Hyperspectral Imaging from Ground Based Mobile Platforms and Applications in Precision Agriculture

Alexander Wendel

A thesis submitted in fulfillment
of the requirements of the degree of
Doctor of Philosophy

Australian Centre for Field Robotics
School of Aerospace, Mechanical and Mechatronic Engineering
The University of Sydney

Submitted August 2018; revised December 2018

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the University or other institute of higher learning, except where due acknowledgement has been made in the text.

I am the primary and corresponding author for all of the following papers, which form the core chapters of this thesis. As such, I drove the data gathering, method implementation, experimental design, data analysis and manuscript writing. I received guidance and supervision from the co-authors of the papers.

Chapter 3 of this thesis is published as Wendel, A. and Underwood, J. (2017b). Illumination compensation in ground based hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129:162–178

Chapter 4 of this thesis is published as Wendel, A. and Underwood, J. (2017a). Extrinsic parameter calibration for line scanning cameras on ground vehicles with navigation systems using a calibration pattern. *Sensors (Switzerland)*, 17(11)

Chapter 5 of this thesis is published as Wendel, A. and Underwood, J. (2016). Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5128–5135. IEEE

Chapter 6 of this thesis is published as Wendel, A., Underwood, J., and Walsh, K. (2018). Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. *Computers and Electronics in Agriculture*, 155:298–313

Alexander Wendel

20 December 2018

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

James Underwood

20 December 2018

Abstract

Hyperspectral imaging or imaging spectroscopy, which captures light in hundreds of narrow bands along the electromagnetic spectrum, has been studied extensively in the remote sensing community for applications ranging from agriculture to mineralogy. Hyperspectral imaging and spectroscopy have been shown to be useful in many agricultural applications, such as disease detection; species/cultivars classification; mapping of plant stress; and predicting specific plant or fruit properties. Advances in technology have brought down the cost, size and weight of hyperspectral cameras, allowing them to be used on smaller platforms and at lower altitudes, increasing spatial resolution and enabling more favourable viewing geometries. While there has been an increase in the number of studies using lower altitude unmanned aerial vehicles (UAVs), there is scant literature addressing the use of hyperspectral cameras from mobile ground based platforms, which have the advantage of being able to carry large payloads and perform long non-stop acquisition campaigns. However, mobile ground-based deployment introduces unique challenges due to lighting and the close viewing proximity.

This thesis focuses on the use of line scanning hyperspectral sensors on mobile ground based platforms and applying them to agricultural applications. While frame hyperspectral imagers are becoming more prevalent, line scan configurations that acquire a single line of pixels per frame are more common, due to their capacity for greater spectral resolution. The first part of the thesis deals with the geometric and spectral calibration and correction of acquired hyperspectral data. These contributions are broadly relevant for all applications of line scanning hyperspectral cameras on ground based mobile platforms. The second part of the thesis applies data captured from the platform to two different agricultural applications: Weed discrimination in a row cropping environment, and mango maturity estimation in orchards.

When operating at low altitudes, changing lighting conditions are common and inevitable. Changes occur even within a short period of time, stemming from unpredictable cloud movements; variations in the colour of sunlight throughout the day; shadows; and reflected or scattered light from surrounding surfaces. This complicates the retrieval of a surface's reflectance, which is solely a function of its physical structure and chemical composition. In spectroscopy, it is desirable to obtain reflectance or other measures like absorbance or transmittance, as they are illumination invari-

ant and therefore more robust in subsequent analysis and prediction tasks. The first contribution of this thesis is the evaluation of an approach to compensate for changes in illumination and obtain reflectance that is less labour intensive than traditional empirical methods. Convenient field protocols are produced that only require a representative set of illumination and reflectance spectral samples. The experimental results confirm that the approach competes well with traditional methods, but with considerably reduced effort by operators in the field.

The second contribution is a method to determine a line scanning camera's rigid 6 degrees of freedom (DOF) offset with respect to a navigation system, enabling accurate mapping of hyperspectral pixels in world coordinates (georegistration) and sensor fusion. The approach requires the scanning of several stationary points, whose position do not need to be known *a priori*, from a number of platform positions and orientations. Uncertainties provided by the navigation system are taken into account in order to compute a likelihood that is maximised to obtain the position and orientation of the camera. In addition, a Markov Chain Monte Carlo (MCMC) algorithm is used to sample the solution space in order to obtain the uncertainty of the optimal offset values. The results demonstrate the effectiveness of the approach with a number of novel visualisations.

The thesis then explores two applications of hyperspectral sensing for agriculture. The first is a self-supervised weed detection framework that allows training of a per-pixel classifier using hyperspectral data without manual labelling. The approach works by making assumptions about the co-linear seeding geometry of commercial row-crops. This enables automatic extraction of training data in order to train a classifier. When trained on the automatically extracted data, the tested classifiers approached the performance of hand labelled training data.

Lastly, this thesis demonstrates the mapping of mango maturity using hyperspectral data on an orchard wide scale using efficient image scanning techniques, which is a world first result. The camera scanned trees in a sideways push-broom configuration, which was challenging because of the complicated tree canopy geometry, which introduced significant lighting variation due to shadows and reflected/scattered light. A novel classification, regression and mapping pipeline is proposed to generate per tree mango maturity averages. The results confirm that maturity prediction in mango orchards is possible in natural daylight using a hyperspectral camera, despite complex micro-illumination-climates under the canopy.

The findings of the research presented in this thesis provide tools for the effective use of line scanning hyperspectral sensors on ground based mobile platforms, which are applicable in a number of application domains. Additionally, two distinct precision agriculture use-cases were developed and evaluated.

Acknowledgements

I would like to extend my heartfelt gratitude to all the people that have made this thesis possible, both directly and indirectly.

First of all I would like to thank my supervisor, James Underwood, for his constant support, guidance and patience over the past four years. I could not have asked for a better advisor and this thesis would not have been possible without his mentorship and the positive environment he created. Thank you for teaching me to be a better researcher, communicator and writer.

I would also like to thank all the people at the ACFR who have helped me with their advice and feedback throughout the years. Whether it be at reading groups, helping with specific thesis related questions or general conversations, I tremendously appreciate everyone sharing their thoughts in their specific areas of expertise. I would like to particularly thank Suchet and John for giving me a gentle introduction to postgraduate research life, for their technical expertise, and for our many thoughtful and interesting discussions.

I would like to thank my friends for supporting and putting up with me throughout my postgraduate studies.

I would like to thank my family, specifically my parents and sister. Without their constant love and support throughout my life and studies this would not have been possible.

Last but not least, I would like to thank Kelly for being there for me every single day. You have enriched my life immeasurably over the last four years!

Contents

Nomenclature

List of Acronyms

ACFR	Australian Centre for Field Robotics
AGV	autonomous ground vehicle
AIS	Airborne Imaging Spectrometer
ANN	artificial neural network
AOTF	acousto-optical tunable filter
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
CARI	chlorophyll absorption ratio index
CASI	Compact Airborne Spectrographic Imager
CCD	charge-coupled device
CHRIS	Compact High Resolution Imaging Spectrometer
CMOS	complementary metal–oxide–semiconductor
CNN	convolutional neural network
CRF	conditional random field
CRISM	Compact Reconnaissance Imaging Spectrometer for Mars
CWT	continuous wavelet transform
DAIS	Digital Airborne Imaging Spectrometer
DEM	digital elevation model
DM	dry matter
DOF	degrees of freedom
DSM	digital surface model
DWT	discrete wavelet transform
EMSC	extended multiplicative scatter correction
EnMAP	Environmental Mapping and Analysis Program
EO-1	Earth Observing One
ESA	European Space Agency
FPI	Fabry-Pérot Inferometer
GA	genetic algorithm
GCP	ground control point
GNSS	global navigation satellite system
GPS	global positioning system
GPS-INS	global positioning system and inertial navigation system

GPU	graphics processing unit
HISUI	Hyperspectral Imager Suite
HSI	hyperspectral imaging
HYDICE	Hyperspectral Digital Imagery Collection Experiment
HyMap	Hyperspectral Mapper
HypXIM	Hyperspectral X Imagery
HyspIRI	Hyperspectral Infrared Imager
IAR	internal average reflectance
ICA	independent component analysis
IFOV	instantaneous field of view
INS	inertial navigation system
IPAR	intercepted photosynthetically active radiation
KNN	k nearest neighbour
LAI	leaf area index
LCTF	liquid crystal tunable filter
LDA	linear discriminant analysis
LDR	linear dimensionality reduction
LIDAR	light detection and ranging
LLE	locally linear embedding
LMM	linear mixing model
MAFE	multidimensional artificial field embedding
MAFE	multidimensional artificial field embedding
MCMC	Markov Chain Monte Carlo
MDS	multidimensional scaling
ML	maximum likelihood
MLR	multiple linear regression
MRF	markov random field
MSC	multiplicative scatter correction
NASA	National Aeronautics and Space Administration
NDVI	normalised difference vegetation index
NDWI	normalised difference water index
NIPALS	nonlinear iterative partial least squares
NIR	near infra-red
NLDR	non-linear dimensionality reduction
NW	Norris-Williams
OMEGA	Observatoire pour la Mineralogie, l'Eau, les Glaces, et I'Activite
PA	precision agriculture
PCA	principal component analysis
PCR	principal component regression
PGP	prism-grating-prism
PLS	partial least squares
PRI	photochemical reflectance index
PRISMA	Precursore IperSpettrale della Missione Applicativa

PROBA-1	Project for On-Board Autonomy
PSRI	plant senescence reflectance index
RANSAC	random sample consensus
RBF	radial basis function
RTK	real time kinematic
RTK-GNSS	real time kinematic global navigation satellite system
RTK-GPS-INS	real time kinematic global positioning system and inertial navigation system
RVI	ratio vegetation index
SAM	spectral angle mapper
SAVI	soil adjusted vegetation index
SfM	structure from motion
SFS	stepwise forward selection
SHALOM	Spaceborne Hyperspectral Applicative Land and Ocean Mission
SNR	signal to noise ratio
SNV	standard normal variate
SPCA	supervised principal component analysis
SRDA	spectrally resolving detector array
SSC	soluble solids contents
SSNE	spherical stochastic neighbor embedding
SSWM	site specific weed management
SVM	support vector machine
SWIR	short-wave infrared
UAV	unmanned aerial vehicle
VNIR	visible and near infrared

Chapter 1

Introduction

For several decades, the remote sensing community and other disciplines have leveraged the detailed spectral information provided by hyperspectral imaging (HSI) for far ranging applications including mineralogy, medicine, agriculture and many others. Satellite and aerial sensing platforms are commonly used, but the increased viewing flexibility and spatial resolution gained by imaging on the ground and much closer to the objects of interest opens up previously unattainable possibilities. For instance, vertical surfaces such as trees in orchards or pit faces in mines can be viewed from the side, and fine details like individual leaves or fruit can be resolved. Despite the advantages, comparatively little research has been done on the use of hyperspectral sensing from ground-based platforms. This knowledge gap motivates the objectives of this thesis, which are to provide tools that enhance the ability to acquire hyperspectral data from ground based mobile platforms, and to examine applications in precision agriculture.

The aims of this thesis can be divided into two main themes. The first theme is to provide convenient techniques that aid in the acquisition of hyperspectral data from ground based mobile vehicles. Specifically, frameworks are examined to deal with changes in illumination, and determining the 6 degrees of freedom (DOF) pose of a line scanning camera, a common type of hyperspectral sensor. The second objective is to apply the insights gained to precision agriculture (PA) applications, an area which

stands to gain great benefit from spectral imaging on the ground. This includes a weed detection framework that does not rely on manually generated training datasets, significantly reducing the labour component in the process. Additionally, in a world first, a novel system for predicting and mapping fruit maturity of mangoes on an orchard-wide scale is introduced.

1.1 Background and Motivation

Hyperspectral cameras are a promising technology that can be viewed as a fusion of camera and spectrometer, as they produce images like a camera, but each pixel in that image contains high resolution colour or spectral information as would be provided by a spectrometer. They are therefore often also referred to as imaging spectrometers. They have proven their value in remote sensing applications, and the prevalent use of spectroscopy in a number of different areas further supports its utility. In many applications both very high spectral and spatial resolution are advantageous, which is where aerial and satellite remote sensing HSI is often inadequate. For example, resolving individual fruit on a tree would be impossible from high altitudes with current technology, due to the lack of spatial camera resolution, and the top-down perspective does not allow fruit on the sides of the trees to be imaged. Using hyperspectral imagers from ground based mobile platforms combines high spatial resolution scanning and flexible viewing geometries with the potential to cover large areas. Hyperspectral sensors have only recently been utilised this way, and therefore literature describing how to use them effectively and demonstrating their potential has been limited.

Compared to aerial platforms, the calibration and correction of HSI data requires different approaches when using ground vehicles. For instance, an inevitable issue encountered when performing HSI at ground level is substantial and often sudden lighting changes, primarily due to cloud movement, shadows and scattering from surrounding surfaces. This is a problem because changes in lighting can significantly affect spectral data and therefore impact performance of subsequent processing. Furthermore, in practical applications it is often necessary to know the position of each

pixel in world coordinates, for example to locate and spray weeds once identified. Determining a pixel's world coordinates is also known as georeferencing, and to do it effectively, a camera's pose must be determined accurately. One of the most common types of hyperspectral imagers are line scanners, which capture a single line of pixels per exposure. Determining the pose for a line scanning camera aboard a mobile ground vehicle imposes additional constraints such as movement restrictions and requiring very high positional accuracy for reference or tie points. The knowledge gap in these areas acts as a barrier to further study of ground based hyperspectral imaging in specific application domains that stand to benefit greatly from its advantages.

One particular area where ground based hyperspectral imaging has the potential to provide a significant contribution is agriculture. Like any business, agricultural operations are constantly attempting to optimise the efficiency of their operations. By increasing yield and decreasing the amount of land, herbicide, pesticide, fertiliser and other inputs used, agribusinesses can maximise their profits. Additionally, because agriculture plays such an integral role for human civilisation, enhanced productivity provides benefits beyond mere business objectives. An expanding population drives an increasing demand for food with limited arable land. PA is one concept that has been proposed to address this problem. In short, its purpose is to measure intra-field variability in order to manage farm operations in a more granular way [157]. It recognises the fact that crops are often affected by factors such as sunlight, soil, nutrients, water, weeds and diseases in a spatially heterogeneous way. For instance, by determining where exactly weed plants are located in the field, they can be targeted directly instead of broadcast spraying, greatly reducing the amount of herbicide used. The sensing aspect of this process is critical, and its precision directly impacts the effectiveness of subsequent actions.

Various types of sensors have been employed for the numerous applications of PA using a number of different platforms. Satellite and high altitude aerial sensing provide great coverage, but are limited in terms of precision due to their distance from the target crop fields. This is where low altitude platforms such as some unmanned aerial vehicles (UAVs) and ground based vehicles can provide a particular advantage,

providing greater resolution in most cases. The types of sensors used in PA applications include cameras, light detection and ranging (LIDAR), spectrometers, soil sensors and others. The "camera" category is particularly broad and encompasses monochrome, RGB, multispectral, thermal, depth, stereo and hyperspectral cameras. The large amount of spectral information in each pixel of a hyperspectral image provides the potential to detect properties that would be difficult to discern with more conventional sensing methods. For example, HSI has been shown to be able to distinguish between plant species on a per pixel basis [207, 264] or detect plant diseases [193, 219]. It can therefore play an important role in PA.

1.2 Thesis Objectives

The purpose of this thesis is to provide tools that specifically address the challenges of correcting for varying illumination and spatial calibration when acquiring HSI data from ground based mobile platforms, and to demonstrate applications in PA.

To counter the effects of variable illumination when operating under cloud cover, one aim of this thesis is to develop a method that effectively compensates for such illumination changes while being convenient in a field setting. The solution should limit or eliminate in-field labour, be robust to rapid and significant changes in lighting and minimise the complexity of the overall system.

To enable accurate georeferencing, another objective of this thesis is to accurately determine a line scanning camera's 6 DOF pose, i.e. position and orientation, with respect to the vehicle it is mounted to. The approach should also be convenient in a field setting, and therefore should limit the area required for the acquisition of calibration data. Any reference patterns should be simple to produce or print, and surveyed reference points (i.e. ground control points) should be avoided as they would need to be extremely accurate for ground based sensors, which can be prohibitive in practice. Also, to reduce the complexity of the system and maximise the generalisability of the calibration approach, auxiliary sensors or cameras should not be required.

Lastly, this thesis applies ground based HSI to two distinct PA problems. The first application focused objective is to detect weeds in corn crop rows using hyperspectral imaging. The aim is to perform this task without the use of manually labelled data, which can be laborious to obtain. The second PA problem is to non-destructively predict the maturity of mango fruit on an orchard scale using ground based HSI. Obtaining and processing hyperspectral data in fruit orchards is challenging due to the complex canopy geometry, which can alter the light seen by the camera in many ways. The objective is to overcome this and estimate an indicator metric for the maturity of mangoes before mapping the predictions to allow growers to more precisely manage harvesting operations.

1.3 Principal Contributions

The contributions of this thesis are tools to spectrally correct and spatially calibrate line scan hyperspectral data on ground based mobile platforms, and applications of these tools in precision agriculture. This includes illumination compensation methods with convenient field protocols suitable for ground based applications, and a calibration approach that determines a line scan camera's pose relative to the platform. The tools are applied to two precision agriculture applications, including a self-supervised weed detection system for row crops, and a world first orchard-scale mango maturity mapping system. This thesis examines two very different push broom sensing geometries, one suitable for row crops by pushing the scan line along the ground in front of the platform, and the other is designed for tree crops by sweeping the scan line across the trees in a sideways facing configuration. The specific contributions from each component of the thesis are itemised below:

Illumination compensation

Published journal paper: Wendel, A. and Underwood, J. (2017b). Illumination compensation in ground based hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129:162–178

- Evaluating the suitability of a previously developed logarithm subspace method for illumination and reflectance extraction [56] for use on large, high spatial and spectral resolution agriculture field datasets.
- The development of multiple field protocols for acquiring training data for the illumination compensation method by Drew and Finlayson [56]. These present different trade-offs between the accuracy of illumination compensation and the logistical complexity of the field work.

Geometric calibration

Published journal paper: Wendel, A. and Underwood, J. (2017a). Extrinsic parameter calibration for line scanning cameras on ground vehicles with navigation systems using a calibration pattern. *Sensors (Switzerland)*, 17(11)

- An approach for estimating a line scanning camera's 6 DOF relative pose with respect to the platform. The method is attuned to the constraints imposed by scanning from a ground based vehicle.
- An approach for estimating the uncertainty of the pose solution using Markov Chain Monte Carlo (MCMC), and associated visualisation methods.
- Open source code available (https://github.com/acfr/calibrate-line-camera).

Weed detection

Published conference paper: Wendel, A. and Underwood, J. (2016). Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5128–5135. IEEE

- A self-supervised framework that uses *a priori* knowledge of commercial farm seeding geometry to automatically generate training sets for per-pixel crop/weed discrimination using HSI data.

Mango maturity mapping

Published journal paper: Wendel, A., Underwood, J., and Walsh, K. (2018). Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. *Computers and Electronics in Agriculture*, 155:298–313

- An analysis of the illumination related challenges encountered when acquiring HSI in fruit orchards due to the complex canopy geometry.
- A world-first end-to-end pipeline for detecting mangoes in an orchard, estimating dry matter (DM) (a commonly used fruit maturity indicator) and georeferencing the predictions to produce orchard-scale maturity maps.

Contribution statement

I am the primary author for all papers that form the core chapters of this thesis (Chapters 3 through 6). As such, I drove the data gathering, method implementation, experimental design, data analysis and manuscript writing. I received guidance and supervision from the co-authors of the papers, particularly from my thesis supervisor James Underwood, who was involved in the development all papers.

1.4 Thesis Structure

Chapter 2 provides an overview of the literature as it pertains to HSI from ground based platforms. This includes an introduction to HSI in general and its applications, followed by the steps required to calibrate, correct, process and georeference the data.

The next two chapters cover the use of hyperspectral imaging on a ground based vehicle, which presents unique challenges compared to the more common satellite and aerial platforms. The proposed techniques are generalisable to any application where a hyperspectral sensor is being used from a ground based mobile platform. Chapter 3 addresses the problem of variable illumination when acquiring hyperspectral data near the ground. It presents an illumination compensation method that considerably

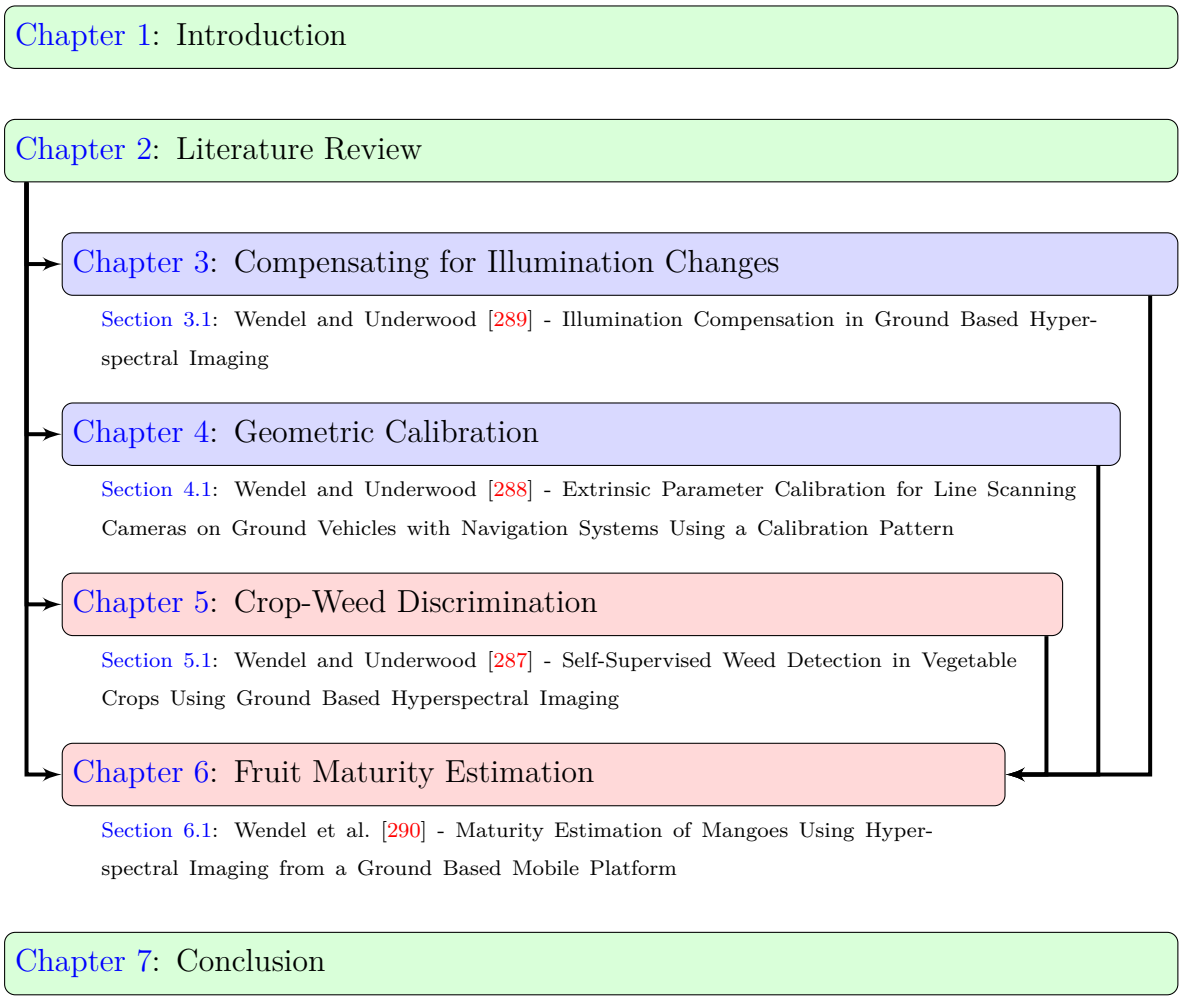


Figure 1.1 – Thesis structure summary. The literature review exposes a number of gaps that motivate the the four core chapters. Chapters 3 and 4 provide solutions for the effective use of HSI on ground based mobile platforms and Chapter 5 introduces a self-supervised crop/weed discrimination framework. The findings and lessons learned from Chapters 3, 4 and 5 feed into Chapter 6 which introduces a pipeline for mango maturity estimation on an orchard-scale.

reduces the amount of manual work required in the field compared to more traditional empirical methods. Chapter 4 develops a novel method for determining a line scanning camera's pose with respect to the platform, which is required for accurate georeferencing. The contribution in this chapter is generalisable to any line scanning camera, not just hyperspectral sensors.

Chapters 5 and 6 then apply HSI from ground based mobile platforms to two distinct

PA applications. Chapter 5 introduces a novel method to perform weed detection using HSI in commercial crop rows where the seeding geometry is known *a priori*. The proposed approach automatically generates training datasets to allow crop/weed discrimination.

Relying heavily upon the lessons learned from previous chapters, Chapter 6 addresses the topic fruit maturity estimation. An end-to-end pipeline is introduced which detects mangoes, predicts their maturity, and georeferences the data to produce orchard-scale maps.

The contributions in Chapters 3 through 6 enable researchers and operators of hyperspectral imagers on ground based mobile platforms to acquire data that is spectrally and spatially useful and accurate. The findings of this thesis demonstrate how such data can be utilised in agricultural applications to allow for more precise farm management and consequently greater yields. The general ideas generated by the work in this thesis push forward the state of the art of HSI on ground based mobile platforms, also enabling its use in domains other than agriculture.

Chapter 2

Literature Review

Examining the reflected, transmitted or absorbed light of plants in order to infer their properties has been extensively studied in both the remote sensing and plant science literature. Light spanning the electromagnetic spectrum beyond the visible, has been leveraged to estimate, predict and study the health, status, coverage, biomass and other attributes of plants or areas of vegetation. Traditionally, the examination of the light signatures, i.e. spectra, after interacting with objects was performed using spectroscopy, which yields a single high resolution spectrum per measurement. Hyperspectral imaging (HSI) or imaging spectroscopy is an extension to this by acquiring individual high resolution spectra for each spatial pixel in an image frame. It has gained popularity since its first uses in the 1980s, due to the amount of information obtained per frame and technological progress allowing ever higher resolutions and increasingly advanced processing techniques.

The use of HSI in precision agriculture (PA) applications has received significant attention in the remote sensing literature. This primarily includes data from high altitude platforms, both satellite and aerial, which allow broad coverage of large areas within a short period of time. Due to the distance between the sensing platforms and farms, these methods generally do not tend to directly interfere with farming operations. Nevertheless, spatial resolution is still limited, and obtaining hyperspectral data at the scale of individual leaves or fruit is not feasible yet from high altitude.

In addition, the geometry with which these platforms can view agricultural fields is restricted to an approximately top down (nadir) perspective, which is suitable for broadacre or row cropping environments, but particularly restrictive for taller crops such as orchard trees.

Low altitude and ground based HSI address some or all of these issues. Significantly higher spatial resolution is gained by reducing the distance between sensors and surfaces of interest, and some low altitude platforms such as multi-rotor helicopters and ground based platforms are capable of offering considerably more flexibility in viewing geometry. For example, when imaging tree crops these platforms can move in between rows of trees allowing sideways acquisition of canopies and the fruit contained within. However, as sensor and platform technology has only recently advanced to the point where data acquisition in such a manner is practically feasible, research addressing ground and near-ground based platforms with HSI is scarce. Most of the literature in this area covers the use of hyperspectral cameras on low altitude unmanned aerial vehicles (UAVs), and studies that specifically investigate the use of hyperspectral cameras on ground based platforms are far less common.

Despite the scarcity of literature on mobile ground based HSI, the topic of how to process hyperspectral data in order to produce meaningful results is an extensively investigated field of research, which includes work not only from the remote sensing community, but also from fields such as plant science, particularly in spectroscopy. An important step when dealing with hyperspectral data is to compensate for varying lighting conditions, in order obtain spectra that are solely a product of the properties of the surface being imaged and not the time varying viewing conditions. Following this, there are number of preprocessing steps that can be taken to maximise the ability of machine learning algorithms to make predictions, such as estimating the concentration of chemical components in vegetation or distinguishing weed plants from crops. Simple linear models to multi-layer deep learning networks have been employed to obtain predictions, and can produce either discrete (classification) or continuous results (regression). For spectral data, there are also some specifically designed techniques, such as vegetation indices and spectral unmixing.

It is often desired to accurately locate hyperspectral pixels or spectra in the world frame, for example to spatially map chemical concentrations or weed density across a field. Usually this process, known as georeferencing, requires both precise and accurate knowledge of a camera's pose. If a platform can be located in world coordinates, for example using an accurate navigation system, then the fixed offset between the platform and the hyperspectral sensor needs to be determined. However, there is very little literature addressing this problem for proximal hyperspectral sensors, which is particularly challenging when using line scanning cameras. Line scanning cameras (also 1D or linear cameras) are a common type hyperspectral camera, which capture a single line of spatial pixels per exposure.

This chapter provides an overview of the literature relating to the use of hyperspectral cameras on ground based mobile platforms in an agricultural context, with a particular focus on line scanning cameras. Section 2.1 provides an introduction to HSI, outlining the different types of sensors, lighting options, platforms and applications. In Section 2.2, common HSI and spectroscopic processing techniques are summarised. Finally, Section 2.3 reviews georeferencing approaches that have been explored in previous studies, and geometric calibration of hyperspectral cameras, which is necessary for some of these georeferencing methods.

2.1 Introduction to Hyperspectral Sensing

Hyperspectral imagers produce an image where each pixel contains the spectral signature of the corresponding spatial location in the world in hundreds of contiguous narrow spectral bands. Therefore, each hyperspectral image has as many channels as there are bands, in contrast to grayscale or RGB images which only have one or three channels respectively. Because of this, hyperspectral data are often referred to as hypercubes, with two spatial dimensions and one spectral dimension. Figure 2.1 illustrates this concept. Hyperspectral cameras are often compared to multispectral sensors. The main differences between these types of sensors is that multispectral cameras capture far fewer and broader bands, which, in contrast to hyperspectral

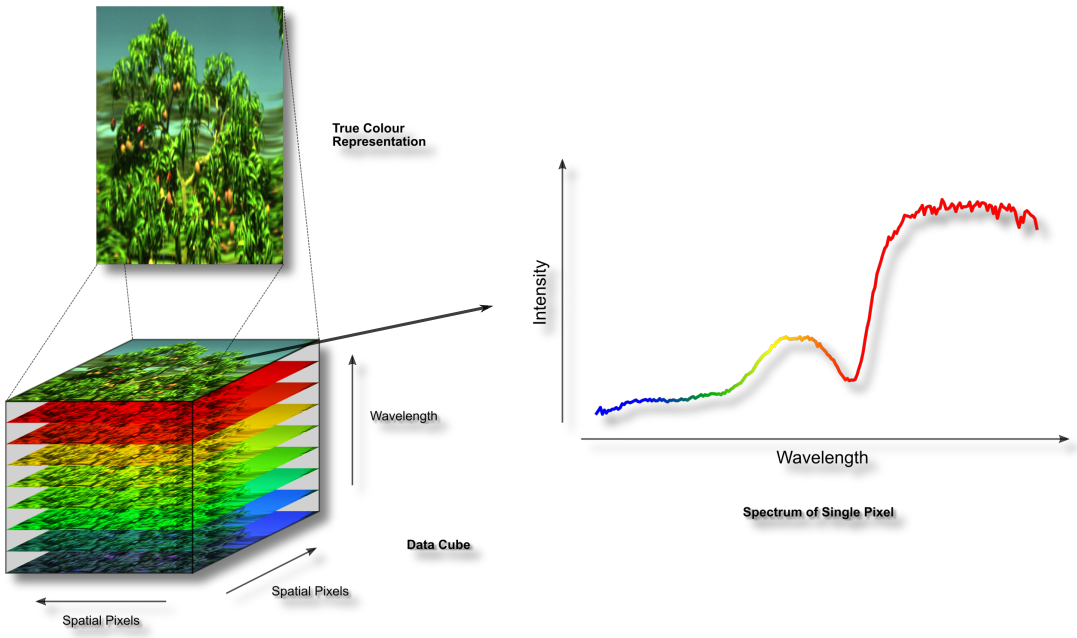


Figure 2.1 – Illustration of a hyperspectral data cube or "hypercube". Each hyperspectral image can be thought of as a cube of data, with two spatial dimensions and one spectral dimension. Each spatial pixel in the image therefore contains a full spectrum of all available bands.

data, are not necessarily contiguous [272, 88].

There are a number of different sensors available that can produce hyperspectral imagery, which will be briefly introduced in this section, as well as some of the lighting options which can be artificial or natural. The applications where hyperspectral imaging has been applied are as broad as the range of platforms, and these will also be summarised in this part of the literature review.

2.1.1 Types of Hyperspectral Cameras

An intuitive type of hyperspectral sensor are tunable filters, which capture each band one image at a time. A simple version of this is the filter wheel, where band filters are physically moved in front of the sensor or camera [69]. As the rotational speed of such a filter is limited, electronically tunable filters, such as acousto-optical tunable filter (AOTF), liquid crystal tunable filter (LCTF) and Fabry-Pérot Inferom-

eter (FPI), have been developed [154, 146, 94, 209]. All these tunable filter cameras capture a single wavelength at a time, which means that bands are not synchronised with each other and will not match unless both the camera and scene are completely stationary. There is a trade-off between the number of available bands and the time required for a full cycle of all bands. LCTF based hyperspectral cameras are limited by their wavelength switching time [239], and are therefore less suitable for mobile applications where the camera is continually moving. Both AOTF and LCTF technologies require polarised light, reducing optical throughput for randomly polarised signals, such as fluorescence and diffuse reflectance [3, 239]. Recently, light-weight FPI based hyperspectral imagers have found increased use in small single- or multi-rotor helicopter UAV applications [120, 119, 275].

One of the most common types of hyperspectral sensors are line scanning cameras, which capture a single line of pixels at a time. Here, a lens focuses light onto an entrance slit. The light that passes through this slit is spread over the desired spectral range onto a frame sensor, such as a charge-coupled device (CCD), using a diffraction grating or prism-grating-prism (PGP) element [281]. Each frame obtained from a line scanning hyperspectral camera is therefore two dimensional, with one spectral and one spatial dimension. Full data cubes of a scene can be obtained by scanning an area in a "push broom" manner [320]. All bands in line scan HSI data are synchronised, but consecutive line scans are not, which can cause spatial warping of the resulting image. While there are several techniques to compensate for non-linear camera motion, imaged surfaces must remain stationary during acquisition. Line scanning hyperspectral imagers are economical, readily available off-the-shelf and have been used in a large variety of applications.

Recently, snapshot HSI cameras which capture a full 2D hyperspectral image frame in one exposure have been developed [146, 114, 99]. These sensors overcome the timing issues faced by both line scanning and tunable filter sensors. They have narrow band filters deposited as filter arrays directly on the sensor, also known as spectrally resolving detector arrays (SRDAs), extending the traditional Bayer colour filter, which is pervasive in RGB imaging. FPI filters integrated over complementary

metal–oxide–semiconductor (CMOS) image sensors are one technology used in this application [77, 7]. The trade-off inherent to SRDAs cameras is that a greater number of bands requires more pixels to deposit filters on, which in turn reduces spatial resolution. The type of imager must therefore be carefully chosen for a particular application.

2.1.2 Illumination

Appropriate lighting is essential in hyperspectral imaging to maximise signal to noise ratio (SNR), consistency and uniformity in the data. When acquiring data close to the surface of interest, it is possible to use artificial means for lighting the scene. Additionally, when operating indoors, such as in lab or factory settings, artificial lighting is unavoidable. Compared to variable natural illumination, controlled artificial light sources can considerably enhance performance [26, 320]. Some systems therefore opt for controlled lighting solutions in field applications. For instance, Slaughter et al. [253] and Zhang et al. [321] used tungsten-halogen based illumination with a blue filter to provide a more uniform SNR across the spectrum, as the bulbs' light was biased toward the red and near infra-red (NIR). Busemeyer et al. [29] also used a halogen lighting system to perform hyperspectral image acquisition. A particular disadvantage of artificial lighting is that outdoors, data must either be acquired at night or underneath a shroud to exclude competing natural light, which can be cumbersome or in some cases infeasible.

In most traditional hyperspectral imaging applications that use aerial and satellite platforms the only option is to rely on natural lighting. There are also several examples where hyperspectral data were used with natural illumination in a ground based context [252, 48]. The advantages of using natural instead of artificial illumination include requiring fewer hardware components and reduced power requirements. One significant disadvantage is that natural light can change considerably throughout the day and even within short periods of time, depending on sun position, atmospheric composition and cloud movements, which change the signature and relative contribu-

tions of direct sunlight and skylight (i.e. sunlight scattered by the atmosphere) [112]. Such changes can affect prediction performance [284] if not properly compensated for, and consequently lead to additional complexity in data preprocessing. Natural illumination also exhibits absorption bands at several wavelengths, where light is significantly reduced. Consequently, the resulting SNR in those narrow regions of the spectrum is severely affected.

2.1.3 Platforms

Since the first uses of HSI on higher altitude manned aircraft, imaging spectroscopy has been applied in various settings. In the early 2000s, the first space based satellite platforms carrying hyperspectral imagers were launched. Recently there has been an increase in the use of smaller and lower altitude UAVs to capture hyperspectral data, as well as ground based mobile platforms, both manned and unmanned.

Satellite and High Altitude Aerial

Hyperspectral imagery obtained from satellite or high altitude aerial platforms is common in the remote sensing literature. One of the first aerial hyperspectral imagers, the Airborne Imaging Spectrometer (AIS), was developed in the early 1980s, which was able to capture data in the short-wave infrared (SWIR) (1200-2400 nm) with up to 128 bands, and a spatial pixel resolution of 32, translating to an instantaneous field of view (IFOV) of 10 m on the ground [88]. It was mainly used for mineralogy applications and acquired the first hyperspectral imagery that showed unambiguous evidence of mineral identification [88]. In the mid to late 1980s, Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) was developed by National Aeronautics and Space Administration (NASA), which flew at altitudes of tens of kilometres, and provided 224 bands from 400 to 2400 nm. Other hyperspectral imagers designed to be carried aboard aircraft include [88] Digital Airborne Imaging Spectrometer (DAIS) [228], Compact Airborne Spectrographic Imager (CASI) [50], Hyperspectral Digital

Imagery Collection Experiment (HYDICE) [18] and Hyperspectral Mapper (HyMap) [38].

The Hyperion Imaging Spectrometer is the first high resolution hyperspectral imager to orbit the earth [272] and was launched by NASA in 2000 aboard the Earth Observing One (EO-1) satellite mission, offering 220 bands from 400 to 2500 nm and a spatial resolution of 30 m on the ground [278]. A year later, European Space Agency (ESA) launched the Project for On-Board Autonomy (PROBA-1) satellite, which carried the Compact High Resolution Imaging Spectrometer (CHRIS), which acquires hyperspectral data with a ground spatial resolution of 17-40 m and provides up to 62 spectral channels from 400 to 1050 nm [17]. The TianGong-1 is another example of a space based hyperspectral imager, providing 128 bands from 400 to 2500 nm and 10-20 m spatial resolution [153]. A number of spaceborne hyperspectral imagers are planned to be launched in the near future, including Precursore IperSpettrale della Missione Applicativa (PRISMA), Environmental Mapping and Analysis Program (EnMAP) hyperspectral imager, Hyperspectral Imager Suite (HISUI), Spaceborne Hyperspectral Applicative Land and Ocean Mission (SHALOM), Hyperspectral Infrared Imager (HyspIRI) and Hyperspectral X Imagery (HypXIM) [272]. Recently, the Aalto-1 nano-satellite was launched, carrying an FPI based hyperspectral camera [216].

In addition to the above, there are also two hyperspectral cameras orbiting the planet Mars [52], the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) on board the Mars Reconnaissance Orbiter [196] and Observatoire pour la Mineralogie, l'Eau, les Glaces, et I'Activite (OMEGA) on board the Mars Express mission [24].

Low Altitude Unmanned Aerial Vehicles

Two oft-cited downsides of higher altitude aerial and satellite platforms are the high cost and limited resolution [5]. As a result of technological advances that the reduced the size and weight of sensors, UAV based hyperspectral imaging on smaller and lower altitude platforms has become popular [327]. Dehaan et al. [49] investigated the use of a Headwall Photonics Hyperspec VNIR line scanning hyperspectral imager

on both a helicopter and fixed wing UAV for the detection of invasive plants. Zarco-Tejada et al. [313] also used a Headwall Micro-Hyperspec VNIR on board a fixed wing UAV for acquiring hyperspectral data in vineyards. Lucieer et al. [166] tested the same hyperspectral sensor mounted to an octo-rotor helicopter. In another example, Hruska et al. [121] tested a Resonon Pika II line scanning hyperspectral imager on a fixed wing UAV.

Frame based and snapshot hyperspectral cameras are very frequently used for smaller single- and multi-rotor helicopter platforms. For instance, Aasen et al. [1] mounted a snapshot hyperspectral camera (Cubert UHD 185-Firefly) on an octorotor helicopter for vegetation monitoring. Honkavaara et al. [119] and Tuominen et al. [275] used FPI based frame hyperspectral cameras on multi-rotor helicopter UAV platforms for tree species and surface moisture recognition.

Proximal Sensing from Ground Vehicles

As sensors have increased in spatial resolution, satellite and aerial HSI is able to resolve more detail on the ground, especially from low altitude UAVs. Nevertheless, for applications where sub-centimetre accuracy is required, such that individual leaves or fruit can be identified, it is necessary to acquire data at close proximity to the surface of interest, often referred to as proximal sensing. In addition, depending on altitude, remote sensors are also limited to wavelengths above the ultraviolet and blue range due to atmospheric scattering [26]. Proximal sensor orientation is also more flexible, whereas remote sensors are usually confined to an approximately downward facing perspective. Furthermore, ground based sensors enjoy the advantage of being able to use artificial illumination [26]. While multirotor UAVs are able to fly at very low altitudes, they are limited in terms of flight time and payload. Ground based platforms, on the other hand, can carry much greater payloads, larger batteries, and greater amounts of fuel, allowing for longer acquisition campaigns without interruption. Mobile ground platforms can include manually driven vehicles, such as cars, trucks tractors or simple hand pushed frames, and autonomous platforms [322, 48, 283]. Zhang et al. [322] designed a tractor towed platform that identified

weeds with a hyperspectral camera and treated them with hot food grade oil. Manually driven "buggies", such as BreedVision [29], PhenoMobile [48], and GPhenoVision [129] can carry more weight and supply more power due to their larger size, and therefore tend to include a greater array of equipment in addition to hyperspectral sensors, including 3D time-of-flight, light curtains, and thermal imaging. On the other hand, autonomous platforms or autonomous ground vehicles (AGVs) can be more operationally convenient and examples that carry hyperspectral sensors include Ladybird [277], Shrimp, and Bonirob [235, 138]. Figure 2.2 shows the Ladybird and Shrimp AGVs and the mounting configuration of their hyperspectral cameras.

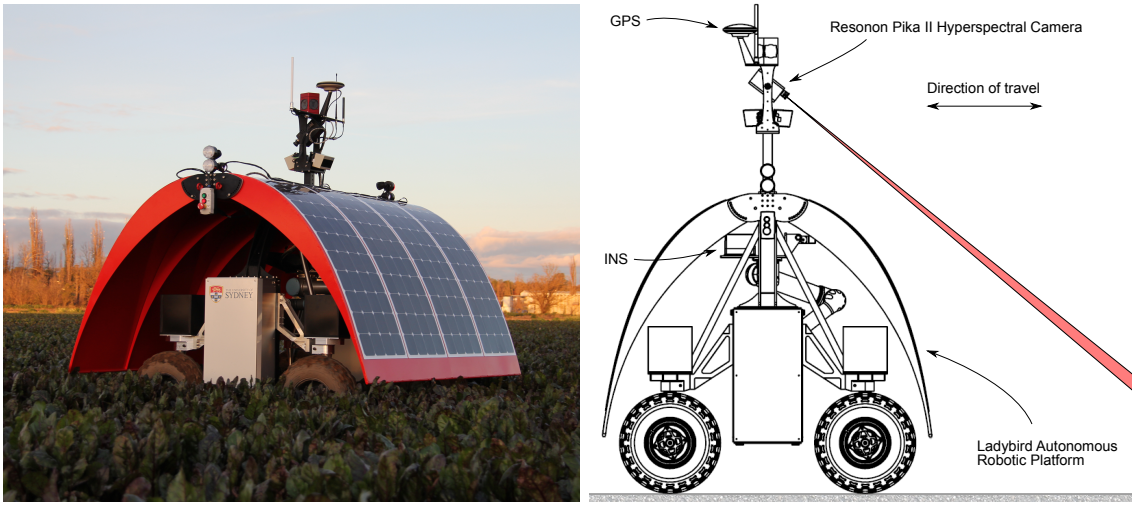
Outside agriculture, there are only a few studies addressing HSI using ground based platforms. In mining applications, HSI has been investigated in identifying minerals on mine faces, allowing more precise planning. For example, Schneider et al. [243] used HSI to detect and classify a mine face's geology. While the experiment was conducted with a tripod mounted sensor, the authors discuss the advantages of mounting the equipment on an AGV, allowing data capture from different angles, at various times throughout the day, and even while mining.

2.1.4 Applications

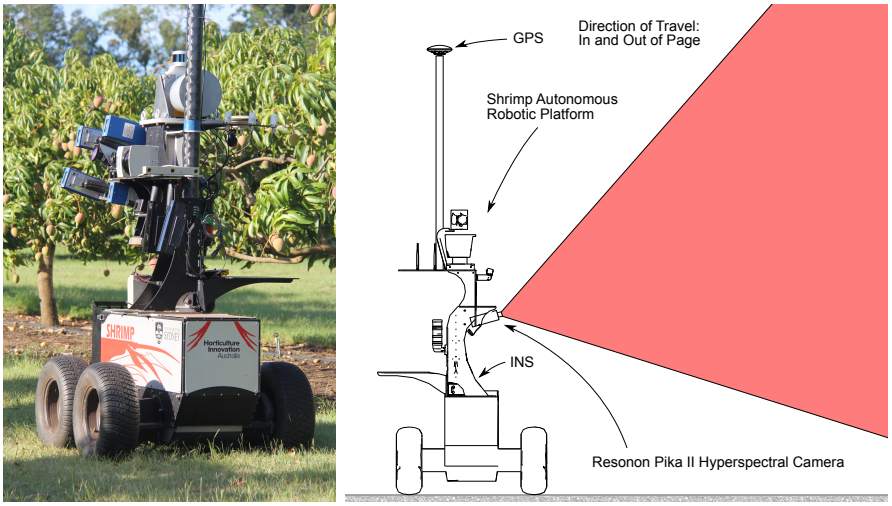
HSI has been used in a number of diverse applications. The earliest hyperspectral imagery was used for mineralogy and geology, and was later followed by other applications such as the identification and analysis of vegetation. While early on HSI was primarily a technology used for remote sensing, more recently proximal sensing has been explored in mineralogy, agriculture, medicine, food quality control, archaeology and flood detection. This section provides a broad overview of the applications where HSI has been studied or employed.

Geology and Mineralogy

Mining and mineralogy are two areas where HSI has been applied both at altitude and proximally, and represent some of the earliest uses of hyperspectral imaging [88].



(a) Ladybird



(b) Shrimp

Figure 2.2 – The ACFR's Ladybird [277] (a) and Shrimp (b) autonomous ground based mobile platforms. While both carry a Resonon Pika II hyperspectral line scanning camera, their viewing geometries are set up for two different PA applications. On Ladybird, the line scan is pushed in front of the platform to create a hyperspectral image of row crops on the ground, while on Shrimp the line scan is swept in a sideways configuration to produce hyperspectral data of orchard trees.

In this application, it is desired to measure the quantities of the different mineral components, which is made possible by absorption features in the visible and near infrared (VNIR) and SWIR reflectance spectra due to electronic and vibrational processes in the various substances [52].

There are numerous examples of aerial hyperspectral remote sensing being used in mineralogy and geology related applications. For example, Bedini et al. [19] used SWIR data from the HyMap imaging spectrometer to map the surface mineralogy in the Rodalquilar caldera region of Spain. Riaza and Müller [227] also used HyMap to monitor mine tailings in the Iberian Pyrite Belt, Spain. Hyperspectral data captured from satellites has also been examined for this application, for example for mapping hydrothermally altered rocks [81] and stratigraphic and lithologic mapping [43] using EO-1 Hyperion data. However, there has been some concern that the limited SNR of current spaceborne sensors prevents them from achieving the same level of mineralogic mapping accuracy as aerial solutions [142, 52].

Proximal examples in this field include Schneider et al. [243], who used two different hyperspectral sensors, one VNIR and one SWIR, fusing classification results from both to detect and classify geological structures on a mine face, and Murphy et al. [198], who applied SWIR between 970 and 2500 nm to the identification mineralogical properties of a mine face. Both of these studies used tripod mounted ground based hyperspectral imagers.

Vegetation and Agriculture

There has been extensive use of HSI in the analysis and mapping of vegetation [268]. Some applications include wetland mapping [92, 4, 5], forestry [314, 115, 83], plant ecology [189] and agriculture [258, 9]. In the remote sensing literature, there are numerous examples of hyperspectral aerial and satellite imagery being examined for agricultural applications, including identifying invasive plant species or different cultivars; mapping diseases and plant stress; predicting yield; measuring the concentration of biochemical constituents, such as nitrogen or chlorophyll; and biophysical proper-

ties, such as leaf area index (LAI).

Leaf nitrogen content is an important measure for fertilizer management. Airborne hyperspectral data has been examined for measuring nitrogen content, for example to predict nitrogen content in rice [125, 237], corn [87] and pastures (carbon to nitrogen ratio) [20]. In another example, Vigneau et al. [283] acquired proximal HSI data using a camera mounted to a tractor, to estimate leaf nitrogen content in wheat. Chlorophyll content is another measure used to gauge plant health, and its estimation has also been explored using airborne [102, 51] and satellite [299] HSI data of crops.

LAI is used to measure crop growth, yield and foliage cover [9], and has been measured using both satellite [155, 299] and airborne [101, 57] HSI data. There is some evidence to suggest that hyperspectral imaging is more effective than broad band multispectral data for estimating yield [303]. Aerial HSI has been studied to estimate yield or yield variability in wheat sorghum fields [303, 302], citrus orchards [307, 308] and cotton [315].

Identifying, measuring and quantifying plant disease, stress and deficiencies is another important area of study in PA, as it allows early detection and targeted application of treatments. Both near and remote sensing techniques have shown potential in that context [177], where the most common sensor technologies include thermography, fluorescence and HSI. HSI in particular has been shown to be highly suitable for detecting crop growth anomalies, as it allows detailed examination of stress dependent changes in targeted spectral ranges [188]. Leaf reflectance is influenced by plant stress via changes in pigmentation, hypersensitive reaction and cell wall degradation [177]. Additionally, properties specific to particular diseases may also be detectable in the spectral domain [177]. Aerial and satellite hyperspectral imaging has been studied for detecting plant stress and various diseases [176, 188, 173, 248, 11, 319]. There are only a few cases of proximal HSI being used for the detection of plant disease, even though it has been suggested that it is preferable to remote sensing in that application [205]. For example, HSI has been used from a buggy for early disease detection (yellow rust) in wheat fields [25, 193], with a portable imaging system (enclosed box with artificial light source) to identify citrus canker [219], and in the lab to reveal tomato spotted

wilt virus in capsicum plants [192].

Differentiating plant species is another challenging task which is useful for applications such as site specific weed management (SSWM). There have been several demonstrations of the use of conventional imagery for discriminating between different plant species using features such as leaf or flower shape, texture, branching structure or vein patterns [252, 41, 199, 113, 214, 255, 126, 160, 37], or more recently using convolutional neural networks (CNNs) and deep learning [238, 190]. However, the aforementioned techniques can be adversely affected by several difficulties. Plant morphology can be altered significantly by flexing and bending, or growth stage [54, 41]. Within-class variation can be very high for some plants, complicating the discrimination between groups [41]. Other external factors, such as weather, insect damage or soil deposits can have a significant detrimental impact on results [321]. Occlusion and unfavourable plant or leaf orientation may also severely impact classification performance in a field setting [252], although some recent methods have shown robustness to partial occlusion [107].

While early papers concluded that differentiation of plant species using their spectral signature may be very difficult or impossible [145], there are now a number of studies indicating that spectral techniques can be used for discrimination of plant species owing to the use of advanced multivariate statistical methods [285, 208, 252]. Spectral classification addresses many of the issues faced by shape or texture based techniques. For example, as many such approaches operate on a per-pixel basis, they are robust to partial occlusions and changes in orientation [252, 320]. In an early example of plant species discrimination, Lacar et al. [144] used airborne HSI to discriminate between grape cultivars. Later, using data from an in-field hyperspectral imager, Vrindts et al. [284] demonstrated separation of sugar beet and maize crop from weeds, and similarly Okamoto et al. [207] classified four species of weed among sugar beet crop. Suzuki et al. [264] used hyperspectral data to map herbage mass and botanical composition in pastures, which is important for understanding the behaviour and performance of grazing animals. Recently, classification over multiple seasons has been investigated and spurred development of more sophisticated classifier pipelines that can deal with

long-term temporal variation [321].

Food quality control

HSI has been explored as a tool to evaluate the quality of agricultural and food products, including fruit, vegetables [159, 218, 94, 220] and meat products [45]. The prediction of internal quality parameters of fruit, such as soluble solids contents (SSC), firmness, sugar content, water content, acidity and other measures, has been demonstrated using various HSI techniques in several fruits [159, 218, 94], including apples [202, 186], strawberries [59], oranges [161], peaches [165], grapes [14, 65] and mangoes [236, 247]. HSI has also been used in detecting external quality defects in fruits and vegetables [159, 218, 94], such as bruises [301, 95], chilling injuries [60], contamination [137] and rottenness [90]. The hyperspectral imaging techniques used in these applications include reflectance, absorption, transmittance, scattering and fluorescence spectroscopy. The literature in this space mainly relates to lab and production line environments, as many of the aforementioned techniques are more easily implemented in such settings, where lighting and scanning geometry can be tightly controlled.

Medical

HSI has found increasing use in medical diagnosis and image-guided surgery [164]. Hyperspectral imaging techniques have been studied for the detection of cancers [66, 53], identification of heart and circulatory disease [36, 147], retinal examination [131, 245], monitoring of shock [84] and in other medical diagnosis applications and research [164]. In addition, HSI could be used as an aid during surgery, by seeing tissue through layers of blood, detecting residual tumours, perceiving viable tissue and preventing its removal, and generally visualising anatomy [164].

Other applications

HSI has also been suggested for in rescue robotics. Trierscheid et al. [273] showed that the spectra of skin in NIR are very characteristic, even when covered by thin layers

of ash, which would prevent recognition with colour cameras. Other applications include security, surveillance and target acquisition [312]; detection of psychological stress [32]; urban land cover estimation [271]; conservation of works of art [67, 154]; archaeology [154]; flood detection [92]; water quality analysis [92]; pharmaceutical process monitoring and quality control [93]; and forensics [58, 179].

2.2 Hyperspectral Data Processing

The processing of HSI data in order to achieve a desired outcome, usually the prediction of classes, or estimation of a chemical or physical measure, normally requires several steps. First the data must be radiometrically calibrated and corrected in order to minimise the effects of both the sensor and illumination conditions. A number of preprocessing operations are then often applied to improve the ability of subsequent operations to extract information from them. Due to the large number of bands available in HSI, dimensionality reduction is an important and common step, which aids both in terms of computational demands and further inference performance. Finally, machine learning techniques are applied in order to predict a desired quantity or class. Many of the processing techniques mentioned in this section fall under the umbrella term chemometrics, which is the study of statistical and mathematical methods applied to the field of chemistry [184], and is particularly applicable to inferring chemical quantities in vegetation.

2.2.1 Radiometric Calibration and Correction [1]

There are various factors that affect the light as it leaves the source, passes through the atmosphere, reflects from a surface, passes through the atmosphere again, and enters the sensor. In many cases it is desirable to obtain a spectrum that is independent of lighting conditions, atmospheric scattering and camera/lens effects.

[1] Contains parts adapted from Wendel and Underwood [289] with minor modifications.

As a first step, to remove any effects the camera and lens combination may have on the image, the hyperspectral image may be converted to at-sensor radiance [262]. This can be achieved by recording a flat field image using a uniform light source, such as an integrating sphere. An integrating sphere can be thought of as a special diffuser that has an entrance aperture and an exit aperture, where light enters and exits respectively. The interior surface of the sphere is coated with a highly reflective diffuse material, and therefore the multiple reflections within the sphere result in a uniform and unpolarised light [211, p. 224]. The spectral characteristic of the light source in the integrating sphere is also known, allowing not only for correction of spatial non-uniformity (e.g. vignetting), but also spectral non-uniformity. If the at-sensor radiance during data acquisition of the uniform light source is known exactly, it is possible to obtain at-sensor spectral radiance in absolute units (i.e. $W \cdot sr^{-1} \cdot m^{-2} \cdot nm^{-1}$), but often it is sufficient to obtain pseudo-radiance, for example by assuming the light source's spectral distribution is equal to 1 [262], which still allows correction for spectral and spatial non-uniformity. The calculation must also subtract dark current measurements, obtained during the integrating sphere measurement and during field acquisition, from the raw values .

When measuring real-world objects, ideally the obtained at-sensor radiance would be independent of any effects introduced by the instrument itself. However, it is not an independent measure of the properties of the surface being imaged, but also the illumination and atmospheric conditions at the time of acquisition, which can vary considerably. Therefore, it is also necessary to obtain reflectance, which is a property of the imaged surface only [31, 8]. The literature covers many different methods for obtaining surface reflectance, but they can be loosely split into two major groups: empirical methods and radiative transfer codes. The former uses information in the captured image, while the latter obtains spectral illumination and atmospheric scattering information independently, given the time of the day and composition of the atmosphere.

Empirical methods

There are several early scene-based approaches to reflectance retrieval from the 1980s [73], including the internal average reflectance (IAR) [141] and flat field [231] correction approaches. The former divides a hyperspectral image by the average spectrum for the whole scene, while the latter assumes that there is an area with spectrally neutral reflectances (little variation with wavelength) in the scene, which can be averaged and used to retrieve reflectance. While these methods are convenient, because no in field reference measurements are required, they often do not provide accurate results [73].

Using a reference panel that is measured in the same lighting conditions as the surface of interest (i.e. in the same scene or image) is a common way to determine reflectance of a surface [305, 280]. Ideally this reference should be a Lambertian scatterer with uniform reflectance within the spectral range of the sensor, such as Spectralon by Labsphere, which exhibits a very flat reflectance curve at a wide wavelength interval from about 300 to 2400 nm [80]. The reference surface's reflectance is known, for instance by measurement in the lab using a spectrometer. To perform the correction, the reference panel is located in the scene to be imaged and its at-sensor radiance at the time of acquisition is determined from the hyperspectral data, along with all other surfaces in the scene, whose reflectances are unknown. The reflectances of any of these other surfaces can be calculated by dividing their radiance spectra by the reference's radiance spectrum and multiplying by the reference's known reflectance. This method is effective in situations where lighting conditions are constant and the sensor is close to the object being measured, such as in the laboratory, factory, low altitude aerial and ground based applications, where atmospheric path radiance is negligible. Interpolation has been used previously to take into account gradual lighting changes [262]. This is useful over shorter durations, where lighting conditions change approximately linearly. However, this method is less suitable where abrupt or non-linear changes occur during the data collection period.

In higher altitude aerial (above approx. 100 m) and satellite applications, any light that is scattered back to the sensor from the atmosphere (path radiance) cannot be

ignored [191]. More rigorous methods, such as the empirical straight line method take this into account. This requires at least two readings of known calibration targets that have different reflectances at all wavelengths [40, 133]. A best fit line model can then be obtained to derive a function for each band. As with the simple reference panel method, this approach is susceptible to changes in lighting that deviate from what was measured at the reference targets.

As an alternative, reference targets can be substituted with a sensor that continuously measures downwelling illumination directly [201, 120, 48, 103]. The downside is that they are usually mounted to the platform and may therefore not exactly represent the lighting incident upon the objects being imaged. For example, illumination at higher altitude aerial platforms will be different to the light on the ground, as the length of propagation through the atmosphere is different, or because of the geometry between clouds, sun, imaging platform and imaged surface. In ground based imaging platforms this may be less of an issue, depending on the geometry of the platform, while in aerial platforms the problem can be partially overcome by using a ground reflectance panel reading that is adjusted based on the continuous downwelling sensor data [143]. In addition, there is an increased cost and complexity of having an appropriately calibrated downwelling sensor, which may be restrictive for some applications. Some manufacturers of higher cost sensors use fibre-optics to pipe downwelling light into a portion of the pixel space of the sensor to avoid the need for multi-sensor calibration (e.g. Jan et al. [127]), however, this is often not available in cheaper sensors.

Another reflectance retrieval method takes advantage of overlapping tie points between consecutive images. These can be used to determine radiometric model parameters via least squares optimisation [120, 103], and may also make use of radiative transfer codes [217]. The prerequisite is that a proportion of the imagery overlaps for every change in lighting. This makes it more suitable for data from full frame multi or hyperspectral cameras, such as Fabry-Perot interferometer-based units [77, 239], which are starting to find increased interest in UAV applications. Theoretically, the method should be extensible to line scanning cameras if there is spatial overlap at times that span the full range of illumination.

Radiative transfer codes

In remote sensing applications, atmospheric modelling from first principles (atmospheric radiative transfer codes) [256] is a widely used method to derive an estimate of reflectance. Radiative transfer codes take as input atmospheric parameters, such as column water vapour and aerosol content to model the propagation of electromagnetic radiation as it is affected by scattering, absorption and emission on its path to the surface and then to the sensor. Several codes have been developed, including ATREM [72], MODTRAN [22] and 6S [282].

Atmospheric transfer codes are known to be computationally intensive [98], and therefore some efforts have been made to produce more efficient algorithms, such as Simple Model for Atmospheric Radiative Transfer (SMART) [246]. Alternatively, look up tables are commonly computed, that allow more efficient retrieval of reflectance using a number of input parameters (e.g. wavelength, pixel position, atmospheric water vapour content, aerosol optical depth, and terrain elevation) and interpolation [257]. An obvious advantage of using radiative transfer codes is that no reference targets are required. On the other hand, obtaining atmospheric parameters is not simple [133], and therefore "standard atmospheres" are often used, introducing inaccuracies [182]. Some solutions, such as FLAASH [6], HATCH [221], and ACORN [97], are able to estimate some atmospheric parameters, including column water vapour. However, commonly these require data in wavelength ranges that are well beyond the capabilities of the cheaper hyperspectral sensors that are normally used on mobile ground and lower altitude UAV platforms (which usually operate below 1000 nm). In addition, reliable parameter estimation often requires favourable conditions. For example, FLAASH does not deal well with heavy haze or water vapour content [183]. Belluardo et al. [21] performed an analysis on the SDISORT radiative transfer model [44] that quantified the estimated spectrum's uncertainty based on uncertainty of the input parameters. To improve accuracy, Thompson et al. [269] have recently examined combining radiative transfer models with empirical reference measurements via Bayesian inference.

2.2.2 Preprocessing

There are various ways in which data obtained using spectroscopic methods are commonly preprocessed. One problem with measured spectra is that there are often differences in overall intensity due to scattering effects [229], which can remain even in radiometrically calibrated and corrected spectra. The most common ways to address this in NIR spectroscopy are multiplicative scatter correction (MSC) and standard normal variate (SNV), but there are also numerous other ways to normalise spectra [229, 64]. Many of these methods have the same basic formula in common, which applies an offset, a, and scaling factor, b:

$$\hat{x}_\lambda = \frac{x_\lambda - a}{b}. \tag{2.1}$$

where x_λ and $\hat{x}_\lambda$ are the original and transformed spectra at wavelength λ. For SNV a is the mean of the spectrum, and b the standard deviation [16]. Other normalization approaches set $a = 0$, and b to the L^1 or L^2 (Euclidean) norms [229, 223].

MSC [79] is performed in two steps, by first calculating correction coefficients with respect to a reference spectrum, which is usually the mean spectrum of the calibration/training set, as follows:

$$\mathbf{x} = a + b\mathbf{x}_{ref} + \mathbf{e}, \tag{2.2}$$

where $\mathbf{x}_{ref}$ is the reference (mean) spectrum and $\mathbf{e}$ is the unmodelled part of the spectrum. Ordinary least squares regression is used to determine a and b which can then be used with Equation 2.1. Additionally, extended multiplicative scatter correction (EMSC) has been introduced with a number of augmentations that include polynomial fitting to the reference spectrum, fitting of a baseline on the wavelength axis, and incorporates prior knowledge of the analytes and interferent spectra [229, 180, 181].

Detrending is another popular method that is applied to spectral data, often in addition to other methods such as SNV [16]. It involves fitting a polynomial to the

spectrum and subtracting it. For example, Buddenbaum and Steffens [28] fitted a second order polynomial to SNV transformed spectra to analyse soil profiles using HSI.

Smoothing filters are commonly applied to remove some of the noise in the spectral signal. While a simple moving average is sometimes used (e.g. Gomez et al. [89]), the most frequently cited technique is Savitzky-Golay smoothing [241]. The method takes a window of given size around each wavelength and fits a polynomial of given order to the points within the window. The fitted polynomial is used to calculate the new value at each centre wavelength. In addition to smoothing, the Savitzky-Golay approach can also be used to compute the derivative of the signal, as it can be determined analytically from the fitted polynomial. Derivatives are popular, because they have the ability to remove both a baseline offset and linear trend Rinnan et al. [229]. The derivative can also be derived using the Norris-Williams (NW) technique [203, 204], which first smooths using a moving average and then calculates the first and second derivatives based on a gap on either side of the centre wavelength. The NW method is less prone to high-frequency noise than Savitzky-Golay, but the use of a fixed gap is difficult to defend in spectroscopy applications, because it assumes a fixed frequency component, such as in a time series [229, 223].

In addition to the above, the limited quantum efficiency of hyperspectral sensors at the edges of the frequency range can introduce significant amounts of noise. In cases where noise is extensive, it is recommended that bands at the highest and lowest ends of the available spectrum are removed [305].

Preprocessing of spectra should be done with care, as incorrect or too severe processing can remove valuable information [229]. Similarly, if multiple approaches are combined, thought should be given to their individual contributions and order. Generally, lower complexity is desirable in a pre-processing pipeline and subsequent modelling [229].

2.2.3 Dimensionality Reduction

Due to the large number of bands available in hyperspectral imagery, the data are high dimensional, often with several hundreds of available features (spectral bands). However, due to their contiguity along the spectral dimension, bands are often highly correlated. Using such high dimensional data directly in machine learning can be problematic due to the "curse of dimensionality", necessitating exponentially increasing amounts of data and amplifying the risk of overfitting, as well as escalating computational requirements. For these reasons, it is common to use some kind of dimensionality reduction algorithm. Carefully chosen bands or projections thereof can help with separation of class clusters or correlation with continuous output variables, improving machine learning performance [320, 210].

There are two broad methods of dimensionality reduction: Feature selection and feature extraction. The former group employs statistical methods to select bands that optimise performance in subsequent processing. On the other hand, feature extraction transforms or projects the data into a lower dimensional feature space [210]. A significant amount of work has gone into studying and improving both types of methods, and exploring their advantages and disadvantages.

Feature selection isolates specific statistically significant bands that optimise subsequent performance. Kavzoglu and Mather [134] compared several band selection methods for use in a neural network classifier, comparing two search algorithms, stepwise forward selection (SFS) and genetic algorithm (GA), and various statistical separability measures (Wilks' Λ, Hotelling's T^2, and separability indices). On the balance, GA produced better classification performance than SFS, but not significantly and conclusively so. It was found that Hotelling's T^2 separability measure generally produced the best classification accuracies. Tibshirani [270] proposes the use of the Lasso operator for feature selection, which involves least squares regression with an L1 penality. The penalty has the effect of automatically setting insignificant coefficients to exactly zero, effectively performing selection without a threshhold. Sparse partial least squares (PLS) algorithms make use of the Lasso to perform feature selection [63]. Alissou and Zhang [10] applied Lasso after principal component analysis (PCA)

to achieve improved hyperspectral data compression.

In contrast to feature selection, extraction seeks to determine a transformation of the original data that optimises performance. Feature extraction can be further subdivided into linear and non-linear dimensionality reduction. In HSI, one of the most common linear dimensionality reduction (LDR) methods is PCA, which employs an eigenvalue decomposition of the feature/band covariance matrix [232]. When the data are projected to the directions of the resulting eigenvectors, high eigenvalues indicate the highest variance, the assumption being that this represents most of the statistically useful information. In the case of HSI data, a few PCA bands with highest eigenvalues carry the majority of the information [232]. Suzuki et al. [263] used PCA prior to an artificial neural network (ANN) and linear discriminant analysis (LDA) classifier for weed detection. This was compared to using original bands that have been selected via step-wise statistical method. PCA resulted in slightly better performance in most cases. Conde et al. [39] compared traditional PCA (known as "m-method") and PCA alternatives, such as supervised principal component analysis (SPCA), for foreign object detection, and found all to perform with greater than 99% accuracy, reducing computational complexity from 1024 dimensions down to 16-19 (depending on the method used). Cheriyadat and Bruce [35] noted, however, that there are many cases where hyperspectral data exhibit significant intraclass variance. As PCA is blind to training data labelling, data may be projected along dimensions maximising intraclass variance, which is counter to classification objectives. This is analogous to regression tasks, where high variance along a PCA dimension may not necessarily be an indicator for information that is related to the target variable.

While PCA is the most commonly discussed LDR technique, others have been investigated and applied, such as multidimensional scaling (MDS) and independent component analysis (ICA). MDS aims to map inter-point distances to a lower dimensional space [42, 267], while ICA ensures all extracted dimensions are not only uncorrelated (as in PCA) but also statistically independent, by utilising higher order statistics [130]. It is frequently used to separate data into its original sources with no other prior information, known as blind source separation [130].

Several past papers have detailed the use of the discrete wavelet transform (DWT) and continuous wavelet transform (CWT) for feature extraction and analysis of hyperspectral datasets [291, 27, 207, 317]. In a similar fashion to the Fourier function, wavelet transforms represent a function in terms of wavelet functions (also known as mother wavelets). In the case of the Fourier transform, these are sinusoid with infinite length. In contrast, wavelet mother functions have limited length, and consequently a major advantage is that both frequency and spatial information is retained [91]. Okamoto et al. [207] applied DWT to crop and weed discrimination. After a DWT transform with the Daubechie's wavelet ($n = 4$), five plant species (one crop and four weed) were classified using LDA. Zhang et al. [317] compared features in a CWT to a number of more conventional hyperspectral features (raw and projections) to estimate yellow rust disease severity. CWT produced superior results when followed up with a multivariate regression model.

Non-linear dimensionality reduction (NLDR) has received increased attention over the last two decades. Remote sensing HSI datasets may exhibit non-linearity due to multi-path scattering, variations in sensing geometry, non-homogenous composition of pixels, attenuating properties of media and the variable presence of water [170, 13]. Non-linearity may manifest in a way that negatively affects performance when applying LDR methods, such as PCA and MDS. Several non-linear dimensionality reduction techniques (also known as manifold learning) have been developed, including locally linear embedding (LLE) [234], Isomap [267], spherical stochastic neighbor embedding (SSNE) [168], and multidimensional artificial field embedding (MAFE) [169].

Roweis and Saul [234] devised LLE, which maintains a k-nearest neighbour reconstruction of each point while mapped to a lower dimensional space. This method was then further developed by Ma et al. [171] and Fang et al. [62] to improve robustness and include spatial information. In a comparison between LLE, variations of LLE, and other dimensionality reduction techniques, using a acSVM classifier, LLE exhibited competitive performance.

Tenenbaum et al. [267] introduced Isomap, which works by placing data points on a

graph, where each point connects to its nearest neighbours. Points are then mapped into a lower dimensional space by minimising the error of geodesic distances to the original feature space. Wang et al. [286] applied Isomap to aerial hyperspectral data to discriminate between seven different classes (but tested on only four) using a Gaussian mixture model (GMM). Accuracy was high between classes of relatively dissimilar spectral signatures, while performance of the pipeline decreased markedly when attempting to discriminate more similar spectra, such as trees and grass.

Lunga et al. [170] provided a detailed recent review of NLDR methods applied specifically to HSI. Of particular interest is the fact that some manifold learning approaches, such as Isomap and LLE often did not perform better than PCA in the particular dataset being tested. On the other hand, two more recent, state of the art methods, MAFE [169] and SSNE [168] gained significant improvements over the other methods.

2.2.4 Machine Learning Models

Machine learning models may be categorised into those that perform classification or regression. The difference between the two is that for classification predictions are categorical, whereas for regression predictions are continuous. Both classification and regression approaches are used in a broad variety of applications spanning different research fields. This section will concentrate on the methods that have been successfully employed in spectroscopy and HSI applications.

The spectral angle mapper (SAM) is a popular distance measure with specific application to hyperspectral data classification. It considers each spectrum to be a vector in the feature/band space, where the angle between vectors is calculated as shown in Eq. (2.3):

$$\alpha = arccos\left[\frac{\sum_{b=1}^{B} s_{ib}s_{jb}}{\left(\sum_{b=1}^{B} s_{ib}^2\right)^{\frac{1}{2}}\left(\sum_{b=1}^{B} s_{jb}^2\right)^{\frac{1}{2}}}\right] \tag{2.3}$$

where B is the number of bands, s_i and s_j are two spectra and α is the angle between them. The angle between the vectors is inversely proportional to their similarities.

Test spectra are classified according to the smallest angle to known average class spectra. Garcia-Allende et al. [76] utilised SAM to successfully discriminate between tobacco leaves and foreign objects. On the other hand, Shafri et al. [249] showed that SAM exhibits significantly inferior performance compared to ANN or maximum likelihood (ML) classification when classifying tree species in order to map tropical forest cover in aerial data. This indicates that SAM may not be suitable for discriminating between classes of very similar spectra, such as different types of vegetation. Moreover, Murphy et al. [197] demonstrated that SAM is not necessarily insensitive to variations in illumination or albedo, as was previously believed.

While SAM was specifically developed for use in hyperspectral data, many other types of classification methods have successfully been applied to HSI. For instance, LDA is a statistical method that uses a ratio of intra- and inter-class variance to predict class membership [263]. Suzuki et al. [264] applied LDA to discriminate three plant species groups (perennial ryegrass, white clover and other plants) in order to map botanical composition and herbage mass in pastures, and Gray et al. [96] used LDA to differentiate soybean from six weed species. The k nearest neighbour (KNN) algorithm has also been used successfully to classify hyperspectral data [74, 206, 106]. Its main advantage is that training does not require additional computation, as classification is achieved based on distances to existing known data points [75]. However, prediction can be more computationally resource intensive than some other classifiers [206]. ML classification assumes that statistics for each class in each band are normally distributed, and calculates the probability that a given data point belongs to a particular group. Each data point is classified according to the highest probability. Good performance has been demonstrated when discriminating between classes of vegetation [249]. Another statistical machine learning technique that has also shown promise in classifying hyperspectral data are support vector machines (SVMs), which attempt to obtain a hyperplane that maximises the distance to two classes [82]. If classes are not linearly separable, SVMs can be used with a kernel function that maps the input variables into a higher dimensional feature space that is linearly separable, also known as the “kernel trick” [244, 63]. Murphy et al. [197] found that while SVM

performed very well when classifying hyperspectral data sourced from the same set as the training data, it exhibited poor results when classifying a mine face from a separate library training set, implying that further work is required if it is to be used in that context.

The most commonly used regression methods used in spectroscopy are multiple linear regression (MLR), principal component regression (PCR) and PLS [175, 223]. MLR is one of the oldest regression approaches, but while it is simple and fast, it does not deal well with highly dimensional, collinear, and noisy input data like VNIR spectra [298, 240]. PLS [78], on the other hand, overcomes these downsides and has consequently become one of the most frequently employed tools in VNIR spectroscopy and for chemometrics in general [266, 175]. PLS attempts to find latent variables (known as components or factors) that capture most of the information from the input variables (e.g. spectra) that is useful for predicting the output variables [223]. These latent variables are determined iteratively, for example by using the nonlinear iterative partial least squares (NIPALS) algorithm [266]. Another popular regression approach is PCR, which simply performs MLR after a PCA dimensionality reduction [78]. PCR inherits PCA's downsides, namely that it performs a transformation that maximises the variance in the input data within in the first number of components, and unlike PLS does not jointly consider maximising output variable variance [223]. PLS is generally considered to perform better than MLR or PCR [175].

MLR, PCR and PLS all assume that there is a linear relationship between the input and output variables. Therefore this relationship must be inherent in the data, or the data must be preprocessed in some way to remove non-linearities (e.g. $log(1/R)$ transform or manifold dimensionality reduction). There are, however, alternative non-linear supervised machine learning methods. The previously mentioned "kernel trick" can also be used for regression to map the input variables into a higher dimensional space that is linearly related to the output variable [175]. For example, Nicolai et al. [200] examined kernel PLS with Gaussian, quadratic and cubic polynomial kernel functions to predict the sugar content in apples from NIR spectra. However, in this particular case the authors found that regardless of the kernel used, kernel PLS

performed no better than ordinary PLS.

ANNs are a broad category of machine learning approaches that have found increased use due to advancements in computational performance. While there are many types of ANNs, the most common types used for HSI are feed forward networks with one or more layers, each with a collection of nodes (i.e. "neurons"), which multiply the input from nodes in the previous layer with a set of weights, sum the result and output a value via a non-linear activation function [158, 111, 12]. The weights and biases are the model parameters to be learned, and are obtained by via back-propagation against a training set [111]. While feed forward ANNs may have only input and output layers, it is common to add a number of hidden layers in-between. ANNs methods can be employed for classification [206, 39] or regression [279, 174].

A particular variety of ANN, CNNs and deep learning have recently garnered much interest in the computer vision community, mainly for classification [140, 250, 265, 109] and object detection [226, 162, 108] applications using RGB images. In contrast to traditional fully connected (or dense) networks, CNNs apply a convolutional operation using a filter of a given size to each layer. The convolutional filter or kernel contains the weights to be optimized by using back-propagation. Convolutional operations are highly parallelisable, allowing for efficient computation on processors with many cores, such as graphics processing units (GPUs). Compared to fully connected networks, CNNs have far fewer weights, which are the bulk of all parameters to be optimised, increasing computational performance, reducing memory usage and the risk of over-fitting.

Various types of CNNs have had a disruptive impact on image classification and object detection over the past decade. Work in this area is often referred to as "deep learning", indicating the large number of layers used. While in one of the earliest works, LeCun et al. [148] introduced a neural network for handwriting recognition, Krizhevsky et al. [140] designed a CNN that raised the bar considerably over previous models. Since then, several studies have made great strides, increasing classification performance by growing the number of layers [250, 265, 109]. Zhong et al. [328] in particular introduced residual blocks with shortcut connections to combat a degrada-

tion problem, which previously decreased accuracy in deeper networks. CNNs have also enabled accurate object detection, with the advent of region based object detectors, such as R-CNN, Fast R-CNN and Faster R-CNN [86, 85, 226], and single shot detectors like SSD and YOLO [162, 224]. Recently, CNN based networks have been used to perform image segmentation instead of simply outputting bounding boxes [108].

CNNs and deep learning have also been investigated specifically for classifying hyperspectral data. In one of the earliest papers on the topic, Chen et al. [34] used stacked autoencoders to extract useful high level features, obtaining greater accuracy when coupled with SVM or logistic regression compared to other established feature extraction methods. The approach first applied PCA on the spectral data, then flattened a region around a central pixel, which was then passed to the autoencoder, taking into account spatial context. Stacked auto-encoders have also been applied to learn illumination invariant features in order to improve classification performance outdoors [292, 295]. Hu et al. [122] proposed a simple network that worked by convolving over individual spectra, using a single convolutional and pooling layer, and two fully connected layers each, and achieved better classification performance compared to an SVM with a radial basis function (RBF) kernel. Using a more traditional approach, Makantasis et al. [178] employed a network that used 2D spatial convolutions on hyperspectral images that were reduced to 10-30 channels using randomised PCA, also achieving superior performance compared to SVM methods. By using 3D convolutions over both spatial and spectral dimensions, Chen et al. [33] proposed a network that achieved competitive classification accuracy. Lee and Kwon [149] demonstrated a much deeper network with a total of nine layers, first stacking the output of three different size 2D filters, and then making use of residual connections in order to allow a larger number of layers. Similarly, Zhong et al. [328] tested residual neural networks with varying depths on hyperspectral data, and showed that they increased performance over more traditional CNN networks.

The training of neural networks using back-propagation can be tuned in many ways, and several optimisations and improvements have been proposed. One insight is

that using weights that have been pre-trained on a similar, often larger, dataset, as a starting point can speed up convergence and allow training with smaller datasets [310, 293]. This is also known as transfer learning. Another technique commonly used in image classification and object detection is augmentation, which artificially increases the training set size and diversity to combat overfitting and increase accuracy [212]. Windrim et al. [294] have demonstrated this with hyperspectral data by artificially introducing additional lighting differences.

As mentioned previously, some CNN networks apply filters over spatial as well as spectral dimensions and can therefore be considered to be a subset of a more general category of spectral-spatial models. These types of models seek to improve per-pixel predictions by taking spatial context into account [110]. In a classification task, for instance, it is likely that a pixel is of the same class as its surrounding pixels. While several spectral-spatial approaches have been explored [110], graphical models including markov random fields (MRFs) [261, 311, 30] and conditional random fields (CRFs) [325, 326] have received considerable attention in the literature.

2.2.5 Spectral Unmixing

In situations where the resolution of the hyperspectral sensor is such that a single pixel covers an area with multiple types of objects, the light seen by the camera can be a mixture of the spectra of these surfaces. Since this is a particular issue for high altitude aerial and satellite platforms, where a sensor's IFOV on the ground can be several metres, a significant body of work in the remote sensing literature addresses the identification of the pure components and their abundances within each pixel [116, 135]. This is commonly referred to as spectral unmixing, and the pure components are referred to as endmembers. This can also be useful on smaller scales, in applications such as mineralogy, where samples can be mixtures of several minerals [276].

The most basic and frequently employed unmixing model is the linear mixing model (LMM), which models a given spectrum as a linear combination of endmember spec-

tra plus an additional uncorrelated noise vector [116]. This formulation is applicable where the pure components exist on a flat surface and do not interact with each other through reflections. If this assumption is not reasonable and a significant amount of light reflects of multiple surfaces before reaching the sensor, then bilinear models are more appropriate. The two-endmember bilinear model introduced by Singer and McCord [251] added an additional artificial endmember, whose spectrum is the product of the original two pure endmembers. A specialised two-endmember bilinear unmixing model that incorporates soil and vegetation interactions has also been proposed, which additionally accounts for light transmitted though leaves [318]. Multiple endmember bilinear models have also been proposed more recently (e.g. Halimi et al. [104], Fan et al. [61], Halimi et al. [105], Meganem et al. [185]). Heylen et al. [116] cite a number of shortcomings of bilinear methods, including multiplication of endmember spectra that are smaller than one, diminishing their contribution; sensitivity to noise; exclusion of self-interactions in many models; and the large number of parameters in some models, making them prone to overfitting. Some bilinear models also exclude bilinear contributions of endmembers not present as linear components, which does not allow for external reflections present in some scenes. Heylen et al. [116] also note that the interpretation of abundances can become cumbersome since the abundances of the bilinear interaction terms form a significant part of the signal, and the linear abundances do not sum to one.

Spectral unmixing methods require prior knowledge of endmember spectra. They can be determined manually from the data by finding pixels with the maximum abundances of the desired endmembers [135]. Alternatively, endmembers can be selected from a library of known spectra [135]. This allows for better comparisons with other studies done at different times or in different illumination conditions. In addition, there are approaches that can determine endmember signatures automatically [135].

Given a spectral mixing model and a set of endmembers, the relative abundances must then be computed for each pixel in a scene. This process is known as inversion and several methods have been developed for this purpose [135]. They include least squares methods, which attempt to minimize the squared error between the

known and modelled spectrum, minimum variance methods, and variable endmember methods, which reduce the number of endmembers within each pixel [135]. Some approaches are able to determine the endmembers and their abundances simultaneously [116].

2.2.6 Vegetation Indices

A common method for extracting information from vegetation spectra is to compute carefully selected indices, which are designed to estimate various biochemical and biophysical properties [324, 5]. Indices can be categorised into broadband and narrowband varieties, where the latter is suitable for hyperspectral sensing applications. Many indices have been studied and compared in the literature [230], and include narrowband versions of broadband indices, e.g. ratio vegetation index (RVI) [132], normalised difference vegetation index (NDVI) [233], soil adjusted vegetation index (SAVI) [123] and normalised difference water index (NDWI) [71]; and indices developed strictly for narrowband use, e.g. plant senescence reflectance index (PSRI) [187], chlorophyll absorption ratio index (CARI) [136] and photochemical reflectance index (PRI) [70]. Vegetation indices are used in a range of applications, from the estimation of LAI and green biomass to pigments, carotenoids and light use efficiency [230].

Choosing the optimal wavelength centres and bandwidths for indices is an important research topic. Consider, for example, one of the most common vegetation indices, NDVI, which has been used for quantifying LAI, percent vegetation cover, intercepted photosynthetically active radiation (IPAR), and green biomass [324]. It has both broadband and narrowband versions, and can be calculated as follows:

$$NDVI = \frac{NIR - Red}{NIR + Red}. \tag{2.4}$$

A common narrowband wavelength combination used for red and NIR is 670 and 800 nm [254, 300]. However, some studies have suggested that for estimating chlorophyll

content and LAI different wavelength centres, such as 705 nm for red, are more suitable [324, 300].

2.3 Mapping of Line Scan Hyperspectral Data

For many applications it is desirable to obtain the position of each HSI pixel in world coordinates, a process called georeferencing. This section outlines some techniques used to achieve this. The main focus will be on data obtained from line scanning cameras, which is one of the most common types of hyperspectral imager, but also more challenging, depending on the georeferencing approach used. For example, methods that require significant overlap of the image data would be very difficult to apply to line scan sensors. Some georeferencing methods require precise knowledge of the sensor offset with respect to another sensor (e.g. navigation system). This section summarises some of the approaches available in the literature to obtain these measurements.

2.3.1 Mapping of Line Scan Data

A common way to georeference line scan data is to project it directly to a digital surface model (DSM) using known camera poses, referred to as direct georeferencing [242]. In order to achieve this, the position and orientation of the camera must be known as accurately as possible with respect to the world frame. Usually this requires the use of an accurate real time kinematic global navigation satellite system (RTK-GNSS) and knowledge of the camera's offset with respect it [262]. These offset values are a combination of the camera's relative position and orientation with respect to the platform, and are often referred to as lever arm lengths and boresight angles respectively. In addition, the DSM must be known with adequate accuracy.

The direct georeferencing process involves first converting the line scan image data to sensor pixels in the world frame using the well known pinhole camera model. For each pixel, view rays can be defined by the camera centre and sensor pixel points

in world coordinates. The georeferenced image coordinates are located where these ray lines intersect with the DSM. There are several examples of direct georeferencing being used for line scan data acquired using various platforms [166, 121, 128, 68, 127].

Indirect georeferencing, on the other hand, does not require positional information of the sensor. Instead other information is used, such as ground control points (GCPs), which are locations that can be identified in the data for which world coordinates are known [15, 213]. However, indirect approaches are far more challenging to implement for line scan data, as it is very difficult to have GCPs in every line scan. Therefore, direct and indirect approaches are often combined to achieve better results in cases where the requirements for adequate direct georeferencing cannot be met [15, 225]. For line scanning cameras, several such methods have been proposed in the literature. For example, Reguera-Salgado and Martín-Herrero [225] used a GCP based particle swarm optimisation routine to refine projection of line scan imagery to a known DSM. Suomalainen et al. [262] used a standard RGB frame camera in addition to a line scanning hyperspectral sensor to reduce the accuracy required by the navigation system. In this case, the authors only employed simple global positioning system and inertial navigation system (GPS-INS) with 4-10 m accuracy. A structure from motion (SfM) algorithm was used to generate a DSM from the frame camera images and GPS-INS positions, as well as camera poses. Using known offsets between the line scan camera and frame camera, the hyperspectral pixels could then be projected to the concurrently generated DSM. Barbieux et al. [15] applied a similar method, first correcting navigation data and obtaining a DSM using bundle adjustment of RGB frame images, and then projecting line scans. In this case, the procedure used tie points between the RGB reference and line scan data to optimise boresight angles. In another study, Tuo and Liu [274] first performed direct georeferencing in a coarse rectification step and then refined the result using a reference image with labelled GCPs and a two degree polynomial model.

2.3.2 Geometric Calibration [2]

Direct georeferencing requires precise calibration of the sensor's intrinsic (e.g., focal length and principal point) and extrinsic parameters (i.e., camera pose with respect to the vehicle body frame). In the remote sensing literature, determination of extrinsic parameters is known as lever arm (translation) and boresight (orientation) alignment.

Extrinsic and intrinsic calibration for 2D frame cameras has been studied extensively due to their ubiquitous use across many different fields, and established solutions exist [323, 194, 163, 117]. Calibration of 1D cameras has not received as much attention. Methods can be loosely grouped into two categories: scan-based calibration and line-based calibration [152]. Scan-based calibration requires an accurate rig with a linear actuator that moves the camera orthogonally to the line scan at a constant speed over a calibration pattern, such as a checker board [55, 124]. This method is suitable for industrial inspection applications in a controlled laboratory or factory setting, where a linear actuator, manipulator arm or other rig is capable of moving the sensor through a precisely specified trajectory. Line-based calibration methods, on the other hand, allow calibration from a single line scan of a 3D target with a carefully designed pattern of lines [151, 167]. Su et al. [259] proposed a similar approach to Li et al. [151] that also obtained radial distortion by capturing multiple line scans. Line-based approaches require that the dimensions of the calibration pattern are known precisely, and that the whole pattern has been imaged in one exposure. Recently, a variation of this method using multiple line scans of a planar calibration pattern has been proposed [306], and the use of an additional auxiliary frame camera has also been explored [152, 260]. All the aforementioned approaches are suitable for well controlled environments: for scan-based calibration the movement of the sensor needs to be accurately controlled, while for line-based methods, the position of the pattern with respect to the sensor is critical. However, in a mobile field platform, where the camera is rigidly mounted in a particular position to the platform, it is often difficult to meet either of those requirements.

[2] Contains parts adapted from Wendel and Underwood [288] with minor modifications.

In the aforementioned methods, extrinsic parameters are usually determined with respect to the calibration pattern or an auxiliary frame camera. Therefore to determine the camera to navigation system transform either requires accurate knowledge of a pattern or points in world coordinates or an additional step such as "hand-eye" calibration [172]. Hand-eye calibration involves determining the transformation from a camera to an end effector (a robotic hand for instance), where these are rigidly linked, and is a thoroughly covered topic in the robotics literature. The problem is generally solved by imaging a calibration pattern from many different locations, where the transformations between the different end effector positions and camera to calibration pattern transformations are known using standard frame camera calibration techniques.

As remote sensing most commonly involves imaging from an aerial or satellite platform, translation (lever arm) offsets have a smaller effect on imaging accuracy, and can be measured manually [150, 213]. Accurately geolocated GCPs are commonly used to determine boresight alignment [195, 68, 316, 118], which can also be adopted for ground based applications [2]. Efforts have been made to avoid the use of GCPs, by detecting points of interest in separate scans of the same area and determining their 3D position using a known digital elevation model (DEM) [150]. Similarly, Knauer and Seiffert [139] matched line scan hyperspectral data to a georeferenced RGB image using edges in order to determine boresight parameters. Non-surveyed tie-points between overlapping acquisition runs have also been used in combination with bundle adjustment to determine boresight parameters [297]. The use of GCPs has been combined with DEMs to improve accuracy and allow self-calibration [309]. Frame cameras have been used to aid in determination of boresight misalignments [296], and additionally in combination with a DEM [15].

While there are several examples of line scanning cameras being examined in low altitude UAV and mobile ground based applications [222, 48, 287, 273], there are fewer studies addressing the extrinsic calibration requirements that closer proximity to the scene implies. Requirements include obtaining a 6 degrees of freedom (DOF) extrinsic parameter solution including translation, which has a greater influence on

mapping when proximal; avoiding GCPs, which need to be more accurately geolocated when viewed from nearby; and a need for smaller survey areas for calibration, because it is more difficult to obtain data over large areas with mobile ground vehicles.

2.4 Summary

HSI research has evolved significantly over the years since its first applications in the early 1980s. This chapter reviewed some of the types of hyperspectral sensors that have been developed since that time, which include tunable filters, line scan and snapshot cameras. Each type has particular strengths and weaknesses in terms of integration time, number of bands, spatial resolution, light throughput, size and weight. These need to be weighed against the particular intended application. Equally important is the choice of illumination; artificial lighting ensures consistency, while natural sun light can be more convenient or unavoidable in many applications.

HSI has been employed on a multitude of mobile platforms over the years. Early applications were dominated by high altitude aerial imaging. As technology advanced, however, hyperspectral cameras have been used on satellites, and lower altitudes, especially onboard smaller UAVs, providing increased ground resolution. Even more spatial detail can be gained in proximal sensing from ground based mobile platforms, which have garnered more interest over the past decade, particularly in agricultural applications. In addition, ground vehicles enable a greater variety of scanning geometries. However, there are few studies directly addressing the use of hyperspectral cameras on such platforms.

This chapter also provided an overview of the many applications for which HSI has been studied. There is a considerable body of literature studying HSI for geology and mineralogy, one of its first uses. Other applications include food quality control, medicine, art conservation, archaeology to name a few. However, a particularly active research area is the use of HSI for analysing vegetation. Findings in this domain can be applied in agriculture by, for example, determining crop health, growth and yield, and detecting weed species.

Processing hyperspectral data to produce estimates or predictions requires several steps. First, the data must be radiometrically calibrated and corrected to remove camera, lens, lighting and atmospheric effects, compensating for sensor non-uniformities and illumination changes during acquisition. The correction for lighting and atmospheric effects are commonly achieved with empirical methods or radiative transfer codes. The former uses information in the data, as well as known reference surfaces. Radiative transfer codes, on the other hand, require information about the atmosphere and sun geometry.

A number of further processing techniques have been proposed in the literature to maximise the predictive ability of the subsequent machine learning models. These include chemometric techniques adopted from NIR spectroscopic applications, smoothing, derivatives, and dimensionality reduction among others. Finally, several machine learning models have been reviewed in this chapter ranging from statistical approaches to various types of neural networks. In specific applications, alternative or additional processing steps such as vegetation indices or spectral unmixing are useful.

In many instances, the position of each hyperspectral pixel in the world frame must be determined, known as georeferencing. This chapter summarised the main categories in this area, direct and indirect georeferencing, and combinations thereof. Direct georeferencing in particular requires accurate camera pose estimates, which can be obtained from vehicle pose estimates (e.g. RTK-GNSS) if a camera's relative offsets with respect to the platform are known. Determining these position and orientation parameters is critical for accurate direct georeferencing. However, this can be a challenging task for line scanning cameras, particularly on mobile ground vehicles, where it is required to obtain a full 6 DOF pose, and the use of GCPs may not be practical. This chapter reviewed some of the existing approaches, but there are not many studies that focus on the use of line scanning cameras aboard mobile ground platforms.

In conclusion, this literature review exposed a number of research areas where further study is required. There are very few studies concentrating directly on the optimal use of hyperspectral sensors aboard mobile ground based platforms. This is

the primary impetus that motivates this thesis. To address the specific challenges encountered when acquiring HSI data at very low altitudes, Chapter 3 examines the use of an illumination compensation approach that is both effective and convenient in a field setting. Chapter 4 focuses on the problem of obtaining the relative pose of a line scanning camera, proposing a novel approach that requires only a small printed calibration pattern, and no GCPs or auxiliary sensors. The thesis then tackles two separate PA problems: weed detection and fruit maturity estimation. Chapter 5 proposes a self-supervised method for discriminating crops from weeds using HSI, leveraging regular crop seeding patterns that are prevalent at commercial farms. Finally, in Chapter 6 the maturity of mangoes was non-destructively estimated while still on the tree and mapped on an orchard-wide scale. These maturity maps can provide crucial information to growers that allows them to optimise harvest operations and timing.

Chapter 3

Compensating for Illumination Changes

When acquiring hyperspectral imaging (HSI) data in natural daylight, changes in illumination are inevitable due to the gradual movement of the sun over the course of a day, cloud variability, subtle variations in atmospheric composition and scattered light from surrounding surfaces. The movement of clouds is a particular issue for ground based and low altitude platforms that operate below cloud level. The resulting changes in lighting colour the measured spectra and can drown out the signal. Therefore, when processing hyperspectral data it is desirable to correct for the effects of environmental illumination and obtain reflectance, which is an illumination invariant property of the surface being measured. This also allows for comparing data acquired in dissimilar conditions.

As summarised in Section 2.2, common approaches to obtain reflectance include empirical methods and radiative transfer codes. When operating mobile ground platforms with hyperspectral sensors, conventional empirical methods can be inconvenient, as they require frequent scanning of reference panels with known reflectance. The logistics of placing these panels can become tedious for operators as ground platforms inherently require more time to cover the same area as higher altitude platforms. An alternative is to permanently mount a reference panel to the vehicle such that it

is in view of the camera at all times. However, this introduces other difficulties in the field due mounting geometry and transport logistics. Downwelling sensors, which measure illumination directly, are another option, but introduce undue complexity to the system. Due to their remoteness from the surfaces of interest, downwelling measurements are also not necessarily representative. An extreme example is tree canopies, where internal scattering from the tree itself significantly affects the signal, which would not be captured by the downwelling sensor. Reference panels are also affected in a similar way if they are not placed in close proximity to the surfaces to be measured. Due to rapid cloud variability, radiative transfer codes would be very challenging to use in this application. They require measurements of atmospheric composition, which are often difficult to obtain, and would need to be updated to match cloud cover and relative contributions of scattered and direct sunlight. In addition, modelling of local light scattering, such as in the aforementioned tree canopy example, would be very challenging.

To address these challenges, in Section 3.1 this chapter proposes an alternative approach to illumination compensation, which only requires a few representative reference panel readings. The approach is based on previous work by Drew and Finlayson [56], who introduced a method to derive illumination and reflectance spectra from the light seen by the sensor. It requires a set of known illumination and reflectance spectra for training, which can be obtained using reference panel readings. As long as the spectra are representative of the data being processed, no further information is required. This means that the reference data are not temporally tied to the HSI data to be processed. The approach was validated by comparing its recovery accuracy with traditional methods, evaluating its effects on vegetation indices, and trialling various training and test dataset combinations. Based on these results, a number of convenient field protocols are proposed to minimise manual labour for operators.

ISPRS Journal of Photogrammetry and Remote Sensing 129 (2017) 162–178

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Illumination compensation in ground based hyperspectral imaging

Alexander Wendel *, James Underwood

Australian Centre for Field Robotics, The Rose Street Building J04, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 20 June 2016
Received in revised form 12 April 2017
Accepted 14 April 2017
Available online 11 May 2017

Keywords:
Hyperspectral
Atmospheric correction
Illumination compensation
Reflectance retrieval
Ground based
Robotics

ABSTRACT

Hyperspectral imaging has emerged as an important tool for analysing vegetation data in agricultural applications. Recently, low altitude and ground based hyperspectral imaging solutions have come to the fore, providing very high resolution data for mapping and studying large areas of crops in detail. However, these platforms introduce a unique set of challenges that need to be overcome to ensure consistent, accurate and timely acquisition of data. One particular problem is dealing with changes in environmental illumination while operating with natural light under cloud cover, which can have considerable effects on spectral shape. In the past this has been commonly achieved by imaging known reference targets at the time of data acquisition, direct measurement of irradiance, or atmospheric modelling. While capturing a reference panel continuously or very frequently allows accurate compensation for illumination changes, this is often not practical with ground based platforms, and impossible in aerial applications. This paper examines the use of an autonomous unmanned ground vehicle (UGV) to gather high resolution hyperspectral imaging data of crops under natural illumination. A process of illumination compensation is performed to extract the inherent reflectance properties of the crops, despite variable illumination. This work adapts a previously developed subspace model approach to reflectance and illumination recovery. Though tested on a ground vehicle in this paper, it is applicable to low altitude unmanned aerial hyperspectral imagery also. The method uses occasional observations of reference panel training data from within the same or other datasets, which enables a practical field protocol that minimises in-field manual labour. This paper tests the new approach, comparing it against traditional methods. Several illumination compensation protocols for high volume ground based data collection are presented based on the results. The findings in this paper are applicable not only to robotics or agricultural applications, but most very low altitude or ground based hyperspectral sensors operating with natural light.

1. Introduction

An elevated awareness of environmental issues, food security and sustainability, coupled with an ever-present desire to reduce costs and waste, maximise quality and increase productivity has highlighted precision agriculture (PA) as an important tool for optimising farming practices (Tey and Brindal, 2012). Mapping and analysing the reflected light spectrum of vegetation has emerged as an important method for various PA objectives (Lee et al., 2010). Multispectral imaging, which can capture image data in several wavelength bands, has been used in various mapping applications (Zhang and Kovacs, 2012; Mulla, 2013), such as the estimation of soil properties (Gomez et al., 2008), weed management (López-Granados et al., 2016), pest management (Du et al., 2008), and crop classification (Panigrahy and Sharma, 1997).

Hyperspectral imaging, which is able to sense spectra of objects in hundreds of narrow bands, provides even more detailed information. This allows for precise measurement of plant health indicators (Thenkabail et al., 2002; Behmann et al., 2014), as well as classification of individual plant species based on spectra alone (Okamoto et al., 2014). There is a substantial body of research covering hyperspectral imaging in the remote sensing community, where both satellite and aerial imaging have been used to map vegetation for various research and farming applications, for example for mapping cotton field variability (Yang et al., 2004), vegetation cover estimation (Zhang et al., 2013), biomass estimation (Marshall and Thenkabail, 2015), vegetation/crop classification (Oldeland et al., 2010; Xue et al., 2017), disease mapping (MacDonald et al., 2016) and nutrient/chlorophyll concentrations

* Corresponding author.
E-mail addresses: a.wendel@acfr.usyd.edu.au (A. Wendel), j.underwood@acfr.usyd.edu.au (J. Underwood).

http://dx.doi.org/10.1016/j.isprsjprs.2017.04.010
0924-2716/

(Sims et al., 2013; Rao et al., 2008; Moharana and Dutta, 2016; Pullanagari et al., 2016).

More recently, hyperspectral imaging solutions operating at lower altitudes have begun to appear on both unmanned aerial vehicles (UAVs) (Uto et al., 2013; Honkavaara et al., 2013, 2016; Aasen et al., 2015) and mobile ground vehicles (Deery et al., 2014; Klose et al., 2010; Underwood et al., 2017), which are increasingly being used to provide high spatial resolution data. Mobile ground platforms for agricultural applications have ranged from simple hand pushed frames to manually driven motorised tractors or buggies to autonomous systems (Zhang et al., 2012; Deery et al., 2014). Recent examples of autonomous platforms include Ladybird (Underwood et al., 2017), Bonirob (Ruckelshausen et al., 2009), and the tracked Armadillo (Nielsen et al., 2012). Larger manually driven "buggies", such as BreedVision (Busemeyer et al., 2013) and PhenoMobile (Deery et al., 2014), can carry more weight and supply more power, and therefore tend to include a greater array of sensors, including 3D time-of-flight, light curtains, and thermal imaging.

Retrieving reflectance, which is a property of the imaged surface only (Chandra and Healey, 2008; Ahlberg, 2010), by compensating for environmental illumination is a particular consideration for hyperspectral sensors. Because higher altitude and satellite imagery generally require relatively clear skies, lighting is more consistent, being only dependent on the time of the day and atmospheric composition. Low altitude and ground based platforms can operate under cloud cover, allowing data acquisition whenever there is sufficient light, but this increases the amount of dynamic lighting variation, due to fluctuating cloud cover density. Additionally, because these configurations image smaller regions of the scene at a time, total scan durations are longer, increasing likelihood that not only the intensity but also the spectrum of light on the ground changes. This makes it difficult to obtain reflectance accurately, but it is often impractical to wait for opportune lighting conditions.

These difficulties highlight the need for autonomous platforms, which ensure that data acquisition is both consistent and fast, while minimising disruption to crops. Factors such as the trajectory of the platform, including its orientation and velocity can be tightly controlled, allowing data to be obtained in a regular manner, which is suitable to feed into automated processing frameworks. Autonomous systems allow very high resolution data to be obtained practically over large areas of a farm, breaking the trade-off between resolution and coverage.

In this paper, we examine the use of an unmanned ground vehicle (UGV) to gather high resolution hyperspectral data of crops, which are post processed to compensate for illumination changes in order to retrieve reflectance.

The contributions of this paper are:

- The development of several different field protocols for gathering the necessary training data for the illumination compensation method by Drew and Finlayson (2007). These present different trade-offs between the accuracy of illumination compensation and the logistical complexity of the field work.
- Testing the applicability of using historical reference data to correct for illumination in future datasets.
- An analysis of the sensitivity to illumination compensation of several metrics/indices that are commonly used in agricultural applications.
- Evaluating the suitability of a previously developed logarithm subspace method for illumination and reflectance extraction (Drew and Finlayson, 2007) for use on a large, high spatial and spectral resolution agriculture based field dataset.

By using the approach detailed in Drew and Finlayson (2007), the following important advantages can be realised:

- No reference target readings need to be tied to imaged pixels.
- Significantly reduced number of reference target readings.
- Feasibility to recover reflectance and illumination from previously acquired training data, where no reference panel readings are available.
- No need to obtain or estimate atmospheric parameters.

In Section 2, we briefly review the literature on illumination compensation. Particular focus is given to the subspace model method by Drew and Finlayson (2007), which approximates both illumination and reflectance spectra based on sets of training data. We posit this method as a basis for more convenient and practical novel field protocols that facilitate compensation for lighting. In Sections 3 and 4, experiments are documented that use high spatial (3 mm by 9 mm) and spectral (2 nm) resolution hyperspectral data cubes, covering 2.75 hectares of a plant phenomics trial. The experiments highlight the magnitudes of reflectance error that can occur when illumination compensation is ignored, and demonstrate the effectiveness of several illumination compensation approaches. Based on these results we provide some clear guidelines for obtaining reflectance in hyperspectral data from ground based field robotics systems (Section 5).

2. Surface reflectance retrieval methods

In this section, we summarise the most common methods used for atmospheric correction and illumination compensation in order to retrieve surface reflectance. For brevity, we use "reflectance" and "surface reflectance" synonymously, as opposed to "at-sensor reflectance" or "top of atmosphere (TOA) reflectance" (Teillet, 2015).

2.1. Empirical methods

There are several early scene-based approaches to reflectance retrieval from the 1980s (Gao et al., 2009), including the Internal Average Reflectance (IAR) (Kruse, 1988) and flat field (Roberts et al., 1986) correction approaches. The former divides a hyperspectral image by the average spectrum for the whole scene, while the latter assumes that there is an area with spectrally neutral reflectances (little variation with wavelength) in the scene, which can be averaged and used to retrieve reflectance. While these methods are convenient, because no in field reference measurements are required, they often do not provide accurate results (Gao et al., 2009).

Using a reference panel that is measured in the same lighting conditions (i.e. in the same scene or the same image as the surface of interest) is a common way to determine reflectance of a surface (Yao and Lewis, 2010; Uto et al., 2013). Ideally this target should be a Lambertian scatterer with uniform reflectance in the spectral range of the sensor, such as Spectralon by Labsphere, which exhibits a very flat reflectance curve at a wide wavelength interval from about 300 to 2400 nm (Geladi, 2007). Once the radiance of the reference target has been measured, the reflectance of a surface in the same lighting conditions can be obtained by dividing its radiance spectrum by the reference's and multiplying by the target's known reflectance (see Section 3.4). This method is effective in situations where the sensor is close to the object being measured, such as laboratory, factory, low altitude aerial and ground based applications, as long as lighting does not change from the conditions measured at the reference panel. Interpolation has been used in the past to take into account gradual lighting changes (Suomalainen et al., 2014). This is useful over shorter durations, where lighting conditions change approximately linearly. However, this method is less

suitable where abrupt or non-linear changes occur throughout the data collection period.

In higher altitude aerial (above approx. 100 m) and satellite applications, any light that is scattered back to the sensor along its line of view (path radiance) cannot be ignored (Mobley, 2016). More rigorous methods, such as the empirical straight line method take this into account. This requires at least two readings of known calibration targets that have different reflectances at all wavelengths (Conel et al., 1987; Karpouzli and Malthus, 2003). A best fit line model can then be obtained to derive a function for each band. As with the simple reference panel method, this approach is susceptible to changes in lighting that deviate from what was measured at the reference targets.

As an alternative, reference targets can be substituted with a sensor that continuously measures downwelling illumination directly (Noble et al., 2002; Honkavaara et al., 2013; Deery et al., 2014; Hakala et al., 2013). The downside is that they are usually mounted to the platform and may therefore not exactly represent the lighting incident upon the objects being imaged. For example, illumination at higher altitude aerial platforms will be different to the light on the ground, as the length of propagation through the atmosphere is different, or because of the geometry between clouds, sun, imaging platform and imaged surface. In ground based imaging platforms this may be less of an issue, depending on the geometry of the platform, while in aerial platforms the problem can be partially overcome by using a ground reflectance panel reading that is adjusted based on the continuous downwelling sensor data (Kuusk and Kuusk, 2010). In addition, there is an increased cost and complexity of having an appropriately calibrated downwelling sensor, which may be restrictive for some applications. Some manufacturers of higher cost sensors use fibre-optics to pipe downwelling light into a portion of the pixel space of the sensor to avoid the need for multi-sensor calibration (e.g. Jan et al. (2008)), however, this is often not available in cheaper sensors.

Another reflectance retrieval method takes advantage of overlapping tie points between consecutive images. These can be used to determine radiometric model parameters via least squares optimisation (Honkavaara et al., 2013; Hakala et al., 2013), and may also make use of radiative transfer codes (Prosa et al., 2013). The prerequisite is that a proportion of the imagery overlaps for every change in lighting. This makes it more suitable for data from full frame (snapshot) multi- or hyperspectral cameras, such as Fabry-Perot interferometer-based units (Geelen et al., 2014; Saari et al., 2009), which are starting to find increased interest in unmanned aerial vehicle (UAV) applications. Theoretically, the method should be extensible to line scanning cameras if there is spatial overlap at times that span the full range of illumination.

2.2. Radiative transfer codes

In remote sensing applications, atmospheric modelling from first principles (atmospheric radiative transfer codes) (Staenz et al., 2002) is a widely used method to derive an estimate of reflectance. Radiative transfer codes take as input atmospheric parameters, such as column water vapour and aerosol content to model the propagation of electromagnetic radiation as it is affected by scattering, absorption and emission on its path to the surface and then to the sensor. Several codes have been developed, including ATREM (Gao et al., 1993), MODTRAN (Berk et al., 1999) and 6S (Vermote et al., 1997).

Atmospheric transfer codes are known to be computationally intensive (Guanter et al., 2009), and therefore some efforts have been made to produce more efficient algorithms, such as Simple Model for Atmospheric Radiative Transfer (SMART) (Seidel et al., 2010). Alternatively, look up tables are commonly computed, that allow more efficient retrieval of reflectance using a number of input parameters (e.g. wavelength, pixel position, atmospheric water vapour content, aerosol optical depth, and terrain elevation) and interpolation (Staenz and Williams, 1997). An obvious advantage of using radiative transfer codes is that no reference targets are required. On the other hand, obtaining atmospheric parameters is not simple (Karpouzli and Malthus, 2003), and therefore "standard atmospheres" are often used, introducing inaccuracies (Mather and Koch, 2011). Some solutions, such as FLAASH (Adler-Golden et al., 1998), HATCH (Qu et al., 2001), and ACORN (Green, 2001), are able to estimate some atmospheric parameters, including column water vapour. However, commonly these require data in wavelength ranges that are well beyond the capabilities of cheaper hyperspectral sensors commonly used on mobile ground and UAV platforms (which usually operate below 1000 nm). In addition, reliable parameter estimation often requires favourable conditions. For example, FLAASH does not deal well with heavy haze or water vapour content (Matthew et al., 2002). Belluardo et al. (2016) performed an analysis on the SDISORT radiative transfer model (Dahlback and Stamnes, 1991), that quantified the estimated spectrum's uncertainty based on uncertainty of the input parameters. To improve accuracy, Thompson et al. (2016) have recently examined combining radiative transfer models with empirical reference measurements via Bayesian inference.

2.3. Subspace model recovery of illumination and reflectance

Approximating the constituent components of an at-sensor radiance spectrum (illumination, reflectance, path scattered radiance etc.) as linear combinations of distinct basis spectra is an alternative approach, that was originally introduced by Ho et al. (1990) in order to achieve colour constancy. Most commonly, basis vectors are estimated from existing training data of illuminants and reflectances (and sometimes path scattered radiance) using a Singular Value Decomposition (SVD) (Chandra and Healey, 2005; Drew and Finlayson, 2007). For each given input signal, these algorithms attempt to find an optimal combination of basis function coefficients, either using an iterative method or analytically. Being able to use a finite set of training data to compensate for illumination in a dataset opens up an opportunity for applications in ground based and low altitude platforms.

The standard use of reference targets ideally requires each pixel to be tied to a corresponding target measurement. With the subspace model method, however, reflectance and illumination spectra can be estimated without this requirement, as long as suitable training reflectance and illumination data are available. This permits simple field protocols where only a few representative reflectance targets need to be measured along with representative surface pixels. Furthermore, if a diverse enough database of prior training spectra is available, it is feasible that no reference measurements need to be made during future campaigns.

Drew and Finlayson (2007) propose an analytic solution using logarithms and includes an additional regression step to refine the results. The authors demonstrated that their method is over 100 times faster than non-analytical iterative approaches with 10 base spectra for both illumination and reflectance, while being at least as accurate. There is an obvious practical benefit in being able to rapidly compute results in both post-processing and real-time applications, such as weed detection and control (Wendel and Underwood, 2016; Underwood et al., 2015).

One downside of this method is that absolute intensity of illumination and reflectance are not recoverable (Drew and Finlayson, 2007). This is a consequence of the inherent ambiguity in intensity of the two results. However, for many applications, such as indices and classification, this is not a disadvantage, as in

these cases only relative intensities (i.e. spectral shape) within a given reflectance spectrum are of importance.

3. Materials and methods

This section outlines the equipment, materials and methods used to acquire data and carry out experiments. A hyperspectral line scan camera and reference target mounted to an unmanned ground vehicle (UGV) were used to scan thousands of different plots of cereal and legume crops for a phenotyping trial. After acquisition, hyperspectral data were first thinned and treated for saturation (except for maps). Then radiance was calculated using dark current and integrating sphere measurements. The data were then ready to be tested using the various lighting compensation methods. Experimental results were compared to reference results from a target that was in view at all times during data acquisition.

3.1. The autonomous ladybird system

The Ladybird robot (Fig. 1) was designed and built in 2014 at the Australian Centre for Field Robotics (ACFR) at The University of Sydney as a flexible tool to support research and development of robotics applications for the commercial vegetable production industry (Underwood et al., 2017). The Ladybird is mechanically adjustable for width, height of the central sensing tower, and height and angle of the solar covers, allowing adaptation for different farm configurations and crop heights. Forward and rear facing lidar and a spherical camera support obstacle avoidance and crop row detection, while real time kinematic (RTK)/global positioning system (GPS)/interial navigation system (INS) allows for map-based farm traversal. Crop sensing is provided by hyperspectral line-scanning, stereo vision (with strobe), thermal infra-red vision and the same lidars used for obstacle avoidance.

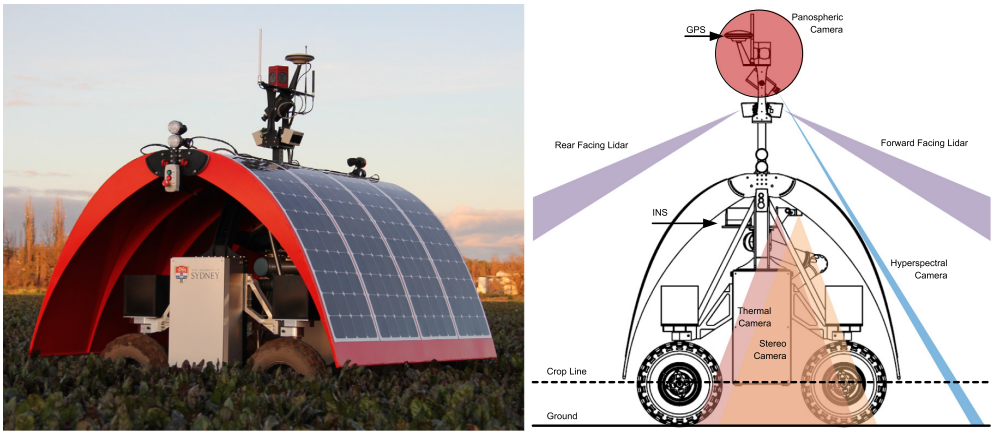
The Ladybird achieves autonomous traversal of crop rows and headland, by navigating over a pre-constructed farm-map. The map specifies the geometry of rows and headland access areas and how they are connected. Using the RTK/GPS/INS, a control system guides the Ladybird along these composite path segments with centimetre precision while maintaining constant speed and orientation. This is particularly important for spectroscopy, where a surface's reflectance may change depending on the sun's and observing instrument's relative angles.

3.2. Data acquisition

Hyperspectral data were acquired with a Resonon Pika II visible to near-infrared (VNIR) line scanning camera that was mounted to the Ladybird robot, pointing towards the ground at approximately 34°below horizontal, such that the visible line intersects the row of crops perpendicularly approximately 2.9 m in front of the Ladybird (see Fig. 1(b)). The camera produces hyperspectral line-images of 648 spatial pixels with 244 spectral channels in each pixel (spectral resolution of 2 nm from 390.9 to 887.4 nm) at a rate of 133 frames per second and native bit depth of 12 (saved as 16 bit binary files). A Schneider Cinegon 8 mm objective lens was used at an aperture setting of approximately f/2.5, and manually focused with a checker board at the expected crop heights. The lens provides a 33°field of view translating to a 3.0 mm/pixel spatial resolution per line scan, across the row, imaging just wider than a single crop row. There is a linear trade-off between vehicle speed and hyperspectral resolution along the row. At the top speed of 1.2 m/s, which allows approximately 0.72 ha/h and was used in this study, a resolution of 9.0 mm/pixel was achieved. This is well matched to the camera-lens combination's instantaneous field of view (IFOV) of 1.9 mrad, translating to a ground resolution of 11.8 mm.

For all scans examined in this study, a reference panel (QPcard 102, MFR #GQP102) was mounted permanently at the edge of the field of view of the camera (see Fig. 2). Spectral smile and keystone can be of particular concern when the panel is mounted at the image border, but calibration documentation from Resonon indicates that for this camera these effects have been optically minimised to less than half a spectral pixel peak to peak, therefore not requiring further correction. This was corroborated via personal correspondence with Resonon (R Swanson 2016, personal communication, 25 August). Dark current measurements were recorded regularly (approx. every 2–4 h), by completely blocking all light to the sensor, while averaging the data over approximately 10 s.

The data were acquired from a number of different phenotype trials operated by South Australian Research and Development Institute (SARDI). Throughout the campaign a large number of datasets (continuous logs) were collected, spread over three sites (Mallala, Pinery and Turretfield), which are located within a radius of 41 km in South Australia. For a complete description of all data gathered for phenotyping, and details regarding the automation of the platform, the reader is referred to Underwood et al. (2017). The



(a) Scanning a field (b) Sensor configuration

Fig. 1. The ladybird and mounted sensor configuration.

(a) (b)

Fig. 2. Permanently mounted calibration target (a) mounted to the platform and (b) as seen in a high resolution close up example of hyperspectral data in true colour. The calibration target can be seen at the bottom of (b) and appears saturated due to the gain applied for aesthetic purposes. The undulating appearance is a result of swinging and vibration of the target mount as Ladybird moves. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Dataset summary.

Dataset Name	Date	Length	Auto/ Manual	Plant Types	Environmental Conditions	% Sat.[a]	% Sat. Target [b]
DDPA-1	19/08/2015	17 min 11 s	Auto	Pulses	Cloudy, low light with some variation due to cloud movement.	0.0%	0.0%
DDPA-2	19/08/2015	21 min 37 s	Auto	Pulses	Cloudy, low light with some variation due to cloud movement.	0.0%	0.0%
CDA-0	19/08/2015	52 min 55 s	Auto	Cereal Varieties	Cloudy, moderate brightness with some variation due to cloud movement.	0.0%	0.0%
NMAD-1	20/08/2015	56 min 11 s	Auto	Various	Cloudy, clearing near the end of the set. Low light at beginning of log to bright near the end.	0.0%	0.1%
ML-0	18/08/2015	28 min 25 s	Manual	Lentils	Bright with a number of clouds causing heavy variation in lighting conditions throughout the log.	0.3%	27.8%
SUL-0	18/08/2015	36 min 30 s	Manual	Lentils	Bright with a number of clouds causing heavy variation in lighting conditions throughout the log.	0.2%	20.9%
DBVL-0	18/08/2015	45 min 5 s	Manual	Lentils	Bright with a number of clouds causing heavy variation in lighting conditions throughout the log.	0.1%	16.9%

[a] Percentage of total pixels that are saturated (excluding line scan pixels 500–648, which mostly correspond to the reference target).
[b] Percentage of line scans where the calibration target is saturated and therefore excluded from calculation of metrics.

sets used in this paper for experimentation and testing were selected based on how much lighting variation due to cloud movement they exhibit. The individual datasets are summarised in Table 1 with some of their properties.

3.3. Thinning and saturation removal

An average dataset of 40 min duration contained approximately 100 GB of raw 16 bit data, which expands to 200 GB when radiometrically corrected due to the use of a 32 bit floating point representation. Experiments were performed on data that were thinned line scan wise by a factor of 100, to allow faster processing of experimental iterations (while still retaining sufficient spatio-temporal coverage). High resolution snapshots such as Fig. 2(b) and voxelised maps such as those in Fig. 9 were produced using the full unthinned dataset.

There are also several instances in the data where pixels saturated (see Table 1). Scan lines where any bands in the reference target spectra were equal to the maximum value of the range of the sensor were excluded from any calculations or statistics due to saturation, the processing of which is outside the scope of this paper. It should be noted that this was done to allow calibration with a continuously visible target. As most of the non-target pixels were not saturated, compensation methods that do not require continuous target readings, such as the subspace based method tested in this paper, can still be used.

3.4. At-sensor radiance calculation

Raw hyperspectral imaging data from the Pika II camera come in digital counts. To account for the effects of non-uniform lens transmittance and sensor quantum efficiency, both spatially and spectrally, these should first be converted to sample at-sensor radiance (L_{Sample}), as follows (Suomalainen et al., 2014),

$$L_{Sample}(\lambda) = \frac{DN_{Sample}(\lambda) - DN_{SampleDarkCurrent}(\lambda)}{DN_{FlatField}(\lambda) - DN_{FFDarkCurrent}(\lambda)} L_{FlatField}(\lambda), \quad (1)$$

where DN_{Sample} and $DN_{SampleDarkCurrent}$ are raw digital number measurements of the sample and nearest dark current respectively. $DN_{FlatField}$ and $DN_{FFDarkCurrent}$ are flat field digital number measurements and corresponding dark current. Flat field images were acquired by averaging approximately 10 s of data in an integrating sphere, where the light level was set to about 95% of saturation. The corresponding dark current was measured in the same environmental conditions. The integrating sphere's detector also provides internal radiance measurements at each wavelength, represented by $L_{FlatField}(\lambda)$. It should be noted here that the camera's parameters,

including integration time, offset, gain and lens aperture must be the same for sample, dark current, and flat field measurements.

3.5. Reflectance quantities

In most natural light situations, including satellite, aerial and ground based sensing, incoming electromagnetic flux is hemispherical and reflected flux is conical, where the solid angle depends on the instantaneous field of view (IFOV), which is known as the hemispherical conical reflectance factor (HCRF) (Schaepman-Strub et al., 2006). However, where reflected solid angles are very small, as is the case with many imaging sensors, including the one used in this paper, the HCRF closely approximates the hemispherical directional reflectance factor (HDRF) (infinitesimally small reflected cone solid angle) (Schaepman-Strub et al., 2006). All reflectance quantity results in this paper refer to the HCRF or HDRF for the geometric configuration (zenith and azimuth angles) at that point in time. The target reflectance panels are assumed to be Lambertian, and consequently their laboratory reflectance measurements were used to directly calculate the illumination portion of the at-sensor radiance spectrum.

3.6. Reflectance and illumination calculation using reference panel

Benchmark (ground truth) sample reflectance (R_{Sample}) measurements were obtained per line scan with the permanently mounted reference panel with the following equation,

$$R_{Sample}(\lambda) = \frac{L_{Sample}(\lambda)}{L_{Ref}(\lambda)} R_{Ref}(\lambda), \tag{2}$$

where R_{Ref} and L_{Ref} are reference panel reflectance and radiance values. This assumes that path radiance is negligible due to the sensor's close proximity to the ground. We implemented an algorithm that automatically extracted ten pixels from the centre of the "mid grey" step of the QPcard 102 reference panel (which varied spatially due to swing and vibration), and averaged them to obtain L_{ref} for each line scan. Calibration target reflectance (R_{Ref}) was measured prior to data collection using an ASD Fieldspec 3 spectrometer with controlled halogen illumination.

From Eq. (2), one can deduce that the incident illumination can be estimated using

$$L_{Illum}(\lambda) = \frac{L_{Ref}(\lambda)}{R_{Ref}(\lambda)}. \tag{3}$$

As mentioned in Section 3.5, since we are actually measuring HCRF >(≈HDRF), L_{Illum} represents the amount of light that would be reflected by a perfectly Lambertian and reflective surface in the direction of the camera, given the hemispherical illumination at the time. It therefore does not represent the full hemispherical electromagnetic flux incident upon the reference panel surface. Computing illumination for every single line scan individually is referred to as REF, and is considered to be the ground truth benchmark against which all other methods are compared.

In addition, other simplistic illumination compensation approaches were simulated using partial reference target data. For the most basic approach, the illumination spectrum was fixed for a whole dataset, based on the first line scan reference (referred to as CONST for brevity). This emulates a field protocol where only one initial reference measurement is taken. Linear interpolation was included as an intermediate method that fits between the above two methods in terms of expected accuracy. Calibration target samples were extracted at different fixed intervals, and the spectra were interpolated with (Suomalainen et al., 2014)

$$L_{Ref}(\lambda) = \frac{t - t_{Ref_1}}{t_{Ref_2} - t_{Ref_1}} L_{Ref_1}(\lambda) + \frac{t_{Ref_2} - t}{t_{Ref_2} - t_{Ref_1}} L_{Ref_2}(\lambda), \tag{4}$$

where t_{Ref_1}, t_{Ref_2} and t are timestamps for the first, second and interpolated measurements, and L_{Ref_1}, L_{Ref_2} and L_{Ref} are radiance spectra vectors for the first, second and interpolated measurements. Saturated target samples were simply removed, with the exception of the first and last samples, which were moved to the first or last unsaturated instance respectively. For comparison, several target extraction intervals were simulated from the continuously visible reference, emulating a field protocol where a reference target might be periodically placed on the ground rather than being permanently mounted to the vehicle. Interpolation methods are referred to as INT-n, where n is the interval between target readings. A special case, INT-BE refers to interpolating between the first and last (non-saturated) line scan only. The interpolation intervals simulate manual placement of reference targets on the ground at regular intervals. Shorter intervals result in more manual labour for the operators.

3.7. Log subspace illumination and reflectance recovery

While several subspace illumination and reflectance recovery variations have been proposed, the method described as "REGLOGSEP" by Drew and Finlayson (2007) was implemented, due to its computational advantages, as described in Section 2.3. For brevity, in this paper it is referred to as LOGSEP. For this section we also adopt that paper's naming convention as follows, which denotes absolute non-logarithmic quantities with hats.

$$\begin{aligned} \hat{C}(\lambda) &= L_{Sample}(\lambda) & C(\lambda) &= \log(L_{Sample}(\lambda)) \\ \hat{E}(\lambda) &= L_{Illum}(\lambda) & E(\lambda) &= \log(L_{Illum}(\lambda)) \\ \hat{S}(\lambda) &= R_{Sample}(\lambda) & S(\lambda) &= \log(R_{Sample}(\lambda)) \end{aligned} \tag{5}$$

Ignoring the effect of upward atmospheric transmittance and path scattered radiance, which is reasonable for very low altitude applications, the at-sensor sample radiance $\hat{C}(\lambda)$ can be modelled as

$$\hat{C}(\lambda) = \hat{E}(\lambda)\hat{S}(\lambda) \tag{6}$$

The foundation of the subspace model is that both $\hat{E}$ and $\hat{S}$ can be approximated by a much smaller number of basis vectors than the dimensionality of the spectra. Therefore we have

$$\hat{E}(\lambda) = \sum_{i=1}^{m} \hat{\epsilon}_i \hat{E}_i(\lambda), \tag{7}$$

$$\hat{S}(\lambda) = \sum_{j=1}^{n} \hat{\sigma}_j \hat{S}_j(\lambda), \tag{8}$$

where m and n are the number of illumination and reflectance basis functions respectively, $\hat{E}_i$ and $\hat{S}_j$ are the individual basis functions for illumination and reflectance respectively, and $\hat{\epsilon}_i$ and $\hat{\sigma}_j$ are the corresponding coefficients.

Drew and Finlayson (2007) take the logarithm of $\hat{C}$. Dropping the hats for logarithmic quantities, this turns the at-sensor radiance into a sum of reflectance and illumination, changing Eq. (6) into

$$C(\lambda) = \sum_{i=1}^{m} \epsilon_i E_i(\lambda) + \sum_{j=1}^{n} \sigma_j S_j(\lambda), \tag{9}$$

The optimum coefficients ϵ_i and σ_i can be found by solving the following equation,

$$\mathbf{A}\alpha = \mathbf{h} \tag{10}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{M} & \mathbf{N} \\ \mathbf{N}^{\mathbf{T}} & \mathbf{P} \end{bmatrix}, \qquad \alpha = \begin{bmatrix} \epsilon \\ \sigma \end{bmatrix}, \qquad \mathbf{h} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix} \tag{11}$$

where,

$$\mathbf{M} = \mathbf{E}^{\mathbf{T}}\mathbf{E}, \quad \mathbf{N} = \mathbf{E}^{\mathbf{T}}\mathbf{S}, \quad \mathbf{O} = \mathbf{N}^{\mathbf{T}}, \\ \mathbf{P} = \mathbf{S}^{\mathbf{T}}\mathbf{S}, \quad \mathbf{f} = \mathbf{E}^{\mathbf{T}}\mathbf{L}, \quad \mathbf{g} = \mathbf{S}^{\mathbf{T}}\mathbf{L} \tag{12}$$

Eq. (10) can be solved in the standard way by performing the matrix product of the inverse of **A** on the left of both sides to obtain α and thus ϵ and σ. See Drew and Finlayson (2007) for details on the derivation of Eq. (10).

E and **S** are $s \times m$ and $s \times n$ matrices containing the respective basis vectors in their columns, where s is the number of bands. Correspondingly, ϵ and σ are the coefficient vectors for illumination and reflectance basis functions.

Drew and Finlayson (2007) further refine the method by including an additional regression step, which serves to take the initial separation result and guide it using a known set of at-sensor radiances, which is the set of all training illuminant and reflectance combinations. A regularisation step is also added by penalising large norm solutions during the regression. The regression and regularisation are achieved with the following equations:

$$\mathbf{W} = \frac{1}{(m+n)} \sum diag(\tilde{\alpha}\tilde{\alpha}^{\mathbf{T}})\mathbf{I}_{(m+n)}, \tag{13}$$

$$\mathbf{Q} = \tilde{\epsilon}\tilde{\alpha}^{\mathbf{T}}(\tilde{\alpha}\tilde{\alpha}^{\mathbf{T}} - \lambda\mathbf{W})^{-1}, \tag{14}$$

$$\mathbf{R} = \tilde{\sigma}\tilde{\alpha}^{\mathbf{T}}(\tilde{\alpha}\tilde{\alpha}^{\mathbf{T}} - \lambda\mathbf{W})^{-1}, \tag{15}$$

where $\mathbf{I}_{(m+n)}$ is the $(m+n) \times (m+n)$ identity matrix and λ is a tuning parameter. $\tilde{\epsilon}, \tilde{\sigma}$ and $\tilde{\alpha}$ are the sets of $m \times N$ illuminant, $n \times N$ reflectance, and $(m+n) \times N$ combined coefficients respectively for all N illuminant and reflectance combinations.

The resulting regression matrices can be applied to any coefficient estimate α with,

$$\alpha_{\mathbf{new}} = \mathbf{T}\alpha, \quad \mathbf{T} = \begin{bmatrix} \mathbf{Q} \\ \mathbf{R} \end{bmatrix} \tag{16}$$

Note that **T** differs from Eq. 15 of Drew and Finlayson (2007), due to an apparent typographical error. Two methods of training data selection were used: from line scans that are evenly spaced in time, and line scans that were manually selected at times where the greatest lighting variation occured. The first simulates placing a calibration target on the ground at regular intervals (as with the INT-n methods). This will be referred to as LOGSEP-n, where n is the interval between readings. The second method simulates a policy where training data are collected manually when qualitative lighting changes are observed by operators in the field, spanning as much variation as possible, while minimising the number of training samples (only 5–8 line scans per dataset). This method will be referred to as LOGSEP-M. Full line scans were used as input training data, where reference panel pixels were used to obtain both illumination and reflectance training spectra. Basis functions were obtained separately for illumination and reflectance spectra, using Singular Value Decomposition (SVD) on the logarithm of the training spectra. To ensure tractability, the number of illumination and reflectance spectra was thinned to a maximum of 1000 each (i.e. $N_{max} = 10^6$) for determination of the regression matrices only. Scan lines where the reference target was saturated were removed from the training data.

Reflectance using the LOGSEP methods can be recovered in two ways: directly from the reflectance base spectra coefficients or by dividing the at-sensor radiance by the recovered illumination spectrum. In our experiments we consider both groups of results. We differentiate the two groups of methods as LOGSEP and LOGSEP-IND (see Table 2).

Table 2
Summary of reflectance and illumination calculation methods.

Short Name	Description
REF	Illumination spectrum derived from target in view at each linescan. Benchmark reference method used to compare all other methods against (for example when calculating spectral angle)
CONST	Illumination spectrum derived from first target observation and assumed to be constant for entire duration
INT-BE	Linear interpolation between illumination spectra at the start and end of the datasets
INT-n	Linear interpolation of illumination spectra at n second intervals
LOGSEP-n	Direct subspace based separation with training spectra at n second intervals
LOGSEP-M	Direct subspace based separation with a few hand selected training spectra
LOGSEP-IND-n	Like LOGSEP-n but using recovered illumination to indirectly obtain reflectance. Only applies to reflectance recovery
LOGSEP-IND-M	Like LOGSEP-M but using recovered illumination to indirectly obtain reflectance. Only applies to reflectance recovery

When taking the logarithm of spectral values, very low, negative or zero values must be dealt with. For all processing in this paper, any input spectra were set to an appropriate floor value prior to processing with LOGSEP.

3.8. Georeferencing, voxelisation and mapping

One important advantage of acquiring hyperspectral data with a mobile ground based platform is the ability to create very high resolution orthorectified maps (Abd-Elrahman et al., 2016). In this paper, this enables us to visually demonstrate the result of compensating for lighting changes. Georeferencing allows the placement of hyperspectral pixels in real world 3D coordinates in order to generate orthorectified maps. Each pixel is represented as an infinite ray in Cartesian space using the standard pinhole camera model. The rays are georeferenced using the vehicle's location and pose estimates from a Novatel SPAN RTK/GPS/INS. The intersection of the rays and local horizontal ground planes define the location of each point in world coordinates, where the local planes are defined by the wheel contact points at each moment in time. This assumption is reasonable where the ground does not slope significantly across the span of the sensor/ground footprint, which was the case for all trials in the this study. It does not take into account plant height, but the accuracy was sufficient for visual interpretation of results. The maps were generated by grouping the resulting 3D points into a regular 3 cm grid (column voxelisation), where the mean was calculated on true colour values (at appropriate red, green and blue wavelengths) and NDVI (see Section 3.9).

3.9. Normalised difference vegetation index

In some of the experiments, the Normalised Difference Vegetation Index (NDVI), calculated with Eq. (17), is used as an example to demonstrate a practical output measure that is very commonly used in precision agriculture (PA) and remote sensing in general.

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{17}$$

Its applications include quantifying leaf area index (LAI), percent vegetation cover, intercepted photosynthetically active radiation (IPAR), and green biomass (Zhao et al., 2007). The NDVI was originally designed to be used with broadband sensors, such as multispectral cameras, whereas its application with narrow band

spectroscopy and hyperspectral imaging requires specific red and NIR bands to be selected. In addition, to reduce noise, it is often necessary to average (bin) adjacent bands to maximise the signal to noise ratio (SNR) of the input values, which requires an additional parameter, averaging width.

All NDVI values in this paper were obtained using red and NIR values at wavelengths 670 nm and 800 nm respectively, which is one of the most common combinations in the literature (Slonecker, 2012; Wu et al., 2008). An averaging width of 10 nm was found experimentally to provide adequate amelioration of noise. In practice different parameters may be chosen depending on the desired outcome.

4. Results

In this section experimental results of the reflectance retrieval methods (Table 2) are outlined in detail. We demonstrate the significant effect lighting variation can have, and we assess the effectiveness of the various compensation methods. The overall accuracy of each method was measured using its spectral angle (SA) with REF for recovered spectra within each of the tested datasets. These are also plotted over time to demonstrate the stability of each method. Maps are presented to illustrate the results and show the improvement of illumination compensation when using ratios and indices. The possibility of compensating for lighting variation in one dataset using data from another or several others is also evaluated using the LOGSEP recovery method.

4.1. Effect of lighting variation on hyperspectral data

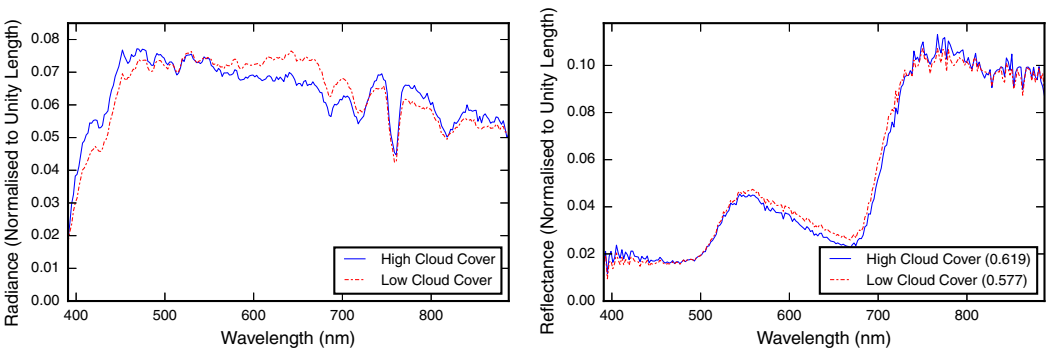
The light incident upon the ground is a combination of direct sunlight and light scattered through the atmosphere (skylight) and from other objects. Skylight is blue shifted compared to direct sunlight, because the atmosphere scatters the light more at lower wavelengths than at higher wavelengths. When the amount of direct sunlight is reduced, for instance by cloud cover, the total illumination incident upon the ground is also blue shifted. This can be seen in Fig. 3(a), where the high cloud cover illumination spectrum falls off more steeply from 500 to 700 nm, compared to the low cloud cover spectrum. The two spectra were taken from the same dataset (ML-0), highlighting the additional difficulty experienced when operating under changing cloud cover conditions.

Fig. 3(b) demonstrates the effect on a recovered reflectance spectrum if illumination is not determined accurately. The same sample reflectance spectrum was multiplied by the two illumination radiance spectra to simulate at-sensor sample radiance spectra. The two radiance spectra were then divided by one illumination spectrum, simulating using one constant illumination measurement for all data (i.e. CONST compensation, see Table 2), which does not account for lighting changes. A clear discrepancy can be seen in almost all wavelengths, which affects the outcome of any further analysis, such as the calculation of Normalised Difference Vegetation Index (NDVI) (see Section 3.9). The NDVI variation over time due to changes in lighting can be observed in Fig. 4 for the spectrum in Fig. 3(b) for every point in time throughout the ML-0 dataset (again compensated with CONST). NDVI values peak at about 0.63 and dip below 0.57, changing by over 9% within only minutes. Closer examination of the illumination spectra in Fig. 3(a) suggests that an even greater effect may be seen in indices, classification and other analysis that rely more heavily on the wavelengths approximately below 500 nm, between 600 and 700 nm, and between 740 and 800 nm. Subtle effects in plant spectra can easily be overshadowed by environmental lighting variation, which highlights the necessity for accurate lighting compensation.

4.2. Choice of basis functions and regularisation parameter

The LOGSEP method is parameterised by three variables that must be selected appropriately for the data. Fig. 5(a), (b) and (c) illustrates the effect of varying the number of illumination and reflectance basis functions on the LOGSEP illumination estimate accuracy, shown as average spectral angle from REF, and computation time. To generate the results, 235 scan lines (152280 pixels) were taken from evenly spaced locations in the ML-0 dataset. Error levels out after approximately 12 reflectance and three illumination basis functions, while computation time increases with the number of basis functions. This presents a trade-off, and the relatively small increase in accuracy gained by increasing the number of basis functions further was not considered worth the corresponding computational cost. Consequently, all subsequent experiments were done based on three illumination and 12 reflectance basis functions.

Fig. 5(d) shows the effect of the regularisation parameter λ on accuracy. Overall the effect is minimal, but after about $\lambda = 10^{-6}$



(a) Illumination Spectra Example (b) Reflectance Spectra Example (NDVI in brackets)

Fig. 3. Sample (a) illumination and (b) reflectance spectra demonstrating the significant differences between a low and high cloud cover scenario. A representative plant sample reflectance spectrum was multiplied by the illumination spectra in (a), and retrieved using one constant reference (CONST) to yield the reflectance spectra in (b). All spectra were taken from the same dataset (ML-0, see Table 1) at the times indicated in Fig. 4. The spectral vectors were normalised to unity length to allow a direct shape comparison (in reality, the high cloud cover spectrum is significantly lower in magnitude). Note: The plots use different *y*-axes as they are individually normalised spectra, and therefore a direct comparison between plots is not meaningful.

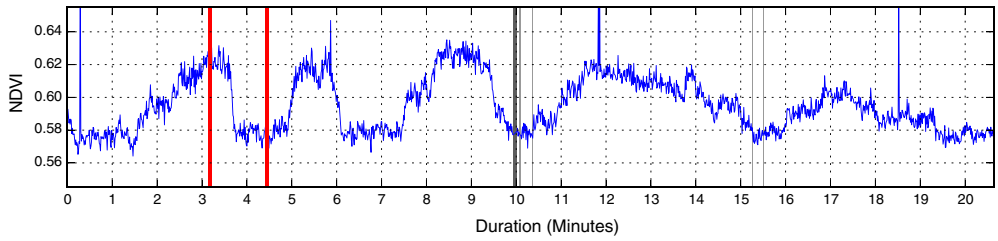


Fig. 4. NDVI values plotted over time for a representative plant sample reflectance spectrum, multiplied by illumination spectra in a high lighting variation dataset (ML-0, see Table 1), and retrieved using one constant reference (CONST). The two red vertical lines indicate the times of the spectra in Fig. 3. The vertical grey regions indicate locations where the reference target was saturated and which were therefore excluded from training, interpolation and calculation of any statistics. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

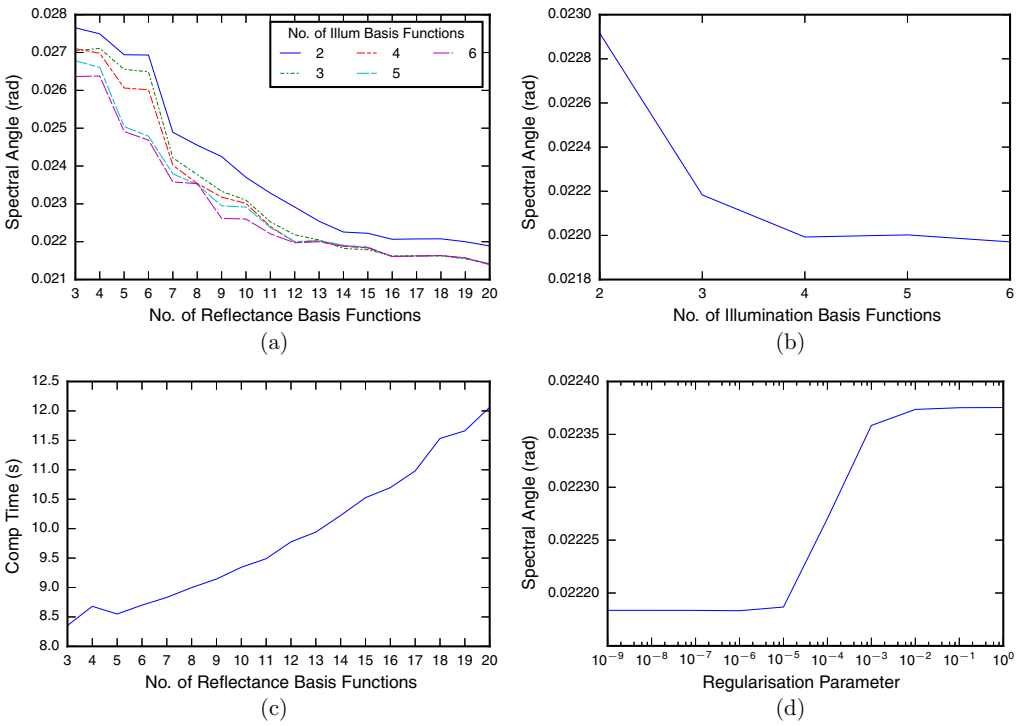


Fig. 5. Basis function and regularisation parameter test results. In (a), results obtained with varying numbers of reflectance basis functions are plotted against spectral angle (as measured against REF). Each curve represents a different number of illumination basis functions. The effect of the number of illumination basis functions on accuracy is shown in (b), plotted for a fixed number of reflectance basis functions of 12. In (c), computational time is graphed for a varying number of reflectance basis functions (illumination basis functions fixed at three). In (d), the effect of varying the regularisation constant with respect to accuracy is shown (note the logarithmic x axis).

Table 3
Spectral angles to REF ($\times 10^{-2}$ radians).

Dataset	CONST	INT-BE	INT-180s	INT-30s	LOGSEP-30	LOGSEP-M	LOGSEP-IND-30	LOGSEP-IND-M
NMAD-1	6.55	7.73	3.92	3.71	8.25	8.39	4.54	4.79
DDPA-2	4.09	3.90	3.72	3.63	8.72	8.69	3.65	3.88
CDA-0	4.98	3.32	2.57	2.47	6.20	6.32	2.98	3.18
DDPA-1	3.12	3.23	2.86	2.73	7.12	7.12	2.78	3.03
ML-0	2.80	2.71	2.29	1.46	4.50	4.59	2.08	2.29
DBVL-0	3.00	4.07	2.23	1.63	5.51	5.61	2.37	2.53
SUL-0	3.22	2.77	1.99	1.57	5.67	5.35	2.47	2.63
Average	3.97	3.96	2.80	2.46	6.56	6.58	2.98	3.19

the error increases more noticeably, and thus all subsequent experiments were performed with this value of λ.

4.3. Recovery accuracy

In order to evaluate the accuracy of retrieved reflectances, the SA of the resulting reflectance spectra for each calibration method were calculated against the continuous reference panel result REF (see Section 3.6) as an error measure. All line scans were calibrated using the methods outlined in Table 2. Average SA error results are shown in Table 3. In rare cases, where INT-*n* reference readings coincide with corrupted calibration target readings (see Fig. 8), SA results within the affected time intervals were omitted from averaging for all methods in Table 3 to facilitate a fair comparison. Similarly, affected line scans were removed from LOGSEP-*n* training data.

The two interpolation methods, perform best on average, though the LOGSEP-IND methods produce comparable results. LOGSEP-M and LOGSEP-IND-M perform similarly to LOGSEP-30 and LOGSEP-IND-30 respectively.

An example of a typical radiance spectrum and its recovered illumination and reflectance are shown in Fig. 6, along with spectral angle values in the legend. Both LOGSEP-30 and LOGSEP-M qualitatively appear to closely model the reference illumination and reflectance. LOGSEP is compared to LOGSEP-IND in Fig. 6(d). Both are similar to each other, yet directly obtaining reflectance as a combination of base spectra (LOGSEP) inherently tends to smooth results, while LOGSEP-IND retains more high frequency detail. Fig. 7 illustrates an example reflectance spectrum that was poorly recovered using LOGSEP. In contrast, LOGSEP-IND appears to provide better recovery performance here.

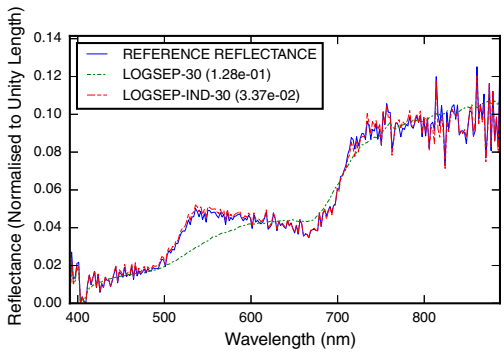


Fig. 7. Example of a poorly recovered reflectance spectrum by LOGSEP. LOGSEP-IND is able to provide a significantly more accurate result here. Spectral angles from REF in radians are shown in brackets.

To demonstrate the evolution of the errors over time, illumination spectral angle values were averaged per line scan and plotted over time in Fig. 8 for four representative example datasets, two with heavy and two with light illumination variation. The interpolated method's SA drops to zero every 30 s, at which time the

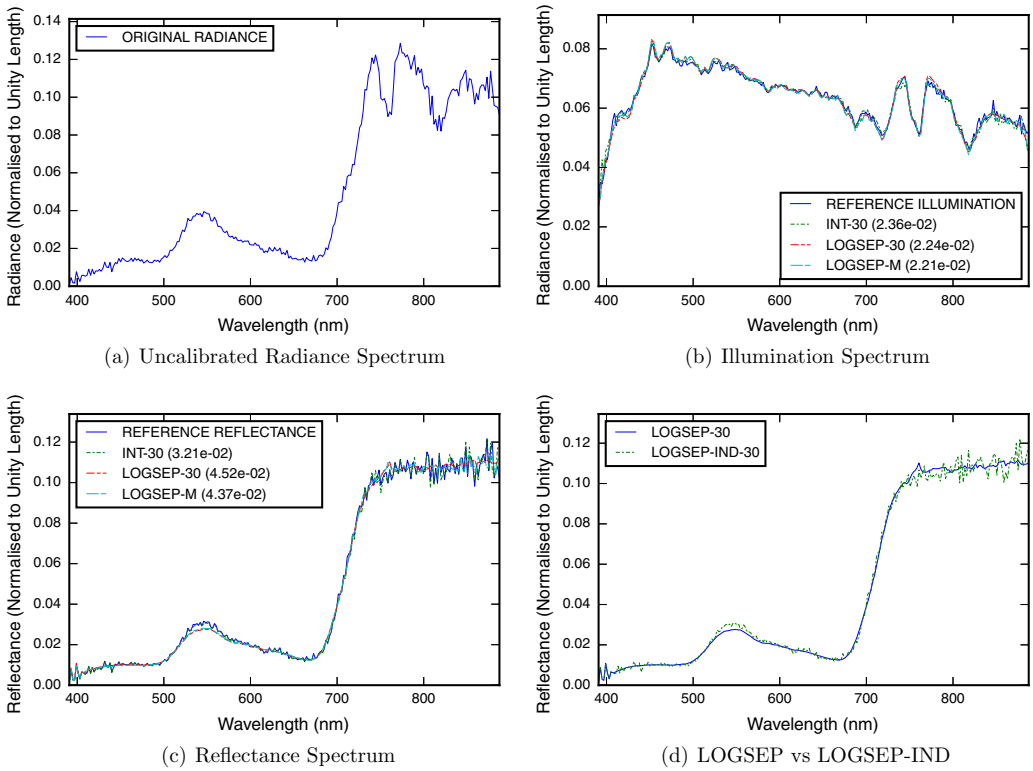
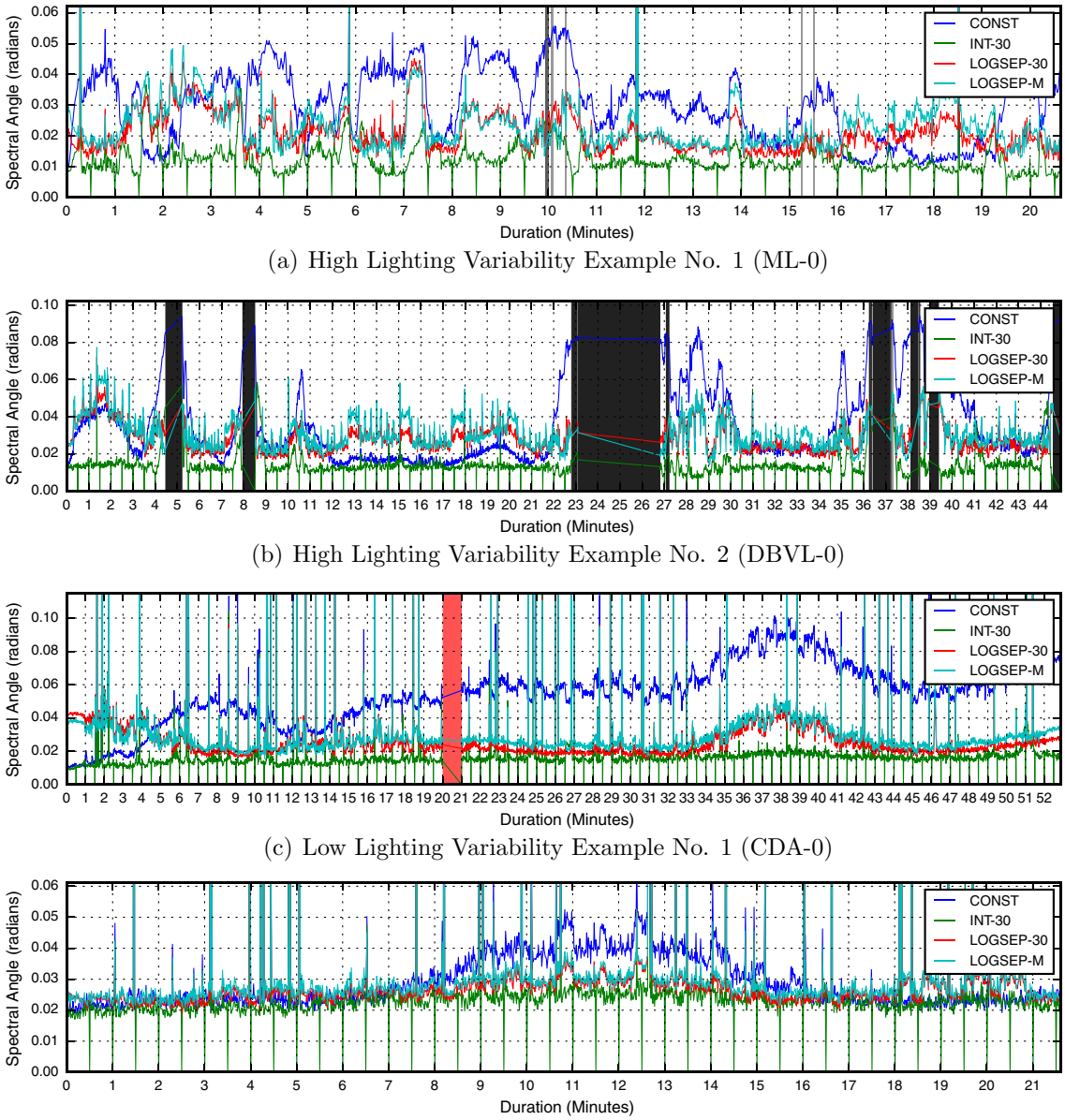


Fig. 6. Reflectance (c and d) and illumination (b) extraction example from a typical at-sensor radiance sample (a). In (d), note the inherent smoothing that occurs when extracting spectra directly with LOGSEP. The noise present in the original reflectance spectrum most likely unfairly inflates LOGSEP reflectance spectral angles. Therefore, illumination and LOGSEP-IND reflectance results are a better gauge of recovery accuracy. Spectral angles from REF in radians are shown in brackets. Note: The plots use different *y*-axes as they are individually normalised spectra, and therefore a direct comparison between plots is not meaningful.



(a) High Lighting Variability Example No. 1 (ML-0)

(b) High Lighting Variability Example No. 2 (DBVL-0)

(c) Low Lighting Variability Example No. 1 (CDA-0)

(d) Low Lighting Variability Example No. 2 (DDPA-2)

Fig. 8. Scan line average spectral angle values calculated for illumination extraction results from CONST, INT-30, LOGSEP-30, and LOGSEP-M against the reference REF. In (a) and (b), environmental lighting was bright and sunny with frequent moving clouds resulting in sudden significant changes. In (c) and (d), however, cloud cover was heavy and variation was limited. The vertical grey regions indicate locations where the reference target was saturated and which were therefore excluded from training, interpolation and calculation of any statistics. The large spikes are caused almost exclusively by the following: excessive swaying of the calibration target, pushing it outside of the field of view of the camera, or confusion of the extraction algorithm with separate calibration targets that were placed directly on the ground (but not used in this paper). All of these affect each of the methods equally, except in rare cases where an INT-*n* reference coincides with one of the spikes, in which case they were excluded, as indicated on the plot as vertical red regions. Spikes were also removed from LOGSEP-*n* training data.

illumination spectrum is exactly equal to the continuous method (REF). It also maintains the lowest spectral angle for the majority of the duration. In most cases, the two LOGSEP methods produce an error that remains close to INT-30 and well below CONST, particularly in Fig. 8(c), whereas in Fig. 8(a) the LOGSEP methods fluctuate more significantly.

Fig. 9(a) shows an uncompensated example of a true colour hyperspectral map of ML-0. Clear lighting differences are visible in several places as clouds moved through the scene on the day of acquisition. In Fig. 9(b), continuous calibration (REF) compensated NDVI values are plotted for the same field, clearly showing differences in vegetation densities, but no obvious illumination

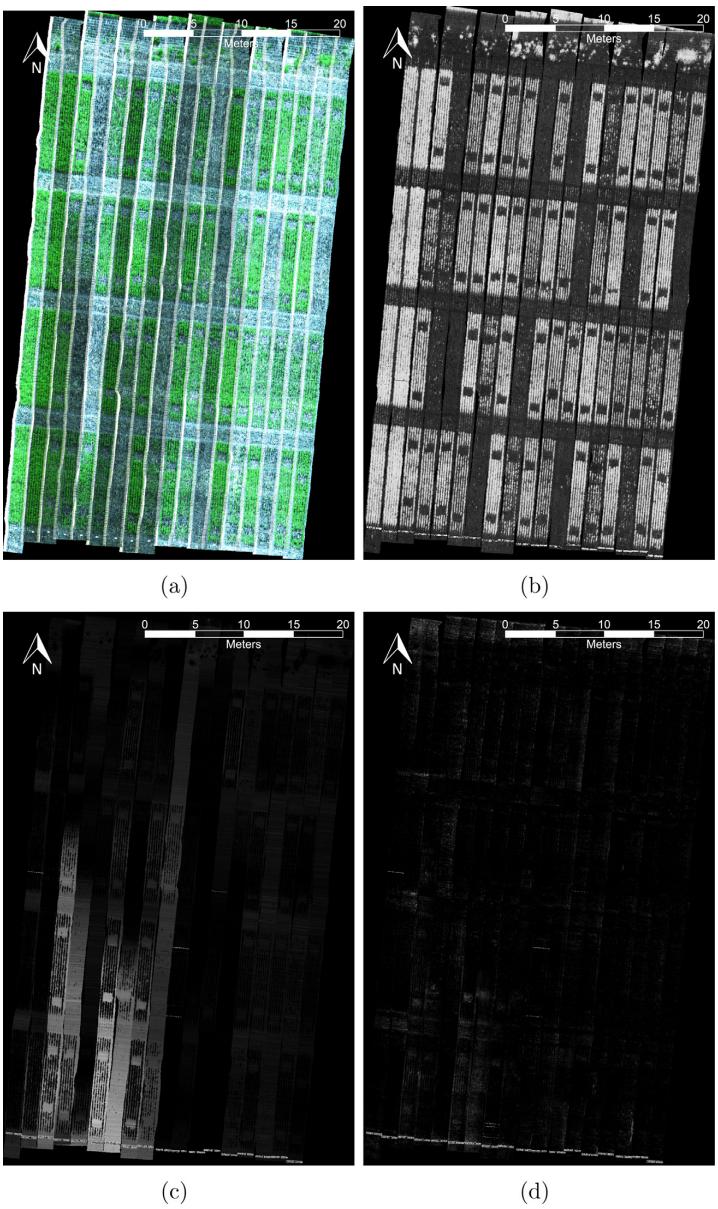


Fig. 9. Orthorectified hyperspectral maps for ML-0. An uncorrected true colour hyperspectral map is shown in (a). Absolute lighting differences due to moving clouds can be seen in several places, as lighter and darker regions along the rows. An NDVI map is shown in (b), based on continuous calibration compensated data (REF). (c) illustrates the square difference between the NDVI results of CONST and REF compensated NDVI, while (d) gives the difference between LOGSEP and REF compensated NDVI. Both (c) and (d) are scaled to 0–0.0051. Difference values were scaled by the same amount in both images. Because CONST does not correct for variations in illumination that occur in the scan, errors occur at the times during the scan when clouds moved past the sun. Given the progressive scan motion of the vehicle, this causes a strong spatial correlation of the errors, which can cause ambiguity. In the LOGSEP compensated image, these differences are much smaller, as quantified in Table 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

effects, as expected. Fig. 9(c) and (d) compares REF NDVI values with CONST and LOGSEP-M results respectively by taking the square difference and mapping it. The CONST result is clearly much more susceptible to lighting variation when compared to the LOGSEP-M square differences, which are barely visible at the same scale.

4.4. Ratio stability

For analysis and mapping, it is very common to calculate vegetation indices, such as NDVI, from hyperspectral imaging data. Compensation for illumination should aim to ensure accuracy when computing these indices, which are generally ratios of

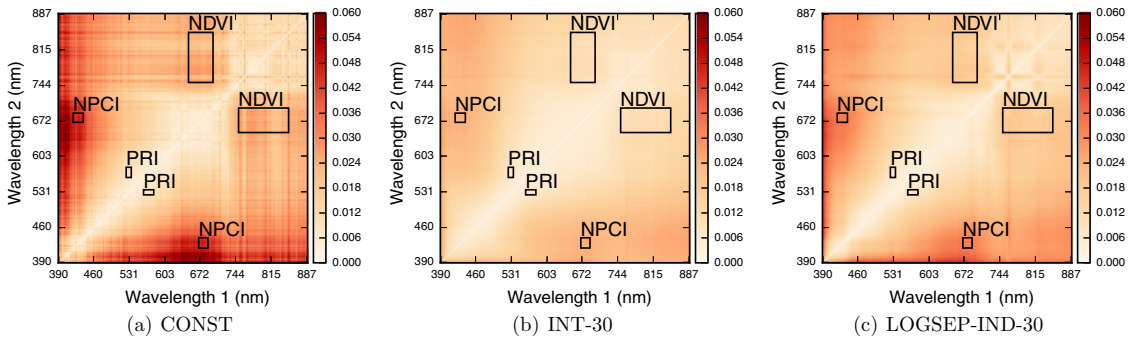


Fig. 10. RMS error matrices for ratios calculated on the ML-0 (high variation) dataset. Variation affected areas occur in the blue, between 390 and 450 nm, especially in combination with 600–700 nm. The rectangles bound common wavelength combinations used to calculate NDVI, which are located in a moderately affected region above 740 nm.

reflectance values at one wavelength to another. To illustrate how the various calibration methods perform, ratios were calculated for all combinations of wavelengths. A generic normalised ratio of reflectances at two wavelengths is expressed as

$$r(\lambda_1, \lambda_2) = \frac{R(\lambda_2) - R(\lambda_1)}{R(\lambda_2) + R(\lambda_1)}, \tag{18}$$

where $R(\lambda)$ is the reflectance value at wavelength λ. For appropriate selections of λ_1 and λ_2 this equation covers a number of vegetation metrics from the literature, including NDVI, normalised pigment chlorophyll index (NPCI), and photochemical reflectance index (PRI) (Slonecker, 2012).

The root mean square error (RMSE) for each ratio was calculated for CONST, INT-30 and LOGSEP-IND-30, with respect to REF. The data are the same used in Section 4.3, and results are shown as a matrix in Fig. 10.

As expected, CONST exhibits greater RMSE values for a large proportion of wavelength combinations. As demonstrated previously, interpolating every 30 s generally corrects for illumination changes well, while LOGSEP-IND-30 performed much better than CONST, as expected, but slightly worse than interpolation. The greatest deviation can be seen in ratios that include wavelengths at the blue extremes (<470 nm), which can be attributed to a greater amount of noise, though the effect is not as bad at the near infra-red (NIR) end of the spectrum (>800 nm). Importantly, however, the ratio region that commonly corresponds to NDVI has been improved markedly by both LOGSEP and INT.

4.5. LOGSEP cross dataset testing

An interesting application of subspace based illumination and reflectance extraction is using a prior known set of training spectra to build a subspace model that can be used on any data where a reference target is not available. Previous sections in this paper have focused on both training and testing on separate data from within the same dataset. This relates to a field protocol where training illumination and reflectance spectra are obtained during the acquisition of each dataset, and compensation is performed as a post-processing step. This section investigates the results from training on illumination and reflectance spectra from one dataset and testing on a different one, which would mean no reference target is needed once training is complete.

Tables 4 and 5 summarise the illumination extraction results with LOGSEP-30 and LOGSEP-M respectively, when trained and tested on all possible dataset combinations. In most cases, best or close to best performance is achieved when training and test spectra came from the same dataset (diagonal). It can also be observed that datasets that came from similar conditions result in better performance. For example, DBVL-0, ML-0 and SUL-0 combinations (high variation, sunny) perform well, as do DDPA-2, DDPA-1 and CDA-0 combinations (low variation, cloudy). Conversely, training on a low variation, cloudy dataset and testing on a high variation, sunny one generally produces lower accuracy.

To explore this further, Fig. 11 compares the highest and lowest magnitude LOGSEP-30 illumination training spectra from representative sunny and cloudy datasets (ML-0 and CDA-0). Both cloudy spectra are blue shifted with respect to the sunny ones (higher relative magnitude between 400 and 500 nm, and lower relative magnitude between 600 and 700 nm). As a result, LOGSEP, and in particular the regression step, can be expected to perform poorly when attempting to reproduce the original illumination (and by extension reflectance) spectrum.

To address this, we combined the training data sets from one higher overall illumination and one lower overall illumination dataset, for the purpose of training on a more diverse range of data.

Table 4
LOGSEP-30 cross dataset illumination extraction spectral angles ($\times 10^{-2}$ radians)

		Source of Training Data						
		NMAD-1	DDPA-2	CDA-0	DDPA-1	ML-0	DBVL-0	SUL-0
Test Data	NMAD-1	4.47	6.79	5.91	5.51	6.72	8.80	9.67
	DDPA-2	5.72	3.31	4.09	4.07	6.37	6.77	7.88
	CDA-0	5.15	4.52	3.08	4.50	4.98	5.14	5.91
	DDPA-1	3.77	3.31	3.61	2.62	5.94	7.01	8.06
	ML-0	5.00	8.84	5.47	7.71	2.17	3.11	3.23
	DBVL-0	5.87	7.38	4.55	7.12	3.92	2.85	2.92
	SUL-0	6.17	8.45	5.35	8.07	4.10	3.14	2.80
	Average	5.17	6.08	4.58	5.66	4.89	5.26	5.78

Table 5
LOGSEP-M cross dataset illumination extraction spectral angles ($\times 10^{-2}$ radians).

		Source of Training Data						
		NMAD-1	DDPA-2	CDA-0	DDPA-1	ML-0	DBVL-0	SUL-0
Test Data	NMAD-1	4.74	7.25	6.06	5.42	6.64	8.77	9.65
	DDPA-2	5.58	3.43	4.52	4.46	6.17	7.18	8.27
	CDA-0	5.10	4.63	3.32	4.67	4.98	5.37	6.03
	DDPA-1	3.86	3.60	4.17	2.84	5.66	7.21	8.07
	ML-0	5.09	9.48	5.10	7.43	2.36	3.09	3.19
	DBVL-0	5.77	7.54	4.40	7.04	4.24	3.07	3.07
	SUL-0	6.10	8.68	5.05	7.89	4.45	3.34	2.94
	Average	5.18	6.37	4.66	5.68	4.93	5.43	5.89

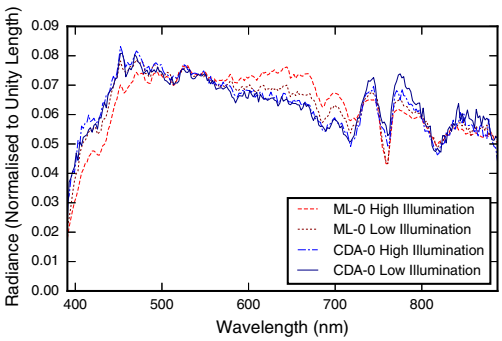


Fig. 11. Comparison of illumination spectra in high and low variation datasets.

As there are a large number of possible combinations, only three were selected on the following criteria:

- Range: DBVL-0 and DDPA-2 cover a large combined minimum to maximum range of illumination magnitude.
- Overlap: SUL-0 and CDA-0 have large combined illumination magnitude overlap.
- Best individual training sets (lowest average SA) combined: ML-0 and CDA-0 each produced the best average SA in their respective high variation and low variation groups.

The results are shown in Table 6. Compared to Tables 4 and 5, average scores have clearly improved. In most cases, individual accuracies have improved to approximately the better of the two datasets in the training combinations. For instance, consider ML-0-CDA-0-LOGSEP-30 SA results: Test dataset DDPA-2 has a score of 4.13×10^{-2} rad, which correlates well with the CDA-0 LOGSEP-30 training set score of 4.09×10^{-2} rad, and test dataset DBVL-0 has an SA of 4.10×10^{-2} rad, which also closely matches ML-0's LOGSEP-30 training set score of 3.92×10^{-2} rad.

5. Discussion

The results in Section 4 demonstrate that the LOGSEP methods are able to extract illumination spectra with an accuracy similar to short time interval interpolation methods. LOGSEP and LOGSEP-M provide very similar performance overall, suggesting that for LOGSEP to extract illumination and reflectance spectra accurately, it is not as important to have a large quantity of training data, but a comprehensive sampling of different illumination conditions.

It is important to note that all REF, CONST and INT illumination spectra are inherently smoothed, as reference panel readings were averaged over ten pixels. This in turn results in high frequency noise being transferred to reflectance when dividing into the radiance spectrum. On the other hand, reflectances obtained directly from LOGSEP (as a combination of reflectance base functions) exhibit a smoothing effect themselves. This can artificially inflate their spectral angle scores as computed against REF, resulting in the inferior results of LOGSEP-30 and LOGSEP-M seen in Table 3. On the other hand, LOGSEP *illumination* results are also smoothed by the limited number of basis functions, and so LOGSEP-IND (calculated by dividing illumination into radiance) provides a fairer quantitative comparison of reflectance results to REF, CONST and INT.

Additionally, there is also the possibility that if certain reflectance spectra are not well represented in the training data (e.g. uncommon non-vegetation spectra), LOGSEP will provide a poor reflectance estimate directly (as a combination of basis functions). Because illumination spectra are modelled in a lower dimensional subspace (three in this case), its spectrum is more strictly constrained, resulting in much better reflectance recovery with LOGSEP-IND, as shown in Fig. 7. For this reason, it is generally preferable to use LOGSEP-IND over LOGSEP.

Table 6
Combination cross dataset illumination extraction spectral angles ($\times 10^{-2}$ radians).

		Source of Training Data					
		Large Range		Best Ind. Perf.		Large Overlap	
		DBVL-0 DDPA-2 LOGSEP-30	DBVL-0 DDPA-2 LOGSEP-M	ML-0 CDA-0 LOGSEP-30	ML-0 CDA-0 LOGSEP-M	SUL-0 CDA-0 LOGSEP-30	SUL-0 CDA-0 LOGSEP-M
Test Data	NMAD-1	6.67	6.50	5.86	6.20	6.41	6.89
	DDPA-2	4.17	4.23	4.13	4.67	4.43	5.14
	CDA-0	3.78	3.77	3.28	3.60	3.36	3.75
	DDPA-1	4.15	4.07	4.04	4.60	4.27	4.83
	ML-0	3.44	3.67	2.81	2.45	3.40	3.12
	DBVL-0	3.34	3.58	4.10	3.93	3.38	3.16
	SUL-0	3.58	4.00	4.51	4.23	3.67	3.25
	Average	4.16	4.26	4.11	4.24	4.13	4.31

The results also demonstrate that it is possible to compensate for illumination changes in any datasets, where a reference target is not available, if a sufficient range of historical illumination training spectra is present. By combining training data from datasets with differing lighting conditions, LOGSEP was universally applicable to all datasets, in contrast to other methods that require in field reference measurements (e.g. Uto et al. (2013),Suomalainen et al. (2014), and Karpouzli and Malthus (2003)).

Based on the results, we are able to provide guidance on which compensation method and matching field protocol to use and when. If accuracy of data is of critical importance and it is practically feasible to permanently mount a calibration target to the platform, continuously imaging a calibration target panel in every line scan will lead to the best results (REF), as no other method was able to perfectly reconstruct the data. However, mounting the target correctly can be tedious and restrictive, or impracticable (e.g. for aerial applications). For example, high crop growth made it very difficult to mount the target while maintaining an optimal sensing geometry in our case, as the covers of the Ladybird needed to be adjusted. We also found that this approach increased the footprint of the vehicle, raised potential safety concerns and was prone to damage. In some cases, the target had to be completely removed to allow for taller crops. In general, it would be a challenge to universally cater for all possible crop heights with a single configuration. Another consideration is the loss of pixels where the target is in view. On some fields, uneven ground caused excessive swing, sometimes even moving the target out of view (as can be seen by the spikes in Fig. 8). Furthermore, as shown in Table 1, there are a number of occasions where the reference target was completely saturated (although the majority of vegetation/soil data were not saturated). This can be avoided in the field by adjusting the exposure, lens aperture and/or reflectance of the target itself, however, there is a trade-off, as signal to noise ratio (SNR) is reduced as the aperture or target reflectance is decreased in low light conditions. Alternatively a downwelling illumination sensor may be considered (e.g. Deery et al. (2014) and Kuusk and Kuusk (2010)), which introduces increased cost and complexity as previously discussed.

A next best alternative is to image reference panels in close succession (INT-x) and interpolate (as in Suomalainen et al. (2014)). This is also tedious and time consuming, as targets need to be manually placed on the ground in line with the intervals (which could correspond to the end of rows, for example). If shorter intervals cannot be maintained, accuracy will deteriorate as sudden changes in illumination are not captured.

On the other hand, field protocols based on the subspace model reflectance extraction method (LOGSEP) can provide a convenient alternative that avoids many of the aforementioned shortcomings. If no historical reference data are available, we have shown that LOGSEP can perform well when only a few separate ($\geqslant 6$) reference panel (and therefore both illumination and reflectance) samples are acquired from each field that is being scanned (LOGSEP-M). This would require operators to monitor conditions (visually or with a light meter, for instance) and image reference targets when lighting has changed significantly. It is important to capture as much lighting variation as possible. We have shown that by combining training samples from datasets with differing lighting conditions, illumination compensation is possible for a wider range of environmental illumination.

If historical illumination and reflectance spectra spanning a wide range of lighting conditions is available, no further in-field training samples are necessary. It would therefore be possible to compile a training set of illumination and reflectance spectra based on previously acquired data. No additional calibration panel measurements would be required, which carries several advantages. The complexity of the system is reduced, because no mounted panels or additional downwelling sensors are required. Completely autonomous operation is made possible, because there is no need to manually place reference panels on the ground, increasing convenience for operators, which is an important factor both for research and commercial applications. In addition, with the exception of a continuously imaged panel, all other reference panel methods require the extraction of reference measurements as a post processing step. A pre-trained system based on prior training spectra enables processing of pixels as they are acquired, allowing real time reflectance retrieval. An additional benefit is that compensation can be performed retrospectively on datasets where a reference target was not available. In contrast to approaches that utilise radiative transfer codes (e.g. Berk et al. (1999) and Qu et al. (2001)), specific atmospheric parameters and their variation with time do not need to be known or estimated explicitly when using LOGSEP.

6. Conclusions

This work examined the use of an unmanned ground vehicle (UGV) for acquiring large quantities of high resolution hyperspectral data of agricultural crops, and compensating for natural illumination variation, in order to retrieve surface reflectance. A previously introduced fast logarithm subspace method for reflectance and illumination extraction was employed and tested against more traditional approaches using reference target panels. We demonstrated that the new approach competes well with frequently (every 30 s) interpolated illumination spectra, while requiring far fewer reference readings. In light of this, we compared and recommended a number of in field policies to obtain appropriate data for the various correction methods. Our experiments show that calibration can be achieved with only six calibration target readings per continuous scan of 60 min. Furthermore, we show that it is possible to train a dataset entirely on historical data, which could allow compensation for lighting changes without any further in field reference data. Future work will validate this idea with a wider array of data in order to understand how a training set would be optimally compiled.

Acknowledgements

The work performed in this study was funded by the South Australian Research and Development Institute (SARDI). The development of the Ladybird and associated infrastructure has been supported by the Australian Centre for Field Robotics at The University of Sydney, AUSVEG and Horticulture Innovation Australia through project VG12104: An Intelligent Farm Robot for the Vegetable Industry.

Thanks to the ongoing support of the ACFR software development team led by Vsevolod Vlaskine, with particular thanks to Andrey Sokolov and Vinny Do. Thanks to Salah Sukkarieh for his direction and support of Agriculture Robotics at the ACFR.

Thanks also to Brooke Schofield, Rohan Kimber and Larn McMurray from SARDI for their help and support while acquiring data at their research plots.

References

Aasen, H., Burkart, A., Bolten, A., Bareth, G., 2015. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: from camera calibration to quality assurance. ISPRS J. Photogramm. Remote Sens. 108, 245–259.

Abd-Elrahman, A., Sassi, N., Wilkinson, B., Dewitt, B., 2016. Georeferencing of mobile ground-based hyperspectral digital single-lens reflex imagery. J. Appl. Remote Sens. 10 (1), 14002.

Adler-Golden, S., Berk, A., Bernstein, L.S., Richtsmeier, S., Acharya, P.K., Matthew, M. W., Anderson, G.P., Allred, C.L., Jeong, L.S., Chetwynd, J.H., 1998. FLAASH, a

MODTRAN4 atmospheric correction package for hyperspectral data retrievals and simulations. In: Proc. 7th Ann. JPL Airborne Earth Science Workshop. pp. 9–14.
Ahlberg, J., 2010. Estimating atmosphere parameters in hyperspectral data. In: SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, pp. 76952A–76952A.
Behmann, J., Steinrücken, J., Plümer, L., 2014. Detection of early plant stress responses in hyperspectral images. ISPRS J. Photogramm. Remote Sens. 93, 98–111.
Belluardo, G., Barchi, G., Baumgartner, D., Rennhofer, M., Weihs, P., Moser, D., 2016. Uncertainty analysis of a radiative transfer model using Monte Carlo method within 280–2500 nm region. Sol. Energy 132, 558–569.
Berk, A., Anderson, G.P., Bernstein, L.S., Acharya, P.K., Dothe, H., Matthew, M.W., Adler-Golden, S.M., Chetwynd Jr., J.H., Richtsmeier, S.C., Pukall, B., et al., 1999. MODTRAN4 radiative transfer modeling for atmospheric correction. In: SPIE's International Symposium on Optical Science, Engineering, and Instrumentation. International Society for Optics and Photonics, pp. 348–353.
Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., Maurer, H.P., Reif, J.C., Würschum, T., Müller, J., et al., 2013. Breedvision – a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors 13 (3), 2830–2847.
Chandra, K., Healey, G., 2005. Using coupled subspace models for recovery of reflectance spectra from airborne images. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, vol. 2. IEEE, pp. 1146–1151.
Chandra, K., Healey, G., 2008. Using coupled subspace models for reflectance/illumination separation. IEEE Trans. Geosci. Remote Sens. 46 (1), 284–290.
Conel, J.E., Green, R.O., Vane, G., Bruegge, C.J., Alley, R.E., Curtiss, B.J., 1987. AIS-2 radiometry and a comparison of methods for the recovery of ground reflectance.
Dahlback, A., Stamnes, K., 1991. A new spherical model for computing the radiation field available for photolysis and heating at twilight. Planet. Space Sci. 39 (5), 671–683.
Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., Furbank, R., 2014. Proximal remote sensing buggies and potential applications for field-based phenotyping. Agronomy 4 (3), 349–379.
Drew, M.S., Finlayson, G.D., 2007. Analytic solution for separating spectra into illumination and surface reflectance components. JOSA A 24 (2), 294–303.
Du, Q., Chang, N.-B., Yang, C., Srilakshmi, K.R., 2008. Combination of multispectral remote sensing, variable rate technology and environmental modeling for citrus pest management. J. Environ. Manage. 86 (1), 14–26.
Gao, B.C., Heidebrecht, K.B., Goetz, A.F.H., 1993. Derivation of scaled surface reflectances from AVIRIS data. Remote Sens. Environ. 44 (2–3), 165–178.
Gao, B.-C., Montes, M.J., Davis, C.O., Goetz, A.F.H., 2009. Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. Remote Sens. Environ. 113, S17–S24.
Geelen, B., Tack, N., Lambrechts, A., 2014. A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic. In: SPIE MOEMS-MEMS. International Society for Optics and Photonics, pp. 89740L–89740L.
Geladi, P.L.M., 2007. Calibration Standards and Image Calibration. In: Techniques and Applications of Hyperspectral Image Analysis. John Wiley & Sons, Ltd, pp. 203–220.
Gomez, C., Rossel, R.A.V., McBratney, A.B., 2008. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: an Australian case study. Geoderma 146 (3), 403–411.
Green, R., 2001. Atmospheric correction now (ACORN). developed by ImSpec LLC, available from Analytical Imaging and Geophysics LLC.
Guanter, L., Richter, R., Kaufmann, H., 2009. On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing. Int. J. Remote Sens. 30 (6), 1407–1424.
Hakala, T., Honkavaara, E., Saari, H., Mäkynen, J., Kaivosoja, J., Pesonen, L., Pölönen, I., 2013. Spectral imaging from UAVs under varying illumination conditions. UAV-g2013, 4–6 September 2013, Rostock, Germany. Edited by G. Grenzdörffer, & R. Bill. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (XL-1/W2). International Society for Photogrammetry and Remote Sensing.
Ho, J., Funt, V., Drew, M.S., 1990. Separating a color signal into illumination and surface reflectance components: Theory and applications. IEEE Trans. Pattern Anal. Mach. Intell. 12 (10), 966–977.
Honkavaara, E., Eskelinen, M.A., Pölönen, I., Saari, H., Ojanen, H., Mannila, R., Holmlund, C., Hakala, T., Litkey, P., Rosnell, T., et al., 2016. Remote sensing of 3-D geometry and surface moisture of a peat production area using hyperspectral frame cameras in visible to short-wave infrared spectral ranges onboard a small unmanned airborne vehicle (UAV). IEEE Trans. Geosci. Remote Sens. 54 (9), 5440–5454.
Honkavaara, E., Saari, H., Kaivosoja, J., Pölönen, I., Hakala, T., Litkey, P., Mäkynen, J., Pesonen, L., 2013. Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture. Remote Sens. 5 (10), 5006–5039.
Jan, H., Zbyněk, M., Lucie, H., Kaplan, V., Lukeš, P., Cudlín, P., 2008. Potentials of the VNIR airborne hyperspectral system AISA Eagle. GIS Ostrava 27, 1–6.
Karpouzli, E., Malthus, T., 2003. The empirical line method for the atmospheric correction of IKONOS imagery. Int. J. Remote Sens. 24 (5), 1143–1150.
Klose, R., Möller, K., Vielstädte, C., Ruckelshausen, A., 2010. Modular system architecture for individual plant phentotyping with an autonomous field robot. In: Proceedings of the 2nd International Conference of Machine Control & Guidance. pp. 299–307.
Kruse, F.A., 1988. Use of airborne imaging spectrometer data to map minerals associated with hydrothermally altered rocks in the northern Grapevine Mountains, Nevada, and California. Remote Sens. Environ. 24 (1), 31–51.
Kuusk, J., Kuusk, A., 2010. Autonomous lightweight airborne spectrometers for ground reflectance measurements. In: 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). IEEE, pp. 1–4.
Lee, W.S., Alchanatis, V., Yang, C., Hirafuji, M., Moshou, D., Li, C., 2010. Sensing technologies for precision specialty crop production. Comput. Electron. Agric. 74 (1), 2–33.
López-Granados, F., Torres-Sánchez, J., Serrano-Pérez, A., de Castro, A.I., Mesas-Carrascosa, F.-J., Peña, J.-M., 2016. Early season weed mapping in sunflower using UAV technology: variability of herbicide treatment maps against weed thresholds. Precision Agric. 17 (2), 183–199.
MacDonald, S.L., Staid, M., Staid, M., Cooper, M.L., 2016. Remote hyperspectral imaging of grapevine leafroll-associated virus 3 in cabernet sauvignon vineyards. Comput. Electron. Agric. 130, 109–117.
Marshall, M., Thenkabail, P., 2015. Advantage of hyperspectral EO-1 Hyperion over multispectral IKONOS, GeoEye-1, WorldView-2, Landsat ETM+, and MODIS vegetation indices in crop biomass estimation. ISPRS J. Photogramm. Remote Sens. 108, 205–218.
Mather, P.M., Koch, M., 2011. Computer Processing of Remotely-Sensed Images: An Introduction. John Wiley & Sons, Ltd..
Matthew, M.W., Adler-Golden, S.M., Berk, A., Felde, G., Anderson, G.P., Gorodetzky, D., Paswaters, S., Shippert, M., 2002. Atmospheric correction of spectral imagery: evaluation of the FLAASH algorithm with AVIRIS data. In: Applied Imagery Pattern Recognition Workshop, 2002. Proceedings. 31st. IEEE, pp. 157–163.
Mobley, C., 2016. Remote Sensing: The Atmospheric Correction Problem. <http://www.oceanopticsbook.info/view/remote_sensing/the_atmospheric_correction_problem>
Moharana, S., Dutta, S., 2016. Spatial variability of chlorophyll and nitrogen content of rice from hyperspectral imagery. ISPRS J. Photogramm. Remote Sens. 122, 17–29.
Mulla, D.J., 2013. Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. Biosyst. Eng. 114 (4), 358–371.
Nielsen, S.H., Jensen, K., Bøgild, A., Jørgensen, O.J., Jacobsen, N.J., Jæger-Hansen, C. L., Jørgensen, R.N., 2012. A low cost, modular robotics tool carrier for precision agriculture research. In: 11th International Conference on Precision Agriculture.
Noble, S.D., Brown, R.B., Crowe, T.G., 2002. The use of spectral properties for weed detection and identification? A review the use of spectral properties for weed detection and identification a review reflectance properties of plants. In: Csae/Scgr. pp. 0–17.
Okamoto, H., Suzuki, Y., Noguchi, N., 2014. Field applications of automated weed control: asia. In: Automation: The Future of Weed Control in Cropping Systems. Springer, pp. 189–200.
Oldeland, J., Dorigo, W., Lieckfeld, L., Lucieer, A., Jürgens, N., 2010. Combining vegetation indices, constrained ordination and fuzzy classification for mapping semi-natural vegetation units from hyperspectral imagery. Remote Sens. Environ. 114 (6), 1155–1166.
Panigrahy, S., Sharma, S.A., 1997. Mapping of crop rotation using multidate Indian Remote Sensing Satellite digital data. ISPRS J. Photogramm. Remote Sens. 52 (2), 85–91.
Prosa, A., Colominaa, I., Navarroa, J.A., Antequerab, R., Andrinalb, P., 2013. Radiometric block adjusment and digital radiometric model generation. ISPRS-Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. 1 (1), 293–298.
Pullanagari, R.R., Kereszturi, G., Yule, I.J., 2016. Mapping of macro and micro nutrients of mixed pastures using airborne AisaFENIX hyperspectral imagery. ISPRS J. Photogramm. Remote Sens. 117, 1–10.
Qu, Z., Kindel, B., Goetz, A.F.H., 2001. Aerosol retrieval and spectral calibration of AVIRIS with HATCH. In: 2001 AVIRIS Geoscience Workshop.
Rao, N.R., Garg, P.K., Ghosh, S.K., Dadhwal, V.K., 2008. Estimation of leaf total chlorophyll and nitrogen concentrations using hyperspectral satellite imagery. J. Agric. Sci. 146 (01), 65–75.
Roberts, D.A., Yamaguchi, Y., Lyon, R.J.P., 1986. Comparison of various techniques for calibration of AIS data. NASA STI/Recon Technical Report N 87.
Ruckelshausen, A., Biber, P., Dorna, M., Gremmes, H., Klose, R., Linz, A., Rahe, F., Resch, R., Thiel, M., Trautz, D., et al., 2009. BoniRob–an autonomous field robot platform for individual plant phenotyping. Precision Agric. 9 (841), 1.
Saari, H., Aallos, V.-V., Akujärvi, A., Antila, T., Holmlund, C., Kantojärvi, U., Mäkynen, J., Ollila, J., 2009. Novel miniaturized hyperspectral sensor for UAV and space applications. In: SPIE Europe Remote Sensing. International Society for Optics and Photonics, pp. 74741M–74741M.
Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S., Martonchik, J.V., 2006. Reflectance quantities in optical remote sensing – definitions and case studies. Remote Sens. Environ. 103 (1), 27–42.
Seidel, F.C., Kokhanovsky, A.A., Schaepman, M.E., 2010. Fast and simple model for atmospheric radiative transfer. Atmos. Meas. Tech. 3 (4), 1129–1141.
Sims, N.C., Culvenor, D., Newnham, G., Coops, N.C., Hopmans, P., 2013. Towards the operational use of satellite hyperspectral image data for mapping nutrient status and fertilizer requirements in Australian plantation forests. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 6 (2), 320–328.
Slonecker, T., 2012. Analysis of the effects of heavy metals on vegetation hyperspectral reflectance properties. In: Thenkabail, P.S., Huete, A., Lyon, J.G. (Eds.), Hyperspectral Remote Sensing of Vegetation. CRC Press, pp. 561–578.

Staenz, K., Secker, J., Gao, B.-C., Davis, C., Nadeau, C., 2002. Radiative transfer codes applied to hyperspectral data for the retrieval of surface reflectance. ISPRS J. Photogramm. Remote Sens. 57 (3), 194–203.
Staenz, K., Williams, D.J., 1997. Retrieval of surface reflectance from hyperspectral data using a look-up table approach. Can. J. Remote Sens. 23 (4), 354–368.
Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R., Kooistra, L., 2014. A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles. Remote Sens. 6 (11), 11013–11030.
Teillet, P.M., 2015. Overview of satellite image radiometry in the solar-reflective optical domain. In: Remotely Sensed Data Characterization, Classification, and Accuracies. Remote Sensing Handbook. CRC Press, pp. 87–108. oct.
Tey, Y.S., Brindal, M., 2012. Factors influencing the adoption of precision agricultural technologies: a review for policy implications. Precision Agric. 13 (6), 713–730.
Thenkabail, P.S., Smith, R.B., De Pauw, E., 2002. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. Photogramm. Eng. Remote Sens. 68 (6), 607–622.
Thompson, D.R., Roberts, D.A., Gao, B.C., Green, R.O., Guild, L., Hayashi, K., Kudela, R., Palacios, S., 2016. Atmospheric correction with the Bayesian empirical line. Opt. Express 24 (3), 2134–2144.
Underwood, J.P., Calleija, M., Taylor, Z., Hung, C., Nieto, J., Fitch, R., Sukkarieh, S., 2015. Real-time target detection and steerable spray for vegetable crops. In: Proceedings, Workshop on Robotics in Agriculture at International Conference on Robotics and Automation (ICRA).
Underwood, J., Wendel, A., Schofield, B., McMurray, L., Kimber, R., 2017. Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. J. Field Rob. (in press).
Uto, K., Seki, H., Saito, G., Kosugi, Y., 2013. Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 6 (2), 851–860.
Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M., Morcrette, J.J., 1997. Second simulation of the satellite signal in the solar spectrum, 6s: an overview. IEEE Trans. Geosci. Remote Sens. 35 (3), 675–686.
Wendel, A., Underwood, J., 2016. Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In: 2016 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 5128–5135.
Wu, C., Niu, Z., Tang, Q., Huang, W., 2008. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. Agric. Forest Meteorol. 148 (8), 1230–1241.
Xue, Z., Du, P., Li, J., Su, H., 2017. Sparse graph regularization for robust crop mapping using hyperspectral remotely sensed imagery with very few in situ data. ISPRS J. Photogramm. Remote Sens. 124, 1–15.
Yang, C., Everitt, J.H., Bradford, J.M., Murden, D., 2004. Airborne hyperspectral imagery and yield monitor data for mapping cotton yield variability. Precision Agric. 5 (5), 445–461.
Yao, H., Lewis, D., 2010. Spectral preprocessing and calibration techniques. In: Sun, D.-W. (Ed.), Hyperspectral Imaging for Food Quality Analysis and Control. Academic Press, San Diego, pp. 45–78.
Zhang, C., Kovacs, J.M., 2012. The application of small unmanned aerial systems for precision agriculture: a review. Precision Agric. 13 (6), 693–712.
Zhang, X., Liao, C., Li, J., Sun, Q., 2013. Fractional vegetation cover estimation in arid and semi-arid environments using HJ-1 satellite hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 21, 506–512.
Zhang, Y., Staab, E.S., Slaughter, D.C., Giles, D.K., Downey, D., 2012. Automated weed control in organic row crops using hyperspectral species identification and thermal micro-dosing. Crop Prot. 41, 96–105. nov.
Zhao, D., Huang, L., Li, J., Qi, J., 2007. A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. ISPRS J. Photogramm. Remote Sens. 62 (1), 25–33. may.

Chapter 4

Geometric Calibration

One of the most common types of hyperspectral sensor are line scanning cameras, which acquire a single row of pixels per exposure. Line scanning cameras are prevalent in both remote sensing [304, 23] and industrial inspection [156, 45, 215], and allow capturing images at higher frame rates or spatial resolution [152]. However, as outlined in Section 2.3.1, accurate georeferencing of line scan data is challenging. Direct georeferencing requires accurate camera pose estimates in the world frame. For systems that use an accurate navigation system (e.g. real time kinematic global positioning system and inertial navigation system (RTK-GPS-INS)), the fixed relative pose of the camera with respect to the platform or vehicle must be determined. The full 6 degrees of freedom (DOF) relative camera pose solution is often referred to as lever-arm lengths (position) and boresight angles (orientation).

A number of existing solutions for determining a line scanning camera's pose were discussed in Section 2.3.2. However, when using such cameras on mobile ground vehicles, these approaches can be cumbersome. For example, some methods require the use of ground control points (GCPs), which are difficult to obtain with the accuracy required when scanning in close proximity. Many methods, particularly in the remote sensing literature, focus on orientation and measure positional offsets by hand, but greater accuracy is required in proximal sensing applications. In ground based implementations, it is also advantageous to minimise the space required to perform

calibration, as it is more difficult to obtain data over large areas and obstacles are a greater consideration.

With the aim to address these concerns, this chapter provides a practical approach for determining the relative pose of a line scanning camera when operating aboard mobile ground vehicles. It requires a calibration pattern with a number of visually distinct points, whose dimensions do not need to be known. This calibration pattern needs to be imaged from a variety of vehicle positions and orientations. After manually labelling the pattern points in the line scan data, the algorithm then uses information from the on board navigation system to estimate the 3D pattern point locations and computes reprojection errors. Taking into account uncertainties of the system in the form of covariance matrices, including those provided by the navigation system, a likelihood is then computed. This likelihood is then iteratively minimised by varying the 6 DOF rigid relative camera offset.

In addition to providing an estimate of the relative camera pose, an Markov Chain Monte Carlo (MCMC) algorithm is used to explore the likelihood in the 6 DOF space surrounding the solution, allowing the calculation of its uncertainty. The resulting covariance matrix can be propagated in subsequent georeferencing operations to estimate mapping uncertainty. The approach was validated using data from two different mobile platforms with geometrically differing mounting configurations. The resulting solutions are then visualised in several novel ways, adding to the contributions of this thesis.

sensors

MDPI

Article

Extrinsic Parameter Calibration for Line Scanning Cameras on Ground Vehicles with Navigation Systems Using a Calibration Pattern

Alexander Wendel * and James Underwood

The Australian Centre for Field Robotics (ACFR), Department of Aerospace, Mechanical and Mechatronic Engineering (AMME), The University of Sydney, Sydney, NSW 2006, Australia; j.underwood@acfr.usyd.edu.au
* Correspondence: a.wendel@acfr.usyd.edu.au; Tel.: +61-2-9114-0895

Received: 1 September 2017; Accepted: 25 October 2017; Published: 30 October 2017

Abstract: Line scanning cameras, which capture only a single line of pixels, have been increasingly used in ground based mobile or robotic platforms. In applications where it is advantageous to directly georeference the camera data to world coordinates, an accurate estimate of the camera's 6D pose is required. This paper focuses on the common case where a mobile platform is equipped with a rigidly mounted line scanning camera, whose pose is unknown, and a navigation system providing vehicle body pose estimates. We propose a novel method that estimates the camera's pose relative to the navigation system. The approach involves imaging and manually labelling a calibration pattern with distinctly identifiable points, triangulating these points from camera and navigation system data and reprojecting them in order to compute a likelihood, which is maximised to estimate the 6D camera pose. Additionally, a Markov Chain Monte Carlo (MCMC) algorithm is used to estimate the uncertainty of the offset. Tested on two different platforms, the method was able to estimate the pose to within 0.06 m/1.05° and 0.18 m/2.39°. We also propose several approaches to displaying and interpreting the 6D results in a human readable way.

Keywords: line scan cameras; extrinsic calibration; camera pose; navigation system; GPS; ground vehicles; georeferencing

1. Introduction

Line scanning (also 1D or linear) cameras, which produce a single line of pixels for each exposure, have been used widely in areas such as remote sensing [1,2] and industrial inspection [3–5]. While 2D frame cameras offer the benefit of imaging a larger scene with each exposure, linescan cameras allow capturing of images at higher frame rates or spatial resolution [6]. One specific but common example is hyperspectral line scanning cameras, which provide both high spatial and spectral resolution. Many applications require accurate and direct determination of the real world coordinates of line scan image data, also known as georeferencing or mapping. This requires precise calibration of the sensor's intrinsic (e.g., focal length and principal point) and extrinsic parameters (i.e., camera pose with respect to the vehicle body frame). In the remote sensing literature, determination of extrinsic parameters is known as lever arm (translation) and boresight (orientation) alignment. More recently line scanning cameras have also been studied for low altitude unmanned aerial vehicle (UAV) and mobile ground based applications [7–10], but there are fewer studies addressing the extrinsic calibration requirements that closer proximity to the scene implies. Requirements include obtaining a 6 degree of freedom (DOF) extrinsic parameter solution including translation, which has a greater influence on mapping when proximal; avoiding ground control points (GCPs), which need to be more accurately geolocated when viewed from nearby; and a need for smaller survey areas for calibration, because it is more difficult to obtain data over large areas with mobile ground vehicles. This paper addresses these requirements by

providing a novel method to estimate line scanning camera pose with respect to the platform body frame, where the location and orientation of the platform is itself provided in world coordinates from a navigation system. The method uses the data from the navigation and line scanning camera only, avoiding the need for auxiliary sensors.

Extrinsic calibration for 2D frame cameras has been studied extensively due to their ubiquitous use across many different fields, and established solutions exist [11–14]. Calibration of 1D cameras has not received as much attention. Methods can be loosely grouped into two categories: scan-based calibration and line-based calibration [6]. Scan-based calibration requires an accurate rig with a linear actuator that moves the camera orthogonally to the line scan at a constant speed over a calibration pattern, such as a checker board [15,16]. This method is suitable for industrial inspection applications in a controlled laboratory or factory setting, where a linear actuator, manipulator arm or other rig is capable of moving the sensor through a precisely specified trajectory. Line-based calibration methods, on the other hand, allow calibration from a single line scan of a 3D target with a carefully designed pattern of lines [17,18]. Line-based approaches require that the dimensions of the calibration pattern are known precisely, and that the whole pattern has been imaged in one exposure. Recently, a variation of this method using multiple line scans of a planar calibration pattern has been proposed [19], and the use of an additional auxiliary frame camera has also been explored [6,20]. All the aforementioned approaches are suitable for well controlled environments: for scan-based calibration the movement of the sensor needs to be accurately controlled, while for line-based methods, the position of the pattern with respect to the sensor is critical. However, in a mobile ground based field platform, where the camera is rigidly mounted in a particular position to the platform, it is difficult to meet either of those requirements.

In previous methods, extrinsic parameters are usually determined with respect to the calibration pattern or an auxiliary frame camera. Therefore to determine the camera to navigation system transform either requires accurate knowledge of a pattern or points in world coordinates or an additional step such as "hand-eye" calibration [21]. Hand-eye calibration involves determining the transformation from a camera to an end effector (a robotic hand for instance), where these are rigidly linked, and is a thoroughly covered topic in the robotics literature. The problem is generally solved by imaging a calibration pattern from many different locations, where the transformations between the different end effector positions and camera to calibration pattern transformations are known using standard frame camera calibration techniques. Comparisons can be made with the problem in this paper, where the navigation system positions (and therefore any transformations between them) are known, and camera to calibration pattern transformations can be determined using any of the previously discussed methods.

As remote sensing most commonly involves imaging from an aerial or satellite platform, translation (lever arm) offsets have a smaller effect on imaging accuracy, and can be measured manually [22,23]. Accurately geolocated GCPs are commonly used to determine boresight alignment [24], which can also be adopted for ground based applications [25]. Efforts have been made to avoid the use of GCPs, by detecting points of interest in separate scans of the same area and determining their 3D position using a known digital elevation model (DEM) [22]. Similarly, non-surveyed tie-points between overlapping acquisition runs have been used in combination with bundle adjustment to determine boresight parameters [26]. The use of GCPs has also been combined with DEMs to improve accuracy and allow self-calibration [27]. Frame cameras have been used to aid in determination of boresight misalignments [28], and additionally in combination with a DEM [29]. Frame camera images have also been used to improve the geometric characteristics of processed hyperspectral linescan images from a UAV [30].

This paper provides a method for the determination of the relative 6 DOF pose of a rigidly mounted line scanning camera with respect to a navigation system on a ground based mobile platform. With this approach many of the previously outlined requirements and limitations are mitigated:

- The dimensions of the calibration pattern do not need to be known, and so it does not need to be printed to any particular accuracy, nor even measured.

- GCPs do not need to be surveyed.
- Auxiliary sensors, such as 2D frame cameras, are not required to aid the calibration.
- A single, compact calibration pattern can be used rather than widely distributed GCPs.
- Translational (lever arm) offsets are determined in addition to rotations (boresight), due to their increased significance when at close proximity to the scene.

The remainder of the paper is organised as follows. In Section 2 the theory of the proposed method is outlined in detail. Then Section 3 provides practical implementation details and the experimental method. Experimental results using the Ladybird and Shrimp robotic platforms are produced and discussed in Sections 4 and 5.

2. Overview of Approach

In this section, the theoretical approach used for estimating the camera pose with respect to the platform body is outlined in detail. Initially, an overview of the line scanning camera model is provided, which is an adaptation of the widely used pinhole model. This allows defining lines or rays in 3D space that intersect both the camera centre and a pixel on the sensor. When combined with the Cartesian coordinate transformations between camera, body and world frames, rays can be projected onto a surface, and conversely a world 3D point can be reprojected to a point on the 2D sensor. It is desirable to minimise any errors in the camera pose, as they directly affect mapping accuracy.

We propose a method that estimates the relative camera pose using image and navigation system data. The data are obtained by moving the platform in order to observe a calibration pattern with multiple point targets from different perspectives. The calibration pattern point locations are then manually labelled in the image data. Starting from an initial hand measured camera pose, image pixel locations of the observed pattern points and corresponding platform poses are combined, and all of the resulting rays are used to triangulate the pattern point locations in world coordinates. These point estimates are then reprojected to the sensor frame for each observation. The reprojection error is calculated as the distance between each observed and reprojected pixel. The reprojection error uncertainty is calculated by propagating the input uncertainties through each calculation as variance-covariance matrices (henceforth referred to as covariance matrices for brevity). Assuming a normal distribution of the reprojection error over input parameters, the likelihood of the data given a relative camera pose hypothesis can be estimated. By maximising the likelihood, the six relative camera pose parameters can be optimised. Following this, a random sampling based procedure is provided to estimate the uncertainties of the optimal camera pose using Markov Chain Monte Carlo (MCMC).

Throughout this paper, superscripts represent the reference frame of a particular variable. Subscripts refer to a descriptor (e.g., which pose is being referred to), axis reference, and instance identifiers for that variable, in that order. For example, $r^b_{c,x}$ refers to the camera centre location along the x axis relative to the body frame.

2.1. Line Scanning Camera Model

Using the pinhole camera model with homogeneous coordinates, a point $\mathbf{p}^w = [x, y, z, 1]^T$ in world coordinates is projected to the camera sensor at $[u, v, 1]^T$ with the following Equation [31]:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} s = \mathbf{P} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \tag{1}$$

where s is a scale factor and $\mathbf{P}$ can be broken down into,

$$\mathbf{P} = \mathbf{K}\mathbf{R}^{w-1}_c [\mathbf{I}_{3\times3} | - \mathbf{p}^w_c]. \tag{2}$$

$\mathbf{R}_c^w$ is the rotation matrix of the camera with respect to the world frame. Joined horizontally are $\mathbf{I}_{3\times3}$ and $\mathbf{p}_c^w$, which are the identity matrix and the world camera position (i.e., the camera centre $[r_{c,x}^w, r_{c,y}^w, r_{c,z}^w]^T$) respectively. $\mathbf{K}$ is the intrinsic camera matrix:

$$\mathbf{K} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

where f, u_0 and v_0 are the focal length (in pixels) and principal points respectively (we neglect skew because there is only one spatial axis). For a line scanning camera, we assume that $v_0 = 0$ and so it follows that for a 3D world point to be visible in the 1D pixel array, it must be located near the plane that intersects the scan line on the sensor (i.e., where $v = 0$) and the camera centre (focal point). How closely a point must be located to that plane depends on the instantaneous field of view (IFOV) and distance from the sensor. The IFOV is the angle over which each pixel is sensitive to radiation. While linescan image data is by definition at $v = 0$, reprojection errors can occur in both u and v as will be shown later. Therefore, even though the model allows for two spatial dimensions on the image sensor, it describes the projection of points for individual 1D line scan frames only.

Each pixel point $[u, v, 1]^T$ maps to a ray or line in 3D space, which connects the sensor pixel, camera centre and object being viewed. While that ray may be defined by any two points that lie on it, the following are mathematically convenient to obtain: the camera centre $\mathbf{p}_c^w$ and $\mathbf{p}_s^w = \mathbf{P}^+[u, v, 1]^T$, where $\mathbf{P}^+$ is the pseudo-inverse of $\mathbf{P}$ [31].

2.2. Rotation and Transform Conventions

In this paper, we use both Euler and axis-angle conventions to represent rotations compactly. The navigation system on the platforms used in this work provide platform pose estimates using the Euler zyx intrinsic convention (also known as Tait-Bryan or yaw-pitch-roll), which are represented as $[\phi_x, \phi_y, \phi_z]$, and may be converted to rotation matrices as per Berner [32] or Section 3.1 in Underwood [33]. While Euler angle representations are commonly used in robotics applications, they present the following ambiguities. Some different combinations of ϕ_x ϕ_y and ϕ_z can represent the same rotation [34]. Similarly, a small freedom of rotation about a non-orthogonal axis can result in a large correlated degree of freedom spread over two Euler angles, which is difficult to interpret when estimating parameter uncertainty. For these reasons, while navigation and hand measured pose data is provided as Euler angles, we favour the axis-angle representation for all internal calculations and results. An axis-angle rotation is given as a unit length vector $\mathbf{e}$ and a rotation θ around it:

$$(\theta, \mathbf{e}) = \left(\theta, \begin{bmatrix} e_x \\ e_y \\ e_z \end{bmatrix} \right). \tag{4}$$

Since rotations only have three degrees of freedom, an axis-angle rotation may be expressed as a length three vector:

$$\theta\mathbf{e} = \begin{bmatrix} \theta e_x \\ \theta e_y \\ \theta e_z \end{bmatrix}. \tag{5}$$

Axis-angle rotations may be converted to rotation matrices as follows [32]:

$$\mathbf{R} = \left[\mathbf{I}_{3x3} + sin(\theta)\mathbf{S}_n + (1 - cos(\theta))\mathbf{S}_n^2 \right], \tag{6}$$

where

$$\mathbf{S}_n = \begin{bmatrix} 0 & -e_z & e_y \\ e_z & 0 & -e_x \\ -e_y & e_x & 0 \end{bmatrix}. \tag{7}$$

A complete 6 DOF pose transform can be compactly represented with the three translation and three orientation parameters:

$$\mathbf{t} = [r_x, r_y, r_z, \phi_x, \phi_y, \phi_z]^T, \tag{8}$$

or

$$\mathbf{t} = [r_x, r_y, r_z, \theta e_x, \theta e_y, \theta e_z]^T, \tag{9}$$

depending on whether Euler or axis-angle conventions are used. The pose transforms of importance in this paper are the world to platform body transform $\mathbf{t}_b^w$, platform body to camera transform $\mathbf{t}_c^b$, and, combining these, the world to camera transform $\mathbf{t}_c^w$ (see Figure 1). Note the sub- and superscripts: e.g., $\mathbf{t}_c^b$ denotes the translation and rotation of the camera axes with respect to the platform body.

By splitting the world pose of the camera $\mathbf{t}_c^w$ into a combination of the body pose $\mathbf{t}_b^w$ and the camera relative pose $\mathbf{t}_c^b$, $\mathbf{P}$ from Equation (2) can be shown as a function of the camera rotation and translation with respect to the body frame $\mathbf{R}_c^b$ and $\mathbf{p}_c^b$, and platform body rotation and translation with respect to the world frame $\mathbf{R}_b^w$ and $\mathbf{p}_b^w$:

$$\mathbf{P} = \mathbf{K}\left(\mathbf{R}_c^{b^{-1}}[\mathbf{I}_{3\times3}| - \mathbf{p}_c^b]\right)\left(\left[\begin{array}{c|c} \mathbf{R}_b^{w^{-1}} & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array}\right]\left[\begin{array}{c|c} \mathbf{I}_{3\times3} & -\mathbf{p}_b^w \\ \hline \mathbf{0} & 1 \end{array}\right]\right) \tag{10}$$

In our case, $\mathbf{R}_b^w$ and $\mathbf{p}_b^w$ are provided by the navigation system, and $\mathbf{R}_c^b$ and $\mathbf{p}_c^b$ are the relative camera pose parameters we would like to estimate.

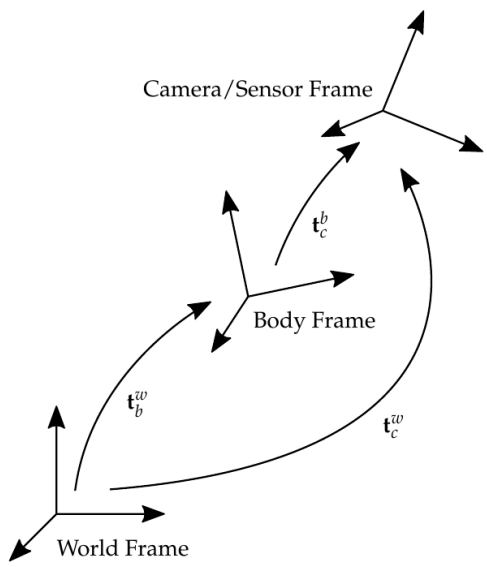


Figure 1. Summary of transforms referenced in this paper.

2.3. Estimation of Calibration Pattern Points

The first step of the proposed method involves estimating the location of calibration pattern points in world coordinates, as these are unknown and must be computed from the data. As shown in Figure 2, rays are calculated for each pixel observation of each calibration pattern point, given the concurrent navigation system solution and camera pose proposal. Average point locations are determined by triangulating all rays corresponding to the same calibration pattern point. Uncertainties for all inputs (pixel locations, navigation solutions and intrinsics) in the form of covariance matrices are propagated using the Jacobian of the point calculation function, yielding an uncertainty estimate (covariance matrix) for each calibration point estimate.

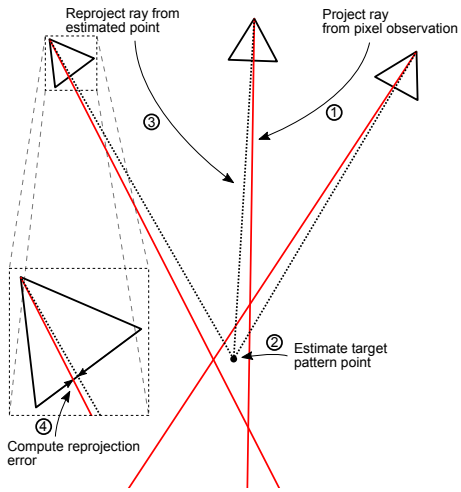


Figure 2. Method summary. Rays corresponding to individual calibration pattern point observations are determined from pixel observations and camera poses. Calibration pattern point locations are then triangulated from all rays, and subsequently reprojected to the camera sensor. A reprojection error can then be computed by calculating the difference between the reprojected point and the pixel observation. Uncertainties are propagated through at each step, which facilitates the calculation of the uncertainty for the reprojection error, and subsequently a likelihood value, which is maximised by the optimiser.

The proposed method starts with repeated imaging of points that can be uniquely identified. The use of a regular calibration pattern ensures points can be easily distinguished and is therefore recommended. The location of each pattern point $\mathbf{p}_k^w$ for $k \in \{1, 2, ..., M\}$ is estimated from all of its observation rays $i \in \{1, 2, ..., N\}$. There are M points on the calibration pattern and the whole pattern is viewed N times. For each point, we calculate the nearest points between all pairs of observation rays (i, j) and apply a weighted average. Nearest points between rays are calculated as follows [35]:

$$\begin{aligned}\mathbf{p}_{k,ij}^w =& \mathbf{p}_{c,k,i}^w + \\ & \frac{(\mathbf{p}_{c,k,j}^w - \mathbf{p}_{c,k,i}^w) \cdot \mathbf{n}_{k,ij}}{(\mathbf{p}_{s,k,i}^w - \mathbf{p}_{c,k,i}^w) \cdot \mathbf{n}_{k,ij}} (\mathbf{p}_{s,k,i}^w - \mathbf{p}_{c,k,i}^w),\end{aligned} \tag{11}$$

where

$$\begin{aligned}\mathbf{n}_{k,ij} =& (\mathbf{p}_{s,k,j}^w - \mathbf{p}_{c,k,j}^w) \times \\ & [(\mathbf{p}_{s,k,i}^w - \mathbf{p}_{c,k,i}^w) \times (\mathbf{p}_{s,k,j}^w - \mathbf{p}_{c,k,j}^w)].\end{aligned} \tag{12}$$

We could estimate $\mathbf{p}_k^w$ as the unweighted mean of all $\mathbf{p}_{k,ij}^w$ for a given pattern point k, but some estimates are more certain than others given the conditions of how they were measured. A more accurate estimate is obtained using a weighted average according to the uncertainty. The uncertainty of each point $\mathbf{p}_{k,ij}^w$ can be obtained by computing its Jacobian $\mathbf{J}_{\mathbf{p}_{k,ij}^w}$ with respect to all input values. Also required are the uncertainties of the pixel and platform pose observations for each ray, expressed as covariance matrices, $\mathbf{Q}_{uv,k,i}$, $\mathbf{Q}_{\mathbf{t}_{b,k,i}^w}$, $\mathbf{Q}_{uv,k,j}$ and $\mathbf{Q}_{\mathbf{t}_{b,k,j}^w}$, as well as intrinsic and extrinsic parameter covariances, $\mathbf{Q}_{int}$ and $\mathbf{Q}_{\mathbf{t}_c^b}$. Although line scan cameras have only one pixel coordinate (u), there is also uncertainty in the second coordinate v, because a point elicits a pixel response if it is located within the camera's IFOV, not necessarily directly on the scan line. $\mathbf{Q}_{int}$ and $\mathbf{Q}_{\mathbf{t}_c^b}$ contain variances and covariances of the intrinsic camera parameters and the relative camera pose respectively. All the input covariance matrices are combined into one matrix $\mathbf{Q}_{k,ij}$:

$$\mathbf{Q}_{k,ij} = \begin{bmatrix} \mathbf{Q}_{uv,k,i} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\mathbf{t}_{b,k,i}^w} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{uv,k,j} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_{\mathbf{t}_{b,k,j}^w} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_{int} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_{\mathbf{t}_c^b} \end{bmatrix}. \tag{13}$$

No correlation between the navigation solutions of the two rays is assumed, which is reasonable if the two observations are sufficiently separated in time. $\mathbf{Q}_{k,ij}$ and $\mathbf{J}_{\mathbf{p}_{k,ij}^w}$ can now be used to compute the uncertainty of $\mathbf{p}_{k,ij}^w$ (Equation (11)) as covariance matrix $\mathbf{\Sigma}_{\mathbf{p}_{k,ij}^w}$:

$$\mathbf{\Sigma}_{\mathbf{p}_{k,ij}^w} = \mathbf{J}_{\mathbf{p}_{k,ij}^w} \mathbf{Q}_{k,ij} \mathbf{J}_{\mathbf{p}_{k,ij}^w}^T. \tag{14}$$

Because we wish to estimate both $\mathbf{t}_c^b$ and $\mathbf{Q}_{\mathbf{t}_c^b}$ with respect to all error sources other than the camera pose, we set all elements of the 6×6 covariance matrix $\mathbf{Q}_{\mathbf{t}_c^b}$ to zero temporarily [33,36]. Each point $\mathbf{p}_k^w$ on the calibration pattern can then be estimated by computing an average that is weighted according to the covariances [37]:

$$\mathbf{W}_{k,ij} = \mathbf{\Sigma}_{\mathbf{p}_{k,ij}^w}^{-1}, \tag{15}$$

$$\mathbf{\Sigma}_{\hat{\mathbf{p}}_k^w} = \left(\sum_i^N \sum_j^N \mathbf{W}_{k,ij} \right)^{-1}, \tag{16}$$

$$\hat{\mathbf{p}}_k^w = \mathbf{\Sigma}_{\hat{\mathbf{p}}_k^w} \left(\sum_i^N \sum_j^N \mathbf{W}_{k,ij} \mathbf{p}_{k,ij}^w \right). \tag{17}$$

This ensures that the contribution of each closest point for each ray pair ($\mathbf{p}_{k,ij}^w$) is weighted according to its certainty, taking into account navigation system uncertainty or challenging viewpoint geometry (such as a small angle between the two rays).

2.4. Calculation of Reprojection Error and Likelihood Optimisation

Once estimates and uncertainties of each calibration pattern point have been obtained, they are reprojected to the camera for each observation, which allows calculating an error between each of the observed pixel locations and the reprojected pixels (see Figure 2). The uncertainties of all inputs and calibration pattern point estimates are also propagated through, which yields an uncertainty value for each reprojection error. This enables the calculation of an overall likelihood value of the data given a camera pose proposal. The optimiser maximises this likelihood by varying the camera pose to arrive at an estimate.

For each observation i, $\hat{\mathbf{p}}_k^w$ can be reprojected according to Equation (1), given a $\mathbf{t}_c^b$ and corresponding navigation system solution $\mathbf{t}_{b,k,i}^w$. The reprojection error is calculated as the Euclidean distance between the reprojected and observed pixel locations:

$$e_{k,i} = \sqrt{(u_{k,i} - \hat{u}_{k,i})^2 + (v_{k,i} - \hat{v}_{k,i})^2}. \tag{18}$$

The reprojection is two dimensional, because non-optimal $\mathbf{t}_c^b$ can result in reprojected pixels that deviate from the one dimensional scan line ($\hat{v}_{k,i} \neq 0$), but $v_{k,i}$ is assumed to be 0. The variance of the reprojection error can also be computed using the input covariance matrix and Jacobian $\mathbf{j}_{e_{k,i}}$:

$$\sigma_{e_{k,i}}^2 = \mathbf{j}_{e_{k,i}}^T \mathbf{Q}_{k,i} \mathbf{j}_{e_{k,i}}, \tag{19}$$

where

$$\mathbf{Q}_{k,i} = \begin{bmatrix} \mathbf{\Sigma}_{\hat{\mathbf{p}}_{k,l}^w} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{uv,k,i} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{\mathbf{t}_{b,k,i}^w} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_{int} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_{\mathbf{t}_c^b} \end{bmatrix}. \tag{20}$$

The Jacobian $\mathbf{j}_{e_{k,i}}$ is lower case because it is only one dimensional in this instance, since $e_{k,i}$ is a scalar value. As in Equation (13), we again set all elements of $\mathbf{Q}_{\mathbf{t}_c^b}$ to zero. The log likelihood of a transform $\mathbf{t}_c^b$ given the observations can then be estimated as,

$$log\Lambda = -\sum_k^M \sum_i^N \frac{e_{k,i}^2}{2\sigma_{e_{k,i}}^2}. \tag{21}$$

The objective is to maximise $log\Lambda$, by varying the 6-DOF $\mathbf{t}_c^b$ vector. This can be achieved using standard optimisation methods to minimise the negative log likelihood:

$$\underset{\mathbf{t}_c^b}{argmin} \; -log\Lambda = \underset{\mathbf{t}_c^b}{argmin} \sum_k^M \sum_i^N \frac{e_{k,i}^2}{2\sigma_{e_{k,i}}^2}, \tag{22}$$

$log\Lambda$ is fully recalculated at each optimisation iteration, which includes the triangulation of calibration pattern points and calculation of their reprojection error.

2.5. Variance-Covariance Matrix Estimation

Once the relative camera pose $\mathbf{t}_c^b$ has been determined, it is desirable to approximate the covariance matrix of the solution, which provides an estimate of how uncertain the six relative camera pose parameters are. In combination with covariances of other parameters, such as the navigation system solution, this also allows mapping accuracy to be quantified. In other words, the result provides values for $\mathbf{Q}_{\mathbf{t}_c^b}$, completing the full covariance matrix (see Equation (29) in Section 3.6). Note that all elements of $\mathbf{Q}_{\mathbf{t}_c^b}$ are set to zero for its estimation and optimisation, as previously mentioned in Sections 2.3 and 2.4.

The proposed approach is based on similar work done with lidar sensors [33,36], but the details differ because 1D cameras do not directly provide depth information. We propose a random sampling based method, where a set of sample sensor to body transforms are selected using a MCMC algorithm [38], which differs from the Monte Carlo (MC) importance sampling approach in [33,36]. This provides greater sampling efficiency and avoids the need to manually define a sampling region. The algorithm is guided by the likelihood of each relative camera pose sample, which governs the selection of the next sample.

There are several MCMC variations, but they all share the property that each sample is selected based on the previous. For a large number of samples, the distribution tends towards the probability

distribution that is being sampled from (i.e., Λ in this paper) [39]. For further details about MCMC sampling, the reader is referred to the numerous resources available on the topic [38,39]. The MCMC algorithm provides a list of samples $\{t^b_{c,1}, t^b_{c,2}, ..., t^b_{c,l}, ..., t^b_{c,r}\}$, which are distributed according to Λ, from which the covariance can be computed as:

$$\mathbf{Q}_{\mathbf{t}^{b*}_c} = \frac{1}{r-1}\sum_{l=1}^{r}\left(\mathbf{t}^b_{c,l} - \bar{\mathbf{t}}^b_c\right)\left(\mathbf{t}^b_{c,l} - \bar{\mathbf{t}}^b_c\right)^T, \tag{23}$$

where

$$\bar{\mathbf{t}}^b_c = \frac{1}{r}\sum_{l=1}^{r}\mathbf{t}^b_{c,l}, \tag{24}$$

3. Materials and Methods

This section outlines the equipment and methods used to obtain the data and analyse the results. A planar calibration pattern was placed in the environment and imaged from several different orientations using a line scanning camera mounted to two different ground based robotic platforms. A navigation system mounted to each platform recorded the 6 DOF position and orientation of the platforms ($\mathbf{t}^w_b = [r^w_{b,x}, r^w_{b,y}, r^w_{b,z}, \phi^w_{b,x}, \phi^w_{b,y}, \phi^w_{b,z}]$) throughout the acquisition period. Image pixel locations of calibration pattern points and matching robot poses were then used to estimate the relative camera pose using an iterative optimisation algorithm. Finally, the uncertainty of the camera pose estimate in the form of a covariance matrix was approximated using MCMC.

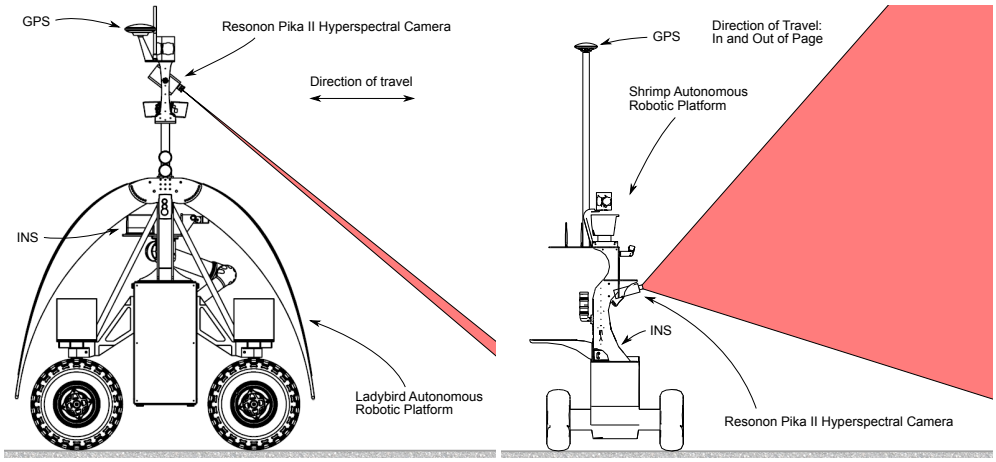
First the ground based mobile platforms and associated sensors used to acquire data are introduced, followed by a description of the data acquisition process and extraction of pattern point observations. The implementation of the method presented in Section 2 is outlined, which includes the optimisation and an outlier removal process. Methods for mapping image data and comparing camera poses are presented, as required for the analysis of the results, and a method is presented to calculate the basin of attraction, to assess the sensitivity of the process to the initial camera pose.

3.1. Mobile Sensing Platforms

A line scanning hyperspectral camera was mounted to two different robotic platforms, Ladybird and Shrimp (Figure 3). Both were designed and built at the Australian Centre for Field Robotics (ACFR) at The University of Sydney as flexible tools to support a range of research applications [40–45]. The sensor suite on both platforms includes a real time kinematic (RTK)/global positioning system (GPS)/inertial navigation system (INS), which provides platform pose and covariance estimates (details in Table 1). The GPS units on both platforms are identical, but the Shrimp platform uses a lower grade inertial measurement unit (IMU) than the Ladybird platform.

Line scan image data were acquired with a Resonon Pika II visible to near infrared (VNIR) line scanning camera that was mounted to the Ladybird and Shrimp robots in a push broom configuration. For the Ladybird, the camera was oriented such that the scan line is horizontal, pitched down for scanning the ground surface (Figure 3a). On Shrimp, the camera was mounted such that the scan line is vertical, and pitched upwards slightly to allow scanning of upright objects (Figure 3b). The camera produces hyperspectral images of 648 spatial by 244 spectral pixels (spectral resolution of 2 nm from 390.9–887.4 nm) at a rate of 133 frames per second and native bit depth of 12. For the purposes of this paper, the spectral dimension was averaged to produce 648 pixel monochrome scan lines. Apart from this averaging step, the method described in this paper is not particular to hyperspectral cameras and may be applied equally to other types of line scanning imagers. Schneider Cinegon 6 mm and 8 mm objective lenses were used for Shrimp and Ladybird respectively, and manually focused with a checker board at the typical distance to the scene. The principal point of the camera/lens combination was assumed to be at the centre of the line scan ($u_0 = 323$), the focal length was assumed to be as

per the manufacturer supplied measurements (see Table 1), and distortion was assumed to be zero. Hand measured pose estimates and manufacturer supplied lens details are shown in Table 1.



(**a**) Ladybird robotic platform (**b**) Shrimp robotic platform

Figure 3. The Ladybird (**a**) and Shrimp (**b**) robotic platforms and sensor configurations.

Table 1. Platform configurations.

	Ladybird	Shrimp
Manually Measured Camera Pose t_c^b		
$r_{c,x}^b, r_{c,y}^b, r_{c,z}^b$ (m)	$0.2, 0.0, -0.8$	$0.0, -0.2, -0.5$
$\phi_{c,x}^b, \phi_{c,y}^b, \phi_{c,z}^b$ (°)	$-56.0, 0.0, -90.0$	$0.0, 105.0, -90.0$
Camera Lens Details		
Manufacturer	Schneider	Schneider
Model	Cinegon 8 mm	Cinegon 6 mm
Focal length	8.2 mm	6.2 mm
Approx. aperture	f/2.5	f/3.0
IFOV	1.88 mrad	2.5 mrad
Navigation System Details		
Manufacturer	Novatel	Novatel
GPS receiver	ProPak-G2plus	ProPak-G2plus
IMU	Honeywell HG1700	IMU-CPT

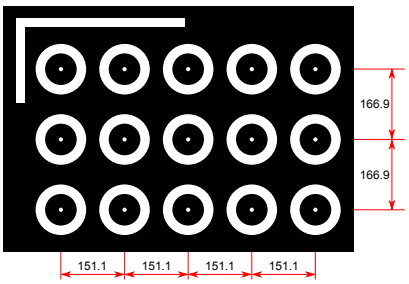
Initial pose estimates were measured by hand with the mobile platforms on a level surface using measuring tape for translational offsets, and a digital inclinometer (SPI Pro 3600) for angular offsets around the robots' horizontal x and y axes. Angular offsets around the robots' vertical z axis were assumed to be the intended mounting orientations, which are in increments of 90° for both platforms. Note that if the camera is mounted at angles that are clearly not in 90° increments, referring to a CAD model is recommended. Hand measured translation parameters ($r_{c,x}^b$, $r_{c,y}^b$ and $r_{c,z}^b$) were assumed to have a standard deviation (σ) of 0.1 m and orientation parameters ($\phi_{c,x}^b$, $\phi_{c,y}^b$ and $\phi_{c,z}^b$) were assumed to have a σ of 2°.

3.2. Data Acquisition

A calibration pattern with 15 points arranged in a 3×5 pattern was printed to an A1 size sheet of paper and mounted to a flat rigid plywood board (see Figure 4c). The pattern was designed to maximise contrast for efficient extraction of pattern points. A corner shape was added to one side of

the pattern to facilitate unique identification of each point. It is not necessary to know the pattern's dimensions for recovery of the platform to camera pose, as each point is treated independently during the calibration. This also means that theoretically a single point with sufficient observations could be used for calibration. However, we added more pattern points since there is no significant practical cost, efficiently increasing the amount of data obtained. For the Ladybird platform, the pattern board was placed on relatively flat ground (see Figure 4a). As shown in Figure 5, the pattern was scanned from several directions around a circle with the calibration pattern in the centre. Two types of scans were performed, one with the robot's wheels flat on the ground and one with one side of the robot elevated by driving over an aluminium channel. This raised two of the wheels by approximately 100 mm, inducing a roll of approximately 4°. For the Shrimp platform, the same calibration pattern was mounted to a ladder in an approximately vertical orientation (see Figure 4b). In this case data were acquired next to a hill with various orientations and positions with respect the pattern, where the hill caused continuously variable roll and pitch, up to approx. 17° (see Figure 5b). For both platforms, body orientation was intentionally varied as much as possible in an attempt to maximise observability of parameters [33]. The robots were manually operated throughout the acquisition period, and care was taken to move slowly and smoothly while the calibration pattern was imaged.

(**a**) Ladybird data acquisition (**b**) Shrimp data acquisition



(**c**) Calibration pattern

Figure 4. Data acquisition configuration and location for Ladybird (**a**) and Shrimp (**b**). The calibration pattern used in both instances is shown in (**c**). Dimensions are in mm. The centre dots are 10 mm in diameter, sized to be as small as possible while still being visible in the image data in order to maximise labelling accuracy. The outer rings help with locating the points in the data for labelling, and are 75–120 mm, inner to outer diameter. Note that these dimensions do not need to be known for the optimisation procedure, nor is it necessary to print them in any particular strict arrangement, though a regular pattern is recommended to allow easy identification of each point.

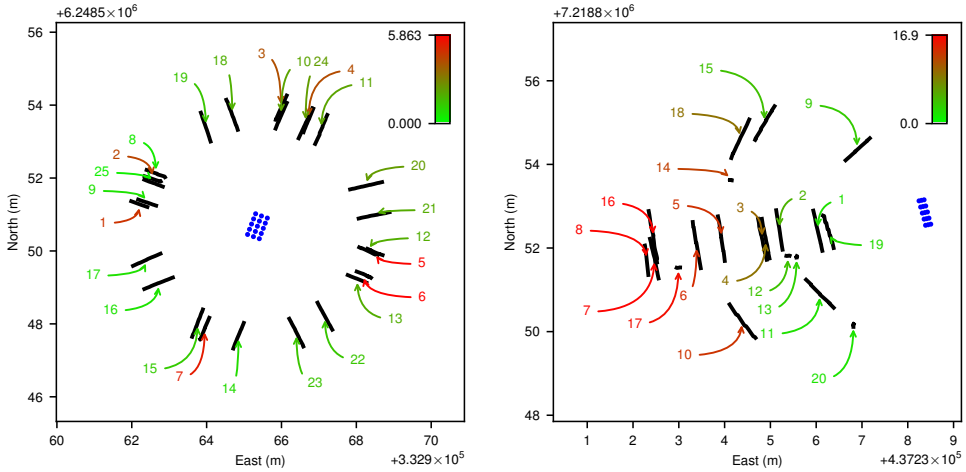
All data was timestamped allowing association between individual scan lines and platform pose solutions. Localisation uncertainties reported by the navigation system are shown in Table 2 as median standard deviations (i.e., square root of the diagonals of the covariance matrices only) for

Sensors **2017**, *17*, 2491

the acquisition runs, which illustrates that the navigation system in the Ladybird platform is able to provide body pose estimates with much greater certainty than the navigation system on Shrimp, due to the higher grade IMU.

Table 2. Median navigation system uncertainties as 1 standard deviation.

Platform	$\sigma_{r^w_{b,x}}$ (m)	$\sigma_{r^w_{b,y}}$ (m)	$\sigma_{r^w_{b,z}}$ (m)	$\sigma_{\phi^w_{b,x}}$ (°)	$\sigma_{\phi^w_{b,y}}$ (°)	$\sigma_{\phi^w_{b,z}}$ (°)
Ladybird	1.052×10^{-2}	1.305×10^{-2}	1.118×10^{-2}	2.362×10^{-1}	2.636×10^{-1}	1.053×10^{-1}
Shrimp	4.520×10^{-2}	4.369×10^{-2}	4.887×10^{-2}	7.534×10^{-1}	7.284×10^{-1}	8.416×10^{-1}



(**a**) Ladybird (**b**) Shrimp

Figure 5. Top down view of platform positions (from navigation system) during each observation of the calibration pattern for Ladybird (**a**) and Shrimp (**b**). Each observation provides a complete view of all points on the calibration pattern, as well as concurrent navigation system solutions. The calibration pattern points are indicated with blue dots. The observation runs are numbered, and coloured according to the norm of vehicle body pitch and roll (i.e., $\sqrt{{\phi^w_{b,x}}^2 + {\phi^w_{b,y}}^2}$), as indicated by the colour bars in degrees. Note that in (**b**) the calibration pattern was upright (mounted to a ladder), which is why the points appear more closely clustered from the top-down perspective. Some of the observation runs appear very short in (**b**). This is because in these instances the platform scanned the pattern by rotating on the spot.

3.3. Pattern Pixel Extraction

Approximate pixel locations of points on the calibration pattern were selected manually by appending successive line scans to form a rectangular image and selecting individual pattern points in order (see Figures 6 and 7). Note that line scans were concatenated naively, ignoring camera or body pose data (i.e., not mapped or georeferenced). This worked well for this purpose, because the platforms were moved slowly and smoothly while the calibration pattern was scanned. Particular care was taken to ensure that point ID numbers were consistent for all observations of the calibration pattern. Pixel locations were then refined to sub pixel precision by extracting a 10×10 patch around the selected points and resizing it to 100×100 pixels using bi-cubic interpolation. The intensity peak closest to the centre was taken as the pattern point pixel location. Along-track, the closest time stamp was used to obtain the corresponding navigation solution. This provides pixel position $u_{k,i}$ and platform pose $[r^w_{b,x}, r^w_{b,y}, r^w_{b,z}, \phi^w_{b,x}, \phi^w_{b,y}, \phi^w_{b,z}]^T$, which are necessary for calibration according to Equation (22).

Figure 6. Example grey scale image obtained by appending successive line scans from Ladybird (without pose compensation). A single view of all points on the calibration pattern is referred to as an observation in this paper, including concurrent navigation system data. The red box indicates one such observation, shown as a more detailed close-up view in Figure 7. The two views of the calibration pattern shown in this figure were obtained by driving the platform forward, producing the first view, and then backwards, giving the second view, which is therefore a mirrored version of the first. The second view appears more stretched because it was scanned more slowly, generating more line scans, given the same fixed frame rate. Note that the observations shown in Figure 5a only include one observation from each forward-reverse pair, because they provide almost the same information (i.e., similar navigation system solutions).

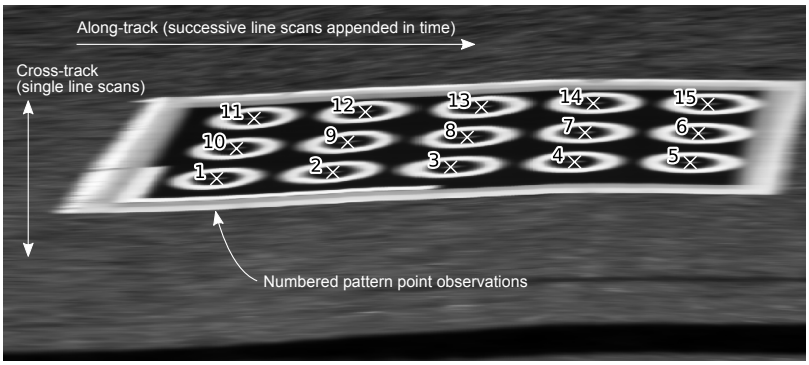


Figure 7. Close-up view of the observation indicated in Figure 6, showing manually labelled calibration point locations (white crosses) with ID numbers. Each 648 pixel line scan is vertical and concatenated horizontally. A similar image was generated for each observation of the calibration pattern prior to manually labelling each point. Care was taken to ensure the numbering scheme remained consistent for all observations of the pattern.

3.4. Optimisation and Uncertainty Estimation

Optimisation was performed using the Powell optimiser algorithm provided by the SciPy python package [46,47]. While other optimisers may be suitable for this task, as long as they minimize a scalar (negative log likelihood), while varying a vector (relative camera pose), the Powell algorithm achieved acceptable performance with the following tolerance values: $tolx = 1 \times 10^{-5}$ and $ftol = 1 \times 10^{-8}$. The objective function that was provided to the optimiser takes the relative camera pose parameters ($\mathbf{t}_c^b = [r_{c,x}^b, r_{c,y}^b, r_{c,z}^b, \theta_c^b e_{c,x}^b, \theta_c^b e_{c,y}^b, \theta_c^b e_{c,z}^b]^T$) and computes the negative log likelihood $-log\Lambda$ (see Equation (22)) given all pixel locations and navigation system solutions. The optimiser repeatedly calls this function, updating $\mathbf{t}_c^b$ in order to find a relative camera pose $\mathbf{t}_c^b*$ that minimises $-log\Lambda$.

As described in Section 2.5 we use MCMC to estimate uncertainties in the form of a covariance matrix. MCMC was performed with the emcee python package [38], which was given a function that computes the log likelihood (Equation (21)). The algorithm was initialised with the previously optimised relative camera pose $\mathbf{t}_c^b*$, and run with 250 walkers and 100 iterations, yielding 25,000 samples. A burn in run was also performed with 100 iterations to allow the function to explore the local region prior to performing the actual sampling run. Each sample represents one hypothetical parameter vector $\mathbf{t}_c^b$. The distribution of the samples generated by the MCMC algorithm correspond to

$log\Lambda$, so the uncertainty of the relative camera pose estimate, $\mathbf{Q}_{\mathbf{t}_c^{b*}}$, can be estimated by computing Equation (23).

For all calculations of $log\Lambda$, 6×6 covariances for the platform pose were provided by the navigation system at each time stamp. It was assumed that a u pixel point location could be estimated to within one standard deviation of 0.5 pixels (i.e., $\sigma_u = 0.5\,pixels$). If a point is visible it must also be within the IFOV of the sensor (see Table 1), which is approximately 2 pixels for both platforms. We assumed this to span two standard deviations (95%), and so one standard deviation is 0.5 pixels ($\sigma_v = 0.5\,pixels$). Principal point and focal length were assumed to have a standard deviation of 2 pixels and 0.1 mm respectively. As previously mentioned, the uncertainty of the relative camera pose, $\mathbf{Q}_{\mathbf{t}_c^b}$, was temporarily set to zero (see Sections 2.3 and 2.4).

3.5. Outlier Removal

Unusually high reprojection errors were removed by an iterative process of outlier rejection. First optimisation was performed on all observations shown in Figure 5 for each platform. Reprojection errors, $e_{k,i}$ were calculated for each observation i of each pattern point k (see Equation (18)). These were then averaged per observation:

$$e_i = \frac{\sum_k^M e_{k,i}}{M}. \tag{25}$$

The observation i with the largest mean reprojection error was then removed from the data set and the process was repeated several times (i.e., optimise, calculate reprojection error, remove observation with largest reprojection error). The removal process may be stopped once all mean reprojection errors are below a threshold.

3.6. Mapping

To demonstrate mapping performance, rays were projected to a plane that was fitted to the estimated pattern point coordinates. Utilising the method in Section 2.3, the pattern points ($\mathbf{p}_k^w = [r_{k,x}^w, r_{k,y}^w, r_{k,z}^w]^T$) were first calculated given the data and relative camera pose . Using the general form of the equation of a plane $ax + by + cz + d = 0$, a best fit plane can be found in a linear least squares fashion (setting $c = -1$):

$$\mathbf{Ax} = \mathbf{b}, \qquad \begin{bmatrix} r_{1,x}^w & r_{1,y}^w & 1 \\ r_{2,x}^w & r_{2,y}^w & 1 \\ \vdots & \vdots & \vdots \\ r_{M,x}^w & r_{M,y}^w & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ d \end{bmatrix} = \begin{bmatrix} r_{1,z}^w \\ r_{2,z}^w \\ \vdots \\ r_{M,z}^w \end{bmatrix}, \tag{26}$$

where the plane parameters $\mathbf{x}$ can be solved for by left multiplying $\mathbf{b}$ with the pseudo-inverse of $\mathbf{A}$, $\mathbf{A}^+$.

The rays for observation i of pattern point k, defined by $\mathbf{p}_{c_{i,k}}^w$ and $\mathbf{p}_{s_{i,k}}^w$ as calculated in Section 2.1, can then be projected to the plane by computing their point of intersection:

$$\mathbf{p}_{proj_{i,k}}^w = \frac{([0, 0, -d/c]^T - \mathbf{p}_{c_{i,k}}^w) \cdot [a, b, c]^T}{(\mathbf{p}_{s_{i,k}}^w - \mathbf{p}_{c_{i,k}}^w) \cdot [a, b, c]^T}. \tag{27}$$

Knowing all input covariance matrices, including the covariance of the relative camera position and orientation, as obtained using MCMC (see Section 2.5), the uncertainty of each point projection $\mathbf{p}_{proj,i,k}^w$ may also be calculated. First the partial derivatives of Equation (27) with respect to all inputs were computed to yield the Jacobian $\mathbf{J}_{p_{proj,i,k}}$ of the x, y and z position of each point. The uncertainty of each projected point can then be calculated:

$$\mathbf{\Sigma}_{\mathbf{p}_{proj_{i,k}}^w} = \mathbf{J}_{p_{proj_{i,k}}} \mathbf{Q}_{k,i} \mathbf{J}_{p_{proj_{i,k}}}^T, \tag{28}$$

where

$$\mathbf{Q}_i = \begin{bmatrix} \mathbf{Q}_{uv,k,i} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\mathbf{t}^w_{b,k,i}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{int} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_{\mathbf{t}^b_c} \end{bmatrix}. \tag{29}$$

3.7. Comparing Poses

To assess how close a solution is to the optimal, we require a distance metric between two different 6 DOF pose transforms. As described in Section 2.2, each pose vector is composed of three translation and three orientation parameters. Given two unique poses $\mathbf{t_1} = [r_{1,x}, r_{1,y}, r_{1,z}, \theta_1 e_{1,x}, \theta_1 e_{1,y}, \theta_1 e_{1,z}]^T$ and $\mathbf{t_2} = [r_{2,x}, r_{2,y}, r_{2,z}, \theta_2 e_{2,x}, \theta_2 e_{2,y}, \theta_2 e_{2,z}]^T$, we can compare the translation parts easily by computing their Euclidean distance:

$$d_{1,2} = \sqrt{(r_{2,x} - r_{1,x})^2 + (r_{2,y} - r_{1,y})^2 + (r_{2,z} - r_{1,z})^2}. \tag{30}$$

However, measuring the distance or difference between two rotations is more complicated, and the readers are referred to Huynh [34] for an in-depth analysis of the topic. Huynh [34] presents and recommends a number of metrics for comparing rotations. Of these, the geodesic on the unit sphere was chosen, because it represents the magnitude of the rotation angle required to align the two rotations, which was deemed to be an intuitive measure. It can be computed as follows:

$$\Phi_{1,2} = 2 \arccos(|\mathbf{q}_1 \cdot \mathbf{q}_2|), \tag{31}$$

where $\mathbf{q}_1$ and $\mathbf{q}_2$ are unit quaternion equivalents of $[\theta_1 e_{1,x}, \theta_1 e_{1,y}, \theta_1 e_{1,z}]^T$ and $[\theta_2 e_{2,x}, \theta_2 e_{2,y}, \theta_2 e_{2,z}]^T$ respectively, computed as:

$$\begin{aligned} \mathbf{q}_i =& a_i + b_i \mathbf{i} + c_i \mathbf{j} + d_i \mathbf{k} \\ =& \cos\left(\frac{\theta_i}{2}\right) + e_{i,x} \sin\left(\frac{\theta_i}{2}\right) \mathbf{i} + \\ & e_{i,y} \sin\left(\frac{\theta_i}{2}\right) \mathbf{j} + e_{i,z} \sin\left(\frac{\theta_i}{2}\right) \mathbf{k}. \end{aligned} \tag{32}$$

$\Phi_{1,2}$ can also be interpreted to be equal to the absolute value of the angular magnitude θ (in the range $[-\pi, \pi]$) of the axis-angle rotation required to align the two orientations. For this reason, the metric will be simply referred to as the axis-angle difference. Combined, $d_{1,2}$ and $\Phi_{1,2}$ form a 2D pose distance that is convenient for visualisation.

3.8. Basin of Attraction

Because the method in this paper requires an approximate initial camera pose, it is important to numerically quantify how precise this initial camera pose must be to yield an accurate optimised estimate. To measure how far an initial hand measured camera pose can be from the optimum, while still resulting in correct global convergence, we test a set of starting conditions that are altered by different amounts, and measure how close the optimal result is from the known global optimum. Due to the high dimensionality of the search space, a random sub-sampling of initial poses is performed. Deviation of initial values from the known optimum is quantified in the two dimensional pose distance space defined in Section 3.7. The 2D space was sampled uniformly in a grid, and for each location a 6D initial parameter vector was randomly generated at the corresponding Euclidean and axis-angle distance from the known optimal value.

First an x, y and z translation vector was generated at random, uniformly distributed over an equal negative to positive range for all three parameters. The vector was normalised to unity and then multiplied by the Euclidean distance value of the corresponding grid position. The resulting vector was added to the optimised translation parameters, yielding the initial coordinates. Similarly, three axis-angle orientation values were randomly generated in the same way, normalised to unity and multiplied by the corresponding axis-angle difference value at the given grid location, yielding an axis-angle rotation. The optimised orientation parameters were then rotated by this difference rotation, producing the initial orientation values.

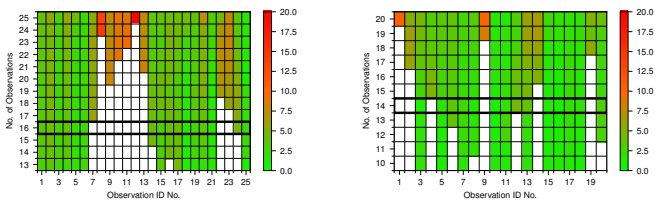
This yields a set of sparse random 6 DOF samples that are uniformly spaced in terms of pose distance from the known optimal camera pose, allowing the basin of attraction to be mapped. Optimisation was performed for each randomly generated initial camera pose on the grid. For each result, the Mahalanobis distance to the reference optimum was calculated, given the covariance matrix resulting from the MCMC uncertainty estimation on the optimised parameters.

4. Results

This section presents the results of line scan camera pose estimation for two different platforms and configurations, including outlier rejection, resulting camera pose and uncertainty, in-depth analysis of the uncertainty, the impact of platform pose diversity on the accuracy, and finally the combined mapping uncertainty.

4.1. Outlier Rejection

The iterative results of outlier removal based on reprojection errors are shown in Figure 8. Each row represents an outlier removal iteration, labelled by the number of remaining observations, where the top row includes all observations. Each column represents one of the observations from Figure 5, where each observation is one view of all 15 points on the calibration pattern. The colour of each cell indicates the mean reprojection error of the 15 points within the single observation of the calibration pattern, for a particular outlier removal iteration. In each figure, the black rectangle highlights the row with the greatest number of observations that all have mean errors less than a 5 pixel threshold. Ladybird exhibited a greater number of outliers and higher worst-case reprojection errors than Shrimp, with 9 compared to 6 outliers respectively. This resulted in 16 inliers for Ladybird and 14 for Shrimp.



(**a**) Ladybird (**b**) Shrimp

Figure 8. Outlier removal and average reprojection errors for (**a**) Ladybird and (**b**) Shrimp. Each row represents an iteration as described in Section 3.5, labelled with the number of remaining observations, and each column represents an observation of the calibration pattern as per Figure 5. White cells indicate that an observation has been removed for that iteration. Otherwise, the cell colour indicates the mean reprojection error in pixels for that particular observation of the calibration pattern at a given outlier removal iteration. For example, in (**a**), observation No. 12 exhibited the greatest mean reprojection error with 25 observations, and was therefore the first observation to be removed. The highlighted rectangle points to the row with the greatest number of remaining observations with a mean reprojection error of less than 5 pixels.

4.2. Pose and Uncertainty Results

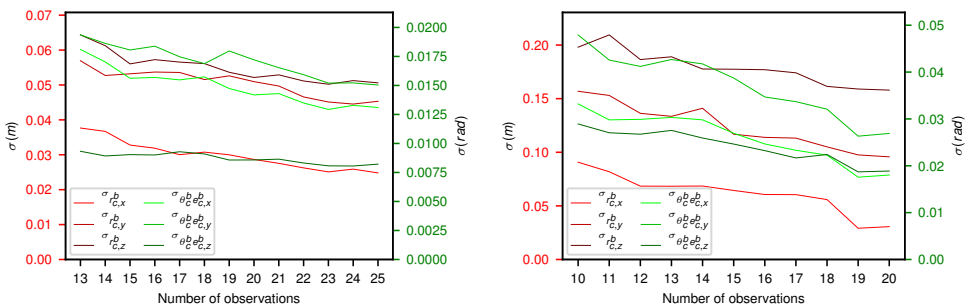
The relative camera pose transforms and associated uncertainties (one standard deviation) for both platforms are shown in Table 3 before and after outlier removal. The hand measured estimate is also shown for reference, where the tolerances reflect the difficulty of measurement. The results for the Shrimp platform exhibit greater uncertainty compared to the Ladybird platform.

Table 3. t_c^b estimates and uncertainties (axis-angle rotations).

	$r_{c,x}^b$ (m)	$r_{c,y}^b$ (m)	$r_{c,z}^b$ (m)	$\theta_c^b e_{c,x}^b$ (rad)	$\theta_c^b e_{c,y}^b$ (rad)	$\theta_c^b e_{c,z}^b$ (rad)
Ladybird						
All observations	0.147 ± 0.025	-0.128 ± 0.045	-0.630 ± 0.051	-0.849 ± 0.013	0.768 ± 0.015	-1.420 ± 0.008
16 observations	0.189 ± 0.032	-0.142 ± 0.054	-0.794 ± 0.057	-0.822 ± 0.016	0.738 ± 0.018	-1.429 ± 0.009
Hand measured	0.200 ± 0.100	0.000 ± 0.100	-0.800 ± 0.100	-0.762 ± 0.039	0.762 ± 0.039	-1.433 ± 0.037
Shrimp						
All observations	0.044 ± 0.031	-0.133 ± 0.096	-0.660 ± 0.158	1.409 ± 0.018	1.400 ± 0.027	-1.078 ± 0.019
14 observations	-0.010 ± 0.069	-0.080 ± 0.141	-0.579 ± 0.178	1.380 ± 0.030	1.427 ± 0.042	-1.093 ± 0.026
Hand measured	0.000 ± 0.100	-0.200 ± 0.100	-0.500 ± 0.100	1.399 ± 0.026	1.399 ± 0.088	-1.074 ± 0.071

Note: Hand measured orientation uncertainties are 2° for each parameter in Euler representation, converted to axis-angle representation by propagating the covariance matrix using the Jacobians of the conversion function.

In Figure 9 each outlier removal stage is plotted against each parameter's standard deviation. As would be expected, increasing the number of observations decreases the uncertainty of the estimate.



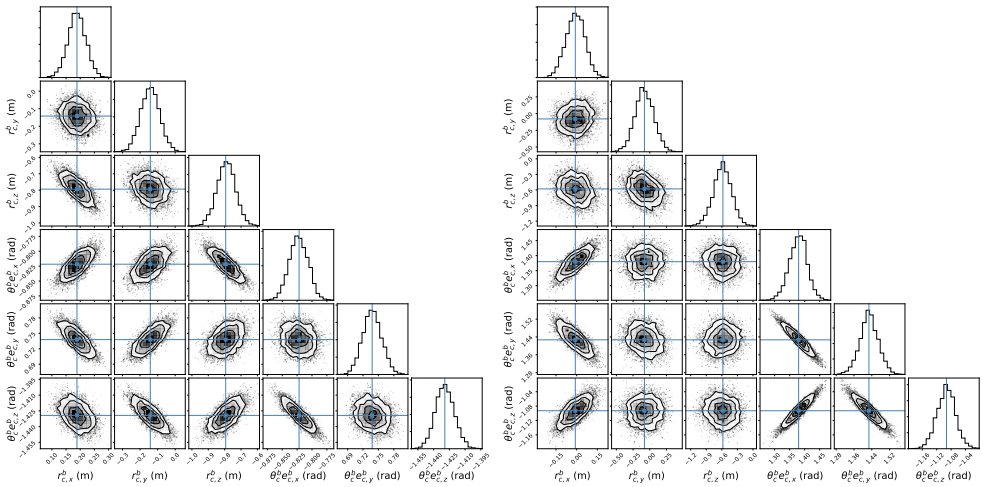
(**a**) Ladybird (**b**) Shrimp

Figure 9. The number of observations vs. standard deviation for (**a**) Ladybird and (**b**) Shrimp respectively. The x axis mirrors the outlier removal stages shown in Figure 8.

The number of observations affects the number of computations and therefore has a significant effect on calibration and MCMC run time. For Ladybird, optimisation and MCMC took approx. 15 min and 7 h respectively for all 25 observations. For 16 observations, this was reduced to just over 7 min and 4 h. For Shrimp, the respective optimisation and MCMC times were reduced from approx. 5.5 min and 5 h for all 20 observations to just over 2 min and 3 h with 14 observations.

4.3. In-Depth Uncertainty Analysis

Examining uncertainty in more detail, MCMC samples are shown on a corner plot in Figure 10 [48]. Each sub-plot below the diagonal provides a 2D histogram, showing the MCMC sample density between two parameters (i.e., the marginal likelihood distribution for a parameter pair), and on the diagonal a 1D histogram, giving the sample density for the marginal likelihood distribution for each single parameter.



(**a**) Ladybird (**b**) Shrimp

Figure 10. Corner plots for (**a**) Ladybird and (**b**) Shrimp platforms for 16 and 14 observations respectively. Each sub-plot below the diagonal provides a 2D histogram of MCMC sample values for a pair of relative camera pose parameters. The sub-plots on the diagonal show 1D histograms for each parameter.

For human interpretation of the uncertainty, in Figure 11a,d the MCMC sample values for $[\theta e^b_{c,x}, \theta e^b_{c,y}, \theta e^b_{c,z}]^T$ are plotted on a sphere. For visualisation only, each point is coloured according to a kernel density estimate (KDE) performed on all MCMC samples to give an indication of the axis-angle vector marginal likelihood. Hand measured pose estimates are shown as red crosses. Likewise, in Figure 11c,f, values for θ^b_c are presented in a histogram, showing the marginal likelihood of the axis-angle magnitude. While both figures indicate a clustering to within two degrees, the Shrimp platform's distribution exhibits a more elongated elliptical shape, while for Ladybird, it is more uniformly spread in all directions. Also apparent for Ladybird is that the manual measurements are well outside the region of high likelihood, both in terms of the axis-angle unit vector and magnitude. Conversely, the hand measured pose for the Shrimp platform is highly likely given the data, suggesting the initial manual measurement may have been more accurate than for Ladybird.

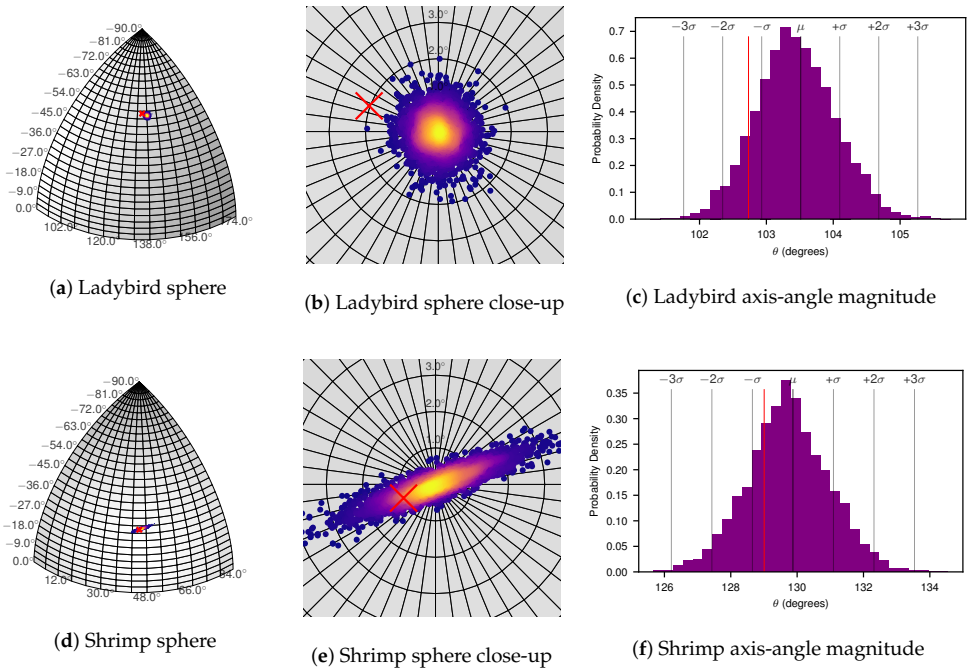
In Figure 12, the distributions of MCMC samples is shown superimposed on the corresponding platform model. Translation parameters are presented as a 2D histogram, similar to Figure 10, demonstrating the marginal density of the likelihood distribution from each orthogonal viewpoint. To present the orientation parameter distribution, line segments coincident with the camera's viewing direction, and anchored to the optimized camera centre, are rotated by each MCMC rotation sample. Each line is semi-transparent, and so as all samples build up, the density distribution of the camera orientation is visualised. It is evident that there is greater variance in the MCMC samples for Shrimp when compared to the Ladybird platform, particularly for the translation parameters, as corroborated by the numerical results in Table 3.

4.4. Effect of Angular Diversity

To investigate how angular diversity of body poses in a dataset affects the certainty of the result, two experimental subsets were compiled from the outlier-rejected dataset with 16 observations for Ladybird. The first includes only ten observations with small roll angles $\phi^w_{b,x} < 1.9°$. The second includes five observations with small roll $\phi^w_{b,x}$ and five with roll angles $3.3° < \phi^w_{b,x} < 5.8°$. Both datasets therefore contain ten total observations, representing low and high angular diversity. Optimisation results for these subsets are shown in Table 4. Despite containing the same number of observations,

Sensors **2017**, *17*, 2491

the dataset with high angular diversity results in significantly lower uncertainty for the optimal camera pose, compared to the low diversity set.



(**a**) Ladybird sphere (**b**) Ladybird sphere close-up (**c**) Ladybird axis-angle magnitude

(**d**) Shrimp sphere (**e**) Shrimp sphere close-up (**f**) Shrimp axis-angle magnitude

Figure 11. Ladybird and Shrimp platform MCMC axis-angle unit vector ($\mathbf{e_s^b}$) samples plotted on a sphere (**a**,**d**) and close-up, centred to a pole (**b**,**e**) (using 16 and 14 observations for Ladybird and Shrimp respectively). Each point on the sphere is coloured according to a KDE to give an indication of the axis-angle vector marginal likelihood for visualisation purposes only. Points were randomly thinned by ten to facilitate plotting. A histogram of the magnitudes of the axis-angle rotation, θ_s^b, is shown in (**c**,**f**). The hand measured pose orientation is shown with red crosses in (**a**,**b**,**d**,**e**), and a vertical red line in (**c**,**f**).

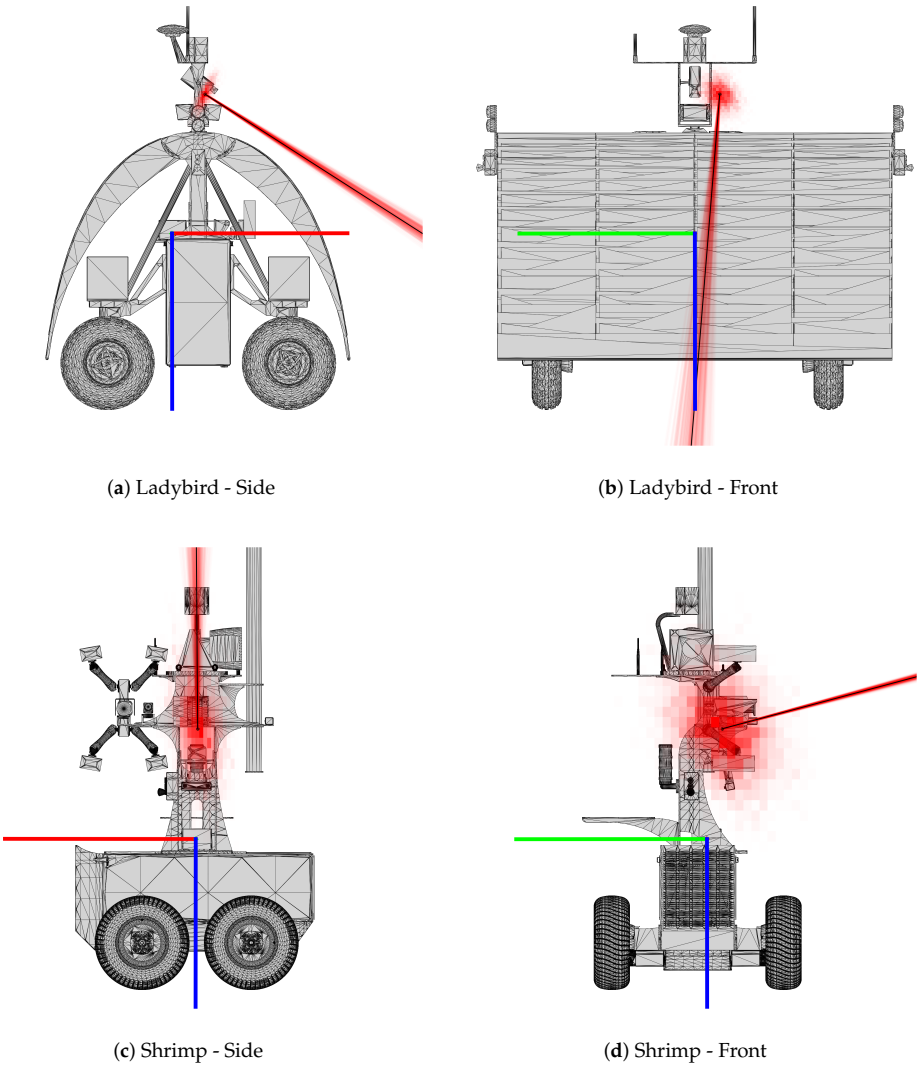
Table 4. Angular diversity comparison.

	$r_{c,x}^b$ (m)	$r_{c,y}^b$ (m)	$r_{c,z}^b$ (m)	$\theta_c^b e_{c,x}^b$ (rad)	$\theta_c^b e_{c,y}^b$ (rad)	$\theta_c^b e_{c,z}^b$ (rad)
High ang. diversity	0.166 ± 0.041	-0.155 ± 0.083	-0.717 ± 0.074	-0.831 ± 0.023	0.744 ± 0.023	-1.420 ± 0.015
Low ang. diversity	0.147 ± 0.118	0.058 ± 0.107	-0.269 ± 0.281	-0.797 ± 0.051	0.796 ± 0.059	-1.452 ± 0.026

4.5. Repeatability

To assess the repeatability of the approach, another two experimental subsets were compiled from the outlier-rejected dataset with 16 observations for Ladybird. They each contained five observations with small roll $\phi_{b,x}^w$ and three with roll angles $3.3° < \phi_{b,x}^w < 5.8°$. The two subsets do not share any observations, making them independent. Optimisation results for these subsets are shown in Table 5. The two results are consistent with each other, as the two distributions overlap to a significant degree. Due to the small dataset size, uncertainty values are higher than the result in Table 3 with all 16 observations, which is expected as demonstrated in Section 4.3.

Sensors **2017**, *17*, 2491



(**a**) Ladybird - Side

(**b**) Ladybird - Front

(**c**) Shrimp - Side

(**d**) Shrimp - Front

Figure 12. Optimised camera pose and MCMC sample poses for Ladybird (**a**,**b**) and Shrimp (**c**,**d**). The optimised pose is shown as an *xyz* axis with black lines. The MCMC samples are shown as red, green and blue *xyz* axes, where greater colour intensity corresponds to greater sample density, approximating the marginal likelihood of the poses in the viewing direction. The body (i.e., navigation system) pose is also shown as a red, green and blue *xyz* axis.

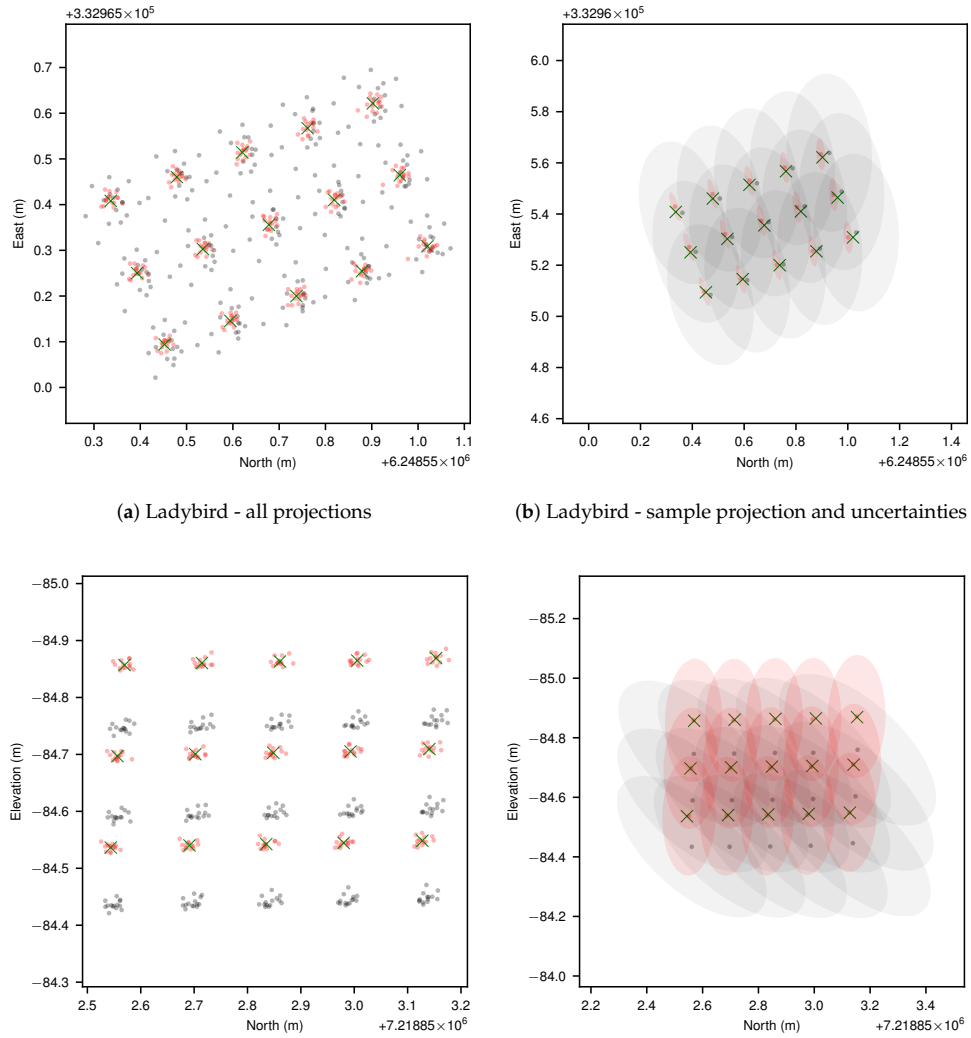
Table 5. Repeatability.

	$r^b_{c,x}$ (m)	$r^b_{c,y}$ (m)	$r^b_{c,z}$ (m)	$\theta^b_c e^b_{c,x}$ (rad)	$\theta^b_c e^b_{c,y}$ (rad)	$\theta^b_c e^b_{c,z}$ (rad)
Subset 1	0.192 ± 0.046	-0.170 ± 0.095	-0.702 ± 0.093	-0.828 ± 0.024	0.735 ± 0.028	-1.420 ± 0.015
Subset 2	0.206 ± 0.071	-0.306 ± 0.123	-0.819 ± 0.112	-0.839 ± 0.034	0.693 ± 0.040	-1.408 ± 0.022

Sensors **2017**, *17*, 2491

4.6. Mapping Accuracy

Figure 13 shows Ladybird's and Shrimp's observations, after outlier removal, projected to the best fit plane of predicted point locations before and after calibration (see Section 3.6). Figure 13a,c give projections of all observations, while Figure 13b,d only show the projections for one observation, but include uncertainty ellipses. Post calibration average pattern point estimates are also shown as green crosses.



(**a**) Ladybird - all projections

(**b**) Ladybird - sample projection and uncertainties

(**c**) Shrimp - all projections

(**d**) Shrimp - sample projection and uncertainties

Figure 13. Point to plane projection comparisons for both platforms. In (**a**,**c**) all scans of all pattern points are plotted to best fit planes before (grey) and after (red) calibration, while in (**b**,**d**) a single projected observation of each pattern point is shown with 1-σ uncertainty ellipses. In all figures, post-calibration pattern point estimates are marked with green crosses for reference. Because the calibration pattern was positioned flat on the ground for Ladybird, a top down (North-East) view was selected for (**a**,**b**). Conversely, for Shrimp the calibration pattern was positioned almost vertically on a ladder, facing west, and therefore a side view (North-Elevation) was selected for (**c**,**d**).

For both platforms, the calibrated camera pose exhibited more densely bound projected points. The change in spread is less pronounced for the Shrimp platform, because as mentioned previously the manual measurements were by chance much closer to the optimum values, though a significant offset can be observed between manual and calibrated results. The plots also demonstrate the effect of relative camera pose uncertainty on mapping uncertainty, which were significant for both Ladybird and Shrimp.

Given the hand measurement for shrimp happened to be close to the optimal, the effect of adding just one degree of error to the measured axis-angle vector is shown in Figure 14, which compares the mapped points from the optimised and erroneous camera pose. The optimisation improves the cluster significantly compared to a hand measured pose with one degree error.

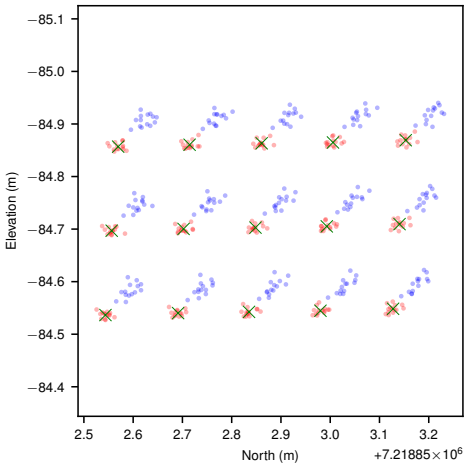
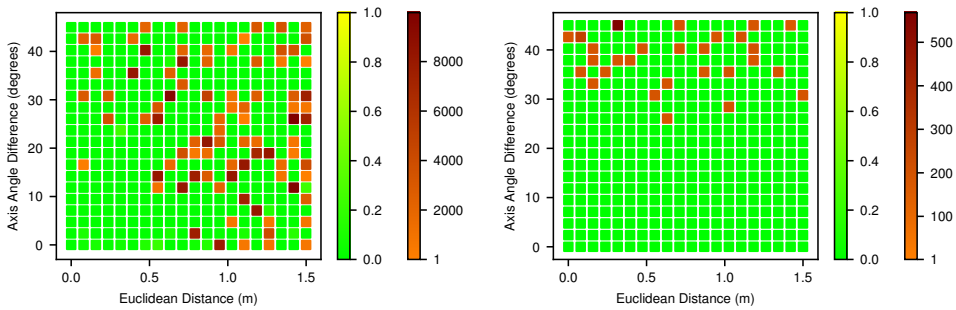


Figure 14. All pattern points plotted to best fit planes for Shrimp, red for post calibration (same as in Figure 13c) and blue for projected points resulting from a relative camera pose where the optimised orientation was altered by 1°. This demonstrates the significant effect small changes in the camera pose can have on mapping performance.

4.7. Basin of Attraction

Basin of attraction plots, which were generated as described in Section 3.8 using the Powell optimiser, are shown in Figure 15 for both platforms. The Mahalanobis distances were generally either close to zero or very large, so they are colour coded into two tiers, below and above 1.0, to improve readability. The basin for the Ladybird platform (Figure 15a) shows success can be expected in the triangular region with less than 20° and 0.5 m of hand measurement error. The basin for Shrimp (Figure 15b) shows a greater immunity to translation errors and successful results can be expected with initialisation errors less than 20° and as high as 1.5 m. Both figures indicate that when the initialisation error is higher, there is still a high chance (approx. 60%) of a solution within a Mahalanobis distance of 0.1, but it cannot be relied upon, and deteriorates as the distance from optimum increases.



(**a**) Ladybird (**b**) Shrimp

Figure 15. Basins of attractions for the (**a**) Ladybird and (**b**) Shrimp platforms. The x and y axes of the plots are the Euclidean distances and axis-angle rotational differences between the initial values and the optimal reference solution respectively. Each grid cell is colour coded into two tiers based on the Mahalanobis distance of the result to the known optimum: below 0.1 and above, which was chosen as a suitable threshold for optimisation success.

5. Discussion

The results show that the proposed method was able to to reliably estimate the relative line scan camera pose on a mobile ground vehicle, resulting in a reduction in mapping error, as long as the calibration data includes sufficient viewpoint variability. An uncertainty of 0.06 m/0.018 rad (1.05°) for Ladybird, and 0.18 m/0.042 rad (2.39°) for Shrimp was achieved. This result is dependent on the certainty of input parameters, which include pixel observations, navigation system solutions, and camera intrinsics. For example, confidence in the calibrated pose parameters for the Ladybird platform was significantly greater than for the Shrimp platform, due to Ladybird's higher grade IMU, which allowed the navigation system to provide more certain solutions.

The results show that it is necessary to examine reprojection errors and remove outliers, as is common with many camera calibration approaches. Outliers statistically fall outside the assumptions encoded in their respective error models, and so the mean of the final camera pose distribution is pulled in the wrong direction. A number of observations for both platforms exhibited high reprojection errors relative to other observations (>approx. 8 pixels). These errors could be caused by manual labelling inaccuracies (e.g., due to limited resolution), navigation system solution errors that incorrectly fall outside the reported navigation uncertainty, or a combination thereof. Removing outliers had significant effects on the results, evident particularly in the correction of some parameters such as the z-offset $\mathbf{t}^b_{c,z}$ for both platforms (see Table 3). The results shown in Figure 8 also support the iterative removal of outliers at each stage. For instance, reprojection errors of Ladybird observation 20 improve as other observations are removed, while Shrimp observation 4 degrades, and was eventually removed. Conversely, as shown in Figure 9 a larger number of observations, outliers or not, allows for greater certainty in the final result. Thus, there are two competing factors when performing outlier rejection. A sufficiently large number of observations is required to maintain an acceptable level of certainty, yet removing outliers is important to minimise reprojection errors. It is, therefore, desirable to obtain a sufficient number of observations to allow for subsequent outlier removal. The additional computational time required when increasing the number of observations is also a consideration. The ending condition threshold should be chosen such that significant outliers are removed, but a sufficient number of observations remain. In this paper a value of 5 pixels was empirically determined as an appropriately balanced threshold given the data.

The main product of the MCMC uncertainty analysis is a covariance matrix (Equation (23)), which can be used to estimate mapping accuracy (e.g., Figure 13a,c). However, covariance matrices represent

uncertainty in a compressed format, given the assumption that the likelihood function is normally distributed. The corner plots (Figure 10) provide a direct view of the MCMC sampling result, which qualitatively confirm that for both vehicles the normality assumption is justified: specifically that the distributions behave linearly within the sampled region, and the 1D histograms are qualitatively Gaussian in shape.

In this paper we propose methods of visualising sensor pose distributions in a human interpretable way, as depicted in Figures 11 and 12. The sphere plots and associated axis-angle magnitude histograms in Figure 11 present the same underlying data in the MCMC sample plots (Figure 10), but focus on human interpretability of the orientation parameters. The sphere provides a relatable reference that demonstrates how closely clustered the pose orientation is. Similarly, the visualisation in Figure 12 allows for human interpretation of the resulting camera pose and uncertainty (likelihood distribution) with respect to the platform models. These figures particularly highlight the greater uncertainty in translation parameters compared to orientation. They also confirm that the solutions are qualitatively "sensible" with respect to the physical platform models.

The primary objective of optimising the camera pose is to reduce mapping errors. This was demonstrated in the results by the tighter clustering of mapped calibration target points that was achieved post-calibration. The improvement was particularly noticeable for the Ladybird platform, and to a lesser extent for the Shrimp platform. By chance, the manually measured camera pose on Shrimp was much closer to the optimal result than it was for Ladybird, and so the mapping improvement for Shrimp was less pronounced. The camera location on the Shrimp platform was easier to access, due to the lower height and smaller footprint, compared to the Ladybird platform, which likely explains the better manual estimate. Nevertheless, such accurate manual measurements can typically not be guaranteed, and Figure 14 reveals the sensitivity of the map to small errors in camera pose, highlighting the need for calibration.

The results reinforce the importance of acquiring a calibration dataset that exhibits a wide variety of platform poses with respect to the calibration pattern. This was tested by optimising both with and without large body roll ($\phi^{w}_{b,x}$) observations. Removing high $\phi^{w}_{b,x}$ observations had a considerable effect, as shown in Table 4, where uncertainty approximately doubled and even tripled for some parameters.

The proposed method is able to deal with a wide range of initial hand measured values when paired with the Powell optimisation algorithm. In Figure 15 we propose an intuitive approach to visualising the basins of attraction, by reducing the 6 DOF initial parameter space to form a 2D pose-distance space, comprising Euclidean and axis-angle distance. The plots demonstrate that initial estimates that deviate up to 0.5 m or 20° are likely to result in a successful optimisation. Additionally, with even larger deviations there is still a better than even (60%) chance of success. However, this is highly dependent on the geometry of the sensor and platform, the acquired data and the chosen optimisation algorithm. In our case, the solution for the Shrimp platform was surprisingly robust to initial translation errors (up to 1.5 m), while some failures can be seen at over 0.5 m for the Ladybird platform. This may be the result of the greater platform roll and pitch angles (up to 17°) achieved with Shrimp during data acquisition. In addition, different optimisers will have varying abilities to deal with local minima and "flat" regions in the 6 DOF parameter space. Nevertheless, measurement error tolerances of ±0.5 m and ±20° should be practically achievable for most applications.

As shown in Table 5, the camera pose estimates obtained from two independent datasets for the same platform were consistent with each other. The results therefore verify that the proposed approach is repeatable by demonstrating that different sets of data from the same platform yield consistent results.

An important advantage of the proposed method is that exact dimensions of the calibration pattern do not need to be known. As such, a planarity assumption or assumptions about the distances or geometry between points are not required. This simplifies the method in the field because it is not affected by printing errors or damage/warping of the pattern which affects the relative geometry of the points. Furthermore, a calibration pattern could be manually produced in the field if necessary.

One important condition, however, is that individual points are uniquely distinguishable in the line scan image data.

6. Conclusions

This paper demonstrated a novel method for estimating a rigidly mounted line scanning camera's fixed 6 DOF pose relative to a mobile platform with a navigation system. The method is appropriate for ground or very low altitude applications, where the scene is relatively near the platform, as it does not require GCPs and uses a compact calibration pattern, the dimensions of which do not need to be known. Furthermore, it does not require data from auxiliary sensors such as full frame cameras. The approach involves imaging a calibration pattern with distinctly identifiable points from various platform poses, and using the navigation system and image data to triangulate their positions in the world frame. Reprojecting the points to the camera yields reprojection errors, which are used as a basis for outlier rejection, and then to calculate the likelihood given a candidate camera pose. By minimising the negative log likelihood, the optimal relative camera pose can be obtained. Given the likelihood function, an MCMC algorithm is able to estimate the certainty of the camera pose. The results demonstrate the effectiveness of the approach using two different mobile platforms with differing mounting configurations. The method was shown to be robust to relatively inaccurate initial hand measurements (within 0.5 m and 20°). Additionally, a number visualisations have been proposed to aid in human interpretation of the results. Future work will attempt to precisely specify platform pose requirements prior to data collection, automate and improve the pattern point extraction process, and explore the application of a robust optimisation routine or loss function to simplify the outlier rejection process.

Acknowledgments: This work is supported by the Australian Centre for Field Robotics (ACFR) at The University of Sydney and by funding from the Australian Government Department of Agriculture and Water Resources as part of its Rural R&D for profit programme. For more information about robots and systems for agriculture at the ACFR, please visit http://sydney.edu.au/acfr/agriculture.

Author Contributions: Both A.W. and J.U. gathered the relevant data in the field. A.W. conceived and implemented the method, while J.U. supervised the work and provided significant conceptual input. The paper was written by A.W. and reviewed by J.U.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Yang, C.; Everitt, J.H.; Bradford, J.M.; Murden, D. Airborne hyperspectral imagery and yield monitor data for mapping cotton yield variability. *Precis. Agric.* **2004**, *5*, 445–461.
2. Bethel, J.; Lee, C.; Landgrebe, D.A. Geometric registration and classification of hyperspectral airborne pushbroom data. *Int. Arch. Photogramm. Remote Sens.* **2000**, *33*, 183–190.
3. Lim, M.S.; Lim, J. Visual measurement of pile movements for the foundation work using a high-speed line-scan camera. *Pattern Recognit.* **2008**, *41*, 2025–2033.
4. Dale, L.M.; Thewis, A.; Boudry, C.; Rotar, I.; Dardenne, P.; Baeten, V.; Pierna, J.A.F. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Appl. Spectrosc. Rev.* **2013**, *48*, 142–159.
5. Pfaff, F.; Maier, G.; Aristov, M.; Noack, B.; Gruna, R.; Hanebeck, U.D.; Längle, T.; Beyerer, J.; Pieper, C.; Kruggel-Emden, H.; et al. Real-time motion prediction using the chromatic offset of line scan cameras. *at-Automatisierungstechnik* **2017**, *65*, 369–380.
6. Li, D.; Wen, G.; Qiu, S. Cross-ratio—Based line scan camera calibration using a planar pattern. *Opt. Eng.* **2016**, *55*, 14104.
7. Ramirez-Paredes, J.P.; Lary, D.J.; Gans, N.R. Low-altitude terrestrial spectroscopy from a pushbroom sensor. *J. Field Robot.* **2015**, *33*, 837–852.

8. Deery, D.; Jimenez-Berni, J.; Jones, H.; Sirault, X.; Furbank, R. Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* **2014**, *4*, 349–379.
9. Wendel, A.; Underwood, J. Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 16–21 May 2016; pp. 5128–5135.
10. Trierscheid, M.; Pellenz, J.; Paulus, D.; Balthasar, D. Hyperspectral imaging or victim detection with rescue robots. In Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics (SSRR), Sendai, Japan, 21–24 October 2008; pp. 7–12.
11. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334.
12. Mostafa, M.M.R. Boresight calibration of integrated inertial/camera systems. In Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS 2001), Banff, AB, Canada, 5–8 June 2001; pp. 5–8.
13. Lobo, J.; Dias, J. Relative pose calibration between visual and inertial sensors. *Int. J. Robot. Res.* **2007**, *26*, 561–575.
14. Hol, J.D.; Schön, T.B.; Gustafsson, F. Modeling and calibration of inertial and vision sensors. *Int. J. Robot. Res.* **2010**, *29*, 231–244.
15. Draréni, J.; Roy, S.; Sturm, P. Plane-based calibration for linear cameras. *Int. J. Comput. Vis.* **2011**, *91*, 146–156.
16. Hui, B.; Wen, G.; Zhao, Z.; Li, D. Line-scan camera calibration in close-range photogrammetry. *Opt. Eng.* **2012**, *51*, 53601–53602.
17. Li, D.; Wen, G.; Hui, B.W.; Qiu, S.; Wang, W. Cross-ratio invariant based line scan camera geometric calibration with static linear data. *Opt. Lasers Eng.* **2014**, *62*, 119–125.
18. Luna, C.; Mazo, M.; Lázaro, J.L.; Vázquez, J.F. Calibration of line-scan cameras. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 2185–2190.
19. Yao, M.; Zhao, Z.; Xu, B. Geometric calibration of line-scan camera using a planar pattern. *J. Electron. Imaging* **2014**, *23*, 13028.
20. Sun, B.; Zhu, J.; Yang, L.; Yang, S.; Niu, Z. Calibration of line-scan cameras for precision measurement. *Appl. Opt.* **2016**, *55*, 6836–6843.
21. Ma, W.; Bioucas-Dias, J.; Chan, T.; Gillis, N.; Gader, P.; Plaza, A.; Ambikapathi, A.; Chi, C. A signal processing perspective on hyperspectral unmixing: Insights from remote sensing. *IEEE Signal Process. Mag.* **2014**, *31*, 67–81.
22. Lenz, A.; Schilling, H.; Perpeet, D.; Wuttke, S.; Gross, W.; Middelmann, W. Automatic in-flight boresight calibration considering topography for hyperspectral pushbroom sensors. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2981–2984.
23. Perry, J.; Ahmed, M.; Abd-Elrahman, A.; Bowman, S.; Kaddoura, Y.; Watts, A. Precision directly georeferenced unmanned aerial remote sensing system: Performance evaluation. In Proceedings of the Institute of Navigation National Technical Meeting, San Diego, CA, USA, 28–30 January 2008; pp. 680–688.
24. Muller, R.; Lehner, M.; Muller, R.; Reinartz, P.; Schroeder, M.; Vollmer, B. A program for direct georeferencing of airborne and spaceborne line scanner images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2002**, *34*, 148–153.
25. Abd-Elrahman, A.; Sassi, N.; Wilkinson, B.; Dewitt, B. Georeferencing of mobile ground-based hyperspectral digital single-lens reflex imagery. *J. Appl. Remote Sens.* **2016**, *10*, 14002.
26. Wohlfeil, J.; Bucher, T. A modular, interactive software-concept for radiometric and geometric correction of airborne and spaceborne linescanner images. In Proceedings of the Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IX (SPIE), Berlin, Germany, 31 August–3 September 2009; Volume 7478.
27. Yeh, C.K.; Tsai, V.J.D. Self-calibrated direct geo-referencing of airborne pushbroom hyperspectral images. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, BC, Canada, 24–29 July 2011; pp. 2881–2883.
28. Wohlfeil, J. Optical Orientation Determination for Airborne and Spaceborne Line Cameras. Ph.D. Thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Berlin, Germany, 2012.

29. Barbieux, K.; Constantin, D.; Merminod, B. Correction of airborne pushbroom images orientation using bundle adjustment of frame images. In Proceedings of the The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, Prague, Czech Republic, 12–19 July 2016; 2016 XXIII ISPRS Congress; Volume XLI-B3, pp. 813–818.
30. Habib, A.; Xiong, W.; He, F.; Yang, H.L.; Crawford, M. Improving orthorectification of UAV-based push-broom scanner imagery using derived orthophotos from frame cameras. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 262–276.
31. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: Cambridge, UK, 2003.
32. Berner, P. Technical Concepts: Orientation, Rotation, Velocity and Acceleration, and the SRM. Available online: http://www.sedris.org/wg8home/Documents/WG80485.pdf (accessed on 21 March 2017).
33. Underwood, J. *Reliable and Safe Autonomy for Ground Vehicles in Unstructured Environments*; University of Sydney: Sydney, Australia, 2009.
34. Huynh, D.Q. Metrics for 3D rotations: Comparison and analysis. *J. Math. Imaging Vis.* **2009**, *35*, 155–164.
35. Gellert, W.; Gottwald, S.; Hellwich, M.; Kästner, H.; Künstner, H. *VNR Concise Encyclopedia of Mathematics*, 2nd ed.; Van Nostrand Reinhold: New York, NY, USA, 1989.
36. Underwood, J.P.; Hill, A.; Peynot, T.; Scheding, S.J. Error modeling and calibration of exteroceptive sensors for accurate mapping applications. *J. Field Robot.* **2010**, *27*, 2–20.
37. James, F. *Statistical Methods in Experimental Physics: 2nd Edition*; World Scientific: Singapore, 2006.
38. Foreman-Mackey, D.; Hogg, D.W.; Lang, D.; Goodman, J. Emcee: The MCMC Hammer. *Publ. Astron. Soc. Pac.* **2013**, *125*, 306–312.
39. MacKay, D.J. *Information Theory, Inference, and Learning Algorithms*; Cambridge University Press: Cambridge, UK, 2008.
40. Underwood, J.; Wendel, A.; Schofield, B.; McMurray, L.; Kimber, R. Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. *J. Field Robot.* **2017**, *34*, 1061–1083.
41. Underwood, J.P.; Calleija, M.; Taylor, Z.; Hung, C.; Nieto, J.; Fitch, R.; Sukkarieh, S. Real-time target detection and steerable spray for vegetable crops. In Proceedings of the International Conference on Robotics and Automation: Robotics in Agriculture Workshop, Seattle, WA, USA, 26–30 May 2015.
42. Wendel, A.; Underwood, J. Illumination compensation in ground based hyperspectral imaging. *ISPRS J. Photogramm. Remote Sens.* **2017**, *129*, 162–178.
43. Stein, M.; Bargoti, S.; Underwood, J. Image based mango fruit detection, localisation and yield estimation using multiple view geometry. *Sensors* **2016**, *16*, 1915.
44. Bargoti, S.; Underwood, J.P. Image segmentation for fruit detection and yield estimation in apple orchards. *J. Field Robot.* **2017**, doi:10.1002/rob.21699.
45. Underwood, J.P.; Hung, C.; Whelan, B.; Sukkarieh, S. Mapping almond orchard canopy volume, flowers, fruit and yield using LiDAR and vision sensors. *Comput. Electron. Agric.* **2016**, *130*, 83–96.
46. Powell, M.J.D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* **1964**, *7*, 155–162.
47. Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open source scientific tools for Python.
48. Foreman-Mackey, D. corner.py: Scatterplot matrices in Python. *J. Open Source Softw.* **2016**, *24*, doi:10.21105/joss.00024.

Chapter 5

Crop-Weed Discrimination

One specific problem in precision agriculture (PA) is the treatment of weed species, which can affect yield and contaminate harvested crop. Traditional broadcast spraying approaches, which apply herbicides over entire fields, are economically inefficient and pose an environmental risk due to over-application [46]. site specific weed management (SSWM), which aims to only spray affected areas, is therefore an attractive alternative [26].

In order to perform SSWM, it is necessary to identify weeds and discriminate them from crop, which do not need to be sprayed with herbicide. Section 2.1.4 summarised some examples of hyperspectral imaging (HSI) being used for discrimination of plant species in general [144, 264] and specifically for SSWM [284, 207, 321]. The majority of the work in this area uses supervised methods, which requires the generation of manually labelled training datasets. These datasets are susceptible to temporal variability due to physiological or environmental changes over days or seasons. Consequently, more labelled data would need to added if new conditions are not yet present in a training dataset, which can be a labour intensive task.

The study in this chapter addresses this problem by introducing a self-supervised training method for detecting weed species using HSI data. While weed detection using more common imaging methods has been successful, they are usually based on spatial features such as texture and shape. This makes them susceptible to partial

occlusion, flexing or other detrimental effects. On the other hand, classification using HSI can be performed on a pixel-wise basis, making it largely immune to these concerns.

Integrating ideas from a previous paper that uses RGB imagery [47], the proposed approach automatically generates separate training sets for weeds and crops by leveraging prior knowledge about regular seeding patterns in row crop fields. This allows the classifier to be continually updated making it robust to the aforementioned physiological and environmental changes. The method was evaluated on data acquired in corn crop rows, using linear discriminant analysis (LDA) and support vector machine (SVM) classifiers, by comparing results to those obtained using manually generated datasets.

2016 IEEE International Conference on Robotics and Automation (ICRA)
Stockholm, Sweden, May 16-21, 2016

Self-Supervised Weed Detection in Vegetable Crops Using Ground Based Hyperspectral Imaging

Alexander Wendel and James Underwood

Abstract— **A critical step in treating or eradicating weed infestations amongst vegetable crops is the ability to accurately and reliably discriminate weeds from crops. In recent times, high spatial resolution hyperspectral imaging data from ground based platforms have shown particular promise in this application. Using spectral vegetation signatures to discriminate between crop and weed species has been demonstrated on several occasions in the literature over the past 15 years. A number of authors demonstrated successful per-pixel classification with accuracies of over 80%. However, the vast majority of the related literature uses supervised methods, where training datasets have been manually compiled. In practice, static training data can be particularly susceptible to temporal variability due to physiological or environmental change. A self-supervised training method that leverages prior knowledge about seeding patterns in vegetable fields has recently been introduced in the context of RGB imaging, allowing the classifier to continually update weed appearance models as conditions change. This paper combines and extends these methods to provide a self-supervised framework for hyperspectral crop/weed discrimination with prior knowledge of seeding patterns using an autonomous mobile ground vehicle. Experimental results in corn crop rows demonstrate the system's performance and limitations.**

I. INTRODUCTION

With an elevated awareness of environmental issues, food security and sustainability, and an ever-present desire to reduce costs and waste, and increase quality and productivity, precision agriculture (PA) has emerged as an important tool for managing variability in order to achieve these aims [1]. In this context, robotics, automation and image processing are becoming indispensable to providing high resolution data and facilitating timely and accurate response and treatment.

One particular focus of PA is site specific weed management, for which one of the greatest challenges is the accurate identification of plant species [2]. Various methods have been proposed, using different sensors including monochrome, colour, multispectral and hyperspectral cameras. Recently, hyperspectral imaging (HSI), which captures data in hundreds of narrow bands, has shown promising results when performing per-pixel classification, which is of particular advantage in situations where morphological features may perform poorly due to occlusion or other detrimental effects [3]. Furthermore, simpler, less computationally intensive methods may be employed, because unlike colour (RGB) imaging, individual pixel spectra already provide sufficient information for plant classification.

In PA a particular challenge for all methods is dealing with variable conditions, such as morphological and illumination changes, the consequence being that reference training data can not be universally applied. As a result, if conditions differ over days, months and seasons, and from location to location, a large volume of training data is required to cover the resulting variance. One way to deal with this issue is to automatically update the classifier in a self-supervised manner, allowing continuous adaptation to changing conditions.

Recent work has demonstrated a self-supervised framework, which learns the characteristics of each class by relying on the fact that commercial crops are typically planted in a strict linear and constant seeding pattern [4]. Thus training data can be extracted by assuming that crop pixels are concentrated near seed lines, while vegetation pixels between rows are highly likely to correspond to weeds. This paper applies these general principles to crop/weed discrimination with hyperspectral data, to achieve the best of both worlds; highly accurate per-pixel classification, without the need for manual labelling, and adaptability to illumination, seasonal, geographical and species variations.

II. BACKGROUND

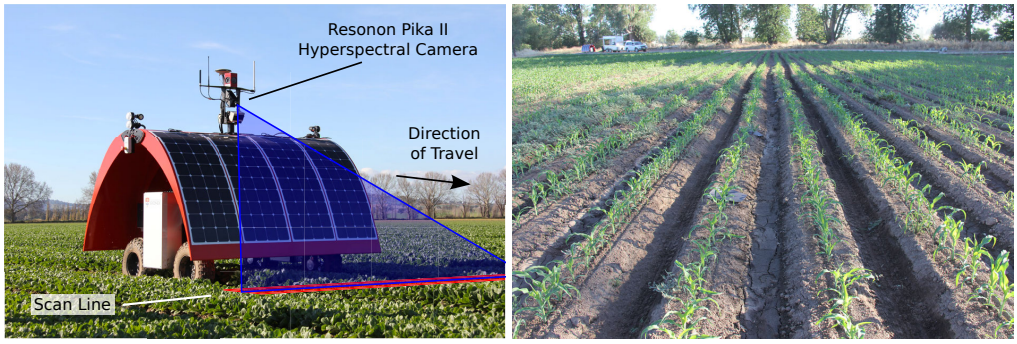
The last decade has seen much research supporting the feasibility of differentiating plant species using their spectra only. There is now significant research indicating good classification performance when using spectral imaging techniques for discrimination of plant species [6, 7, 8, 3]. There are many papers covering classification using data from both multispectral and hyperspectral sensors. Multispectral imaging provides relatively few broad spectral bands, which can be specifically targeted to a particular application. HSI, on the other hand, often involves several hundreds of very narrow spectral bands.

In one of the earliest works, a multispectral imaging system was used to classify weeds and sugar beet crop [9]. Classifiers including k-Nearest Neighbour (kNN), minimal distance, and multi-layer neural network with non-linear mapping (MLNLM) were compared and MLNLM produced the highest success rates. Reflectances at 441, 446, 459, 883, 924 and 988 nm yielded MLNLM classification success rates of 80.1% and 91.4% for crop and weed classes respectively.

Later, four species of weed among sugar beet crop were discriminated with success rates from 74.7% to 97.3% for weeds and 81.3% for crops [10]. The data was transformed

This work is supported by the Australian Centre for Field Robotics at the University of Sydney, AUSVEG and Horticulture Innovation Australia through project VG12104 An Intelligent Farm Robot for the Vegetable Industry.

Alexander Wendel and James Underwood are with the Australian Centre for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering at The University of Sydney, NSW 2006 Australia a.wendel, and j.underwood@acfr.usyd.edu.au



(a) Platform Overview (b) Typical Corn Crop Rows

Fig. 1. (a) The Ladybird robot used to autonomously gather the hyperspectral data [5]. (b) Rows of corn that are typical of the data in this paper. Each hyperspectral scan line crosses two rows at a time.

using wavelet analysis, followed by stepwise selection of the most significant features. Classification was achieved using Linear Discriminant Analysis (LDA). LDA has also been compared against a neural network (NN) based classifier on similar data (sugar beet crop and four species of weed) [11]. In addition, raw data (no transformation) and data transformed by Principal Component Analysis (PCA) were compared by feeding both into the two classifiers. While the PCA-NN pipeline performed best in a two class discrimination scenario (weed and crop), with success rates from 90.2% to 99.1% depending on the weed type, PCA-LDA was similarly accurate with success rates from 89.3% to 97.0%.

Comparing NN to maximum likelihood classification (MLC) with canola, pea and wheat crops, and two species of weed, NN demonstrated best performance with overall accuracies of 73.7% to 95.4% when trained and tested on the same date [12]. Partial least squares discriminant analysis (PLS-DA) has been implemented for classifying wheat crop and weeds, achieving a total accuracy of 72% [13].

There is also a large body of literature supporting the use of monochrome or RGB imaging for weed detection. High classification rates can be obtained by extracting features based on leaf shape, texture or vein patterns [3]. For example, Gabor wavelet and Gradient Field Distribution (GFD) techniques have been combined to detect weeds with an overall accuracy of 94% [14]. In other research, Active Shape Models (ASM), which are deformable templates that can only change according to a statistical model learned from the training data, have also been leveraged to successfully discriminate weed from crop with accuracies from 65% to over 90% depending on exact method and plant species [15, 16].

Random Forest classifier has been implemented to estimate crop/weed certainty at sparse pixel positions, which is then smoothed using a Markov Random Field method [17]. Shape, contour, and statistical features were calculated for image tiles extracted from a sparse grid. Fully connected plant regions were interpolated to obtain the final classified image. In testing, a high average classification accuracy of 93.8% was achieved.

One issue that many of the papers to date do not address yet is coping with temporal and geographical variability of plant characteristics. All methods mentioned so far require a process of training. Therefore, final classification performance is closely linked to how representative a training dataset is of current conditions. Both spectral signatures and morphological features change as plants grow. Also, due to changing uncontrolled conditions present at commercial farms, plant colour can be subject to considerable variability over time and from location to location. [18] found that changing conditions over multiple seasons can cause unacceptable degradation of performance (up to 36% total error rate). While it was determined that a global calibration procedure improved performance, the authors went further to design a multi-classifier system with three canonical Bayesian classifiers optimized for three seasons individually. A total discrimination accuracy of 95.8% was demonstrated with the multi-classifier system. However, the method requires historical data from several seasons.

More recently, morphological features extracted from colour images have been utilised to perform self-supervised weed detection in corn and soybean fields [4]. The authors leverage the fact that in modern agriculture most crops are sown in strict patterns by automated methods. By employing the Hough transform, linear crop rows were automatically detected. Pixels are labelled as outside or inside the row using rules based on their distance from the row centres, and probability distributions are calculated for each group via the Parzen-Rosenblatt window method [19, 20]. By assuming data points located outside rows are exclusively weeds, the weed probability distribution in feature space is known. The crop probability distribution was inferred by subtracting the known weed distribution from the inside-row distribution, which is a mixture of crop and weed. A high classification rate was achieved on fields with varying levels of infestation, using a combination of Naive Bayes and Gaussian Mixture

Model (GMM) classification. The system would therefore be able to adjust to conditions present at individual farms at a given time (no historical datasets are required).

To the best knowledge of the authors, similar self-supervised training techniques have not been employed for ground based weed detection with hyperspectral data, which can offer some useful advantages over colour or monochrome imaging. Relying on morphological features can adversely affect classification where plants occlude each other, are damaged by insects or weather, or bent into unidentifiable shapes [21, 22, 18]. Classification on HSI, which is done for each individual pixel, is mostly invariant to these effects. Additionally, for standard imaging methods, feature vectors must first be generated, which can be computationally intensive. Conversely, raw hyperspectral data already contains a rich set of features (bands), which can require less complex preprocessing.

The methodology in this paper draws upon previously established concepts to automatically detect and pre-process training data from hyperspectral images obtained with a ground based robot [4]. The subsequent classification pipeline itself is based on prior work in supervised hyperspectral crop/weed discrimination [11]. Therefore, the main contribution of this work is a self-supervised training data generation and weed detection system that is suitable for hyperspectral data obtained from vegetable fields. To demonstrate the system, it was tested on rows of corn.

III. MATERIALS AND METHODS

A. Overview

The objective of the system is to classify the pixels in ground based hyperspectral imagery, as either belonging to crop or weed. A line scan hyperspectral camera mounted to the autonomous mobile ground vehicle Ladybird™[5] (see Fig. 1), traversing corn crop fields was used to gather the raw hyperspectral data. The resulting line scans ($N_{bands} \times N_{pixels}$) were stacked to create three dimensional data cubes ($N_{bands} \times N_{pixels} \times N_{linescans}$). A software pipeline was implemented to pre-process, extract training data, and classify each pixel (see Fig. 2).

Training data was generated automatically using the row detection and extraction method described in Section III-E. A manually labelled training and testing data set was also assembled as described in Section III-C.

The hyperspectral data preprocessing and classification methodologies take cues from previous work in the literature [11]. First, vegetation pixels are separated from background pixels via Normalised Difference Vegetation Index (NDVI) thresholding. Individual spectra means and minima are then normalised to 1 and 0 respectively, prior to feature extraction using PCA. Finally, classification is performed using an LDA classifier. For the implementation, the open source machine learning package scikit-learn [23] and image processing package scikit-image [24] were used.

Classification using manual training data was performed to form a benchmark basis against which to compare the automatically collected training data set.

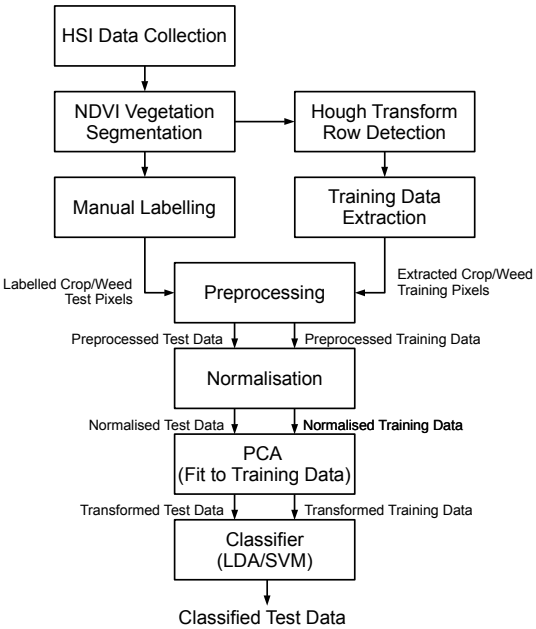


Fig. 2. Summarised self-supervised processing pipeline for the weed detection method presented in this paper. Non-vegetation pixels are first removed using NDVI. Portions of the data are then manually labelled as either weed or crop to be used as test data. Separately, the NDVI separated data is processed to automatically extract the training dataset. Both the test and training datasets are preprocessed, normalised and transformed with PCA. Finally the classifer is trained with the automatically extracted data, and prediction is subsequently performed on the test data.

B. Data Aquisition

All data used in this analysis were collected on 25th November, 2014, from rows of corn crop (*Zea mays everta*) at Mulyan Farms near Cowra, NSW, interspersed with weeds of the following varieties:

- Caltrop (*Tribulus terrestris*),
- Curly dock (*Rumex crispus*) or red dock (*rumex sanguineus*)
- Baryardgrass (*Echinochloa crusgalli*)
- Types of *Ipomoea* spp. and/or *Polymeria* spp.

A Resonon Pika II VNIR hyperspectral line scanning camera was mounted to the Ladybird robot, pointing towards the ground at about 52° from the horizontal (see Fig. 1). The camera produces hyperspectral images of 648 spatial by 244 spectral pixels (spectral resolution of 2nm from 390.9-887.4nm) at a frame rate of 133 fps and bit depth of 12. A Schneider Cinegon 6mm objective lens was used at an aperture setting of f/2.1, and manually focused with a checkerboard at average height of the corn. The lens provides a 43.5° field of view translating to a 2.8mm/pixel cross track spatial resolution. The along track resolution is limited by the robot's velocity, which at 1.2m/s equates to 9.0mm/pixel.

Two crop rows were scanned at a time (i.e. each scan

line crosses two crop rows). In addition, pairs of crop rows were covered twice in one scan, in the outward and return direction.

The sky was clear (no clouds moving in front of the sun), individual scans were short (less than 10 min), and all scans were made within two hours. Therefore illumination changes within scans were almost non-existent, while there was little variation from scan to scan.

An RGB camera and strobe combination mounted underneath the robot and facing downwards simultaneously gathered colour images. In addition, black cardboard was used to highlight a number of crop and weed examples that were separately photographed with a DSLR camera (see Fig. 7).

C. Manual Data Labelling

Both the outward and return traverses of two representative scans were sparsely labelled by hand (i.e. number of labelled pixels $\ll$ total number of pixels), yielding a total of four labelled sets. The label locations were chosen semi-randomly, while trying to avoid bias and background pixels were removed using the NDVI thresholding method described in Section III-D (see Fig. 3). The process yielded thousands of vegetation pixels from tens of distinct crop and weed plants.

Black cardboard backings with physical labels as shown in Fig. 7 helped to identify a few of the plants, while the majority of box selected regions were distinguished by physical appearance in HSI false colour or the separate RGB images. The extracted data was labelled as either crop or weed and care was taken to ensure class assignment is correct to avoid any cross-class contamination. Dataset details are summarised in Table I.

TABLE I
HAND LABELLED DATASET DETAILS

		No. of Sample Pixels	
Name	**Details**	**Crop**	**Weed**
DS1-1 HL	300m outward scan	65706	51255
DS1-2 HL	300m return scan	50457	30026
DS3-1 HL	80m outward scan	10175	10167
DS3-2 HL	80m return scan	11899	10755

D. Vegetation Separation

Background pixels, which include soil, dead vegetation and other non-vegetation were masked out prior to classification of vegetation type. Vegetation spectra are very distinct from non-vegetation, allowing simple discrimination techniques to separate them. Vegetation indices with careful waveband selection are particularly useful for this purpose. While there are several indices that could be employed, the NDVI is pervasive in the literature and has also previously produced good results in this particular application [25]. Vegetation exhibits high absorption at red wavelengths, while high reflectance is observed at near infrared (NIR) wavelengths. The NDVI builds on this fact to calculate a vegetation measure:

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

Red and NIR wavelengths were carefully chosen to ensure good separation. Average Red-NIR values from wavelength ranges 679-685nm and 740-744nm produced good results, and is in line with the literature [26, 27]. To generate a vegetation mask, a threshold was applied to the computed NDVI index values, where values above the threshold were considered to be vegetation.

Otsu's method [28] was employed to dynamically compute an optimum threshold that minimises intra-class variance. Fig. 3 illustrates the background separation process.

It was found that near boundaries between classes (e.g. the edge of a leaf against the background soil) spectral mixing can occur, whereby the intensity profile (hyperspectral colour) of a pixel is influenced by the spectra of both classes. To minimise this effect, binary erosion was applied to the vegetation mask. Two iterations with a square kernel size of three were determined to offer good classification performance without excessively reducing the vegetation mask.

E. Automatic Training Dataset Generation

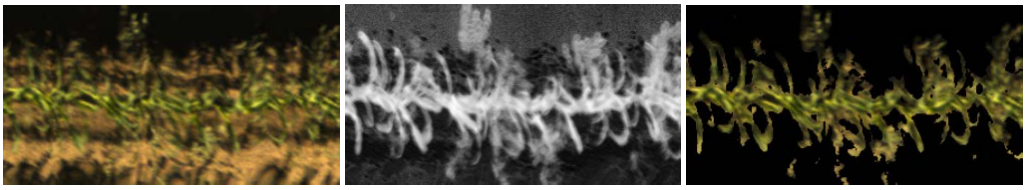
In order to collect crop and weed training data in a self-supervised manner, row centres need to be detected. Fig. 4 shows the vegetation mask of a typical section of a corn crop row. Rows are distinct and run in a consistent parallel direction. In order to obtain good representations of crop row centres, the Hough transform is used to extract two lines that are sufficiently far apart.

The next step leverages the consistency with which corn plants adhere to the detected row centre lines. To determine crop and weed training pixels, the vegetation mask is first broken up into distinct connected regions. Any of these regions that do not touch either row line are assumed to exclusively consist of weed pixels. Crop training data are collected from a fixed region around the row centre, where the probability of encountering crop pixels is highest. In this work, crop training samples were extracted from one pixel thick row centre lines. While these pixels are in reality likely to be a mix of mostly crop and a few weed spectra, for the purposes of training they were all assumed to be crop. The extraction rules are slightly different to ones introduced in [4], where left and right row borders were determined first and any blob having at least one pixel within the borders was considered to belong *inside the row*. The modified rules were determined to be more reliable for the dataset tested in this paper.

F. Spectral Preprocessing

In order to mitigate the effects of varying environmental illumination intensity at different times of day and between shade and direct sunlight, a simple normalisation method was used [11]:

$$N_i = \frac{S_i - S_{min}}{S_{mean} - S_{min}} \tag{2}$$



(a) Cropped Sample Image (False Colour) (b) NDVI Image (c) Resulting Masked Image (False Colour)

Fig. 3. The above illustrates the vegetation separation process. The images are close-ups of stacked hyperspectral scan lines, which are vertical, with crop rows running horizontally. The sample image (a) is converted to an NDVI intensity map (b). An image mask is generated using all NDVI values above a threshold to generate the filtered image in (c). Note that false colour representations are generated with RGB values at wavelengths 600, 530 and 475nm respectively.

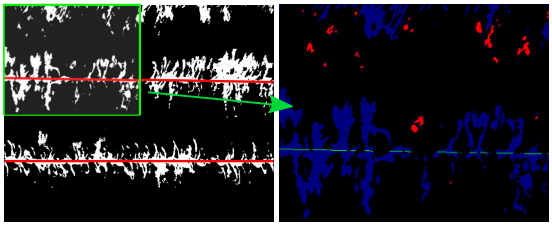


Fig. 4. Binary vegetation mask (see Section III-D) with predicted row centres (left) and cropped close up of detected training samples (right, see Section III-E). Red regions represent weed sample pixels, whereas green regions represent crop sample pixels. Blue regions are discarded vegetation pixels.

where S_i is the pixel value at band i, S_{mean} and S_{min} are the mean and minimum spectral intensity values respectively per pixel, and N_i is the normalised pixel intensity.

Savitzky-Golay smoothing was applied to the spectra [29]. A window length of 11 with polynomial order 2 was determined to be appropriate. Also, excessive noise due to lower quantum efficiency at the edges of the spectral range was mitigated by cropping the first and last 20 bands from all spectra.

After centring the data, PCA was used for feature extraction and dimensionality reduction, which assumes that information is maximised in the direction of greatest variance. PCA was performed on the training data, and then both training and test data were transformed based on the resulting reduced dimensions. The 20 most explanatory PCA dimensions were retained. The transformed data was whitened, as it positively affected the Support Vector Machine (SVM) classifier, with no effect either way on the LDA classifier.

G. Classification Methods

Both LDA and SVM based classifiers were implemented and compared using the scikit-learn package [23]. LDA was chosen as it previously provided competitive results for crop/weed discrimination [11] and fast processing, which is applicable particularly in the context of robotics, including real time applications such as targeted weed spraying [5]. SVM is a more sophisticated but well known classifier, which has been employed with HSI in the past [30]. For this paper, a Radial Basis Function (RBF) kernel was used with the SVM classifier.

Prior to classification, both training and test sets were balanced to prevent bias in the classifiers and metrics.

IV. RESULTS AND DISCUSSION

A. Supervised Classification

Supervised training and classification was initially performed to optimise the classification pipeline parameters and set a performance benchmark. Ten-fold cross validation was applied to each of the four datasets independently. Each whole dataset was also used for training and tested against each of the other three datasets (e.g. not including itself), yielding 12 training/testing dataset combinations. Results are shown in Fig. 5, and also in Table II (table does not include cross validation results for brevity).

These results confirm that good performance can be achieved when discriminating plant species with hyperspectral data. SVM achieved marginally better results than LDA.

B. Classification Using Automatically Collected Data

Automatic collection of training data was achieved using the method described in Section III-E, applied to windows of 1000 scan lines at a time. Datasets were auto-generated on outward and return traverses of the same two scans used for manual labelling (see Section III-C), again yielding four datasets, which were individually tested against each of the four hand labelled datasets, resulting in 16 training/testing dataset combinations. Table III details the four datasets.

To evaluate the accuracy of the automatic dataset generation algorithm, it was compared to the hand labelled datasets. To achieve this, matching pixels between corresponding hand labelled and auto-generated dataset pairs (for instance "DS1-1 HL" and "DS1-1 Auto") were extracted and their labels compared. The last two columns in Table III summarise the results, where the "HL Overlap" column gives the number of pixels shared by each auto-generated dataset and its corresponding hand labelled dataset. These accuracy numbers are very high, but it should be noted that this is not an exact measure of extraction performance, because it represents the intersection between two extraction methods, which is less likely to include challenging regions due to the human bias towards exemplars that are easier to define in the

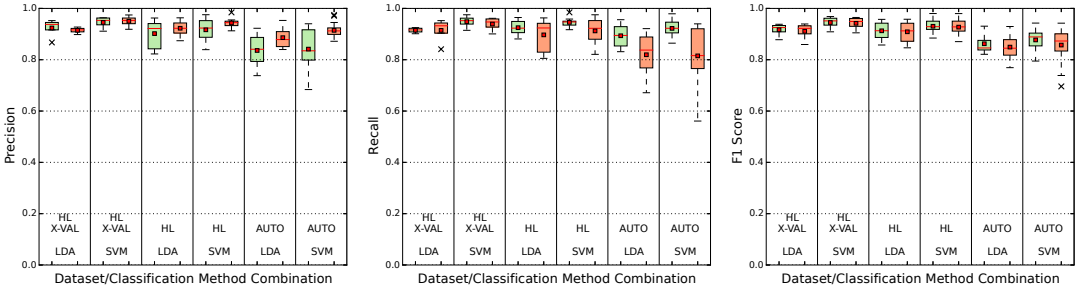


Fig. 5. Summarised results from both hand labelled and auto generated training sets. Green and red boxes represent results for crop and weed classes respectively. Values are summarised in Table II (not including 10-fold cross validation). Short hand column labels are as follows: X-VAL - 10-fold cross validation, HL - Hand labelled datasets, AUTO - Automatically generated datasets.

TABLE II

MANUAL AND AUTO TRAINED CLASSIFIER PERFORMANCE

Training Set Type[1]/ Classifier	Min	Median	Mean	Max
Precision				
HL LDA Crop	0.82	0.92	0.90	0.96
HL LDA Weed	0.87	0.92	0.92	0.96
HL SVM Crop	0.84	0.92	0.92	0.98
HL SVM Weed	0.91	0.94	0.94	0.98
AUTO LDA Crop	0.74	0.84	0.84	0.92
AUTO LDA Weed	0.84	0.88	0.89	0.95
AUTO SVM Crop	0.68	0.83	0.84	0.94
AUTO SVM Weed	0.87	0.91	0.91	0.97
Recall				
HL LDA Crop	0.88	0.92	0.92	0.96
HL LDA Weed	0.81	0.92	0.90	0.96
HL SVM Crop	0.92	0.95	0.95	0.98
HL SVM Weed	0.82	0.92	0.91	0.98
AUTO LDA Crop	0.83	0.89	0.89	0.96
AUTO LDA Weed	0.67	0.84	0.82	0.92
AUTO SVM Crop	0.86	0.92	0.92	0.98
AUTO SVM Weed	0.56	0.82	0.82	0.94
F1 Score				
HL LDA Crop	0.86	0.91	0.91	0.96
HL LDA Weed	0.85	0.91	0.91	0.96
HL SVM Crop	0.88	0.93	0.93	0.98
HL SVM Weed	0.87	0.93	0.93	0.98
AUTO LDA Crop	0.82	0.85	0.86	0.93
AUTO LDA Weed	0.77	0.84	0.85	0.93
AUTO SVM Crop	0.79	0.89	0.88	0.94
AUTO SVM Weed	0.70	0.87	0.86	0.94

[1] HL - Hand labelled datasets, AUTO - Auto-generated datasets.

TABLE III

AUTO-GENERATED DATASET DETAILS

		Compared to HL Data	
Name	**No. of Pixels[1]**	**HL Overlap[2]**	**Ext. Accuracy[3]**
DS1-1 Auto	32515 / 3153	3447 / 4045	98.52 / 100.00%
DS1-2 Auto	46096 / 3566	2875 / 2961	100.00 / 100.00%
DS3-1 Auto	8348 / 8110	641 / 2791	100.00 / 98.35%
DS3-2 Auto	9680 / 4983	732 / 4668	100.00 / 100.00%

[1] Number of crop/weed pixels per dataset (after erosion).
[2] Number of crop/weed pixels shared with the corresponding hand labelled dataset. Note that these numbers were calculated without erosion to maximise the overlap between between auto and hand labelled datasets.
[3] Crop/weed percentage of shared pixels that have been correctly identified with the automatic method.

manual labelling process. Nevertheless, it can be viewed as an indication that the algorithm was functioning correctly.

Classification results are shown on Table II and summarised in the "AUTO DS" columns in Fig. 5. SVM and LDA performed similarly. While median and mean scores were marginally higher for the SVM classifier, LDA exhibited slightly more consistent results when trained with the auto-generated datasets. In most cases F1 scores remained above 0.75 for both classes. While it is evident that these per-pixel scores do not match the level of performance obtained with hand labelled training sets, this is to be expected, due to the higher levels of contamination from the opposite class in the training sets.

A large dip can be seen in the box plot for the self-supervised SVM weed recall results (and to a lesser extent the corresponding crop precision scores). This was caused by two of the 16 training/test set combinations (0.56 and 0.61 recall), and future work will seek to diagnose the reasons for the poor performance from those particular training/test set pairs. The other combinations performed better, as demonstrated by a median and mean recall of 0.82.

The results appear to have a bias towards high recall/low precision for crop pixels and the reverse for weed, particularly in cases of lower performance. This indicates a bias in the classifiers towards labelling points as crop; in other words, more weed pixels are falsely classified as crop than crop pixels as weed. Dealing with this may depend on the final application. For example, a lower chance of falsely eradicating crop plants may be desirable in targeted herbicide applications. In this case, LDA prior probabilities (which are set to 0.5 for both classes in this paper) could be manually adjusted to ensure high crop recall.

Fig. 6 outlines some typical classification results, comparing a hand labelled training dataset to an auto-generated one (DS1-1 HL and DS1-1 Auto). It can be seen that the majority of falsely labelled pixels occurred around the edges of crop plants and as speckle within weeds. Both issues may be overcome by incorporating a form of spatial smoothing or segmentation. Examples 2 and 3 demonstrate the separation

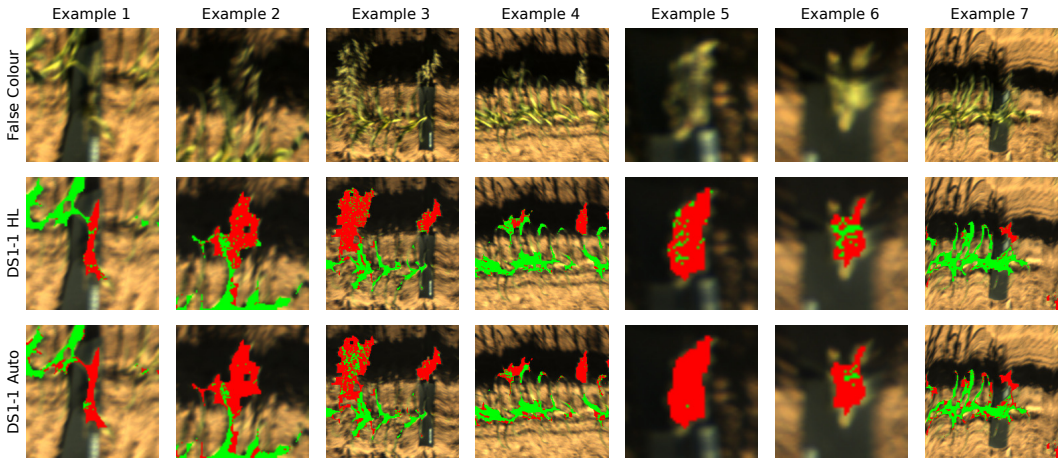


Fig. 6. Sample classification results showing false colour RGB images (first row), manually supervised classification (second row) and self-supervised classification (third row). Green and red indicates crops and weeds respectively. Hand labelled and auto-generated training sets are DS1-1 HL and DS1-1 Auto respectively. Black cardboard backings in some of the images explicitly label plants as crop or weed (as identifiable in corresponding DSLR photographs, see Fig. 7).

(a) Example 1 - Weed (b) Example 3 - Crop (c) Example 5 - Weed (d) Example 6 - Weed (e) Example 7 - Crop

Fig. 7. DSLR photos of five of the seven examples shown in Fig. 6. Weed species are as follows: (a) curled or red dock, (c) caltrop and (d) *Ipomoea* sp. or *Polymeria* sp.

between overlapping plants of different classes, a particular strength of HSI based classification. Some of the examples also demonstrate that due to the limited spatial resolution and morphological erosion, very small plant pixels will not be classified at all. This can be improved by decreasing the scanning speed, or increasing the resolution by scanning closer to the ground or using a higher resolution line scanning camera.

In general, both classification methods yielded relatively high performance when performing self-supervised training, in some cases rivalling scores from hand labelled datasets, yet the automated method requires no human labelling effort or intervention. Additionally, self-supervision is particularly useful when conditions change, as the classifier can be continually updated with new training data, and it is expected that the performance in this case will remain constant. On the other hand, a manually generated, static training set performs well in matching conditions, but would degrade as these vary considerably [18].

Finally, as an indication of speed, training and classification was timed and averaged over ten iterations with DS1-1 Auto training and DS1-1 HL test sets. LDA training only took 0.0079s on 6306 total pixels, while SVM took 0.3665s. Testing on 102510 pixels was timed at 0.0036s and 4.2455s with LDA and SVM respectively. This suggests that using a simple classifier like LDA might be preferable in situations where processing time is critical.

V. FUTURE CONSIDERATIONS

This paper demonstrates how crop/weed discrimination can be achieved without labour intensive manual labelling, with a framework that is well equipped to handle appearance variability. Further testing is required to confirm and demonstrate how robust the system is to changing conditions compared to static training data.

For weed management, a complete system would use classification results to perform corrective actions, such as spraying or removal of identified weeds, which requires a number of additional steps. Firstly, the hyperspectral data and classification results should be geographically registered. This will allow real world location of identified weeds and subsequent treatment. Alternatively, if real time treatment is desired, weed locations must be translated to treatment locations relative to the robot in a fraction of a second as the

vehicle passes. Spatial smoothing or segmentation will also be necessary to reduce the speckled nature of the result, and help convert per-pixel classification results to identification of whole plants, as some treatment options will require centre locations of plants.

VI. CONCLUSION

This paper presented a self-supervised method for discriminating weeds in crop fields, without the need for manual labelling. The technique, which is based on prior work using RGB images [4], gathers training data automatically to form a self-supervised classification framework that is resistant to variation. This would allow the classifier to adapt to changing class appearance without manually generating new datasets. Good overall weed/crop discrimination performance was achieved with all of the extracted datasets (mean F1 scores of 0.85 or better), approaching performance of hand labelled training data (mean F1 scores of 0.91 or better), yet not requiring any manual labelling. Future work will seek to improve classification performance by investigating techniques that refine the training data, as well as the use of more sophisticated classification methods. The inclusion of spatial information, real-time performance, and online loop closure for weed control are all opportunities for future work. The framework's ability to deal with changing conditions should also be confirmed in datasets that span greater variability.

REFERENCES

[1] Y. S. Tey and M. Brindal, "Factors influencing the adoption of precision agricultural technologies: a review for policy implications," *Precision agriculture*, vol. 13, no. 6, pp. 713–730, 2012.

[2] R. B. Brown and S. D. Noble, "Site-specific weed management: sensing requirements-what do we need to see?" *Weed Science*, vol. 53, no. 2, pp. 252–258, 2005.

[3] D. Slaughter, "The Biological Engineer: Sensing the Difference Between Crops and Weeds," in *Automation: The Future of Weed Control in Cropping Systems*. Springer, 2014, pp. 71–95.

[4] F. M. De Rainville, A. Durand, F. A. Fortin, K. Tanguy, X. Maldague, B. Panneton, and M. J. Simard, "Bayesian classification and unsupervised learning for isolating weeds in row crops," *Pattern Analysis and Applications*, vol. 17, no. 2, pp. 1–14, 2012.

[5] J. P. Underwood, M. Calleija, Z. Taylor, C. Hung, J. Nieto, R. Fitch, and S. Sukkarieh, "Real-time target detection and steerable spray for vegetable crops," in *in proceedings, Workshop on Robotics in Agriculture at International Conference on Robotics and Automation (ICRA)*, 2015.

[6] S. Hong, L. Minzan, and Q. Zhang, "Detection system of smart sprayers: Status, challenges, and perspectives," *International Journal of Agricultural and Biological Engineering*, vol. 5, no. 3, pp. 10–23, 2012.

[7] C. Mccarthy, S. Rees, and C. Baillie, "Machine vision-based weed spot sprying: A review and where next for sugarcane?" in *32nd Annual Conference of the Australian Society of Sugar Cane Technologists*, 2010, p. 7.

[8] H. Okamoto, Y. Suzuki, and N. Noguchi, "Field Applications of Automated Weed Control: Asia," in *Automation: The Future of Weed Control in Cropping Systems*. Springer, 2014, pp. 189–200.

[9] F. Feyaerts and L. Van Gool, "Multi-spectral vision system for weed detection," *Pattern Recognition Letters*, vol. 22, no. 6-7, pp. 667–674, may 2001.

[10] H. Okamoto, T. Murata, T. Kataoka, and S. I. Hata, "Plant classification for weed detection using hyperspectral imaging with wavelet analysis," *Weed Biology and Management*, vol. 7, no. 1, pp. 31–37, 2007.

[11] Y. Suzuki, H. Okamoto, and T. Kataoka, "Image Segmentation between Crop and Weed using Hyperspectral Imaging for Weed Detection in Soybean Field," *Environment Control in Biology*, vol. 46, no. 3, pp. 163–173, 2008.

[12] P. Eddy, A. Smith, B. Hill, D. Peddle, C. Coburn, ., and R. Blackshaw, "Hybrid Segmentation Artificial Neural Network Classification of High Resolution Hyperspectral Imagery for Site-Specific Herbicide Management in Agriculture," *Photogrammetric Engineering and Remote Sensing*, vol. 74, no. 10, pp. 1249–1257, 2008.

[13] I. Herrmann, U. Shapira, S. Kinast, a. Karnieli, and D. J. Bonfil, "Ground-level hyperspectral imagery for detecting weeds in wheat fields," *Precision Agriculture*, vol. 14, no. 6, pp. 637–659, 2013.

[14] A. J. Ishak, A. Hussain, and M. M. Mustafa, "Weed image classification using Gabor wavelet and gradient field distribution," *Computers and Electronics in Agriculture*, vol. 66, no. 1, pp. 53–61, 2009.

[15] M. Persson and B. Åstrand, "Classification of crops and weeds extracted by active shape models," *Biosystems Engineering*, vol. 100, no. 4, pp. 484–497, 2008.

[16] H. T. Søgaard, "Weed classification by active shape models," *Biosystems engineering*, vol. 91, no. 3, pp. 271–281, 2005.

[17] S. Haug, A. Michaels, P. Biber, and J. Ostermann, "Plant classification system for crop /weed discrimination without segmentation," in *2014 IEEE Winter Conference on Applications of Computer Vision, WACV 2014*. IEEE, 2014, pp. 1142–1149.

[18] Y. Zhang, D. C. Slaughter, and E. S. Staab, "Robust hyperspectral vision-based classification for multi-season weed mapping," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 69, pp. 65–73, apr 2012.

[19] M. Rosenblatt and Others, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, 1956.

[20] E. Parzen, "On estimation of a probability density function and mode," *The annals of mathematical statistics*, pp. 1065–1076, 1962.

[21] J. A. Dille, "Plant Morphology and the Critical Period of Weed Control," in *Automation: The Future of Weed Control in Cropping Systems*. Springer, 2014, pp. 51–69.

[22] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin, "Plant species identification using digital morphometrics: A review," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7562–7573, 2012.

[23] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine Learning in Python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[24] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "Scikit-image: image processing in Python." PeerJ PrePrints, Tech. Rep., 2014.

[25] C. McCarthy, N. Hancock, and S. Raine, "Applied machine vision of plants: A review with implications for field deployment in automated farming operations," *Intelligent Service Robotics*, vol. 3, no. 4, pp. 209–217, 2010.

[26] D. A. Sims and J. A. Gamon, "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages," *Remote sensing of environment*, vol. 81, no. 2, pp. 337–354, 2002.

[27] S. Sankaran, J. M. Maja, S. Buchanon, and R. Ehsani, "Huanglongbing (citrus greening) detection using visible, near infrared and thermal imaging techniques," *Sensors*, vol. 13, no. 2, pp. 2117–2130, 2013.

[28] N. Otsu, "A threshold selection method from Gray-level," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-9, no. 285-296, pp. 62–66, 1979.

[29] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures." *Analytical chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.

[30] R. J. Murphy, S. T. Monteiro, and S. Schneider, "Evaluating classification techniques for mapping vertical geology using field-based hyperspectral sensors," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 8, pp. 3066–3080, 2012.

Chapter 6

Fruit Maturity Estimation

In horticulture, estimating the optimal time for harvest is important to maximise revenue and minimise waste. Harvested fruit need to have optimal maturity to allow for storage and transport while maximising eating quality. By tracking maturity of fruit on tree as they develop, growers use their expertise and prior knowledge to try and predict a suitable harvest date. While maturity can be approximated by visual inspection, more accurate methods have been developed. Dry matter (DM) content has been shown to be a reliable indicator of maturity, as it correlates closely with a fruit's carbohydrate content. The most direct way to measure DM is by oven-drying which requires sacrificing the fruit sample. However, more convenient and non-destructive methods using spectroscopy have been developed instead. One approach currently in use for some fruit involves holding a hand-held spectrometer with partial transmittance optics in contact with sample fruit to be measured. Several such measurements can be made per block and tracked over time via weekly or bi-weekly measurements.

While using the hand-held spectrometer method is accurate on a per fruit basis, it is very labour intensive if a high level of data density and resolution is required, e.g. measuring fruit in every single tree of an orchard block. This would be useful if an orchard block exhibits significant maturity variability. For example, if specific regions receive more sun light they will mature more quickly. To maximise yield in this case,

it would be necessary to harvest these areas earlier to avoid overripe fruit or fruit that drop from the tree prior to picking. This is also referred to as selective harvesting. Alternatively, knowing the fruit maturity profile across a block instead of a simple average allows a more accurate estimate of a single optimal harvest date.

The need for more granular maturity estimates presents an opportunity for hyperspectral imaging, which motivates the work in this chapter. The study proposes the use of a hyperspectral sensor and navigation system aboard a mobile ground vehicle to efficiently predict and map maturity on an orchard scale. An entire orchard block was scanned, producing hyperspectral data for hundreds of trees. A novel classification, regression and mapping pipeline is introduced to process the data. The system extracts pixels belonging to mango fruit before providing a DM estimate for each pixel. Using the navigation system, the data are then georeferenced, producing maps that provide DM predictions that are averaged per tree.

The system was presented with challenging conditions as the hyperspectral camera was set up to scan trees in a sideways push-broom configuration using natural daylight. Lighting variations were extensive due to the complicated canopy geometry, introducing shadows and scattered light from surrounding surfaces. Despite this, experimental validation of the system supports its ability to predict DM and its repeatability over two days.

Computers and Electronics in Agriculture 155 (2018) 298–313

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform

Alexander Wendel[a,*], James Underwood[a], Kerry Walsh[b]

[a] Australian Centre for Field Robotics, The Rose Street Building J04, The University of Sydney, NSW 2006, Australia
[b] Central Queensland University, Building 361, 1.22, Rockhampton, QLD 4700, Australia

ARTICLE INFO

Keywords:
Hyperspectral
Mango
Dry matter
Unmanned ground vehicle
Maturity

ABSTRACT

Monitoring the maturity of fruit in commercial orchards can help growers optimise the time of harvest. Dry matter content (DM) of fruit is used as an indicator of mango maturity, measured in-field with a hand-held spectrometer. This approach is labour intensive, limiting the extent to which DM variability can be measured across an orchard block, which would enable selective harvesting. This paper proposes an alternative approach that utilises a hyperspectral camera, LIDAR sensor and navigation system mounted to a ground vehicle to predict fruit DM individually for hundreds of trees in a mango orchard block. First, the challenges faced due to tree geometry and shadows in mango orchards are addressed. Then the ability to predict DM at a distance using hyperspectral imaging (411.3–867.0 nm) was demonstrated. Two regression methods, partial least squares (PLS) and a convolutional neural network (CNN) were compared and tested against DM results from a hand-held NIR spectrometer using harvested ($n = 468$, $\sigma = 2.32\%w/w$) and on-tree fruit ($n = 662$, $\sigma = 1.79\%w/w$). The CNN achieved a cross validation $R^2_{CV} = 0.64$ and $RMSE_{CV} = 1.08\%w/w$ in fruit on tree, while PLS achieved $R^2_{CV} = 0.58$ and $RMSE_{CV} = 1.17\%w/w$. In order to discriminate mango and non-mango pixels, PLS discriminant analysis (PLS-DA) and a CNN were also compared, where both methods achieved good classification performance with a mean $F1 > 0.97$. Having established mango classification and DM prediction performance, hyperspectral data were processed for a full orchard block and projected to world coordinates using AGV position and orientation as provided by the navigation system. Trees were segmented using corresponding LIDAR data, which allowed association of projected DM predictions to individual trees. Repeated scans of the orchard block over two days allowed a measure of repeatability, which was achieved with an $RMSE < 0.29\%w/w$. The results provide strong evidence that predicting maturity at a distance for all trees in an orchard is feasible using a hyperspectral camera, which will be an important management tool for growers to optimise harvest timing and yield.

1. Introduction

In commercial orchards, monitoring the maturity of fruit as they develop is important because it influences the optimal time to harvest. Growers aim to harvest fruit at "harvest maturity", which allows the fruit to ripen off the tree. The best tasting fruit would be ripened on the tree, as they would have the longest time to accumulate sugars and starch, but would have no allowance for transport and shelf life. Because of this, growers try to optimise the balance between on-tree ripening and transport/shelf-life, known as "harvest maturity", which includes achieving physiological maturity and some on-tree ripening. While fruit maturity can be approximated by visual inspection (Subedi and Walsh, 2011), much progress has been made to more accurately measure it. For some types of fruit, such as mangoes and bananas, dry matter (DM) content is considered an important indicator for fruit maturity, correlating with a fruit's carbohydrate content (starch and sugars) (Walsh and Subedi, 2014). For example, in Walsh and Subedi (2014) DM was measured for mangoes over eight weeks from the stone hardening stage leading up to harvest, and ranged from just over 9% w/w to almost 22% w/w. The Australian Mango Industry Association (AMIA) recommends a harvest DM of ranging from 13% w/w (R2E2 cultivar) to 15% w/w (Kensington Pride, Calypso and Honey Gold cultivars) (AMIA, 2015). DM can be determined destructively, but recent advances have yielded spectroscopic methods and portable hand-held instruments that allow accurately estimating fruit dry matter content on intact fruit, even while still on the tree.

DM can be measured in the field using a hand-held spectrometer with partial transmittance optics by holding it in contact with

* Corresponding author.
E-mail addresses: a.wendel@acfr.usyd.edu.au (A. Wendel), j.underwood@acfr.usyd.edu.au (J. Underwood), k.walsh@cqu.edu.au (K. Walsh).

https://doi.org/10.1016/j.compag.2018.10.021
Received 20 June 2018; Received in revised form 26 August 2018; Accepted 13 October 2018
Available online 28 October 2018

individual fruit. Several such measurements can be averaged per block and tracked over time via weekly or bi-weekly measurements. A number of past studies have shown the potential of using near infra-red (NIR) spectroscopy for the prediction of mango fruit DM (Guthrie and Walsh, 1997; Saranwong et al., 2004; Walsh et al., 2004; Subedi et al., 2007). Subedi and Walsh (2011) examined the applicability of NIR spectroscopy using partial transmittance optics, as proposed by Greensill and Walsh (2000), for the estimation of DM and total soluble solids (TSS) of mesocarp tissue of both banana and mango. Using a wavelength range of 500–1050 nm and partial least squares (PLS) regression, mango DM could be successfully predicted across all stages of maturity with a cross validation correlation coefficient $R^2_{CV} > 0.75$ and root mean square error $RMSE_{CV} < 0.70\%w/w$. Walsh and Subedi (2014) and Subedi et al. (2013) used hand-held spectrometers with PLS regression to estimate mango fruit DM in the field, and confirmed its applicability in making harvesting decisions. In addition, machine vision approaches have been proposed for the prediction of fruit properties in production lines. For instance, Nandi et al. (2014) used an RGB camera and support vector machine (SVM) to classify harvested mangoes during packing into four maturity groups with an accuracy of >90% to allow the fruit to be sorted for transport to different locations. There are also many examples of dry matter or maturity estimation in other fruit, such as banana (Rajkumar et al., 2012; Surya Prabha and Satheesh Kumar, 2015), persimmon (Mohammadi et al., 2015), and tomato (Acharya et al., 2017). In another related study, Rungpichayapichet et al. (2017) utilised imaging spectroscopy (also known as hyperspectral imaging) in a lab setting to map firmness, titratable acidity (TA) and TSS across a whole fruit.

The aforementioned methods allow accurate estimation of fruit quality, including DM, on a per fruit basis, where environmental conditions such as illumination can be controlled. However, these approaches become logistically intractable when a higher level of granularity is required in the data, such as averages of fruit properties per-tree across a whole orchard block. Non-destructively predicting pre-harvest fruit properties on an orchard scale, which involves thousands of trees and hundreds of thousands or millions of fruit requires a prohibitive amount of labour using current best practice methods. Nevertheless, such detail would be useful to growers, as fruit maturity can vary significantly within orchard blocks, due to external factors such as available sunlight, soil quality and water availability. Additionally, mangoes can have multiple flowerings per year, leading to significant variability in maturation times from tree to tree and even within single tree canopies (Subedi et al., 2013). Growers could use knowledge of this variation in maturity to optimise harvest timing or to selectively harvest specific areas within orchard blocks. This would increase yield by reducing harvested fruit at suboptimal maturity and minimising mature fruit that drop from trees prior to harvest.

There is a substantial body of research examining the use of satellite and aerial remote sensing in predicting vegetation properties, including for vegetation cover estimation (Zhang et al., 2013), biomass estimation (Marshall and Thenkabail, 2015), vegetation/crop classification (Oldeland et al., 2010; Xue et al., 2017), disease mapping (MacDonald et al., 2016) and nutrient/chlorophyll concentrations (Sims et al., 2013; Rao et al., 2008; Moharana and Dutta, 2016; Pullanagari et al., 2016). In fruit orchards, such remote sensing techniques may be useful for determining properties that are discernible at the canopy level, such as volume, mineral nutrition content, yield, diseases/pests and stress (Usha and Singh, 2013; Robson et al., 2017; Salgadoe et al., 2018), however, they do not currently provide the spatial resolution needed to directly predict fruit DM. Low altitude unmanned aerial vehicle (UAV) platforms can provide the resolution needed, but the geometry of many fruit crops is such that they must be viewed at a shallow angle with respect to the ground. As a result, to be ideally located for viewing individual fruit on trees, a sensor should generally be situated at tree level, moving in-between orchard rows. Navigating between tree rows has been demonstrated with multi-rotor helicopter UAVs (Verbeke et al., 2014; Stefas et al., 2016; Hulens et al., 2017), however, while these types of platforms are convenient in many ways, their limited battery life and payload can be prohibitive in some situations.

Recently, several studies have explored the use of mobile ground platforms to acquire data of large areas for the purpose of phenotyping and yield estimation (Deery et al., 2014; Andrade-Sanchez et al., 2014; Busemeyer et al., 2013; Comar et al., 2012; Montes et al., 2011; Underwood et al., 2017; Bargoti and Underwood, 2017; Wang et al., 2013; Payne et al., 2014; Qureshi et al., 2017; Aquino et al., 2018). These platforms allow the use of multiple sensors to acquire high resolution data over large regions, which can be used to measure various plant properties. Bargoti and Underwood (2017) developed a convolutional neural network (CNN) based image segmentation approach, which used data from an unmanned ground vehicle (UGV) to detect individual apples on trees and estimate yield per row. Stein et al. (2016) developed an approach that tracked CNN detections from multiple views of mango trees to predict yield, and compared it to single view (one side of a tree only) and dual view approaches (two opposing sides of a tree). Linker (2018) performed yield estimation in apple trees using a pipeline that included Upright SpeededUp Robust Features (U-SURF) descriptors and a SVM classifier. Payne et al. (2014) developed a pipeline using a number of thresholds that allowed detection of mango fruit and subsequent yield count. Several other studies demonstrated fruit detection and yield estimation (e.g. Dorj et al., 2017; Wang et al., 2013; Gongal et al., 2016). In addition, estimating fruit size (Wang et al., 2017), and assessing water status using thermal imaging (Gutiérrez et al., 2018) and NIR spectroscopy (Diago et al., 2018) have also been demonstrated in fruit orchards or vineyards. However, the authors are not aware of any published studies that use imaging sensors from a platform that moves along the inter-row for the estimation of fruit maturity.

This study examines the use of a UGV with a hyperspectral imaging sensor and navigation system to efficiently predict and map DM on an orchard scale. The system is coupled to our prior work on light detection and ranging (LIDAR) tree segmentation to allow association of the estimates per tree. While hyperspectral imaging, which captures spectra in hundreds of narrow bands for each pixel in a given image frame, has been extensively studied in the remote sensing community, there are only a few examples of the use of hyperspectral imaging on ground based mobile platforms for agricultural applications, such as phenotyping (Deery et al., 2014; Klose et al., 2010; Underwood et al., 2017). To the best of the authors' knowledge, however, there are no published studies yet of hyperspectral imaging being used in fruit orchards to predict fruit properties, from mobile ground platforms or otherwise. In this work hyperspectral images were captured using natural light, which presents some challenges due to lighting variations as will be detailed later. Using artificial illumination (e.g. halogen lights at night) may be a potential option, but introduces other difficulties such as increased power draw and timing logistics. The system presented in this work enables the generation of orchard-scale fruit maturity (dry matter) maps to facilitate granular farm management and selective harvesting. Specifically, this paper makes the following contributions:

- An efficient system for orchard-scale maturity (dry matter) mapping with tree and fruit-level resolution, using mobile ground-based hyperspectral sensing, including;
- A novel, unified approach to detect fruit in hyperspectral imagery and estimate its maturity (dry matter) simultaneously.
- A system for associating and averaging dry matter estimates per tree, to generate spatial orchard-scale dry matter maps.
- A solution to illumination compensation derived from prior work (Wendel and Underwood, 2017b; Drew and Finlayson, 2007) and an analysis of the challenges faced when acquiring hyperspectral data in fruit orchards due to variations in lighting resulting from the complex acquisition and tree canopy geometry.

2. Materials and methods

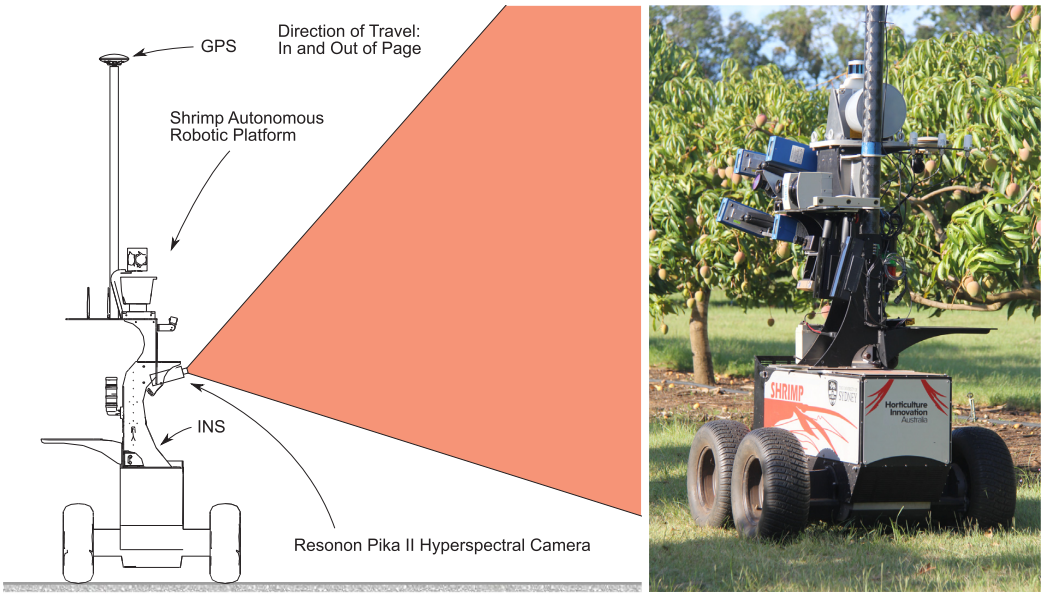
This section details the acquisition of the hyperspectral data and its processing in order to achieve DM estimation on an orchard scale and the experimental protocol for validation. All data were acquired at an orchard in Bundaberg, Queensland, Australia, using sensors mounted to a UGV. Orchard-scale DM estimation requires classifying each hyperspectral image pixel (i.e. individual spectra) as being mango or non-mango. Each classified mango pixel spectrum was then regressed to a DM value. Classification and regression models were trained and validated using manually labelled ground truth data. Finally, the resultant DM estimates were summarised in spatial maps of the average DM per tree. Each step of this process (classification, regression and mapping), was evaluated experimentally to demonstrate the effectiveness of the proposed method.

2.1. Data acquisition

Data were collected using a UGV called "Shrimp", developed at the Australian Centre for Field Robotics (ACFR), which is equipped with a Resonon Pika II visible to near infrared (VNIR) line scanning hyperspectral camera, a Velodyne HDL64E 3D LIDAR and a Novatel SPAN-CPT real time kinematic, global positioning and inertial navigation system (RTK/GPS/INS). as indicated in Fig. 1(a). The Resonon Pika II camera was mounted to the platform in a push broom configuration such that the scan line is vertical, and pitched upwards slightly to allow scanning of upright objects such as orchard trees (Fig. 1(a)). The camera produces hyperspectral images of 648 spatial by 244 spectral pixels (spectral resolution of 2.1 nm from 390.9 to 887.4 nm) at a rate of 133 frames per second and native bit depth of 12. A Schneider Cinegon 6 mm objective lens was manually focused with a checker board at the typical distance to the scene. The Velodyne HDL64E 3D LIDAR was configured to spin at 10 Hz with a data rate of 1.3 million points per second. Contrary to typical practice, the Velodyne was mounted "sideways", as pictured in Fig. 1(b), to capture the full height of the tree canopies. The Novatel SPAN system combines an OEM3 GPS receiver with an IMU-CPT inertial measurement unit (IMU), to provide six degree-of-freedom location and orientation of the platform at 50 Hz.

The data were collected from three cultivar B74 (Calypso) mango orchard blocks ("A", "B" and "C", see Fig. 2(a)) at Simpson Farms in Bundaberg,Queensland, Australia on 6th and 7th December 2017. Block A comprises 10 rows ranging from 60 m to 260 m in length, with a rectangular cut-out where two sheds are located, containing a total of 494 trees. Block B is comprised of five rows ranging from aprox. 25 m to 210 m in length, and contains 121 trees. Block C is comprised of four rows ranging from 175 m to 335 m in length, and contains 266 trees. Rows in block A run approximately south to north, while they run approximately west to east in blocks B and C. All trees, which were typically 3–4 m in height, were scanned from both sides, by traversing every row in both directions while continuously recording the sensor data. In addition, to allow repeatability testing, all trees in block A were scanned twice over two days. Several illumination reference panels (QPcard 102, MFR #GQP102) were placed at the beginning of as many rows as logistically possible. The panels were mounted to tripods as shown in Fig. 2(b) and placed both in direct sunlight and shadowed regions underneath trees to measure as wide a range of reference illumination conditions as possible. Scans in block A for both sides of the trees were timed to have good solar illumination behind the hyperspectral camera given in-field logistical constraints. This meant that the west side was scanned in the afternoon, and the east side in the morning, while data acquisition at mid-day, when the sun is at its zenith, was avoided whenever possible. Nevertheless, due to logistical constraints one repeated scan of the west side of block A had to be started just before midday, even though the west side of trees would optimally be scanned after midday, providing an opportunity to examine how this affected the results. Because blocks B and C run west to east, it was not possible to locate the camera optimally with respect to the sun at any time of the day, and therefore acquisition was timed to ensure that overall illumination was adequate (i.e. not too early or late in the day). See Table 1 for a detailed acquisition summary. Dark current measurements for the hyperspectral camera were recorded regularly (before or after acquisition runs), by completely blocking all light to the sensor and averaging the data over approximately 10 s.

A total of 78 fruit were selected to span a wide range of DM values



(a) Sensor Configuration (b) Shrimp in Orchard

Fig. 1. The Shrimp robotic platform and sensor configuration.

(a) Map of plots (b) Reference panels

Fig. 2. Google Earth map of the three blocks A, B and C from which data were acquired (a) and QPcard 102 reference panels used for illumination compensation (b).

Table 1
Acquisition summary.

Block	Date	Tree sides	Start time	Finish time	Conditions
A	6/12/2017	East	9:52:02	11:26:11	Clear
A	6/12/2017	West	13:34:05	14:30:25	Some very light clouds
A	7/12/2017	East	8:35:13	9:29:27	Partly cloudy
A	7/12/2017	West	10:09:09	11:21:39	Partly cloudy
B	6/12/2017	North & South	11:56:37	12:31:16	Clear
C	7/12/2017	North & South	9:48:43	12:52:32	Partly cloudy

($\mu = 12.98\%w/w$, $\sigma = 2.39\%w/w$) as measured with a hand-held Felix F-750 produce quality meter that was calibrated prior to use (validation $R_2 = 0.95$, $RMSE = 0.56\%w/w$). This instrument employs a Zeiss MMS1 NIR enhanced spectrometer, which has a pixel resolution of approx. 3.3 nm, optical resolution of 10 nm, wavelength range 400–1100 nm and an internal halogen lamp, referencing for background illumination and lamp illumination on every sample. The measurements were taken at the equator of the fruit on two opposing sides. The fruit were picked and placed in trays such that the measured location faced the camera with adjacent illumination reference panels, as shown in Fig. 3(a). The trays were then scanned by Shrimp at different times throughout one day. The fruit were also turned after a period of time to allow scanning of the opposing measured area. This provided an additional "fruit-in-tray" ground truth dataset, with controlled fruit orientation and illumination conditions.

Additionally, individual fruit from two rows in A and one row each in B and C were tagged and measured with the hand-held produce quality meter ($\mu = 12.94\%w/w$, $\sigma = 1.79\%w/w$). The measurements were taken at the equator on the fruit region facing out from the tree. Up to three fruit were selected per tree-side, with high, medium and low DM values, and coloured tags were unambiguously placed on the stem above the fruit, to allow field measurements to be associated to the hyperspectral data, as shown in Fig. 3(b). Care was taken to place tags

(a) Fruit-in-tray (b) Fruit-on-tree

Fig. 3. Photos of "fruit-in-tray" with calibration panels (a) and example "fruit-on-tree" with pink, blue and orange tags (b). The photo in (a) was taken earlier in the day, and the shadow across the fruit was not present later during acquisition when the sun was in a more favourable position. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Summary of data used for training and validation.

Dataset Name	No. of Fruit	No. of Groups	No. of Spectra[c]	DM% w/w μ/σ/min/max[d]	Purpose
Fruit-in-tray	78	468[a]	22932	12.90/2.32/ 9.20/21.57	Regression cross validation
Fruit-on-tree	662 (block A/B/C split: 420/93/149)	662	33018	12.94/1.79/ 8.98/19.95	Regression cross validation Regression training for mapping Positive data for classification cross validation Positive data for classification in mapping
Non-mango	204 (all from block A)	204[b]	33048	N/A	Negative data for classification cross validation Negative data for classification in mapping

[a] Fruit-in-tray were scanned twice on one side, and four times on the other.
[b] Negative training data were grouped by blocks of adjacent pixels.
[c] Each spectrum is a single sample for training and cross-validation.
[d] Calculated on groups.

as far as possible above the fruit to minimise their effect on the colour of each mango. This produced a "fruit-on-tree" ground truth dataset. While there are examples of mangoes with DM of near 20% w/w, the average for both "fruit-in-tray" and "fruit-on-tree" is closer to 13% w/w as the blocks were not mature enough for harvesting yet.

2.2. Labelling

Manual labelling was performed on true colour hyperspectral image data pertaining to individual fruit, non-fruit examples, and calibration panels, as summarised in Table 2. Hyperspectral line scans were appended in time with no motion compensation or orthorectification applied and converted to RGB true colour images using wavelengths at 605, 555 and 445 nm (see Fig. 4(a)). The RGB conversion was performed to aid in manual labelling only, and full spectra were preserved and used downstream. Individual fruit were selected with tight bounding boxes. Then a representative region (patches of size 4×4–13×13) on each fruit was selected as the best pixels to be used for regression. These patches were chosen to be both well illuminated to maximise signal to noise ratio (SNR) and closest to the fruit cheek area measured by the hand held produce quality meter (see Fig. 4(b)). A similar number of rectangular patches of non-mango pixels (e.g. grass, tree branches, leaves…) were also labelled for use as negative training samples for classification. For illumination compensation, pixels on the calibration panel and temporally adjacent (within 5 s) were also manually labelled.

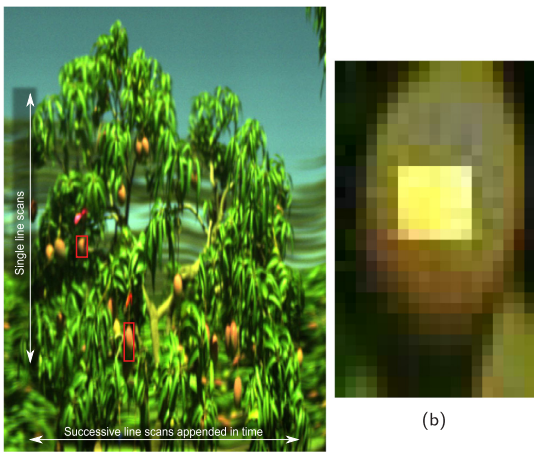


Fig. 4. Example true colour RGB image with fruit labels (a) and close-up fruit with patch label (b). The images were created by appending consecutive line scans from the Resonon Pika II hyperspectral camera and selecting three appropriate RGB wavelengths. As the camera is oriented such that each scan line is vertical, each column in these images represents pixels from a single scan line. Full spectrum data between 390.9 and 887.4 nm is available for each pixel (not shown).

Labelled fruit hyperspectral data needed to be matched to their respective ground truth dry matter values. "Fruit-in-tray" mangoes were individually numbered with a marker corresponding to individual manual DM measurements. "Fruit-on-tree" data, on the other hand, required assigning labelled data to individual trees. This was automated by projecting the data to the world reference frame, using camera extrinsics obtained as per the approach in Wendel and Underwood (2017a), and assigning them to the closest trees using segmented LIDAR data (Wellington et al., 2012; Underwood et al., 2015, 2016; Stein et al., 2016).

2.3. Spectral preprocessing

Illumination compensation was performed as per Wendel and Underwood (2017b) and Drew and Finlayson (2007). First the spectra were converted to at-sensor sample pseudo-radiance, $\mathbf{l}_s$, to account for the effects of non-uniform lens transmittance and sensor quantum efficiency, both spatially and spectrally, using the following equation (Suomalainen et al., 2014):

$$l_{s,\lambda} = \frac{dn_{s,\lambda} - dn_{sdc,\lambda}}{dn_{ff,\lambda} - dn_{ffdc,\lambda}} l_{ff,\lambda}, \tag{1}$$

where $dn_{s,\lambda}$ and $dn_{sdc\lambda}$ are raw digital number (DN) measurements of the sample and nearest in time dark current at wavelength λ respectively; $dn_{ff,\lambda}$ and $dn_{ffdc,\lambda}$ are flat field raw digital number measurements, acquired using an integrating sphere, and corresponding dark current.

For the "fruit-on-tree" data, the "LOGSEP" method described in Wendel and Underwood (2017b) was then applied to convert the spectra to reflectance. The LOGSEP algorithm requires reference illumination radiance and surface reflectance spectra for training, which were produced using the reference panels. When a reference panel is seen in the hyperspectral data, the illumination that the panel is exposed to at that point in time, $l_{i,\lambda}$, can be estimated with

$$l_{i,\lambda} = \frac{l_{r,\lambda}}{r_{r,\lambda}}, \tag{2}$$

where $l_{r,\lambda}$ and $r_{r,\lambda}$ are radiance and reflectance of the reference panel respectively. Reference panel pixels were manually selected as patches between 5 × 5 and 9 × 9 in size from the "mid grey" step of the panel (i.e. the second darkest step), and averaged to produce a single spectrum per reference panel observation. Patches were made as large as allowed by the apparent size of the panel in the hyperspectral data, which varied depending on the panel's distance from the camera. The calibration panel reflectance, $r_{r,\lambda}$, was measured in the lab prior to data collection using an ASD Fieldspec 3 spectrometer with controlled halogen illumination.

To generate reference reflectance spectra for LOGSEP training, sample surface pixel radiances, $l_{s,\lambda}$ (e.g. of mangoes, grass, leaves, etc.), that were imaged within a few seconds of the a calibration panel to ensure the illumination seen by the panel is representative of the light hitting other surfaces, were extracted from the data. Care was also taken to ensure that sample surfaces had a similar geometry with respect to the sun. These were then converted to reflectance using the following equation:

$$r_{s,\lambda} = \frac{l_{s,\lambda}}{l_{r,\lambda}} r_{r,\lambda}. \tag{3}$$

Using the resultant training data, the LOGSEP model was trained and used to estimate illumination for each pixel used in further processing. By dividing the at-sensor radiance of each pixel by its corresponding LOGSEP illumination estimate, $l^*_{i,\lambda}$, the reflectance for each pixel, $r^*_{s,\lambda}$, can be recovered:

$$r^*_{s,\lambda} = \frac{l_{s,\lambda}}{l^*_{i,\lambda}}. \tag{4}$$

A grid search was performed to determine the number of reflectance and illumination basis functions required by the LOGSEP model. Radiance-at-sensor values were simulated by multiplying ground truth vegetation reflectance spectra (from Eq. (3)) with all known illumination spectra (from Eq. (2)). Vegetation spectra, where normalised difference vegetation index (NDVI) was greater than 0.5, were selected as their reflectance recovery accuracy is of particular importance for regression. After calibrating LOGSEP models within the grid search ranges, reflectance spectra were recovered from the simulated data, and predicted reflectance values were compared with their known counterparts by calculating the spectral angle (SA) between them. The number of basis functions was chosen where the resultant average SA asymptotes as they are increased. The regularisation parameter λ was determined separately, by also performing a grid search in the same manner.

For "fruit-in-tray" data, where reference panels were placed near the fruit (see Fig. 3(a)), linear interpolation across time (referred to as "INT" in Wendel and Underwood (2017b)) was used to obtain the illumination estimate, $l^*_{i,\lambda}$, at the time stamp of a given pixel (Suomalainen et al., 2014), instead of the LOGSEP approach. This was done because all "fruit-in-tray" data were obtained within a short period of time of imaging the reference panels, which allows for accurate interpolation (Wendel and Underwood, 2017b).

Due to the camera's low quantum efficiency at the extreme ends of the spectrum, which results in a considerably decreased SNR, the first and last ten bands of the reflectance spectrum were removed, resulting in a reduced range from 411.3 to 867.0 nm. Then, all reflectance spectra were normalised with a standard normal variate (SNV) transform:

$$r_{SNV,\lambda} = \frac{r_\lambda - \bar{\mathbf{r}}}{\sigma_{\mathbf{r}}}, \tag{5}$$

where $\mathbf{r}$ is the reflectance spectrum, $\bar{\mathbf{r}}$ its mean and $\sigma_{\mathbf{r}}$ its standard deviation. The biased standard deviation was used for SNV calculations in this paper. Savitzky-Golay (Savitzky and Golay, 1964) filtering and derivatives were considered, but found to have no additional beneficial effect, as supported by previous findings (Rungpichayapichet et al., 2017). Binning of spectra also appeared to have no particular benefit, so all remaining 224 bands were retained prior to further processing.

2.4. Classification and regression

Two essential steps in the processing pipeline are classification, identifying individual mango pixels/spectra, and regression, estimating DM of identified mango pixels. For both steps, two approaches were examined in this paper: PLS and CNN.

Previous work has shown that PLS is a promising approach to perform regression with high dimensional, collinear (correlated) and noisy input data, making it ideally suited for spectroscopic applications (Wold et al., 2001). Several previous studies have successfully used PLS for predicting fruit properties from spectroscopic data (eg. Walsh et al. (2004), Acharya et al. (2017), Guthrie and Walsh (1997), Rungpichayapichet et al. (2017)). Partial least squares discriminant analysis (PLS-DA) is a variant of PLS that can be used when the output is categorical. Therefore, PLS was tested to be used for both classification and regression. For the experiments performed in this paper, the implementation from the Python scikit-learn package was used (Pedregosa et al., 2011). The model was trained with a maximum of 500 nonlinear iterative partial least squares (NIPALS) iterations and 20 components/factors. For PLS-DA, the two classes in the ground truth data were represented as either −1 or 1 for training of the PLS model, and a threshold of zero determined the predicted class.

CNNs and deep learning have garnered much interest in the computer vision community, mainly for classification (Krizhevsky et al., 2012; Simonyan and Zisserman, 2015; Szegedy et al., 2015) and object detection (Ren et al., 2015; He et al., 2016; Liu et al., 2016; He et al., 2017) applications using RGB images. More recently CNNs and deep learning have also been investigated for classifying hyperspectral data (Hu et al., 2015; Makantasis et al., 2015; Chen et al., 2016; Lee and Kwon, 2017; Zhong et al., 2017; Windrim et al., 2018a). The work by Windrim et al. (2016b,a, 2018b) in particular investigates deep learning to deal with variations in illumination. Due to the promise shown by these approaches for hyperspectral data and to provide a nonlinear alternative, a CNN was also implemented for classification and regression tasks. Moreover, as explained later in this section, the CNN allows simultaneous classification and prediction, simplifying the overall processing pipeline.

The general architecture used for the CNN model is based on the two convolutional layer plus two fully connected layer model examined by Windrim et al. (2016b) (see Fig. 5). No padding was used for the convolutions. It is important to understand that unlike most CNN architectures designed for computer vision applications, as per Windrim et al. (2016b) this model only uses one dimensional filters, convolving along the spectral dimension only. Its input therefore consists of individual spectra (individual hyperspectral image pixels), which are treated independently from adjacent pixels. Three output variations of this architecture were implemented: classification only (CNN-CLASS), regression only (CNN-REG) and combined classification/regression (CNN-COMB). For classification only, the network outputs two values; the "confidences" for the two classes (mango and non-mango). The cost function for the optimiser was the softmax cross entropy (Murphy, 2012), as is common for classification with neural networks:

$$L_{ce} = \frac{1}{N}\sum_{k=1}^{N}\left[-\sum_{i=1}^{M} y_{k,i}\log\left(\frac{e^{z_{k,i}}}{\sum_{j=1}^{N} e^{z_{k,j}}}\right)\right], \tag{6}$$

where y represents the true values (ground truth) in one-hot representation, z is the logits output of the last layer of the network, M is the number of classes (two in this case), and N is the number of samples in a given batch.

For regression to DM, there is a single scalar output, and the cost was calculated as the mean squared error (MSE) with respect to the ground truth:

$$L_{mse} = \frac{1}{N}\sum_{k=1}^{N}(z_k - y_k)^2, \tag{7}$$

where y is the ground truth DM value. For the combined classification/regression model, the last layer of the network outputs $M + 1$ values (i.e. three in this case), the DM estimate y_1 and the classification logits output y_i for $i \in 2, \ldots, M + 1$. Naturally, the cost function becomes a combination of Eqs. (6) and (7):

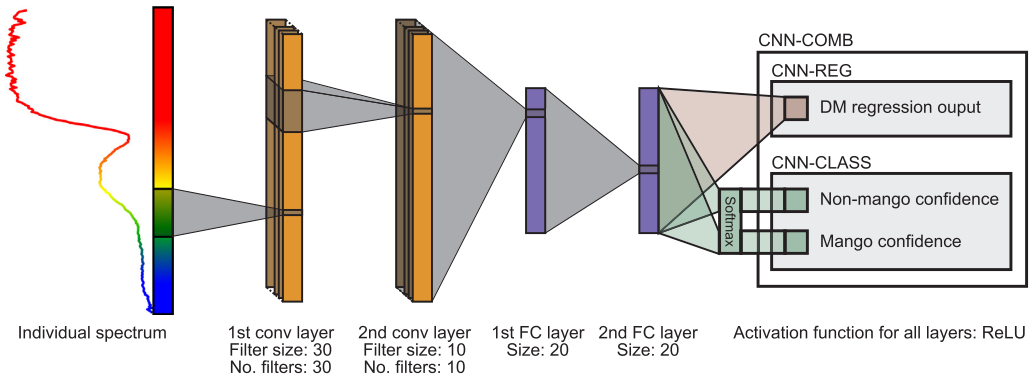


Fig. 5. CNN architecture used in this study. The architecture is the two convolutional layer plus two fully connected layer model presented by Windrim et al. (2016b). No padding was used for the convolutions. Three different variations were tested: CNN-REG, which only has one output for regressing to DM, CNN-CLASS, which provides two outputs representing the softmax "confidence" of mango and non-mango classes, and CNN-COMB, which is a combination of both with three output neurons.

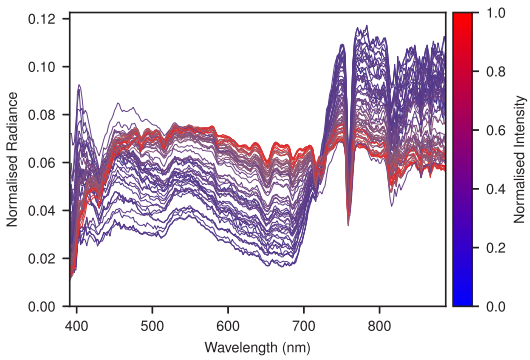
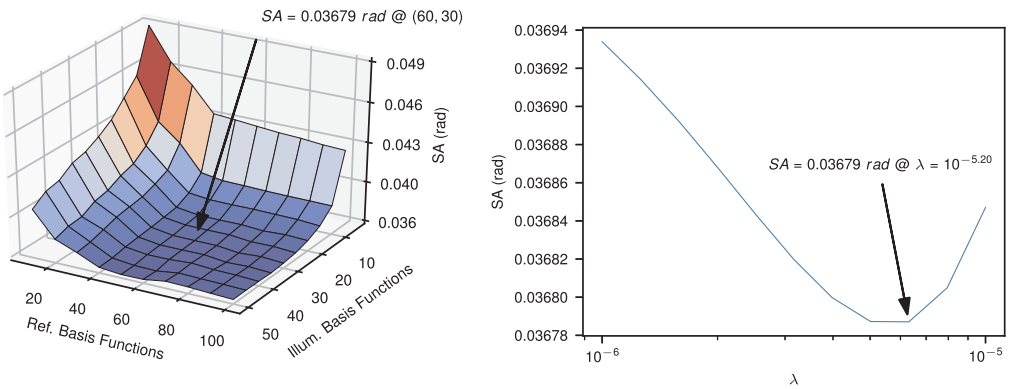


Fig. 6. Illumination spectra obtained from calibration panels (normalised to unity L^2 norm), colour coded according to their original mean intensity (normalised such that the brightest spectrum has a mean intensity of one). A clear red edge region can be seen in the darkest spectra. There is also a raised region between 500 and 600 nm in all spectra. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$L_{comb} = \frac{1}{N}\sum_{k=1}^{N}\left[y_{k,u}(z_{k,1}-y_{k,1})^2 - \sum_{i=2}^{M+1} y_{k,i}\log\left(\frac{e^{z_{k,i}}}{\sum_{j=1}^{N} e^{z_{k,j}}}\right)\right]. \tag{8}$$

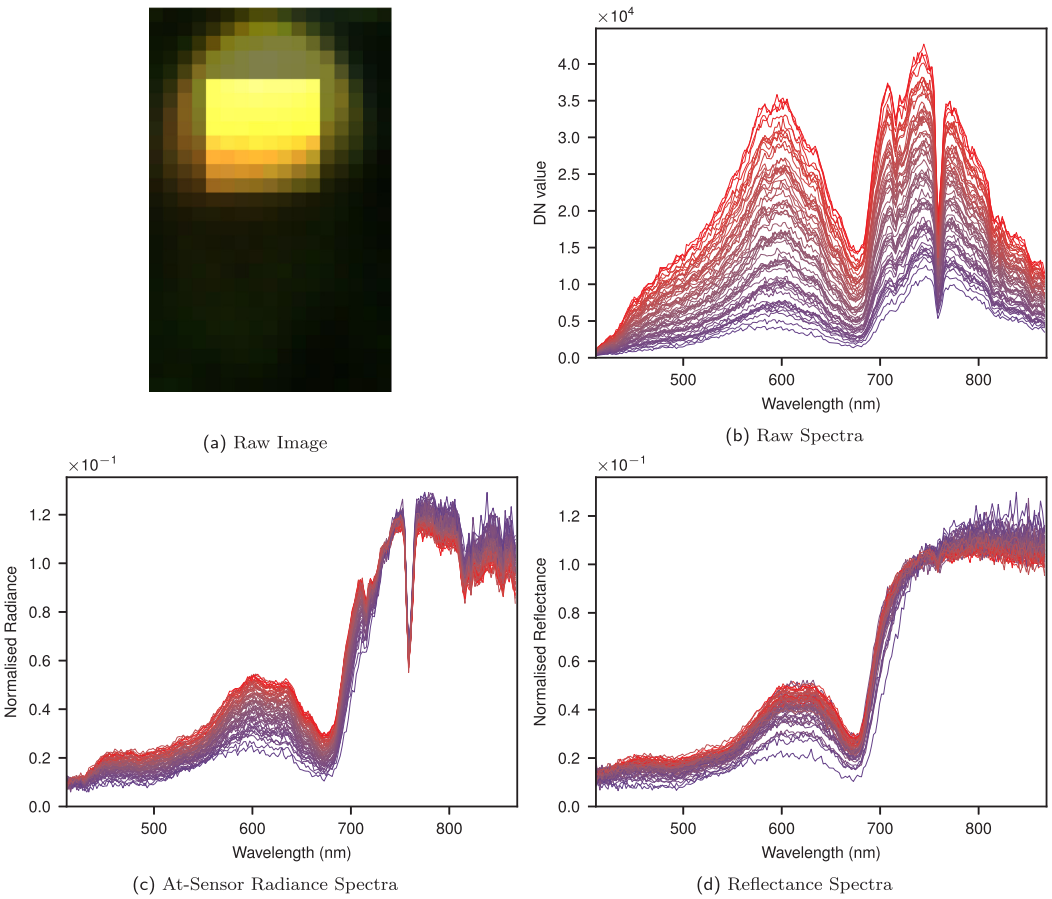
Note that the square error is now multiplied by $y_{k,u}$, where u is the one-hot index corresponding to the mango class. This ensures that the regression cost is only applied when the training data represents a mango, and zero otherwise. The models were implemented using Tensorflow (Abadi et al., 2016), and optimised using the Adam optimiser (Kingma and Ba, 2015) with a batch size of 100 over 20,000 iterations. Weight variables were initialised using the "Xavier" method (Glorot and Bengio, 2010), while biases were initialised to zero.

To test all classification and regression approaches, the data were first mean centred and scaled to unit variance. To test classification and regression performance, fivefold cross-validation was performed, and repeated five times. All metrics were computed per fold and averaged across all folds and repeats. To prevent data incest, spectra from multiple pixels in a single fruit were forced to exist wholly within single splits. For classification, the dataset consisted of labelled mangoes and negative data. Mango and non-mango samples (individual spectra) were balanced in each training and test set of each fold. For regression, the dataset comprised labelled mango pixels only, all of which included DM ground truth. Regression models were trained on and predicted



(a) Basis functions (b) λ

Fig. 7. Result of grid search to determine the number of basis functions (a) and regularisation parameter λ (b) for LOGSEP illumination compensation. After 60 and 30 reflectance and illumination basis functions respectively, accuracy, measured as the SA from the known ground truth reflectance, levels off. SA even increases slightly for reflectance basis functions greater than 60. For the regularisation parameter, a minimum SA can be found at $\lambda = 10^{-5.2}$.



(a) Raw Image (b) Raw Spectra (c) At-Sensor Radiance Spectra (d) Reflectance Spectra

Fig. 8. Raw digital number (DN) (b), at-sensor radiance (c) and reflectance (d) example spectra from a labelled patch on one mango fruit as shown in (a) in true colour. The spectra are colour coded according to their raw mean intensity. Radiance and reflectance spectra were normalised to unity L^2 norm. At 600 nm the spectra in (d) are more tightly clustered than prior to reflectance calibration (c), demonstrating the effectiveness of the illumination compensation method.

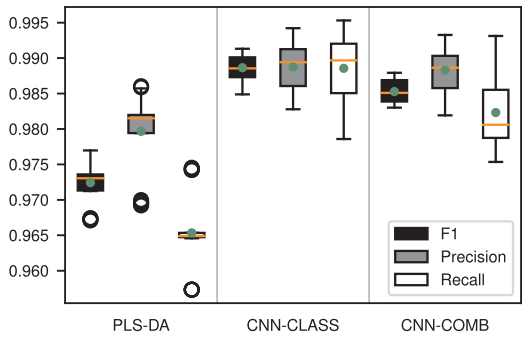


Fig. 9. Box plot of F1 score, precision and recall of the mango classification cross-validation results for the three classification methods: PLS-DA, CNN-CLASS and CNN-COMB.

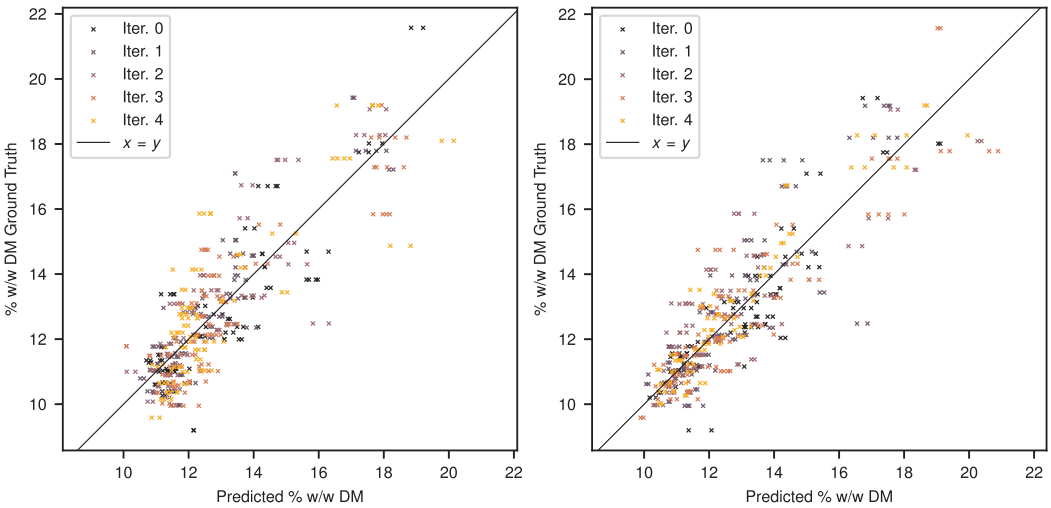
individual spectra, which were then averaged per group. The folds were generated in order to preserve the overall DM distribution in both training and test splits, using sorted stratification (Lowe, 2016). This was achieved by sorting the dataset by DM, and randomly withholding N mod k samples, where N is the number of samples and k is the number of desired folds (five in this case). The remaining samples were divided it into $floor(N/k)$ equally sized consecutive groups (i.e. each group is size k). Exactly one sample was then taken randomly, without replacement, from each group to generate a split, and this was repeated k times to generate k splits. The withheld samples were then randomly distributed into the splits. To test the regression performance of CNN-COMB, the negative samples were added to the mango training set of each fold, and thinned to balance the number of positive and negative samples.

Table 3
"Fruit-in-tray" CV regression results.

	R^2 (± 1 σ)	*RMSE* (% w/w) (± 1 σ)
PLS	0.743 (± 0.046)	1.169 (± 0.120)
CNN-REG	0.744 (± 0.086)	1.153 (± 0.208)

2.5. Mapping

Two end to end pipelines were tested to predict and map DM on an orchard scale: classification with CNN plus regression with PLS ("CNN-PLS"), and combined CNN classification and regression ("CNN-COMB"). For CNN-PLS, all pixels from a given orchard block were classified as



(a) PLS (b) CNN

Fig. 10. Scatter plots of all five iterations of the cross validation procedure (one repeat only) for the "fruit-in-tray" dataset.

Table 4
"Fruit-on-tree" CV regression results.

	R^2 (± 1 σ)	*RMSE* (% w/w) (± 1 σ)
PLS	0.580 (± 0.046)	1.168 (± 0.066)
CNN-REG	0.639 (± 0.053)	1.081 (± 0.078)
CNN-COMB	0.631 (± 0.049)	1.094 (± 0.071)

mango or non-mango using a trained CNN-CLASS network and mango pixels were passed to a trained PLS model to predict DM values. As described in Section 2.4, the combined CNN model, CNN-COMB, indicates if a pixel is a mango or not, and simultaneously outputs a DM estimate. For both methods, binary erosion was performed on the detected mango mask with a 3 × 3 square structuring element to remove isolated single pixel detections. All mangoes with ground truth DM data were used for training of all classifiers and regression models used for mapping (see Table 2). Due to the additional available training data (i.e. all labelled data, rather than just 80% for five fold cross validation), the number of CNN training iterations was increased to 25000. For these experiments PLS-DA was not used as a classifier as CNN-CLASS and CNN-COMB offered slightly better performance (see Fig. 9) and to limit the number of classifier-regressor permutations.

Once DM values were predicted using either method, they were then projected from the imagery to the world reference frame using the corresponding navigation system position and orientation data and associated to individual trees using the tree-segmented LIDAR data. The extrinsic camera offsets were determined using the line scan camera calibration method from Wendel and Underwood (2017a). DM values were projected to a plane, which is located at a fixed distance from the camera, and perpendicular to its viewing direction (z-axis) using the pinhole model. The projected points were then associated to the closest tree within a radius of 4 m, where tree centroids were obtained from LIDAR segmentation according to Underwood et al. (2015) and Stein et al. (2016). DM values were then averaged per tree side. The data were separated according to which side was being scanned, because fruit development can differ significantly on opposite sides of a tree due to factors such as varying exposure to sunlight and differing illumination conditions.

3. Results

This section outlines the experimental results. The ability of the LOGSEP method used in this paper to compensate for variable illumination is analysed, which also provides information about the nature of illumination variation in the data. The performance of the classifiers introduced in Section 2.4 is reported, followed by the prediction performance of the regression models from Section 2.4. Then spatial maps that average DM on a per tree basis using all imagery acquired from an entire orchard block were generated for scans taken over two days to test repeatability. This demonstrates the ability of the combined classification-regression pipelines described in Section 2.5 to process data on an orchard scale.

3.1. Illumination compensation

In Fig. 6 a wide variety of illumination spectra obtained from the calibration panels as per Section 2.3 are shown. The spectra are colour coded to signify their brightness. A clear difference in spectral profile can be seen between dark and bright spectra, most visibly where dark spectra exhibit a sharp red edge region between approx. 690 and 750 nm that is almost completely absent in bright illuminant spectra. This suggests the variability in illumination is not only influenced by the ratio of direct sun light to diffuse sky light, e.g. as shown for open field conditions in Wendel and Underwood (2017b), but additionally by secondary reflections within the complex canopy geometry. In layterms, the shaded reference panels are illuminated by light that has a green and NIR 'tinge'; it is coloured by the surrounding canopy, because of light reflected and transmitted by nearby vegetation. This highlights the need for explicit illumination compensation.

Due to the different nature and complexity of the illumination, the parameters determined in previous work using the LOGSEP approach cannot be assumed, and a grid search was performed to determine the optimal parameters. The results of the grid search used to select the number of LOGSEP model basis functions are shown as SA averages on the surface plot in Fig. 7(a). After approx. 60 reflectance basis functions the SA values level off and start to rise slightly. Similarly, no significant SA decrease can be observed after approx. 30 illumination basis functions. As a result, all further processing was performed using 60 reflectance and 30 illumination basis functions. The result of the grid

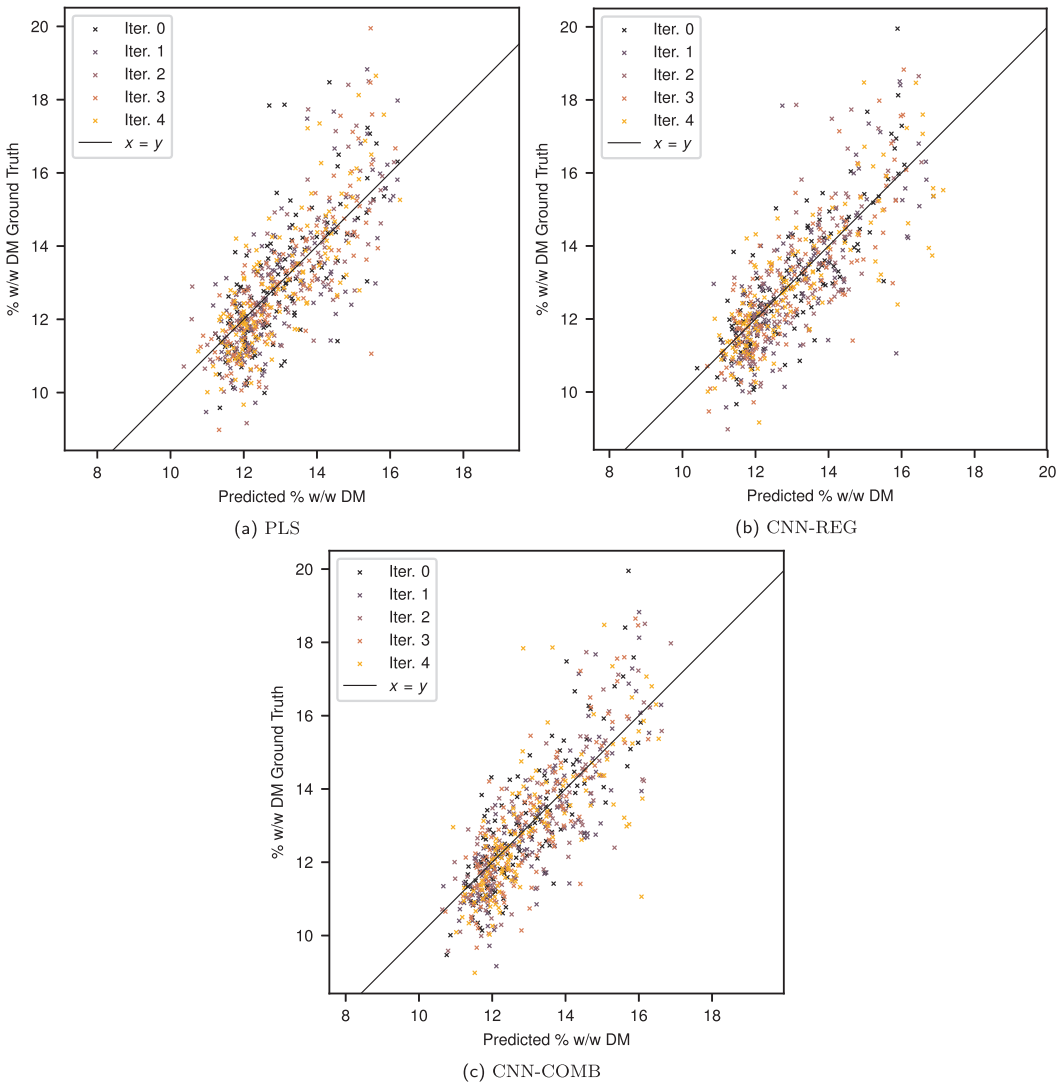


Fig. 11. Scatter plot of all five iterations of the cross validation procedure for the "fruit-on-tree" dataset (one repeat only) for (a) PLS, (b) CNN-REG and (c) CNN-COMB.

search to determine the regularisation parameter λ is shown in Fig. 7(b), where a clear minimum is visible, corresponding to a lambda value of $\lambda = 10^{-5.2}$.

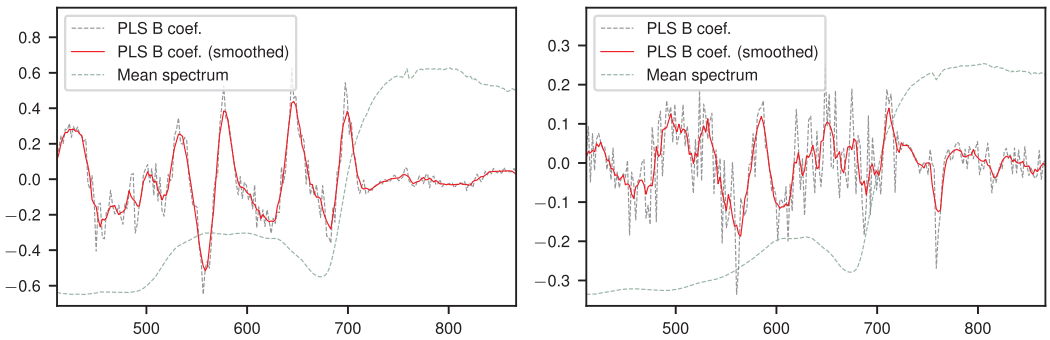
Fig. 8 demonstrates the calibration results for a patch on one mango fruit, selected due to its large range in illumination magnitude. Examining the reflectance spectra before and after compensation, a tighter clustering is observable (e.g. at 600 nm) in the reflectance spectra (Fig. 8(d)) compared to the radiance spectra (Fig. 8(c)) as a result of the LOGSEP illumination compensation algorithm. In addition, the dip corresponding to the O_2 "A" absorption band at approx. 760 nm, and other atmospheric features have been minimised.

3.2. Classification

Fig. 9 shows the classification cross validation (CV) F1, precision and recall results for PLS-DA, CNN, and CNN-COMB. PLS-DA exhibited good performance, with F1 scores for all folds greater than 0.95 and median greater than 0.97, but recall performance is significantly worse than precision (i.e. more false negatives than false positives). Both CNN-CLASS and CNN-COMB performed similarly, with more consistent precision and recall scores. CNN-CLASS, which is only trained to do classification, had slightly higher F1 and recall scores than CNN-COMB by a small margin (mean F1 0.989 vs. 0.985), while PLS-DA performed slightly worse on average than both CNN methods (mean F1 = 0.972).

3.3. Regression

"Fruit-in-tray" (see Table 2) CV regression results are outlined in Table 3, giving means and one standard deviation over all folds and repeats (25 in total). PLS performed similarly to CNN-REG (greater R^2 and lower *RMSE*), but exhibited a lower variance over the CV iterations. Fig. 10 gives scatter plots of the predictions, colour coded by CV iteration. A clear trend can be seen for both methods, generally adhering to the desired $x = y$ line. This is true for most of the DM range



(a) Fruit-in-tray (b) Fruit-on-tree

Fig. 12. PLS B coefficients for "fruit-in-tray" (a) and "fruit-on-tree" (b). The model was fit to the full datasets in both cases (no samples withheld), and the mean spectrum for all samples in each case is shown, scaled to fit the plot. The smoothed coefficient curves were produced by averaging over five bands (approx. 12 nm).

(approx. 10–22% w/w), as expected from the R^2 and *RMSE* values.

CV results for "fruit-on-tree" data (see Table 2) are shown in Table 4. Both CNN-REG and CNN-COMB perform better than PLS (greater R^2, lower *RMSE*), but CNN-REG exhibits a higher variance over the CV folds. The scatter plots in Fig. 11 provide similar looking results, where both CNN methods show slightly less noise. The results deteriorate at the tail ends of the DM distribution. In particular, fruit with higher DM values were underestimated, and fruit with lower DM values were overestimated.

Fig. 12 shows the B coefficients of the PLS model fitted to all samples in the "fruit-in-tray" and "fruit-on-tree" datasets. While the coefficients for the two datasets show some differences, there are a number of common positive peaks at approx. 410, 590, 650 and 700–710 nm, and negative peaks at 450, 560, 600–610 and 690 nm, indicating the wavelengths that are most important for predicting DM, though some of the peaks are separated by up to approx. 15 nm. Overall, the model's emphasis is in the visible region, corresponding to carotenoids and chlorophyll, whereas coefficients in the NIR are relatively small.

To quantify the effect of illumination compensation on performance, PLS was trained and tested using raw, radiance and LOGSEP reflectance spectra from the "fruit-on-tree" data, see Table 5. Both raw and radiance spectra perform similarly, as would be expected given that neither accounts for changes in illumination. The reflectance spectra recovered using the LOGSEP algorithm showed an improvement of approx. 0.025 in R_2 and 0.04% w/w in *RMSE* respectively.

3.4. Mapping

Fig. 13 shows mapping results as produced using the method outlined in Section 2.5. The maps are separated according to which tree side the data were acquired from (east or west), and by classification/regression method, CNN-COMB or CNN-PLS. Each row in Fig. 13 provides the results for one classification/regression method and row side combination. The three figures in each row give a map of day two, a difference map from day one to day two, and a scatter plot comparing mean DM values per tree for the two days of acquisition. By comparing the two days, repeatability of the method can be assessed.

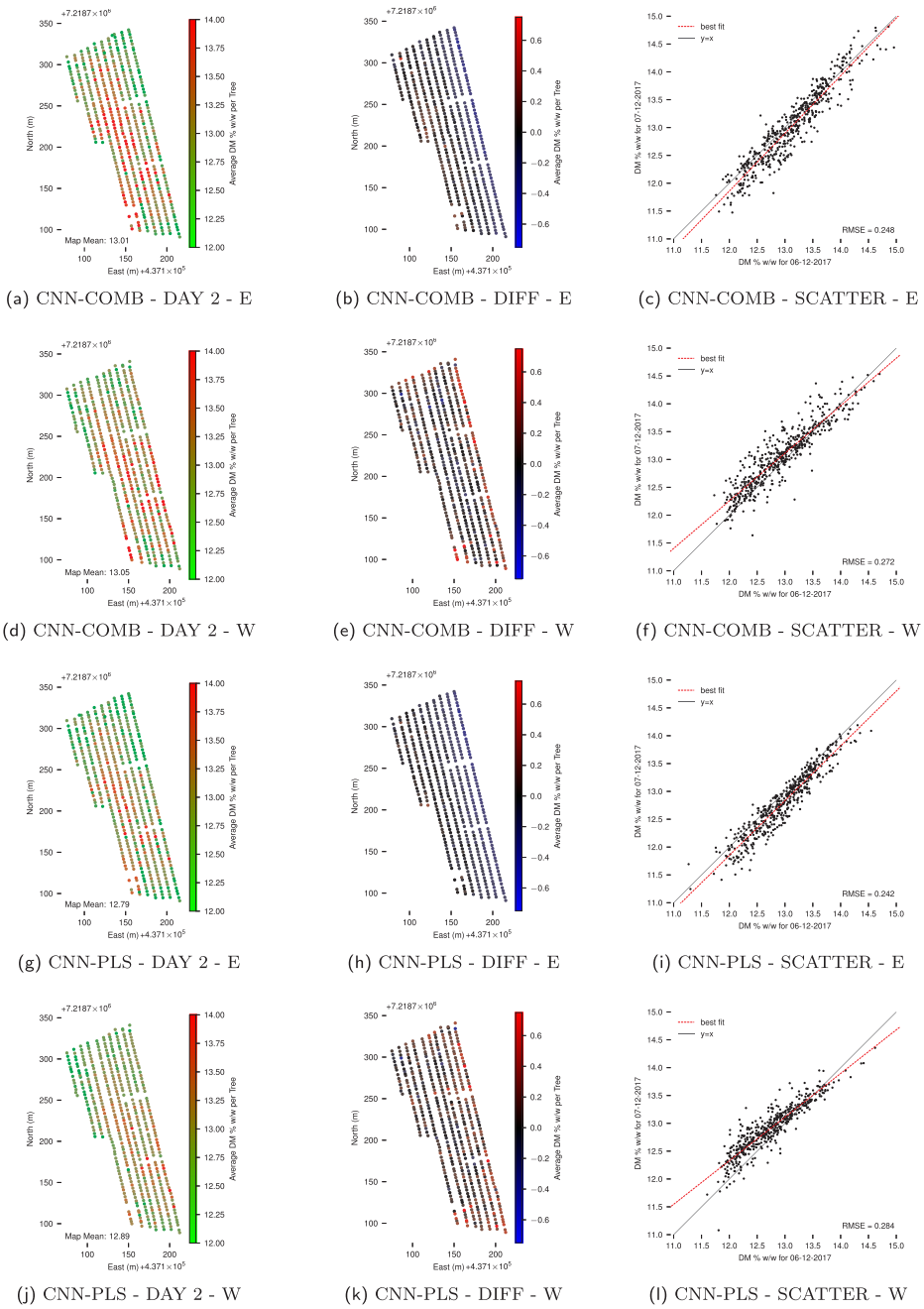
Qualitatively, variation of predicted DM between the two days is minimal as illustrated in Fig. 13(b), (e), (h) and (k), suggesting good repeatability. One particular exception is the eastern most row, where the maps deviate more significantly for both sides, in opposite directions. The scatter plots show good agreement between the two days on the east side ($RMSE = 0.248\%w/w$ for CNN-COMB and $RMSE = 0.242\%w/w$ for CNN-PLS). For the west side, the results from each day deviate somewhat for both methods ($RMSE = 0.272\%w/w$ and $RMSE = 0.284\%w/w$ for CNN-COMB and CNN-PLS respectively). There is still a high level of correlation, as signified by the close adherence to the line of best fit, but the relationship deviates from equality (i.e. the $y = x$ line).

In all maps (Fig. 13(a), (d), (g) and (j)), distinct regions of greater average DM can be observed, particularly towards the central area. Differences between the east and west sides can also be discerned, with the east side exhibiting greater contrast between areas of low and high DM.

Fig. 14 provides a close-up view of the combined classification and regression results for three sample trees with low, medium and high average DM, processed with CNN-COMB (Fig. 14(a), (b) and (c)), and high average DM processed with CNN-PLS (Fig. 14(d)). Most mangos exhibit a gradient, where the top of the fruit was estimated to have a higher DM than the bottom, which is most visible in Figs. 14(b), (c) and (d).

4. Discussion

The results demonstrate the challenges faced when operating a hyperspectral camera in orchards with natural lighting, which was found to be considerably more variable than what was encountered in previous work with row crops (Underwood et al., 2017; Wendel and Underwood, 2017b). Illumination varies not only due to changing cloud conditions, but also between directly lit and shaded regions. While visible shadows can be minimised by timing data acquisition such that the sun is located behind the camera, they can not be eliminated completely. Additionally, this is not possible when orchard rows run in an east-west direction, or when constrained by logistics of scanning a commercial farm. Both limitations were encountered during data acquisition in this study; i.e. blocks B and C ran east to west, while field logistics required starting a scan of the west side of block A just before midday for one of the repeats, even though acquisition after midday would be preferred. Fig. 15 shows an optimally lit tree (sun located behind camera) and one imaged close to midday with the sun overhead, as captured by the hyperspectral camera. In the latter image, extensive shadows can been seen on the mangos. Fig. 6 demonstrates the significant colour difference between direct sunlight and secondary illumination from shaded regions. As the spectra of objects seen by the hyperspectral camera are a combination of their intrinsic reflectance and ambient illumination, this has a significant effect on the apparent colour seen by the sensor. Shadow illumination spectra exhibit a strong red edge, which is likely the result of a significant contribution of light reflected from or transmitted through vegetation, whereas bright illumination spectra are dominated by direct sun light, so the relative impact of light reflected from other objects is minimised. All the illumination spectra also featured a noticeable raised region between



(a) CNN-COMB - DAY 2 - E (b) CNN-COMB - DIFF - E (c) CNN-COMB - SCATTER - E

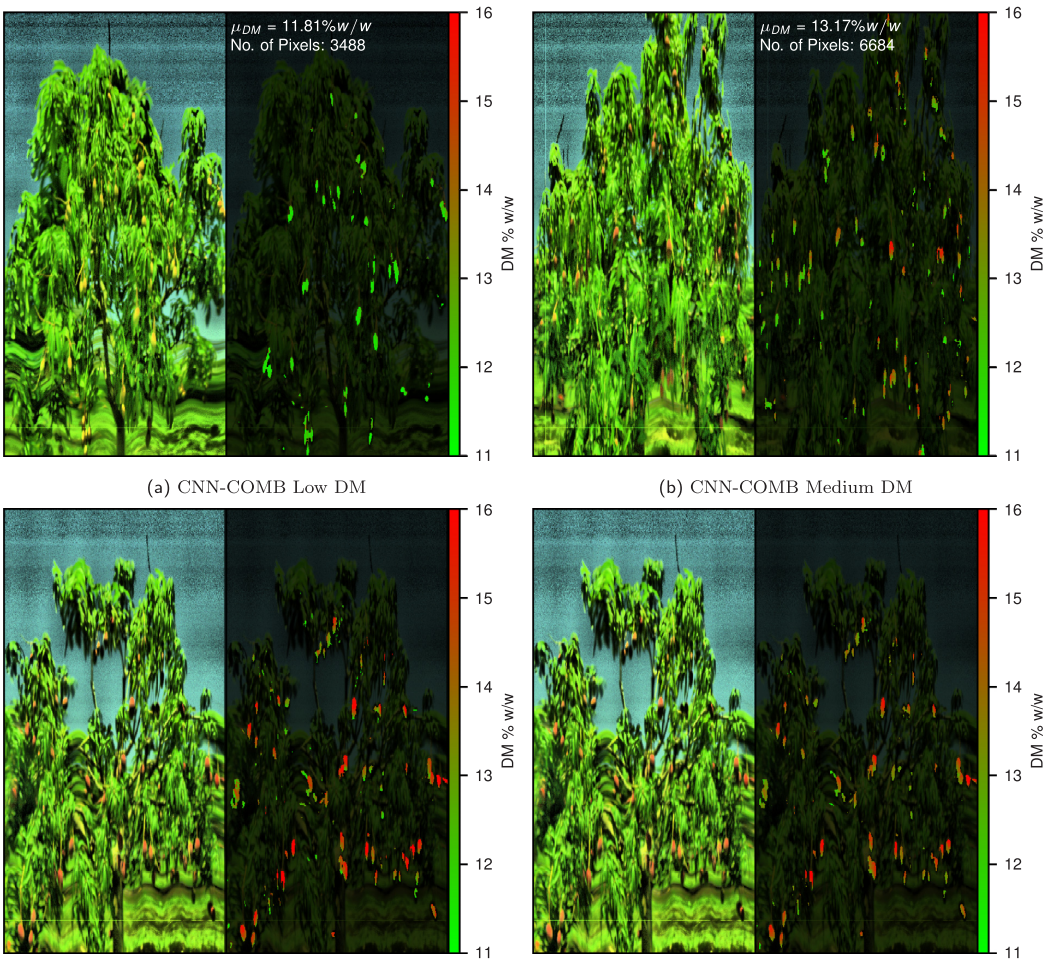
(d) CNN-COMB - DAY 2 - W (e) CNN-COMB - DIFF - W (f) CNN-COMB - SCATTER - W

(g) CNN-PLS - DAY 2 - E (h) CNN-PLS - DIFF - E (i) CNN-PLS - SCATTER - E

(j) CNN-PLS - DAY 2 - W (k) CNN-PLS - DIFF - W (l) CNN-PLS - SCATTER - W

Fig. 13. DM spatial maps generated using CNN-COMB (a–f) and CNN-PLS (g–l). Figs. (a), (d), (g) and (j) show maps for the second day only (7/12/2017), while Figs. (b), (e), (h) and (k) show the differences between the first and second days. Figs. (c), (f), (i) and (l) compare the results for the two days on a scatter plot, where each point represents the average DM for a tree on both days of acquisition.

approx. 500 and 600 nm. This coincides with the "chlorophyll bump" seen in vegetation and may also be caused by light reflected from leaves and grass. The LOGSEP illumination compensation method used in this paper compensated for these conditions reasonably well. Compared to the findings of Wendel and Underwood (2017b), the additional illumination complexity required a significant increase in the number of basis functions for optimal compensation performance (see Fig. 7). An example of the resultant reflectances can be seen in 8(d), which are much improved over 8(c), where a considerable spread is visible, largely due to the variation in lighting. As shown in Table 5, LOGSEP illumination compensation also results in a noticeable improvement in prediction performance.



(a) CNN-COMB Low DM (b) CNN-COMB Medium DM

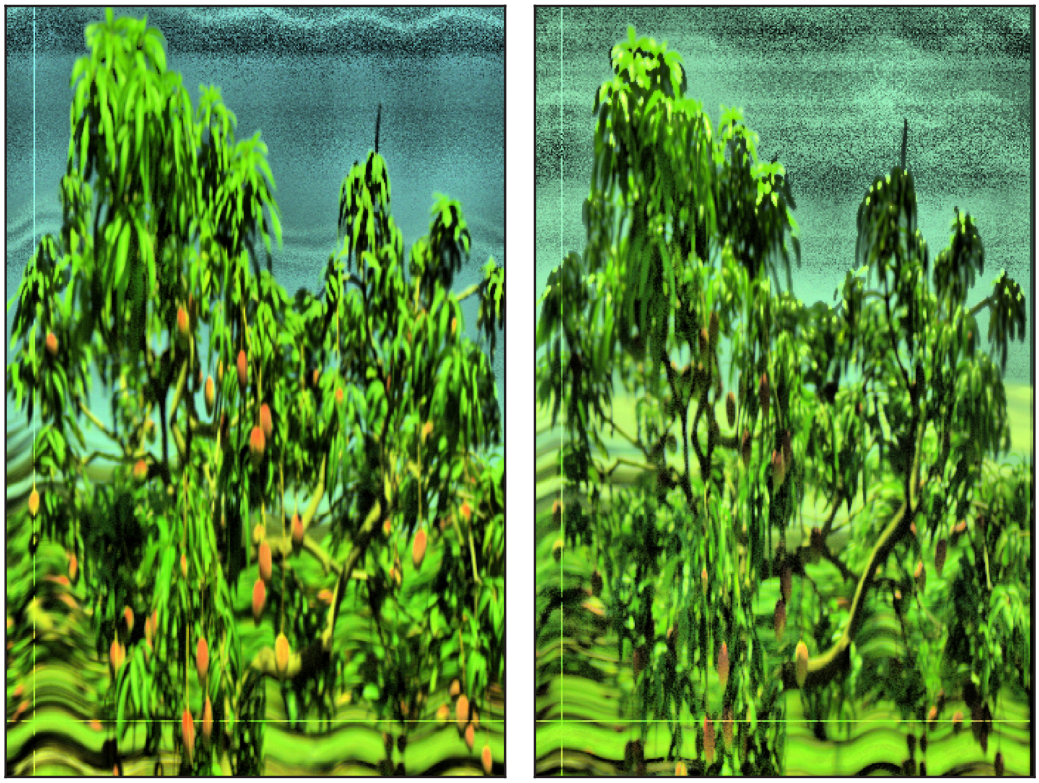
(c) CNN-COMB High DM (d) CNN-PLS High DM

Fig. 14. Sample mango detections and DM estimates for trees with (a) low, (b) medium, (c) high average DM, using CNN-COMB, and high average DM using CNN-PLS (d). In all four cases, the image on the left is a true colour RGB representation created by choosing three appropriate bands from the hyperspectral data. Adaptive histogram equalisation has also been applied to aid in interpretation. In the image on the right, colour coded mango DM values are overlaid on top of a darkened version of the true colour RGB image. (c) and (d) show the same tree on the same side and acquisition time to allow a qualitative comparison between CNN-PLS and CNN-COMB.

The "fruit-in-tray" results provide a first insight into the plausibility of predicting DM using reflectance in a field setting. The environment is as controlled as possible, while still using natural light outdoors. The effect of variable illumination was minimised with three calibration panels in close proximity to the fruit, and the scanned areas of the fruit were optimally sunlit (i.e. no shadows or extreme angles between sun and fruit surface). Mango DM prediction has previously only been demonstrated with hand-held contact sensors or in the lab, and the estimation of other parameters using hyperspectral imaging of mangoes has been shown in the lab only. The results shown in Table 3 and Fig. 10 confirm that predicting DM is possible outside the lab at a distance in natural lighting conditions. In addition, by averaging over several fruit instead of relying on predictions for individual fruit, much of the prediction noise in Fig. 10 may be ameliorated.

The "fruit-on-tree" results further support the notion that prediction of DM at a distance is possible using hyperspectral imaging in conditions that are realistic for eventual deployment at orchard-scale, with promising results ($R^2 = 0.64$ and $RMSE = 1.08\%w/w$). Notably, CNN-COMB performed as well as CNN-REG, while offering similar classification performance to CNN-CLASS (see Fig. 9), suggesting that the model is able to perform both classification and regression simultaneously. The regression algorithms over- and underestimated DM at very low and high values respectively, which was particularly noticeable for fruit with DM over 17% w/w. One reason for this might be the lack of samples with a high DM in the "fruit-on-tree" dataset, suggesting the need for training data with a high variance. Another cause might be greater noise encountered in the "on-tree" setting.

There are several sources of error in these experiments. Hyperspectral sensor noise is most prominent at the edges of the spectral range, where SNR decreases significantly, and this was addressed by cropping the spectral range by ten bands on each side to 411.3–867.0 nm. The DM estimates from the hand-held unit also exhibit some noise (validation $RMSE = 0.56\%w/w$), which affects both training and cross-validation data. In addition, for "fruit-on-tree" there is some



(a) Sun Behind Camera (b) Sun Overhead

Fig. 15. Example hyperspectral scans of the same tree in (a) good and (b) sub-optimal illumination conditions. In (a), the camera is behind the camera, which ensures that most of the mangos are properly lit. In (b), the sun is overhead, causing significant shadows. Both images are true colour RGB created by selecting three representative bands from the hyperspectral data and applying adaptive histogram normalization.

Table 5
"Fruit-on-tree" CV illumination compensation comparison.

	R^2 ($\pm$ 1 σ)	*RMSE* (% w/w) ($\pm$ 1 σ)
Raw	0.556 ($\pm$ 0.054)	1.200 ($\pm$ 0.072)
Radiance	0.555 ($\pm$ 0.046)	1.203 ($\pm$ 0.064)
LOGSEP Reflectance	0.580 ($\pm$ 0.044)	1.167 ($\pm$ 0.068)

All results obtained using PLS regression.

uncertainty when locating the hand-held instrument measurement area when labelling patches in the hyperspectral data. This can affect the results, because the DM measured by the hand-held instrument can vary across a fruit. A significant effort was made to associate the correct fruit with its corresponding manually measured DM value. Nevertheless, a small number of fruit were ambiguously located between trees, and some tag colours were difficult to discern due to lack of visual clarity in the hyperspectral data stemming from the lower spatial resolution or low light. There is also some additional noise introduced by assumptions made when compensating for illumination. Object surface pixels are assumed to be oriented and behave similarly to the calibration panels, which in turn are assumed to be perfectly Lambertian (Wendel and Underwood, 2017b). Mangoes have a relatively matte surface, which supports these assumptions and minimises any variance resulting from it. All of the aforementioned sources of error affect training and cross-validation data, and some also the mapped data. As a result, these sources of error may constitute a significant proportion of the noise seen in the predictions from the hyperspectral data, which should be kept in mind when interpreting the results.

The orchard-scale repeatability study showed good repeatability across two days when illumination conditions were similar, as shown in the scatter plots Figs. 13(c) ($RMSE = 0.248\%w/w$) and 13(i) ($RMSE = 0.242\%w/w$), and the maps in Fig. 13(a), (b), (g) and (h), whereas when scans were repeated at significantly different sun angles, performance deteriorated as shown in the scatter plots Figs. 13(f) ($RMSE = 0.272\%w/w$) and 13(l) ($RMSE = 0.284\%w/w$), and the maps in Fig. 13(d), (e), (j) and (k). On the second day the scan of the west side was started just before midday with sub-optimal illumination. This caused extensive shadows within tree canopies, and few fruit were directly lit (see Fig. 15(b)). Consequently the sensor signal to noise was significantly reduced and the extreme difference was only partly compensated for by LOGSEP leading to a lower repeatability score. There are clearly some limitations, particularly associated with low light conditions that are challenging to compensate for algorithmically. It is therefore always recommended to acquire data under best lighting conditions when possible, while algorithmic compensation may be used to handle factors that can not be controlled, such as the complex shading due to canopy geometry. The exact scope of such limitations requires further examination in future work to define. Closer examination of the PLS models (see Fig. 12) reveals a correlation with carotenoid and chlorophyll levels in the visible region of the spectra, which is likely very field and variety specific. Therefore, future work should also characterise the performance under varying weather, season/time of year, fruit variety and location, in order to determine to

what extent calibrations can be extrapolated to different conditions, and whether site specific calibration is required.

Quantifying the required accuracy of the system in order to be useful to growers is difficult, but experience has shown that a level of accuracy of approx. 0.5% w/w and precision of 1% w/w is a good target for predicting both whole block and tree DM means. The cross validation root mean square error (RMSE) shown in Table 4 ranges from 1.08% w/w to 1.17% w/w for fruit averages. Averaging whole trees with several tens to hundreds of fruit could bring the level of precision to within the required bounds (i.e. reduce the noise around the $x = y$ line in Fig. 11). While accuracy deteriorates at the tail ends of the DM range, it mostly remains approximately within 0.5% w/w. However, as mentioned previously, more testing is required to further study the accuracy of the system over several days with a wider range of variable lighting conditions, seasons, orchards and cultivars.

The results presented in this study support the use of hyperspectral imaging from a ground vehicle to predict and map mango DM, which would be a useful tool for managing orchards. Aggregate measurements of a whole block can be used to track maturity of a block over time to determine a single optimal harvest time. In addition, the distribution of DM within the block can be used quantify the risk of mature fruit dropping prior to a chosen harvest date. In some cases the data will inform economic viability to harvest sub-regions of blocks at two different times, which is not viable today due to a lack of spatial information. Once mechanised/robotic harvesting has been achieved, per-tree selective harvesting may become economically feasible, either by pre-scanning orchard blocks for later harvest or by predicting maturity and harvesting simultaneously on the same platform. The maps in this paper would feed into the algorithm that determines which trees to harvest on which date for maximal yield and fruit quality.

5. Conclusion

In this paper, the use of a hyperspectral imaging sensor and navigation system mounted to an UGV to efficiently estimate and map mango maturity was examined. The experiments revealed that predicting DM, an indicator of mango maturity, at distance using a hyperspectral sensor is possible and repeatable. The maps presented in the results, which estimate average DM per tree for all trees in an orchard block, provide valuable information to growers, allowing precise timing of harvesting operations. The approach would also help in making selective harvesting decisions, and guiding automated harvesting once technologically feasible. Further study will seek to determine how transferable calibration results are among different orchards, seasons, times of the year, and cultivars, as well as establishing the limitations of the algorithmic illumination compensation. As hyperspectral sensors are expensive compared to many other types of cameras, another important avenue of research would be to determine if an acceptable level of accuracy can be achieved with multispectral or even RGB cameras. In those cases, much of the methodology in this paper would still apply and be required.

Acknowledgements

This work is supported by the Australian Centre for Field Robotics (ACFR) at The University of Sydney and by funding from the Australian Government Department of Agriculture and Water Resources as part of its Rural R&D for profit programme. For more information about robots and systems for agriculture at the ACFR, please visit http://sydney.edu.au/acfr/agriculture. Thanks to Simpson Farms and Chad Simpson for their support and to Nicholas Anderson for providing the ground truth dry matter measurements.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., Brain, G., 2016. TensorFlow: a system for large-scale machine learning tensorflow: a system for large-scale machine learning. In: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI'16), pp. 265–284. <https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi>.

Acharya, U.K., Subedi, P.P., Walsh, K.B., 2017. Robustness of Tomato Quality Evaluation Using a Portable Vis-SWNIRS for Dry Matter and Colour. Int. J. Anal. Chem. 2017.

AMIA, 2015. Industry moves on quality standards. URL <https://www.industry.mangoes.net.au/resource-collection/2015/7/17/industry-moves-on-quality-standards>.

Andrade-Sanchez, P., Gore, M.A., Heun, J.T., Thorp, K.R., Carmo-Silva, A.E., French, A.N., Salvucci, M.E., White, J.W., 2014. Development and evaluation of a field-based high-throughput phenotyping platform. Funct. Plant Biol. 41 (1), 68–79.

Aquino, A., Millan, B., Diago, M.P., Tardaguila, J., 2018. Automated early yield prediction in vineyards from on-the-go image acquisition. Comput. Electron. Agric. 144, 26–36.

Bargoti, S., Underwood, J.P., 2017. Image segmentation for fruit detection and yield estimation in apple orchards. J. Field Robotics 34 (6), 1039–1060.

Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., Maurer, H.P., Reif, J.C., Würschum, T., Müller, J., et al., 2013. Breedvision - a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors 13 (3), 2830–2847.

Chen, Y., Jiang, H., Li, C., Jia, X., Ghamisi, P., 2016. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Trans. Geoscience Remote Sens. 54 (10), 6232–6251.

Comar, A., Burger, P., De Solan, B., Baret, F., Daumard, F., Hanocq, J.F., 2012. A semi-automatic system for high throughput phenotyping wheat cultivars in-field conditions: description and first results. Funct. Plant Biol. 39 (11), 914–924.

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., Furbank, R., 2014. Proximal remote sensing buggies and potential applications for field-based phenotyping. Agronomy 4 (3), 349–379.

Diago, M.P., Fernández-Novales, J., Gutiérrez, S., Marañón, M., Tardaguila, J., 2018. Development and validation of a new methodology to assess the vineyard water status by on-the-go near infrared spectroscopy. Frontiers Plant Sci. 9. <http://journal.frontiersin.org/article/10.3389/fpls.2018.00059/full>.

Dorj, U.-O., Lee, M., Yun, S.-s., 2017. An yield estimation in citrus orchards via fruit detection and counting using image processing. Comput. Electron. Agric. 140, 103–112. <http://linkinghub.elsevier.com/retrieve/pii/S0168169916312455>.

Drew, M.S., Finlayson, G.D., 2007. Analytic solution for separating spectra into illumination and surface reflectance components. JOSA A 24 (2), 294–303.

Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, pp. 249–256.

Gongal, A., Silwal, A., Amatya, S., Karkee, M., Zhang, Q., Lewis, K., 2016. Apple crop-load estimation with over-the-row machine vision system. Comput. Electron. Agric. 120, 26–35.

Greensill, C.V., Walsh, K.B., 2000. A remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit. Meas. Sci. Technol. 11 (12), 1674.

Guthrie, J., Walsh, K., 1997. Non-invasive assessment of pineapple and mango fruit quality using near infra-red spectroscopy. Anim. Prod. Sci. 37 (2), 253–263.

Gutiérrez, S., Diago, M.P., Fernández-Novales, J., Tardaguila, J., 2018. Vineyard water status assessment using on-the-go thermal imaging and machine learning. PLoS ONE 13 (2).

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, pp. 2980–2988.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

Hu, W., Huang, Y., Wei, L., Zhang, F., Li, H., 2015. Deep convolutional neural networks for hyperspectral image classification. J. Sens. 2015.

Hulens, D., Vandersteegen, M., Goedemé, T., 2017. Real-time vision-based UAV navigation in fruit orchards. In: Proceedings VISAPP 2017, pp. 1–6.

Kingma, D.P., Ba, J.L., 2015. Adam: a method for stochastic optimization. International Conference on Learning Representations 2015, 1–15.

Klose, R., Möller, K., Vielstädte, C., Ruckelshausen, A., 2010. Modular system architecture for individual plant phentotyping with an autonomous field robot. In: Proceedings of the 2nd International Conference of Machine Control & Guidance, pp. 299–307.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Adv. Neural Inf. Process. Syst. pp. 1097–1105.

Lee, H., Kwon, H., 2017. Going deeper with contextual CNN for hyperspectral image classification. IEEE Trans. Image Process. 26 (10), 4843–4855.

Linker, R., 2018. Machine learning based analysis of night-time images for yield prediction in apple orchard. Biosyst. Eng. 167, 114–125.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C., 2016. Ssd: single shot multibox detector. In: European Conference on Computer Vision. Springer, pp. 21–37.

Lowe, S.C., 2016. Stratified Validation Splits for Regression Problems. URL <http://scottclowe.com/2016-03-19-stratified-regression-partitions>.

MacDonald, S.L., Staid, M., Staid, M., Cooper, M.L., 2016. Remote hyperspectral imaging of grapevine leafroll-associated virus 3 in cabernet sauvignon vineyards. Comput. Electron. Agric. 130, 109–117.

Makantasis, K., Karantzalos, K., Doulamis, A., Doulamis, N., 2015. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In: IGARSS 2015. 2015 IEEE International Geoscience and Remote Sensing

Symposium. Proceedings, pp. 4959–4962.
Marshall, M., Thenkabail, P., 2015. Advantage of hyperspectral EO-1 Hyperion over multispectral IKONOS, GeoEye-1, WorldView-2, Landsat ETM+, and {MODIS} vegetation indices in crop biomass estimation. {ISPRS} J. Photogrammetry Remote Sens. 108, 205–218. <http://www.sciencedirect.com/science/article/pii/S0924271615001926>.
Mohammadi, V., Kheiralipour, K., Ghasemi-Varnamkhasti, M., 2015. Detecting maturity of persimmon fruit based on image processing technique. Scientia Horticulturae 184, 123–128.
Moharana, S., Dutta, S., 2016. Spatial variability of chlorophyll and nitrogen content of rice from hyperspectral imagery. {ISPRS} J. Photogrammetry Remote Sens. 122, 17–29. <http://www.sciencedirect.com/science/article/pii/S0924271616303422>.
Montes, J.M., Technow, F., Dhillon, B.S., Mauch, F., Melchinger, A.E., 2011. High-throughput non-destructive biomass determination during early plant development in maize under field conditions. Field Crops Res. 121 (2), 268–273.
Murphy, K.P., 2012. Machine Learning: A Probabilistic Perspective. The MIT Press.
Nandi, C.S., Tudu, B., Koley, C., 2014. A machine vision-based maturity prediction system for sorting of harvested mangoes. IEEE Trans. Instrum. Meas. 63 (7), 1722–1730.
Oldeland, J., Dorigo, W., Lieckfeld, L., Lucieer, A., Jürgens, N., 2010. Combining vegetation indices, constrained ordination and fuzzy classification for mapping semi-natural vegetation units from hyperspectral imagery. Remote Sens. Environ. 114 (6), 1155–1166.
Payne, A., Walsh, K., Subedi, P., Jarvis, D., 2014. Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and night time imaging. Comput. Electron. Agric. 100, 160–167.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in {P}ython. J. Machine Learning Res. 12, 2825–2830.
Pullanagari, R.R., Kereszturi, G., Yule, I.J., 2016. Mapping of macro and micro nutrients of mixed pastures using airborne AisaFENIX hyperspectral imagery. ISPRS J. Photogrammetry Remote Sens. 117, 1–10.
Qureshi, W.S., Payne, A., Walsh, K.B., Linker, R., Cohen, O., Dailey, M.N., 2017. Machine vision for counting fruit on mango tree canopies. Precision Agric. 18 (2), 224–244.
Rajkumar, P., Wang, N., ElMasry, G., Raghavan, G.S.V., Gariepy, Y., 2012. Studies on banana fruit quality and maturity stages using hyperspectral imaging. J. Food Eng. 108 (1), 194–200.
Rao, N.R., Garg, P.K., Ghosh, S.K., Dadhwal, V.K., 2008. Estimation of leaf total chlorophyll and nitrogen concentrations using hyperspectral satellite imagery. J. Agric. Sci. 146 (01), 65–75.
Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster R-CNN: towards real-time object detection with region proposal networks. In: Adv. Neural Inf. Process. Syst. pp. 91–99.
Robson, A., Rahman, M.M., Muir, J., 2017. Using Worldview Satellite Imagery to Map Yield in Avocado (Persea americana): A Case Study in Bundaberg, Australia. Remote Sens. 9 (12), 1223.
Rungpichayapichet, P., Nagle, M., Yuwanbun, P., Khuwijitjaru, P., Mahayothee, B., Müller, J., 2017. Prediction mapping of physicochemical properties in mango by hyperspectral imaging. Biosyst. Eng. 159, 109–120.
Salgadoe, A., Robson, A., Lamb, D., Dann, E., Searle, C., 2018. Quantifying the severity of phytophthora root rot disease in avocado trees using image analysis. Remote Sensing 10 (2).
Saranwong, S., Sornsrivichai, J., Kawano, S., 2004. Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy. Postharvest Biol. Technol. 31 (2), 137–145.
Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36 (8), 1627–1639.
Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. URL <https://arxiv.org/pdf/1409.1556.pdf>.
Sims, N.C., Culvenor, D., Newnham, G., Coops, N.C., Hopmans, P., 2013. Towards the operational use of satellite hyperspectral image data for mapping nutrient status and fertilizer requirements in Australian plantation forests. IEEE J. Sel. Top. Appl. Earth Observations Remote Sens. 6 (2), 320–328.
Stefas, N., Bayram, H., Isler, V., 2016. Vision-based UAV navigation in orchards. IFAC-PapersOnLine 49 (16), 10–15.
Stein, M., Bargoti, S., Underwood, J., 2016. Image based mango fruit detection, localisation and yield estimation using multiple view geometry. Sensors 16 (11), 1915. <http://www.mdpi.com/1424-8220/16/11/1915>.
Subedi, P., Walsh, K., Purdy, P., 2013. Determination of optimum maturity stages of mangoes using fruit spectral signatures. In: Acta Horticulturae, vol. 992. International Society for Horticultural Science (ISHS), Leuven, Belgium, pp. 521–528. <https://doi.org/10.17660/ActaHortic.2013.992.64>.
Subedi, P.P., Walsh, K.B., 2011. Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy. Postharvest Biol. Technol. 62 (3), 238–245.
Subedi, P.P., Walsh, K.B., Owens, G., 2007. Prediction of mango eating quality at harvest using short-wave near infrared spectrometry. Postharvest Biol. Technol. 43 (3), 326–334.
Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R., Kooistra, L., 2014. A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles. Remote Sens. 6 (11), 11013–11030.
Surya Prabha, D., Satheesh Kumar, J., 2015. Assessment of banana fruit maturity by image processing technique. J. Food Sci. Technol. 52 (3), 1316–1327.
Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., et al., 2015. Going Deeper with Convolutions. CVPR.
Underwood, J., Wendel, A., Schofield, B., McMurray, L., Kimber, R., 2017. Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. J. Field Robotics 34 (6), 1061–1083.
Underwood, J.P., Hung, C., Whelan, B., Sukkarieh, S., 2016. Mapping almond orchard canopy volume, flowers, fruit and yield using lidar and vision sensors. Comput. Electron. Agric. 130, 83–96.
Underwood, J.P., Jagbrant, G., Nieto, J.I., Sukkarieh, S., 2015. Lidar-based tree recognition and platform localization in orchards. J. Field Robotics 32 (8), 1056–1074.
Usha, K., Singh, B., 2013. Potential applications of remote sensing in horticulture-A review. Scientia Horticulturae 153, 71–83. https://doi.org/10.1016/j.scienta.2013.01.008.
Verbeke, J., Hulens, D., Ramon, H., Goedeme, T., De Schutter, J., 2014. The design and construction of a high endurance hexacopter suited for narrow corridors. In: 2014 International Conference on Unmanned Aircraft Systems, ICUAS 2014 - Conference Proceedings, pp. 543–551.
Walsh, K., Golic, M., Greensill, C., 2004. Sorting of fruit using near infrared spectroscopy: application to a range of fruit and vegetables for soluble solids and dry matter content. J. Near Infrared Spectrosc. 12, 141–148.
Walsh, K.B., Subedi, P.P., 2014. In-field monitoring of mango fruit dry matter for maturity estimation. In: XXIX International Horticultural Congress on Horticulture: Sustaining Lives, Livelihoods and Landscapes (IHC2014): 1119, pp. 273–278.
Wang, Q., Nuske, S., Bergerman, M., Singh, S., 2013. Automated crop yield estimation for apple orchards. In: Experimental Robotics, vol. 88. pp. 745–758. https://doi.org/10.1007/978-3-319-00065-7_50.
Wang, Z., Walsh, K., Verma, B., 2017. On-Tree Mango Fruit Size Estimation Using RGB-D Images. Sensors 17 (12), 2738. <http://www.mdpi.com/1424-8220/17/12/2738>.
Wellington, C., Campoy, J., Khot, L., Ehsani, R., 2012. Orchard tree modelling for advanced sprayer control and automatic tree inventory. In: {IEEE/RSJ} International Conference on Intelligent Robots and Systems ({IROS}) Workshop on Agricultural Robotics, pp. 5–6.
Wendel, A., Underwood, J., 2017a. Extrinsic parameter calibration for line scanning cameras on ground vehicles with navigation systems using a calibration pattern. Sens. (Switz.) 17 (11).
Wendel, A., Underwood, J., 2017b. Illumination compensation in ground based hyperspectral imaging. ISPRS J. Photogrammetry Remote Sens. 129, 162–178.
Windrim, L., Melkumyan, A., Murphy, R., Chlingaryan, A., Nieto, J., 2016a. Unsupervised feature learning for illumination robustness. In: Proceedings - International Conference on Image Processing, ICIP Vol. 2016-Augus. pp. 4453–4457.
Windrim, L., Melkumyan, A., Murphy, R.J., Chlingaryan, A., Ramakrishnan, R., 2018b. Pretraining for hyperspectral convolutional neural network classification. IEEE Transact. Geosci. Remote Sens. 56 (5), 2798–2810. https://doi.org/10.1109/TGRS.2017.2783886.
Windrim, L., Ramakrishnan, R., Melkumyan, A., Murphy, R., 2016b. Hyperspectral CNN classification with limited training samples. arXiv preprint arXiv:1611.09007.
Windrim, L., Ramakrishnan, R., Melkumyan, A., Murphy, R.J., 2018b. A physics-based deep learning approach to shadow invariant representations of hyperspectral images. IEEE Trans. Image Process. 27 (2), 665–677.
Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. In: Chemometrics and Intelligent Laboratory Systems, vol. 58. pp. 109–130.
Xue, Z., Du, P., Li, J., Su, H., 2017. Sparse graph regularization for robust crop mapping using hyperspectral remotely sensed imagery with very few in situ data. ISPRS J. Photogrammetry Remote Sens. 124, 1–15.
Zhang, X., Liao, C., Li, J., Sun, Q., 2013. Fractional vegetation cover estimation in arid and semi-arid environments using HJ-1 satellite hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 21, 506–512.
Zhong, Z., Li, J., Ma, L., Jiang, H., Zhao, H., 2017. Deep Residual Networks for Hyperspectral Image Classification. Institute of Electrical and Electronics Engineers.

Chapter 7

Conclusion

This thesis presented novel methods that allow the effective use of hyperspectral sensors from ground based mobile platforms, and demonstrated their use in precision agriculture (PA) applications. In Section 7.1 this chapter first summarises the approaches that were developed and then reiterates the specific contributions made in Section 7.2. There are number of avenues for future research that follow on from the findings in this thesis, which are explored in Section 7.3.

7.1 Summary

In this thesis, tools were developed that allow the effective use of hyperspectral cameras on ground based mobile platforms. First, the problem of variable illumination when operating using natural daylight under cloud cover was addressed with a compensation method that corrects for the spectral effects resulting from the changes in lighting. To allow accurate georeferencing or mapping, a calibration approach for determining the 6 DOF pose of line scanning cameras with respect to the mobile platform was proposed. Ground based mobile data captured using two different viewing geometries were then applied in a PA context. An approach for self-supervised weed detection using hyperspectral data was introduced, which automatically generates training datasets to successfully discriminate weed and crop plants, greatly reducing

the manual labelling effort required. Finally, an end-to-end pipeline that estimates the maturity of mangoes on an orchard-wide scale was developed. It detects fruit pixels, estimates their dry matter content, an indicator for maturity, and produces maps with per-tree averages, allowing for more precise harvest management.

Chapter 2 provided an overview of the literature pertaining to hyperspectral imaging from ground based mobile platforms. It included a general overview of hyperspectral imaging, how it is acquired and its many applications. A summary of commonly used processing techniques was also presented. This includes correcting for illumination and atmospheric effects, as well as preprocessing steps to prepare the data for subsequent machine learning tasks. There are a broad range of machine learning techniques, the most common of which were also reviewed. A brief overview of frequently used georeferencing techniques to determine the world coordinates of hyperspectral pixels was provided, including approaches for geometric calibration of hyperspectral cameras, which is necessary for accurate direct georeferencing. The review revealed a significant gap in the literature pertaining to the use of hyperspectral sensors on ground based mobile platforms. A clear need for approaches that effectively deal with variable illumination and allow for accurate georeferencing was made apparent. Furthermore, examples of hyperspectral sensors on mobile vehicles being used in practical applications are also scarce. The dearth of research in this area is the primary motivation for this thesis.

Compensating for changes in illumination is a significant problem for hyperspectral imaging, particularly when operating below cloud level. Traditional approaches, which rely on the use of reference panels or radiative transfer codes, can be challenging and labour intensive to implement when using a ground based mobile platform. The illumination compensation approach proposed in Chapter 3 addresses this, outlining an alternative method for retrieving reflectance that requires only a few reference panel readings, even when lighting conditions change rapidly. The timing of reference panel readings does not need to be contemporaneous with data acquisition, as long as they are representative of the conditions in the scene. This enables the use of historical reference data from previous campaigns to correct for illumination changes

in the future, making the proposed approach very convenient, not only for ground based platforms but also lower altitude unmanned aerial vehicles (UAVs).

Once hyperspectral data have been processed to produce a desired outcome, e.g. classifying pixels or estimating their chemical composition, it is necessary to determine where those pixels are located in world coordinates. To achieve this, the 6 degrees of freedom (DOF) pose of the camera must be known accurately. Chapter 4 proposed a novel method that achieves this for line scanning cameras, taking into account some of the practical considerations inherent when imaging from a mobile ground vehicle. The approach requires a relatively small calibration pattern that must be imaged from a variety of orientations, and also computes and visualises the uncertainty of the result. The procedure is valid for any line scanning camera, not only hyperspectral sensors. No surveyed ground control points (GCPs) or auxiliary cameras are required, making the approach particularly convenient when imaging from ground based vehicles.

Chapter 5 then applied ground based hyperspectral imaging to a critical PA topic, site specific weed management (SSWM). By leveraging *a priori* knowledge of the regular and consistent seeding pattern in row crops, this thesis proposed a weed detection approach that automatically generates training datasets of weed and crop samples. This has several practical advantages, including eliminating manual labelling and adjusting to changing conditions on the fly, for example variations in plant physiological and chemical properties over time. The approach was validated in corn crop rows, and compared favourably to classification models trained with manually labelled datasets.

Finally, Chapter 6 proposed an end-to-end pipeline that produces orchard scale maps of mango dry matter (DM) content, an indicator of maturity. Previous work in this area mostly used contact spectroscopy, and there are no previous studies that measure the the DM of mangoes remotely while still on the tree, which is a world first result of this thesis. A line scanning hyperspectral camera oriented in a sideways configuration was used to acquire data of a mango orchard in Bundaberg. After training on a number of reference measurements, the system discriminates mango pixels from surrounding surfaces and estimates dry matter content for each one. The pixels are then associated to individual trees, allowing the production of dry matter

maps with per-tree averages. This result enables growers to more precisely predict an optimal harvest time or to selectively harvest specific areas in an orchard as they mature. It could also form a critical component of an automated harvesting system, once such technology is available.

The findings in this thesis represent a significant step forward in the use of hyperspectral imaging on ground based vehicles. The illumination compensation and spatial calibration tools allow the effective use of hyperspectral data to leverage the significant advantages of imaging in close proximity to the objects of interest. The benefits include centimetre spatial resolution or better, and viewing angles that would not be feasible at higher altitudes. For example, individual fruit in orchard trees can be resolved by acquiring data on the ground in a sideways facing push broom configuration, as demonstrated in Chapter 6. The potential of mobile ground based hyperspectral imaging was leveraged in two PA applications, SSWM and mango maturity estimation. Both add significant contributions to PA and demonstrate the utility of using hyperspectral imagers on ground based mobile platforms.

7.2 Contributions

This thesis contributes tools to spectrally correct and spatially calibrate line scan hyperspectral data on ground based mobile platforms, and application of these tools in precision agriculture. A detailed breakdown of the contributions is provided below.

Illumination compensation

- Evaluating the suitability of a previously developed logarithm subspace method for illumination and reflectance extraction [56] for use on large, high spatial and spectral resolution agriculture field datasets.
- The development of multiple field protocols for acquiring training data for the illumination compensation method by Drew and Finlayson [56]. These present different trade-offs between the accuracy of illumination compensation and the logistical complexity of the field work.

Geometric calibration

- An approach for estimating a line scanning camera's 6 DOF relative pose with respect to the platform. The method is attuned to the constraints imposed by scanning from a ground based vehicle.
- An approach for estimating the uncertainty of the pose solution using Markov Chain Monte Carlo (MCMC), and associated visualisation methods.
- Open source code available (https://github.com/acfr/calibrate-line-camera).

Weed detection

- A self-supervised framework that uses *a priori* knowledge of commercial farm seeding geometry to automatically generate training sets for per-pixel crop/weed discrimination using hyperspectral imaging (HSI) data.

Mango maturity mapping

- An analysis of the illumination related challenges encountered when acquiring HSI in fruit orchards due to the complex canopy geometry.
- A world-first end-to-end pipeline for detecting mangoes in an orchard, estimating DM (a commonly used fruit maturity indicator) and georeferencing the predictions to produce orchard-scale maturity maps.

7.3 Future Work

The approaches proposed in this thesis address several of the gaps uncovered by the literature review (Chapter 2). However, there are many areas where more research, data and analysis is required.

The illumination compensation approach proposed in Chapter 3 has been thoroughly evaluated. It is suitable not only when reference panel spectra have been acquired at the same time as the surfaces of interest, but also when the reference data were obtained independently, as long as they are representative of the range of illumination in the scene. However, further study with a wider array of data is required to gain a better understanding of the optimal composition of a training dataset for the approach. The work in Chapter 3 focused on illumination changes due to cloud movement, but did not evaluate the performance of the method in shadows. While the results in Chapter 6 suggest that the algorithm copes well in shaded areas within tree canopies, a comprehensive analysis that includes a comparison to ground truth reflectance and illumination measurements, as in Chapter 3, is required to more rigorously evaluate its performance.

Chapter 4 thoroughly evaluated the performance of the proposed geometric calibration method, demonstrating its effectiveness. Nevertheless, as presented, manual labelling of calibration pattern points in the line scan data is required. This component of the calibration procedure could be simplified with a detector that automatically extracts the pixel location and time stamp of each point. Future work could also endeavour to simplify the outlier rejection process, with the application of a robust optimisation routine such as random sample consensus (RANSAC) or similar. In addition, a further analysis of the amount of platform pose diversity required for scanning the calibration pattern to achieve a satisfactory result would be useful. Ideally, the results of such a study would provide an algorithm that computes the necessary scanning positions and orientations given a desired maximum uncertainty and practical pose constraints.

The self-supervised weed detection system presented in Chapter 5 would also benefit from testing with a greater variety of data, to confirm its robustness to changing conditions. Given the recent rise and popularity of deep learning, it would be essential to perform a follow-up study to determine how a convolutional neural network (CNN) could be implemented to enhance the performance of the framework. In order to tie the approach into a weed management system, some spatial smoothing and segmenta-

tion would likely be required to convert per-pixel classification results to identification of whole plants, followed by georegistration to allow subsequent targeted spay operations.

The end-to-end pipeline presented in Chapter 6 is a promising first step for estimating fruit maturity on an orchard-scale. However, more work is required to determine how transferable the calibration results are among different orchards, seasons, times of the year, and cultivars. As previously mentioned, the illumination compensation method should be evaluated to quantify its performance in an orchard environment against ground truth illumination and reflectance data as in Chapter 3. Also, while previous spectroscopic results suggest that the spectral resolution provided by hyperspectral sensors might be necessary, an in-depth study should examine the use of specifically tuned multispectral sensors or even RGB cameras for estimating fruit maturity. These sensors are considerably cheaper than hyperspectral cameras and would therefore be preferred for commercialisation if acceptable performance can be obtained. Another avenue of research is to apply the fruit classification component presented in Chapter 6 to yield estimation. This was studied in a recently submitted paper using the same mango orchard data with yield prediction results that rival state of the art RGB camera based systems [100].

List of References

[1] Aasen, H., Burkart, A., Bolten, A., and Bareth, G. (2015). Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: from camera calibration to quality assurance. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108:245–259.

[2] Abd-Elrahman, A., Sassi, N., Wilkinson, B., and Dewitt, B. (2016). Georeferencing of mobile ground-based hyperspectral digital single-lens reflex imagery. *Journal of Applied Remote Sensing*, 10(1):14002.

[3] Abdlaty, R., Orepoulos, J., Sinclair, P., Berman, R., and Fang, Q. (2018). High Throughput AOTF Hyperspectral Imager for Randomly Polarized Light. In *Photonics*, volume 5, page 3. Multidisciplinary Digital Publishing Institute.

[4] Adam, E., Mutanga, O., and Rugege, D. (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review. *Wetlands Ecology and Management*, 18(3):281–296.

[5] Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., and Sousa, J. J. (2017). Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sensing*, 9(11):1110.

[6] Adler-Golden, S., Berk, A., Bernstein, L. S., Richtsmeier, S., Acharya, P. K., Matthew, M. W., Anderson, G. P., Allred, C. L., Jeong, L. S., and Chetwynd, J. H. (1998). FLAASH, a MODTRAN4 atmospheric correction package for hyperspectral data retrievals and simulations. In *Proc. 7th Ann. JPL Airborne Earth Science Workshop*, pages 9–14.

[7] Agrawal, P., Tack, K., Geelen, B., Masschelein, B., Moran, P. M. A., Lambrechts, A., and Jayapala, M. (2016). Characterization of VNIR hyperspectral sensors with monolithically integrated optical filters. *Electronic Imaging*, 2016(12):1–7.

[8] Ahlberg, J. (2010). Estimating atmosphere parameters in hyperspectral data. In *SPIE Defense, Security, and Sensing*, pages 76952A—-76952A. International Society for Optics and Photonics.

[9] Alchanatis, V. and Cohen, Y. (2011). Spectral and Spatial Methods for Hyperspectral Image Analysis for Estimation of Biophysical and Biochemical Properties of Agricultural Crops. *Hyperspectral remote sensing of vegetation*, pages 289–308.

[10] Alissou, S. A. and Zhang, Y. (2014). Hyperspectral data compression using lasso algorithm for spectral decorrelation. In *Proc. SPIE*, volume 9124, pages 91240A–91240A–9.

[11] Apan, A., Held, A., Phinn, S., and Markley, J. (2004). Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery. *International Journal of Remote Sensing*, 25(2):489–498.

[12] Atkinson, P. M. and Tatnall, A. R. L. (1997). Introduction neural networks in remote sensing. *International Journal of remote sensing*, 18(4):699–709.

[13] Bachmann, C. M., Ainsworth, T. L., and Fusina, R. A. (2005). Exploiting manifold geometry in hyperspectral imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(3):441–454.

[14] Baiano, A., Terracone, C., Peri, G., and Romaniello, R. (2012). Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes. *Computers and Electronics in Agriculture*, 87:142–151.

[15] Barbieux, K., Constantin, D., and Merminod, B. (2016). Correction of Airborne Pushbroom Images Orientation Using Bundle Adjustment of Frame Images. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41:813–818.

[16] Barnes, R. J., Dhanoa, M. S., and Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied spectroscopy*, 43(5):772–777.

[17] Barnsley, M. J., Settle, J. J., Cutter, M. A., Lobb, D. R., and Teston, F. (2004). The PROBA/CHRIS mission: A low-cost smallsat for hyperspectral multiangle observations of the earth surface and atmosphere. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1512–1520.

[18] Basedow, R. W. and Zalewski, E. (1995). Characteristics of the HYDICE sensor. In *Summaries of the Fifth Annual JPL Airborne Earth Science Workshop. Volume 1: AVIRIS Workshop*, volume 1.

[19] Bedini, E., Van Der Meer, F., and Van Ruitenbeek, F. (2009). Use of HyMap imaging spectrometer data to map mineralogy in the Rodalquilar caldera, southeast Spain. *International Journal of Remote Sensing*, 30(2):327–348.

[20] Beeri, O., Phillips, R., Hendrickson, J., Frank, A. B., and Kronberg, S. (2007). Estimating forage quantity and quality using aerial hyperspectral imagery for northern mixed-grass prairie. *Remote Sensing of Environment*, 110(2):216–225.

[21] Belluardo, G., Barchi, G., Baumgartner, D., Rennhofer, M., Weihs, P., and Moser, D. (2016). Uncertainty analysis of a radiative transfer model using Monte Carlo method within 280–2500nm region. *Solar Energy*, 132:558–569.

[22] Berk, A., Anderson, G. P., Bernstein, L. S., Acharya, P. K., Dothe, H., Matthew, M. W., Adler-Golden, S. M., Chetwynd Jr, J. H., Richtsmeier, S. C., Pukall, B., and Others (1999). MODTRAN4 radiative transfer modeling for atmospheric correction. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pages 348–353. International Society for Optics and Photonics.

[23] Bethel, J., Lee, C., and Landgrebe, D. A. (2000). Geometric registration and classification of hyperspectral airborne pushbroom data. *International Archives of Photogrammetry and Remote Sensing*, 33(B7/1; PART 7):183–190.

[24] Bibring, J.-P., Langevin, Y., Gendrin, A., Gondet, B., Poulet, F., Berthé, M., Soufflot, A., Arvidson, R., Mangold, N., Mustard, J., and Others (2005). Mars surface diversity as revealed by the OMEGA/Mars Express observations. *Science*.

[25] Bravo, C., Moshou, D., West, J., McCartney, A., and Ramon, H. (2003). Early disease detection in wheat fields using spectral reflectance. *Biosystems Engineering*, 84(2):137–145.

[26] Brown, R. B. and Noble, S. D. (2005). Site-specific weed management: sensing requirements-what do we need to see? *Weed Science*, 53(2):252–258.

[27] Bruce, L. M., Koger, C. H., and Li, J. (2002). Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction. *Geoscience and Remote Sensing, IEEE Transactions on*, 40(10):2331–2338.

[28] Buddenbaum, H. and Steffens, M. (2012). The effects of spectral pretreatments on chemometric analyses of soil profiles using laboratory imaging spectroscopy. *Applied and Environmental Soil Science*, 2012.

[29] Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., Maurer, H. P., Reif, J. C., Würschum, T., Müller, J., and Others (2013). Breedvision - A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors*, 13(3):2830–2847.

[30] Cao, X., Zhou, F., Xu, L., Meng, D., Xu, Z., and Paisley, J. (2018). Hyperspectral image classification with Markov random fields and a convolutional neural network. *IEEE Transactions on Image Processing*, 27(5):2354–2367.

[31] Chandra, K. and Healey, G. (2008). Using coupled subspace models for reflectance/illumination separation. *Geoscience and Remote Sensing, IEEE Transactions on*, 46(1):284–290.

[32] Chen, T., Yuen, P., Richardson, M., Liu, G., and She, Z. (2014a). Detection of psychological stress using a hyperspectral imaging technique. *IEEE Transactions on affective computing*, 5(4):391–405.

[33] Chen, Y., Jiang, H., Li, C., Jia, X., and Ghamisi, P. (2016). Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10):6232–6251.

[34] Chen, Y., Lin, Z., Zhao, X., Wang, G., and Gu, Y. (2014b). Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected topics in applied earth observations and remote sensing*, 7(6):2094–2107.

[35] Cheriyadat, A. and Bruce, L. M. (2003). Why principal component analysis is not an appropriate feature extraction method for hyperspectral data. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International*, volume 6, pages 3420–3422. IEEE.

[36] Chin, J. A., Wang, E. C., and Kibbe, M. R. (2011). Evaluation of hyperspectral technology for assessing the presence and severity of peripheral artery disease. *Journal of vascular surgery*, 54(6):1679–1688.

[37] Christensen, S., Søgaard, H. T., Kudsk, P., Nørremark, M., Lund, I., Nadimi, E. S., and Jørgensen, R. (2009). Site-specific weed control technologies. *Weed Research*, 49(3):233–241.

[38] Cocks, T., Jenssen, R., Steward, I., Wilson, I., and Shields, T. (1998). The HyMap airborne hyperspectral sensor: the system, calibration and performance. In *Proceedings of the 1rd EARSeL Workshop on Imaging Spectrometry*, pages 37–42. EARSeL.

[39] Conde, O. M., Amado, M., García-Allende, P. B., Cobo, A., and López-Higuera, J. M. (2007). Evaluation of PCA dimensionality reduction techniques in imaging spectroscopy for foreign object detection. In Shen, S. S. and Lewis, P. E., editors, *Proceedings of SPIE - The International Society for Optical Engineering*, volume 6565, pages 65650M–65650M–11. International Society for Optics and Photonics.

[40] Conel, J. E., Green, R. O., Vane, G., Bruegge, C. J., Alley, R. E., and Curtiss, B. J. (1987). AIS-2 radiometry and a comparison of methods for the recovery of ground reflectance.

[41] Cope, J. S., Corney, D., Clark, J. Y., Remagnino, P., and Wilkin, P. (2012). Plant species identification using digital morphometrics: A review. *Expert Systems with Applications*, 39(8):7562–7573.

[42] Cox, T. F. and Cox, M. A. A. (2000). *Multidimensional Scaling, Second Edition*, volume 88.

[43] Dadon, A., Ben-Dor, E., Beyth, M., and Karnieli, A. (2011). Examination of spaceborne imaging spectroscopy data utility for stratigraphic and lithologic mapping. *Journal of Applied Remote Sensing*, 5(1):53507.

[44] Dahlback, A. and Stamnes, K. (1991). A new spherical model for computing the radiation field available for photolysis and heating at twilight. *Planetary and Space Science*, 39(5):671–683.

[45] Dale, L. M., Thewis, A., Boudry, C., Rotar, I., Dardenne, P., Baeten, V., and Pierna, J. a. F. (2013). Hyperspectral Imaging Applications in Agriculture and Agro-Food Product Quality and Safety Control: A Review. *Applied Spectroscopy Reviews*, 48(2):142–159.

[46] de Castro, A.-I., Jurado-Expósito, M., Gómez-Casero, M.-T., and López-Granados, F. (2012). Applying Neural Networks to Hyperspectral and Multispectral Field Data for Discrimination of Cruciferous Weeds in Winter Crops. *The Scientific World Journal*, 2012:1–11.

[47] De Rainville, F. M., Durand, A., Fortin, F. A., Tanguy, K., Maldague, X., Panneton, B., and Simard, M. J. (2012). Bayesian classification and unsupervised learning for isolating weeds in row crops. *Pattern Analysis and Applications*, 17(2):1–14.

[48] Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., and Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy*, 4(3):349–379.

[49] Dehaan, R. L., Weston, L. A., and Rumbachs, R. (2012). The design and the development of a hyperspectral unmanned aircraft mapping system for the detection of invasive plants. In *Eighteenth Australasian Weeds Conference*, pages 103–107.

[50] Dekker, A. G., Malthus, T. J., Wijnen, M. M., and Seyhan, E. (1992). The effect of spectral band width and positioning on the spectral signature analysis of inland waters. *Remote Sensing Of Environment*, 41(2-3):211–226.

[51] Delegido, J., Alonso, L., González, G., and Moreno, J. (2010). Estimating chlorophyll content of crops from hyperspectral data using a normalized area over reflectance curve (NAOC). *International Journal of Applied Earth Observation and Geoinformation*, 12(3):165–174.

[52] der Meer, F. D., der Werff, H. M. A., Van Ruitenbeek, F. J., Hecker, C. A., Bakker, W. H., Noomen, M. F., Van Der Meijde, M., Carranza, E. J. M., De Smeth, J. B., and Woldai, T. (2012). Multi-and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14(1):112–128.

[53] Dicker, D. T., Lerner, J., Van Belle, P., Guerry 4th, D., Herlyn, M., Elder, D. E., and El-Deiry, W. S. (2006). Differentiation of normal skin and melanoma using high resolution hyperspectral imaging. *Cancer biology & therapy*, 5(8):1033–1038.

[54] Dille, J. A. (2014). Plant Morphology and the Critical Period of Weed Control. In *Automation: The Future of Weed Control in Cropping Systems*, pages 51–69. Springer.

[55] Draréni, J., Roy, S., and Sturm, P. (2011). Plane-based calibration for linear cameras. *International Journal of Computer Vision*, 91(2):146–156.

[56] Drew, M. S. and Finlayson, G. D. (2007). Analytic solution for separating spectra into illumination and surface reflectance components. *JOSA A*, 24(2):294–303.

[57] Duan, S.-B., Li, Z.-L., Wu, H., Tang, B.-H., Ma, L., Zhao, E., and Li, C. (2014). Inversion of the PROSAIL model to estimate leaf area index of maize, potato, and sunflower fields from unmanned aerial vehicle hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation*, 26:12–20.

[58] Edelman, G. J., Gaston, E., Van Leeuwen, T. G., Cullen, P. J., and Aalders, M. C. G. (2012). Hyperspectral imaging for non-contact analysis of forensic traces. *Forensic science international*, 223(1-3):28–39.

[59] ElMasry, G., Wang, N., ElSayed, A., and Ngadi, M. (2007). Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering*, 81(1):98–107.

[60] ElMasry, G., Wang, N., and Vigneault, C. (2009). Detecting chilling injury in Red Delicious apple using hyperspectral imaging and neural networks. *Postharvest biology and technology*, 52(1):1–8.

[61] Fan, W., Hu, B., Miller, J., and Li, M. (2009). Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data. *International Journal of Remote Sensing*, 30(11):2951–2962.

[62] Fang, Y., Li, H., Liang, K., Hu, Y., Zhang, S., and Wang, H. (2014). Dimensionality Reduction of Hyperspectral Images Based on Robust Spatial Information Using Locally Linear Embedding. 11(10):1712–1716.

[63] Fassnacht, F. E., Neumann, C., Förster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P. K., and Koch, B. (2014). Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central european test sites. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):2547–2561.

[64] Fearn, T., Riccioli, C., Garrido-Varo, A., and Guerrero-Ginel, J. E. (2009). On the geometry of SNV and MSC. *Chemometrics and Intelligent Laboratory Systems*, 96(1):22–26.

[65] Fernandes, A. M., Oliveira, P., Moura, J. P., Oliveira, A. A., Falco, V., Correia, M. J., and Melo-Pinto, P. (2011). Determination of anthocyanin concentration in whole grape skins using hyperspectral imaging and adaptive boosting neural networks. *Journal of Food Engineering*, 105(2):216–226.

[66] Ferris, D. G., Lawhead, R. A., Dickman, E. D., Holtzapple, N., Miller, J. A., Grogan, S., Bambot, S., Agrawal, A., and Faupel, M. L. (2001). Multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia. *Journal of Lower Genital Tract Disease*, 5(2):65–72.

[67] Fischer, C. and Kakoulli, I. (2006). Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications. *Studies in Conservation*, 51(sup1):3–16.

[68] Freitas, S., Silva, H., Almeida, J., and Silva, E. (2018). Hyperspectral Imaging for Real-Time Unmanned Aerial Vehicle Maritime Target Detection. *Journal of Intelligent & Robotic Systems*, 90(3-4):551–570.

[69] Frentress, Z., Young, L. C., and Edwards, H. D. (1964). Field photometer with nine-element filter wheel. *Applied Optics*, 3(2):303–308.

[70] Gamon, J., Serrano, L., and Surfus, J. S. (1997). The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia*, 112(4):492–501.

[71] Gao, B.-C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3):257–266.

[72] Gao, B. C., Heidebrecht, K. B., and Goetz, A. F. H. (1993). Derivation of scaled surface reflectances from AVIRIS data. *Remote Sensing of Environment*, 44(2-3):165–178.

[73] Gao, B.-C., Montes, M. J., Davis, C. O., and Goetz, A. F. H. (2009). Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment*, 113:S17–S24.

[74] Garcia-Allende, P. B., Conde, O. M., Amado, M., Quintela, A., and Lopez-Higuera, J. M. (2008a). Hyperspectral data processing algorithm combining principal component analysis and K nearest neighbours. In Shen, S. S. and Lewis, P. E., editors, *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV*, volume 6966, pages 69660H–69660H–9.

[75] García-Allende, P.-B., Conde, O. M., and López-Higuera, J. M. (2011). Hyperspectral Imaging for Raw Material Sorting and Processed Product Quality Control. *Applications and Experiences of Quality Control, Prof. Ognyan Ivanov (Ed.), ISBN*, pages 953–978.

[76] Garcia-Allende, P. B., Conde, O. M., Mirapeix, J., Cubillas, A. M., and Lopez-Higuera, J. M. (2008b). Data processing method applying principal component analysis and spectral angle mapper for imaging spectroscopic sensors. *Sensors Journal, IEEE*, 8(7):1310–1316.

[77] Geelen, B., Tack, N., and Lambrechts, A. (2014). A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic. In *SPIE MOEMS-MEMS*, page 89740L. International Society for Optics and Photonics.

[78] Geladi, P. and Kowalski, B. R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, 185(C):1–17.

[79] Geladi, P., MacDougall, D., and Martens, H. (1985). Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied spectroscopy*, 39(3):491–500.

[80] Geladi, P. L. M. (2007). Calibration Standards and Image Calibration. In *Techniques and Applications of Hyperspectral Image Analysis*, pages 203–220. John Wiley & Sons, Ltd.

[81] Gersman, R., Ben-Dor, E., Beyth, M., Avigad, D., Abraha, M., and Kibreab, A. (2008). Mapping of hydrothermally altered rocks by the EO-1 Hyperion sensor, Northern Danakil Depression, Eritrea. *International Journal of Remote Sensing*, 29(13):3911–3936.

[82] Ghamisi, P., Plaza, J., Chen, Y., Li, J., and Plaza, A. J. (2017). Advanced spectral classifiers for hyperspectral images: A review. *IEEE Geoscience and Remote Sensing Magazine*, 5(1):8–32.

[83] Ghiyamat, A. and Shafri, H. Z. M. (2010). A review on hyperspectral remote sensing for homogeneous and heterogeneous forest biodiversity assessment. *International Journal of Remote Sensing*, 31(7):1837–1856.

[84] Gillies, R., Freeman, J. E., Cancio, L. C., Brand, D., Hopmeier, M., and Mansfield, J. R. (2003). Systemic effects of shock and resuscitation monitored by visible hyperspectral imaging. *Diabetes technology & therapeutics*, 5(5):847–855.

[85] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448.

[86] Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587.

[87] Goel, P. K., Prasher, S. O., Landry, J.-A., Patel, R. M., Bonnell, R. B., Viau, A. A., and Miller, J. R. (2003). Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn. *Computers and electronics in agriculture*, 38(2):99–124.

[88] Goetz, A. F. H. (2009). Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment*, 113:S5–S16.

[89] Gomez, A. H., He, Y., and Pereira, A. G. (2006). Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques. *Journal of food engineering*, 77(2):313–319.

[90] Gómez-Sanchis, J., Martín-Guerrero, J. D., Soria-Olivas, E., Martínez-Sober, M., Magdalena-Benedito, R., and Blasco, J. (2012). Detecting rottenness caused by Penicillium genus fungi in citrus fruits using machine learning techniques. *Expert Systems with Applications*, 39(1):780–785.

[91] Gonzalez, R. C. and Woods, R. E. (2007). *Digital Image Processing (3rd Edition)*. Pearson.

[92] Govender, M., Chetty, K., and Bulcock, H. (2007). A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water Sa*, 33(2).

[93] Gowen, A. A., O'donnell, C. P., Cullen, P. J., and Bell, S. E. J. (2008a). Recent applications of chemical imaging to pharmaceutical process monitoring and quality control. *European journal of pharmaceutics and biopharmaceutics*, 69(1):10–22.

[94] Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., and Frias, J. M. (2007). Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, 18(12):590–598.

[95] Gowen, A. A., O'donnell, C. P., Taghizadeh, M., Cullen, P. J., Frias, J. M., and Downey, G. (2008b). Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (Agaricus bisporus). *Journal of Chemometrics: A Journal of the Chemometrics Society*, 22(3-4):259–267.

[96] Gray, C. J., Shaw, D. R., and Bruce, L. M. (2009). Utility of hyperspectral reflectance for differentiating soybean (Glycine max) and six weed species. *Weed Technology*, 23(1):108–119.

[97] Green, R. (2001). Atmospheric correction now (ACORN). *developed by ImSpec LLC, available from Analytical Imaging and Geophysics LLC.*

[98] Guanter, L., Richter, R., and Kaufmann, H. (2009). On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing. *International Journal of Remote Sensing*, 30(6):1407–1424.

[99] Gupta, N., Ashe, P. R., and Tan, S. (2011). Miniature snapshot multispectral imager. *Optical Engineering*, 50(3):33203.

[100] Gutiérrez, S., Wendel, A., and Underwood, J. (2018). Ground based hyperspectral imaging for extensive mango yield estimation. *Manuscript submitted for publication in Computers and Electronics in Agriculture.*

[101] Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote sensing of environment*, 90(3):337–352.

[102] Haboudane, D., Tremblay, N., Miller, J. R., and Vigneault, P. (2008). Remote estimation of crop chlorophyll content using spectral indices derived from hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 46(2):423–437.

[103] Hakala, T., Honkavaara, E., Saari, H., Mäkynen, J., Kaivosoja, J., Pesonen, L., and Pölönen, I. (2013). Spectral Imaging From Uavs Under Varying Illumination Conditions. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1/W2:189–194.

[104] Halimi, A., Altmann, Y., Dobigeon, N., and Tourneret, J.-Y. (2011a). Nonlinear unmixing of hyperspectral images using a generalized bilinear model. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4153–4162.

[105] Halimi, A., Altmann, Y., Dobigeon, N., and Tourneret, J.-Y. (2011b). Unmixing hyperspectral images using the generalized bilinear model. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 1886–1889. IEEE.

[106] Hasanlou, M. and Samadzadegan, F. (2012). Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using K-NN classifier. *Geoscience and Remote Sensing Letters, IEEE*, 9(6):1046–1050.

[107] Haug, S., Michaels, A., Biber, P., and Ostermann, J. (2014). Plant classification system for crop/weed discrimination without segmentation. In *2014 IEEE Winter Conference on Applications of Computer Vision, WACV 2014*, pages 1142–1149. IEEE.

[108] He, K., Gkioxari, G., Dollar, P., and Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2017-Octob, pages 2980–2988. IEEE.

[109] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

[110] He, L., Li, J., Liu, C., and Li, S. (2018). Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines. *IEEE Transactions on Geoscience and Remote Sensing*, 56(3):1579–1597.

[111] He, X. and Xu, S. (2010). Artificial neural networks. *Process Neural Networks: Theory and Applications*, pages 20–42.

[112] Healey, G. and Slater, D. (1999). Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions. *Geoscience and Remote Sensing, IEEE Transactions on*, 37(6):2706–2717.

[113] Hearn, D. J. (2009). Shape analysis for the automated identification of plants from images of leaves. *Taxon*, 58(3):934–954.

[114] Heist, S., Zhang, C., Reichwald, K., Kühmstedt, P., and Notni, G. (2018). Generation and evaluation of hyperspectral 3D surface models based on a structured light system with hyperspectral snapshot mosaic sensors. In *Dimensional Optical Metrology and Inspection for Practical Applications VII*, volume 10667, page 106670F. International Society for Optics and Photonics.

[115] Hernández-Clemente, R., Navarro-Cerrillo, R. M., and Zarco-Tejada, P. J. (2014). Deriving predictive relationships of carotenoid content at the canopy level in a conifer forest using hyperspectral imagery and model simulation. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8):5206–5217.

[116] Heylen, R., Parente, M., and Gader, P. (2014). A review of nonlinear hyperspectral unmixing methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):1844–1868.

[117] Hol, J. D., Schön, T. B., and Gustafsson, F. (2010). Modeling and calibration of inertial and vision sensors. *The international journal of robotics research*, 29(2-3):231–244.

[118] Holzwarth, S., Müller, R., and Simon, C. (2005). Determination and monitoring of boresight misalignment angles during the hymap campaigns HYEUROPE 2003 and HYEUROPE 2004. In *4th EARSEL Workshop on Imaging Spectrometry, Warschau, Polen*, pages 27–29. Citeseer.

[119] Honkavaara, E., Eskelinen, M. A., Pölönen, I., Saari, H., Ojanen, H., Mannila, R., Holmlund, C., Hakala, T., Litkey, P., Rosnell, T., and Others (2016). Remote sensing of 3-D geometry and surface moisture of a peat production area using hyperspectral frame cameras in visible to short-wave infrared spectral ranges onboard a small unmanned airborne vehicle (UAV). *IEEE Transactions on Geoscience and Remote Sensing*, 54(9):5440–5454.

[120] Honkavaara, E., Saari, H., Kaivosoja, J., Pölönen, I., Hakala, T., Litkey, P., Mäkynen, J., and Pesonen, L. (2013). Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture. *Remote Sensing*, 5(10):5006–5039.

[121] Hruska, R., Mitchell, J., Anderson, M., and Glenn, N. F. (2012). Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle. *Remote Sensing*, 4(9):2736–2752.

[122] Hu, W., Huang, Y., Wei, L., Zhang, F., and Li, H. (2015). Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors*, 2015.

[123] Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote sensing of environment*, 25(3):295–309.

[124] Hui, B., Wen, G., Zhao, Z., and Li, D. (2012). Line-scan camera calibration in close-range photogrammetry. *Optical Engineering*, 51(5):53601–53602.

[125] Inoue, Y., Sakaiya, E., Zhu, Y., and Takahashi, W. (2012). Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sensing of Environment*, 126:210–221.

[126] Ishak, A. J., Hussain, A., and Mustafa, M. M. (2009). Weed image classification using Gabor wavelet and gradient field distribution. *Computers and Electronics in Agriculture*, 66(1):53–61.

[127] Jan, H., Zbyněk, M., Lucie, H., Kaplan, V., Lukeš, P., and Cudlín, P. (2008). Potentials of the VNIR airborne hyperspectral system AISA Eagle. *GIS Ostrava*, 27:1–6.

[128] Jaud, M., Le Dantec, N., Ammann, J., Grandjean, P., Constantin, D., Akhtman, Y., Barbieux, K., Allemand, P., Delacourt, C., and Merminod, B. (2018). Direct georeferencing of a pushbroom, lightweight hyperspectral system for mini-UAV applications. *Remote Sensing*, 10(2):204.

[129] Jiang, Y., Li, C., Robertson, J. S., Sun, S., Xu, R., and Paterson, A. H. (2018). Gphenovision: a ground mobile system with multi-modal imaging for field-based high throughput phenotyping of cotton. *Scientific reports*, 8(1):1213.

[130] Johnson, R. J., Williams, J. P., and Bauer, K. W. (2013). AutoGAD: An improved ICA-based hyperspectral anomaly detection algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6):3492–3503.

[131] Johnson, W. R., Wilson, D. W., Fink, W., Humayun, M. S., and Bearman, G. H. (2007). Snapshot hyperspectral imaging in ophthalmology. *Journal of biomedical optics*, 12(1):14036.

[132] Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4):663–666.

[133] Karpouzli, E. and Malthus, T. (2003). The empirical line method for the atmospheric correction of IKONOS imagery. *International Journal of Remote Sensing*, 24(5):1143–1150.

[134] Kavzoglu, T. and Mather, P. M. (2002). The role of feature selection in artificial neural network applications. *International Journal of Remote Sensing*, 23(15):2919–2937.

[135] Keshava, N. and Mustard, J. F. (2002). Spectral unmixing. *IEEE signal processing magazine*, 19(1):44–57.

[136] Kim, M. S. (1994). *The Use of Narrow Spectral Bands for Improving Remote Sensing Estimations of Fractionally Absorbed Photosynthetically Active Radiation (fapar)*. PhD thesis, University of Maryland at College Park.

[137] Kim, M. S., Chen, Y.-R., Cho, B.-K., Chao, K., Yang, C.-C., Lefcourt, A. M., and Chan, D. (2007). Hyperspectral reflectance and fluorescence line-scan imaging for online defect and fecal contamination inspection of apples. *Sensing and Instrumentation for Food Quality and Safety*, 1(3):151.

[138] Klose, R., Möller, K., Vielstädte, C., and Ruckelshausen, A. (2010). Modular system architecture for individual plant phentotyping with an autonomous field robot. In *Proceedings of the 2nd International Conference of Machine Control & Guidance*, pages 299–307.

[139] Knauer, U. and Seiffert, U. (2013). An approach for automated registration of hyperspectral images for boresight calibration. *DGPF Tagungsband*.

[140] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.

[141] Kruse, F. A. (1988). Use of airborne imaging spectrometer data to map minerals associated with hydrothermally altered rocks in the northern Grapevine Mountains, Nevada, and California. *Remote Sensing of Environment*, 24(1):31–51.

[142] Kruse, F. A., Boardman, J. W., and Huntington, J. F. (2003). Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE transactions on Geoscience and Remote Sensing*, 41(6):1388–1400.

[143] Kuusk, J. and Kuusk, A. (2010). Autonomous lightweight airborne spectrometers for ground reflectance measurements. In *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010 2nd Workshop on*, pages 1–4. IEEE.

[144] Lacar, F. M., Lewis, M. M., and Grierson, I. T. (2001). Use of hyperspectral imagery for mapping grape varieties in the Barossa Valley, South Australia. In *Geoscience and Remote Sensing Symposium, 2001. IGARSS'01. IEEE 2001 International*, volume 6, pages 2875–2877. IEEE.

[145] Lamb, D. W. and Brown, R. B. (2001). Pa—precision agriculture: Remote-sensing and mapping of weeds in crops. *Journal of Agricultural Engineering Research*, 78(2):117–125.

[146] Lapray, P.-J., Wang, X., Thomas, J.-B., and Gouton, P. (2014). Multispectral filter arrays: recent advances and practical implementation. *Sensors (Basel, Switzerland)*, 14(11):21626–59.

[147] Larsen, E. L. P., Randeberg, L. L., Aksnes, A., Svaasand, L. O., Olstad, E., and Haugen, O. A. (2011). Hyperspectral imaging of atherosclerotic plaques in vitro. *Journal of biomedical optics*, 16(2):26011.

[148] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

[149] Lee, H. and Kwon, H. (2017). Going deeper with contextual cnn for hyperspectral image classification. *IEEE Transactions on Image Processing*, 26(10):4843–4855.

[150] Lenz, A., Schilling, H., Perpeet, D., Wuttke, S., Gross, W., and Middelmann, W. (2014). Automatic in-flight boresight calibration considering topography for hyperspectral pushbroom sensors. In *2014 IEEE Geoscience and Remote Sensing Symposium*, pages 2981–2984.

[151] Li, D., Wen, G., Hui, B. W., Qiu, S., and Wang, W. (2014). Cross-ratio invariant based line scan camera geometric calibration with static linear data. *Optics and Lasers in Engineering*, 62:119–125.

[152] Li, D., Wen, G., and Qiu, S. (2016a). Cross-ratio–based line scan camera calibration using a planar pattern. *Optical Engineering*, 55(1):14104.

[153] Li, X., Wu, T., Liu, K., Li, Y., and Zhang, L. (2016b). Evaluation of the Chinese fine spatial resolution hyperspectral satellite TianGong-1 in urban land-cover classification. *Remote Sensing*, 8(5):438.

[154] Liang, H. (2012). Advances in multispectral and hyperspectral imaging for archaeology and art conservation. *Applied Physics A*, 106(2):309–323.

[155] Liang, L., Di, L., Zhang, L., Deng, M., Qin, Z., Zhao, S., and Lin, H. (2015). Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. *Remote Sensing of Environment*, 165:123–134.

[156] Lim, M.-S. and Lim, J. (2008). Visual measurement of pile movements for the foundation work using a high-speed line-scan camera. *Pattern Recognition*, 41(6):2025–2033.

[157] Lindblom, J., Lundström, C., Ljung, M., and Jonsson, A. (2017). Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies. *Precision Agriculture*, 18(3):309–331.

[158] Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE Assp magazine*, 4(2):4–22.

[159] Liu, D., Zeng, X.-A., and Sun, D.-W. (2015). Recent developments and applications of hyperspectral imaging for quality evaluation of agricultural products: a review. *Critical reviews in food science and nutrition*, 55(12):1744–1757.

[160] Liu, J., Zhang, S., and Deng, S. (2009). A method of plant classification based on wavelet transforms and support vector machines. In *Emerging Intelligent Computing Technology and Applications*, pages 253–260. Springer.

[161] Liu, M., Zhang, L., and Guo, E. (2007). Hyperspectral laser-induced fluorescence imaging for nondestructive assessing soluble solids content of orange. In *International Conference on Computer and Computing Technologies in Agriculture*, pages 51–59. Springer.

[162] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C. (2016). Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer.

[163] Lobo, J. and Dias, J. (2007). Relative pose calibration between visual and inertial sensors. *The International Journal of Robotics Research*, 26(6):561–575.

[164] Lu, G. and Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of biomedical optics*, 19(1):10901.

[165] Lu, R. and Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems engineering*, 93(2):161–171.

[166] Lucieer, A., Malenovskỳ, Z., Veness, T., and Wallace, L. (2014). HyperUAS—Imaging spectroscopy from a multirotor unmanned aircraft system. *Journal of Field Robotics*, 31(4):571–590.

[167] Luna, C., Mazo, M., Lázaro, J. L., Vázquez, J. F., and Others (2010). Calibration of line-scan cameras. *Instrumentation and Measurement, IEEE Transactions on*, 59(8):2185–2190.

[168] Lunga, D. and Ersoy, O. (2013). Spherical stochastic neighbor embedding of hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(2):857–871.

[169] Lunga, D. and Ersoy, O. (2014). Multidimensional artificial field embedding with spatial sensitivity. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2):1518–1532.

[170] Lunga, D., Prasad, S., Crawford, M., and Ersoy, O. (2014). Manifold-Learning-Based Feature Extraction for Classification of Hyperspectral Data: A Review of Advances in Manifold Learning. *Signal Processing Magazine, IEEE*, 31(1):55–66.

[171] Ma, L., Crawford, M. M., and Tian, J. (2010). Anomaly detection for hyperspectral images based on robust locally linear embedding. *Journal of Infrared, Millimeter, and Terahertz Waves*, 31(6):753–762.

[172] Ma, S. and Hu, Z. (2014). Hand-Eye Calibration. In Ikeuchi, K., editor, *Computer Vision: A Reference Guide*, pages 355–358. Springer US, Boston, MA.

[173] MacDonald, S. L., Staid, M., Staid, M., and Cooper, M. L. (2016). Remote hyperspectral imaging of grapevine leafroll-associated virus 3 in cabernet sauvignon vineyards. *Computers and Electronics in Agriculture*, 130:109–117.

[174] Maftoonazad, N., Karimi, Y., Ramaswamy, H., and Prasher, S. (2011). Artificial neural network modeling of hyperspectral radiometric data for quality changes associated with avocados during storage. *Journal of Food Processing and Preservation*, 35(4):432–446.

[175] Magwaza, L., Opara, U., Nieuwoudt, H., Cronje, P., Saeys, W., and Nicolaï, B. (2012). NIR Spectroscopy Applications for Internal and External Quality Analysis of Citrus Fruit-A Review. *Food and Bioprocess Technology*, 5(2):425–444.

[176] Mahlein, A.-K. (2016). Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*, 100(2):241–251.

[177] Mahlein, A.-K., Oerke, E.-C., Steiner, U., and Dehne, H.-W. (2012). Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology*, 133(1):197–209.

[178] Makantasis, K., Karantzalos, K., Doulamis, A., and Doulamis, N. (2015). Deep supervised learning for hyperspectral data classification through convolutional neural networks. *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4959–4962.

[179] Malkoff, D. B. and Oliver, W. R. (2000). Hyperspectral imaging applied to forensic medicine. In *Spectral Imaging: Instrumentation, Applications, and Analysis*, volume 3920, pages 108–117. International Society for Optics and Photonics.

[180] Martens, H., Nielsen, J. P., and Engelsen, S. B. (2003). Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Analytical Chemistry*, 75(3):394–404.

[181] Martens, H. and Stark, E. (1991). Extended multiplicative signal correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy. *Journal of pharmaceutical and biomedical analysis*, 9(8):625–635.

[182] Mather, P. M. and Koch, M. (2011). *Computer Processing of Remotely-Sensed Images: An Introduction, Fourth Edition.* John Wiley & Sons, Ltd.

[183] Matthew, M. W., Adler-Golden, S. M., Berk, A., Felde, G., Anderson, G. P., Gorodetzky, D., Paswaters, S., and Shippert, M. (2002). Atmospheric correction of spectral imagery: evaluation of the FLAASH algorithm with AVIRIS data. In *Applied Imagery Pattern Recognition Workshop, 2002. Proceedings. 31st*, pages 157–163. IEEE.

[184] McClure, W. F. (2003). 204 Years of near infrared technology: 1800-2003.

[185] Meganem, I., Deliot, P., Briottet, X., Deville, Y., and Hosseini, S. (2014). Linear–quadratic mixing model for reflectances in urban environments. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1):544–558.

[186] Mendoza, F., Lu, R., Ariana, D., Cen, H., and Bailey, B. (2011). Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. *Postharvest Biology and Technology*, 62(2):149–160.

[187] Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. U. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia plantarum*, 106(1):135–141.

[188] Mewes, T., Franke, J., and Menz, G. (2011). Spectral requirements on airborne hyperspectral remote sensing data for wheat disease detection. *Precision Agriculture*, 12(6):795–812.

[189] Middleton, E. M., Huemmrich, K. F., Cheng, Y.-b., and Margolis, H. a. (2011). Spectral bioindicators of photosynthetic efficiency and vegetation stress. *Hyperspectral remote sensing of vegetation*, pages 265–282.

[190] Milioto, A., Lottes, P., and Stachniss, C. (2018). Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in cnns. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2229–2235. IEEE.

[191] Mobley, C. (2016). Remote Sensing: The Atmospheric Correction Problem.

[192] Moghadam, P., Ward, D., Goan, E., Jayawardena, S., Sikka, P., and Hernandez, E. (2017). Plant Disease Detection Using Hyperspectral Imaging. In *Digital Image Computing: Techniques and Applications (DICTA), 2017 International Conference on*, pages 1–8. IEEE.

[193] Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A., and Ramon, H. (2004). Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks. *Computers and Electronics in Agriculture*, 44(3):173–188.

[194] Mostafa, M. M. R. (2001). Boresight calibration of integrated inertial/camera systems. In *Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS 2001), Banff, AB, Canada*, pages 5–8.

[195] Muller, R., Lehner, M., Muller, R., Reinartz, P., Schroeder, M., and Vollmer, B. (2002). A program for direct georeferencing of airborne and spaceborne line scanner images. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(1):148–153.

[196] Murchie, S., Arvidson, R., Bedini, P., Beisser, K., Bibring, J.-P., Bishop, J., Boldt, J., Cavender, P., Choo, T., Clancy, R. T., and Others (2007). Compact reconnaissance imaging spectrometer for Mars (CRISM) on Mars reconnaissance orbiter (MRO). *Journal of Geophysical Research: Planets*, 112(E5).

[197] Murphy, R. J., Monteiro, S. T., and Schneider, S. (2012). Evaluating classification techniques for mapping vertical geology using field-based hyperspectral sensors. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(8):3066–3080.

[198] Murphy, R. J., Taylor, Z., Schneider, S., and Nieto, J. (2015). Mapping clay minerals in an open-pit mine using hyperspectral and LiDAR data. *European Journal of Remote Sensing*, 48(1):511–526.

[199] Neto, J. C., Meyer, G. E., Jones, D. D., and Samal, A. K. (2006). Plant species identification using Elliptic Fourier leaf shape analysis. *Computers and electronics in agriculture*, 50(2):121–134.

[200] Nicolai, B. M., Theron, K. I., and Lammertyn, J. (2007). Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple. *Chemometrics and intelligent laboratory systems*, 85(2):243–252.

[201] Noble, S. D., Brown, R. B., and Crowe, T. G. (2002). The Use of Spectral Properties for Weed Detection and Identification – A Review. In *CSAE/SCGR*, pages 0–17.

[202] Noh, H. K. and Lu, R. (2007). Hyperspectral laser-induced fluorescence imaging for assessing apple fruit quality. *Postharvest Biology and Technology*, 43(2):193–201.

[203] Norris, K. H. (1983). Extracting information from spectrophotometric curves. Predicting chemical composition from visible and near-infrared spectra. In *Food research and data analysis: proceedings from the IUFoST Symposium, September 20-23, 1982, Oslo, Norway/edited by H. Martens and H. Russwurm, Jr.* London: Applied Science Publishers, 1983.

[204] Norris, K. H. and Williams, P. C. (1984). Optimization of Mathematical Treatments of Raw Near-Infrared Signal in the. *Cereal Chem*, 61(2):158–165.

[205] Oerke, E.-C., Mahlein, A.-K., and Steiner, U. (2014). Proximal sensing of plant diseases. In *Detection and Diagnostics of Plant Pathogens*, pages 55–68. Springer.

[206] O'farrell, M., Lewis, E., Flanagan, C., Lyons, W., and Jackman, N. (2005). Comparison of k-NN and neural network methods in the classification of spectral data from an optical fibre-based sensor system used for quality control in the food industry. *Sensors and Actuators B: Chemical*, 111:354–362.

[207] Okamoto, H., Murata, T., Kataoka, T., and Hata, S. I. (2007). Plant classification for weed detection using hyperspectral imaging with wavelet analysis. *Weed Biology and Management*, 7(1):31–37.

[208] Okamoto, H., Suzuki, Y., and Noguchi, N. (2014). Field Applications of Automated Weed Control: Asia. In *Automation: The Future of Weed Control in Cropping Systems*, pages 189–200. Springer.

[209] Oliveira, R. A., Tommaselli, A. M. G., and Honkavaara, E. (2016). Geometric calibration of a hyperspectral frame camera. *The Photogrammetric Record*, 31(155):325–347.

[210] Pal, M. and Foody, G. M. (2010). Feature selection for classification of hyperspectral data by SVM. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5):2297–2307.

[211] Palmer, J. M. and Grant, B. G. (2010). *The art of radiometry.* SPIE Press Bellingham.

[212] Perez, L. and Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621.*

[213] Perry, J., Ahmed, M., Abd-Elrahman, A., Bowman, S., Kaddoura, Y., and Watts, A. (2001). Precision Directly Georeferenced Unmanned Aerial Remote Sensing System: Performance Evaluation. In *Proceedings of the Institute of Navigation National Technical Meeting, San Diego, CA, USA*, volume 2830, page 9.

[214] Persson, M. and Åstrand, B. (2008). Classification of crops and weeds extracted by active shape models. *Biosystems Engineering*, 100(4):484–497.

[215] Pfaff, F., Maier, G., Aristov, M., Noack, B., Gruna, R., Hanebeck, U. D., Längle, T., Beyerer, J., Pieper, C., Kruggel-Emden, H., and Others (2017). Real-time motion prediction using the chromatic offset of line scan cameras. *at-Automatisierungstechnik*, 65(6):369–380.

[216] Praks, J., Kestilä, A., Tikka, T., Näsilä, A., Riwanto, B., Jovanovic, N., Niemelä, P., Silva, N., Vainio, R., and Janhunen, P. (2017). Aalto-1 Earth Observation nanosatellite mission status and in orbit experiments. In *General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS), 2017 XXXIInd*, pages 1–3. IEEE.

[217] Prosa, A., Colominaa, I., Navarroa, J. A., Antequerab, R., and Andrinalb, P. (2013). Radiometric Block Adjusment and Digital Radiometric Model Generation. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1(1):293–298.

[218] Pu, Y. Y., Feng, Y. Z., and Sun, D. W. (2015). Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: A review. *Comprehensive Reviews in Food Science and Food Safety*, 14(2):176–188.

[219] Qin, J., Burks, T. F., Kim, M. S., Chao, K., and Ritenour, M. A. (2008). Citrus canker detection using hyperspectral reflectance imaging and PCA-based image classification method. *Sensing and Instrumentation for Food Quality and Safety*, 2(3):168–177.

[220] Qin, J., Chao, K., Kim, M. S., Lu, R., and Burks, T. F. (2013). Hyperspectral and multispectral imaging for evaluating food safety and quality. *Journal of Food Engineering*, 118(2):157–171.

[221] Qu, Z., Kindel, B., and Goetz, A. F. H. (2001). Aerosol retrieval and spectral calibration of AVIRIS with HATCH. In *2001 AVIRIS Geoscience Workshop*.

[222] Ramirez-Paredes, J.-P., Lary, D. J., and Gans, N. R. (2016). Low-altitude Terrestrial Spectroscopy from a Pushbroom Sensor. *Journal of Field Robotics*, 33(6):837–852.

[223] Ravikanth, L., Jayas, D., White, N., Fields, P., and Sun, D.-W. (2017). Extraction of Spectral Information from Hyperspectral Data and Application of Hyperspectral Imaging for Food and Agricultural Products. *Food and Bioprocess Technology*, 10(1).

[224] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.

[225] Reguera-Salgado, J. and Martín-Herrero, J. (2012). High performance GCP-based Particle Swarm Optimization of orthorectification of airborne pushbroom imagery. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, pages 4086–4089. IEEE.

[226] Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99.

[227] Riaza, A. and Müller, A. (2010). Hyperspectral remote sensing monitoring of pyrite mine wastes: a record of climate variability (Pyrite Belt, Spain). *Environmental Earth Sciences*, 61(3):575–594.

[228] Richter, R. (1996). Atmospheric correction of DAIS hyperspectral image data. In *Algorithms for Multispectral and Hyperspectral Imagery II*, volume 2758, pages 390–400. International Society for Optics and Photonics.

[229] Rinnan, Å., van den Berg, F., and Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. 28(10):1201–1222.

[230] Roberts, D., Roth, K., and Perroy, R. (2011). Hyperspectral Vegetation Indices. *Hyperspectral Remote Sensing of Vegetation*, pages 309–328.

[231] Roberts, D. A., Yamaguchi, Y., and Lyon, R. J. P. (1986). Comparison of various techniques for calibration of AIS data. *NASA STI/Recon Technical Report N*, 87.

[232] Rodarmel, C. and Shan, J. (2002). Principal Component Analysis for Hyperspectral Image Classification. *Surveying and Land Information Science*, 62(2):115–122.

[233] Rouse Jr, J., Haas, R. H., Schell, J. A., and Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS.

[234] Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science (New York, N.Y.)*, 290(5500):2323–2326.

[235] Ruckelshausen, A., Biber, P., Dorna, M., Gremmes, H., Klose, R., Linz, A., Rahe, F., Resch, R., Thiel, M., Trautz, D., and Others (2009). BoniRob–an autonomous field robot platform for individual plant phenotyping. *Precision agriculture*, 9(841):1.

[236] Rungpichayapichet, P., Nagle, M., Yuwanbun, P., Khuwijitjaru, P., Mahayothee, B., and Müller, J. (2017). Prediction mapping of physicochemical properties in mango by hyperspectral imaging. *Biosystems Engineering*, 159:109–120.

[237] Ryu, C., Suguri, M., and Umeda, M. (2011). Multivariate analysis of nitrogen content for rice at the heading stage using reflectance of airborne hyperspectral remote sensing. *Field Crops Research*, 122(3):214–224.

[238] Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J., and Siegwart, R. (2018). weedNet: Dense Semantic Weed Classification Using Multispectral Images and MAV for Smart Farming. *IEEE Robotics and Automation Letters*, 3(1):588–595.

[239] Saari, H., Aallos, V.-V., Akujärvi, A., Antila, T., Holmlund, C., Kantojärvi, U., Mäkynen, J., and Ollila, J. (2009). Novel miniaturized hyperspectral sensor for UAV and space applications. In *Sensors, Systems, and Next-Generation Satellites XIII*, volume 7474, page 74741M. International Society for Optics and Photonics.

[240] Sáiz-Abajo, M. J., Mevik, B. H., Segtnan, V. H., and Næs, T. (2005). Ensemble methods and data augmentation by noise addition applied to the analysis of spectroscopic data. *Analytica Chimica Acta*, 533(2):147–159.

[241] Savitzky, A. and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639.

[242] Schläpfer, D. and Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: parametric orthorectification. *International Journal of Remote Sensing*, 23(13):2609–2630.

[243] Schneider, S., Melkumyan, A., Murphy, R. J., and Nettleton, E. (2012). A geological perception system for autonomous mining. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 2986–2991. IEEE.

[244] Schölkopf, B., Smola, A. J., Bach, F., et al. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press.

[245] Schweizer, J., Hollmach, J., Steiner, G., Knels, L., Funk, R. H. W., and Koch, E. (2012). Hyperspectral imaging-A new modality for eye diagnostics. *Biomedical Engineering/Biomedizinische Technik*, 57(SI-1 Track-P):293–296.

[246] Seidel, F. C., Kokhanovsky, A. A., and Schaepman, M. E. (2010). Fast and simple model for atmospheric radiative transfer. *Atmospheric Measurement Techniques*, 3(4):1129–1141.

[247] Servakaranpalayam, S. (2006). *Potential applications of hyperspectral imaging for the determination of total soluble solids, water content and firmness in mango*.

[248] Shafri, H. Z. M. and Hamdan, N. (2009). Hyperspectral imagery for mapping disease infection in oil palm plantationusing vegetation indices and red edge techniques. *American Journal of Applied Sciences*, 6(6):1031.

[249] Shafri, H. Z. M., Suhaili, A., and Mansor, S. (2007). The performance of maximum likelihood, spectral angle mapper, neural network and decision tree classifiers in hyperspectral image analysis. *Journal of Computer Science*, 3(6):419.

[250] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[251] Singer, R. B. and McCord, T. B. (1979). Mars-large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. In *Lunar and Planetary Science Conference Proceedings*, volume 10, pages 1835–1848.

[252] Slaughter, D. (2014). The Biological Engineer: Sensing the Difference Between Crops and Weeds. In *Automation: The Future of Weed Control in Cropping Systems*, pages 71–95. Springer.

[253] Slaughter, D. C., Giles, D. K., Fennimore, S. a., and Smith, R. F. (2008). Multispectral Machine Vision Identification of Lettuce and Weed Seedlings for Automated Weed Control. *Weed Technology*, 22(2):378–384.

[254] Slonecker, T. (2012). Analysis of the effects of heavy metals on vegetation hyperspectral reflectance properties. In Thenkabail, P. S., Huete, A., and Lyon, J. G., editors, *Hyperspectral remote sensing of vegetation*, pages 561–578. CRC Press.

[255] Søgaard, H. T. (2005). Weed classification by active shape models. *Biosystems engineering*, 91(3):271–281.

[256] Staenz, K., Secker, J., Gao, B.-C., Davis, C., and Nadeau, C. (2002). Radiative transfer codes applied to hyperspectral data for the retrieval of surface reflectance. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(3):194–203.

[257] Staenz, K. and Williams, D. J. (1997). Retrieval of surface reflectance from hyperspectral data using a look-up table approach. *Canadian Journal of Remote Sensing*, 23(4):354–368.

[258] Stroppiana, D., Fava, F., Boschetti, M., and Brivio, P. (2011). Estimation of Nitrogen Content in Crops and Pastures Using Hyperspectral Vegetation Indices. *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS '09. First Workshop on*, pages 245–262.

[259] Su, D., Bender, A., and Sukkarieh, S. (2018). Improved Cross-Ratio Invariant-Based Intrinsic Calibration of A Hyperspectral Line-Scan Camera. *Sensors*, 18(6):1885.

[260] Sun, B., Zhu, J., Yang, L., Yang, S., and Niu, Z. (2016). Calibration of line-scan cameras for precision measurement. *Applied optics*, 55(25):6836–6843.

[261] Sun, S., Zhong, P., Xiao, H., and Wang, R. (2015). An MRF model-based active learning framework for the spectral-spatial classification of hyperspectral imagery. *IEEE Journal of Selected Topics in Signal Processing*, 9(6):1074–1088.

[262] Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R., and Kooistra, L. (2014). A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles. *Remote Sensing*, 6(11):11013–11030.

[263] Suzuki, Y., Okamoto, H., and Kataoka, T. (2008). Image Segmentation between Crop and Weed using Hyperspectral Imaging for Weed Detection in Soybean Field. *Environment Control in Biology*, 46(3):163–173.

[264] Suzuki, Y., Okamoto, H., Takahashi, M., Kataoka, T., and Shibata, Y. (2012). Mapping the spatial distribution of botanical composition and herbage mass in pastures using hyperspectral imaging. *Grassland Science*, 58(1):1–7.

[265] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions(Inception, GoogLeNet). In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 07-12-June, pages 1–9. Cvpr.

[266] Tang, L., Peng, S., Bi, Y., Shan, P., and Hu, X. (2014). A new method combining LDA and PLS for dimension reduction. *PLoS ONE*, 9(5):1–10.

[267] Tenenbaum, J. B., de Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science (New York, N.Y.)*, 290(5500):2319–2323.

[268] Thenkabail, P. S., Lyon, J. G., and Huete, A. (2012). *Hyperspectral remote sensing of vegetation*. CRC Press.

[269] Thompson, D. R., Roberts, D. A., Gao, B. C., Green, R. O., Guild, L., Hayashi, K., Kudela, R., and Palacios, S. (2016). Atmospheric correction with the Bayesian empirical line. *Optics express*, 24(3):2134–2144.

[270] Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso Robert Tibshirani. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

[271] Tong, X., Xie, H., and Weng, Q. (2014). Urban land cover classification with airborne hyperspectral data: What features to use? *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(10):3998–4009.

[272] Transon, J., D'Andrimont, R., Maugnard, A., and Defourny, P. (2018). Survey of Hyperspectral Earth Observation Applications from Space in the Sentinel-2 Context. *Remote Sensing*, 10(2):157.

[273] Trierscheid, M., Pellenz, J., Paulus, D., and Balthasar, D. (2008). Hyperspectral Imaging or Victim Detection with Rescue Robots. In *Safety, Security and Rescue Robotics, 2008. SSRR 2008. IEEE International Workshop on*, pages 7–12. IEEE.

[274] Tuo, H. and Liu, Y. (2005). A new coarse-to-fine rectification algorithm for airborne push-broom hyperspectral images. *Pattern recognition letters*, 26(11):1782–1791.

[275] Tuominen, S., Näsi, R., Honkavaara, E., Balazs, A., Hakala, T., Viljanen, N., Pölönen, I., Saari, H., and Reinikainen, J. (2017). Tree species recognition in species rich area using UAV-borne hyperspectral imagery and stereo-photogrammetric point cloud. *International archives of the photogrammetry, remote sensing and spatial information sciences; Volume XLII-3/W3*.

[276] Uezato, T., Murphy, R. J., Melkumyan, A., and Chlingaryan, A. (2016). A novel spectral unmixing method incorporating spectral variability within endmember classes. *IEEE Transactions on Geoscience and Remote Sensing*, 54(5):2812–2831.

[277] Underwood, J., Wendel, A., Schofield, B., McMurray, L., and Kimber, R. (2017). Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. *Journal of Field Robotics*, 34(6):1061–1083.

[278] Ungar, S. G. (2002). Overview of the Earth Observing One (EO-1) Mission. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 1, pages 568–571 vol.1.

[279] Uno, Y., Prasher, S. O., Lacroix, R., Goel, P. K., Karimi, Y., Viau, A., and Patel, R. M. (2005). Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. *Computers and electronics in agriculture*, 47(2):149–161.

[280] Uto, K., Seki, H., Saito, G., and Kosugi, Y. (2013). Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 6(2):851–860.

[281] Vaarala, T., Aikio, M., and Keraenen, H. (1997). Advanced prism-grating-prism imaging spectrograph in online industrial applications. In *New Image Processing Techniques and Applications: Algorithms, Methods, and Components II*, volume 3101, pages 322–331. International Society for Optics and Photonics.

[282] Vermote, E. F., Tanré, D., Deuzé, J. L., Herman, M., and Morcrette, J. J. (1997). Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):675–686.

[283] Vigneau, N., Ecarnot, M., Rabatel, G., and Roumet, P. (2011). Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in Wheat. *Field Crops Research*, 122(1):25–31.

[284] Vrindts, E., De Baerdemaeker, J., and Ramon, H. (2002). Weed detection using canopy reflection. *Precision Agriculture*, 3(1):63–80.

[285] Wang, N., Zhang, N., Dowell, F. E., Sun, Y., and Peterson, D. E. (2001). Design of an optical weed sensor usingplant spectral characteristics. *Transactions of the ASAE*, 44(2):409.

[286] Wang, X. R., Kumar, S., Kaupp, T., Upcroft, B., and Durrant-Whyte, H. (2005). Applying ISOMAP to the learning of hyperspectral image. In *Proceedings*

of the 2005 Australasian Conference on Robotics and Automation. Australian Robotics and Automation Association.

[287] Wendel, A. and Underwood, J. (2016). Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5128–5135. IEEE.

[288] Wendel, A. and Underwood, J. (2017a). Extrinsic parameter calibration for line scanning cameras on ground vehicles with navigation systems using a calibration pattern. *Sensors (Switzerland)*, 17(11).

[289] Wendel, A. and Underwood, J. (2017b). Illumination compensation in ground based hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129:162–178.

[290] Wendel, A., Underwood, J., and Walsh, K. (2018). Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. *Computers and Electronics in Agriculture*, 155:298–313.

[291] West, T., Prasad, S., and Bruce, L. M. (2007). Multiclassifiers and decision fusion in the wavelet domain for exploitation of hyperspectral data. In *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, pages 4850–4853. IEEE.

[292] Windrim, L., Melkumyan, A., Murphy, R., Chlingaryan, A., and Nieto, J. (2016a). Unsupervised feature learning for illumination robustness. In *Proceedings - International Conference on Image Processing, ICIP*, volume 2016-Augus, pages 4453–4457.

[293] Windrim, L., Melkumyan, A., Murphy, R. J., Chlingaryan, A., and Ramakrishnan, R. (2018a). Pretraining for hyperspectral convolutional neural network classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(5):2798–2810.

[294] Windrim, L., Ramakrishnan, R., Melkumyan, A., and Murphy, R. (2016b). Hyperspectral CNN classification with limited training samples. *arXiv preprint arXiv:1611.09007*.

[295] Windrim, L., Ramakrishnan, R., Melkumyan, A., and Murphy, R. J. (2018b). A Physics-Based Deep Learning Approach to Shadow Invariant Representations of Hyperspectral Images. *IEEE Transactions on Image Processing*, 27(2):665–677.

[296] Wohlfeil, J. (2012). *Optical orientation determination for airborne and spaceborne line cameras*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II.

[297] Wohlfeil, J. and Bucher, T. (2009). A modular, interactive software-concept for radiometric and geometric correction of airborne and spaceborne linescanner images. In *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IX*, volume 7478, page 74781E. International Society for Optics and Photonics.

[298] Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. In *Chemometrics and Intelligent Laboratory Systems*, volume 58, pages 109–130.

[299] Wu, C., Han, X., Niu, Z., and Dong, J. (2010). An evaluation of EO-1 hyperspectral Hyperion data for chlorophyll content and leaf area index estimation. *International Journal of Remote Sensing*, 31(4):1079–1086.

[300] Wu, C., Niu, Z., Tang, Q., and Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *agricultural and forest meteorology*, 148(8):1230–1241.

[301] Xing, J. and De Baerdemaeker, J. (2005). Bruise detection on 'Jonagold'apples using hyperspectral imaging. *Postharvest Biology and Technology*, 37(2):152–162.

[302] Yang, C., Everitt, J. H., and Bradford, J. M. (2007). Airborne hyperspectral imagery and linear spectral unmixing for mapping variation in crop yield. *Precision Agriculture*, 8(6):279–296.

[303] Yang, C., Everitt, J. H., and Bradford, J. M. (2008). Yield estimation from hyperspectral imagery using spectral angle mapper (SAM). *Transactions of the ASABE*, 51(2):729–737.

[304] Yang, C., Everitt, J. H., Bradford, J. M., and Murden, D. (2004). Airborne hyperspectral imagery and yield monitor data for mapping cotton yield variability. *Precision Agriculture*, 5(5):445–461.

[305] Yao, H. and Lewis, D. (2010). Spectral Preprocessing and Calibration Techniques. In Sun, D.-W., editor, *Hyperspectral Imaging for Food Quality Analysis and Control*, pages 45–78. Academic Press, San Diego.

[306] Yao, M., Zhao, Z., and Xu, B. (2014). Geometric calibration of line-scan camera using a planar pattern. *Journal of Electronic Imaging*, 23(1):13028.

[307] Ye, X., Sakai, K., Garciano, L. O., Asada, S.-I., and Sasao, A. (2006). Estimation of citrus yield from airborne hyperspectral images using a neural network model. *Ecological Modelling*, 198(3-4):426–432.

[308] Ye, X., Sakai, K., Sasao, A., and Asada, S.-i. (2008). Potential of airborne hyperspectral imagery to estimate fruit yield in citrus. *Chemometrics and Intelligent Laboratory Systems*, 90(2):132–144.

[309] Yeh, C.-K. and Tsai, V. J. D. (2011). Self-calibrated direct geo-referencing of airborne pushbroom hyperspectral images. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 2881–2883. IEEE.

[310] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328.

[311] Yuan, Y., Lin, J., and Wang, Q. (2016). Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization. *IEEE transactions on cybernetics*, 46(12):2966–2977.

[312] Yuen, P. W. T. and Richardson, M. (2010). An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition. *The Imaging Science Journal*, 58(5):241–253.

[313] Zarco-Tejada, P. J., Guillén-Climent, M. L., Hernández-Clemente, R., Catalina, A., González, M., and Martín, P. (2013). Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV). *Agricultural and forest meteorology*, 171:281–294.

[314] Zarco-Tejada, P. J., Hornero, A., Hernández-Clemente, R., and Beck, P. S. A. (2018). Understanding the temporal dimension of the red-edge spectral region for forest decline detection using high-resolution hyperspectral and Sentinel-2a imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 137:134.

[315] Zarco-Tejada, P. J., Ustin, S. L., and Whiting, M. L. (2005). Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. *Agronomy Journal*, 97(3):641–653.

[316] Zhang, A., Hu, S., Meng, X., Yang, L., and Li, H. (2015). Toward High Altitude Airship Ground-Based Boresight Calibration of Hyperspectral Pushbroom Imaging Sensors. *Remote Sensing*, 7(12):17297–17311.

[317] Zhang, J., Pu, R., Loraamm, R. W., Yang, G., Wang, J., and Others (2014). Comparison between wavelet spectral features and conventional spectral features in detecting yellow rust for winter wheat. *Computers and Electronics in Agriculture*, 100:79–87.

[318] Zhang, L., Li, D., Tong, Q., and Zheng, L. (1998). Study of the spectral mixture model of soil and vegetation in PoYang Lake area, China. *International Journal of Remote Sensing*, 19(11):2077–2084.

[319] Zhang, M., Qin, Z., Liu, X., and Ustin, S. L. (2003). Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral

remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 4(4):295–310.

[320] Zhang, Y. (2011). *Hyperspectral vision-based machine learning for robust plant recognition in autonomous weed control*. PhD thesis, University of California, Davis.

[321] Zhang, Y., Slaughter, D. C., and Staab, E. S. (2012a). Robust hyperspectral vision-based classification for multi-season weed mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, 69:65–73.

[322] Zhang, Y., Staab, E. S., Slaughter, D. C., Giles, D. K., and Downey, D. (2012b). Automated weed control in organic row crops using hyperspectral species identification and thermal micro-dosing. *Crop Protection*, 41:96–105.

[323] Zhang, Z. (2002). A Flexible New Technique for Camera Calibration (Technical Report). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334.

[324] Zhao, D., Huang, L., Li, J., and Qi, J. (2007). A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(1):25–33.

[325] Zhong, P. and Gong, Z. (2017). A Hybrid DBN and CRF Model for Spectral-Spatial Classification of Hyperspectral Images. *Statistics, Optimization & Information Computing*, 5(2):75.

[326] Zhong, P. and Wang, R. (2011). Modeling and classifying hyperspectral imagery by CRFs with sparse higher order potentials. *IEEE Transactions on Geoscience and Remote Sensing*, 49(2):688–705.

[327] Zhong, Y., Wang, X., Xu, Y., Jia, T., Cui, S., Wei, L., Ma, A., and Zhang, L. (2017a). MINI-UAV borne hyperspectral remote sensing: A review. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 5908–5911.

[328] Zhong, Z., Li, J., Ma, L., Jiang, H., and Zhao, H. (2017b). Deep residual networks for hyperspectral image classification. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1824–1827.