisions during flight. Finally, the intelligent flight-controller records and stores flight settings and data for each acquisition; making it possible to reuse flight parameters for monitoring applications (DJI, 2016c).

On the other hand, the Matrice 600 is a multi-rotor drone that integrates a flight controller, two transmission systems, and a battery management system for maximum performance and quick setup. Its battery management system ensures flight endurance of up to 38 minutes. With a maximum takeoff weight of 15.5kgs, it is able to fly with a range of sensors to meet various monitoring needs. The flight controller can be improved with two additional GNSS receivers and IMUs (Inertial Measurement Units) or with dynamic differential real time kinematic (D-RTK) GNSS for enhanced accuracy (DJI, 2016b). Figure 3.1 shows a complete set up of the two UAVs used in this study in the study site.

Figure 3.1 Phantom 4 with parrot sequoia (left); Matrice 600 pro with micro MCA Tetracam (right)
SOURCE: Author

3.1.2. Parrot Sequoia and Micro MCA Tetracam

The Parrot Sequoia multispectral sensor captures the electromagnetic spectrum in four separate parts: green, red, red-edge and NIR with spectral bands in the range of 530 – 570, 640 – 680, 730 -740, and 770 – 810 nanometers respectively. It incorporates the GPS, IMU & magnetometer thus increased accuracy of data capture; integration of the GPS into the sensor makes it fully self-directed; the IMU compensates for positional and orientation variations of the sensor due to drone movements; and the magnetometer maintains the heading of the sensor, it indicates north to the sensor and thus act like a compass. In addition, Sequoia integrates an "irradiance sensor which is mounted on top of a drone to continuously record light conditions in the same spectral bands as the multispectral sensor" (Micasense, 2017). With a dimension of 59mm x 41mm x 28mm, the camera is suitable for most drones. Figure 3.2 shows a detailed description of the Parrot Sequoia camera.

Figure 3. 2 Parrot Sequoia (MicaSense, 2017)

On the other hand, the micro MCA multispectral camera used in this study has six separate cameras. Each camera is synchronized with the other cameras so that each is able to capture the same scene at the exact same time of exposure. During each exposure instant, six (6) separate channels of visible or NIR radiation move through each lens and filter to form separate monochromatic images on each sensor. The images are simultaneously transferred from each sensor to separate camera internal memories for later post-flight processing (Tetracam Inc., 2017). Figure 3.3 shows the micro MCA Tetracam used in this study and a diagram illustrating the architecture of each spectral sensor.

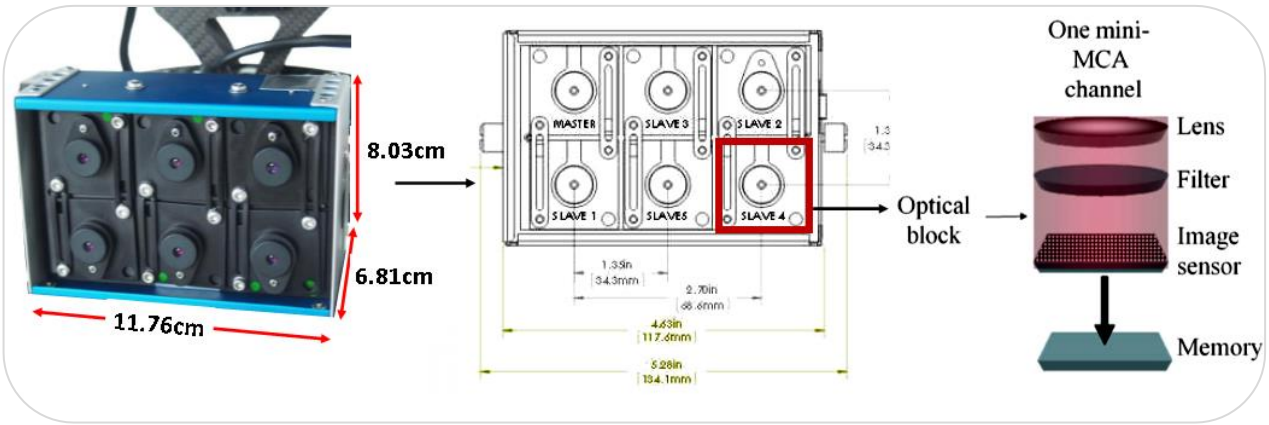


Figure 3.3 Micro MCA Tetracam (Tetracam Inc., 2016)

The Micro MCA is equipped with band-pass filters with a full width of 10nm to restrict the radiation that contacts the sensor to a narrow band of wavelengths. The spectral wavelength specifications of each band of the Micro MCA Tetracam and the Parrot Sequoia is shown in table 3.

BAND	PARROT SEQUOIA	MICRO MCA TETRACAM
BLUE	-	410 - 490nm
GREEN	530 - 570nm	510 - 590nm
RED	640 - 680nm	630 - 710nm
RED	-	660 -740nm
RED EDGE	730 - 740nm	730 -740nm
NEAR INFRARED	770 - 810nm	760 - 800nm

Table 1. Wavelength specifications of Parrot Sequoia and Micro MCA Tetracam

Further comparison of Parrot Sequoia and Micro MCA Tetracam is presented in table 4. As shown in tables 3 and 4, the two cameras differ in spectral band widths and other camera specifications like the focal length.

Specifications	Parrot Sequoia		Micro MCA Tetracam
	RGB	**Multispectral**	**Multispectral**
Lenses	1	4	6
Focal Length	4.88mm	3.98mm	9.6mm
Spectral Range	400 – 700nm	530 – 810nm	450 – 800nm
Pixel Size	1.34µm	3.75µm	5.2 µm
Resolution (Pixels)	4608 x 3456	1280 x 960	1280 x 1024
FOV (H° ×V°) (m)	64.6 × 50.8	62.2 × 48.7	84 x 67
Camera Weight	135g		530g
Shutter type	Global		Rolling
Camera Size (cm)	6 × 4 × 3		12 x 8 x 7

Table 2. Specifications of the Parrot Sequoia and the Micro MCA Tetracam

3.2. Software

The photogrammetric workflow was based on Pix4D, PixelWrench2 (Tetracam Inc.) and Agisoft Photoscan (Agisoft LLC, St. Petersburg, Russia) for first and second acquisitions with the Parrot Sequoia, and third acquisition with the Micro MCA Tetracam respectively. Pix4D and Agisoft photoscan have three major steps: (1) aerial triangulation and image-to-image matching; (2) Point cloud densification and DSM generation; and (3) Orthorectification and orthomosaicking. PixelWrench2 on the other hand was used to produce multi-page TIFF (Tagged Image File Format) files from the RAW (i.e. image as seen by the camera, flat and dark) Micro MCA Tetracam image files. The PixelWrench2 software carries out an inter-band alignment using affine transformation based on a global transformation matrix for the camera.

The image registration framework is based on image processing and computer vision libraries and toolbox in **MATLAB 2017b**. Graphical plots and statistical analysis was done using the Matlab platform as well. Other software packages used are shown in table 5.

Software	***Use***
Pix4d Capture (Mobile App.)	Flight planning
ENVI Classic 5.3	Raster Processing
Erdas Imagine 2016	Raster Processing
ArcGIS 10.5.1 (ArcMap)	Map production
Microsoft Word 2016	Report writing
Microsoft Excel 2016	Statistical reporting and analysis
R Statistical Package	Statistical analysis and graphical visualization

Table 3. List of Software

3.3. Data

3.3.1. Image Acquisition and Data properties

The site of image acquisition was a maize field (approximately 25 acres in area) south west of Gronau city, Germany (52° 10'N, 6° 55'E). Image acquisition was done in three (3) time-steps. All acquisitions took place between 11am and 1pm.

Image acquisition marked the beginning of the workflow in this project. It was done in three (3) time-steps. Three acquisitions were made with the sequoia and one with the Micro MCA Tetracam. Both UAV systems were flown in the third epoch. The flight lines of every epoch are shown in Figure 3.4. The change of flight line pattern in epoch 3 was not an informed decision. However, all the other flying parameters were maintained as in epoch 1 and 2.

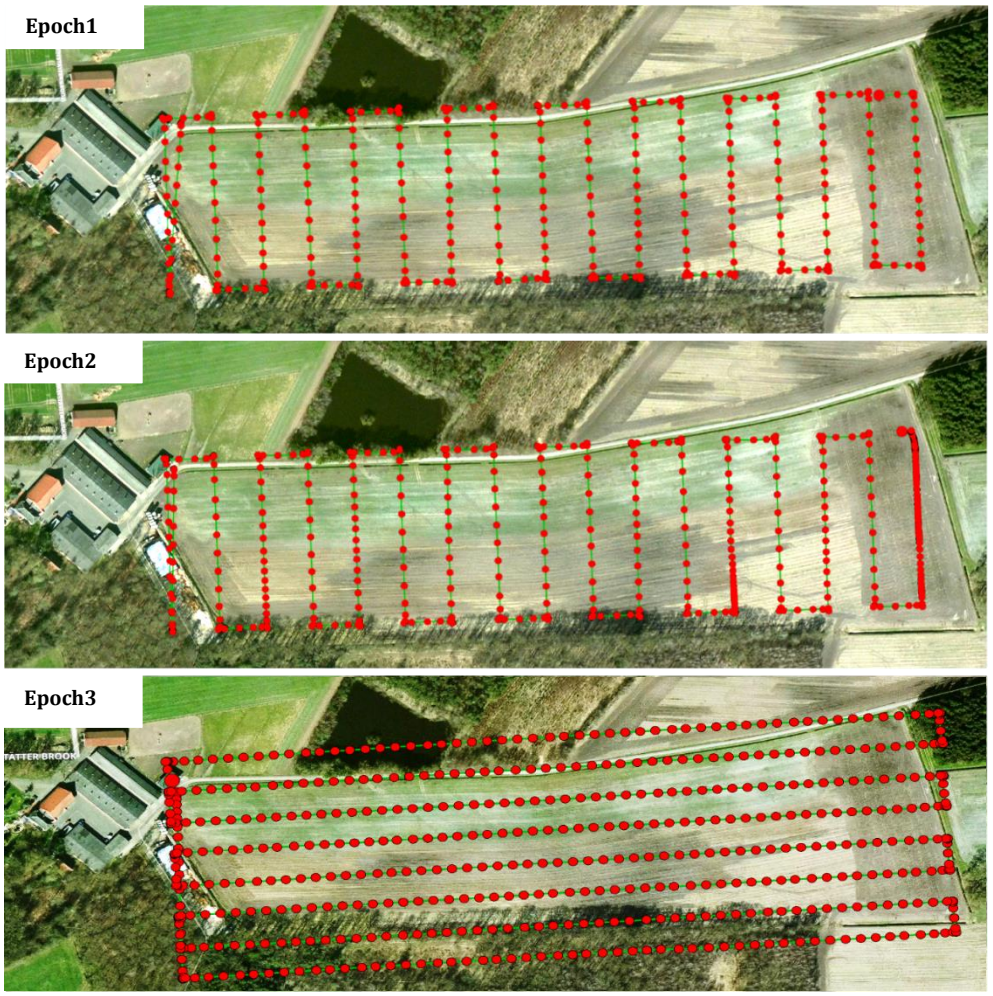


Figure 3.4 Flight lines used for image acquisition superimposed on google satellite basemap
SOURCE: Author

The flight parameters like forward and lateral overlap were optimally set to enable accurate 3D reconstruction. 98% to 100% of all the acquired images were successfully calibrated. Image acquisition details are shown in table 6.

Acquisition Date	Camera	Flying Height	Forward Overlap	Side Overlap	GSD	No. of images
8th Aug, 2017	Sequoia	50m	80%	40%	5.01cm	2092
		70m	80%	40%	6.84cm	1476
11th Aug, 2017	Sequoia	70m	80%	40%	6.69cm	1724
19th Sept, 2017	Sequoia	50m	80%	40%	5.02cm	2094
	Micro MCA	100m	80%	40%	4.64cm	1656

Table 4. Image acquisition details and image properties

Orthomosaics were generated for both Parrot Sequoia and Micro MCA; the orthomosaic for each channel is shown in Figures 3.5 and 3.6 respectively. See appendix 1 and 2 for orthophotos from other acquisitions and the quality report.

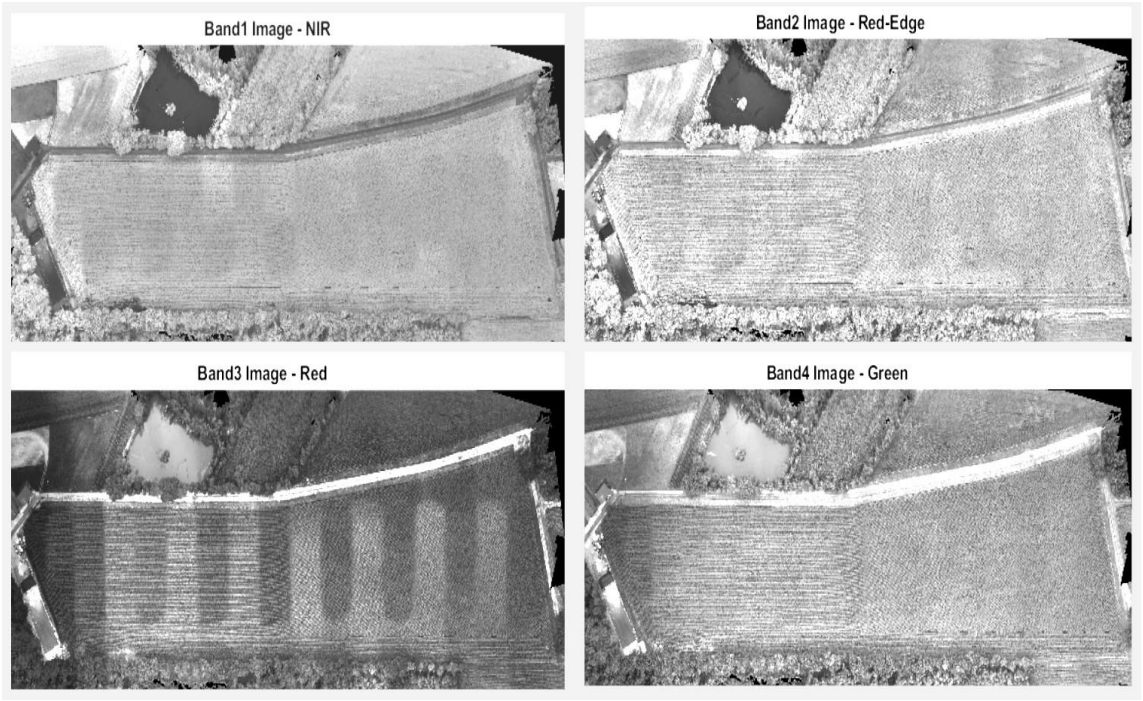


Figure 3.5 Orthomosaics of four channels of Parrot Sequoia (Epoch1) SOURCE: Author

Having failed to initialize the GPS, micro MCA Tetracam images were not geo-tagged. In addition, the acquired images were hazy possibly due to the high speed of the Matrice 600, hence presenting another interesting challenge in this research. Photogrammetric processing using Pix4D was not successful therefore Agisoft Photoscan was used to align the images as shown in Figure 3.6. Automatic tie point detection and image matching was only possible for a set of fifteen (15) images thus a subset of the farm as depicted in Figure 3.6. The non-georeferenced orthomosaic was later co-registered and projected to the spatial reference of the master image.

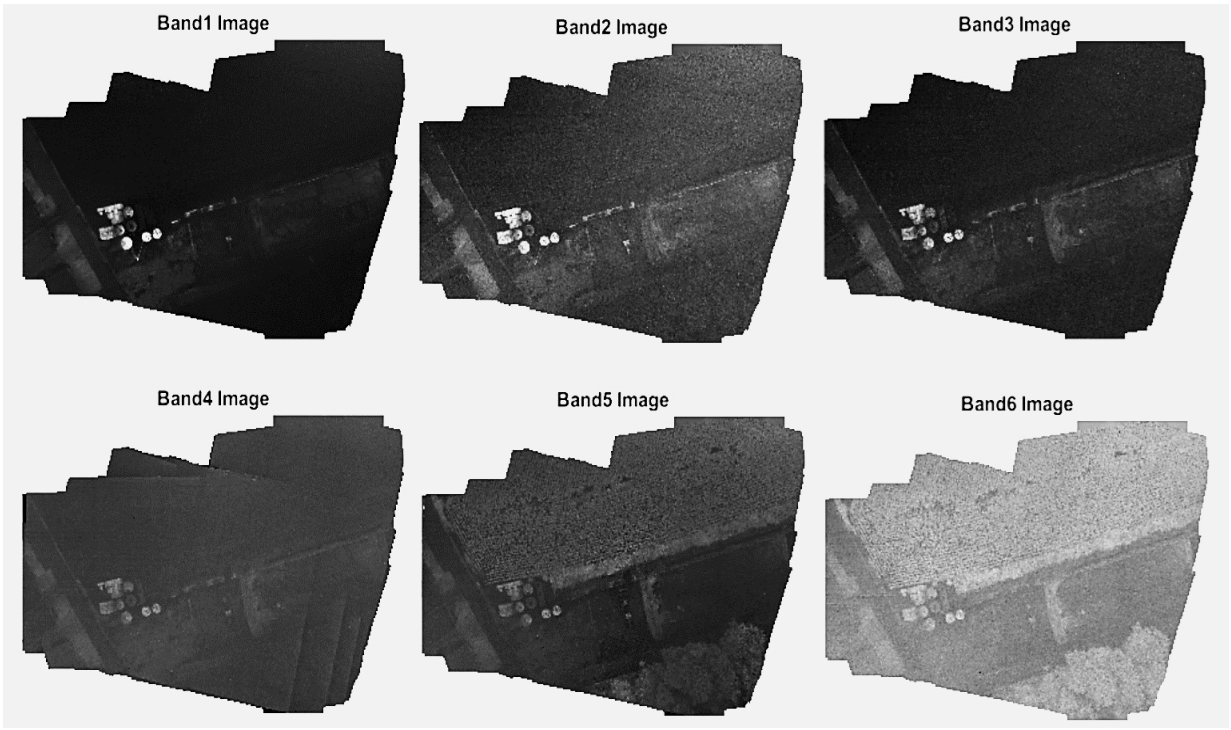


Figure 3.6 Orthomosaics of the six channels of the Micro MCA Tetracam SOURCE: Author

4. METHODOLOGY

This chapter gives a detailed description of the approaches taken and methods used to realize the main objective of co-registration of multitemporal series of multispectral imagery for crop monitoring. The experimental studies are discussed in this chapter as well.

4.1. Adopted Methodology

The general overview of methods, processes, decisions, intermidiate and final outputs are captured in the flowchart in Figure 4.1.

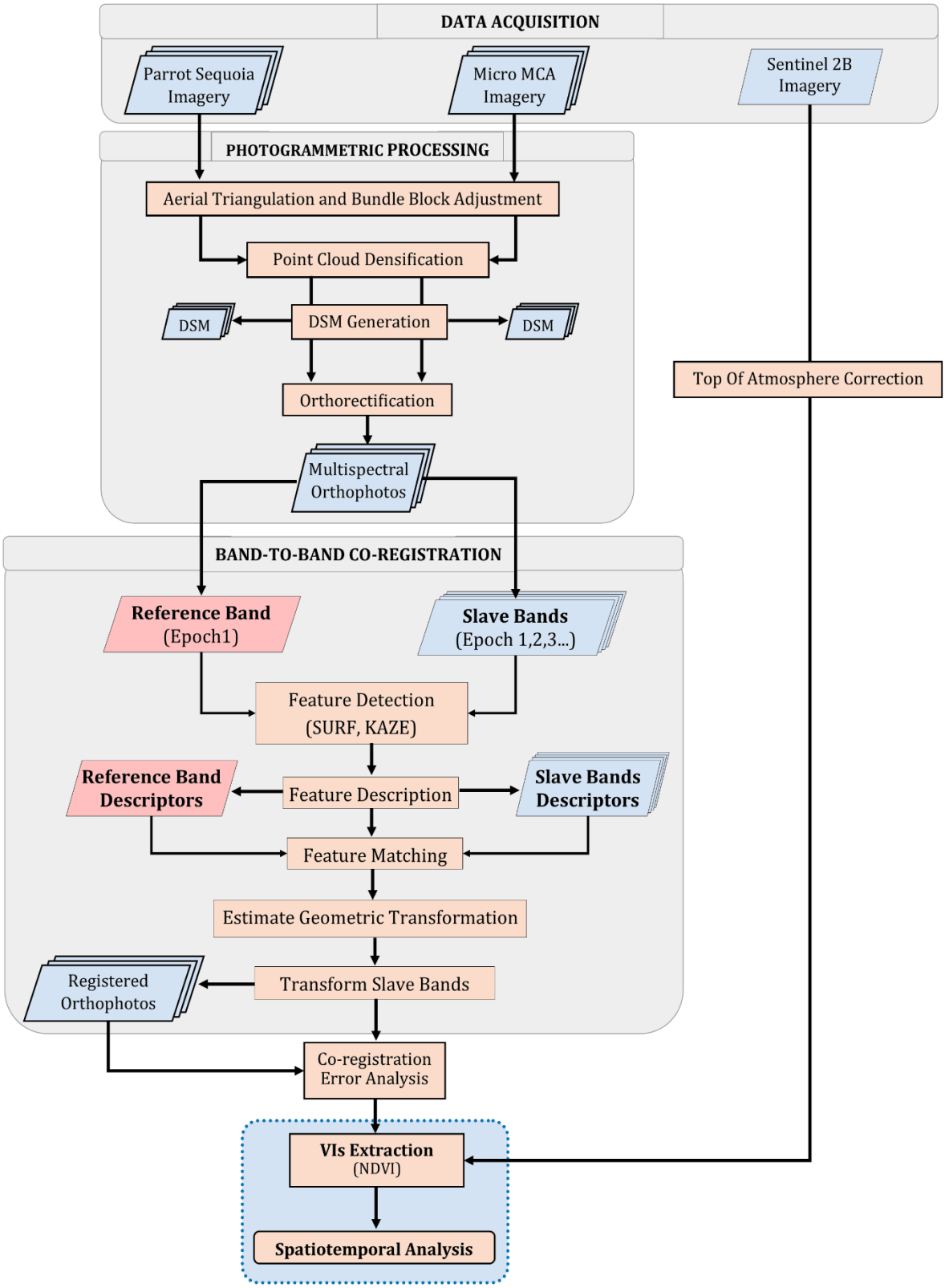


Figure 4.1 General overview of adopted methodology

4.2. Photogrammetric workflow

Pix4D was used for most of the photogrammetric workflow. Two approaches were used to perform the photogrammetric workflow; they include, processing all the bands together using the multispectral template in Pix4d, and using a customized template mimicking the RGB template to process each band separately. The photogrammetric products of these approaches were compared; the misregistration was quantified (See results in chapter 3). Aligning the images of Micro MCA Tetracam was done using Pixel Wrench II (PW2); it uses a calibration file that contains the relative orientation between the master and slave channels to align all the six channels at the image level. PW2 was then used for conversion of RAW images to multipage TIFs which were then processed in Agisoft Photoscan which mainly has four stages like Pix4d; that is, aerial triangulation, point cloud densification, DSM generation, and orthorectification.

4.2.1. Initial Processing

In Pix4D, the aerial triangulation phase is referred to as initial processing. The first image registration is done in this stage; keypoints detection and extraction, and matching of individual images is done. In this stage, Pix4D offers a user interface where desired processing settings can be adjusted to meet specific needs; the keypoints image scale allows the user to define the image size used to extract the keypoints. The options for Keypoints matching are two: matching image pairs, which dictates which pairs of images are matched; and the matching strategy which basically determines the matching method. Finally, calibration is done by optimizing all the internal and external parameters of the camera. The output of this stage is preliminary 3D reconstructed points. Figure 4.2 shows a successful and unsuccessful outputs; the red points indicate images at camera stations that have not been calibrated making it impossible to determine their relative orientation. In this project, the Ground Control Points were not used in processing, due to logistical and technical issues. The GCP's are meant to minimize systematic errors and deformations in images, stabilize bundle solutions, and determine correct 3D reconstruction (Nex & Remondino, 2014). However, the lack of GCPs did not hamper this study since the acquisition of the first epoch was assumed to be the reference epoch, and registration assessments of subsequent acquisitions were based on the first epoch.

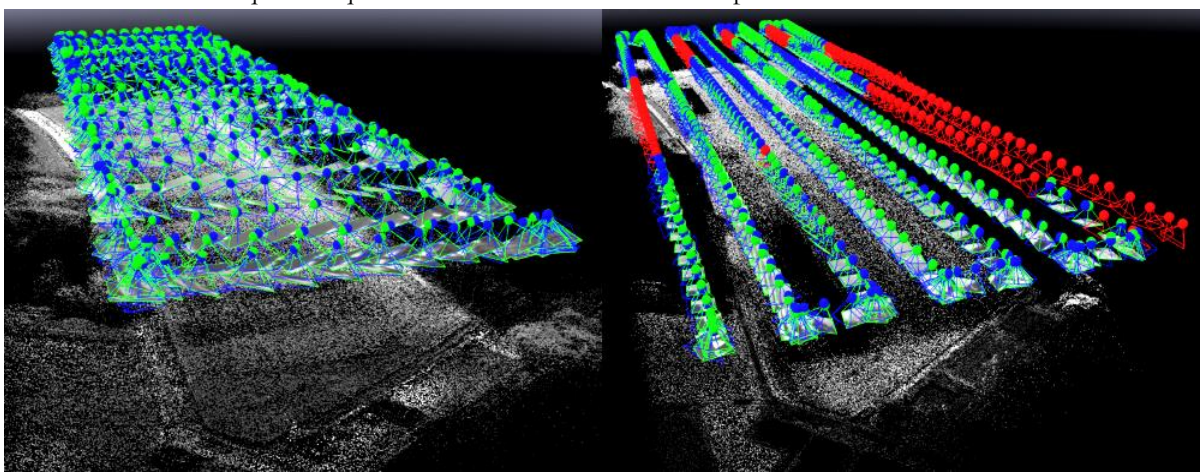


Figure 4. 2. a) A snippet of successful initial processing b) A snippet of unsuccessful initial processing
SOURCE: Author

4.2.2. Point Cloud Densification

The 3D points generated in the initial processing stage are sparse and insufficient to estimate planes and geometry in the image scene. For this reason, there is need to generate dense points sufficient enough for this purpose. For this phase to yield accurate results, the camera network needs to be optimal; all cameras need to have been calibrated in the initial processing stage. Non- calibrated cameras lead to errors in 3D reconstruction. See Figure 4.2 for the effect of non-calibrated camera within the image block. The failure in calibrating some images can be attributed to poor image quality at those specific locations therefore impeding automatic keypoints generation and consequently making it impossible to relatively orient these images with the others. A sample of dense point clouds yielded in this stage is shown in Figure 4.3.

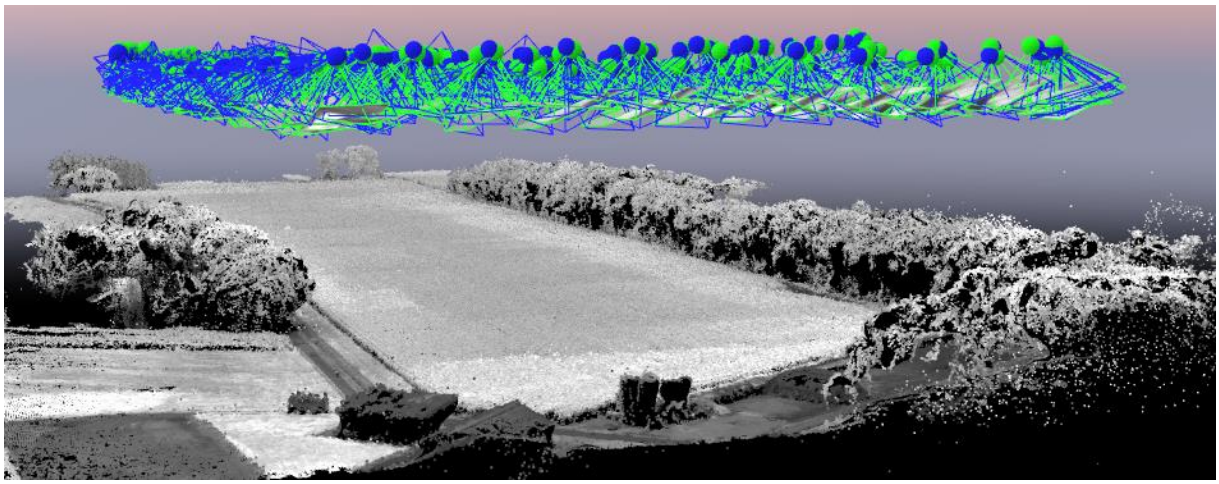


Figure 4.3 Showing densified point cloud and mesh SOURCE: Author

4.2.3. Digital Surface Model and Orthophoto generation

The densified point clouds generated in sub-section 4.2.2 play a role in this stage: the point clouds are used to generate the Digital Surface Model (DSM). The DSM is thereafter used as an input to generate the orthophoto. It is important to note that the multispectral template for image processing in Pix4D yields orthomosaics of each camera/multispectral channel separately, therefore having four orthophotos (for the case of Parrot Sequoia) as shown in Figure 3.5 of chapter three. Figure 4.4 shows the DSM and orthophoto for the green band of epoch 1.

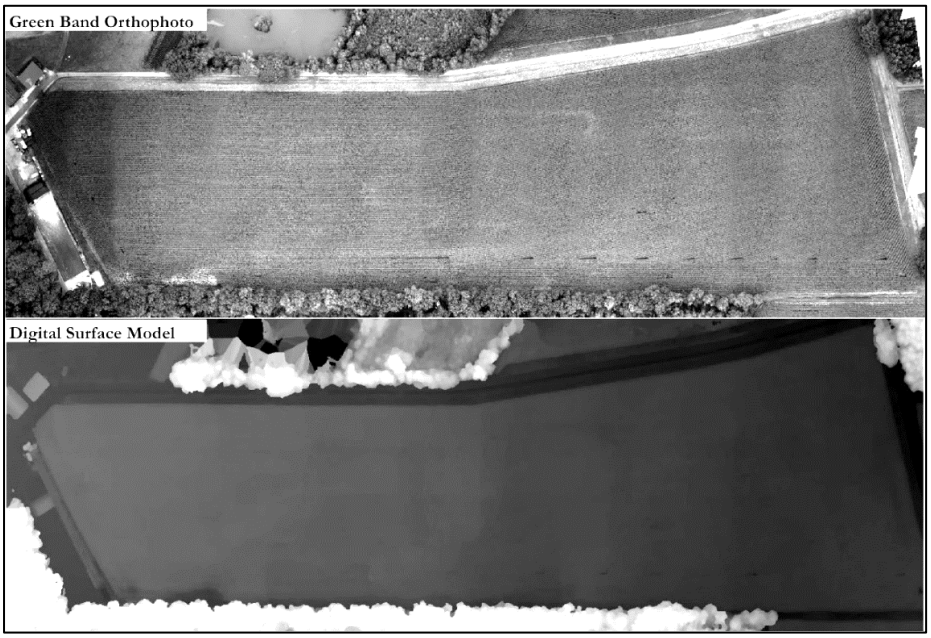


Figure 4.4 A snippet of the sequoia green band orthophoto (top) and DSM (bottom) of epoch 1 SOURCE: Author

4.3. Image Co-registration

Accurate image co-registration is vital for multitemporal analysis. As mentioned earlier, misregistration of spectral bands will result in extraction of erratic vegetation indices; misclassification of crop phenology; and interpolation errors in values between available observations among others. This study investigates intra-epoch and inter-epoch misregistration errors and consequently optimally parameterizes keypoints detection algorithms to perform band-to-band co-registration. It is imperative to align the orthophotos for comparative analysis. Four major region detectors, already implemented in Matlab 2017b, were selected for this research; they include SURF, BRISK, MSER and KAZE. The building blocks of these algorithm are presented in chapter two. Co-registration experiments were done at two levels; the single images and orthophotos per band. The nature of the experiments and their aims are discussed in subsequent sections.

4.3.1. Feature Detection

The first phase of image co-registration is feature detection. The keypoints were detected in all the bands using the default settings shown in table 7.

Algorithm	*Parameters*		*Algorithm*	*Parameters*	
SURF	**MetricThreshold**	1000	*MSER*	**ThresholdDelta**	**2**
	NumScaleLevels	4		**RegionAreaRange**	[30 14000]
	NumOctaves	3		**MaxAreaVariation**	0.25
BRISK	**MinContrast**	0.2	*KAZE*	**Threshold**	0.0001
	MinQuality	0.1		**NumScaleLevels**	4
	NumOctaves	4		**NumOctaves**	3

Table 5. Default feature detection parameters of the selected algorithms

In SURF, the strongest feature threshold is specified by a positive scalar, the *'MetricThreshold'*. The higher the value, the less blobs are detected (See Figure 4.5). This value therefore determines which keypoints will be detected. The number of octaves to implement, *'NumOctaves',* are also specified as a non-negative scalar. Recommended values are between 1 and 4. The higher the number of octaves the larger the detected blobs. Each octave spans a number of scales that are analyzed using varying size filters; higher octaves use larger filters. The scale levels, *'NumScaleLevels',* refer to the number of levels per octave to compute; a positive integer between 3 and 6 is recommended. A default of 4 implies that the input image is successively downsampled by a factor of 2 until an image pyramid with four images is formed. The higher the scale levels (i.e. finer scale increments) the more the number of blobs detected (MathWorks, 2017a).

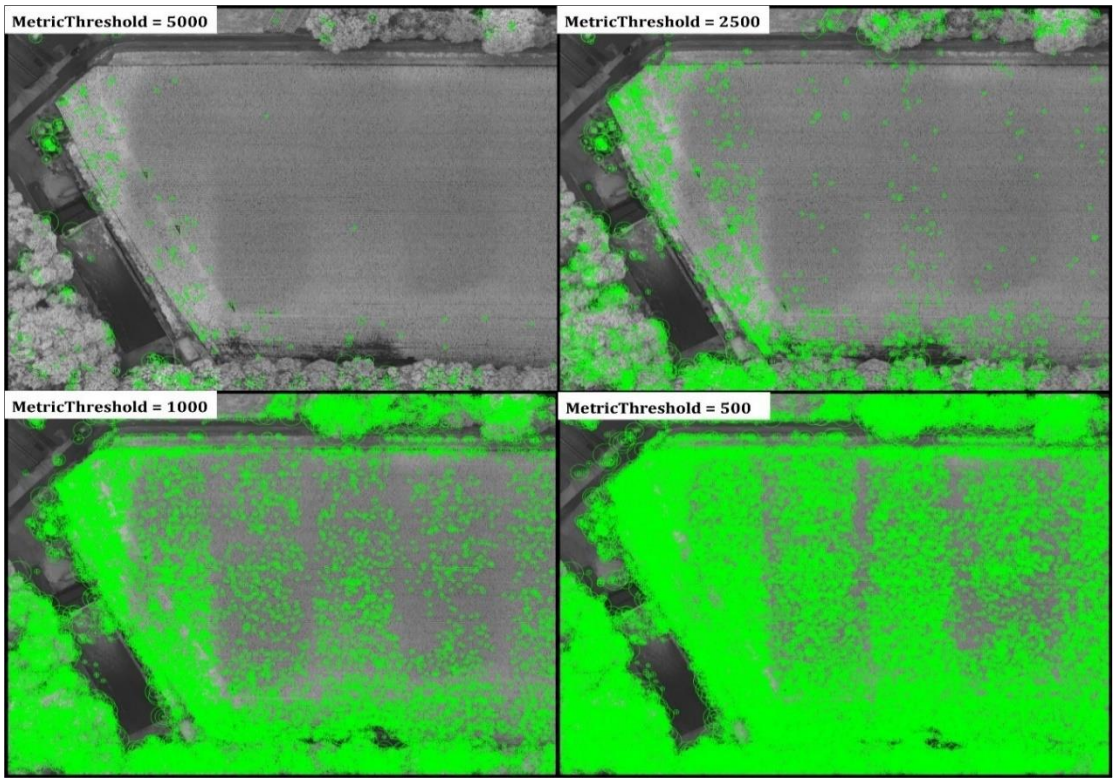


Figure 4.5 Illustrating the effect of different metric thresholds on feature detection. SOURCE: Author

The minimum contrast, *'MinContrast',* specifies the minimum intensity difference between a region and its immediate surrounding. It is a scalar in the range of zero (0) and one (1). An increase in this value would lead to a decrease in the number of blobs detected by BRISK. Similarly, the minimum quality, *'MinQuality',* ranges between zero (0) and one (1); it denotes the minimum accepted quality of detected regions. When the value tends towards one (1) erroneous blobs are removed (MathWorks, 2017a).

The *'ThresholdDelta'* is the step size between intensity threshold levels in the range of [0 100]. It is expressed as a percentage of the input data type used in selecting external regions while testing for their stability. Typical values range from 0.8 to 4. The lower the value, the more regions are detected as blobs. In MSER, the size of the region, *'RegionAreaRange',* is a two-element vector denoting the minimum and maximum areas of regions in pixels to be allowed in the detection process. At varying intensity thresholds, the Maximum area variation between extremal regions, *'MaxAreaVariation',* is specified by a positive scalar between 0.1 to 1. An increase in this value results in detection of more external regions which may be less stable.

The KAZE algorithm uses three methods to compute conductivity; a property that makes diffusion of the image luminance adaptive to the local image structure. The 'region' method is the default setting, it uses the Perona & Malik, 1987 conductivity coefficient (see equation 9); this study employs the 'region' method. The other methods are the 'sharpedge' and the 'edge' which use the Perona & Malik conductivity coefficient in equation 10 and the Weickert conductivity coefficient respectively. Details on the Weickert conductivity coefficient have been discussed in the paper *'Anisotropic Diffusion in Image Processing'* by Weickert, 1998. The conductivity function *g* is described in two formulations as follows:

Eq. 9 $$g_1 = \frac{1}{1+\frac{|\nabla L\sigma|^2}{k^2}}$$ Eq. 10 $$g_2 = \exp\left(-\frac{|\nabla L\sigma|^2}{k^2}\right)$$

$\nabla L\sigma$ is the luminous function; it is the gradient of a Gaussian smoothed version of the original image L; k is a contrast factor that controls the level of diffusion. "The g_1 function promotes wide regions over smaller ones whereas g_2 promotes high-contrast edges" (Alcantarilla, Bartoli, & Davison, 2012).

In KAZE, the local extrema is a function of the Hessian threshold. The *'Threshold'* is specified as a scalar greater than or equal to zero (0). An increase in this value excludes less significant local extrema. The multiscale detection factor, *'NumOctave',* and the scale levels are scalars in the range of 3 to 10. Larger features are detected by increasing the multiscale detection factor whereas smoother scale changes and additional intermediate scales between octaves are realized by increasing the scale levels.

4.3.2. Experimental studies on Feature Detection

The parametric settings applicable in all the selected algorithms were tuned in a bid to finding the optimal parameterization for co-registration of multispectral UAV imagery. The experiments discussed in this section informed the decisions made to develop the final co-registration. The performance of the selected algorithms, SURF, BRISK, MSER and KAZE, was investigated under different independent parametric circumstances to determine the effect of tuning a given parameter, and the performance of each algorithm in feature detection. The intensity threshold is a vital parameter but it was not adjusted due to difference in its scalar range across the selected algorithms. However, the general trend of this parameter has been discussed briefly in sub-section 4.3.1. The two major parameters that were tuned include the number of octaves and the scale levels. Holding all other parameters constant, these two parameters of interest were investigated autonomously.

a) **Number of Octaves: T**his parameter gives the algorithms the scale invariant property by subsequently downsampling images by a scale factor of 2 until a pyramid is formed. Using the defaults of the other parameters as shown in table 7, the octaves were varied from 1 to 4 as recommended by Matlab documentation. The octave number selected refers to the number of image sets needed to create a hierarchical pyramid. The number of octaves is not a parameter for the MSER thus only the performance of three algorithms were compared. Figure 5.4 in chapter five shows the effect of tuning the number of octaves; the results are presented in chapter five.

b) **Scale Levels:** The scale levels to compute per octave have a direct bearing on the number of features to be detected. The number of scales determines the number of images per set of octaves. It means, estimation of LoG at different decomposition levels of the original image. LoG therefore acts as a blob detector in various image scales due to change in the standard deviation which acts as the scaling parameter.

4.3.3. Feature description

The descriptor extraction method can be set automatically or changed each time to match the detector used. The automatic option was preferred; it selects the method based on the class of the input points (which are the outputs of the feature detection phase). For any of the extraction methods automatically detected, the function sets the orientation property of the output object (i.e. valid points) to the orientation of the extracted features in radians.

The rotation invariance is specified as a logical scalar, *true* or *false*. Setting the rotation invariance to *true* constrains the estimation of the orientation of the feature vectors; the feature orientation is set to $\pi/2$. The logical scalar *true* is only used when the intention is not to capture the rotation information of the descriptors. On the other hand, *false* allows for the estimation of orientation of the features and thus becoming invariant to rotation.

The default descriptor size is 64. However, an option of making it 128 exists. This option applies only when the function implements the SURF or KAZE method. The larger feature size of 128 provides greater accuracy, but decreases the feature matching speed (MathWorks, 2017b).

Feature description is a function of the neighbouring pixel; it is done by extracting the intensity gradients of the neighbouring pixels and saving them as the unique identity of the center pixel (detected Keypoint). This neighbourhood can either be 64 or 128 in dimension. Intuitively, a neighbourhood of 128 should give more accurate descriptors, but for this study, the descriptor size of 64 was used since it results into more indexed pairs that can be used for matching. The results of SURF and KAZE are presented in the next chapter.

4.3.4. Feature matching

Matching of descriptors is the penultimate stage of image co-registration. In this stage, the unique identities of the descriptors are indexed and paired. Two matching methods exist, the *Exhaustive method* which computes the pairwise distance between feature vectors 'A' and 'B'; and the *Approximate method* which uses an efficient approximate nearest neighbor search. The latter is appropriate for large feature sets. The *Exhaustive* method was used for experiments in this study.

To select strong matches, a matching threshold is specified; it represents a percent of the distance from a perfect match. Two feature vectors are a match when the distance between them is less than the set threshold. A match is rejected when the distance between the features is greater than the set threshold value. In this research, defaults of 10 for binary feature vectors or 1 for non-binary feature vectors was used. The higher the matching threshold, the more matches obtained (not necessarily 'good' matches).

The feature matching metric is either the sum of absolute differences (SAD) or sum of squared differences (SSD). The SSD was used in all the experiments in this study. The feature matching metric only applies when the input feature sets are not binary feature objects. When the input is binary feature objects, the Hamming distance is used to compute the similarity metric. Something worth noting is that binary inputs (e.g. BRISK features) require a larger match threshold value to return 'good' matches.

The one-to-one and one-to-many matching approaches are parametric arguments in the matching function. The one-to-one matching approach was used; only the unique matches between features 'A' and 'B' were returned. This means that multiple features in 'A' can match to one feature in 'B' for the case of one-to-many matches (See Figure 4.6), and only one feature in 'A' matched to one in 'B' as illustrated in Figure 4.7. Only the unique matches (one-to-one) was preferred because it increases the chances of getting only correct matches and estimating an accurate transformation matrix. Removal of the wrong matches and outliers is discussed in section 4.3.5.

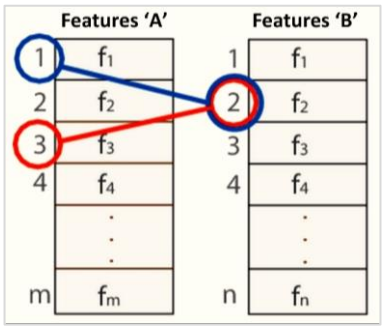


Figure 4.6. Illustrating one-to-many matching

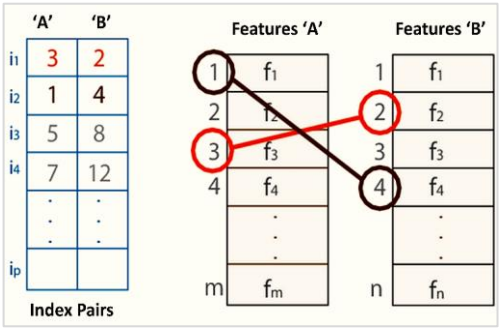


Figure 4.7. Keypoints indexing and one-to-one matching

All experiments employed the *exhaustive* brute force method for matching. The default settings of the matching threshold, intensity ratio threshold, and the feature matching metric were maintained. However, tests on one-to-one and one-to-many matches were done to assess the best matches for estimating the transformation matrix. When the *Unique* argument is set as *true,* the function performs a forward-backward match to select only unique matches between features 'A' and 'B'. Conversely, *false* returns all matches between features 'A' and 'B'. The results are presented in the next chapter informed the use of *Unique* = *true* in this research.

4.3.5. Removal of Outliers and Estimation of transformation matrix

In this study, the outliers in matched points 'A' and 'B' were excluded using the M-estimator Sample Consensus (MSAC) algorithm which is a variant of the RANSAC (Torr & Zisserman, 2000). RANSAC suffers a setback; it is sensitive to the threshold that defines inliers and outlier. A very large threshold tends to rank all the hypotheses equally and qualify them as good for the fitted model. Conversely, a very small threshold tends to be unstable in estimating parameters. That is, adding or removing a datum to the set of inliers may lead to fluctuation of estimated parameters (Torr & Zisserman, 2000). The MSAC (presented in equation 11) partially compensates for this undesirable effect. It penalizes the outliers equally but scores the inliers on how well they fit the data.

$$\rho_2(e^2) = \begin{cases} e^2 & e^2 < T^2 \\ T^2 & e^2 \geq T^2 \end{cases} \qquad \text{(Eq.11)}$$

Where ρ_2 is the robust error term, and T is the threshold for considering inliers.

Using the inliers selected for model fitting by MSAC, the transformation matrix was estimated. The *'Estimate Geometric Transformation'* function in Matlab simultaneously removes the outliers and estimates the transformation matrix. It takes in a set of putative matches, and randomly selects the best set of matches to fit the model, and computes the transformation matrix between the inlying points. The geometric transformation is either affine 2D for the case of similarity and affine transformation methods and projective 2D for projective method. The 2D similarity transformation method was used in this study because it retains angles and length ratios, and because our orthophotos are already planimetric and geometrically similar.

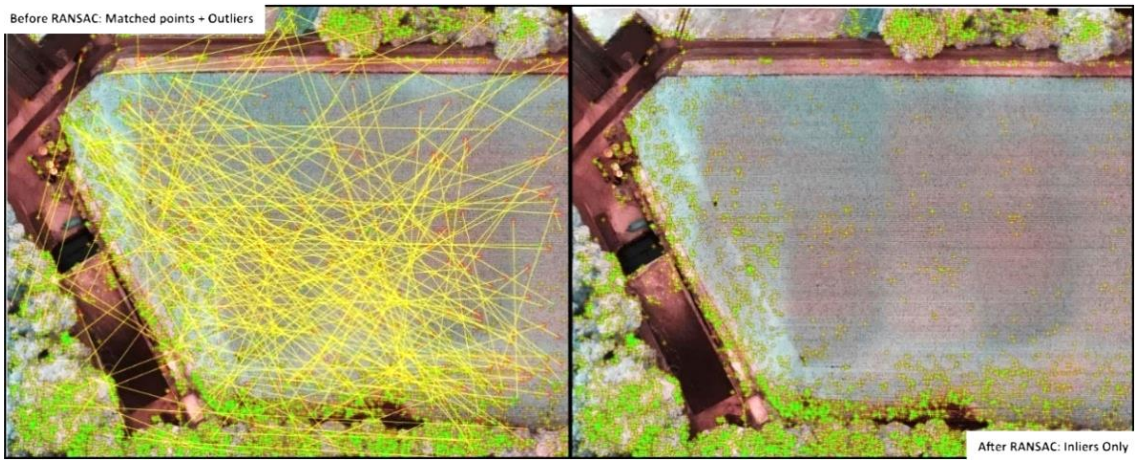


Figure 4.8. Showing putative matched plus outliers (left), and only the correct matches and inliers (right)

Figure 4.8 illustrates a before and after scenarios of invoking the RANSAC algorithm. The first case shows all putative matches including the mismatched pairs shown by linked crisscrossing yellow lines. The second images show only the inliers based on the set distance threshold; these are only the correct matches.

The outlier matches were defined by a distance threshold between features 'A' and 'B' upon inverting the geometric transformation. Only points that meet this threshold were used to compute the transformation matrix. The distance threshold, *MaxDistance* in Matlab, was tuned to understand how it affects the transformation matrix. Figure 4.9 illustrates the *MaxDistance*. The higher the projection distance threshold the higher the residual error (See results in chapter five).

Estimation of the transformation matrix was done at the image and orthophoto level. At both levels, the Transformation matrix was compared element by element, and quantified as a RMSE for comparison between different camera positions. The Transformation matrix was further decomposed to fetch out the rotation and translation (i.e. relative orientation). The displacement between matched inliers exhibit a systematic displacement between cameras. However, transformation matrices indicated that the relative orientation between the cameras is not the same at different camera positions. Results are presented in the next chapter.

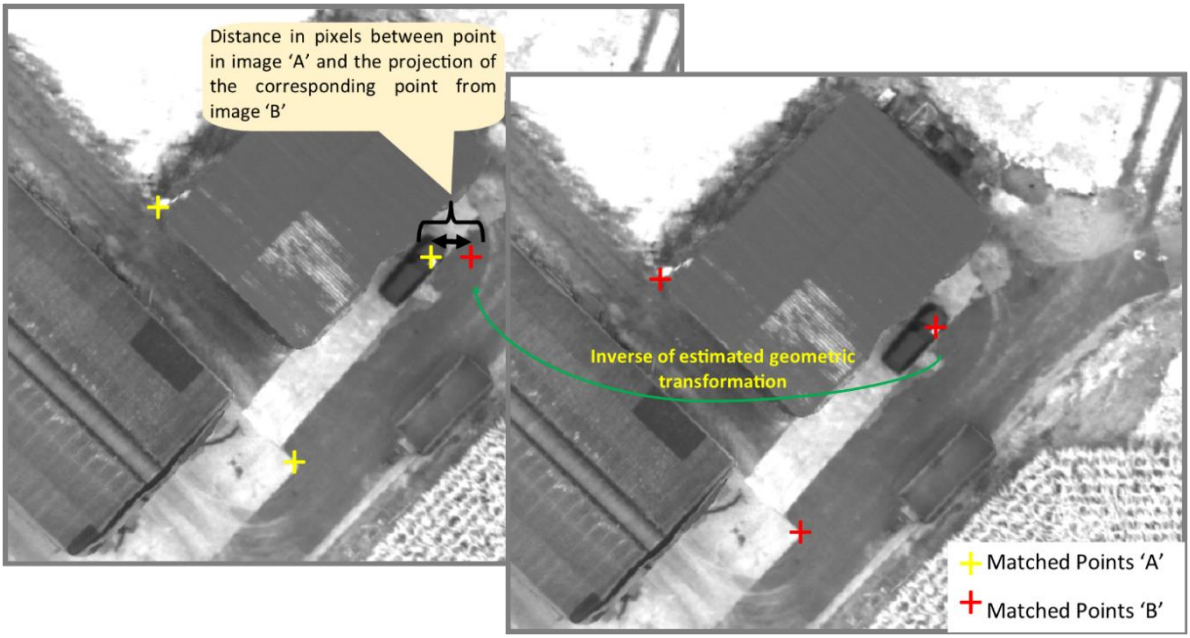


Figure 4.9. Illustrating geometric transformation and the projection distance.
SOURCE: Author

4.3.6. Band-to-band Co-registration

Post processing misregistration assessment revealed significant misalignment between DSMs and orthophotos within the same epoch, as well as between those of different epochs. Intra-epoch misalignment was observed to be systematic (see results in chapter 5) and was attributed to the basis distance between the cameras. Inter-epoch misalignment on the other hand was due to different acquisition conditions (i.e. influence of wind), 'poor' UAV GPS positioning (±5) since GCPs were not used in this study. In both cases, a reference band was selected; all other bands (slaves) were co-registered to the reference (in this case the red edge because most keypoints were detected on it in all epochs). The methodology involved feature detection, feature description, matching of keypoints from other bands to those of the red edge, and autonomously estimating the transformation matrix per image pair. The transformed slave-bands were concatenated with the reference band in epoch one; but in epoch two, the reference band (being the red edge epoch 1) was excluded from the concatenation. The reason for concatenation was to generate a multispectral composite image for all the epochs. The adopted methodology presented in Figure 16 demonstrates a standard approach for both intra-epoch and inter-epoch co-registration.

Since Pix4D's photogrammetric workflow for multispectral images yields separate orthophotos for each spectral channel, band-to-band registration tests were inevitable. The tests were done at the image and orthophoto level. In both cases, intra-epoch and inter-epoch tests were done to establish the best band combination for co-registration of multitemporal series of multispectral UAV imagery.

a) Intra-Epoch registration

In an effort to investigate misregistration within a single acquisition, it was necessary to establish the best band to use as the reference. Feature detection, descriptor extraction and feature matching tests were done to rank the best band. It is important for the reference band to have sufficient number of keypoints to be matched with features extracted from the other bands. The bands sensitive to vegetation spectral response (i.e. Red edge, NIR and Green) yielded comparatively better results than the red band. Tests at the image level were done to see the performance of the bands in a built-up section of the farm. Results are presented in chapter five.

b) Inter-Epoch registration

Registration of multispectral imagery between epochs presents an interesting research problem. Once the registration within the first acquisition was done, it was imperative to accurately align the subsequent acquisitions to the first one. Having established that the red edge was the reference band, the inter-epoch registration involved estimating the transformation matrix between all the subsequent bands of each acquisition with the red edge band of the first acquisition; geometrically transforming them; assessing their pairwise registration accuracy; and stacking them together per epoch. This was termed as 'many-to-one' registration; i.e. there is only one master image (See Figure 4.10a). The aim was to achieve subpixel registration accuracy between the reference image and all the slave images.

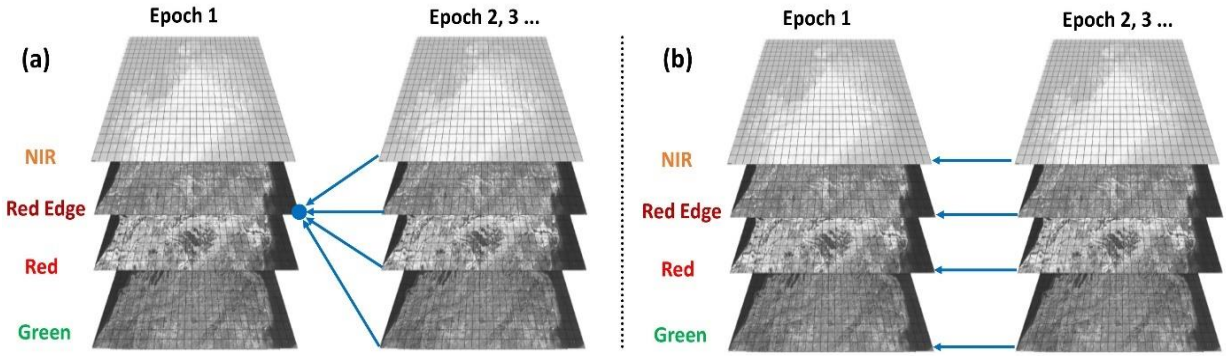


Figure 4.10. (a) Many-to-one registration (b) One-to-one registration. SOURCE: Author

Another approach involved using all the bands in the first acquisition as the reference, and align similar bands in the subsequent acquisitions. Each band in the first acquisition was considered a master to the corresponding subsequent bands (See Figure 4.10b). By the virtue of subsequent acquisitions aligning to the corresponding band at a subpixel level, then the co-registration was considered successful. Accuracy assessments was done by computing the distance between matched features before and after the registration

4.3.7. DSM co-registration and Analysis

The Digital Surface Model is a vital photogrammetric product for crop monitoring. From the DSM, crop height models are retrieved; thus, monitoring the phenology of the particular crop over time. As is the case for orthophotos, multitemporal DSMs are misaligned due to the same reasons highlighted in section 4.3.6. Therefore, for accurate comparative analysis and change detection purposes, co-registration of the DSM is inevitable. Two methods were used to co-register the DSMs from different epochs; manual and automatic. Manual co-registration involved on-screen identification of stable tie points between the two DSMs. The tie points were 'tied' manually using the georeferencing tool in ArcGIS. The automatic method used the corresponding transformation matrix estimated during orthophoto co-registration of each band pair. The DSM generated in epoch one was considered as the reference; and the other as the slave. Longitudinal profiles along the same line of stable features were used to assess the accuracy of the co-registration.

4.3.8. Co-registration Accuracy Assessment

To measure and quantify the co-registration accuracy of the Parrot Sequoia and the Micro MCA Tetracam sensors, the misregistration error among the multispectral band was measured. This was done by computing the projection distance between the inlying point pairs. In essence, the horizontal positional positions of inliers before and after registration was used to assess the co-registration accuracy. This method was preferred because the points used in both cases are not necessarily the same. That is, paired points before co-registration were used to compute the distance apart thus the displacement. After co-registration, new keypoints were extracted and matched. The distance between the second set of matched points was used to assess the co-registration accuracy (See Figure 4.11). The positional accuracy and the distance between detected features between bands was expected to be less than half a pixel after co-registration. An ideal situation would be to have a zero value after co-registration. The accuracy assessment was done at the image and orthophoto level. Both results are presented in chapter three.

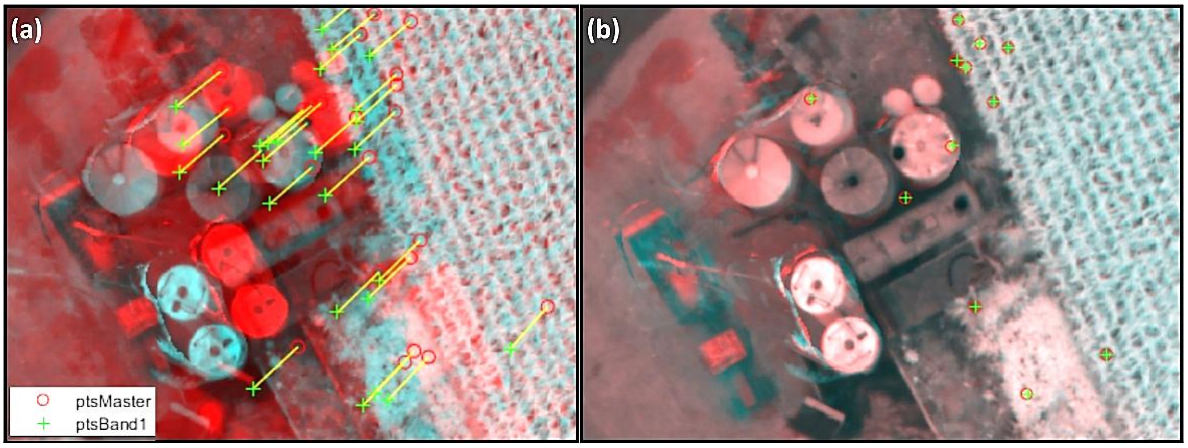


Figure 4.11. Showing distance between matched features before (a) and after (b) co-registration
SOURCE: Author

Accuracy assessment was done for two co-registration approaches: many-to-one and one-to-one. Assessing these two approaches would inform the best approach for co-registration of multitemporal series of multispectral images. Since Ground Control Points (GCPs) were not used in this case, the first acquisition was considered the reference. Horizontal positional root mean square error (RMSE) was used to evaluate the registration accuracy. The RMSE was computed by finding root of the average of the set of squared differences between coordinate values of inlying slave and master keypoints, where the master refers to the first epoch. The positional RMSE was computed as shown in equations 12 – 14.

(Eq. 12) $$RMSE_x = \sqrt{\sum_i \frac{(Xslave - Xmaster)^2}{n}}$$ (Eq. 13) $$RMSE_y = \sqrt{\sum_i \frac{(Yslave - Ymaster)^2}{n}}$$

(Eq. 14) $$RMSE_r = \sqrt{RMSE_x^2 + RMSE_y^2}$$

The $RMSE_x$ and $RMSE_y$ were used to evaluate systematic displacements in either direction as presented in Figure 5.14 in chapter 5. On the other hand, the combined $RMSE_r$ was used for overall registration accuracy assessment. The closer the value is to zero, the more accurate it is. The registration threshold was 0.5 of a pixel, therefore RMSEs less than 0.5 were considered 'good'.

4.4. Supplementary Experiments

Standalone tests to help achieve the main objectives, and support the argumentation of this research were done. The tests were meant to improve the results of initial experiments and improve the understanding of underlying photogrammetric processes in Pix4D.

4.4.1. Feature Detection using different scenes

Feature detection was done in different scenes within the farm to understand how each spectral band performs under different physical conditions. The aim was to assess the trend of feature detection in different scenes per band. The predominant band for feature detection would then be the candidate for the master channel. Three scenes were selected: a predominantly built up scene; a maize crop scene; a mixed vegetation scene. The results are presented in section 5.1.1 in the next chapter.

4.4.2. Relative Orientation Analysis

In an effort to understand the underlying processes of Pix4D's photogrammetric workflow on multispectral cameras, the camera estimated parameters were analyzed. The principal point coordinates (X, Y, Z) and the rotation angles, omega (ω), phi (φ) and kappa (k), used to rotate the geodetic coordinate system and align it with the image coordinate system of every image at every camera position was analyzed in reference to the master. The goal was to establish the relative orientation of the cameras between each other per station. The same relative orientation between all cameras in all positions would imply computing the orientation once and propagating it to all other images in different positions. The rotation matrix that defines the camera orientation is expressed as follows:

$$\mathbf{R} = \mathbf{R}_x(\omega)\,\mathbf{R}_y(\varphi)\,\mathbf{R}_z(\kappa)$$

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega) & -\sin(\omega) \\ 0 & \sin(\omega) & \cos(\omega) \end{pmatrix} \begin{pmatrix} \cos(\varphi) & 0 & \sin(\varphi) \\ 0 & 1 & 0 \\ -\sin(\varphi) & 0 & \cos(\varphi) \end{pmatrix} \begin{pmatrix} \cos(\kappa) & -\sin(\kappa) & 0 \\ \sin(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} \cos(\kappa)\cos(\varphi) & -\sin(\kappa)\cos(\varphi) & \sin(\varphi) \\ \cos(\kappa)\sin(\omega)\sin(\varphi) + \sin(\kappa)\cos(\omega) & \cos(\kappa)\cos(\omega) - \sin(\kappa)\sin(\omega)\sin(\varphi) & -\sin(\omega)\cos(\varphi) \\ \sin(\kappa)\sin(\omega) - \cos(\kappa)\cos(\omega)\sin(\varphi) & \sin(\kappa)\cos(\omega)\sin(\varphi) + \cos(\kappa)\sin(\omega) & \cos(\omega)\cos(\varphi) \end{pmatrix}$$

If $\mathbf{X} = (X, Y, Z)$ is a 3D point in world coordinate system, its position $\mathbf{X'} = (X', Y', Z')$ in camera coordinate system is given by: $\mathbf{X' = R^T (X - T),}$ where $\mathbf{T}$ is the position of the camera projection center in world coordinates (Tx, Ty, Tz).

4.5. Extraction of Vegetation Indices (VIs)

4.5.1. Spectral Indices of Crops

Spectral vegetation indices are generated by adding, subtracting or dividing information from different spectral bands into a single scalable value. Spectral indices are designed to give an approximate measure of vegetation status (Jensen, 2007). The rationale for spectral vegetation indices is to exploit the unique spectral properties of vegetation; in so doing, exposing the vigour and photosynthetic activity of crops. The Normalized Difference Vegetation Index (NDVI) was the selected index to characterize crop health in this research. NDVI is computed as shown below.

$$\mathbf{NDVI} = \frac{(R_{NIR} - R_{Red})}{(R_{NIR} + R_{Red})}$$

Where R_{NIR} and R_{Red} is the near infrared and red spectral channels respectively. NDVI is a measure of vegetation photosynthetic vigour; it ranges from -1 to 1. Dense and healthy vegetation tend to have strong positive values whereas features like clouds, water and snow will be characterized with negative values.

To statistically assess the NDVI performance, thirteen (13) zones, labelled A – M, were selected (see Figure 4.12). A – H are zones within the maize field; 'I' is the water pond; J is bare ground; K is a roof; L is tarmac; and M is concrete. K, L and M acted as control zones since their NDVI values were expected to be the same or differ slightly.

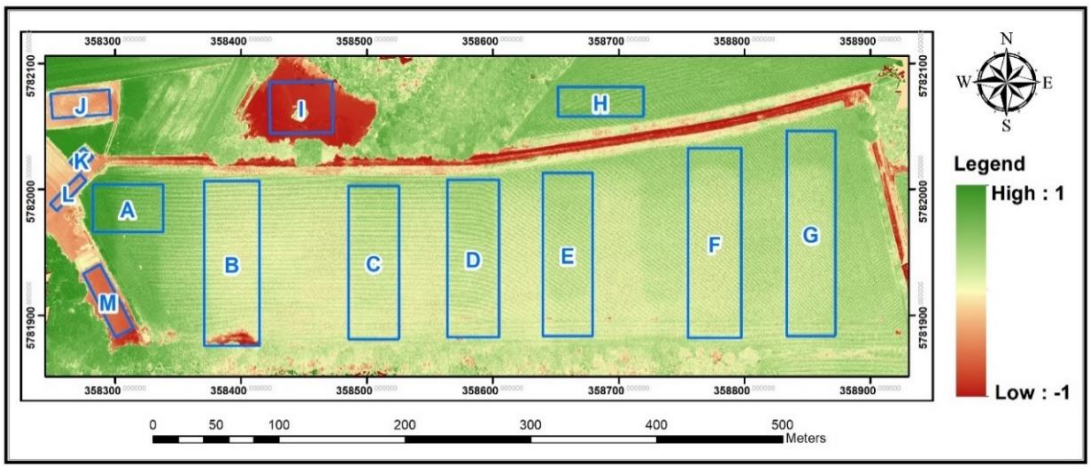


Figure 4.12. Areas of interest used for statistical analysis. SOURCE: Author

Areas A and H cover regions with quite healthy vegetation as depicted by NDVI maps derived from the UAV imagery shown in Figure 4.12. In this study, spectral scaling to reflectance was not done due to the lack of scaling objects on site and lack of ground measurements by the spectro-radiometer. Therefore, the computation of the VIs was based on digital numbers. However, Parrot Sequoia has an irradiance sensor mounted on a drone to continuously record light conditions in the same spectral bands as the multispectral sensor. On the other hand, atmospheric correction is not quite needed for UAV images since they experience minimal atmospheric interference. Since VIs were computed using DNs, the values may not precise, but still provide information on the spatial variability of the vegetation index.

4.5.2. Parrot Sequoia Versus Micro MCA Spectral Analysis

The spectral response of vegetation and the spectral feature (i.e. NDVI) between the two cameras were examined. For spectral feature analysis, NDVI maps for a subset of the farm were derived from both Parrot Sequoia and Micro MCA images acquired in the third epoch. However, they were acquired at different flying heights of 50m and 100m respectively. Thus, the GSD for Sequoia was 5.02cm and that of Micro MCA was 4.64cm, therefore anticipating more or less similar variability in the NDVI values. In this case, only a subset of the farm was used since it was only possible to align a few Micro MCA images covering the section. The Micro MCA images were hazy and were not geotagged. Comparative zonal statistics were computed and compared to those of Parrot sequoia.

To assess the spectral response of vegetation between the two cameras, spectral signatures of two classes of crops (photosynthetically active and less active) were extracted from corresponding composite images of Parrot Sequoia and Micro MCA Tetracam. Spectral sampling was done using the signature editor tool in the classification tool bar in ArcGIS. The signature file was then exported and the mean DN value of every class plotted as shown in section 5.9 of chapter 5.

5. RESULTS AND ANALYSIS

In this chapter, results from the experiments conducted and their interpretation are presented. As described in Chapter 4, the several experiments informed the parameterization of the co-registration algorithm for multitemporal series of multispectral UAV imagery. Experiments were conducted using both default and tuned parameters of registration algorithms.

5.1 Master band Selection and Algorithm performance

Co-registration of a series of images calls for the selection of a reference image; the reference/master is the image upon which subsequent images will be aligned. Since the aim of this research was to investigate intra-epoch and inter-epoch alignment and consequently develop an automated co-registration method as highlighted in objectives two and three of this research, tests were done to rank the bands and choose the master. As mentioned in chapter 4, the tested feature detection and extraction methods include SURF, BRISK, MSER and KAZE algorithms; they are state-of-the-art algorithms used in computer vision applications involving image registration. In addition, these algorithms have already been implemented in Matlab 2017b thus the need to configure Matlab to external libraries did not arise. Moreover, they were sufficient to assess the performance of both binary and float descriptor algorithms in co-registration of multispectral imagery. The master band selection and tests of algorithm performance was done as described in sections 5.1.1 to 5.1.5.

5.1.1. Feature Detection using default parameters

Intuitively, to ensure image pairs that yield maximum similarity between keypoints, the master band should have a sufficient number of valid intensity features to increase the chances of getting correct matches with the slave bands. For this reason, feature detection was done for all the three datasets (i.e. epochs 1-3) using the default parameters of the four feature detection algorithms presented in table 7, page 24. Intensity features were extracted from both Sequoia and Micro MCA images. The performance of the algorithms is shown in Figures 5.1 for Sequoia.

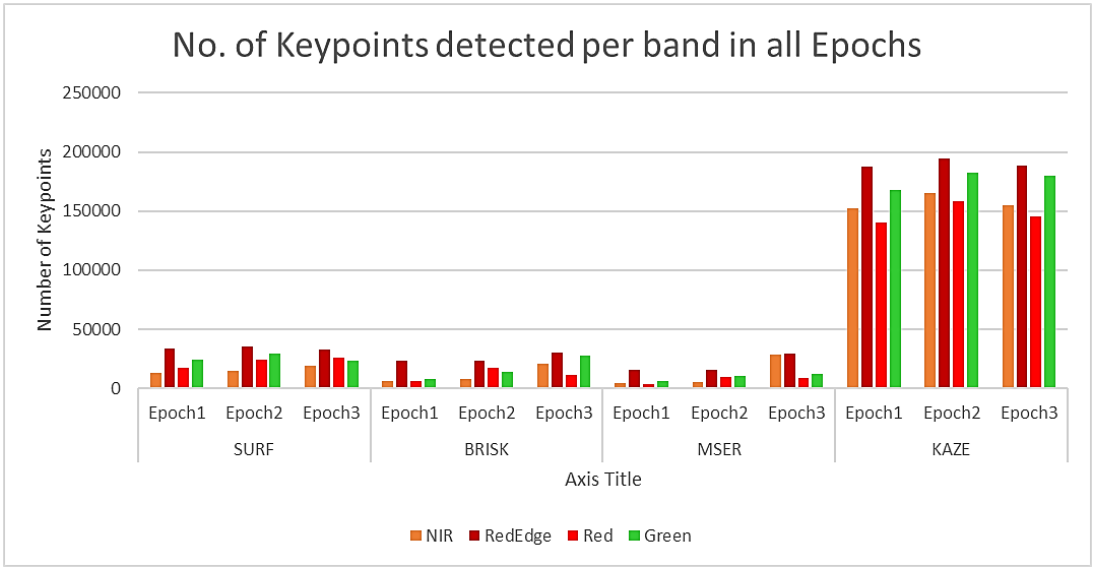


Figure 5.1. Feature Detection per band for Parrot Sequoia images

Figure 5.1 reveals an interesting observation; the bands most sensitive to vegetation spectral response record the highest number of intensity features detected. The red edge stood out as the band from which the most features were detected; the green and near infrared channels follow respectively. The least Keypoints were detected in the red band; spectral response for vegetation is absorbed in the red band thus a possible reason for yielding less intensity features. In addition, Figure 5.1 depicts that KAZE outperformed SURF, BRISK and MSER by detecting three times the number of points detected by other algorithms. KAZE detected over 150,000 points opposed to less than 50,000 points detected by the rest for Sequoia imagery, and over 3000 – 10,000 points against under 1200 points for the Micro MCA Tetracam imagery.

Supplementary tests to validate the detection of features per band was done with the aim to understand the observed trend in Figure 5.1. The tests were performed at a single image level since it is only at this level that the dataset allowed for different scene selection. Feature detection was done in three different scenes; (a) a mixed scene of buildings and crop cover, (b) a homogenous scene of crop cover, (c) a heterogenous scene of vegetation cover (images 146, 246 and 446 of Figure 5.1). Figure 5.2 presents feature detection results obtained per band per scene.

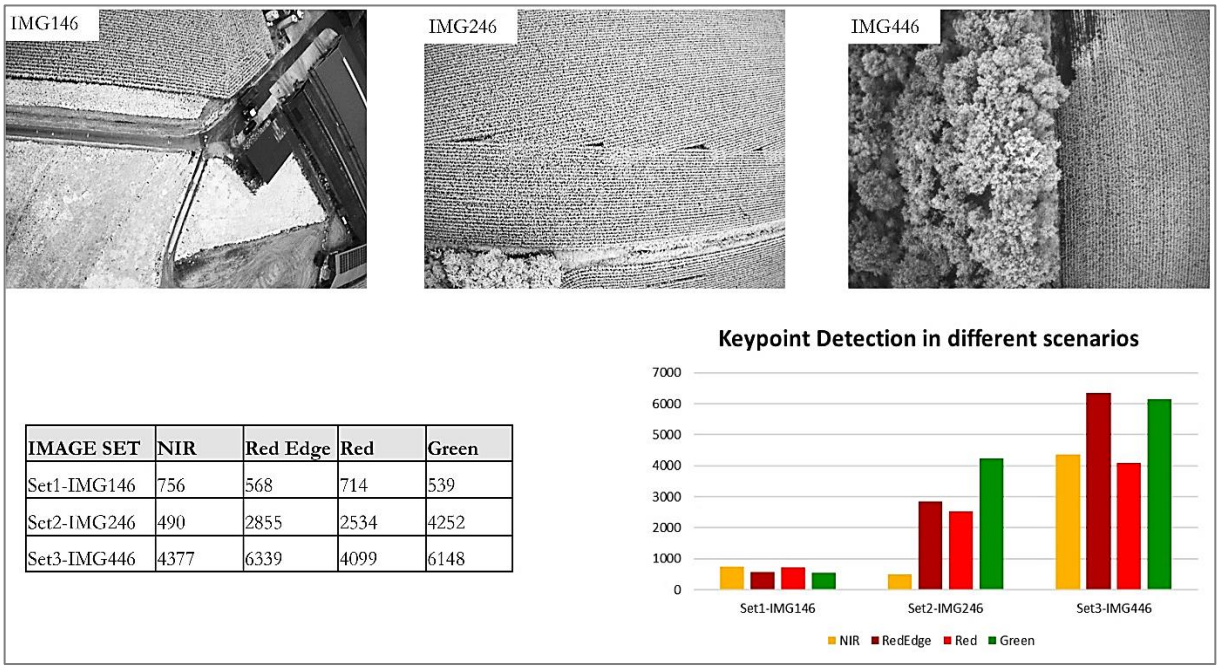


IMAGE SET	NIR	Red Edge	Red	Green
Set1-IMG146	756	568	714	539
Set2-IMG246	490	2855	2534	4252
Set3-IMG446	4377	6339	4099	6148

Figure 5.2. Feature detection per band at different scenes

In scene (a), more Keypoints were detected in the red and NIR bands. However, the performance of the red edge and the green increases in scenes (b) and (c) respectively. The dismal performance of the red edge and green in scene (a) is implicitly due to the evened-out texture and intensity by the building structures thus few local extrema to be detected as keypoints. Conversely, scene (c) offers a highly varying scene in terms of texture and intensity thus detection of more features. The texture forms part of the global features that describe a keypoint. For example, when SURF detects features, it employs vectors describing each of the features shape, orientation, intensity and texture (Hassaballah et al., 2016a).

Similar feature detection experiments to inform the selection of the master band were done using the Micro MCA Tetracam images. The tests were done on an image subset shown in Figure 3.6, page 20; the intention being to investigate if the observation made in Figure 5.1 is replicable using a dataset acquired using a sensor with a slightly different spectral regime. The results for the Micro MCA image are shown in Figure 5.3.

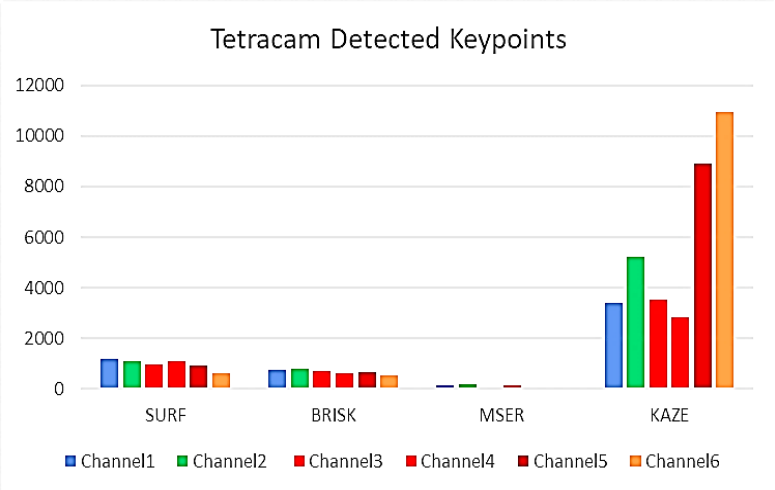


Figure 5.3. Feature Detection per band for Micro MCA

According to Figure 5.3, channel 6 (NIR) was observed to yield most keypoints, followed by channels 5 (Red edge) and 2 (Green) respectively. It was still appreciated that the channels sensitive to vegetation spectral response yield more keypoint features. On the downside, the performance of the feature detection algorithms is seen to have weakened; this could be attributed to the poor image quality of the Micro MCA Tetracam for this particular acquisition. Also, the KAZE algorithm was seen to yield most features in both datasets.

Sub-section conclusion

The red edge, green and NIR channels are the best candidates for the master band since they record the most number of features. Since the next stage of image registration experiments need the master band upon which point matches will be done, the red edge was selected as the master band following the results and analysis in section 5.1.1 and insightful image pairing sequence of the Red edge Micasense and Parrot Sequoia cameras in the work of Jhan, Rau, Haala, and Caramer (2017).

5.1.2. Feature Matching using default parameters

The selected algorithms were tested for feature matching; the aim was to rank the algorithms based on the total number correct matches and inliers it retains. This is important for image registration since a sufficient number of inliers are needed for accurate estimation of the relative orientation between paired bands. The best performing algorithms would then be selected for further co-registration tests; and also, eventually answering research question three of objective three on the best registration method for our research problem. Using the red edge of the first epoch as the master band and others as slaves, inter-epoch feature matching was done to evaluate the algorithms. Figure 5.4 shows the analysis of the matching results obtained for each of the four algorithms.

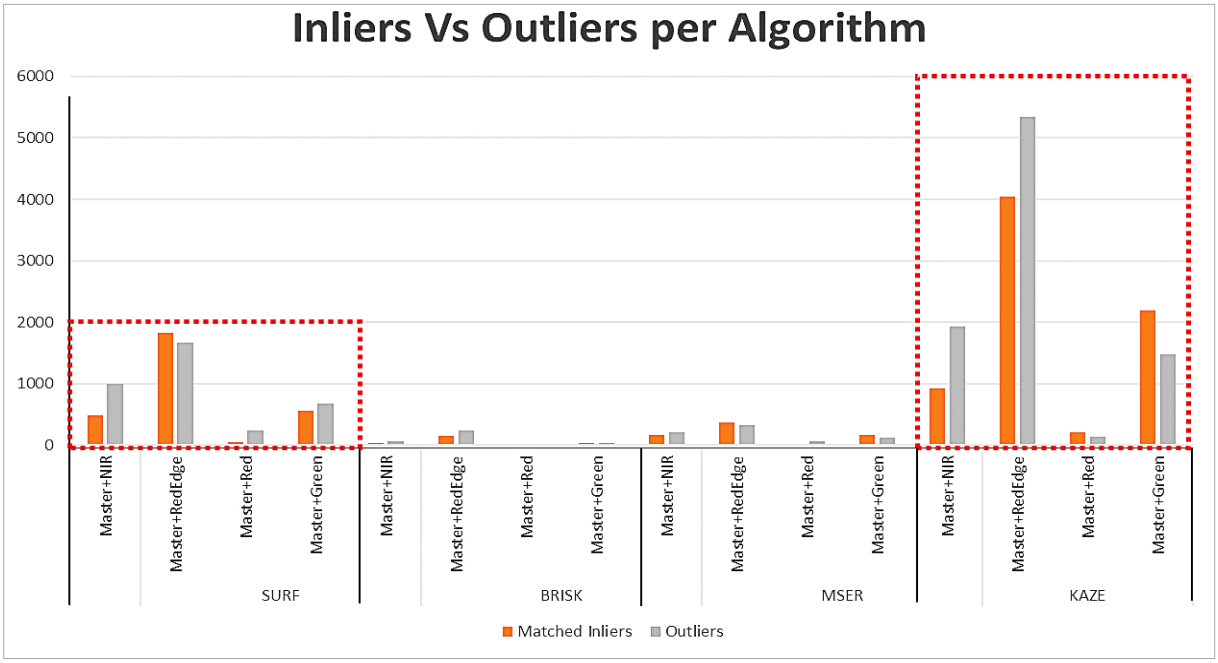
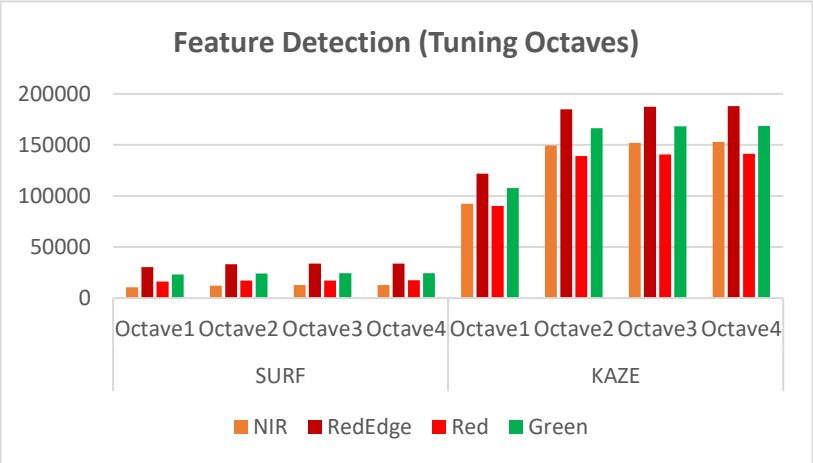


Figure 5.4. Illustrating inliers and outliers of matched Keypoints as obtained per algorithm.

The red edge band combination (i.e. 'Master + Red edge') between two epochs recorded the most number of inliers; this was somewhat expected since they are spectrally similar. The master band combination with green and NIR followed respectively. On matters algorithm, KAZE and SURF algorithms succeeded in finding correct matches within 1000 iterations for all the band combinations (i.e. Master + NIR; Master + Red edge; Master + Red; and Master + Green). In Figure 5.3, KAZE is seen to have more outliers than inliers in the first and second combinations. Nonetheless, the overall performance of KAZE depicts more inliers than the other algorithms in all the band combinations. BRISK and MSER did not converge to find sufficient points after 1000 iterations for the master and red band combination. For this reason, the KAZE and SURF were selected to be the winning algorithms for subsequent tests in this study.

5.1.3. Parameterization for feature detection

The algorithms were tuned and tested under different parameter values. The results presented show the impact of tuning the number of octaves and scale levels. The analysis was done per band. See Figures 5.5 and 5.6 for illustration.



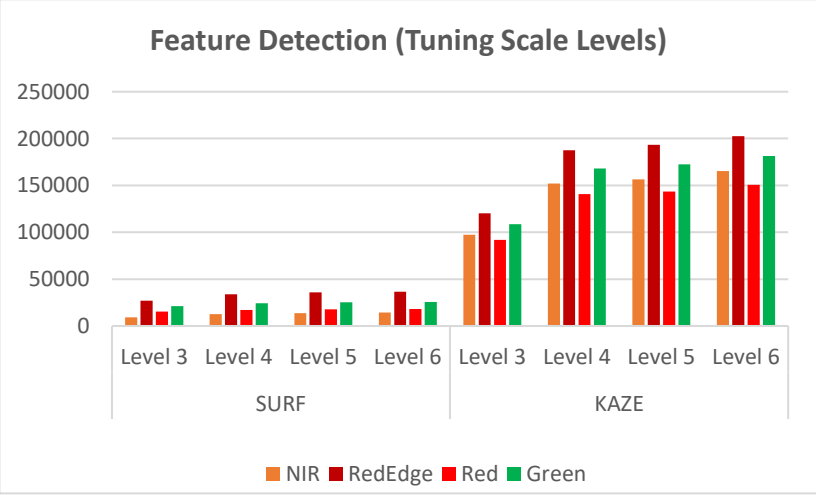
Time (Seconds)		
Octaves	**SURF**	**KAZE**
1	2.15	70.86
2	2.60	147.91
3	2.80	164.09
4	3.02	177.73

Table 6. Time taken per octave for feature detection

Figure 5.5. Showing the impact of tuning number of octaves.

Octaves are a set of images in a Gaussian scale space with the upper images in the layering having been scale down. The pixels of local extrema are detected as keypoints. The results displayed in Figure 5.5 show that the higher the octave the more features are detected cumulatively. However, the difference in the number of features detected per octave shows a decreasing trend in the number of features detected in higher octaves. The KAZE algorithm presents a sharp cumulative increase in points detected from the first to the second octave, but somewhat plateaus by the third octave. The SURF is no match to KAZE in feature detection but is overwhelmingly fast as shown in table 8. SURF recorded a time difference per octave of less than one second. Conversely, KAZE doubles the time between the first to the second octave, and takes 174.71 more seconds than SURF to complete detection in the fourth octave.

Holding other parameters constant, the scale level was tuned from level three to six, and the results of experiments done using sequoia orthophoto are presented in Figure 5.6.



Time (Seconds)		
Levels	**SURF**	**KAZE**
3	2.34	82.23
4	2.94	145.15
5	3.41	175.44
6	3.61	210.62

Table 7. Time per scale level for feature detection

Figure 5.6. Showing the impact of tuning scale levels.

Figure 5.6 shows that the higher the scale levels the more features are detected cumulatively. However, a difference of the number of features per scale level show that most features are detected in level three (3). KAZE exhibits somewhat a steady increase from the second scale level. As seen in tuning the number of octaves, a sharp increase in the number of points detected is seen between level 3 and 4 of KAZE. For both SURF and KAZE a stable condition of feature detection is reached in scale level five and second octave

because the number of features detected per band beyond these points are less than five percent of the total number of features detected. On the other hand, an increase in scale levels increases the computational time; SURF was still overwhelmingly faster than KAZE with KAZE almost doubling the time between level 3 and 4, and taking 207.01 seconds more in level 6. See table 9.

The number of features detected by the KAZE algorithm using the default parameters demanded more computational power to index and match these features. It was therefore necessary to tune the KAZE algorithm to perform at per with the other three algorithms. The *Hessian Threshold* was adjusted from 0.0001 to 0.0012; an increase in this value excludes less significant local extrema thus decreasing the number of Keypoints detected. See Figure 5.7. The results indicate a comparable performance with the contending algorithms, with the red edge yielding over 30,000 points for SURF and KAZE.

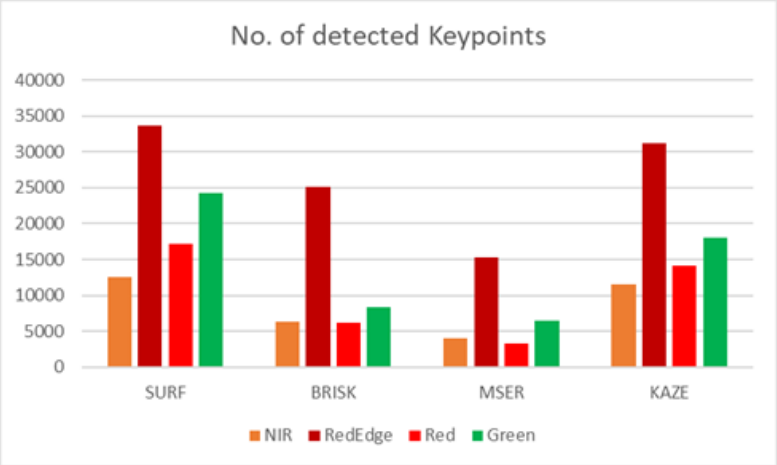


Figure 5.7. Feature Detection with KAZE tuned threshold

5.1.4. Parameterization for feature description and matching

The descriptor size was set to either 64 or 128. A larger feature size of 128 provided a greater accuracy, but decreased the number of matched features thus insufficient matches to estimate the transformation matrix. The results on the effect of tuning the descriptor size are presented in table 10.

Descriptor Size	**Band Pairs**	**SURF**			
		Inliers	Outliers	Inliers	Outliers
64'	Master+NIR	190	1135	431	2420
	Master+RedEdge	661	2594	1822	7568
	Master+Red	25	250	122	234
	Master+Green	187	913	943	2735
128'	Master+NIR	22	80	107	401
	Master+RedEdge	75	264	946	2462
	Master+Red	7	5	13	10
	Master+Green	15	51	264	649

Table 8. Showing effect of descriptor size to number of correct matches

The results presented in table 10 indicate a drastic decrease in the number of matched points when using a descriptor size of 128. It was observed that the number of inliers using a descriptor size of 128 were not sufficient for MSAC to fit the best model to estimate the transformation matrix. Thus, misregistration errors were still evident after registration. On the other hand, with a descriptor size of 64, the number of matched features increased, consequently the inliers as well. The descriptor size of 64 was therefore selected since the number of inliers were sufficient for estimating the transformation matrix. Figures 5.8 (b) and (c) show significant misregistration even after co-registration using a descriptor size of 128, while (d) shows eliminated misregistration while using inliers with a descriptor size of 64.

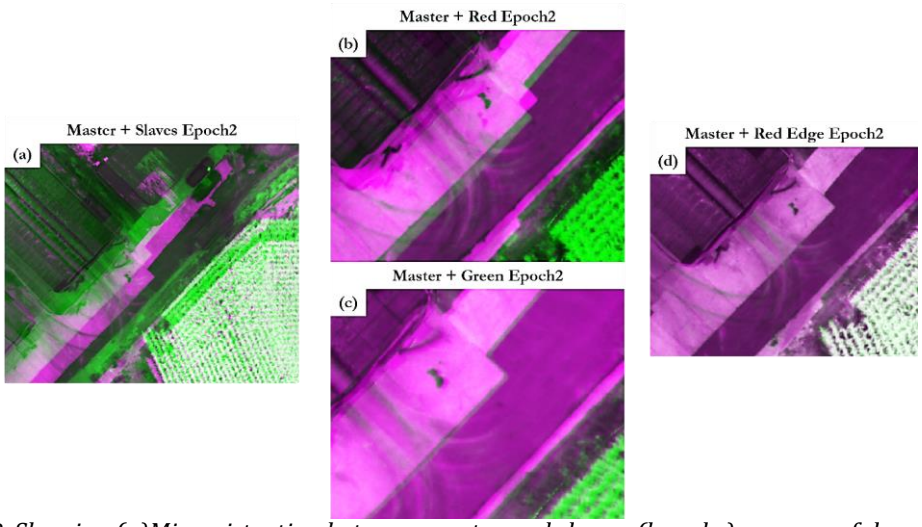


Figure 5.8. Showing (a)Misregistration between master and slaves; (b and c) unsuccessful registration using matched inliers of descriptor size 128; successful registration (d)

For feature matching parameterization, tuning the matching function was also done. The *exhaustive* brute force matching method was used. Two approaches were tested; the one-to-many matching (*unique* = false) resulted in more matches than one-to-one matching (*unique* = true). See table 11. When the *Unique* argument is set as true, the function performs a forward-backward match to select only unique matches between features 'A' and 'B'. Conversely, false returns all matches between features 'A' and 'B'. Since the one-to-one matching selects only unique matches between features 'A' and 'B', it was adopted for this research.

Match Criteria	Index Pairs	SURF	KAZE
	RedEdge+NIR	5069	7386
Unique = True	RedEdge+Red	568	726
	RedEdge+Green	4209	6162
	RedEdge+NIR	5460	7674
Unique = False	RedEdge+Red	669	759
	RedEdge+Green	4346	6370

Table 9. Number of matches per matching method

5.1.5. Outlier removal and Estimation of the geo-transformation matrix

Correct matches between image pairs were obtained by randomly sampling all the putative matches iteratively until the best sample to estimate the rotation and translation between the image pairs was found (i.e. RANSAC). The best sample was considered the set of inlying point pairs. In Figure 5.9, the effect of varying the projection distance threshold is presented.

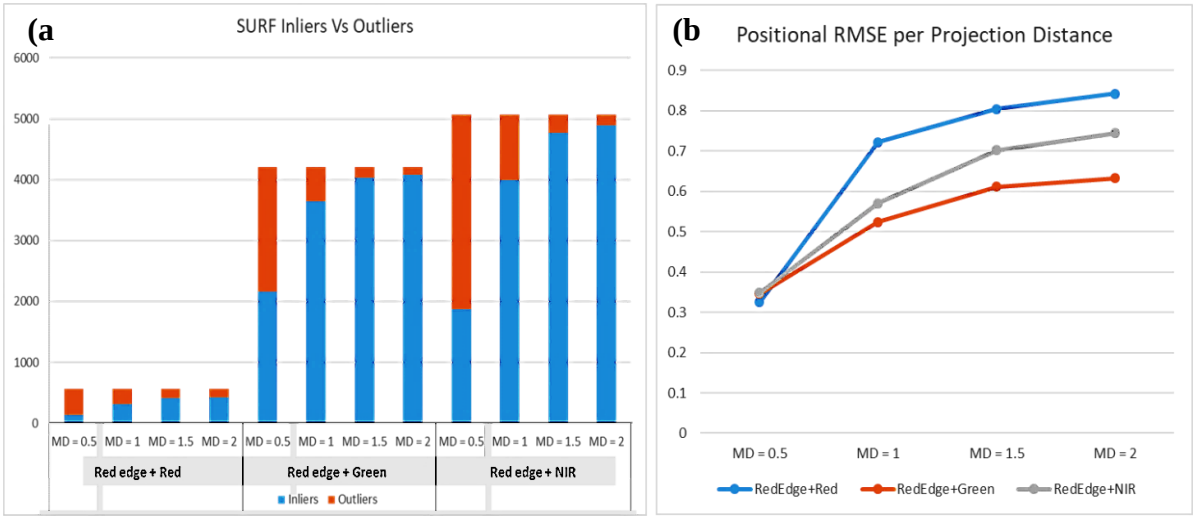


Figure 5.9. (a) Inliers vs outliers per distance threshold (b) Positional RMSE of Keypoints per distance threshold

It was observed that an increase in the projection distance '*MaxDistance',* led to an increase in the number of inliers as well as the residual error. As illustrated in Figure 5.9(b), the positional RMSE of the inliers increased with an increase in the projection distance. For this study, the threshold set for misregistration after geo-transformation was 0.5 of a pixel. Thus, RMSE's greater than 0.5 were unacceptable. Thus, a projection distance of 0.5 was adopted.

5.2. The Adopted Algorithm Parameters

A series of experiments for which results are presented in section 5.1 informed the selection of parameters presented in table 12 in the final registration algorithms. It important to note that these are the parameters for co-registering the whole orthophoto and not single images. For the single images, default parameters proved to be satisfactory.

Algorithm	Parameters	Intra-epoch	Inter-epoch
SURF	MetricThreshold	2500	2700
	NumScaleLevels	4	6
	NumOctaves	3	3
KAZE	Threshold	0.0012	0.0012
	NumScaleLevels	4	6
	NumOctaves	3	3

Table 10. Adopted feature extraction parameters

In addition, a descriptor size of 64 was selected; the *exhaustive* brute force matching method was used with the one-to-many matching strategy; and a maximum projection distance of 0.25 and 0.5 were selected for intra-epoch and inter-epoch respectively.

5.3. Intra-epoch band-to-band registration

5.3.1. Parrot Sequoia

Intra-epoch misregistration analysis was done at the image and orthophoto level. Figure 5.10 illustrates the displacement per band combination using the Red edge as the master band.

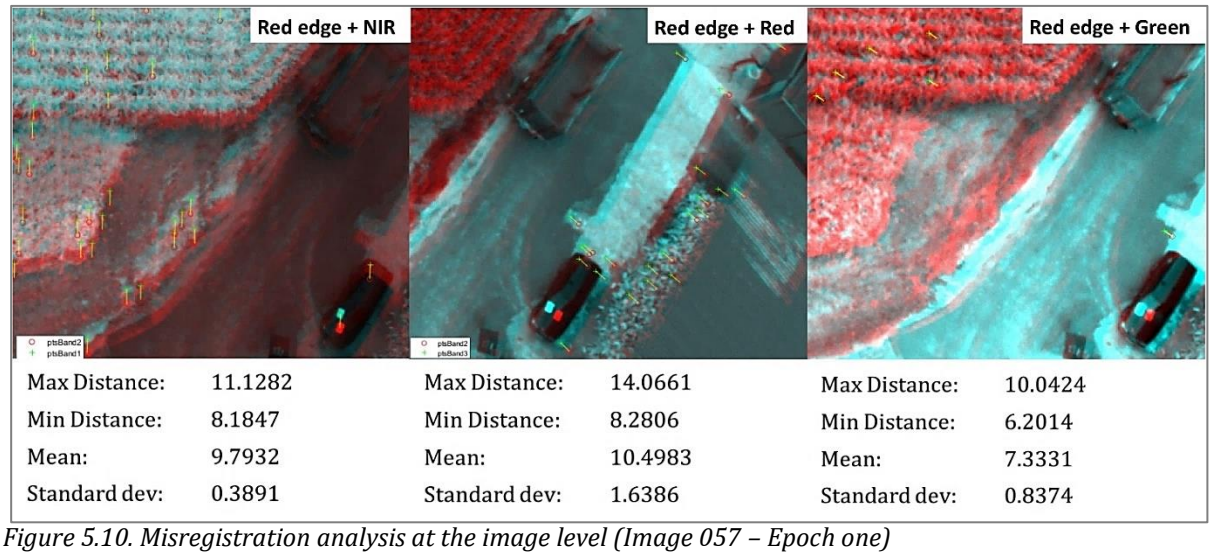
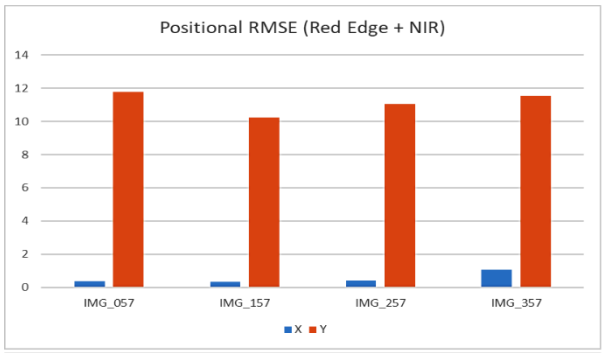
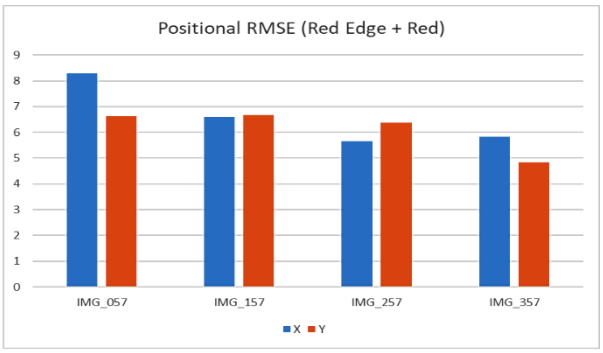
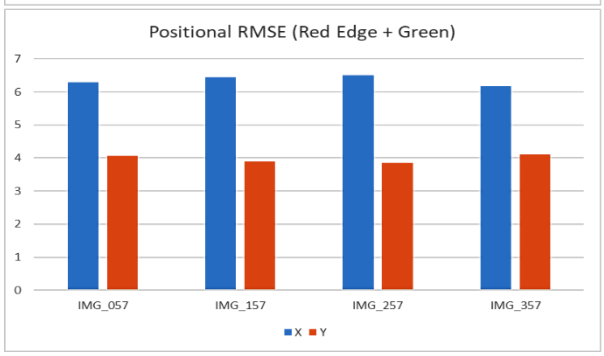


Figure 5.10. Misregistration analysis at the image level (Image 057 – Epoch one)

Keypoint pairs in Figure 5.10 clearly show a 90 degrees systematic displacement between the red edge and the NIR and somewhat a 145 and 120 degrees displacement between the red edge and the red and green bands respectively. The intra-epoch analysis revealed a systematic displacement attributed to the basis distance of the cameras. The displacement was observed to be aligned with the camera position in the sensor. The distances between the point pairs vary from band combination to another. The test was done in four different camera positions within the block in epoch one and for selected camera positions in the subsequent epochs. The results are presented in Figure 5.11.





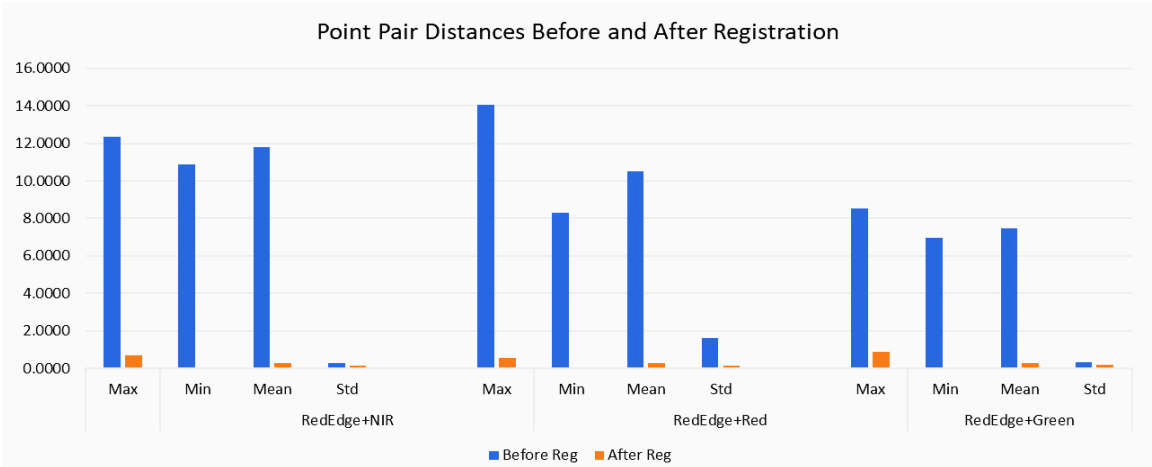


Camera Position	Master+NIR		Master+Red		Master+Green	
	X	Y	X	Y	X	Y
IMG_057	0.37	11.78	8.29	6.64	6.29	4.07
IMG_157	0.34	10.22	6.60	6.68	6.45	3.89
IMG_257	0.43	11.04	5.65	6.38	6.50	3.86
IMG_357	1.07	11.52	5.84	4.84	6.18	4.10

Figure 5.11. Analysis of systematic band displacement within the same epoch

The red edge and NIR combination exhibited a significant displacement of about 11 pixels in the "Y" direction. The red edge and red showed a shift of about 6 pixels in both directions apart from camera position '057' that above 8 pixels in the "X" direction. The last combination, red edge and green, unveiled a uniform displacement of about 6 pixels in the "X" direction. In addition, it was also observed that an increase in the flying height reduced the displacement effect of the basis distance.

The co-registration results at the image level within the same epoch (epoch 1) are presented in Figure 5.12. The corresponding accuracies (in pixels) after registration are tabulated as 'After Reg', which summarizes the maximum (Max), minimum (Min), Mean and Standard deviation (Std) distance values between paired keypoints of the three band combinations used.



	Before Reg	After Reg
Max	12.3714	0.7011
Min	10.8925	0.0111
Mean	11.7801	0.2897
Std	0.2803	0.1515

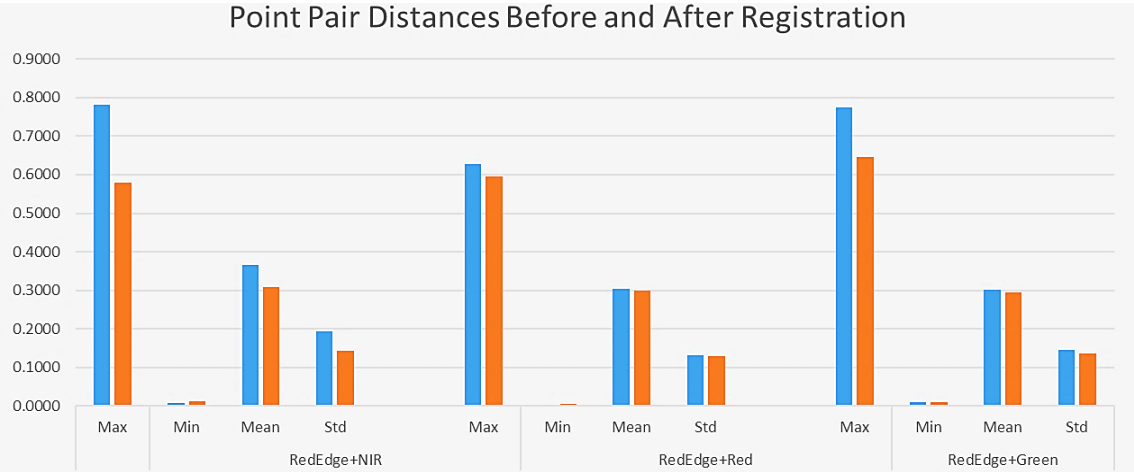
	Before Reg	After Reg
Max	14.0661	0.5729
Min	8.2806	0.0551
Mean	10.4983	0.2690
Std	1.6386	0.1408

	Before Reg	After Reg
Max	8.5456	0.8954
Min	6.9488	0.0330
Mean	7.4799	0.2854
Std	0.3310	0.1658

Figure 5.12. Image level misregistration error before and after co-registration (image. No. 057-Epoch 1)

Figure 5.12 clearly illustrates that the image level band displacement is eliminated upon registration, and that subpixel accuracy was achieved. Minimum distances between paired points close to zero were achieved and the maximum distances were less than a pixel.

On the other hand, the intra-epoch misregistration at the orthophoto level is quite minimal. This can be attributed to corrections during aerial triangulation and orthorectification. As demonstrated in Figure 5.13, the co-registering bands within the same epoch still reduces misregistration that exists, even though with a small margin.



	Before Reg	After Reg
Max	0.7821	0.5798
Min	0.0087	0.0116
Mean	0.3658	0.3095
Std	0.1943	0.1431

	Before Reg	After Reg
Max	0.6283	0.5944
Min	0.0001	0.0049
Mean	0.3034	0.2997
Std	0.1309	0.1287

	Before Reg	After Reg
Max	0.7746	0.6457
Min	0.0105	0.0093
Mean	0.3006	0.2945
Std	0.1457	0.1357

Figure 5.13. Orthophoto level band misregistration error before and after co-registration (Epoch 1)

The mean and standard deviations of the point pair distances show that indeed a reduction is evident after co-registration. This reduction trend is also seen with the Maximum value between inlying point pairs. These registration effects may not be appreciable qualitatively (i.e. visual inspection of band overlays), but are statistically significant; as the value tends to zero the more accurate the registration is.

5.3.2. Micro MCA Tetracam

The intra-epoch co-registration of the Micro MCA Tetracam was done to evaluate the accuracy of the registration process at the image level. The Micro MCA files are in '.RAW' format, in order to process the files, Tetracam's PixelWrench2 software was used to convert the '.RAW' file formats into multipage TIFFs. It was observed that a multipage TIFF was a stack of six (6) channels per camera station during acquisition. This is unlike Sequoia which has all the channels processed separately. The Micro MCA orthophoto was then tested for misregistration. The results are presented in table 13 below.

Algorithm	Bands	Max	Min	Mean	Std	RMSE_X	RMSE_Y	H_RMSE
SURF	CH6 + CH1	0.48	0.00	0.09	0.09	0.10	0.07	0.13
	CH6 + CH2	0.49	0.00	0.10	0.10	0.11	0.08	0.14
	CH6 + CH3	0.51	0.00	0.11	0.11	0.13	0.09	0.15
	CH6 + CH4	0.51	0.00	0.11	0.11	0.13	0.09	0.16
	CH6 + CH5	0.51	0.00	0.13	0.12	0.14	0.10	0.17

Table 11. Misregistration error analysis for Micro MCA Tetracam (unit=pixels)

From table 13, it is apparent that all the six (6) channels of the Micro MCA Tetracam were aligned to the master channel in the range of 0.13 – 0.17 of a pixel. In comparison to the Sequoia, the Micro MCA Tetracam records lower maximum values between the master and slaves. The mean error and the standard deviation is also comparatively low in Micro MCA Tetracam than in Parrot Sequoia.

5.4. Inter-epoch band-to-band registration

Two approaches of co-registration were experimented upon; that is, one master band from epoch one used to align all subsequent acquisitions, termed as 'many-to-one registration', and each band from epoch one used as a master for all corresponding bands in subsequent acquisitions, termed as 'one-to-one registration'.

5.4.1. Many-to-One Registration

The many-to-one band alignment was done with the red edge as the master band. Subpixel accuracies were obtained using a projection distance of 0.7 and 0.5 of a pixel as the projection distance threshold of SURF and KAZE respectively. SURF could not find sufficient point pairs to estimate geometric transformation between orthophotos of epoch 1 and 2 at a projection distance threshold of 0.5. On the other hand, KAZE could find sufficient points, therefore even aligning the bands more accurately. However, the horizontal positional RMSE obtained with SURF and KAZE after registration presented in Figure 5.14. Misregistration errors in the range of 0.02 – 1.16, 0.07 – 0.91, and 0.04 – 0.98 pixels for SURF, against KAZE's 0.02 – 1.32, 0.03 – 1.37, and 0.02 – 1.31 pixels for the same band combinations were evident.

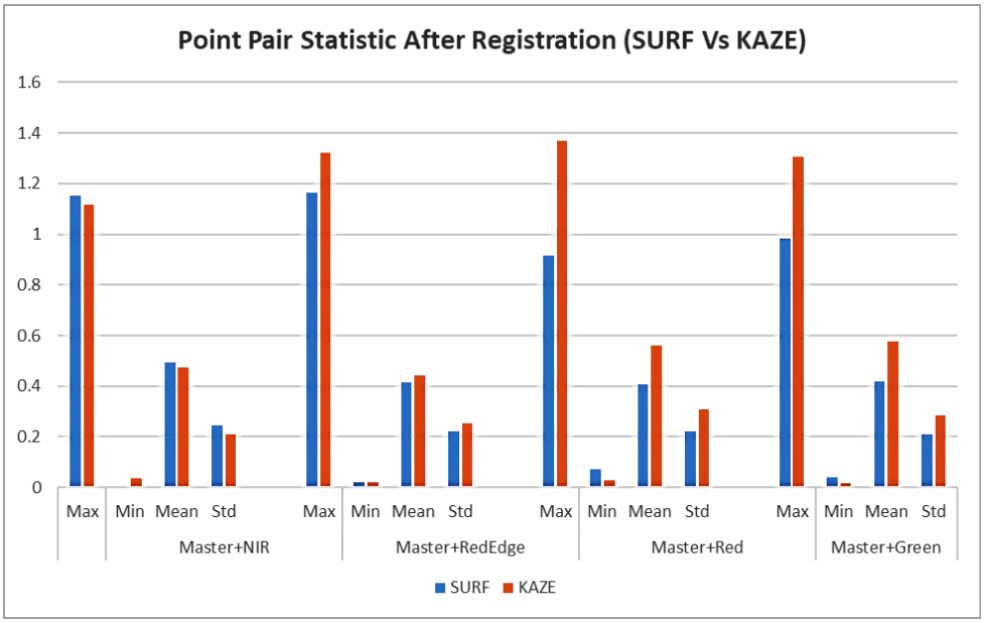


Figure 5.14. The point pair statistic of many-to-one band registration at MaxDistance 0.7

Following the ability of KAZE to find sufficient number of point pairs and consequently estimate the geometric transformation matrix at *MaxDistances* of 0.7 and 0.5, a comparison of the obtained results was done. Table 14 summarizes the descriptive statistics of different spectral band pairs.

Algorithm	Bands	Max	Min	Mean	Std
KAZE 0.7	Master + NIR	1.12	0.04	0.47	0.21
	Master + Red edge	1.32	0.02	0.44	0.25
	Master + Red	1.37	0.03	0.56	0.31
	Master + Green	1.31	0.02	0.57	0.28
KAZE 0.5	Master + NIR	0.81	0.05	0.39	0.19
	Master + Red edge	0.76	0.01	0.33	0.17
	Master + Red	1.48	0.02	0.50	0.34
	Master + Green	0.82	0.01	0.37	0.19

Table 12. KAZE's point pair statistics of many-to-one band registration (unit-pixels)

From the statistics presented in table 14, it is observed that the distance between point pairs reduces further with *MaxDistance* set to 0.5. However, the 'Master + Red' band combination performs equally the same in both cases. As observed in earlier results, this band combination comparatively records the worst registration results.

5.4.2. One-to-One Registration

All spectral channels of epoch one were used as master bands, the corresponding subsequent bands were slaves (see illustration in Figure 4.10, page 29). The displacement between corresponding bands was seen to be systematic across all the bands. Figure 5.15 illustrates the systematic displacement between orthophotos derived from epoch one and two per band. The yellow lines connect the matched Keypoints thus is the distance between paired points.

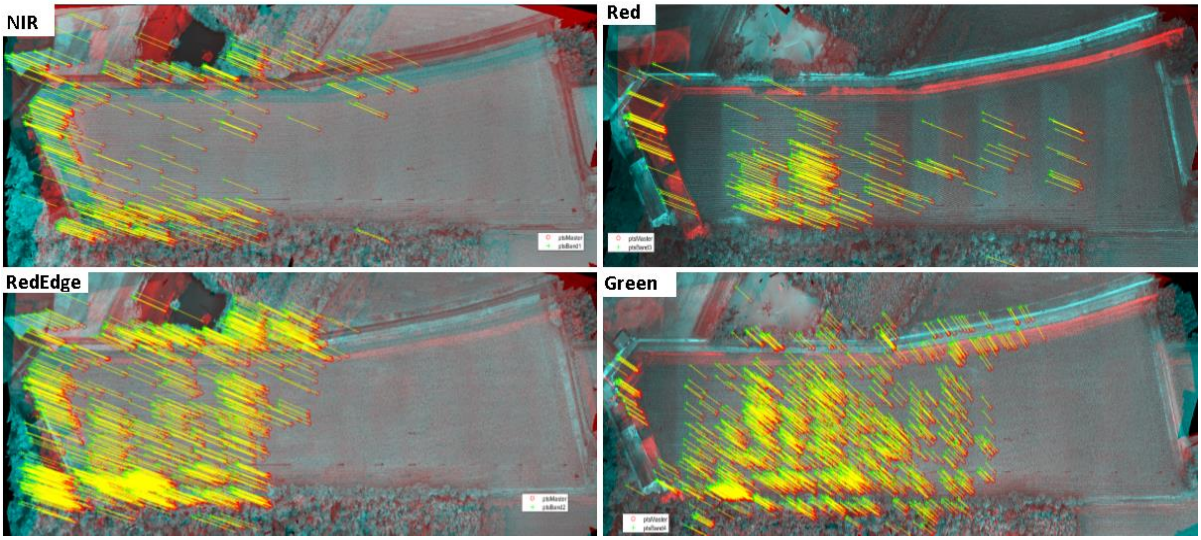


Figure 5.15. Systematic displacement between epoch one and two (left); and eliminated displacement (right)

It was observed that the red band combination recorded the lowest number of inliers whereas green had the highest (See Figure 5.15 & 5.16). Nonetheless, the inliers allowed for accurate estimation of the transformation matrix, and thus subpixel accuracies were obtained. From the statistics presented in boxplots presented in Figure 5.17, it is observed that the mean registration error is in the range of 0.26 – 0.38. See appendix 6 for more descriptive statistics. This is quite acceptable since it is subpixel accuracy and is less than the set threshold of 0.5 of a pixel. Surprisingly, the red band combination is seen to be the best for inter-epoch co-registration since they record the least mean error. A snippet of the one-to-one co-registration results is presented in figure 5.18.

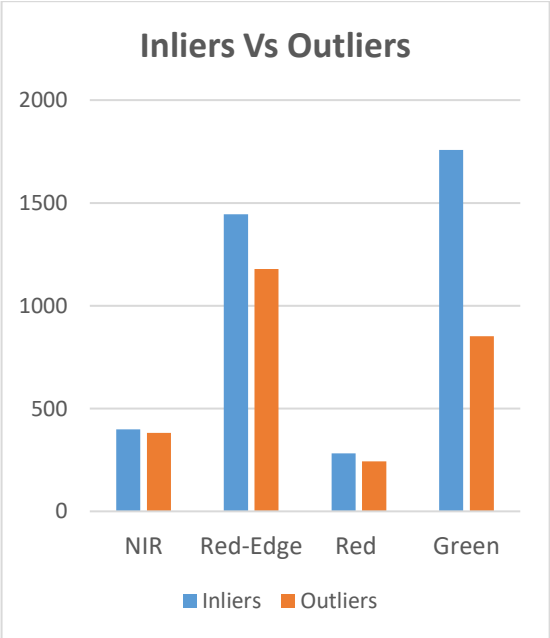


Figure 5.16. Inliers vs outliers for one-to-one approach

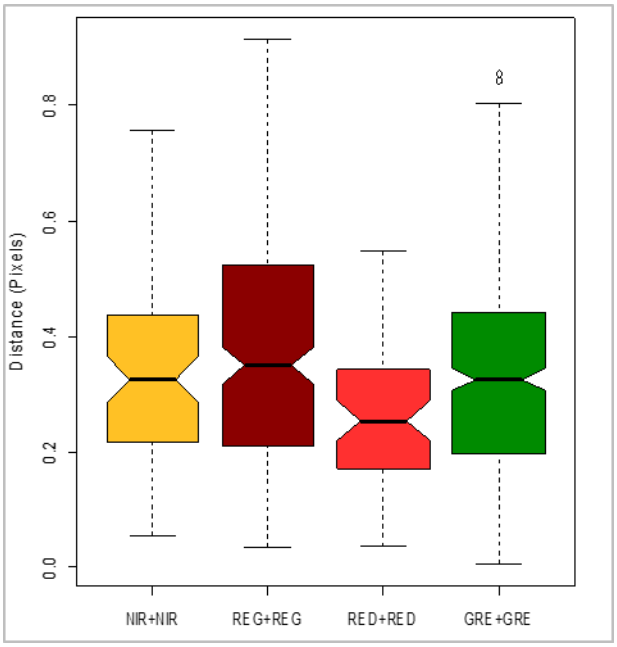


Figure 5.17. Boxplots of paired distances after registration

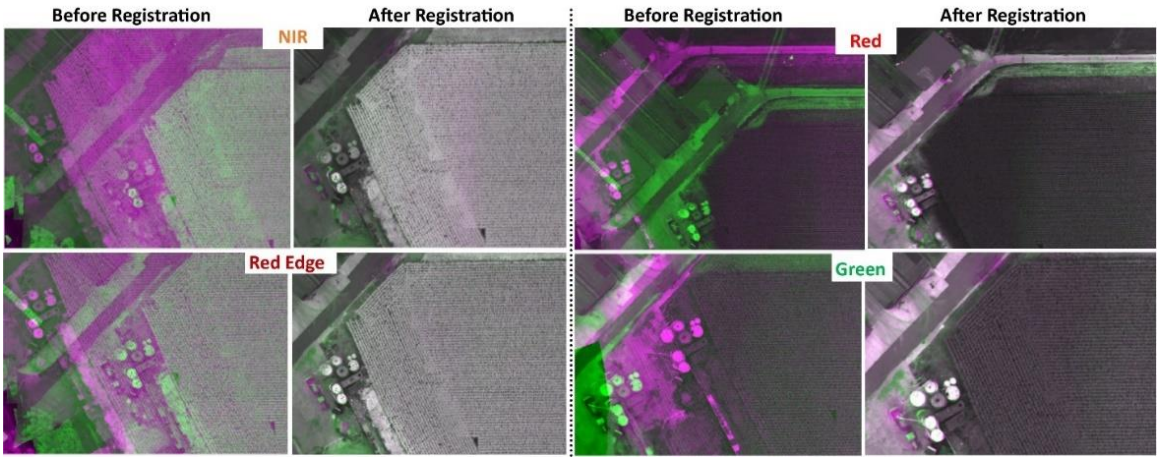


Figure 5.18. Showing inter-band alignment before and after registration.

5.5. Co-registration Accuracy Assessment

The horizontal positional RMSE was used to assess the accuracy of the co-registration. Keypoints pairwise distances were computed before and after co-registration. The keypoints detected on the first epoch orthophoto were regarded as the reference points. After co-registration, new keypoints were detected on the registered slave epoch, and matched to the master/reference keypoints. The positional RMSE was then computed between the point pair coordinates of the master and the slave keypoints as shown in section 4.3.8, page 31. Figure 5.15 in page 43 shows inlying point pairs before and after registration while tables 15 – 21 show the corresponding horizontal positional RMSE.

5.5.1. Intra-epoch accuracy assessment

Accuracy assessment was done at both the image and the orthophoto level. According to the keypoints pairwise distances tabulated in table 15 and 16, misregistration was more pronounced at the image level than the orthophoto level. Individual images suffered a systematic displacement as a function of the basis distance between the cameras within the sensor (refer to Figures 5.10 and 5.11). In contrast, misregistration between the orthophoto bands was at a subpixel level as shown in table 16.

Image Set	Red edge + NIR		Red edge + Red		Red edge + Green	
	Before	After	Before	After	Before	After
IMG_057	11.78	0.33	10.62	0.30	7.49	0.33
IMG_157	10.23	0.29	9.39	0.35	7.53	0.30
IMG_257	11.05	0.31	8.53	0.40	7.56	0.30
IMG_357	11.57	0.30	7.59	0.37	7.42	0.46

Table 13. Horizontal positional RMSE of select individual images, before and after co-registration

The 'Red edge + NIR', 'Red edge + Red' and 'Red edge + Green' band combinations were observed to record an average horizontal positional displacement of 11.16, 9.0 and 7.5 pixels respectively before registration. After registration, they recorded subpixel average horizontal displacement of 0.31, 0.36, and 0.35 respectively. Table 16 - 18 presents the co-registration accuracies of the orthophotos per band combination within all the epochs using SURF.

	Red edge + NIR		Red edge + Red		Red edge + Green	
	Before	After	Before	After	Before	After
Epoch1	0.17	0.17	0.22	0.22	0.18	0.18
Epoch2	0.18	0.18	0.20	0.16	0.21	0.18
Epoch3	0.18	0.18	0.21	0.19	0.20	0.20

Table 14. Horizontal positional RMSE of bands within an epoch using SURF

According to the results, the bands are aligned to a subpixel accuracy. The positional RMSE of the inliers is the equal for epoch one and three, and decimal differences in epoch two, before and after registration. Since the registration procedure was intensity-based, the slight differences in epoch 2 could be attributed to illumination differences and the randomness during outlier removal.

5.5.2. Inter-epoch accuracy assessment

Inter-epoch registration was done at the orthophoto level. The accuracy results presented in this section are those of aligning epochs two and three to epoch one of Parrot Sequoia imagery (i.e. master epoch). Also presented herein are the accuracies obtained from both many-to-one, and one-to-one band registration.

Band Combination	SURF	KAZE
Master + NIR2	0.57	0.52
Master + Red edge2	0.49	0.51
Master + Red2	0.48	0.64
Master + Green2	0.47	0.64

Table 15. Positional RMSE of many-to-one registration between Sequoia epoch 1 and 2

Band Combination	SURF	KAZE
NIR1 + NIR2	0.36	0.34
Red edge1+Red edge2	0.39	0.36
Red1 + Red2	0.28	0.30
Green1 + Green2	0.32	0.32

Table 16. Positional RMSE of one-to-one registration between Sequoia 1 one and 2

Band Combination	SURF	KAZE
Master + NIR3	0.54	0.50
Master + Red edge3	0.46	0.40
Master + Red3	0.47	0.66
Master + Green3	0.44	0.64

Table 18. Positional RMSE of many-to-one registration between Sequoia epoch 1 and 3

Band Combination	SURF	KAZE
NIR1 + NIR3	0.39	0.37
Red edge1+Red edge3	0.33	0.31
Red1 + Red3	0.31	0.31
Green1 + Green3	0.34	0.33

Table 17. Positional RMSE of one-to-one registration between Sequoia epoch 1 and 3

As demonstrated in tables 17 to 20, both SURF and KAZE obtained subpixel registration accuracies. SURF however recorded lower RMSE values than KAZE thus more accurate. On the other hand, the one-to-one registration approach yields better results than many-to-one approach; the many-to-one approach recorded an average RMSE of 0.5 against the one-to-one approach of 0.36 across all the band combinations. The similarity in spectral properties per band combination in the one-to-one approach is one of the possible reasons as to why band pairs are better aligned.

Due to poor image quality and the poorly distributed conjugate points (see Appendix 4 for detailed diagrammatic illustration) between the Micro MCA Tetracam and the master image, the obtained inter-epoch co-registration accuracy was comparatively low. The number of correct matches found between Sequoia's registered epoch 3 and Micro MCA's epoch 3 orthophotos at a distance threshold of 0.5 – 1 pixels was insufficient to estimate geometric transformation, thus a threshold of 2 was used. The co-registration accuracy obtained is presented in table 21.

Band Combination	SURF	KAZE
Master + CH1	2.11	1.51
Master + CH2	1.26	1.18
Master + CH3	1.75	2.08
Master + CH4	1.21	1.37
Master + CH5	1.25	1.29
Master + CH6	1.21	1.25

Table 19. Positional RMSE for Sequoia and Micro MCA Epoch 3

The co-registration error was in the range of 1.21 – 2.11 pixels for both SURF and KAZE. However, this can greatly improve to a subpixel level if image with a better quality and is properly scaled is used.

5.6. Analysis of supplementary test

5.6.1. Relative Orientation Analysis

Pix4D executes aerial triangulation and bundle block adjustment at its first stage. At this stage, camera parameters are optimized and a text file containing the estimated camera parameters per station is generated. The file contains the coordinates of the principal point of each camera, and the rotation angles. To assess the difference in relative orientation of the cameras at each station, the deviation of the NIR, red and green camera parameters from the red edge camera was assessed. Figure 5.19 presents a boxplot analysis of the same for epochs one.

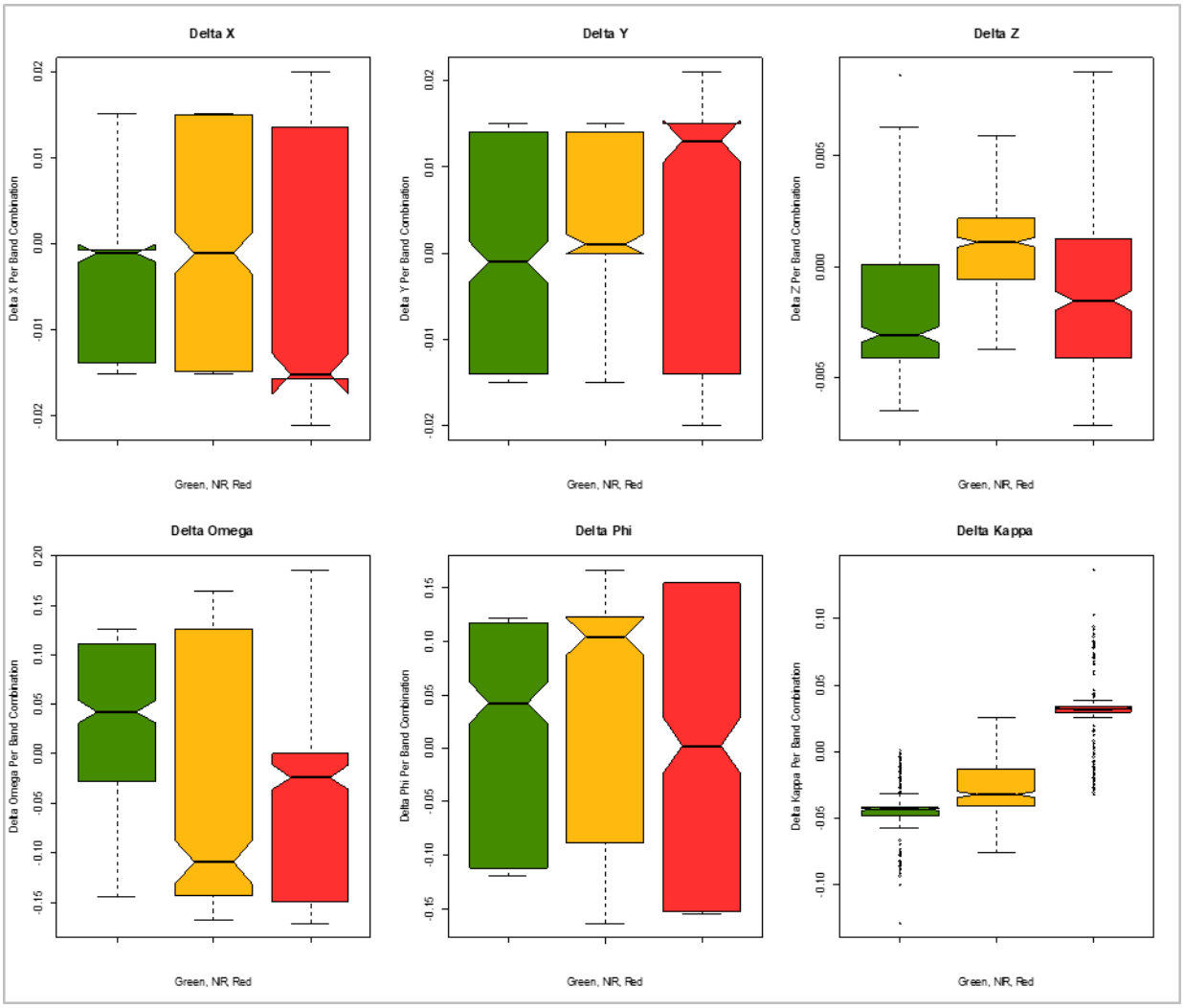


Figure 5.19. Analysis of relative orientation of each camera per station

Delta X, Delta Y, Delta Z, Delta Omega, Delta Phi and Delta Kappa, refer to the difference in the principal point coordinates and the rotation angles between the master band (i.e. Red edge) and the slaves (i.e. green, NIR, red respectively).

It is apparent that the camera parameters are different for each image band. The position of the projection centre (in world coordinate system) changes at each camera position. However, the relative orientation of the cameras to each other remain the same. The differences shown in Figure 5.19 depict a mean error ranging from 0 – 0.11 for all the parameters. For this reason, the relative orientation of the cameras in one station can be propagated to all camera stations. Notably, none of the observed ROP differences between cameras were greater than a pixel. See the appendix 5 for differences in the spatial and rotational components of the relative orientation parameters between the red edge master band and the slave bands.

5.7. DSM Co-registration and Analysis

DSM analysis was done within and between epochs; intra-epoch DSM analysis was done with the aim of exposing the variation in DSMs extracted per band, and consequently establish the best band for DSM extraction. In so doing, research question three of objective two was answered (see results in section 5.7.1). On the other hand, inter-epoch DSM assessment focused on accurate co-registration of multitemporal DSMs; the aim was to align DSMs from all epochs to epoch one so as to monitor the crop health over time. In this study, this was really not possible since the image acquisitions were done at a late stage of the crop's growth. However, the DSMs were still used to assess crop damage after a heavy storm.

5.7.1. Intra-epoch DSM Analysis

In an effort to establish the best band for DSM extraction, the disparities among intra-epoch DSMs was assessed. Longitudinal profiles from two sections of the farm were used for the analysis; the scenes included (a) stable and regular structures such as buildings, and (b) a mix of known stable structures and the maize farm. Having used longitudinal profiles, the analysis was pixel based. Figure 5.20 illustrates the longitudinal profile variations of the DSMs.

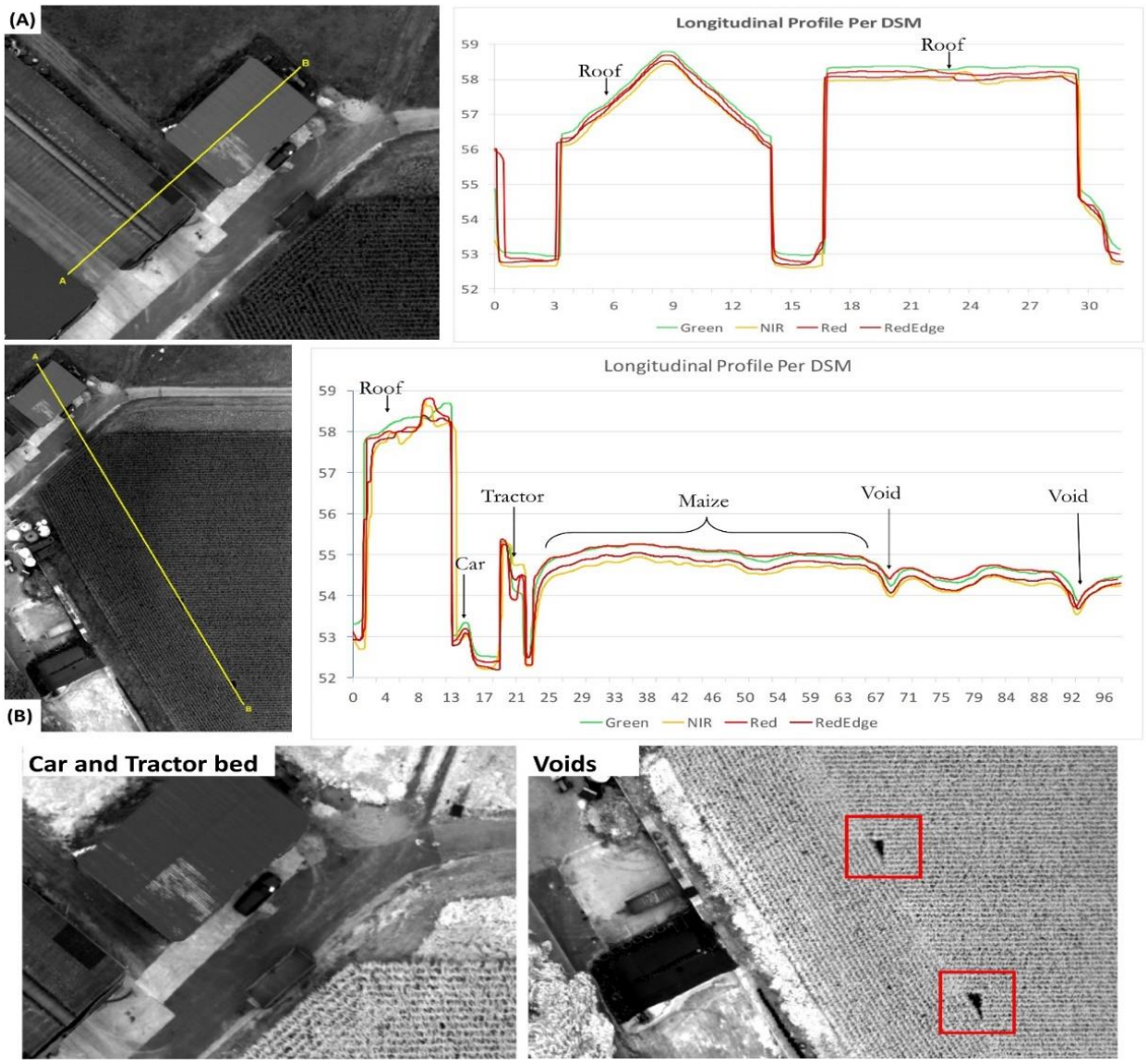


Figure 5.20. Longitudinal profiles per DSM; permanent structures (A), and composite scene (B)

There are significant disparities between intra-epoch DSMs. However, the red band DSM captured the void in the tractor bed more accurately than the others. The green and red DSMs are observed to be almost similar in modelling the maize crop surface. Using the red edge as the master band, the mean deviation of other DSMs was computed and is shown on table 22.

Epoch		Green DSM	NIR DSM	Red DSM
One	Profile A	0.32	0.16	0.20
	Profile B	0.24	0.17	0.25
Two	Profile A	0.46	0.59	0.26
	Profile B	0.20	0.37	0.46
Three	Profile A	0.18	0.10	0.41
	Profile B	0.22	0.12	0.37

Table 20. Intra-epoch DSM mean deviation in meters for all epochs

In addition, it is important to note that the maximum and minimum values of the DSMs varied quite a lot. The maximum values were observed to have a dismal deviation apart from epoch two. However, the deviation of the minimum values were different within and between epochs. Figure 5.21 shows the maximum and minimum DSM values per epoch.

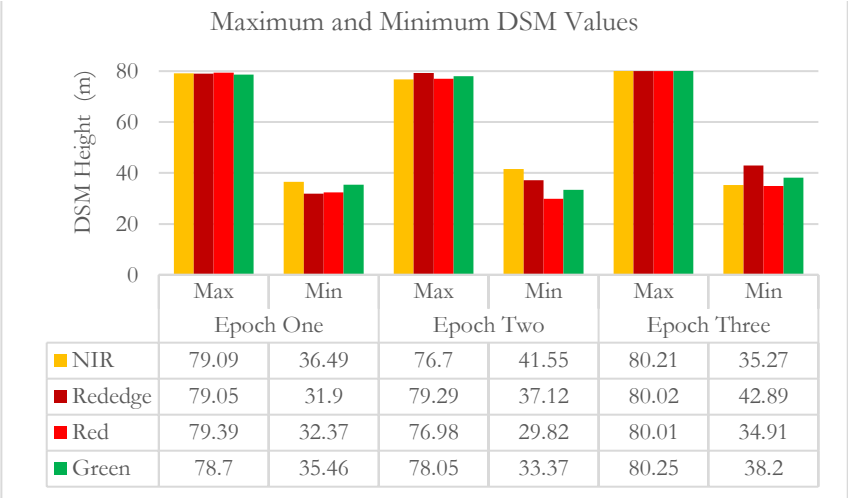


Figure 5.21. Maximum and minimum DSM values per band for all epochs

The disparities seen in epoch two DSM values in Figure 5.21 could be accountable for the wide range of mean error observed in table 21 above. Also, the large variations could be propagated processing errors during image calibration and point cloud densification.

Further analysis to compare the DSMs by correlating spatial means was done. Firstly, grids of 10m by 10m covering the whole farm were created. Secondly, spatial mean of each grid was. Thirdly, the grids' spatial means were correlated. The correlation matrix between DSMs of epoch 1 to 3, rounded off to two decimal places, are presented in table 23.

		Green	NIR	Red	Red edge
Epoch 1	Green	1			
	NIR	1	1		
	Red	1	1	1	
	Red edge	1	1	1	1
Epoch 2	Green	1			
	NIR	0.99	1		
	Red	0.97	0.98	1	
	Red edge	1	0.99	0.96	1
Epoch 3	Green	1			
	NIR	0.96	1		
	Red	0.97	0.97	1	
	Red edge	0.97	0.98	0.97	1

Table 21. Epoch 1 to 3 DSM correlation table of spatial means

The intra-epoch band correlation results show that the DSMs are quite similar. Epoch 1 DSMs recorded a correlation coefficient of 1.00 in all the band combinations. The subsequent DSMs combinations recorded correlation coefficients in the range of 0.96 – 1.00, and 0.96 – 0.98 for epoch two and three respectively. From table 23, it is prudent to conclude that DSMs within an epoch can all be used to model crop heights; but pixel level results show that the red band is the best for DSM extraction (refer to Figure 5.20).

5.7.2. Inter-epoch DSM Analysis

Following the assumption that the relative orientation of the DSMs between epochs is the same as that of their corresponding orthophotos, the geometric transformation matrix estimated between the orthophotos was used to co-register the slave DSMs to the master. Figure 5.22 shows DSM alignment before and after registration.

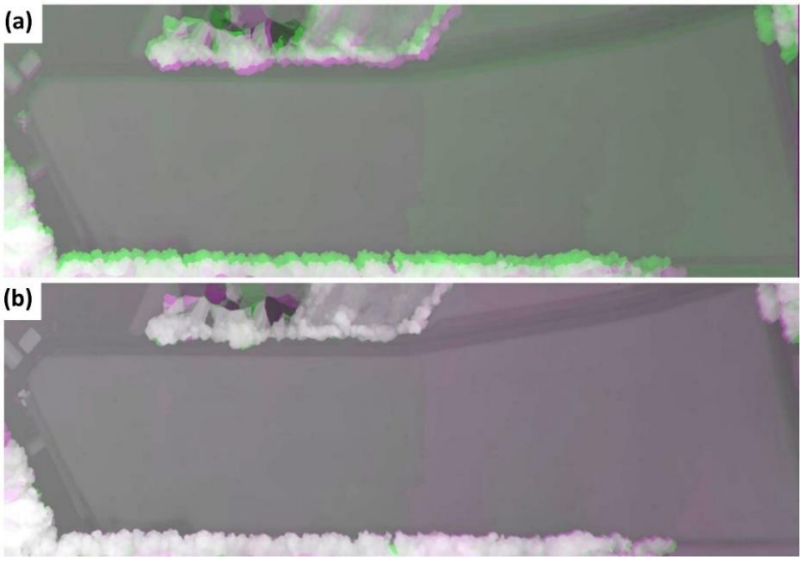


Figure 5.22. (a) Showing misregistration between two DSMs (b) Showing eliminated misregistration

Visual inspection revealed that epoch 2 DSM had been registered. However, misalignments on top canopy remain. Apart from visual inspection, accuracy assessment was done by manual onscreen georeferencing approach. Residual error of conjugate points between the master and the slave were computed. Table 24 presents the residual errors for the conjugate points. The residual error was observed to be between 7-14cm. This could probably be less; manual georeferencing to generate residuals between tie points could have introduced slight errors.

Ref	X1	Y1	X2	Y2	Resid. X	Resid. Y	Resid. XY
1	358283.58	5782025.01	358283.58	5782025.00	-0.035	-0.118	0.12
2	358296.79	5781870.50	358296.76	5781870.43	0.021	0.068	0.07
3	358690.85	5781897.10	358690.85	5781898.56	-0.026	-0.088	0.09
4	358538.78	5782020.64	358538.89	5782020.30	0.04	0.138	0.14

Table 22. Manual Registration Residual errors

5.8. Vegetation Index Analysis

Qualitative comparative analysis was done on NDVI extracted from Sentinel 2B imagery dated 16th August, 2017, and Parrot Sequoia imagery acquired on 8th and 11th August respectively. NDVI was computed using the image analysis tool in ArcGIS; symbology and visualization was also done in ArcGIS. The NDVI calculation was done by applying the formulae shown in section 4.5.1 to the registered bands only. The VI maps were correctly aligned to each other since they were generated from registered bands. The NDVI maps for Sentinel 2B and Parrot Sequoia are shown on Figure 5.23. The zones with healthy vegetation are vividly seen in the field as dark green, while unhealthy vegetation is between light green and cream yellow. Beige and deep red reveal the non-vegetated areas like water pond, roads, bare soil, and building roofs. The lowest values of -0.5 and -0.6 for Sequoia acquisitions are values for the water pond north of the farm. Values in the range of 0.2 – 0.9 correspond to vegetation; the lower values between 0.2 - 0.3 represent shrubs and grassland, while 0.4 – 0.9 represent healthy crops (Herring, 2000).

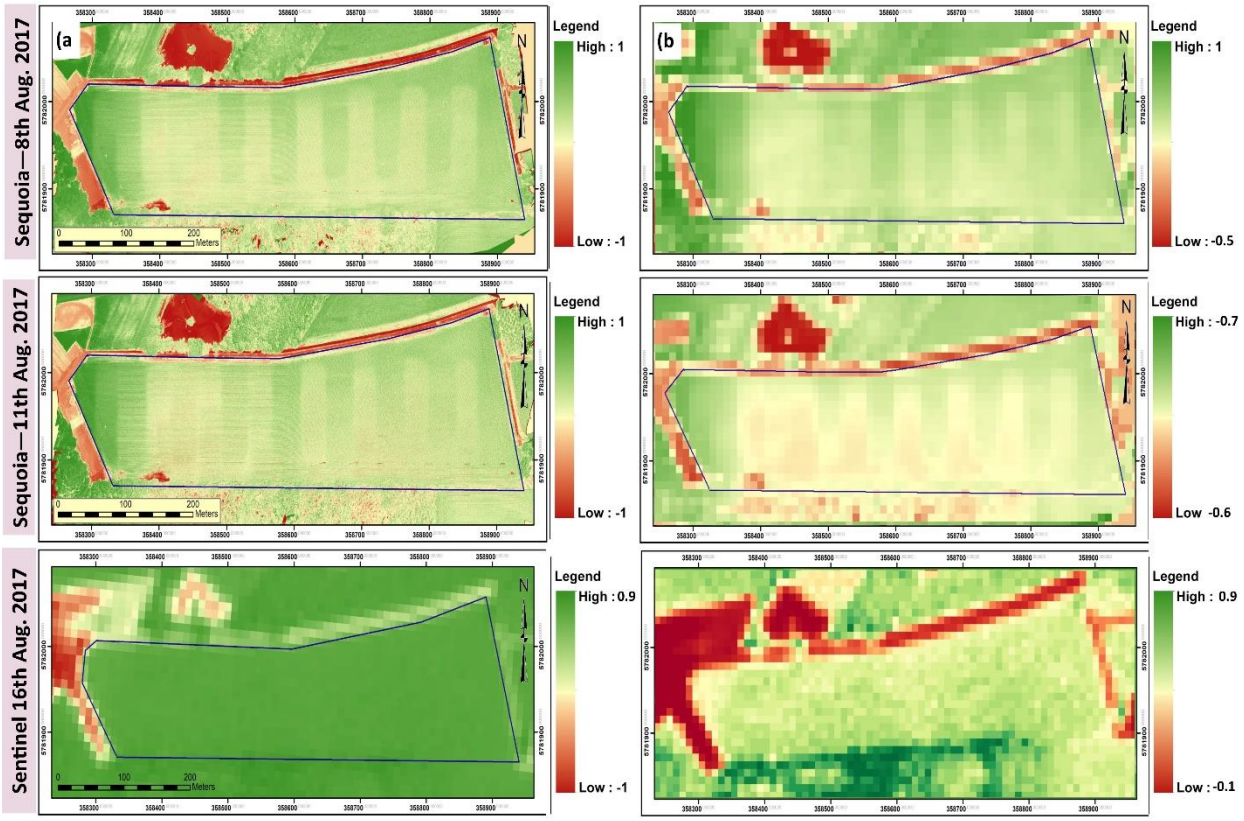


Figure 5.23. The NDVI maps (a) Original imagery (b) Aggregated UAV image and stretched Sentinel image.

NDVI derived from Sequoia and Sentinel 2B was compared. UAV-derived NDVI was aggregated by a cell factor of 147.5 to obtain 10m x 10m cells. The aggregation method was based on the mean value of original pixels that fall within the output cell. As visualized in Figure 5.23, Sentinel 2B depicts a homogeneous crop status across the farm in (a); however, histogram stretching in the range of -0.1 to 1 (b), reveals spatial variability of NDVI; with -0.1 to 0.1 being bare ground. Contrary to expectations, intra-farm spatial variability of NDVI is not the same between Sentinel 2B and Parrot Sequoia. For example, UAV-derived NDVI depicts high values west of the farm (in both acquisitions) whereas Sentinel 2B show medium values. However, it is important to note that the systematic vertical spatial variability seen in the UAV imagery as shown in Figure 5.23 is not attributed to the photosynthetic status of the crops but to the propagated photogrammetric processing error during image stitching per flight strip. It is suspected that the flight parameters used in the first and second acquisitions were not optimal. Irrespective of this setback, UAV imagery reveal a patch of bare soil south west of the farm (approximately 400m^2), which is supposedly due to crop failure since crop rows are visible, but this is not captured by Sentinel 2B. (see Figure 5.24 for a detailed view).

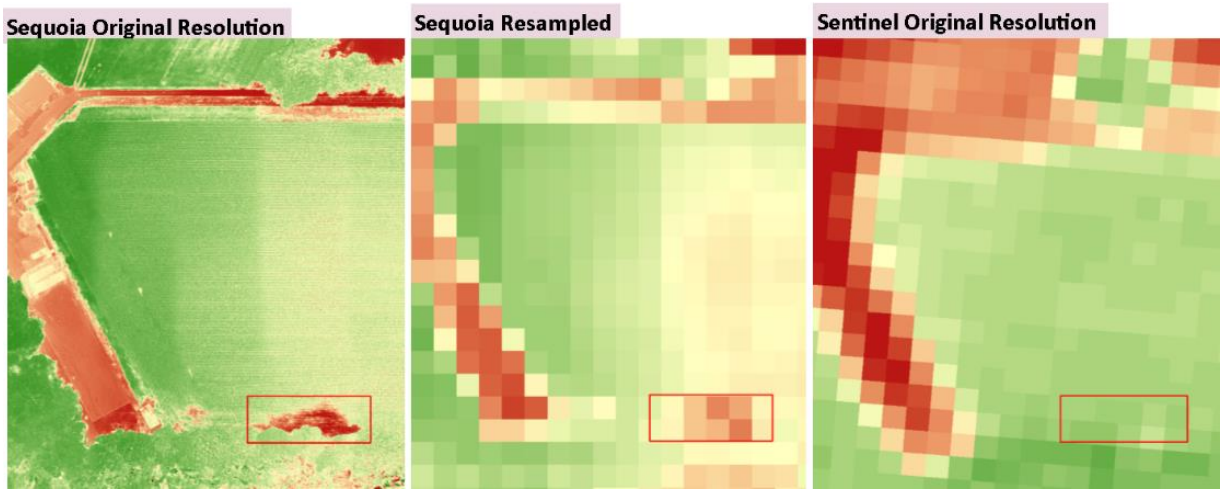
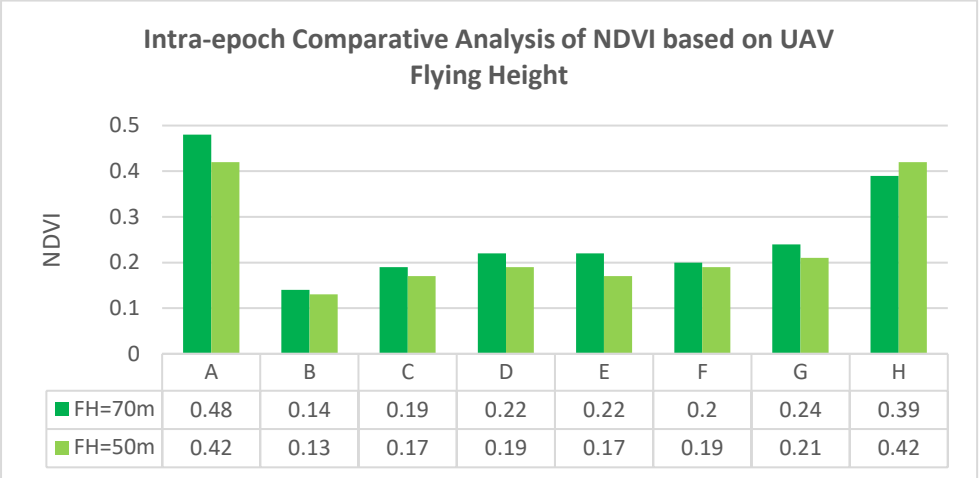


Figure 5.24 Showing NDVI spatial variability between Parrot Sequoia and Sentinel 2B

5.8.1. Intra-epoch NDVI Analysis

For each sampled zone (A-H), descriptive statistics were derived using the zonal statistics tool in ArcGIS. Figure 5.25 presents the zonal mean NDVI values for images acquired on the same day but at different flying heights. The mean NDVI infers mean productivity of the maize crop, and the standard deviations shows the spatial variability in productivity (Candiago et al., 2015). See appendix 8 for complete descriptive statistics.



	A	B	C	D	E	F	G	H
FH=70m	0.48	0.14	0.19	0.22	0.22	0.2	0.24	0.39
FH=50m	0.42	0.13	0.17	0.19	0.17	0.19	0.21	0.42

Figure 5.25 Zonal mean NDVI values of epoch 1 for images acquired with different flying height

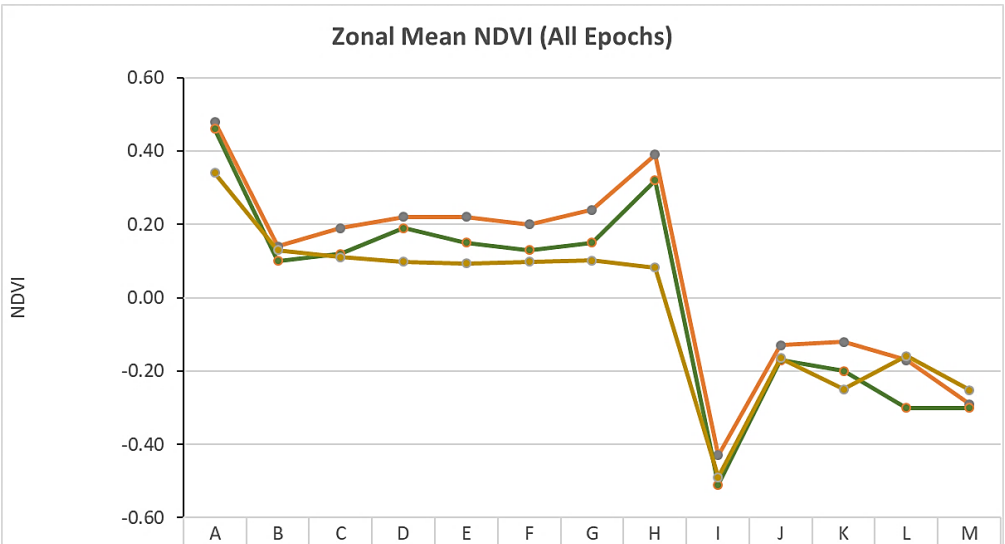
Zones A and H correspond to the most photosynthetically active vegetation, having scored 0.48, 0.42 and 0.39, 0.42 respectively. The two datasets are highly correlated, with the mean NDVI of all the zones scoring a correlation coefficient of 0.97. With a mean deviation of 0.05, zones B and F, both within the maize field, have the lowest deviation of 0.1.

Figure 5.26. Statistical comparison of NDVI values between Epoch 1 orthophotos

According to the zonal statistics, the highest deviations are seen in zones J and K, bare ground and roof respectively (see figure 5.26). This is somewhat counterintuitive since minimal deviation was expected in these zones (validation zones) because the acquisitions are of the same day, within the same hour.

5.8.2. Inter-epoch Analysis of NDVI

Having established the variations in NDVI from images acquired in the same day but with different spatial resolutions, an inter-epoch comparative analysis of vegetation performance was done for all the three epochs. Being a qualitative assessment, the zonal statistics presented in Figure 5.27 illustrate the observed spatiotemporal dynamics of NDVI values.



	A	B	C	D	E	F	G	H	I	J	K	L	M
Ep1(FH=70m)	0.48	0.14	0.19	0.22	0.22	0.20	0.24	0.39	-0.43	-0.13	-0.12	-0.17	-0.29
Ep2(FH=70m)	0.46	0.10	0.12	0.19	0.15	0.13	0.15	0.32	-0.51	-0.17	-0.20	-0.30	-0.30
Ep3(FH=50m)	0.34	0.13	0.11	0.10	0.09	0.10	0.10	0.08	-0.49	-0.17	-0.25	-0.16	-0.25

Figure 5.27. Statistical comparison of NDVI values between epoch 1 to epoch 3

Spatiotemporal analysis of NDVI using UAV images depicted a declining trend over time; this is possibly due reduced photosynthetic activity between fruit development (i.e. Silking) and ripening stages of maize. In addition, the heavy storm event experienced between epoch two and three could be a possible reason for lower NDVI values in epoch three due to damaged crops (See Figure 5.28; more soil is exposed within the farm and Figure 5.29 for diminishing greenness of the crop over the three epochs).

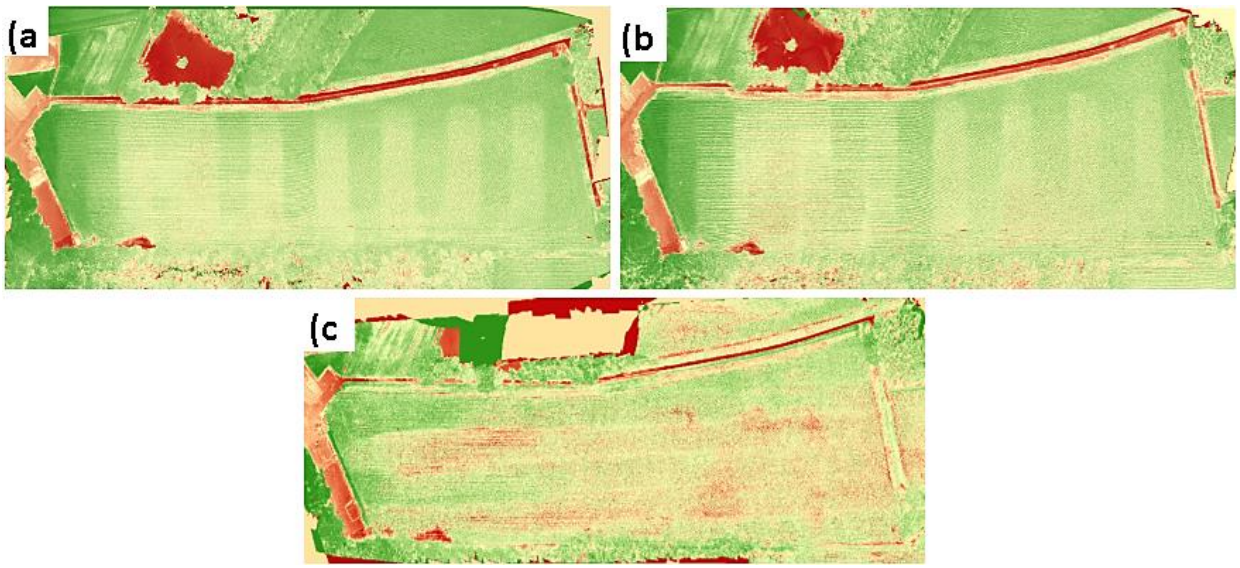


Figure 5.28. NDVI maps of (a) Epoch 1 (b)Epoch 2 (c) Epoch 3 showing aftermath of storm in epoch 3

Figure 5.29. Showing the crop greenness between epoch 1 and epoch 3

The greenness of maize in epoch one and three as shown in Figure 5.29 alludes to a reduction in the chlorophyll content, and as such the NDVI scores in epoch three are lower than those of epoch 1. Zones J, L and M perform better in epoch three; this can be attributed to the post-storm effects since acquisition of epoch three was done after a storm. Grass could have started sprouting on zone J, and patches in zone M.

5.8.3. Sequoia Versus Micro MCA NDVI Analysis

Comparative zonal statistics were computed for the zones marked A – I in both orthophotos as shown in Figure 5.30; A, C, D, E, G, H and I are within the maize field but with different crop densities. There are evident peripheral differences in the NDVI scores that can be attributed to have been propagated from the orthophoto distortions. Nonetheless, central areas without distortion are quite similar.

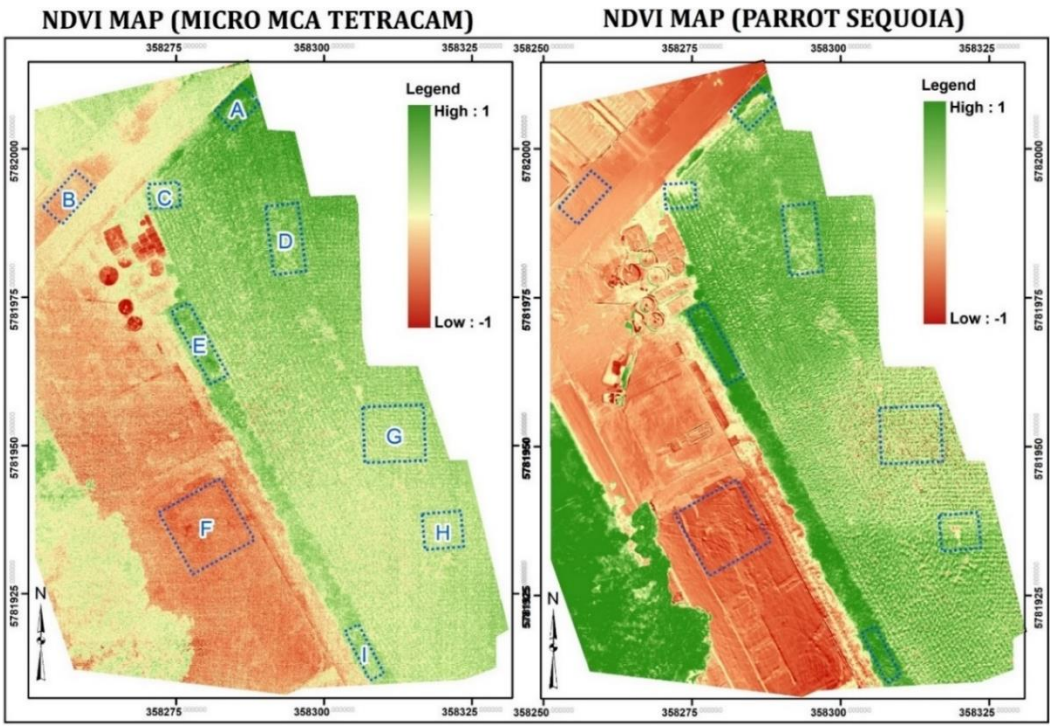
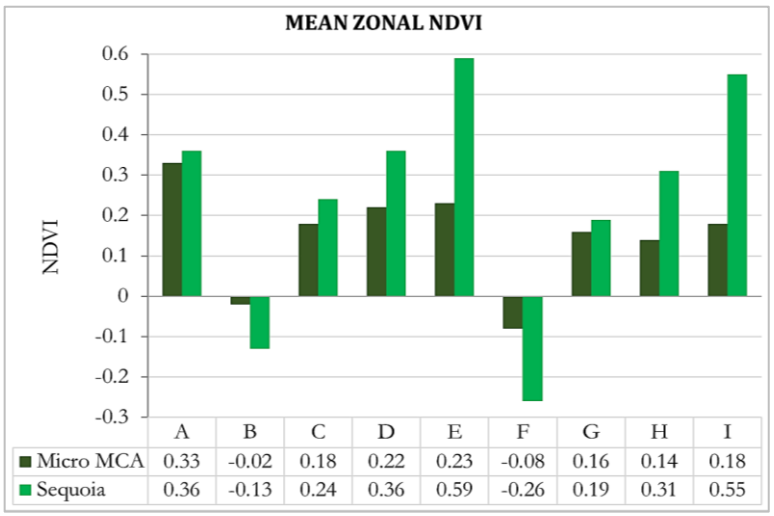


Figure 5.30. NDVI maps of the Micro MCA Tetracam and the Parrot Sequoia

The results suggested that the VI response as extracted with Parrot Sequoia and Micro MCA Tetracam are highly correlated; the spatial averages of each zone recorded a positive correlation of 0.86. Despite the differences in GSD between the two cameras, and the poor image quality of Micro MCA Tetracam, the spatial variability of NDVI is comparable. The average deviation of the mean values between the two datasets is 0.16, with the Parrot Sequoia registering higher scores than Micro MCA Tetracam. As shown in Figure 5.31, the highest average deviations are seen in zones E and I; Micro MCA depicts comparatively low NDVI values than Sequoia which reflects crops with more photosynthetic vigour. See appendix 8b for more descriptive statistics. Based on the observed results, it is practical not to combine different UAV acquisitions for crop monitoring applications.

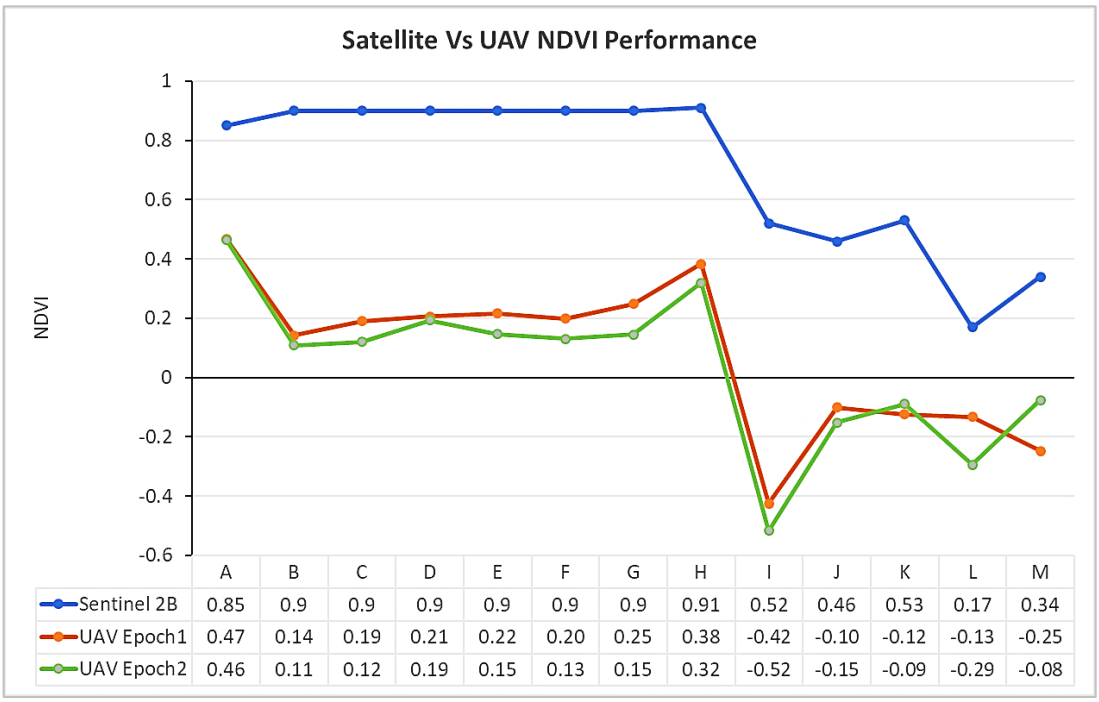


	A	B	C	D	E	F	G	H	I
Micro MCA	0.33	-0.02	0.18	0.22	0.23	-0.08	0.16	0.14	0.18
Sequoia	0.36	-0.13	0.24	0.36	0.59	-0.26	0.19	0.31	0.55

Figure 5. 31. Zonal NDVI average for Micro MCA and Parrot Sequoia

5.8.4. Comparative Analysis of Satellite and UAV-derived NDVI

Satellite-derived NDVI was compared to NDVI values from UAV orthoimages. The goal was to find out whether the spectral features of the two datasets are within the same range; validate the application of multispectral UAV imagery for crop monitoring operations; and to account for or explain the variability between the sensors. Figure 5.32 shows a plot of the zonal NDVI mean value for zones A to M as derived from Sentinel 2B (acquired on 16th, Aug. 2017) and epoch 1 and 2 UAV images acquired on 8th and 11th Aug. 2017 respectively.



	A	B	C	D	E	F	G	H	I	J	K	L	M
Sentinel 2B	0.85	0.9	0.9	0.9	0.9	0.9	0.9	0.91	0.52	0.46	0.53	0.17	0.34
UAV Epoch1	0.47	0.14	0.19	0.21	0.22	0.20	0.25	0.38	-0.42	-0.10	-0.12	-0.13	-0.25
UAV Epoch2	0.46	0.11	0.12	0.19	0.15	0.13	0.15	0.32	-0.52	-0.15	-0.09	-0.29	-0.08

Figure 5.32. Zonal mean values for Sentinel 2B and UAV imagery

Sentinel 2B depicts a homogeneous crop health status across the farm; its NDVI values increase slightly from A to B and plateaus across to H. On the other hand, zone 'I' (water pond) was expected to have the lowest value for all the sensors but this was not the case for Sentinel 2B; J, L and M (bare ground, tarmac and concrete respectively) had lower values. This effect can be attributed to the mixed pixel problem since the probability of having more feature classes per pixel at coarser spatial resolutions is higher. However, although the range of values differ, Sentinel 2B and UAV NDVI record similar trends in most zones; a strong positive correlation of 0.93 and 0.91 was obtained in zones B – H for epoch 1 and 2 respectively, and 0.77 in zones J-M for epoch 2. Zone 'A' was not used in the computation since the UAV value strongly affected the correlation negatively. For pixelwise comparative analysis plots, refer to appendix 9.

To further compare and understand the relationship between the spectral features of Sentinel 2B and Parrot Sequoia, a linear relationship was tested. A scatter plot combined with a trend line was used to model this relationship. As already established with the correlation values, a positive linear relationship is established. Naturally, spectral features derived from images with closer acquisitions dates should be more related. However, R^2 values (0.65 for epoch 1 and 0.59 for epoch 2) show that the relationship between epoch 1 and Sentinel 2B is stronger than that of epoch 2. Figure 5.33 shows a scatter plot of the zonal mean NDVIs for epoch 1 and 2.

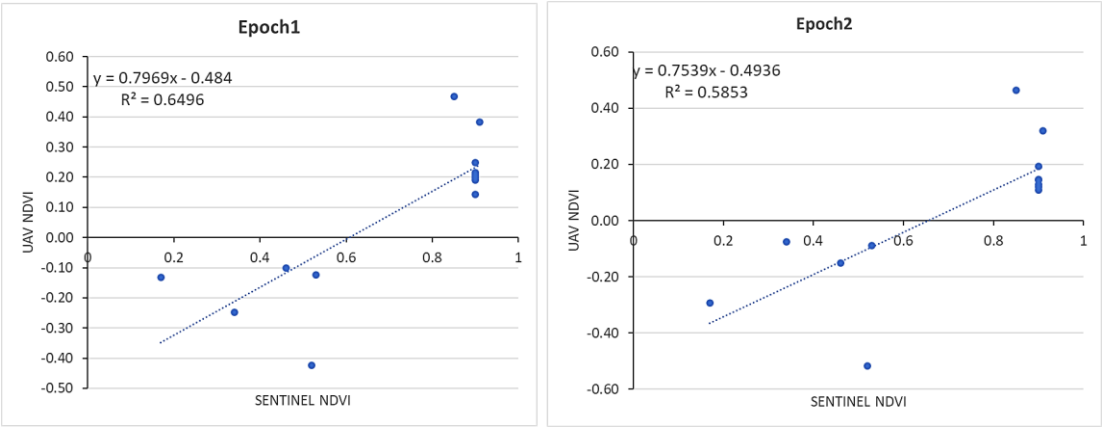


Figure 5.33. Scatter plots for Epoch 1 and Epoch 2 NDVI scores on Sentinel NDVI

5.9. Spectral Variability Analysis

To compare the spectral variability between the two cameras, spectral signatures of two classes of crops (photosynthetically active and less active) were extracted from corresponding composite images of Parrot Sequoia and Micro MCA Tetracam. Based on the results of the NDVI values, photosynthetically active and less active vegetation areas were sampled (see Figure 5.34b) using the signature editor tool in the classification tool bar in ArcGIS. The signature file was then exported and the mean DN value of every class plotted as shown in Figure 5.34a.

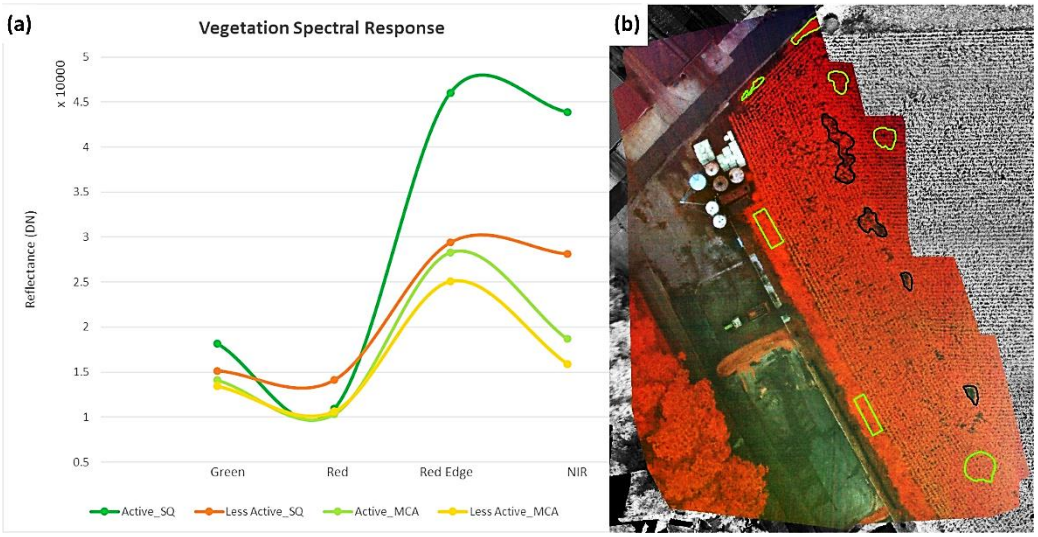


Figure 5.34. (a) Spectral analysis of cameras and crops (b) Sampled areas for both cameras; Sequoia (grayscale) and Micro MCA (false colour)

As depicted in Figure 5.34, both cameras are able to spectrally distinguish between photosynthetically active (green samples) and less active (black samples) vegetation; the green band reflects more than the red which absorbs most reflectance, and a sharp transition is seen in the red edge. Given that both composite images are of the same radiometric resolution (i.e. 16bits), the reflectance (in DN values) curves of the Micro MCA Tetracam perform comparatively low. The average spectral deviation between the cameras was by a factor of 1.6 and 1.3 for photosynthetically active and less active crops. The highest deviation was seen with the NIR bands; this could be attributed to the poor image quality of this particular acquisition as highlighted before; difference in camera calibration and spectral band widths.

6. DISCUSSION

This chapter presents discussions from the methodologies used and the results and analysis obtained in this research. Contributions are made considering published works, and a comparison with existing research is done. The loopholes of the research and possible areas of improvement are highlighted.

6.1. Master band selection

The spectral ranges of each band differ in radiant intensity per unit wavelength thus varying the spectral intensity per band (Zhu & Blackborow, 2011). Since this research employed intensity-based co-registration techniques, feature detection results reveal that detection of keypoints is unique for each band. According to results presented in section 5.1 the number of keypoints detected per band are higher in the bands sensitive to vegetation (red edge, green and near infrared). However, the red edge is selected as the master band since most features were detected on it. This is possibly due to the fact that, the red edge is the most sensitive to photosynthetic properties of vegetation thus sensing vegetation better than the rest. Jhan et al. (2017), use the red edge as the master band for intra-epoch co-registration and investigation of the parallax issue between multispectral cameras. In this study, the red edge was master band for intra-epoch and many-to-one inter-epoch co-registration approaches. As observed in the results and the works of Jhan et al. (2017), using the red edge as the master band gave a band-to-band matching sequence that had maximum point pair similarity.

6.2. Intra-epoch Co-registration

In this research, intra-epoch misregistration was assessed using two approaches; multispectral bands processed individually, one band at a time; and collective processing using Pix4d's multispectral template. In both cases, misregistration was observed to be systematic. This was concluded to be due to the basis (i.e. distance between the camera lenses). The displacement and relative orientation between the master (red edge) and the slaves was seen to be the same as the relative positions of the cameras to each other. In the results presented in Figures 5.10 and 5.11 in section 5.3.1, it is evident that the displacement is directional along the Y axis between the red edge and NIR for all the four sets of images sampled from different camera positions within the image block. The displacement along the X axis is at a subpixel level. Surprisingly, the displacements between the red edge, red and green combinations were not observed to be clearly systematic; probable causes would be the precision of the GPS, and additional displacements during movement of the UAV.

The results of Pix4d's photogrammetric outputs showed that collectively processed bands of Parrot Sequoia achieved sub-pixel co-registration accuracy in the range of 0.16 – 0.22 which is fair enough for crop monitoring; the results of Townshend et al. (1992), demonstrate that to obtain NDVI values with a 90% confidence level, a co-registration accuracy of 0.2 pixels or less should be obtained. On the other other hand, single undistorted images achieved registered using SURF and KAZE obtained co-registration accuracies in the range of 0.2 – 0.46, and individually processed bands in the range of 0.32 – 0.78. This is because in Pix4D's multispectral processing template, image matching is done simulteneously hence higher co-registration accuracies compared to individual processing which utilizes the encoded locational information (i.e. EXIF) in images per band separately. Jhan et al. (2017), record co-registration accuracies in the range of 0.25 – 0.38 using a robust and adaptive band-to-band image transform algorithm, which is comparable to the results achieved in registering undistorted images in this study. In light of these results, Pix4d's photogrammetric outputs of multispectral imagery using the multispectral template is directly usable for monitoring puposes.

Interestingly, intra-epoch co-registration accuracy of Micro MCA Tetracam obtained in this study using PW2 is comparatively higher that what was recorded in the works of Jhan et al. (2016), who observed that PW2 band co-registration errors were in the range of 1.23 – 3.81 pixels. The results of this study showed that PW2 errors were in the range of 0.13 – 0.17. Co-registration accuracy assessment in this study was based on the horizontal positional RMSE of conjugate points. The obtained results could mean that Pixel Wrench improved on the camera calibration model or the accuracies observed are unreliable since the Micro MCA Tetracam image block was not geotagged nor scaled using any GCPs. Nonetheless, the accuracies achieved in both cameras for intra-epoch co-registration were quite satisfactory for crop monitoring and accurate VI extraction.

6.3. Inter-epoch Co-registration

Two image co-registration approaches were tested in this study; many-to-one and one-to-one. The experiments involved estimating the relative orientation between bands from different epochs using inlying conjugate points. For the many-to-one approach, the red edge of epoch one was used as the cardinal master band upon which all other bands were aligned; the co-registration errors were thus only for a matching sequence of red edge and slave bands from epoch one and all other subsequent epochs. The results obtained for this approach recorded co-registration accuracy in the range of 0.47 – 0.57 and 0.51 – 0.64 for SURF and KAZE respectively. Albeit not on multispectral but RGB imagery, the Reference Image Block Co-registration (RIBC) approach by Aicardi et al. (2016) obtained results in the range of 0.4 – 1 for most epochs using both manual and automated approaches. On the other hand, the one-to-one approach proposed having all bands of epoch one as master bands; as such, co-registration would only be to corresponding bands. It is however important to note that all the master bands must also be accurately aligned to each other. In essence, a one-to-one co-registration approach should be preceded by a many-to-one intra-epoch registration of the reference epoch. The one-to-one approach obtained co-registration errors in the range of 0.28 – 0.39 and 0.30 – 0.36 for SURF and KAZE respectively. The better results obtained in the one-to-one approach can be alluded to spectral similarities between corresponding bands. As mentioned earlier, bands of similar spectral domains are expected to yield more correct matches and well distributed point pairs across the image; this is crucial for accurate estimation of relative orientation.

As partly demonstrated in this study, the intensity-based feature detectors are sensitive to texture; drastic scene changes due to heavy storm or simply illumination differences can significantly affect co-registration accuracy. In this research, illumination was not a major problem but due to a heavy storm between the second and third epoch, registration of the third epoch to the reference epoch suffered a setback of higher co-registration errors greater than a pixel. In addition, due to the poor image quality of the Micro MCA Tetracam, and the inability of intensity features to find enough correct matches to estimate geometric transformation in both many-to-one and one-to-one approaches, manual co-registration using ArcGIS georeferencing tool was done. The co-registration accuracy assessment of the manual registration was based on visual assessment thus qualitative.

Co-registration of DSMs from different epochs was done using the geometric transformation parameters of corresponding band pairs. Assumption was made that the inter-band and inter-DSM displacement is similar; hence propagating the co-registration parameters. The co-registration error assessment was qualitative; visual comparison of displacement before and after co-registration showed that subpixel accuracies could be achieved but artificial lines of misregistration were still evident in the extents of the image block. A possible area of pending research would be to compare the estimated geometric transformation between spectral bands using RANSAC and that between DSMs using phase correlation.

In this research, positional RMSE of the inlying conjugate points was used for accuracy assessment. Values of less than 0.5 of a pixel were considered precise since 0.5 of a pixel was the set threshold. As is seen in the works of Liang Tang, (2011); Coulter & Stow, (2008); and Behling et al. (2014), depending on the metrics used, the lower the value the more precise the co-registration. In addition, the distance between point pairs was used to assess the registration accuracy (i.e. minimum, maximum and mean values before and after co-registration). Another possible way to assess the co-registration accuracy could be to compute the epipolar geometry between band-pairs and compute the residual error of the distances of matched points from their corresponding epipolar lines (Onyango, Nex, Peter, & Jende, 2017). It is however important to note that accuracy varies depending on the method used to estimate geometric transformation. Since the image block in this study was orthorectified, assumptions of planarity, parallelism and similarity were made. Thus, the use of similarity transformation method. On the flipside, Jhan et al. (2017), argue that image planes are not exactly parallel thus the use of similarity and affine transformation for band-to-band co-registration is unsuitable. In light of the co-registration results, it is prudent to use one photogrammetric platform to process UAV images for optimal co-registration accuracy. In addition, for successful inter-sensor co-registration (i.e. Parrot Sequoia and Micro MCA Tetracam), image quality is a vital factor to consider.

6.4. Algorithm performance

The results showed that region detectors that use float descriptors are more robust than the ones using binary detectors. SURF and KAZE detected and indexed correct matches in all the spectral channels while the binary descriptors failed to find enough keypoints to qualify as inliers for all the band combinations. Feature detection is a function of the contrast threshold, scale levels and number of octaves. These parameters are all vital for optimal performance of the algorithms. However, different algorithms rely on deferent methods to estimate image intensity gradients (i.e. LoG, DoG, Hessian Matrix etc.). Binary descriptors simply compare pixel differences whereas float descriptors compute intensity gradients; this might be the binary detectors' setback to extracting sufficient conjugate points across all the bands. The parameters of SURF and KAZE were tuned for optimal performance and computational efficiency. Both algorithms obtained subpixel accuracies but SURF was observed to be faster; KAZE employs the additive operator splitting, which has been reported to be quite inefficient (Gerke, Nex, & Jende, 2016). On the other hand, KAZE is more rigorous in feature detection than SURF. The results of this study demonstrated that both KAZE and SURF are effective algorithms for co-registration of multispectral images. However, SURF is more suitable since it is sufficiently robust and faster in feature extraction. Thus, making it efficient and computationally inexpensive.

6.5. Vegetation Index Analysis

The influence of misregistration on the accuracy of VI extraction and analysis will depend on the scale of the crop of interest and the scene properties. According to Townshend et al. (1992), to obtain NDVI values with a 90% confidence level, an intra-epoch co-registration accuracy of 0.2 pixels or less should be obtained; and for heterogeneous and sparse scenes, 0.5 pixels will suffice. In this research, co-registration accuracies of between 0.1 and 0.3 pixels were obtained for both Parrot Sequoia and the Micro MCA. To minimize the effects of remaining misregistration error on extracted VIs, Coulter & Stow (2008) propose spatial aggregation and thresholding of VI.

Spatiotemporal analysis of NDVI using UAV images depicted a declining trend over time; this is possibly due reduced photosynthetic activity between fruit development (i.e. Silking) and ripening stages of maize. In addition, the heavy storm event experienced between epoch two and three could be a possible reason for lower NDVI values in epoch three due to damaged crops. Photogrammetric processing errors were observed to be propagated to the spectral features; the systematic alternating strips in epoch one and two do not depict the crop vigour but flight strips. The strips are not in epoch three due to a change in flight line orientation. Interestingly, the zonal statistics of the alternating strips show that there is little discrepancy in NDVI between these zones (refer to Figures 5.27 and 5.31). On the other hand, a comparison of spectral features extracted from UAV imagery of different GSD revealed that spatial variability increases with increased spatial resolution; this assessment was however qualitative.

Although the spatial resolution and spectral bandwidths of UAVs and satellite imagery differ, some level of correspondence of spectral features is expected. A comparison of NDVI derived from Sentinel 2B and Parrot Sequoia revealed a weak correspondence in spatial variability between the two sensors. Qualitative assessment suggested that the farm's crop vigour was homogeneous. As illustrated in the results and analysis section, it was noted that Sentinel 2B failed to capture certain intra-farm variability. Upon further scaling of NDVI values in the range of -0.1 to 0.9 spatial variability was appreciated. Surprisingly, quantitative assessment of NDVI derived from Sentinel 2B and Parrot Sequoia showed a strong positive correlation of 0.93 on the maize field, and 0.77 on non-vegetated areas. In contrast, comparatively low R^2 values of NDVI between the two sensors suggested a weak relationship and poor variability accountability of spectral features. The differences observed between the satellite and UAV sensors could be attributed to differing spectral bands widths; although the ranges overlap and are somewhat the same, they are not identical. Also, the variation may be due to resolution difference and whether Sentinel 2B data is reflectance or scaled radiance; and propagated errors during UAV image acquisition and processing. In related works, Kelcey & Lucieer, (2012B), compare the spectral features between WorldView-2 and Micro MCA; the correlation of spectral features was higher than the results obtained in this study. Based on the two results, it is apparent that camera calibration of the UAV dataset is key to achieving more comparable results to the satellite derived features.

6.6. Spectral Analysis of cameras

An attempt was made to analyse the vegetation spectral response of each camera. The obtained results were compared to the reflectance curves presented in Jhan et al. (2017) and Candiago et al. (2015). With the understanding that camera calibration and central spectral wavelengths of corresponding bands differ, evaluating the spectral response of vegetation for each camera was essential. It was observed that the spectral response from Micro MCA Tetracam was comparatively too low. This was attributed to poor image quality, filter transmission errors, and relative monochromatic efficiency (Kelcey & Lucieer, 2012B); a correction process is necessary to restore DN relationships between bands. The correction should only operate upon the radiance component of Micro MCA's raw data. Interestingly, although hazy and non-scaled during photogrammetric processing, the Micro MCA was still able to map out the spatial variability of spectral feature (NDVI) in a similar manner as Parrot Sequoia. To this end, the application of UAV multispectral imagery for crop monitoring is seen to be effective, efficient and adaptive. Appropriate camera calibration, correction of image distortions, image noise reduction and other forms of error elimination is key to extraction of reliable spectral features.

7. CONCLUSION AND RECOMMENDATIONS

7.1. Conclusion

The main objective of this research was to investigate intra-epoch and inter-epoch misregistration of multispectral UAV imagery, and to explore the potentials of unmanned aerial systems for crop monitoring. This study proposes an intensity-based feature detection and description method to automatically co-register both intra-epoch and inter-epoch multispectral imagery. SURF and KAZE were tested in this research; both region detectors demonstrated the ability to co-register multispectral imagery to a subpixel accuracy level provided the conjugate points are well distributed. GCPs were not used for co-registration in this study; the band-to-band alignment was based on a master band (i.e. red edge) that acted as the reference. But this does not hamper monitoring applications (Aicardi et al., 2016); upon achieving accurate co-registration, multitemporal orthoimages can used for crop monitoring and any other monitoring application. In light of the results obtained in this study, SURF is equally robust, more efficient, and computationally inexpensive than KAZE. On the other hand, KAZE is more vigorous than SURF and always detects more keypoints hence increasing the chances to get more correct matches per band combination. Both algorithms can therefore be used satisfactorily for intra-epoch and inter-epoch co-registration, although their performance will differ based on parameterization and image quality. The presented co-registration approaches of many-to-one and one-to-one can therefore be used as described. However, there is room to improve the methodology and algorithm to automatically select the master band based of number of valid points detected, and to automatically compute inter-epoch changes in the multi-temporal orthophotos and DSMs.

The comparison of satellites and UAV-derived spectral features revealed a linear relationship between the two. However, the quality of UAV imagery and radiometric calibration is key to obtaining optimal spectral features correlation. The analysis of NDVI results between Sequoia and Micro MCA demonstrated that due to differences in spectral regimes of multispectral UAV imagery, the use of one system throughout the monitoring period is prudent. In summation, both Parrot Sequoia and Micro MCA are applicable for crop monitoring; they both have the spectral bands vital for monitoring photosynthetic activities of crops. Although Micro MCA has two more bands and can therefore sense more spectral properties, these additional spectral features offer a basis for future research.

This study suffered a setback of insufficient image acquisition; from an application point of view, acquisitions should occur from the onset of the growing season and subsequently target the phenological stages of the given crop. In addition, proper UAV system setup is recommended to avoid processing errors that are propagated throughout the workflow. For optimal photogrammetric results, sufficient time and effort should be given to flight planning and image acquisition. An integrated decision support system that employs the use of UAV-based products will go a long way in supporting smart farm management.

7.2. Answers to the Research Questions

The answers to the research questions of this study are as follows:

Objective 1: To review the application of UAVs in crop monitoring and evaluate the contemporary methods of UAV image co-registration.

a) What is the role of UAV imagery in crop monitoring and what milestones have been achieved?

An extensive literature review showed that the use of UAV for agricultural applications has greatly increased in the decade. Due to the drawbacks of satellite remote sensing and aircraft surveys, UAVs imagery bridge the gap by capturing intra-farm heterogeneity, and sensing the intrinsic spatial elements such as diseases, soil nutrients, water content, and crop health among others at a more detailed level. In addition, related works highlight an array of UAV imagery application in crop monitoring including crop phenotyping, early weed detection, and estimation of above ground biomass. The milestones covered over the years include the development of a range of crop monitoring drones with varying capabilities including prolonged flying duration and increased altitude ceiling. the development of UAV-based multispectral sensors with both broad band and narrow band channels is also a major milestone. Lastly, automated methods for processing UAV-based products are continuously being proposed including this research.

b) What is the added value of additional spectral channels for crop monitoring?

The dominant factors affecting the leaf reflectance include the leaf pigments, cell structure, and the water content. For this reason, the chemical state of photosynthetically active vegetation differs from that with decreasing or failing photosynthetic activity. A number of literature and the spectral analysis of this study show that reflectance of healthy vegetation in the visible region of the electromagnetic spectrum is minimal; reflectance is absorbed in both blue and red channels because of chlorophyll absorption. As described in literature, green, red edge, and near infrared bands are the best candidates for vegetation sensing. The green band has the greatest reflectance of crops in the visible region. It is highly correlated to the amount of chlorophyll contained in the plant. On the other hand, the cell structure of a healthy crop is an excellent reflector of NIR wavelengths; thus, measuring NIR reflectance is one way to determine crop health. The Red edge is a transition zone between the red and the NIR; it is a point of abrupt reflectance change. It is very sensitive to plant stress. For these reasons, channels that are more sensitive to the dominant factors affecting leaf reflectance are an added value to crop monitoring.

c) What are the state-of-the-art methods used for UAV image co-registration?

Automated co-registration of images is mainly based on image intensity features. According to literature and the proposed methodology in this research, intensity features that are invariant to scale, rotation and illumination define the state-of-the-art methods. Intensity-based co-registration methods employ four main steps; feature detection, feature description, feature matching and geometric transformation. In this study, region detectors were selected instead of edge and corner detectors. Prior works revealed that region detectors are best suited for aerial image registration. The contemporary region detection algorithms used for this research include SURF and KAZE. Other state-of-the-art methods are described in detail in section 2.3.

Objective 2: To establish the best band combinations for photogrammetric processing and image analysis.

a) What is the effect of different band combinations on the accuracy of band-to-band registration?

The results of this study demonstrated that different band combinations yield different number of inliers. The number and distribution of inliers differs per band combination as presented in Figure 30 in section 5.1.2 and section 5.2.1. The co-registration accuracy results show that different band combinations have different sub-pixel accuracy levels. Bands with similar spectral properties were observed to be the best combination for co-registration (refer to one-to-one co-registration approach).

b) Which is the best band combination for image co-registration within an epoch?

Based on the results of this study, the many-to-one approach presents the 'Red edge + Green' as the best band combination. However, since an RMSE deviation of 0.01 and 0.02 for 'Red edge + Red' and 'Red edge + Red edge' combinations respectively are seen, these combinations are considered equally the best for intra-epoch co-registration. Importantly, the number of inliers and a good distribution of conjugate points across the image scene lead to better co-registration accuracies.

c) Which is the best band for DSM extraction?

Due to lack of ground reference points, this research assessed the best band for DSM extraction using longitudinal profiles. The red band was observed to be the best for DSM extraction; the red band modelled the surface better than the other bands. Although this assessment was qualitative, the red band depicted voids correctly as opposed to the other bands. It is for this reason that the red band is considered the best for DSM extraction. However, a correlation of DSM spatial means demonstrated that DSMs per extracted band are highly correlated in the range of 0.97 – 1. Hence the conclusion that DSMs extracted from each band can be used for appropriate applications; the nature of the application dictates the DSM accuracy needed.

Objective 3: To accurately co-register multitemporal series of multispectral imagery and DSMs, and to assess the co-registration errors.

a) What are the inherent co-registration errors? How can these errors be minimized?

From the experiments conducted in this research, observations were made that the major source of co-registration error is incorrect matches. As illustrated in Figure 22 in section 4.3.5 of this report, most matched points are seen to be on vegetation which is highly dynamic in terms of keypoints position. In addition, uneven and clustered distribution of the inliers within the spatial domain to be registered was noted to be an intrinsic error source; evenly distributed point pairs across the image are better and reliable is estimation of relative orientation and geometric transformation. As demonstrated in the works of Aicardi et al. (2016), image enhancement could result in improved keypoints detection and distribution. Thus, minimized errors.

b) Which is the best band combination for image co-registration between epochs?

The results show that the 'Red edge + Green' is the best band combination for the many-to-one co-registration approach, and 'Red + Red' for the one-to-one approach. It is observed that bands with similar spectral properties are generally the best combination for co-registration. Further work still remains; exploring the transformation parameters of the best band combination and propagating it to the other band combinations to improve the co-registration accuracy.

c) Which registration method best fits this research problem?

The results of this study demonstrated that both SURF and KAZE obtained subpixel accuracies but SURF was seen to be ten-fold faster than KAZE. On the other hand, KAZE is rigorous in feature detection than SURF. The results demonstrated that both KAZE and SURF are effective algorithms for co-registration of multispectral images. However, for optimal co-registration accuracy algorithm parameterization, image quality, and scene properties are factors to consider. Other region detectors like SIFT and AKAZE still remain to be tested.

Objective 4: To evaluate the similarity between spectral signatures from crops as extracted from Micro-MCA Tetracam, Parrot Sequoia imagery and Sentinel 2.

a) What is the spatial variability of the spectral features extracted UAVs and Sentinel 2B?

The analysis of results in this study depicted a declining trend of NDVI over time; this is partly attributed to belated harvesting thus reduced photosynthetic activity after the maturity stage of the maize. Parrot Sequoia and Micro MCA were observed to show significant correspondence in spatial variability of NDVI. Thus, both cameras can be used for crop monitoring applications provided they are well calibrated before image acquisition. Quantitative assessment of NDVI derived from Sentinel 2B and Parrot Sequoia showed a strong positive correlation of spectral features. In contrast, comparatively low R^2 values of NDVI between the two sensors suggested a weak relationship and poor variability accountability of spectral features. The differences observed between the satellite and UAV sensors could be attributed to differing spectral bands widths; although the ranges overlap and are somewhat the same, they are not identical. Also, the variation may be due to resolution difference and whether Sentinel 2B data is reflectance or scaled radiance; and propagated errors during UAV image acquisition and processing.

b) What is the variability in spectral response of vegetation between Parrot Sequoia and Micro MCA Tetracam?

It was observed that the spectral response from Micro MCA Tetracam was comparatively too low. This was attributed to poor image quality, filter transmission errors, and relative monochromatic efficiency. A correction process is necessary to restore DN relationships between bands. Interestingly, although hazy and non-scaled during photogrammetric processing, the Micro MCA was still able to map out the spatial variability of spectral feature (NDVI) in a similar manner as Parrot Sequoia. The spectral response of vegetation between the two cameras is therefore slightly different due to spectral bandwidth difference, but are comparable. For this reason, the use of one system for crop monitoring is emphasized.

In conclusion, the scientific contribution of this thesis includes:

i) Demonstrating that intensity feature detectors and descriptors such as SURF and KAZE are suitable for intra-epoch and inter-epoch co-registration of UAV multispectral imagery for crop monitoring.
ii) Development of a descriptor-based methodology for automatic co-registration of multitemporal series of multispectral UAV imagery.
iii) Evaluating the application of UAV imagery for crop monitoring in relation to satellite imagery; exploring and demonstrating the potentials of UAV spectral features for crop monitoring.

7.3. Recommendations

In light of the results obtained in this study, future research areas emerge. This study did not use GCPs to scale, georeferenced and estimate distortions within the photogrammetric block. Thus, greater inter-epoch displacements. To investigate the misregistration error between orthophotos that have been processed using ground control points in all the epochs would go a long way in contributing to inter-epoch co-registration methods. While systematic displacement was observed at the image level and at the individually processed bands level, an evaluation of misregistration errors of the Micro MCA Tetracam at the image level, and the stability of the relative orientation of each camera per sensor location could shed more light on the accuracy of the system for crop monitoring purposes. On the other hand, this study did not suffer from illumination problems but suffered from poor image quality of the Micro MCA. To alleviate this problem, exploring image enhancement techniques like the Wallis filter to eliminate inter-epoch illumination effects and enhance hazy images could be a possible solution. However, the enhance images only form part of the co-registration process; to estimate geometric transformation between the enhanced images, and then use the transformation matrix to transform the original image. Automating co-registration of inter-epoch DSMs still remains to be done systematically; examining the geometric transformation matrix between two orthophotos, and that between two corresponding DSMs estimated by phase correlation would be a starting point. A similarity in the transformation matrices would mean the transformation between orthophotos can be propagated to register corresponding DSMs as tested in this study. Most importantly, since this study and most that preceded it used RANSAC for outlier removal, exploring or developing outlier removal methods that do not rely on random sampling would ensure a definite co-registration error. Utilizing all point pairs that meet a given distance threshold would be a starting point, instead of sampling a set of points. Lastly, to assess the spatial variability of more spectral features other than the NDVI, and to investigate the spectral variability between Parrot Sequoia and Micro MCA Tetracam.

8. LIST OF REFERENCES

Adel, E., Elmogy, M., & Elbakry, H. (2014). Image Stitching based on Feature Extraction Techniques: A Survey. *International Journal of Computer Applications*, *99*(6), 1–8. https://doi.org/10.5120/17374-7818

Aerospace, P. (2017). Vapor 55 Helicopter UAV / Drone. Retrieved February 5, 2018, from http://www.pulseaero.com/uas-products/vapor-55

Aicardi, I., Nex, F., Gerke, M., & Lingua, A. (2016). An Image-Based Approach for the Co-Registration of Multi-Temporal UAV Image Datasets. *Remote Sensing*, *8*(9), 779. https://doi.org/10.3390/rs8090779

Alcantarilla, P. F., Bartoli, A., & Davison, A. J. (2012). KAZE features. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7577 LNCS, pp. 214–227). https://doi.org/10.1007/978-3-642-33783-3_16

Allahyari, M. S., Mohammadzadeh, M., & Nastis, S. A. (2016). Agricultural experts attitude towards precision agriculture: Evidence from Guilan Agricultural Organization, Northern Iran. *Information Processing in Agriculture*, *3*(3), 183–189. https://doi.org/10.1016/j.inpa.2016.07.001

Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sensing*. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/rs5020949

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 3951 LNCS, pp. 404–417). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11744023_32

Behling, R., Roessner, S., Segl, K., Kleinschmit, B., & Kaufmann, H. (2014). Robust automated image co-registration of optical multi-sensor time series data: Database generation for multi-temporal landslide detection. *Remote Sensing*, *6*(3), 2572–2600. https://doi.org/10.3390/rs6032572

Bin, L., & Samiei, M. (2012). Comparison for Image Edge Detection Algorithms. *IOSR Journal of Computer Engineering*, *2*(6), 1–4. Retrieved from www.iosrjournals.org

Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010). BRIEF: Binary robust independent elementary features. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6314 LNCS, pp. 778–792). https://doi.org/10.1007/978-3-642-15561-1_56

Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., & Gattelli, M. (2015). Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. *Remote Sensing*, *7*(12), 4026–4047. https://doi.org/10.3390/rs70404026

Chapman, S., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M., … Jimenez-Berni, J. (2014). Pheno-Copter: A Low-Altitude, Autonomous Remote-Sensing Robotic Helicopter for High-Throughput Field-Based Phenotyping. *Agronomy*, *4*(2), 279–301. https://doi.org/10.3390/agronomy4020279

Chen, J., Zou, L. hui, Zhang, J., & Dou, L. hua. (2009). The comparison and application of corner detection algorithms. *Journal of Multimedia*, *4*(6), 435–441. https://doi.org/10.4304/jmm.4.6.435-441

Coulter, L. L., & Stow, D. A. (2008). Assessment of the spatial co-registration of multitemporal imagery from large format digital cameras in the context of detailed change detection. *Sensors*, *8*(4), 2161–2173. https://doi.org/10.3390/s8042161

Crop Copter. (2017). Vegetative Indices and Leaf Reflectance | Crop Copter. Retrieved September 25, 2017, from http://www.cropcopter.co/vegetative-indices-and-leaf-reflectance/

Dalla Mura, M., Benediktsson, J. A., Waske, B., & Bruzzone, L. (2009). Morphological attribute filters for the analysis of very high resolution remote sensing images. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS '09)*, *3*(10), 2–3. https://doi.org/10.1109/IGARSS.2009.5418096

Davies, E. R. (2005a). CHAPTER 14 – Polygon and Corner Detection. In *Machine Vision* (pp. 379–411). Elsevier. https://doi.org/10.1016/B978-012206093-9/50017-4

Davies, E. R. (2005b). Edge Detection. In *Machine Vision: Theory, Algorithms, Practicalities* (pp. 131–157). Elsevier. https://doi.org/10.1016/B978-012206093-9/50008-3

DJI. (2016a). Matrice 600 User Manual, 1–72. Retrieved from https://www.dji.com/matrice600-pro/info

DJI. (2016b). Matrice 600 User Manual, 1–72. Retrieved from https://dl.djicdn.com/downloads/m600+pro/20170717/Matrice_600_Pro_User_Manual_v1.0_EN.pdf

DJI. (2016c). Phantom 4 User Manual. Retrieved from https://dl.djicdn.com/downloads/phantom_4/en/Phantom_4_User_Manual_en_v1.0.pdf

EBee. (2015). The survey-grade mapping drone. Retrieved from https://www.geometius.nl/wp-content/uploads/2017/04/Brochure-senseFly-eBee-RTK1.pdf

Elarab, M., Ticlavilca, A. M., Torres-Rua, A. F., Maslova, I., & McKee, M. (2015). Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. *International Journal of Applied Earth Observation and Geoinformation*, *43*, 32–42. https://doi.org/10.1016/j.jag.2015.03.017

Fan, B., Wang, Z., & Wu, F. (2015). Local Image Descriptor : Modern Approaches, 99. https://doi.org/10.1007/978-3-662-49173-7

FAO. (2013). Feeding the world. In *FAO Statistical Yearbook 2013* (pp. 123–158). Retrieved from http://www.fao.org/docrep/018/i3107e/i3107e03.pdf

Fernández Alcantarilla, P., Bartoli, A., & Davison, A. J. (2012). KAZE Features. Retrieved from https://www.doc.ic.ac.uk/~ajd/Publications/alcantarilla_etal_eccv2012.pdf

Fischler, M. A., & Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Apphcatlons to Image Analysis and Automated Cartography Martin. *Communications of the ACM*, *24*(6), 381–395. https://doi.org/10.1145/358669.358692

Francisca López-Granados, Jorge Torres-Sánchez, Ana-Isabel De Castro, Angélica Serrano-Pérez, Francisco-Javier Mesas-Carrascosa, J.-M. P. (2016). Object-based early monitoring of a grass weed in a grass crop using high resolution UAV imagery. *Agronomy for Sustainable Development*, *36*(4), 67. https://doi.org/10.1007/s13593-016-0405-7

Fredrick Arthur, O. (2017). *Automated Image Registration*. University of Twente.

Freeman, P. K., & Freeland, R. S. (2015). Agricultural UAVs in the U.S.: Potential, policy, and hype. *Remote Sensing Applications: Society and Environment*, *2*, 35–43. https://doi.org/10.1016/j.rsase.2015.10.002

Fytsilis, A. L., Prokos, A., Koutroumbas, K. D., Michail, D., & Kontoes, C. C. (2016). A methodology for near real-time change detection between Unmanned Aerial Vehicle and wide area satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *119*, 165–186. https://doi.org/10.1016/j.isprsjprs.2016.06.001

Garcia-Ruiz, F., Sankaran, S., Maja, J. M., Lee, W. S., Rasmussen, J., & Ehsani, R. (2013). Comparison of two aerial imaging platforms for identification of Huanglongbing-infected citrus trees. *Computers and Electronics in Agriculture*, *91*, 106–115. https://doi.org/10.1016/j.compag.2012.12.002

Geosense. (2017). MAPIR Cameras. Retrieved February 5, 2018, from http://www.geosense.gr/en/mapir-cameras/

Gerke, M., Nex, F., & Jende, P. (2016). Co-registration of terrestrial and UAV-based images - Experimental results. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Vol. 40, pp. 11–18). https://doi.org/10.5194/isprsarchives-XL-3-W4-11-2016

Gevaert, C. M., Suomalainen, J., Tang, J., & Kooistra, L. (2015). Generation of Spectral-Temporal Response Surfaces by Combining Multispectral Satellite and Hyperspectral UAV Imagery for Precision Agriculture Applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(6), 3140–3146. https://doi.org/10.1109/JSTARS.2015.2406339

Gonzalez, R., & Woods, R. (2002). *Digital image processing*. *Prentice Hall*. https://doi.org/10.1016/0734-189X(90)90171-Q

Govaerts, B., & Verhulst, N. (2010). The normalized difference vegetation index (NDVI) Greenseeker (TM) handheld sensor: toward the integrated evaluation of crop management. *Part A-Concepts and case studies*. *International Maize and Wheat Improvement Center*. Retrieved from http://plantstress.com/methods/Greenseeker.PDF

Greene, T. (2012). *Inventory and monitoring toolbox: birds* (Vol. 1). Retrieved from http://www.doc.govt.nz/Documents/science-and-technical/inventory-monitoring/im-toolbox-birds-complete-aerial-photo-counts.pdf

Haghighattalab, A., González Pérez, L., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., … Poland, J. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant Methods*, *12*(1), 35. https://doi.org/10.1186/s13007-016-0134-6

Harris Geospatial solutions. (2017). Spectral Indices -- Documentation Center [Harris Geospatial Docs Center]. Retrieved September 27, 2017, from http://www.harrisgeospatial.com/docs/RegistrationImageToImage.html

Hassaballah, M., Abdelmgeid, A. A., & Alshazly, H. A. (2016). Image features detection, description and matching. *Studies in Computational Intelligence*, *630*, 11–45. https://doi.org/10.1007/978-3-319-28854-3_2

Herring, J. W. and D. (2000). Measuring Vegetation (NDVI & EVI) : Feature Articles. Retrieved from https://earthobservatory.nasa.gov/Features/MeasuringVegetation/

Huertas, A., & Medioni, G. (1986). Detection of intensity changes with subpixel accuracy using laplacean-gaussian masks. *IEEE Tran-Sactions on PAMI*, *8*(5), 651–664.

Inoue, Y., Sakaiya, E., & Wang, C. (2014). Capability of C-band backscattering coefficients from high-resolution satellite SAR sensors to assess biophysical variables in paddy rice. *Remote Sensing of Environment*, *140*, 257–266. https://doi.org/10.1016/j.rse.2013.09.001

Issei, H.-Y., Kazunobu, I., & Noboru, N. (2010). Satellite and aerial remote sensing for production estimates and crop assessment. *Environ.Control Biol.*, *48*(2), 51–58. https://doi.org/10.2525/ecb.48.51

Jensen, J. R. (2007). *Remote Sensing of Environment: an earth resource perespective* (2nd Editio). Upper Saddle River, USA: Pearson Prentice Hall.

Jhan, J. P., Rau, J. Y., Haala, N., & Cramer, M. (2017). Investigation of parallax issues for multi-lens multispectral camera band co-registration. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Vol. 42, pp. 157–163). https://doi.org/10.5194/isprs-archives-XLII-2-W6-157-2017

Jhan, J. P., Rau, J. Y., & Huang, C. Y. (2016). Band-to-band registration and ortho-rectification of multilens/multispectral imagery: A case study of MiniMCA-12 acquired by a fixed-wing UAS. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 66–77. https://doi.org/10.1016/j.isprsjprs.2016.01.008

K. C. Swain, S. J. Thomson, H. P. W. J. (2010). Adoption of an Unmanned Helicopter for Low-Altitude Remote Sensing To Estimate Yield and Total Biomass of a Rice Crop. *American Society of Agricultural and*

Biological Engineers, *53*(1), 21–27. https://doi.org/10.13031/2013.29493
Kelcey, J., & Lucieer, A. (2012a). Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, *4*(5), 1462–1493. https://doi.org/10.3390/rs4051462
Kelcey, J., & Lucieer, A. (2012b). Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, *4*(5), 1462–1493. https://doi.org/10.3390/rs4051462
Kong, H., Akakin, H. C., & Sarma, S. E. (2013). A generalized laplacian of gaussian filter for blob detection and its applications. *IEEE Transactions on Cybernetics*, *43*(6), 1719–1733. https://doi.org/10.1109/TSMCB.2012.2228639
Lee, W. S., Alchanatis, V., Yang, C., Hirafuji, M., Moshou, D., & Li, C. (2010). Sensing technologies for precision specialty crop production. *Computers and Electronics in Agriculture*, *74*(1), 2–33. https://doi.org/10.1016/j.compag.2010.08.005
Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011). BRISK: Binary Robust invariant scalable keypoints. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2548–2555). IEEE. https://doi.org/10.1109/ICCV.2011.6126542
Levi, G., & Hassner, T. (2015). LATCH: Learned Arrangements of Three Patch Codes. Retrieved from https://arxiv.org/pdf/1501.03719.pdf
Link, J., Senner, D., & Claupein, W. (2013). Developing and evaluating an aerial sensor platform (ASP) to collect multispectral data for deriving management decisions in precision farming. *Computers and Electronics in Agriculture*, *94*, 20–28. https://doi.org/10.1016/j.compag.2013.03.003
Lobell, D. B. (2013). The use of satellite data for crop yield gap analysis. *Field Crops Research*, *143*, 56–64. https://doi.org/10.1016/j.fcr.2012.08.008
Lopez-Sanchez, J. M., & David Ballester-Berman, J. (2009). Potentials of polarimetric SAR interferometry for agriculture monitoring. *Radio Science*, *44*(2), n/a-n/a. https://doi.org/10.1029/2008RS004078
Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision* (pp. 1150–1157 vol.2). IEEE. https://doi.org/10.1109/ICCV.1999.790410
Lowe, D. G. (2004). Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, *60*, 91–11020042. https://doi.org/http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94
Madhuri, S. (2014). Classification of Image Registration Techniques and Algorithms in Digital Image Processing – A Research Survey. *Ijctt*, *15*(2), 78–82. https://doi.org/10.14445/22312803/IJCTT-V15P118
Matese, A., Toscano, P., Di Gennaro, S. F., Genesio, L., Vaccari, F. P., Primicerio, J., … Gioli, B. (2015). Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. *Remote Sensing*, *7*(3), 2971–2990. https://doi.org/10.3390/rs70302971
MathWorks. (2017a). Detect BRISK features and return BRISKPoints object - MATLAB detectBRISKFeatures - MathWorks Benelux. Retrieved December 5, 2017, from https://nl.mathworks.com/help/vision/ref/detectbriskfeatures.html
MathWorks. (2017b). Extract interest point descriptors - MATLAB extractFeatures - MathWorks Benelux. Retrieved December 5, 2017, from https://nl.mathworks.com/help/vision/ref/extractfeatures.html?searchHighlight=extractFeatures&s_tid=doc_srchtitle
Mazur, M. (2016). Six ways drones are revolutionizing agriculture. Retrieved May 20, 2017, from https://www.technologyreview.com/s/601935/six-ways-drones-are-revolutionizing-agriculture/
Mesas-Carrascosa, F. J., Torres-Sanchez, J., Clavero-Rumbao, I., Garcia-Ferrer, A., Pena, J. M., Borra-Serrano, I., & Lopez-Granados, F. (2015). Assessing optimal flight parameters for generating accurate multispectral orthomosaicks by uav to support site-specific crop management. *Remote Sensing*, *7*(10), 12793–12814. https://doi.org/10.3390/rs71012793
Micasense. (2016). RedEdge Integration Guide – MicaSense Knowledge Base. Retrieved February 5, 2018, from https://support.micasense.com/hc/en-us/articles/226364087-RedEdge-Integration-Guide
Micasense. (2017). What does the Sunshine (Irradiance) Sensor on Sequoia do? – MicaSense Knowledge Base. Retrieved February 6, 2018, from https://support.micasense.com/hc/en-us/articles/216533688-What-does-the-Sunshine-Irradiance-Sensor-on-Sequoia-do-
MicaSense. (2017). Parrot Sequoia. *MicaSense*. Retrieved from https://dronedecole.fr/parrot-sequoia/
Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., … Van Gool, L. (2005). A comparison of affine region detectors. *International Journal of Computer Vision*, *65*(1–2), 43–72. https://doi.org/10.1007/s11263-005-3848-x
Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., & Foley, J. A. (2013). Corrigendum: Closing yield gaps through nutrient and water management. *Nature*, *494*(7437), 390–390. https://doi.org/10.1038/nature11907
Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, *114*(4), 358–371. https://doi.org/10.1016/j.biosystemseng.2012.08.009
Näsi, R., Honkavaara, E., Lyytikäinen-Saarenmaa, P., Blomqvist, M., Litkey, P., Hakala, T., … Holopainen, M. (2015). Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level. *Remote Sensing*, *7*(11), 15467–15493. https://doi.org/10.3390/rs71115467

Nex, F., Gerke, M., Remondino, F., Przybilla, H.-J., Bäumker, M., & Zurhorst, A. (2015). Isprs Benchmark for Multi-Platform Photogrammetry. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II-3/W4*(March), 135–142. https://doi.org/10.5194/isprsannals-II-3-W4-135-2015

Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: A review. *Applied Geomatics*. https://doi.org/10.1007/s12518-013-0120-x

Onyango, F. A., Nex, F., Peter, M. S., & Jende, P. (2017). Accurate estimation of orientation parameters of UAV images through image registration with aerial oblique imagery. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Vol. 42, pp. 599–605). https://doi.org/10.5194/isprs-archives-XLII-1-W1-599-2017

Parrot Sequoia. (2017). Parrot SEQUOIA | Technical Specifications. Retrieved February 5, 2018, from https://www.parrot.com/us/business-solutions/parrot-sequoia#parrot-sequoia-

Peña, J. M., Torres-Sánchez, J., de Castro, A. I., Kelly, M., & López-Granados, F. (2013). Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. *PLoS ONE*, *8*(10), e77151. https://doi.org/10.1371/journal.pone.0077151

Perona, P., & Malik, J. (1987). Scale Space and Edge Detection using Anisotropic Diffusion. *IEEE* Transactions On Pattern Analysis And Machine Intelligence, *12*(7), 16–22. https://doi.org/10.1109/34.56205

Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., & Sarazzi, D. (2011). UAV photogrammetry for mapping and 3d modeling–current status and future perspectives. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *38–1/C22*(1), 25–31. https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-25-2011

Rey, C., Martin, M., Lobo, A., Luna, I., Diago, M., Millan, B., & Tardaguila, J. (2013). Multispectral imagery acquired from a UAV to assess the spatial variability of a Tempranillo vineyard. *Precision Agriculture 2013 - Papers Presented at the 9th European Conference on Precision Agriculture, ECPA 2013*, *55*(June 2015), 617–624. https://doi.org/10.3920/978-90-8686-778-3

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., … Pavek, M. J. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *European Journal of Agronomy*. https://doi.org/10.1016/j.eja.2015.07.004

Shrivakshan, G. T., & Chandrasekar, C. (2012). A Comparison of various Edge Detection Techniques used in Image Processing. *International Journal of Computer Science Issues (IJCSI)*, *9*(5), 269–276. Retrieved from https://www.ijcsi.org/papers/IJCSI-9-5-1-269-276.pdf

Song, Y. (2016). Evaluation of the UAV-Based multispectral imagery and its application for crop intra-field nitrogen monitoring and yield prediction in Ontario, (September). Retrieved from http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=5779&context=etd

Strecha, C., Bronstein, A. M., Bronstein, M. M., Fua, P., & Member, S. (2011). LDAHash : Improved Matching with Smaller Descriptors. *IEEE* Transactions On Pattern Analysis And Machine Intelligence, *33*(10), 1–13. Retrieved from https://cvlab.epfl.ch/files/content/sites/cvlab2/files/publications/publications/2012/BronsteinF12.pdf

Suárez, L., Zarco-Tejada, P. J., Berni, J. A. J., González-Dugo, V., & Fereres, E. (2009). Modelling PRI for water stress detection using radiative transfer models. *Remote Sensing of Environment*, *113*(4), 730–744. https://doi.org/10.1016/j.rse.2008.12.001

Tang, L. (2011). The Automatic Assessment Method of Co-registration Accuracy between Remote Sensing Images, 1690–1693.

Tang, L., Li, Y., Ling, T., He, L., Luo, T., & Guo, D. (2016). UAV image registration algorithm based on overlapping region detrction. In *International Geoscience and Remote Sensing Symposium (IGARSS)* (Vol. 2016–Novem, pp. 6706–6709). IEEE. https://doi.org/10.1109/IGARSS.2016.7730751

Tetracam Inc. (2016). Tetracam Micro-MCA. Retrieved August 6, 2017, from http://www.tetracam.com/Products-Micro_MCA.htm

Tetracam Inc. (2016). Tetracam Micro-MCA. Retrieved February 6, 2018, from http://www.tetracam.com/Products-Micro_MCA.htm

Tilly, N., Hoffmeister, D., Cao, Q., Huang, S., Lenz-Wiedemann, V., Miao, Y., & Bareth, G. (2014). Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *Journal of Applied Remote Sensing*, *8*(1), 83671. https://doi.org/10.1117/1.JRS.8.083671

Torr, P. H. S., & Zisserman, A. (2000). MLESAC: A New Robust Estimator with Application to Estimating Image Geometry. *Computer Vision and Image Understanding*, *78*(1), 138–156. https://doi.org/10.1006/cviu.1999.0832

Townshend, J. R. G., Gurney, C., McManus, J., & Justice, C. O. (1992). The Impact of Misregistration on Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(5), 1054–1060. https://doi.org/10.1109/36.175340

Turner, D., Lucieer, A., Malenovský, Z., King, D. H., & Robinson, S. A. (2014a). Spatial co-registration of ultra-high resolution visible, multispectral and thermal images acquired with a micro-UAV over antarctic moss beds. *Remote Sensing*, *6*(5), 4003–4024. https://doi.org/10.3390/rs6054003

Turner, D., Lucieer, A., Malenovský, Z., King, D. H., & Robinson, S. A. (2014b). Spatial co-registration of ultra-high resolution visible, multispectral and thermal images acquired with a micro-UAV over antarctic moss

beds. *Remote Sensing*, *6*(5), 4003–4024. https://doi.org/10.3390/rs6054003

Wei, Z., Han, Y., Li, M., Yang, K., Yang, Y., Luo, Y., & Ong, S. H. (2017). A small UAV Based multi-temporal image registration for dynamic agricultural terrace monitoring. *Remote Sensing*, *9*(9). https://doi.org/10.3390/rs9090904

Weickert, J. (1998). Anisotropic diffusion in image processing. *Image Rochester NY*, *256*(3), 170. https://doi.org/10.1.1.11.751

World Bank. (2008). Agriculture for Development. *Agriculture*, *54*, 1–6. https://doi.org/10.1596/978-0-8213-7233-3

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., … Yang, H. (2017). Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Frontiers in Plant Science*, *8*(June). https://doi.org/10.3389/fpls.2017.01111

Yemane Tumlisan, G. (2017). Monitoring Growth Development and Yield Estimation of Maize using Very High-Resolution UAV- Images in Gronau, Germany. University Of Twente Faculty of Geoinformation Science and Earth Observation. Retrieved from http://www.itc.nl/library/papers_2017/msc/nrm/tumlisan.pdf

Zhu, A. H., & Blackborow, P. (2011). *Understanding Radiance (Brightness), Irradiance and Radiant Flux*. Retrieved from http://www.energetiq.com/TechLibrary/Technical_Notes/Understanding Radiance (Brightness), Irradiance and Radiant Flux - Mar 2011.pdf

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Myneni, R. B. (2013). Global data sets of vegetation leaf area index (LAI)3g and fraction of photosynthetically active radiation (FPAR)3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (NDVI3G) for the period 1981 to 2. *Remote Sensing*, *5*(2), 927–948. https://doi.org/10.3390/rs5020927

Zitová, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, *21*(11), 977–1000. https://doi.org/10.1016/S0262-8856(03)00137-9

APPENDICES

Appendix 1: Snippet of *Photogrammetric Processing Quality Report of Epoch 1*

Quality Report

Generated with Pix4Dmapper Pro version 3.2.23

Important: Click on the different icons for:

Help to analyze the results in the Quality Report

Additional information about the sections

Click here for additional tips to analyze the Quality Report

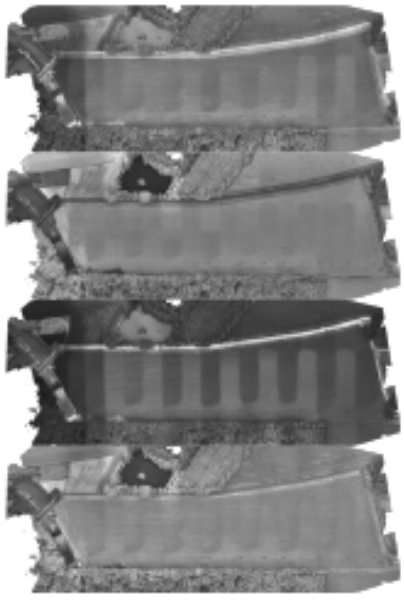
Summary

Project	70m_DayOne
Processed	2017-09-20 10:13:42
Camera Model Name(s)	Sequoia_4.0_1280x960 (Green), Sequoia_4.0_1280x960 (Red), Sequoia_4.0_1280x960 (Red edge), Sequoia_4.0_1280x960 (NIR)
Rig name(s)	«Sequoia»
Average Ground Sampling Distance (GSD)	6.84 cm / 2.69 in
Area Covered	0.2087 km^2 / 20.8653 ha / 0.0806 sq. mi. / 51.586 acres
Time for Initial Processing (without report)	54m:01s

Quality Check

Images	median of 10000 keypoints per image	✓
Dataset	1476 out of 1476 images calibrated (100%), all images enabled	✓
Camera Optimization	0.06% relative difference between initial and optimized internal camera parameters	✓
Matching	median of 5709.45 matches per calibrated image	✓
Georeferencing	yes, no 3D GCP	⚠

Preview



Appendix 2: *Snippet of Photogrammetric Processing Quality Report of Epoch 2*

Quality Report

Generated with Pix4Dmapper Pro version 3.2.23

Important: Click on the different icons for:

Help to analyze the results in the Quality Report

Additional information about the sections

Click here for additional tips to analyze the Quality Report

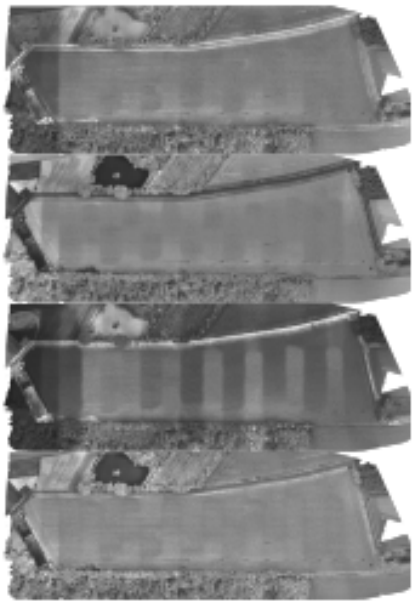
Summary

Project	1stFlight_DayTwo
Processed	2017-09-19 10:04:16
Camera Model Name(s)	Sequoia_4.0_1280x960 (Green), Sequoia_4.0_1280x960 (Red), Sequoia_4.0_1280x960 (Red edge), Sequoia_4.0_1280x960 (NIR)
Rig name(s)	«Sequoia»
Average Ground Sampling Distance (GSD)	6.69 cm / 2.63 in
Area Covered	0.1961 km^2 / 19.8113 ha / 0.0765 sq. mi. / 48.98 acres
Time for Initial Processing (without report)	01h:04m:08s

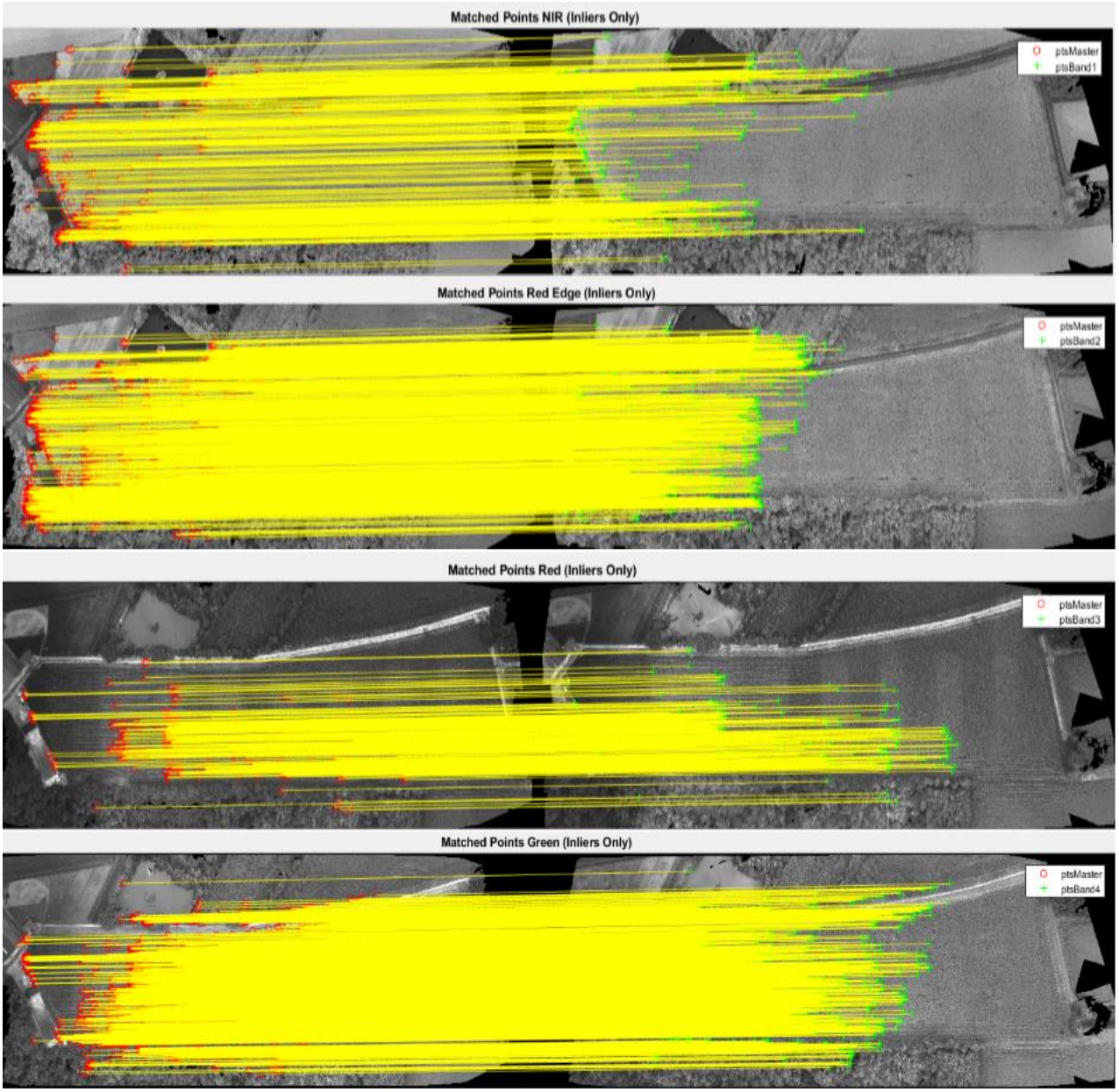
Quality Check

Images	median of 10000 keypoints per image	✓
Dataset	1724 out of 1724 images calibrated (100%), all images enabled	✓
Camera Optimization	0.09% relative difference between initial and optimized internal camera parameters	✓
Matching	median of 5612.78 matches per calibrated image	✓
Georeferencing	yes, no 3D GCP	⚠

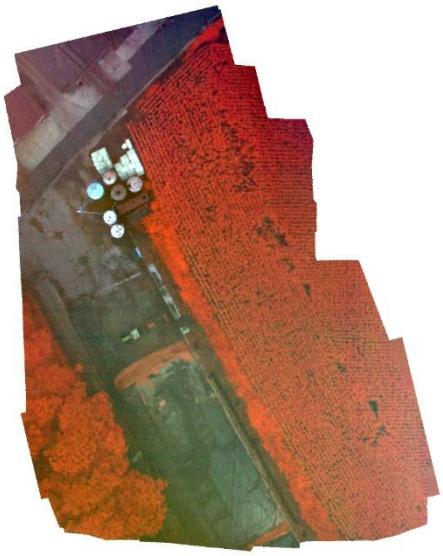
Preview

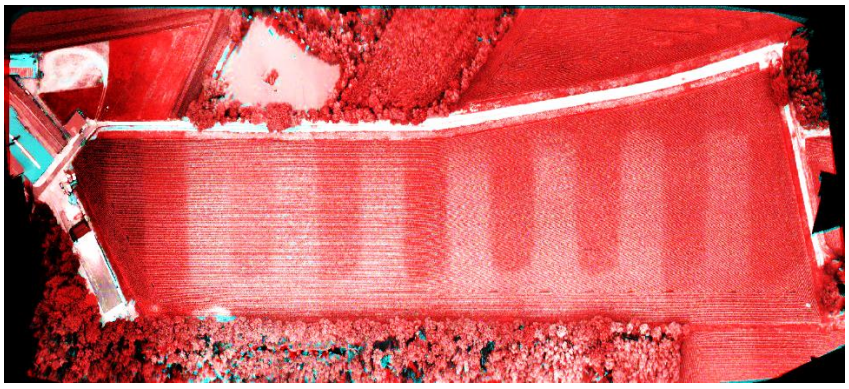


Appendix 3a: *Correct matches between corresponding orthophotos between Epoch 1 and 2*

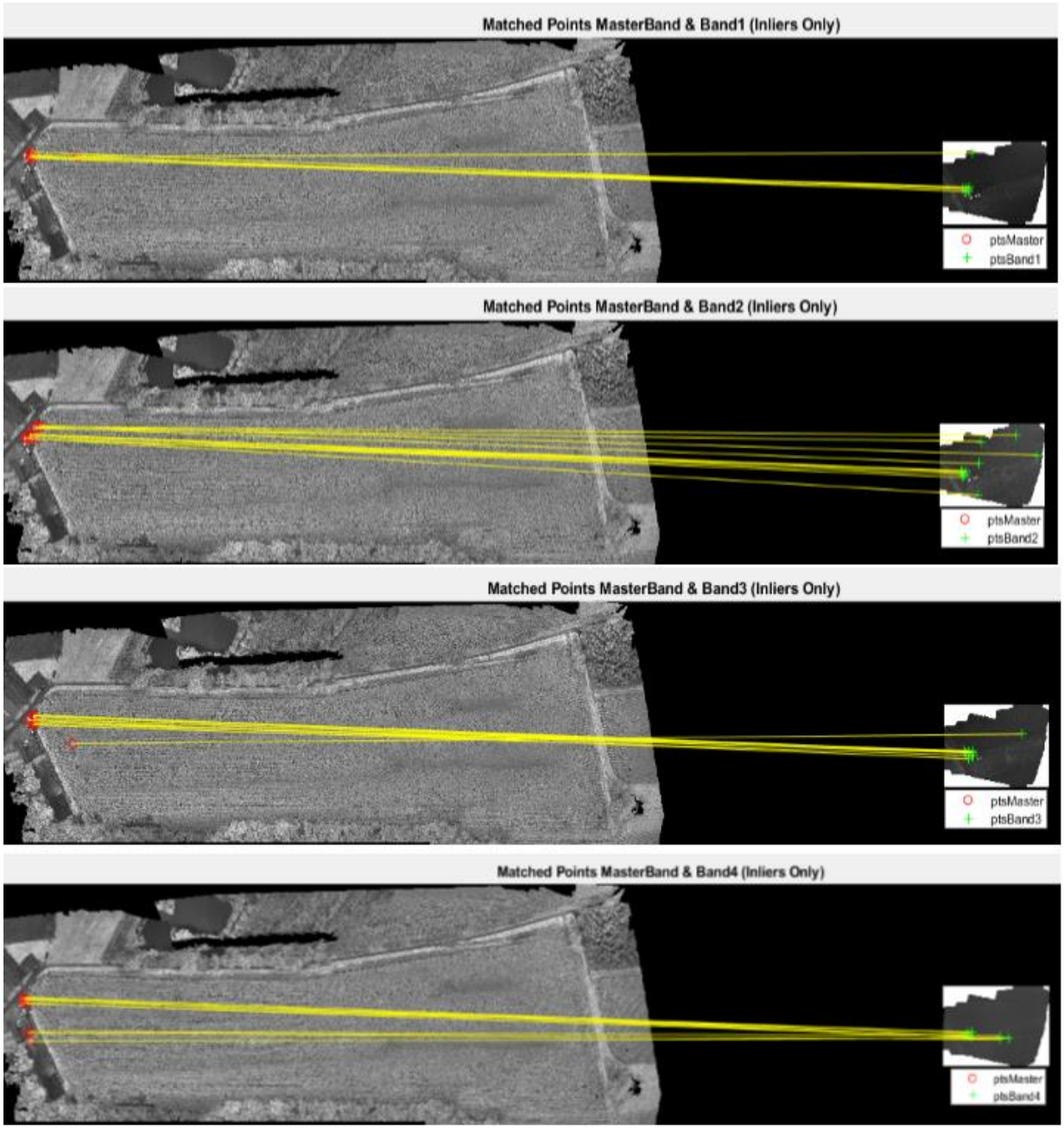


Appendix 3b: *Composite Orthophotos of Micro MCA Tetracam (left) and Parrot Sequoia (Right)*

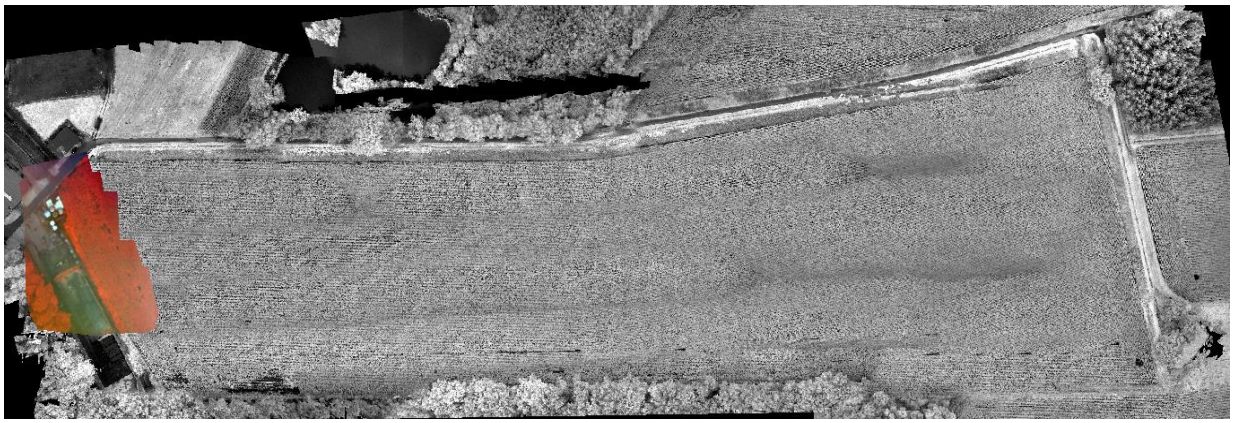




Appendix 4a: *Correct matches between corresponding orthophotos between Sequoia and Micro MCA Epoch 3*



Appendix 4b: *Registered Micro MCA Tetracam overlaid on Sequoia orthophoto of Epoch 3*



Appendix 5: *Registered Micro MCA Tetracam overlaid on Sequoia orthophoto of Epoch 3*

	Epoch One			*Epoch Two*			*Epoch Three*		
	MEAN	STD	MAX	MEAN	STD	MAX	MEAN	STD	MAX
ΔX_1	0.00	0.01	0.02	0.00	0.01	0.01	0.00	0.01	0.01
ΔX_2	0.00	0.01	0.02	0.00	0.01	0.02	0.00	0.01	0.02
ΔX_3	-0.01	0.01	0.02	-0.01	0.01	0.02	-0.01	0.01	0.02
ΔY_1	0.00	0.01	0.02	0.00	0.01	0.01	0.00	0.01	0.02
ΔY_2	0.00	0.01	0.02	0.00	0.01	0.01	0.00	0.01	0.01
ΔY_3	0.00	0.01	0.02	0.00	0.01	0.02	0.00	0.01	0.02
ΔZ_1	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01
ΔZ_2	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01
ΔZ_3	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01
$\Delta\omega_1$	0.03	0.06	0.13	0.02	0.04	0.10	0.02	0.04	0.11
$\Delta\omega_2$	-0.04	0.12	0.16	-0.07	0.13	0.19	-0.07	0.13	0.19
$\Delta\omega_3$	-0.05	0.07	0.19	-0.06	0.08	0.19	-0.06	0.07	0.19
$\Delta\varphi_1$	0.02	0.10	0.12	0.02	0.07	0.08	0.02	0.07	0.08
$\Delta\varphi_2$	0.04	0.10	0.17	0.01	0.13	0.20	0.01	0.13	0.20
$\Delta\varphi_3$	0.00	0.13	0.15	-0.02	0.15	0.18	-0.02	0.15	0.18
$\Delta\varkappa_1$	-0.04	0.02	0.00	-0.04	0.01	0.00	-0.04	0.01	0.00
$\Delta\varkappa_2$	-0.03	0.02	0.03	-0.03	0.02	0.04	-0.03	0.02	0.04
$\Delta\varkappa_3$	0.03	0.02	0.14	0.02	0.03	0.12	0.02	0.03	0.12

Appendix 6: *Point pair statistics of one-to-one band registration*

Algorithm	**Band Combination**	**Max**	**Min**	**Mean**	**Std**
SURF	NIR1 + NIR2	0.76	0.05	0.34	0.17
	RedEdge1+RedEdge2	0.91	0.03	0.38	0.20
	Red1 + Red2	0.55	0.04	0.26	0.12
	Green1 + Green2	0.85	0.01	0.33	0.17

Appendix 7: *Impact of Distance Threshold on Outlier Removal and Registration Accuracy*

Algorithm	**Band Combination**	**Inliers**	**Outliers**		**[X, Y]**
MaxDistance 0.5					
SURF	RedEdge(B2)+NIR(B1)	1878	3191	RMSE1	[0.2243, 0.2661]
	RedEdge(B2)+Red(B3)	139	429	RMSE2	[0.2290, 0.2304]
	RedEdge(B2)+Green(B4)	2157	2052	RMSE3	[0.2640, 0.2241]
KAZE (Thres=0.0012)	"	2661	4725	RMSE1	[0.2896, 0.2828]
	"	235	491	RMSE2	[0.3764, 0.2217]
	"	3262	2900	RMSE3	[0.2986, 0.3124]
MaxDistance 1					
Algorithm	Band Combination	Matchled (Inliers)	Outliers		
SURF	RedEdge(B2)+NIR(B1)	3996	1073	RMSE1	[0.3899, 0.4157]
	RedEdge(B2)+Red(B3)	315	253	RMSE2	[0.5743, 0.4372]
	RedEdge(B2)+Green(B4)	3650	559	RMSE3	[0.3718, 0.3686]
KAZE (Thres=0.0012)	"	5656	1730	RMSE1	[0.4210, 0.4339]
	"	577	149	RMSE2	[0.5227, 0.4359]
	"	5631	531	RMSE3	[0.3758, 0.3868]
MaxDistance 1.5					
Algorithm	Band Combination	Matchled (Inliers)	Outliers		
	RedEdge(B2)+NIR(B1)	4774	295	RMSE1	[0.4808, 0.5118]
SURF	RedEdge(B2)+Red(B3)	409	159	RMSE2	[0.6371, 0.4904]
	RedEdge(B2)+Green(B4)	4037	172	RMSE3	[0.4227,0.4423]
	"	7143	243	RMSE1	[0.4862, 0.5169]
KAZE (Thres=0.0012)	"	686	40	RMSE2	[0.6199, 0.5092]
	"	6077	85	RMSE3	[0.4144, 0.4304]
MaxDistance 2					
Algorithm	Band Combination	Matchled (Inliers)	Outliers		
	RedEdge(B2)+NIR(B1)	4900	169	RMSE1	[0.5123, 0.5393]
SURF	RedEdge(B2)+Red(B3)	420	148	RMSE2	[0.6626, 0.5202]
	RedEdge(B2)+Green(B4)	4081	128	RMSE3	[0.4402, 0.4546]
	"	7316	70	RMSE1	[0.5221, 0.5378]
KAZE (Thres=0.0012)	"	711	15	RMSE2	[0.6461, 0.5448]
	"	6138	24	RMSE3	[0.4357, 0.4443]

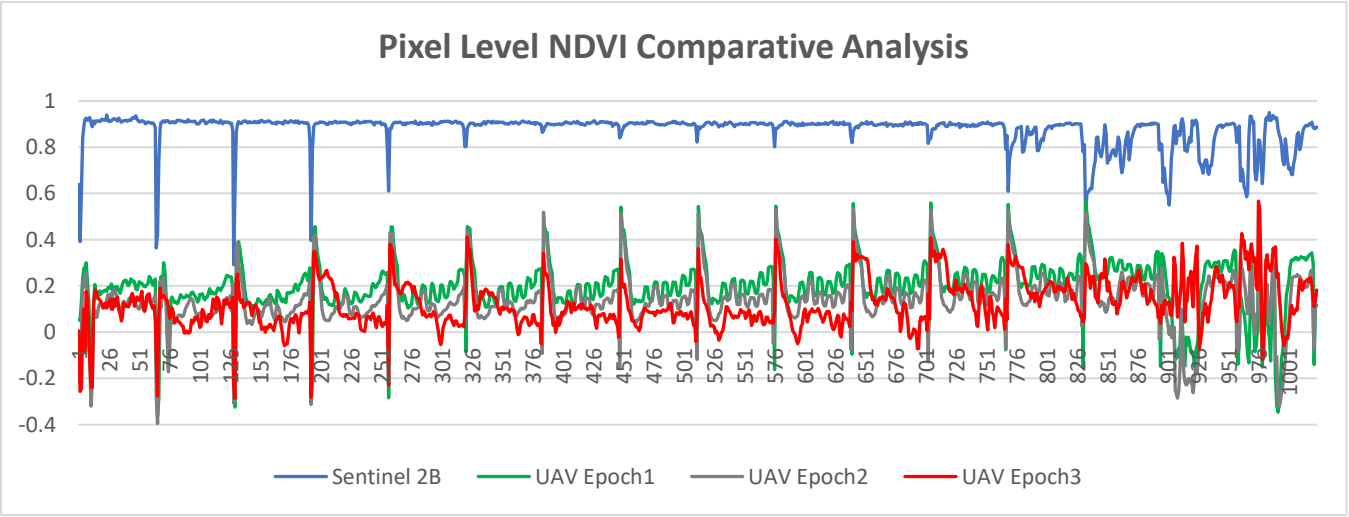
Appendix 8a: *Zonal Statistics for Intra-epoch NDVI analysis*

Epoch 1 (Flying Height = 70m)					Epoch 1 (Flying Height = 50m)			
Zone	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**
A	-0.08	0.70	0.48	0.10	-0.29	0.76	0.42	0.13
B	-0.68	0.57	0.14	0.15	-0.73	0.64	0.13	0.18
C	-0.35	0.57	0.19	0.10	-0.53	0.64	0.17	0.14
D	-0.29	0.54	0.22	0.09	-0.62	0.65	0.19	0.15
E	-0.27	0.52	0.22	0.09	-0.56	0.66	0.17	0.14
F	-0.29	0.57	0.20	0.11	-0.54	0.63	0.19	0.15
G	-0.28	0.61	0.24	0.10	-0.48	0.64	0.21	0.14
H	-0.08	0.61	0.39	0.07	-0.19	1.00	0.42	0.09
I	-1.00	1.00	-0.43	0.16	-1.00	1.00	-0.36	0.20
J	-0.27	0.26	-0.13	0.05	-0.43	0.15	-0.28	0.06
K	-0.28	0.08	-0.12	0.03	-0.48	0.06	-0.24	0.04
L	-0.28	0.09	-0.17	0.03	-0.54	0.23	-0.22	0.07
M	-0.54	0.03	-0.29	0.03	-0.60	0.61	-0.21	0.09

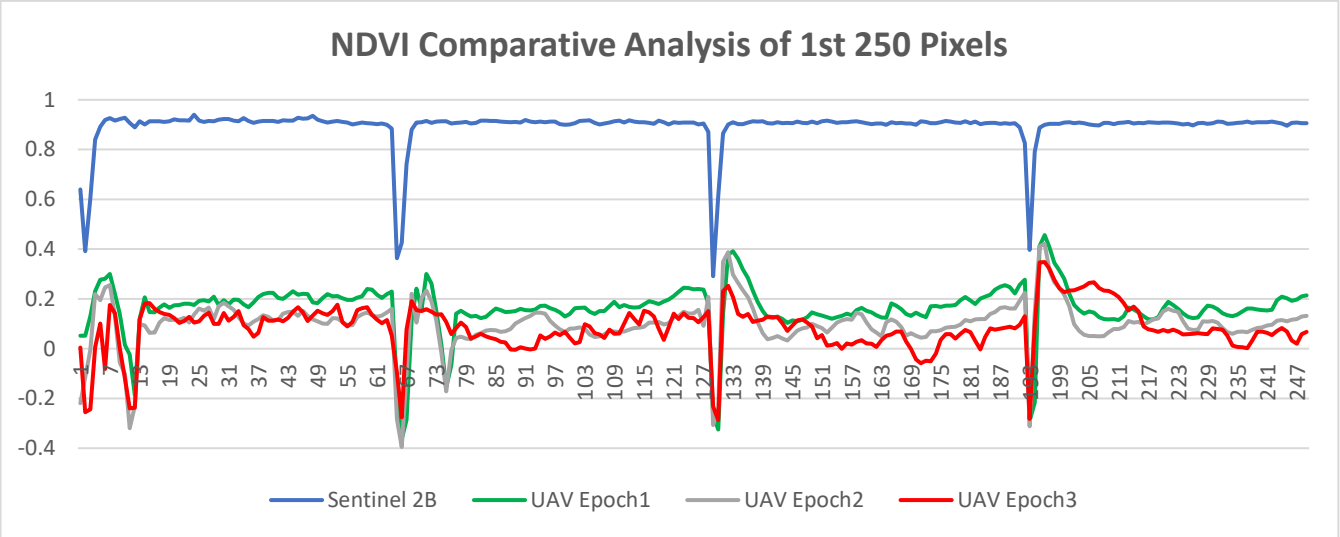
Appendix 8b: *Zonal statistics for Micro MCA Tetracam and Parrot Sequoia*

Zone	**Micro MCA Tetracam**				**Parrot Sequoia**			
	Min	**Max**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**
A	0.05	0.62	0.33	0.07	-0.25	0.65	0.36	0.17
B	-0.18	0.17	-0.02	0.04	-0.21	0.13	-0.13	0.03
C	-0.04	0.50	0.18	0.07	-0.15	0.61	0.24	0.15
D	-0.04	0.51	0.22	0.06	-0.23	0.63	0.36	0.12
E	-0.15	0.47	0.23	0.06	0.18	0.71	0.59	0.06
F	-0.28	0.18	-0.08	0.05	-0.53	0.13	-0.26	0.06
G	-0.15	0.39	0.16	0.06	-0.57	0.69	0.19	0.20
H	-0.30	0.38	0.14	0.06	-0.25	0.68	0.31	0.14
I	-0.08	0.51	0.18	0.06	0.20	0.71	0.55	0.07

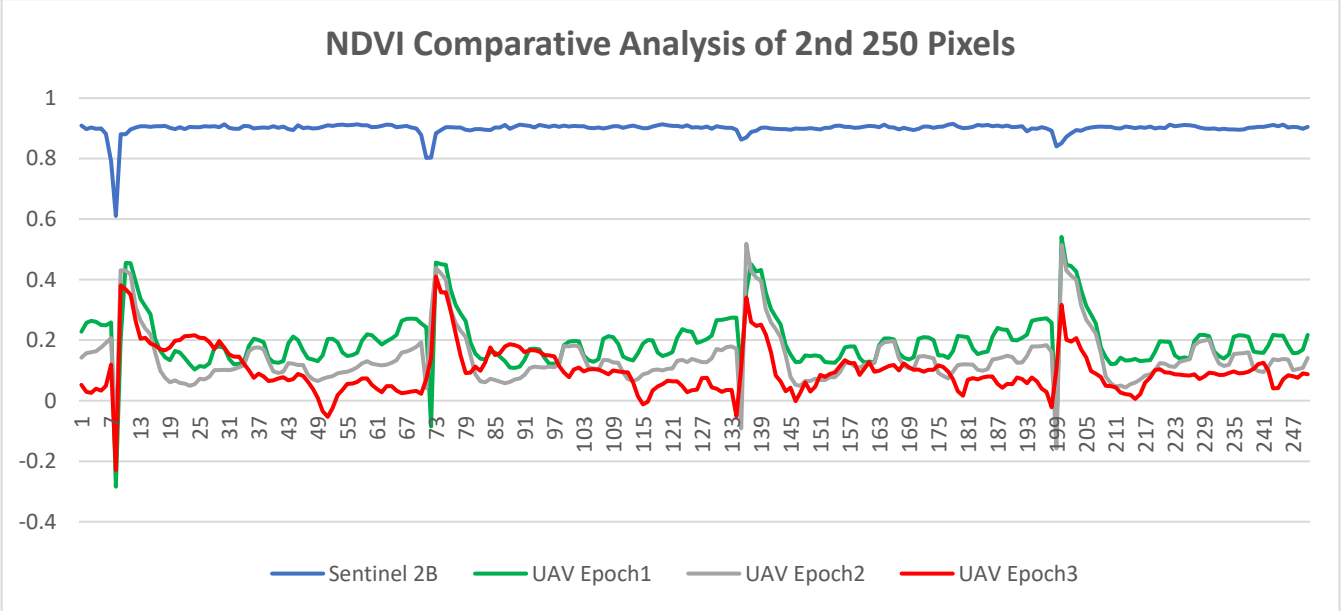
Appendix 9a: *Pixelwise NDVI Comparative Analysis*



Appendix 9b: *Pixelwise NDVI Comparative Analysis of the first set of 250 pixels*



Appendix 9c: *Pixelwise NDVI Comparative Analysis of the second set of 250 pixels*



Appendix 9d: *Pixelwise NDVI Comparative Analysis of 500 pixels (Upper Half)*