UNIVERSITY OF STRATHCLYDE

Techniques for capture and analysis of hyperspectral data

by

Timothy George Harold Kelman

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

Centre for excellence in Signal & Image Processing
Department of Electronic and Electrical Engineering

December 2015

Declaration of Authorship

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: _____

Date: _____

Abstract

The work presented in this thesis focusses on new techniques for capture and analysis of hyperspectral data. Due to the three-dimensional nature of hyperspectral data, image acquisition often requires some form of movement of either the object or the detector. This thesis presents a novel technique which utilises a rotational line-scan rather than a linear line-scan. Furthermore, a method for automatically calibrating this system using a calibration object is described. Compared with traditional linear scanning systems, the performance is shown to be high enough that a rotational scanning system is a viable alternative.

Classification is an important tool in hyperspectral image analysis. In this thesis, five different classification techniques are explained before they are tested on a classification problem; the classification of five different kinds of Chinese tea leaves. The process from capture to pre-processing to classification and post-processing is described. The effects of altering the parameters of the classifiers and the pre and post-processing steps are also evaluated.

This thesis documents the analysis of baked sponges using hyperspectral imaging. By comparing hyperspectral images of sponges of varying ages with the results of an expert tasting panel, a strong correlation is shown between the hyperspectral data and human determined taste, texture and appearance scores. This data is then used to show the distribution of moisture content throughout a sponge image.

While hyperspectral imaging provides significantly more data than a conventional imaging system, the benefits offered by this extra data are not always clear. A quantitative analysis of hyperspectral imaging versus conventional imaging is performed using a rice grain classification problem where spatial, spectral and colour information is compared.

Acknowledgements

First and foremost, I wish to thank Professor Stephen Marshall for supervising me over the past five years. From inspiring me during an undergraduate class which he lectured right through to the viva, he has been a great help. Whenever I was feeling anxious about my research, Steve was always available for a meeting, and every single time I felt better after speaking to him than I did before. Although the project took me longer to complete than planned, Steve secured several contracts for me to continue as a researcher. I doubt I would have ever been able to complete my Ph.D had I not been able to remain in the department.

My second supervisor, Dr. Jinchang Ren, had a huge impact on my research. His expertise in the mathematical side of statistical classification was invaluable. He always made himself available to help and gave me the push I often needed to get my research published.

A third major influence was my friend and colleague, Dr. Paul Murray. From helping me with my first conference paper, to securing funding to continue my research, I have always been able to rely on him for both his technical advice and his friendship.

I wish to thank all the staff in CeSIP, specifically Christine Bryson, for their support - administrative, technical and social.

Gilden Photonics were part funders of my Ph.D and I wish to thank them for introducing my to hyperspectral imaging and the support they provided. All the staff members I worked with were friendly, knowledgeable and keen to help wherever they could.

I would like to thank my two examiners, Professor Stephen Morgan and Dr. Vladimir Stankovic for their helpful suggestions which significantly improved the quality of this thesis.

Finally, I offer my thanks to my family and friends for their support. They never seemed to doubt me, even when I did.

Contents

List of Figures

List of Tables

Abbreviations

ANN	artificial neural network
AOTF	acousto-optic tunable filter
BIL	band-interleaved-by-line
BIP	band-interleaved-by-pixel
BSQ	band-sequential
CCD	charge-coupled device
CMOS	complementary metal-oxide-semiconductor
DAU	data acquisition unit
HDD	hard disk drive
HO	horizontal offset
HSI	hyperspectral imaging
HgCdTe/MCT	mercury cadmium tellurium
InGaAs	indium gallium arsenide
LCTF	liquid crystal tunable filter
MDC	minimum distance classifier
MLC	maximum likelihood classifier
MLPP	multi-layer perceptron
MSE	mean squared error
NIR	near infra-red
PCA	principal component analysis
PTFE	polytetrafluoroethylene
RBF	radial basis function
RGB	red green blue
SNR	signal-to-noise ratio
SSD	solid state drive

SSIM	structural similarity index
SVM	support vector machine
SWIR	short wave infra-red
Si	silicon
USB	universal serial bus
VIS	visible
VNIR	visible near infra-red
VO	vertical offset

Chapter 1

Introduction

Hyperspectral imaging provides both spatial and spectral information about a scene. Each spatial location in a hyperspectral image contains a contiguous portion of the electromagnetic spectrum, usually with some part extending outside the human visible range. This wealth of data gives substantially more information about a scene compared to conventional digital imaging, making it a useful tool in a number of areas such as food and drink quality assessment [1], forensic science [2], and in the military [3].

The number of spectral bands captured varies, but tends to be in the region of 100s of bands [1]. As a result, hyperspectral images are much larger than their standard greyscale or colour imaging counterparts. This wealth of data provides valuable information, but the large quantities can cause problems during capture and subsequent analysis. One aim of this thesis is therefore to provide new techniques which aid in the capture and analysis of hyperspectral data.

The other main motivation for this work was that it was felt that was a need to demonstrate the potential of hyperspectral imaging to non-expert users who may be unaware or unconvinced that it could be potential solution to a problem they have. Originally, hyperspectral imaging systems were bulky and expensive, so mainly used in a military context. The price and size has come down as the technology improves, and therefore the savings required from implementing a system are reduced, meaning hyperspectral imaging can realise its potential in a wider range of applications. For a business to invest in a hyperspectral imaging system, it must be convinced of its worth. Any processing and analysis required may be beyond the capabilities of those in the business, and the

camera manufacturers may also lack the will or ability to help. By collaborating with system manufacturers and end users this gap can be bridged.

This thesis aims to determine the capability of hyperspectral imaging through a classification problem using Chinese tea leaves, shelf-life prediction of baked sponges and a quantitative analysis of any improvements offered over conventional digital imaging using classification of four different types of rice as an example. A novel technique for data capture using rotational, rather than linear, movement is also proposed.

1.1 Organisation of Thesis

The remainder of this thesis is organised as follows:

- **Chapter 2:** This chapter is split into three sections. The first provides a background to the different hyperspectral imaging technologies, imaging methodologies and wavelength ranges available. The second part looks at some of the current signal processing techniques used in hyperspectral image analysis specifically focussing on dimensionality reduction. Finally, an overview is given of the main applications areas where hyperspectral imaging technology is utilised.

- **Chapter 3:** In this chapter, a novel alternative to the linear line-scan normally used in pushbroom acquisition is presented. The key difference is in the use of rotational motion rather than linear motion. Two offsets associated with a rotational line-scan are explained and an algorithm which compensates for the errors they cause is proposed. A method for determining these offsets using a calibration object is explained before a quantitative comparison is made with images captured using a regular line scan system.

- **Chapter 4:** Classification is a common process to implement after capture of hyperspectral data. In this chapter, five of the most popular classification techniques are explained in detail. Each of these classifiers is then demonstrated in a classification problem; the classification of five different types of Chinese tea. The process of capturing and manipulating the data into a suitable format is described. Results are then produced showing the performance of each classifier and the effects of changing the most important parameters.

- **Chapter 5:** Chapter 5 focusses on the analysis of baked sponges, specifically how their eating quality changes with time. An experiment that utilises hyperspectral imaging and human tasting is described. It is demonstrated that the moisture content appears to reduce as the sponges age showing correlation with the also decreasing eating quality. The spatial distribution of moisture throughout each sponge is then estimated using a maximum likelihood classifier. Finally, the potential benefits of implementing a hyperspectral imaging system on a bakery production line are discussed.

- **Chapter 6:** This chapter presents a quantitative analysis of the benefits offered by hyperspectral imaging compared to conventional greyscale or colour digital imaging. A technique for reducing the storage and processing requirements for hyperspectral images with many unwanted background pixels is proposed. Spatial features are combined with the hyperspectral data allowing five subsets from the data to be created: spatial data, colour data, spectral data, spatial + colour data, and spectral + spatial data. Classification is performed using a support vector machine with each of these datasets, giving quantitative results showing the improvement offered by hyperspectral imaging over conventional techniques.

- **Chapter 7:** In Chapter 7, the main results from this thesis are summarised along with discussion about their impact and the direction any subsequent research could take.

1.2 Original Contributions

It is believed that the following novel contributions have been made:

- In Chapter 3, a technique is proposed for reconstructing a rotational linescan back into Cartesian space based on two offsets inherent within a rotational scan. Furthermore, two methods for determining theses offsets based on a calibration object are presented. The first method requires a perfect alignment of the calibration object while the second method relaxes this alignment requirement.

- Chapter 4 presents a methodology to image, correct, crop, pre-process, classify, and spatially filter samples of Chinese tea using a hyperspectral system.

- In Chapter 5 hyperspectral data of a baked sponge is shown to correlate with results from a human tasting panel. The distribution of moisture throughout a sponge is also estimated using a maximum likelihood classifier.

- Chapter 6 documents a technique for the quantitative analysis of spectral, spatial and colour information when used for classification, specifically in the case of rice grains. A sparse representation technique is also proposed, drastically reducing the storage and processing requirements of the data.

Chapter 2

Background

2.1 Introduction

As explained in the previous chapter, hyperspectral imaging is an imaging technique which can provide both spatial and spectral information about a scene simultaneously. The term hyperspectral imaging covers a broad range of technologies, imaging methodologies and wavelength ranges. Whilst detailing each of these individually is beyond the scope of this thesis, the first part of this chapter aims to explain the principle behind each along with some associated advantages and disadvantages.

The second part of this chapter looks at a common signal processing technique known as dimensionality reduction. Multiple detailed examples will be used to demonstrate how it can be implemented and the benefits it can offer.

The final section in this chapter will give an overview of the main applications areas that take advantage of hyperspectral imaging, provide some specific examples and justify the choice of the topics chosen for the rest of this thesis.

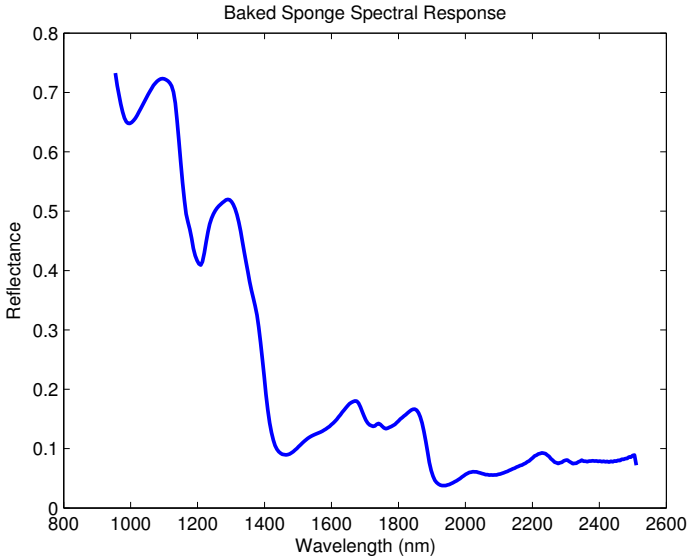


Figure 2.1: Spectral response of a baked sponge

2.2 Technologies, Imaging Methodologies and Wavelength Ranges

2.2.1 Imaging, Spectroscopy and Hyperspectral Imaging

Hyperspectral imaging can be thought of as a combination of spectroscopy and conventional digital imaging. This section will briefly explain both of these before showing how their principles can be combined to create hyperspectral imaging. The term 'band' will be used throughout in different contexts, but it will always refer to a single channel of a multichannel dataset.

Spectroscopy is the study of matter emitting or interacting with electromagnetic radiation. This information can be represented by a plot of the response against wavelength, known as a spectrum. Normally, only a single spectrum is produced with each measurement, i.e. there is no spatial information associated with the signal. This is an average of all the spectral information for a specific area. Figure 2.1 shows a spectroscopic measurement of a baked sponge. The y-axis here refers to the intensity of the reflected light measured by the sensor as a proportion of the intensity that would be reflected by a white calibration tile with near-perfect reflectance.

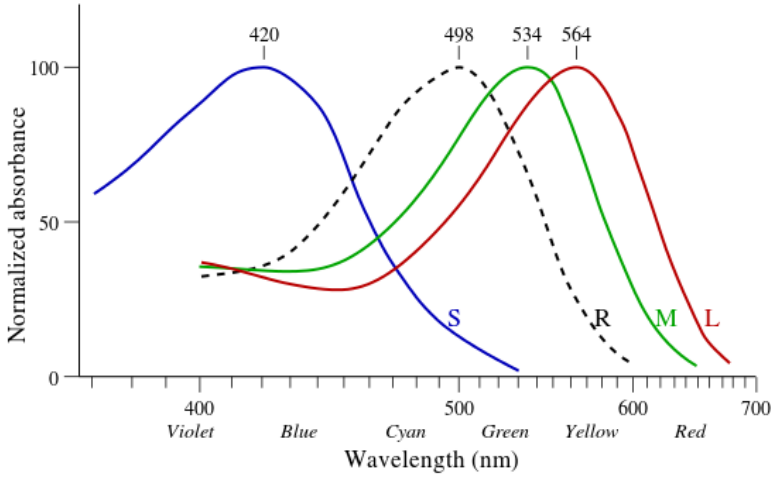


Figure 2.2: Cone Responses. S, M, and L refer to short, medium, and long, respectively. The dashed line (R) shows the response of rods, more suited for night vision [4].

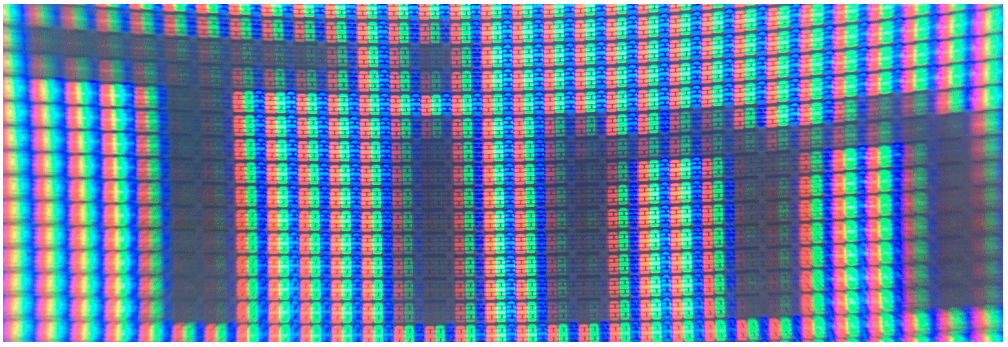


Figure 2.3: Close up view of pixels on an RGB display

Conventional digital imaging on the other hand, captures reflected light from multiple locations in a scene, building a spatial image. Each of these spatial locations corresponds to a single pixel on the sensor. In a greyscale image, each pixel has a single value corresponding to the intensity of the captured light, i.e., there is no wavelength information. In colour imaging, each pixel is represented by three components; red, green and blue (RGB). Each component is a weighted sum of light from a specific region of the electromagnetic spectrum, corresponding to the three types of receptor cones in the human eye. Figure 2.2 shows the absorbance at different wavelengths of light for each type of receptor cone. Note that although colour information is contained within an RGB image, there is no information about specific wavelengths, only weighted sums of them. Figure 2.3 shows an example of pixels on an RGB display.

Hyperspectral imaging combines spectroscopy and conventional imaging to produce a spectrum at each pixel in the image. If you consider a greyscale image to be single band and an RGB image to have three bands, then a hyperspectral image may have a

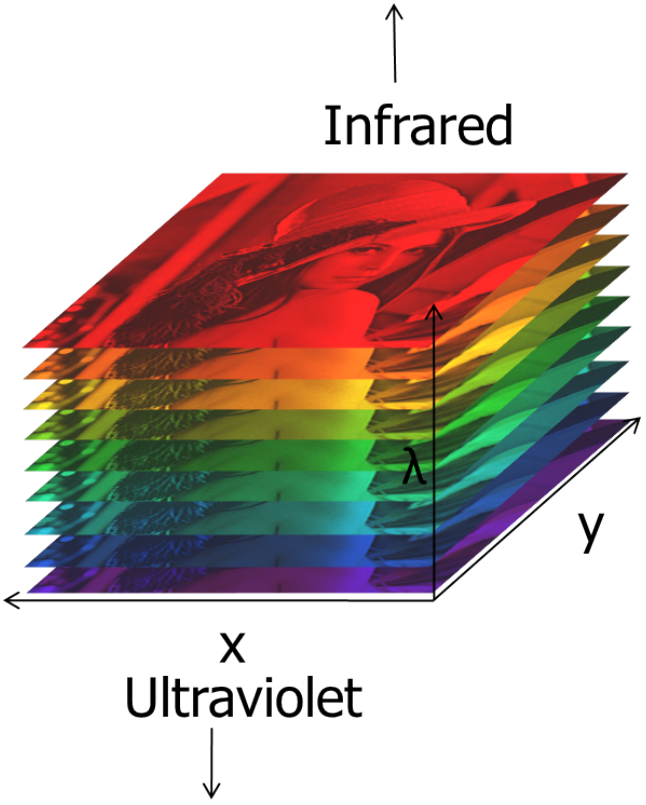


Figure 2.4: Hyperspectral image consisting of images at different bands

hundred bands or even a thousand bands. (The actual number of bands required for an image to be 'hyperspectral' is discussed in the following section.) Figure 2.4 shows how a hyperspectral image is made up of many bands.

2.2.2 Spectral Imaging Classes

For an image to be considered 'hyperspectral', it generally has to consist of more than 20 bands. Any fewer than this, and it would be considered a 'multispectral' image. There is no strict definition, however, so often data containing fewer than 20 bands is referred to as hyperspectral, especially in a marketing context. More than several thousand bands is sometimes referred to as 'ultraspectral'. All three of these terms are often irrelevant as they are not related spectral resolution. For example, a 'multispectral' image consisting of 20 bands in the range 400 nm to 600 nm has a spectral resolution of 10 nm, whereas a 'hyperspectral' image consisting of 100 bands in the wavelength range 1000 nm to 2500 nm has an inferior spectral resolution of 17 nm. Which of these is better could only be decided based on the application.

2.2.3 Hyperspectral image data formats and display methods

Since hyperspectral images are three-dimensional (two spatial dimensions, one spectral dimension), the data is in the form of a cube. In this thesis, a hyperspectral image is referred to as a hypercube, although in other literature they may be called a hyperspectral cube, spectral cube, spectral volume, datacube or data volume. The two spatial dimensions are referred to as x and y and the spectral dimension is referred to as λ (wavelength). The three dimensions allow the data to be displayed in three separate ways. Using the MATLAB notation where a colon (:) represents all values in that dimension [5] and a hypercube, I, with dimensions (x, y, λ), we can represent the data as follows:

$I(:, :, \lambda)$ would return every pixel at a specific λ. This would produce a greyscale image of the scene. For example, Figure 2.5(a) shows a hyperspectral image of sections of a yellow pepper (left), a red pepper (middle) and a green pepper (right) at a wavelength of 690 nm. Since 690 nm corresponds to a red wavelength, the red pepper appears brighter than the other two. A colour representation of a hypercube can be created in what is known as 'false colour', by simply selecting three bands corresponding to red, green and blue. Figure 2.5(b) shows a false colour image of the peppers with the red, green and blue wavelengths roughly corresponding to the peaks shown in Figure 2.2. As expected, the pepper in the centre is indeed red. Since this image has been created using only three bands, it looks unnatural when compared to the viewing the object in real life. A true colour image generated using a spectral sensitivity function is shown in Figure 2.5(c) giving a more realistic representation of the scene. To achieve the most accurate representation of the scene, a spectral sensitivity function of the camera and spectrograph used would be required. Since this was not available, the spectral sensitivity function of the Nikon D70 camera was used, shown in Figure 2.5(d). More information about the use of spectral sensitivity functions to create colour images can be found in [6].

$I(x, y, :)$ would return every λ value for a specific (x, y) location in the hypercube. This would produce a spectrum of a single point in the scene. Figure 2.6(a) shows the reflectance spectrum at three points in the three peppers hypercube. The locations of these points is indicated in Figure 2.6(b). From this reflectance information, we can see how the colours of the red and green peppers correspond to higher reflectance

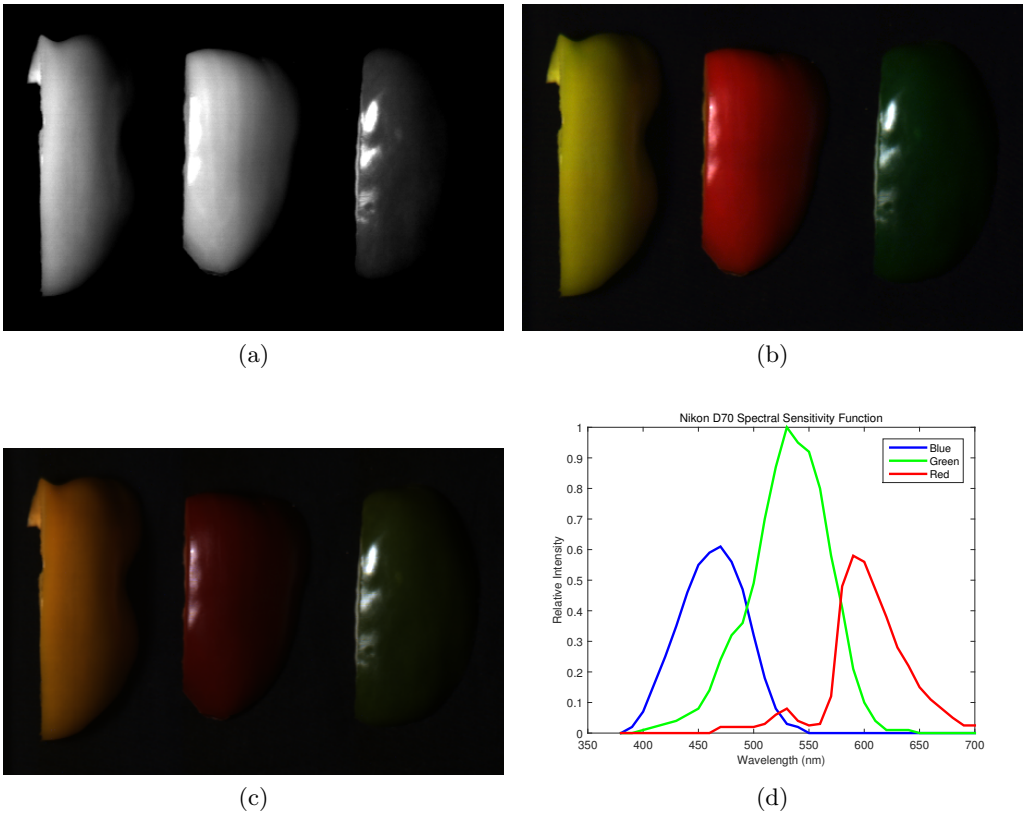
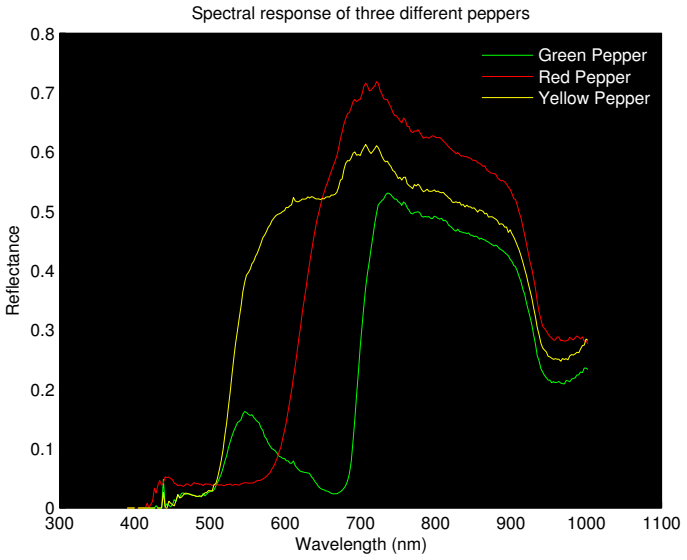


Figure 2.5: Sections of a yellow pepper (left), a red pepper (middle) and a green pepper (right) (a) At 690 nm (b) In false colour generated from three spectral bands (c) In true colour generated from a spectral sensitivity function (d)Spectral sensitivity function of the Nikon D70 [7].

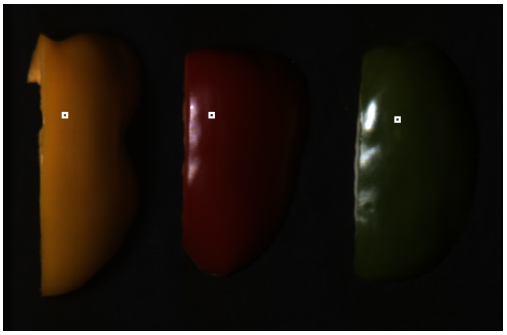
values at their corresponding wavelengths. The yellow pepper's colour is due to the high reflectance at green and red wavelengths.

$I(:, y, :)$ would return every value at a specific y location. Since only y is specified and not x, this is a plane rather than a single point. If $I(:, :, \lambda)$ can be considered as a horizontal slice then $I(:, y, :)$ can be considered a vertical slice. Figure 2.7(a) shows a slice through the three peppers hypercube. Figure 2.7(b) shows the spatial location of this line in the hypercube. It is difficult to extract any meaningful information from this view, so it is rarely used to visualise the data. By comparing Figure 2.6(a) and Figure 2.7(a), however, similarities are evident: Figure 2.7(a) is effectively a 'top down' view of data very similar to Figure 2.6(a). This format demonstrates how the pushbroom acquisition technique works, as explained in Section 2.2.4.

$I(x, :, :)$ produces a plane orthogonal to $I(:, y, :)$ and is not different in any practical way after acquisition since the order of x and y is arbitrary.



(a)



(b)

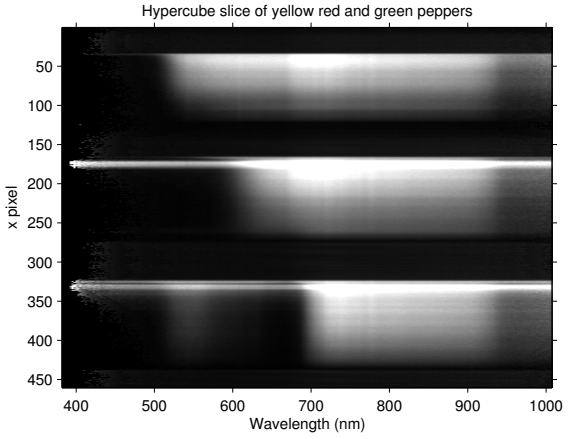
Figure 2.6: Spectral responses of different types of pepper (a) Plotted spectra (b) Locations of plotted spectra

Table 2.1: Interleave Methods

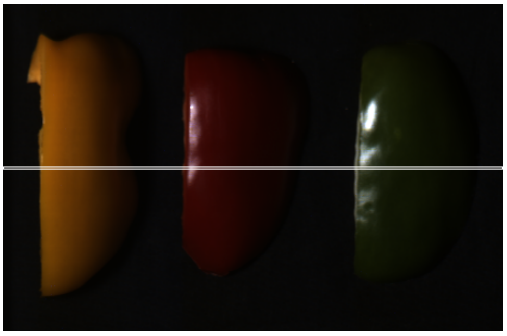
Method	**Example**
Band-Interleaved-by-Line	AAAAABBBBBCCCCCAAAAA...
Band-Interleaved-by-Pixel	ABCABCABCABCABCABCAB...
Band-Sequential	AAAAAAAAAAAAAAABBBBB...

2.2.3.1 Interleave Methods

As well as several methods to visualise a hypercube, there are several ways to store one. These are known as Band-Interleaved-by-Pixel (BIP), Band-Sequential (BSQ) and Band-Interleaved-by-Line (BIL). Using data shown in Figure 2.8, a synthetic hypercube containing letters rather than values, Table 2.1 shows the difference in how each value is stored in a hypercube file.



(a)



(b)

Figure 2.7: Slice of hypercube containing yellow, red and green peppers (a) All values at all wavelengths at a specific y location (b) Location of line slice in hypercube

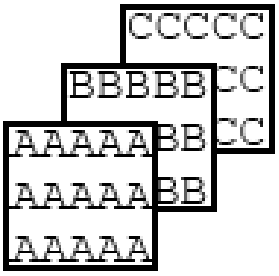


Figure 2.8: Synthetic hypercube made up of letters. Each letter corresponds to a different band

Each of these methods is suited to a different display method. BIP stores each pixel at all bands one after the other, so all values at $I(x, y, :)$ would be next to each other. BSQ stores each band separately so all values at $I(:, :, \lambda)$ would be next to each other and BIL stores each line at all wavelengths separately so all values at $I(:, y, \lambda)$ would be next to each other. Using a display method with its corresponding interleave method results in much faster reading and writing of the file as the file pointer does not need to skip around the file after each value has been read.

2.2.4 Image Acquisition Techniques

2.2.4.1 Area Scan

An area scan is where the entire scene is captured at a specific wavelength band, then the wavelength band is changed and the entire scene is captured again. This is repeated for all bands required, building up the hypercube. Figure 2.9(a) shows how the changing the wavelength band and recapturing the scene builds up the hypercube. The most common way to capture data in this format is to use a sensor sensitive to the whole wavelength range required and placing different band-pass filters in front of it. Each filter used corresponds to a wavelength band in the produced hypercube. Changing the filter is often achieved using a filter wheel, where the wheel spins a set amount after each band has been captured. Another form of area scan imaging uses a tunable filter where an electric current changes the band-pass capability of the filter. Similar to the filter wheel, the filter is tuned to a different wavelength band after the previous band has been captured.

Since all spatial information at a specific wavelength band is captured simultaneously, BSQ is used as the data is captured in the correct order for this format.

Advantages

The main advantage of an area scan is the simplicity, both in terms of understanding and implementation when a low number of wavelength bands are required. Since each wavelength band is captured separately, the exposure time for the sensor can be adjusted independently. Also, unlike some of the other methods described below, both the sample and the detector can remain stationary.

Disadvantages

Poor spectral resolution and a long acquisition time are the main disadvantages of the area scanning technique and improving one of these will tend to be at the expense of the other. The spectral resolution is poor because a discrete filter (or in the case of a tunable filter, a discrete setting) is required for each wavelength band. The acquisition time is long because there is an exposure time for each wavelength band captured as well as the time taken to change the filter. Adding more filters can increase the spectral resolution but will increase the acquisition time and vice-versa.

2.2.4.2 Whiskbroom

In whiskbroom acquisition, all wavelength bands are captured for a single spatial point in the scene. Then, either the detector or the object is moved and another spectrum is acquired. This is repeated for all spatial points in the scene. Figure 2.9(b) shows how a hypercube is generated using this technique. BIP is the method used to store the captured data as all the information for each pixel is captured one after the other. The optics in a whiskbroom scan are simpler than the alternatives, as no spatial information is captured. Two dimensions of movement (of the detector, object or both) are required, however, and care must be taken to ensure that the movement is accurate, consistent and repeatable.

Advantages

The whiskbroom method is equivalent to using a spectrometer (which captures a spectrum at a single point) and scanning it across the scene in both spatial dimensions. Spectrometers can be much cheaper than the alternatives discussed in this section. This could mean that a greater wavelength range could be covered, or they could be covered at a higher spectral resolution.

Disadvantages

Having to scan in two directions has several drawbacks. Firstly, it is more time consuming as either the object or the detector has to be repositioned for every capture. If the object is not temporally stable, then the hypercube captured will not be an accurate representation of the scene. Secondly, advanced positioning methods are required to ensure that each pixel captured is registered to the right spatial location.

2.2.4.3 Pushbroom

Pushbroom is by far the most popular hyperspectral imaging acquisition technique. Instead of capturing a single pixel at all wavelength bands like the whiskbroom technique, the pushbroom technique captures an entire line of pixels at all wavelength bands simultaneously. The detector or object is then moved along the other spatial dimension. This is shown in Figure 2.9(c). This produces data in the BIL format. Each spatial line is a two-dimensional slice of the hypercube and can be captured on a two-dimensional array detector, such as a greyscale camera (assuming the camera is responsive in the all of the

spectral bands of interest). Images from this camera are then captured sequentially, like a video. While most pushbroom systems use linear motion, a major contribution in this thesis is a pushbroom system which uses rotational motion. This is explained in detail in Chapter 3.

Advantages

Although an accurate way of moving either the detector or the object is required, this movement is often naturally available. For example, objects on a conveyor belt on a production line moving underneath a hyperspectral camera or an aircraft holding a downward facing hyperspectral camera flying in a straight line over a field. The pushbroom technique combines the high spectral resolution of a whiskbroom system with the high spatial resolution of an area scan system.

Disadvantages

Since all wavelength bands are captured simultaneously, the exposure time for each wavelength band cannot be set independently. The scanning direction and speed need to be accurately set to match the exposure time and the spatial resolution of the dispersive element to ensure that square pixels are generated. If non-square pixels are produced, then the generated hypercube will need to be post-processed, requiring some interpolation and will never be an exact representation of the scene.

2.2.4.4 Snapshot

As the name implies, a snapshot system is able to capture both spatial and spectral information about a scene in a single exposure, i.e. all three dimensions captured at once, as seen in Figure 2.9(d). No scanning is involved in this process.

Advantages

The main advantage of snapshot systems is their ability to capture full hypercubes in a very short time. The exposure time is comparable with that of a single line in a pushbroom system. This allows frame-rates suitable for recording hyperspectral videos.

Disadvantages

Snapshot systems are still relatively early in development, and resolutions both spatially and spectrally are poorer than the more popular pushbroom systems.

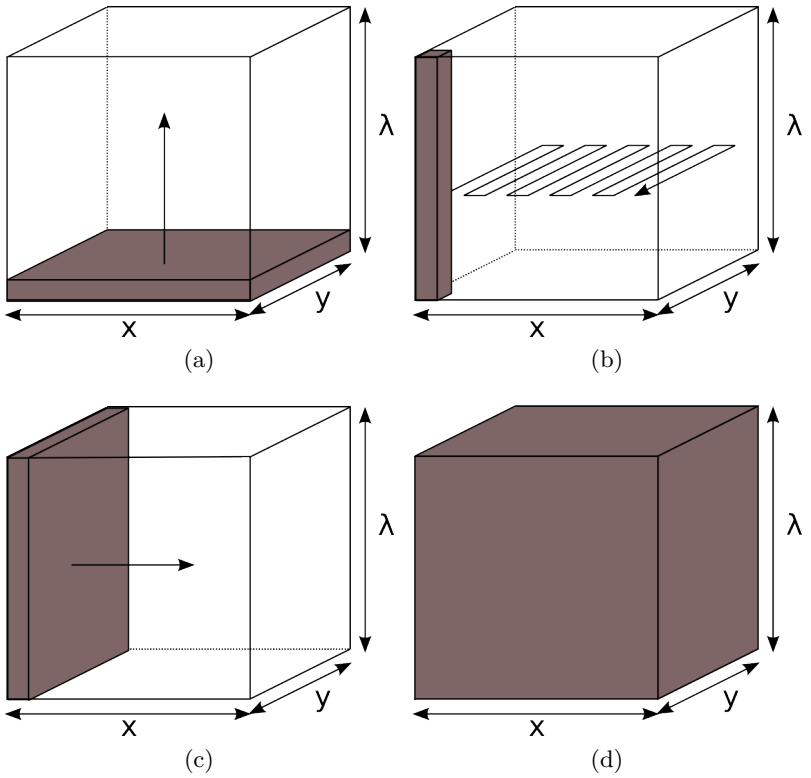


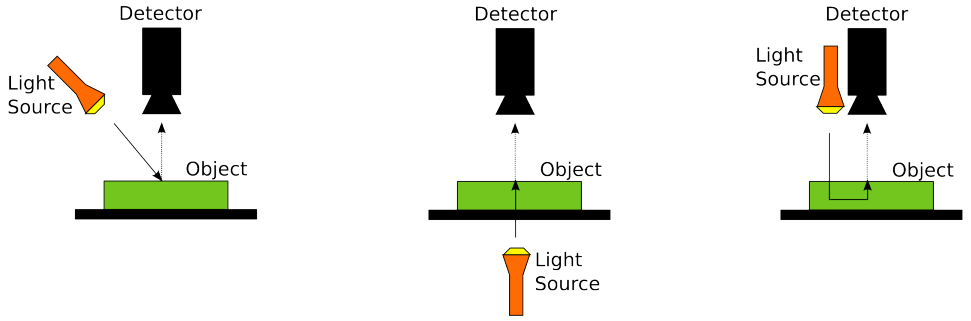
Figure 2.9: Different image acquisition techniques (a) Area Scan (b) Whiskbroom (c) Pushbroom (d) Snapshot

2.2.5 Image Sensing Modes

Reflectance, transmittance and interactance are the three different sensing modes for hyperspectral imaging. The difference between each comes down to the position of the object, detector and light source. Figure 2.10 shows the differences between these imaging modes [8].

Reflectance is the most popular hyperspectral imaging mode and all of the hypercubes in the work shown in this thesis were captured in reflectance mode. Simply, a light source illuminates an object and the hyperspectral camera records how much of this light is reflected off the object and into the detector.

In transmittance, the object is placed between the light source and the detector. The amount of light transmitted through the object reaching the detector is recorded. This allows more detailed information about the inside of the object to be captured such as internal defects. The amount of light penetrating the object, however, is often very small making measurement difficult. This is especially true for objects with low transparency.



(a) (b) (c)

Figure 2.10: Different imaging modes (a) Reflectance mode (b) Transmittance modes (c) Interactance mode

Interactance provides a compromise between reflectance and transmittance where the light source and the detector are on the same side of the object [9]. The detector measures light backscattered through the object, reducing the influence of the thickness of the object and any surface effects [8]. Interactance requires a more complicated set-up than reflectance or transmission, however, as a light seal must used to ensure that no specular reflections are picked up by the detector.

[9] provides a detailed comparison of these three methods, and concludes that interactance is the most effective method for the specific case of measuring the internal properties of kiwifruit. It should be noted, however, that this study only looked at spectroscopy techniques, not hyperspectral imaging. Interactance and transmission are harder to implement with hyperspectral imaging, so reflectance mode is still the most popular.

2.2.6 Visible, Near Infrared, Short Wave Infra-red

Much in the same way that different manufacturers have different definitions of the terms 'multispectral' and 'hyperspectral', the same is true when it comes to naming wavelength range capabilities of hyperspectral systems. Some combine the terms to describe a system which covers multiple ranges while others simply refer to the region which is most prominent in the system. In this thesis, three terms will be used: 'visible', 'near infra-red' and 'short wave infra-red'. This section will briefly explain what is meant by each of these terms.

Table 2.2: Comparison of different hyperspectral wavelength ranges and their names.

Wavelength Range	**Name**	**Abbreviation**
400 nm to 700 nm	Visible	VIS
400 nm to 1000 nm	Visible Near Infra-red	VNIR
900 nm to 1700 nm	Near Infra-red	NIR
1000 nm to 2500 nm	Short Wave Infra-red	SWIR

Visible (VIS)

The term 'visible' in this sense refers to the wavelengths of light that can be detected by the human eye (around 390 nm to 700 nm [10]). A VIS range system will be responsive at most of these wavelengths (however, the shortest wavelength detected is often only around 400 nm).

Near Infra-red (NIR)

'Near infra-red' tends to refer to the region of the spectrum with wavelengths 750 nm to 1400 nm. Note that the NIR region starts as soon as the wavelength of the light is longer than what is detectable by the human eye.

Short Wave Infra-red (SWIR) SWIR is the region after NIR that covers the range 1400 nm to 3000 nm. In practice, most hyperspectral imaging systems detect wavelengths no longer the 2500 nm.

Some example systems

Shown in Table 2.2 is a list of several different wavelength ranges and how they will be referred to in this thesis.

2.2.7 Instruments Required for Hyperspectral Imaging

Despite the different acquisition techniques, sensing modes and wavelength ranges, nearly every type of hyperspectral imaging needs the same five components: a light source, a lens, a wavelength dispersion device, an image detector and a data acquisition unit. This section will explain the purpose of each of these components and the different technologies behind them.

2.2.7.1 Light Sources

Unless the object being examined with a hyperspectral imaging system can emit its own light in some way (for example, measuring the spectral response of a new type of light bulb), a light source is required to illuminate the target. For field applications such as crop monitoring, this light source is the sun. Other light sources include halogen lamps, light emitting diodes and tunable light sources.

Halogen Lamps

For most hyperspectral imaging, a broadband light source such as a halogen lamp is required. A halogen lamp consists of a tungsten filament enclosed in a glass bulb filled with a halogen gas, usually iodine or bromine. When the filament is heated to a sufficiently high temperature by passing an electric current through it, it produces light. The spectral response of a halogen lamp is smooth and wide, covering the full range for VIS, VNIR, NIR and SWIR (less than 380 nm to greater than 2500 nm). Due to their low cost and the wide spectrum produced, halogen lamps have been used extensively for hyperspectral imaging applications [11]. They do, however, suffer from several drawbacks including shorter life time than the alternatives, high heat output and spectral shifts dues to temperature change. The high heat output is often a major factor, especially when the light source must be close to the objects. If the acquisition time is too long, then the heat produced by the halogen lamp may modify the sample in an undesired way.

Light Emitting Diodes (LEDs)

An LED is a semiconductor device that emits light when its anode has a voltage that is more positive than its cathode by at least the forward voltage drop. The wavelength of the light emitted is determined by the energy band gap of the semiconductor. First used as early as 1962 [12] LEDs have rapidly developed in recent years due to their small size, low cost, fast response time, long lifetime, low heat generation and low energy consumption. Another major advantage is that all photons emitted from an LED can be in the one direction with no energy loss. The material used in an LED determines the spectrum of the light it emits. Narrowband light at different wavelengths in the VIS/VNIR/SWIR range or broadband white light can be produced. The broadband light is particularly suited to hyperspectral imaging. Many of the drawbacks from halogen lamps are addressed by LEDs, although currently they have a much lower light intensity.

LED technology is expected to continue to improve and is expected to become the main light source for hyperspectral imaging applications in the future [8].

Tunable Light Sources

A tunable light source emits light of a specified wavelength, but this wavelength can be changed during operation. Rather than placing the dispersive element between the detector and the object like most other methods, it is combined with a broadband illumination source so that only a specific wavelength of light illuminates the object at any one time. This approach is particularly suited to area scan methods rather the whiskbroom or pushbroom. Since only narrowband light is ever illuminating the source, the intensity is usually relatively weak reducing the chance of damage being caused by heat [13].

2.2.7.2 Lenses

To ensure that a sharp image is acquired, a lens is required to focus the light onto the imaging detector, similar to how a conventional camera works. A major problem, however, is the effect of chromatic aberration. Chromatic aberration is the distortion caused by the optical power of a lens varying with wavelength [14]. This especially relevant to hyperspectral imaging as often a wide range of wavelengths is needed. A specialist lens designed for the wavelength range required must be used.

2.2.7.3 Wavelength Dispersion Devices

Every hyperspectral imaging system must contain a wavelength dispersion device of some sort. This is the part of the system that splits broadband light into individual wavelength bands. This section will explain the operation of some of the more popular wavelength dispersion devices, namely: filter wheels, imaging spectrographs, tunable filters and Fourier transform imaging spectrometers.

Filter Wheels

A filter wheel is made up of several band pass filters, each one allowing only light from a narrow band of the spectrum to pass through. An example of a filter wheel is shown in Figure 2.11. Unlike many other dispersive elements, filter wheels can be easily reconfigured to operate with different wavelength ranges (VIS,VNIR,NIR and SWIR).

Figure 2.11: A filter wheel [15]

Filter wheel systems use area scan acquisition and tend to have fewer than 20 filters in them, so are more like to be considered part of a multispectral rather than hyperspectral system. The time taken to change wavelengths is also significantly longer compared to other methods.

Imaging Spectrographs

An imaging spectrograph is a device which disperses broadband light into different wavelength components simultaneously and retains one dimension of spatial information. By scanning in a perpendicular direction to this spatial dimension, a three-dimensional hypercube can be generated used the pushbroom acquisition method, explained in more detail in Section 2.2.4.

Imaging spectrographs usually use diffraction gratings to disperse the light. A diffraction grating is an optical device which refracts light based on its wavelength. These gratings tend to be either reflective or transmissive. Reflective gratings consist of a grating superimposed on a reflective surface while transmissive gratings consist of a grating superimposed on a transparent surface.

Imaging spectrographs using a reflective grating usually require an entrance slit and two concentric spherical mirrors. The reflective grating must be aberration-corrected. After light enters through the slit, it is reflected onto the grating by one of the mirrors. The angle that the light reflects at depends on its wavelength. This dispersed light is then reflected by the other mirror out of the spectrograph where it can be recorded by an image detector.

An imaging spectrograph using a transmissive grating tends to use a prism-grating-prism (PGP) configuration [8]. After light passes through the entrance slit, it is collimated

by a front lens before being dispersed by the PGP. The back lens then projects the dispersed light out the back of the spectrograph where it can be recorded by an image detector.

Reflective gratings are thought to be most suited for hyperspectral imaging systems due to the high image quality, lack of high-order aberrations, low distortion, low f-number and large field of view [16]. Reflective optical components such as mirrors tend to be more efficient than transmissive components such as prisms, resulting in higher signal-to-noise ratios (SNRs) for hyperspectral systems using reflective grating imaging spectrographs. The main disadvantage with reflective grating systems is that some distortion is inherent while transmissive grating systems use on-axis optics, so naturally have less distortion.

Tunable Filters

Tunable filters are wavelength dispersion devices suitable for acquiring hyperspectral images using the area scan method, and share similarities with filter wheels. The difference is that an electronic signal is used to change the bandpass property of the filter, rather than to rotate a filter wheel selecting a new discrete filter. There are two main types of tunable filters: acousto-optic tunable filters (AOTFs) and liquid crystal tunable filters (LCTFs). AOTFs work by using the acousto-optic effect [17] where a birefringent crystal's optical properties (i.e. the wavelength of light it allows to pass through it) are modified by an acoustic signal. A piezoelectric transducer is used to convert an electric signal into the acoustic one required. LCTFs use a liquid crystal which is electronically controlled to only transmit a chosen wavelength band whilst blocking others. Compared to filter wheels, tunable filters are much faster at changing from one wavelength band to another since they do not have any moving parts. AOTFs are faster than LCTFs (microseconds compared to milliseconds) however, they suffer from poorer image quality. Both types of tunable filter, however, suffer from several shortcomings: high F-number leading to small light collection angle and low light collection efficiency, requiring linearly polarised incident light causing significant losses and needing a longer exposure time than imaging spectrographs under similar illumination [8].

Fourier Transform Imaging Spectrometers

By using an interferometer to self-interfere a broadband light source, an interferogram containing its spectral data is produced. Performing an inverse Fourier transform on this interferogram produces a spectrum of the light source. Fourier transform imaging

spectrometers are based on this principle. An incoming beam of light is split into two identical beams by a partially reflecting mirror known as a beam splitter. Each of these beams follows a different path before being combined and arriving at a detector. The difference in distance between the two paths creates a phase difference between the two beams resulting in an interference pattern at the detector. By altering the path difference, different interference patterns are produced (an interferogram). Performing an inverse Fourier transform on the interferogram results in a spectrum of the incident light beam. The Michelson and Sagnac interferometers are the two main interferometer designs [8].

In the Michelson configuration, one mirror and the beam splitter are held in a fixed position and another mirror is moved to change the path length of one of the beams. In the Sagnac configuration, two mirrors are fixed whilst the beam splitter can be rotated. By keeping the mirrors in a fixed position, the Sagnac interferometer has good mechanical stability and compactness, but relatively low resolution. The moving mirror in the Michelson configuration results in sensitivity to vibrations. The Sagnac configuration operates in a similar way to an imaging spectrometer where a single spatial dimension is captured along with the spectral information for that spatial line, therefore a pushbroom scan is required to build up a three-dimensional hypercube. The Michelson configuration has a pixel based interferogram allowing both spatial dimensions to be captured simultaneously. However, rather than having to move the detector or object like the Sagnac configuration, one of the mirrors must be moved, resulting in a time delay. Slower movement produces a better spectral resolution and higher SNR, but at the expensive of acquisition speed. Mainly used in bioanalytical chemistry in medicine, Fourier transform imaging spectrometers are now starting to become more popular in other hyperspectral imaging fields [8].

Single Shot Imagers

Single shot imagers acquire all spatial and spectral information simultaneously (snapshot method). This is the only method for capturing hyperspectral images of fast moving objects/scenes. In recent years, several different types of single shot imaging systems have been developed including using a multiple lens and filter approach [18], a coded aperture and compressive sensing approach [19], a multiple thin mirror approach [20], a generalisation of the Lyot filter [21] and a monolithic filter array approach [22]. While the

spectral resolutions of these systems tend to be poorer than the pushbroom alternatives, their ability to acquire all data in a single exposure means they are the preferred option for time sensitive acquisition. As the technology (and therefore spectral resolution) improves, single shot imagers will become more popular in areas where pushbroom is currently the main hyperspectral imaging acquisition method.

2.2.7.4 Image Sensors

A sensor is required for measuring the intensity of the incident light. Sensors used in hyperspectral imaging convert the incident photons into electrical signals, quantifying the intensity of the light. These sensors are usually sensor arrays, i.e. they measure at multiple different locations simultaneously. This sensor array may be one or two-dimensional. A one-dimensional sensor array may be used in a whiskbroom hyperspectral imaging system where a single point of light is dispersed into its wavelength components. Each of these components is detected by a different pixel in the sensor array, producing the spectrum of the incident light. A two-dimensional image sensor array is most often used in conventional (greyscale) digital imaging where each array pixel detects light from a different part of the scene. By viewing the response of each array pixel as a pixel in an image, we get a spatial representation of the scene, with each pixel value corresponding to the intensity of light at that point. An area scan hyperspectral imaging system uses a two-dimensional image sensor in this way. Pushbroom systems also require a two-dimensional sensor array, however, they are used slightly differently. Rather than capturing a spatial image at a specific wavelength, an image of the spectral dimension and one spatial dimension is captured.

Two main types of image sensor are charge-coupled device (CCD) and complementary metal-oxide-semiconductor (CMOS). These consist of light sensitive photodiodes which convert incident light into an electrical signal. These photodiodes tend to be made up of one of three materials: silicon (Si), indium gallium arsenide (InGaAs) or mercury cadmium tellurium (HgCdTe or MCT). Si detectors have a response in the wavelength range 300 nm to 1100 nm, making it suitable for VIS and VNIR imaging, and even capable of being used for some ultra-violet or NIR applications. InGaAs is responsive in the NIR region (900 nm to 1700 nm) but by altering the percentages of indium arsenide (InAs) and gallium arsenide(GaAs), this detection range can be extended up to 2600

nm. MCT has a much wider range of 800 nm up to 25 µm, although the exact range is dependent on the percentages of mercury(Hg) and cadmium(Cd). MCT is the most popular choice for SWIR hyperspectral imaging systems.

2.2.7.5 Data Acquisition Unit

To record the hyperspectral data captured by the sensor, a data acquisition unit (DAU) is required. Data acquisition has relatively low processing requirements compared to manipulation and analysis of hyperspectral data. The minimum requirements for a DAU are as follows:

- Hard disk drive (HDD) or solid-state drive (SSD) with sufficient write speed and storage capacity
- Sensor interface (USB, Firewire, Camera Link, etc.)
- Processor capable reading in frames and writing them to the HDD/SSD

The specification above is for a very simple hyperspectral system. The DAU may be required to perform extra functions such as controlling movement of a stage/mirror or navigation of an unmanned aerial vehicle and would have to be specified appropriately. Of course, data analysis/processing is often required to be performed on the DAU, so a more powerful computer is often used.

2.2.8 A Laboratory-Based Pushbroom Hyperspectral Imaging System

As explained earlier, there are five main components required for a pushbroom hyperspectral system:

- Light Source
- Lens
- Dispersive Element
- Imaging Sensor
- Data Acquisition Unit

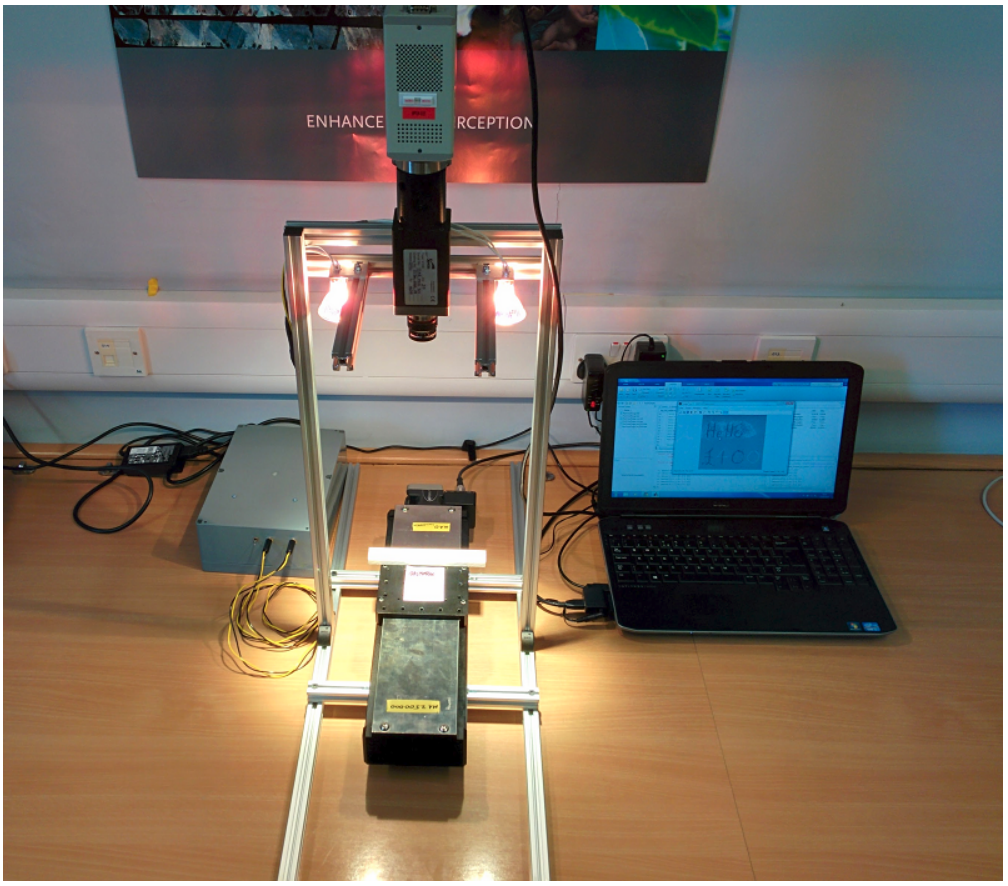


Figure 2.12: VNIR HSI System in the Strathclyde HSI Centre.

This section will describe each of these components as part of VNIR system located in the Hyperspectral Imaging Centre at the University of Strathclyde. The complete system is shown in Figure 2.12. Since this is a pushbroom system, there is a need to move the object or the detector. In the HSI Centre, there are three different devices which can provide this movement, as detailed below.

2.2.8.1 Light Source

The light source used in the HSI Centre is two 12V 50 W halogen lamps. Previously, a single 400 W halogen lamp was used. This caused problems of heating up the sample and an uneven light distribution. The two lower wattage lamps angled to illuminate a narrower line in the field of view solved both of these problems. These two lamps can be seen in Figure 2.12.

Figure 2.13: Schneider Lens sensitive in the range 400 nm to 1000 nm.

Figure 2.14: Specim V10e imaging spectrograph with spectral range of 400 nm to 1000 nm.

2.2.8.2 Lens

Since the system is VNIR with a spectral range of 400 nm to 1000 nm, a VNIR specific lens is required. A Schneider Compact VIS-NIR Lens with C-Mount is used. The lens can be seen in Figure 2.13.

2.2.8.3 Dispersive Element

The dispersive element used in the system is a Specim V10E imaging spectrograph which has a spectral resolution of 2.8 nm and a spectral range of 400 nm to 1000 nm. It accepts a c-mount lens and a c-mount detector. Figure 2.14 shows the imaging spectrograph.

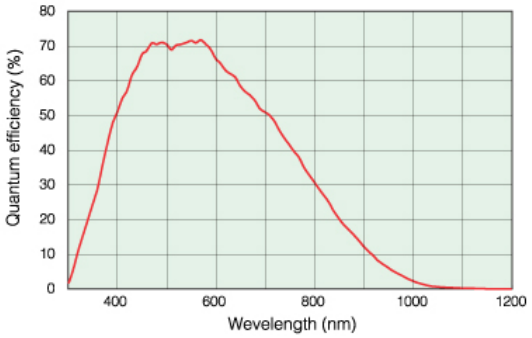


Figure 2.15: Spectral response of the Hamamatsu ORCA-05G image detector [23].

Figure 2.16: Hamamatsu ORCA-05G image sensor.

2.2.8.4 Imaging Sensor

A Hamamatsu ORCA-05G CCD camera is used as the image detector. This camera has a maximum resolution of 1344 × 1024 pixels. The readout speed is 8.9 frames/s at full resolution, but this can be increased to 43 frames/s when using 8 × 8 binning (168 × 128 resolution). Exposure times range from 10 µs to 1 s. A 12-bit analogue-to-digital converter is used giving a range of integer outputs of 0 to 4095. Connection to the DAU is via Firewire. The spectral response of this camera is shown in Figure 2.15. As can be seen from the spectral response, the detector is sensitive to the entire range of the imaging spectrograph (400 nm to 1000 nm). The sensitivity, however, is not the same for every wavelength. This means that the range of intensity values depends of the wavelength being measured. This is one of a number of reasons why correction of data is required as described in Section 2.2.9. The camera can be seen in Figure 2.16.

2.2.8.5 Data Acquisition Unit

Although the minimum requirements for the DAU are relatively low, manipulation and analysis of data requires a higher specified machine. In the system shown in Figure 2.12, a laptop with the following specifications is used:

- Intel Core i5-3340M 2.7 GHz processor
- 8 GB RAM
- 500 GB hard disk
- Firewire card
- 2 TB External USB hard disk
- Windows 7 64-bit

This machine is capable of capturing hyperspectral data of enormous proportions. Theoretically, a hypercube the entire size of the external hard disk could be captured. At the optimum resolution of the spectrograph and camera, this hypercube would have dimensions approximately $336 \times 11,500,000 \times 256$. At a spatial resolution of 5 pixels/mm, this hypercube would cover an area 2.3 km long.

2.2.8.6 Object/Detector Movement Devices

To provide the movement required for a pushbroom scan, three different devices have been used with the set up described above.

Linear Stage

The main movement device used in this system is a Zolix motorised linear stage. It consists of a 120 mm platform which can move up to 200 mm with speeds ranging from less than 0.1 mm/s to around 50 mm/s. It is mainly used for capturing small samples which can be easily be placed on the stage. It can simulate an object on a production line conveyor belt. This is the stage shown in Figure 2.12.

Motorised Trolley

To facilitate scanning of objects which are longer than the 200 mm offered by the linear

Figure 2.17: Motorised trolley.

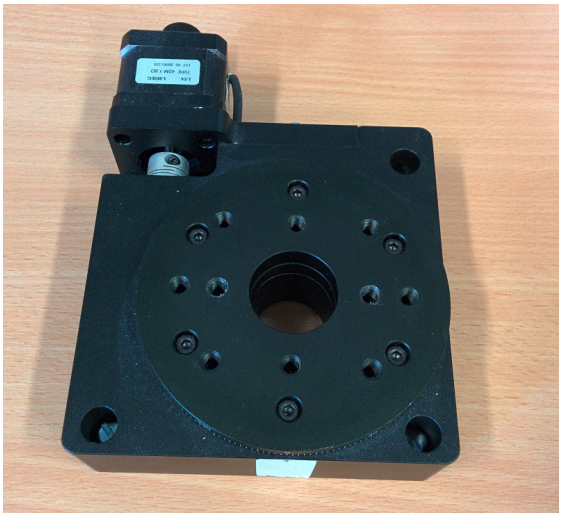


Figure 2.18: Rotational stage.

stage, a motorised trolley was built out of extruded aluminium. The trolley is moved by a motor which winds an anchored wire onto a reel. Changing the size of this reel and/or the motor speed alters the speed of the trolley. Figure 2.17 shows the trolley being used to scan 10 m long sections of power cable.

Rotational Stage

A rotational stage is also available. To capture data with a rotational stage as part of a pushbroom system, some post processing is required. This is explained in Chapter 3. By attaching the rotational stage to the linear stage, the system can image using a whiskbroom set up. Rather than capturing in Cartesian space (x, y), the image would be in polar form (r, θ). Figure 2.18 shows this rotational stage on its own.

2.2.9 Correction of Captured Data and Calibration of a System

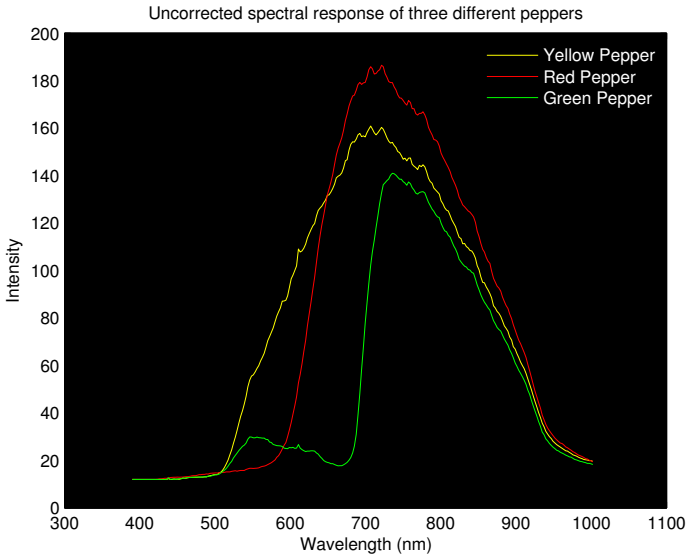
In a typical hyperspectral imaging system such as the one detailed in Section 2.2.8, the raw data captured is only an indication of the absolute intensity of the light detected by the image sensor. The spectral response of the light source and the imaging sensor will greatly affect the intensity of the output data, possibly more so than the actual object being imaged. The process required to account for this is called correction. Another factor which must be considered is calibration. If the spectrograph and image sensor have not been calibrated to each other, then there is no way of knowing which spatial locations of each wavelength on the spectrograph correspond to which pixels on the sensor.

2.2.9.1 Correction

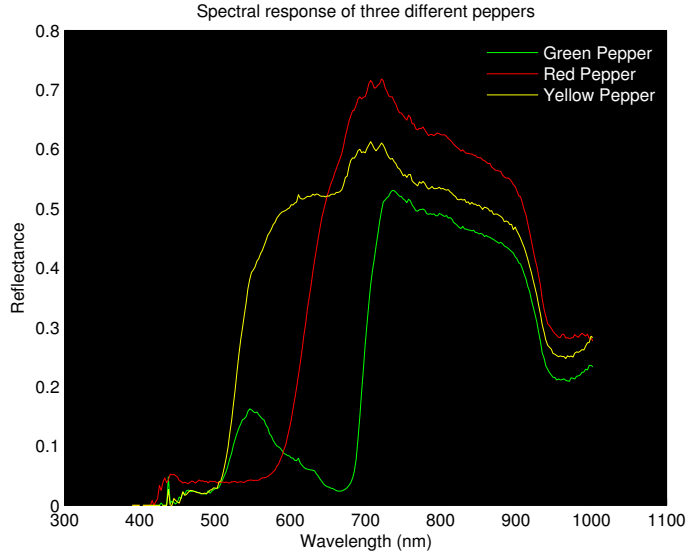
The aim of correction is to turn the intensity values at each pixel at each wavelength in a raw hypercube into reflectance values in the range 0 to 1. 0 should correspond to no light entering the sensor and 1 should correspond to all light from the light source being reflected into the sensor. To do this, 'dark' and 'white' references must be taken. The dark reference is taken by capturing a small hypercube with the lens cap on preventing any light from reaching the sensor. The white reference requires a hypercube of an object that is ideally 100% reflective. Spectralon and polytetrafluoroethylene (PTFE) are two popular materials used to make calibration tiles. Spectralon is >99% and >95% reflective over the ranges 400-1000 nm and 250-2500 nm, respectively. PTFE is >98% and >95% reflective over the ranges 250-1500 nm and 250-2200 nm, respectively [24]. By taking averages of the reference hypercubes, the dimensions are reduced from $(x \times y \times \lambda)$ to $(x \times \lambda)$. Each spatial line $(x \times \lambda)$ in a raw hypercube is corrected as follows:

$$C_L = \frac{U_L - D_m}{W_m - D_m} \tag{2.1}$$

where C_L is the corrected line, U_L is the uncorrected line, D_m is the mean of the dark reference and W_m is the mean of the white reference. After this correction, every pixel at every wavelength has a value between 0 and 1, removing the impact of the non-uniform spectral response of the image sensor and light source. Figure 2.19 shows a comparison



(a)



(b)

Figure 2.19: Spectral response of three different types of pepper (a) Uncorrected (b) Corrected

between corrected and uncorrected spectra for the three peppers used in Section 2.2. The corrected spectra give a true representation of the intensity of light reflected at each wavelength band. The spectra in figure 2.19(a) are clearly influenced by response of the sensor (Figure 2.15) but not in Figure 2.19(b).

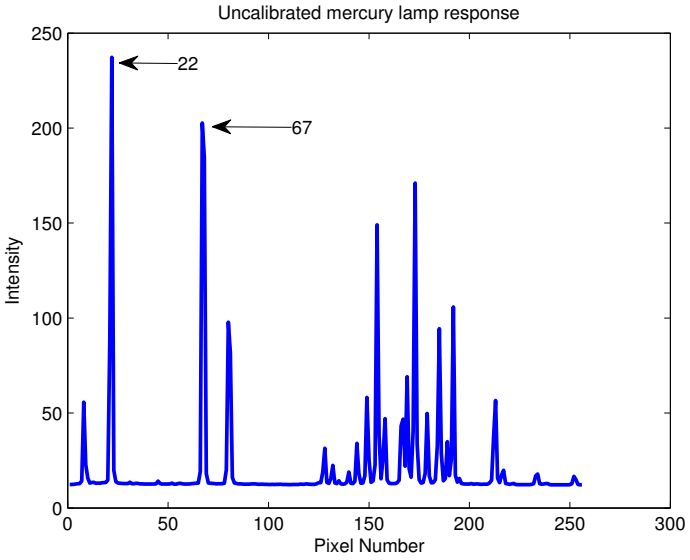


Figure 2.20: Uncalibrated mercury lamp response.

2.2.9.2 Calibration

In the system explained in Section 2.2.8, the imaging spectrograph and the imaging sensor are connected directly to each other. Although they are designed to work together, the entire output of the spectrograph is not projected exactly onto the CCD array in the imaging sensor. This means that exactly which wavelength band corresponds to which pixel is not known, so calibration must be performed. For this, a light source with a known spectral response is required. A mercury argon lamp is often chosen, as it offers several known peaks at specific wavelengths. Two main peaks that are in a useful range for a VNIR system are at 435.84 nm and 546.07 nm [25]. Exact intensities are not important for calibration, all that is required is knowledge of the peaks in the spectrum. An average spectrum of an uncalibrated hypercube of a mercury lamp is shown in Figure 2.20. Since no exact wavelength information is known at this stage, the x-axis in the graph is in pixels, rather than wavelengths. The aim of calibration is to take two (or more) peaks in the mercury lamp hypercube and determine which known peaks they correspond to. Since wavelengths are equally dispersed by the imaging spectrograph, the straight line that fits between the points can be used to determine the wavelength of each pixel. Shown below is an example:

From Figure 2.20 it can be seen that there are peaks at 22 pixels and 67 pixels. Although no exact wavelength information is known, the ranges of the spectrograph and detector mean that these peaks must correspond to the known peaks of 435.84 nm and 546.07 nm,

respectively. The equation of a straight line that fits between them can be determined as follows:

$$
\begin{aligned}
y &= mx + c && (2.2)\\
m &= \frac{y_2 - y_1}{x_2 - x_1}\\
\Rightarrow m &= \frac{(546.07 - 435.84)}{67 - 22}\\
\Rightarrow m &= 2.45\\
\Rightarrow 546.07 &= 2.45 \times 67 + c\\
\Rightarrow c &= 546.07 - 2.45 \times 67\\
\Rightarrow c &= 381.92\\
\Rightarrow y &= 2.38 \times x + 381.92 && (2.3)
\end{aligned}
$$

where y is wavelength in nm and x is pixel location in pixels. Equation 2.3 can then be used to workout the corresponding wavelength for each pixel on the image sensor.

2.3 Dimensionality Reduction

The main advantage of hyperspectral imaging over standard imaging and multispectral imaging is the quantity of data available. Whilst this provides more information about a scene, there is significant redundancy in the data. For example, the difference between an image of a scene at 700 nm compared to the same scene at 701 nm is minimal. By removing this redundancy, computational complexity can be reduced, allowing results to be obtained faster on lower specification hardware. Storage requirements are also a significantly reduced. The performance of many classification methods is also affected by the dimensionality of the input data [26]. For these reasons, many of the often cited journal papers on hyperspectral image processing involve some method of dimensionality reduction [27–30].

The main aim of dimensionality reduction in a hyperspectral imaging context is to take a hypercube of dimensions (x, y, λ), and to convert it into the form (x, y, b) where $b < \lambda$.

Reducing either x or y could also be considered dimensionality reduction, but that is more of a standard image processing technique so will not be considered here.

As explained in Section 2.2.2, hyperspectral images tend to have more than 20 bands, often hundreds. Most uses of dimensionality reduction tend to reduce the number of bands to fewer than 20.

There are a number of different methods of dimensionality used in hyperspectral image processing such as wavelet transforms [31], orthogonal subspace projections [29], extended morphological transformations [32], and principal component analysis [33]. Rather than explaining each of these techniques, PCA was chosen as the focus for this section as it is the most widely applied dimensionality reduction method for hyperspectral data [34–39].

2.3.1 Principal Component Analysis

First documented by Pearson [40] and later developed independently by Hotelling [41], Principal Component Analysis (PCA) is a method used to reduced the dimensionality of a dataset with a large number of interrelated variables, while retaining as much of the variation present in the dataset as possible. This is achieved by transforming the data into a new set of uncorrelated variables known as principal components, ordered such that the first few components retain most of the variation present in all of the original data [42]. Since most of the information is stored in the first few components, the other components can be discarded with very little loss of information. In a hyperspectral imaging context, this allows a hypercube to be greatly reduced in size, whilst still retaining most of the information, allowing efficient processing.

Going into specific details of how PCA is derived is beyond the scope of this section (see [42] for a definitive reference). Instead, a brief explanation of the steps required will be followed by three examples which demonstrate how PCA works and how powerful it can be when used to analyse hyperspectral data.

Assuming a hypercube I, consists of N band images and has dimensions $m \times n$ then each band image has $M = m \times n$ pixels and the i^{th} pixel in an image can be represented as the vector $x^i = [x^i_1, x^i_2, ..., x^i_N]^T, i = 1, 2, ..., M$. A hypercube can therefore be represented

as an $N \times M$ matrix, where $X = (x^1, x^2, ..., x^M)$. This is often known as unfolding the hypercube. PCA is then performed on this matrix as follows:

Algorithm 1: PCA Algorithm[43]

Input: The unfolded hypercube X

Mean centre the matrix as $\bar{X} = [x^1 - \mu, x^2 - \mu, ..., x^M - \mu]$ where $\mu = \frac{1}{M}\sum_{i=1}^{M} x^i$ is the mean vector in the image.

Calculate the covariance matrix $\Sigma = \frac{1}{M}\sum_{i=1}^{M}(x^i - \mu)^T = \bar{X}\bar{X}^T$.

Perform eigenvalue decomposition of the covariance matrix as $\Sigma = U\Lambda U^T$ where $\Lambda = diag(\lambda_1, \lambda_2, ...\lambda_N)$ is the diagonal matrix composed of the eigenvalues $\lambda_1, \lambda_2, ...\lambda_N$ and $U = [u^1, u^2, ..., u^N]^T$ is the orthonormal matrix composed of the corresponding eigenvectors $u^1, u^2, ..., u^N$.

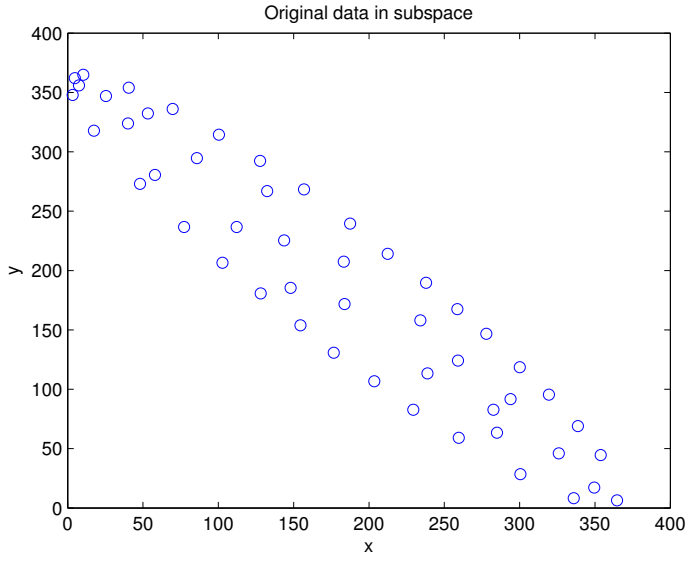
Sort the eigenvectors in descending order and retain the first K eigenvectors $U_K = (u^1, u^2, ..., u^K)$ where $K<N$ to approximate the original images: $z^i = [z_1^i, z_2^i, ..., z_K^i]^T = U_K^T x^i$, where $z^i, i = [1, 2, ...M]$ forms the first K bands of the transformed hypercube.

Output: The principal components z

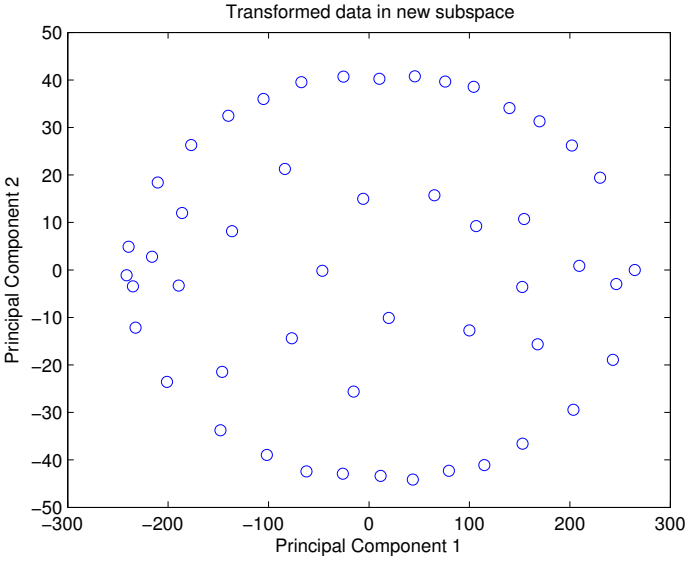
The first example will use simple synthetic two-dimensional data, showing how data can be represented in a different subspace. The second example will use a synthetic hypercube with 10 bands, showing how PCA can be used to highlight differences between similar band images. The third example will use a real hyperspectral image of different types of pen ink, with the aim of distinguishing between visually similar writings. This simulates, for example, the problem of detecting a fraudulently altered cheque.

Example 1: Synthetic two-dimensional example

Figure 2.21(a) shows a plot of a two dimensional dataset. Each point consist of an x value and a y value, producing an angled oval shape. PCA transforms this data into a new subspace which highlights the variance in the data; the first principal component corresponds to the major axis of the oval while the second principal component corresponds to the minor axis. In this example, it is clear that simply rotating and stretching the data would achieve the same result shown in Figure 2.21(b). As the number of dimensions increases, however, manually determining where and how the data should be manipulated becomes much harder. It is in these higher dimensional subspaces that PCA excels. To use the data in this example with Algorithm 1, the unfolded data



(a)



(b)

Figure 2.21: Oval shape represented in two different subspaces (a) Original subspace (b) Principal component subspace

would contain just two columns and all of the principal components would be retained, $K = N = 2$.

Example 2: Synthetic hypercube example

Consider a hypercube of size 100 x 100 x 10 (x x y x λ) pixels where each band is identical, an image made up of random noise with values between 0 and 1, shown in Figure 2.22(a). Then consider that at band number 8, the values located within the red box in Figure 2.22(b) are increased by 0.01 and at band number 3, the values located

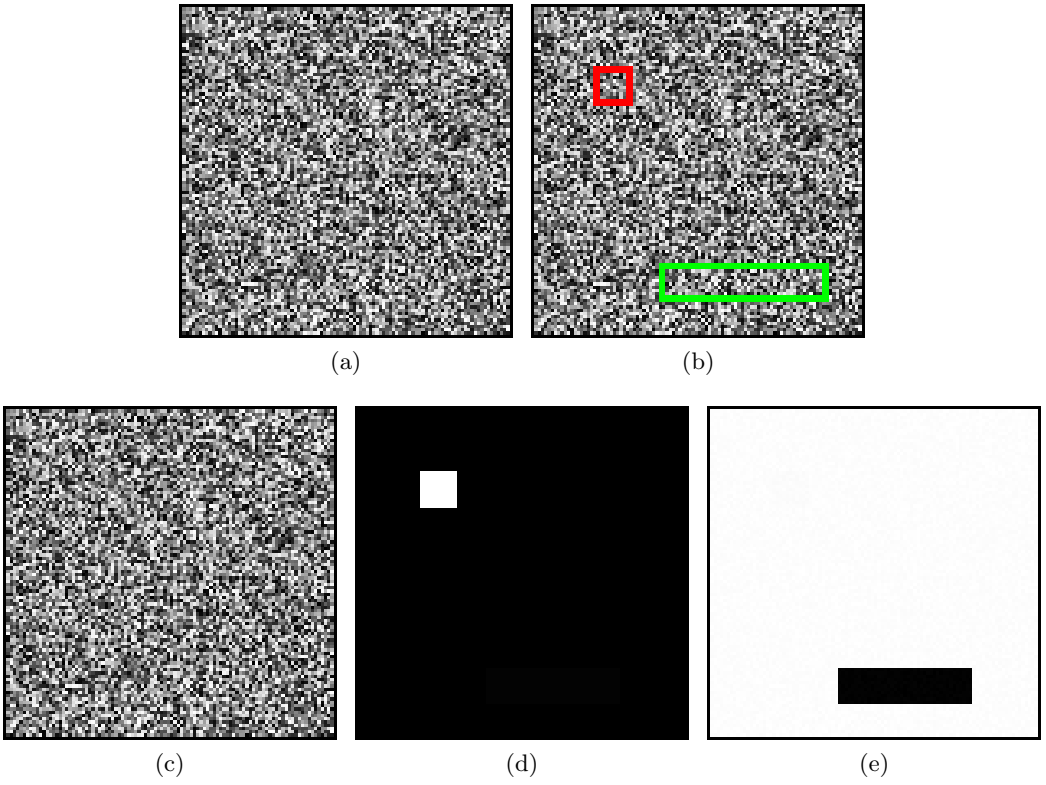
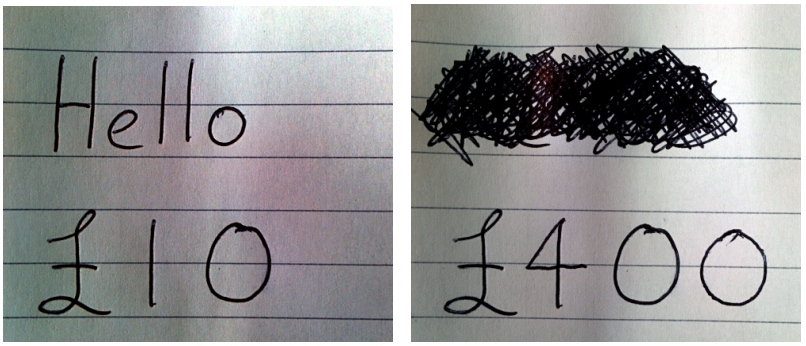


Figure 2.22: Synthetic hypercube PCA (a) Image of random values (b) Locations of edited points (c) Principal component 1 (d) Principal component 2 (e) Principal component 3

within the green box are decreased by 0.001. Now the hypercube is made up of data where each band is very similar to every other and determining where the variance lies is much harder than in the previous example. Transforming this hypercube into a subspace determined by PCA rearranges the data in such a way that the variances present are immediately obvious. The first principal component, shown in Figure 2.22(c), gives an average of the whole dataset. The second and third components, shown in Figures 2.22(d) and 2.22(e) respectively, show exactly where the variances in the data lie. This example, while synthetic, shows how multi-dimensional data can be reduced to a few components that still represent nearly all of the variance within the dataset.

Example 3: Real hypercube example - alteration detection

Figure 2.23(a) shows a small section of a simple hand-written document consisting of the word 'Hello' and '£10', written with an ordinary black pen. Figure 2.23(b) shows this same section after it has been altered by a different, but still black, pen. The word 'hello' has been completely crossed out and the amount has been changed to '£400'.



(a) (b)

Figure 2.23: Document alteration simulation captured with standard digital camera(a) Original writings (b) Modified writings

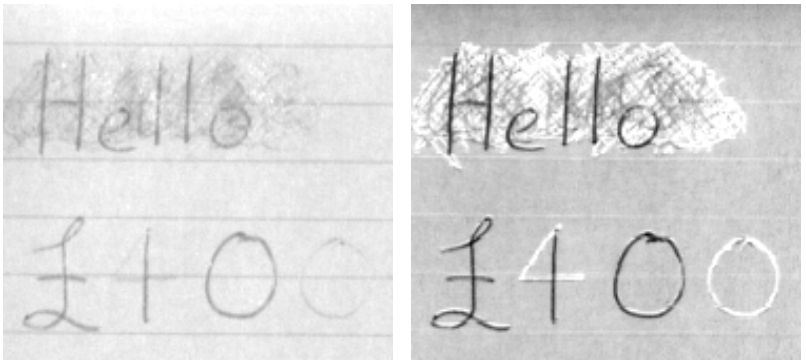
These two images were captured using a standard RGB camera. Clearly, determining what was originally written using only the second image is very hard, if not impossible.

Since the alterations were made with a slightly different ink, however, it is possible that hyperspectral imaging should be able to detect the difference. Figure 2.24(a) shows one band (700 nm) from a hyperspectral image of the altered document. From this, it can be seen that the ink used to make the alterations is partially transparent at 700 nm, revealing the changes. The low contrast, however, makes these alterations hard to see, especially where the '1' has been changed to a '4'. By transforming the data into principal components, these differences are highlighted. Figure 2.24(b) shows the fourth principal component which gives a much higher contrast image, clearly separating the two inks.

Although this example is just a simulation, it shows how PCA can be used to show differences that would be very hard to detect otherwise. The altered document could be an altered bank cheque that could have gone unnoticed if hyperspectral imaging and PCA had not been employed.

2.4 Applications

As explained in the Chapter 1, the price of hyperspectral imaging technology has decreased over recent years, resulting in a wider range of applications of the technology. This section will outline some of the major application areas and give some referenced examples of how hyperspectral imaging has been beneficial.



(a) (b)

Figure 2.24: Hyperspectral data of document alteration simulation (a) At 700 nm (b) Principal component 4

2.4.1 Food And Drink

Food and drink analysis is a major area for hyperspectral imaging. Consumers demand high standards of quality and safety so it is important for the industry that quick, accurate, and objective inspection systems are developed. Many current inspection systems rely on human visual analysis which fails to meet all of these needs. Hyperspectral imaging has therefore become a financially viable solution in a great number of situations. Since it is such a popular area for hyperspectral imaging, this thesis focusses on a number of food and drink related problems. Chapter 4 looks at the classification of Chinese tea, Chapter 5 looks at improving the baking process in sponge cakes and Chapter 6 examines how well different imaging technologies are able to classify different types of rice.

In [44], classification of Chinese tea using hyperspectral imaging was performed, however only three principal components were retained for classification and the different classes were different qualities of the same tea. In the research presented in Chapter 4, the number of principal components retained was one of the parameters whose influence was measured and the different classes were different tea types.

Hyperspectral imaging was used to assess the quality of butter cookies in [45]. In this experiment, different qualities of cookies were produced by altering the baking times. A sensory evaluation panel of six people was used to categorise each cookie as either *under-baked*, *adequately baked* or *over-baked* based purely on the browning. Moisture was measured by weighing each cookie, drying in a drying cabinet then weighing again. The lost mass was assumed to be moisture and could therefore be used to determine the

moisture content. Strong correlation was shown between the visible range hyperspectral data and the browning scores, as would be expected since the browning scores were based on visual observation. A regression model was then used to predict moisture content based on a mean spectrum for each cookie. The analysis of baked sponges in Chapter 5 differs from the above cookie analysis in that it uses different ages of sponges to achieve the different qualities of sponges, an expert panel is used to produce scores based on taste, texture and appearance and the spatial information is used show the distribution of the moisture in each sponge.

The moisture distribution in mangoes was investigated using hyperspectral imaging in [46]. As with the cookies experiment described above, ground truth for moisture was obtained by weighing before and after drying. A regression model was used map the moisture distribution in each imaged mango slice. This study, however, used an NIR system for moisture visualisation, unlike the baked sponges experiment in Chapter 5 which uses a VNIR system.

In [47], hyperspectral imaging was used for bruise detection on pickling cucumbers. Different levels of bruising were created by dropping and rolling under load, simulating the damage caused by mechanical harvesting systems. An NIR system was used and PCA, band ratio and band difference techniques were used to segregate bruised from non-bruised cucumbers.

Seafood can be successfully analysed using hyperspectral imaging as demonstrated in [48]. In this review, nematode contamination, chemical composition, freshness, microbial spoilage, adulteration and edible meat content are all shown to be determinable using hyperspectral data.

Chilling injury in Red Delicious apples was successfully detected using five VNIR bands and an artificial neural network [49] and a similar method was used for peaches in [50].

Beef tenderness has been predicted using hyperspectral imaging [51, 52]. Before this analysis, tenderness measurements were based on slice shear force - a destructive test. Hyperspectral imaging has enabled rapid, accurate, non-destructive measurements of beef tenderness and therefore quality.

The agriculture sector, specifically precision farming, has seen a similar increase in the use of hyperspectral imaging as a useful analytical tool such as in topsoil mapping [53], in

measurement of weed infestation in corn [54] and in the identification, classification and measurement of disease severity of rice crop [55, 56]. The studies on rice classification have been remote sensing based problems looking at whole fields of crops, rather than using a lab based system as in Chapter 6 where individual rice grains can be identified and analysed.

[1] provides a comprehensive review of hyperspectral imaging as a tool for food quality analysis.

2.4.2 Forensic Science

As shown by the simple example in Section 2.3.1, forgery detection is possible with hyperspectral imaging. The forensic science field has had a number of successes using hyperspectral imaging as an analytical tool. Examples include estimating the age of blood stains at a crime scene [2], analysis/detection of dynamite residue [57] and document analysis [58].

2.4.3 Military

Like with many technologies, hyperspectral imaging owes much of its progress to military funded projects. Although now used in a much wider range of areas, hyperspectral imaging is still crucial in a number of military applications mainly focussing on target detection [3], tracking [59] and terrain mapping [60].

2.4.4 Medical Imaging

Use of hyperspectral imaging in medical applications is another area where adoption continues to increase. It has been used for early diagnosis of various types of cancer [61], detection and evaluation of peripheral arterial disease [62] and retinal oxygen measurement [63, 64]. Several other medical applications are described in the extensive review in [61].

2.5 Summary

This chapter has presented a comprehensive review of the main aspects of hyperspectral imaging specifically focussing on the areas related to research documented in later chapters. The first part of the chapter focussed on the different hyperspectral technologies, imaging methodologies and the components needed for a full hyperspectral system. An example system was described which is identical to the one used to capture most of the data that is analysed in the rest of this thesis. The second part of the chapter looked at dimensionality reduction, specifically Principal Component Analysis, an important tool in hyperspectral image processing with some synthetic and real examples demonstrating its capability. Finally, four of the main application areas of hyperspectral imaging were outlined with numerous examples of where hyperspectral imaging has been a useful tool. Due to the number of successes in the food and drink industry, Chinese tea, baked sponges and rice are investigated in the later chapters of this thesis.

Chapter 3

Rotational Scanning Techniques

3.1 Introduction

As explained in Section 2.2.4, most hyperspectral imaging acquisition techniques require two of the hypercube dimensions to remain constant, whilst the third is varied. A popular method for capturing hyperspectral data is linear pushbroom scanning [47, 65]. This technique builds an image by capturing all available wavelengths on a spatial line-scan translated across the image over time. A major drawback of this technique is that an accurate way of linearly moving the camera or the object is required [66]. Rotational movement is often easier to achieve or naturally available, hence the work presented here investigates an alternative approach to the traditional linear pushbroom method.

Since linear pushbroom scanning captures a line of pixels perpendicular to the direction of motion, a Cartesian representation of the image is created by simply placing these lines side by side. The speed of the motion and the exposure time of the camera must be set accurately to produce square pixels. If the exposure time is too long or the motion is too fast, the pixels will be squashed. If the exposure time is too short, or the motion is too slow, the pixels will be stretched.

Constructing a Cartesian representation of an image from a rotational scan is not as simple, since the lines from the scan cannot just be placed side by side. Some form of algorithm is needed to perform this construction. Figure 3.1 shows how a rotational scan differs from a linear scan.

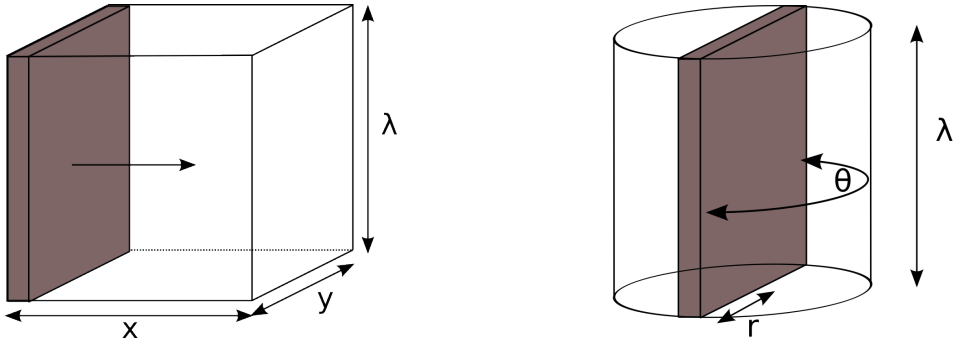


Figure 3.1: Two different types of line-scan. (a) Linear. (b) Rotational.

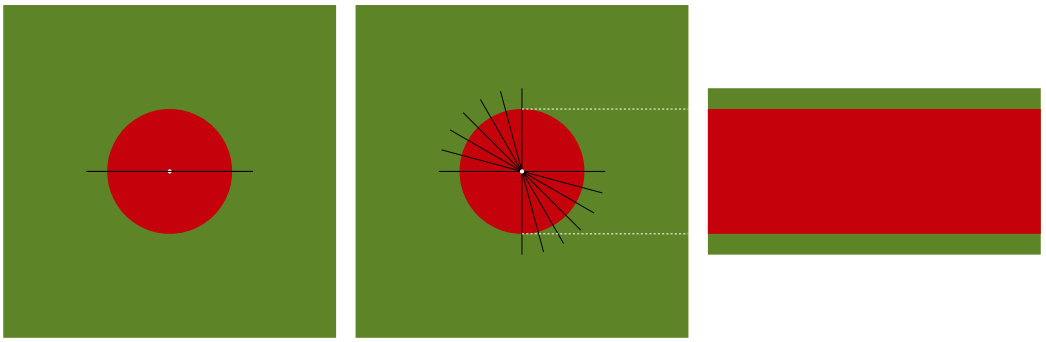
There are numerous cases of hyperspectral image capture using rotational movement [67–70], but in each of these cases, the axis of rotation has been perpendicular to optical axis. The method described above uses a rotational scan where the axis of motion is parallel to the optical axis which is believed to be entirely novel.

All images acquired for this study were captured using the system described in Section 2.2.8 with the rotational stage as the movement device.

The remaining parts of this chapter are organised as follows: Sections 3.2, 3.3 and 3.4 describe how a rotational scan is performed and the problems associated with it; Section 3.5 shows how a rotational scan can be converted to Cartesian form; Section 3.6 proposes two techniques for calculating the two offsets associated with a rotational scan; Section 3.7 shows the implementation of three interpolation techniques; Section 3.8 contains the results of a comparison with linear scanning using image similarity measures; Section 3.9 discusses some other considerations when using a rotational scan over a linear one; and summarising remarks are presented in Section 3.10.

3.2 The Ideal Scan

The ideal rotational scan has the centre of each line scanned exactly coincident with the axis of rotation. An example of an ideal scan is shown (from a top down view) in Figures 3.2(a) and 3.2(b) where the white dot represents the axis of rotation and the black line represents an imaged line. Figure 3.2(c) shows what the output of the scan would look like. In this case, the height of the red bar is exactly the diameter of the red circle and there is an equal amount of green on either side of it. Converting an ideal



(a) (b) (c)

Figure 3.2: An ideal line-scan of a red circle on a green background. (a) First line of the scan. (b) More lines of the scan. (c) The output of the scan.

scan to Cartesian representation is a relatively simple task, since each half of the scan represents the image in polar form. There are publicly available toolboxes which can perform this conversion[71].

3.3 Offset Scans

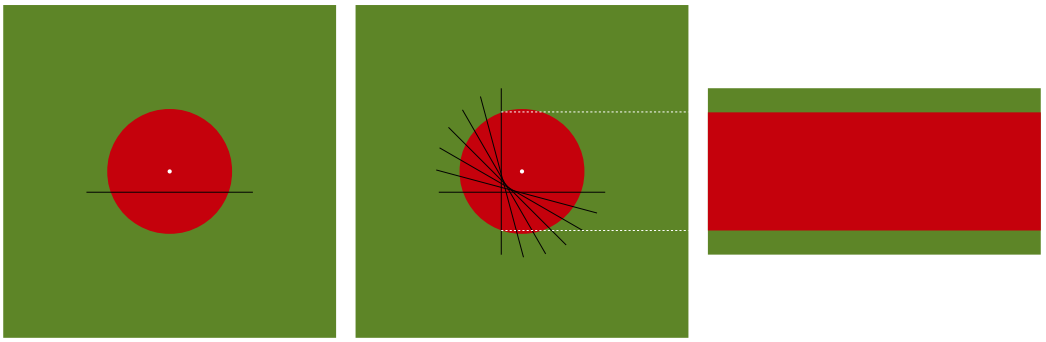
In practice, however, achieving an ideal scan is almost impossible, i.e. the axis of rotation will not coincide with the centre of each line scanned. This error can be thought of as the combination of two different offsets: a vertical offset and a horizontal offset.

3.3.1 Vertical Offset

Figure 3.3(a) shows the first line of a vertically offset scan. The axis of rotation is still the white dot, but the line scan does not pass through it. Since the line scan does not pass through the centre of the red circle, the red bar produced in Figure 3.3(c) has a slightly smaller height.

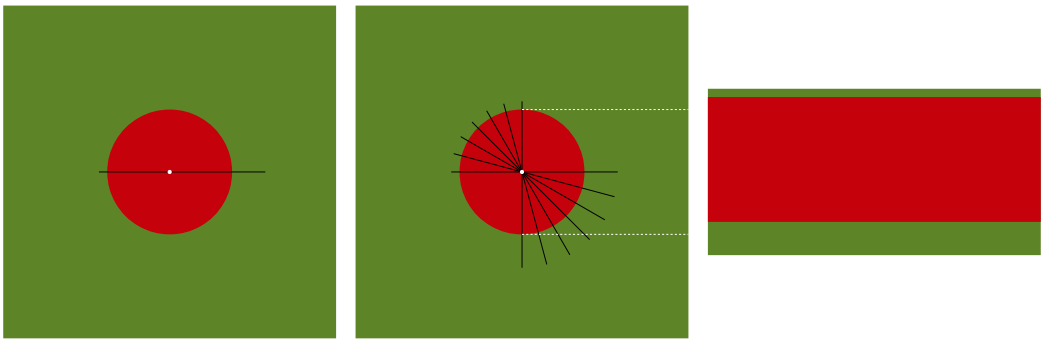
3.3.2 Horizontal Offset

A horizontal offset is easier to accommodate as it simply causes a shift in the output of the scan. A horizontal offset would cause Figure 3.2(c), for example, to be shifted up or down such that there is an unequal amount of green on either side of the red bar. This is shown in Figure 3.4.



(a) (b) (c)

Figure 3.3: A vertically offset line-scan of a red circle on a green background. (a) First line of the scan. (b) More lines of the scan. (c) The output of the scan.



(a) (b) (c)

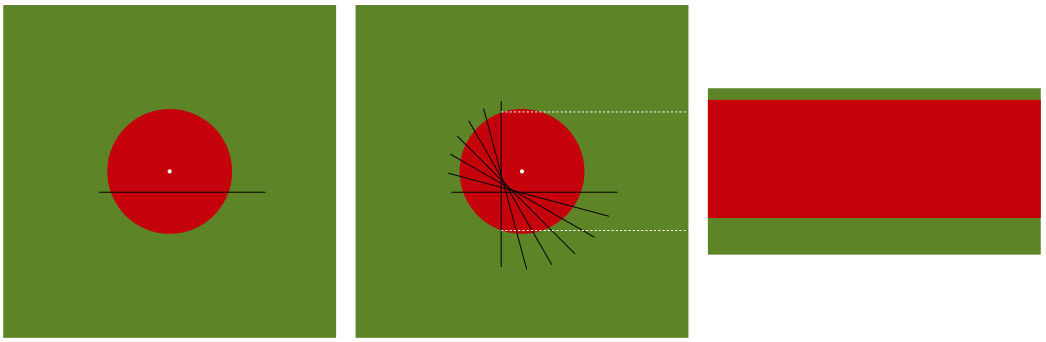
Figure 3.4: A horizontally offset line-scan of a red circle on a green background. (a) First line of the scan. (b) More lines of the scan. (c) The output of the scan.

3.3.3 Vertical Offset and Horizontal Offset

When both of these offsets are present, they operate independently. Figure 3.5 shows a scan that is both vertically and horizontally offset. Figure 3.5(c) shows that the output of this scan is a combination Figures 3.3(c) and 3.4(c); the height of the red bar has been shortened due to the vertical offset and it has been shifted upwards due to the horizontal offset.

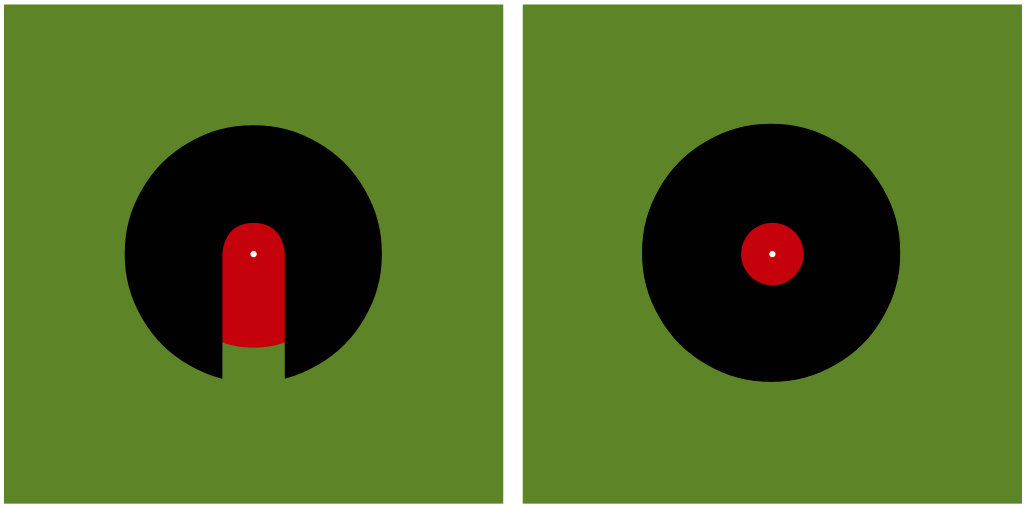
3.4 Half Rotation versus Full Rotation

In an ideal scan, only half a rotation is needed to ensure that every possible point on the object in the field of view is scanned. This is not the case with an offset scan. The vertical offset means that a circle of radius equal to the offset will never be scanned, regardless of how many rotations are completed. To ensure every possible point on



(a) (b) (c)

Figure 3.5: A vertically and horizontally offset line-scan of a red circle on a green background. (a) First line of the scan. (b) More lines of the scan. (c) The output of the scan.



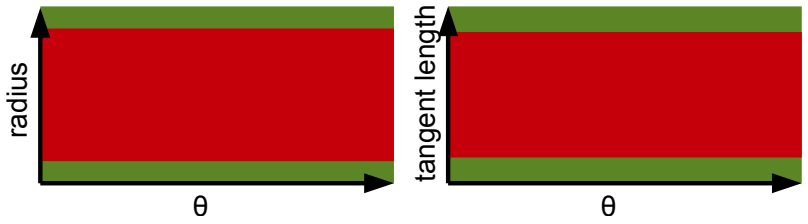
(a) (b)

Figure 3.6: The difference between vertically offset half rotation and full rotation scans. (a) Half rotation. (b) Full rotation.

the object in the field of view is scanned, one full rotation is required. There is some redundancy in this scan, however, as many points will be scanned twice. Half a rotation will scan most of the image in the field of view leaving a section of the object unscanned and another section scanned twice. The size of both of these sections is dependent on the vertical offset. Figure 3.6 shows a half rotation and a full rotation scan.

3.5 Mathematical Representation and Conversion

Each line in a vertically offset scan can be thought of as a tangent to a circle with a radius equal to the offset value. Since achieving a vertical offset of precisely zero is very



(a) (b)

Figure 3.7: The difference coordinate systems. (a) Ideal Scan. (b) Vertically Offset Scan.

unlikely, every scan can be represented by tangents to a circle, even if this circle is very small. The horizontal offset determines how long each tangent extends from the circle.

3.5.1 Converting to Cartesian Representation

An offset scan of an object is similar to representing it in polar form. Each pixel in polar form corresponds to a radius and an angle but in an offset scan, each pixel corresponds to a tangent length and an angle, see Figure 3.7. Cartesian co-ordinates can easily be determined for a polar representation of an image using

$$x = r\cos\theta \tag{3.1}$$

$$y = r\sin\theta \tag{3.2}$$

Converting from an offset scan to Cartesian co-ordinates is not so simple.

Since a horizontal offset is just a shift and can be easily compensated for at a later stage, this section will only consider a vertical offset. As explained earlier, a half scan will not cover every point in the field of view so a scan of 2π radians will be used.

Before the Cartesian representation can be calculated, the vertical offset must be known. This is equal to the radius of the circle whose tangents make up the scan. Let this radius be r. The equation of this circle is then

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \tag{3.3}$$

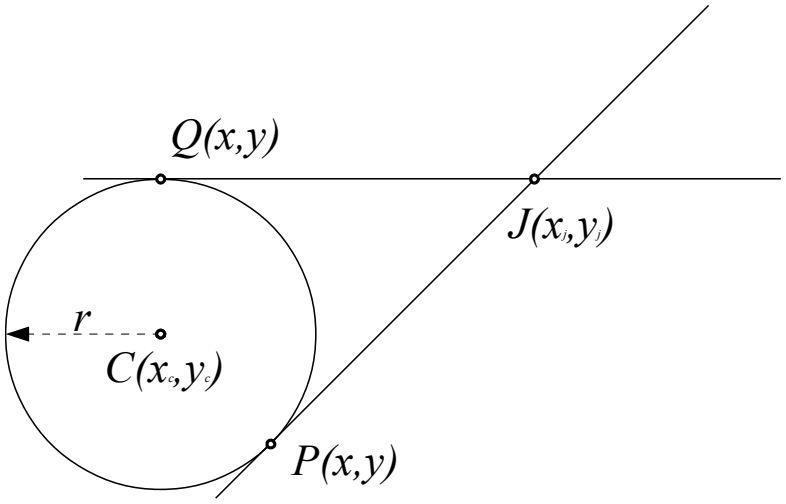


Figure 3.8: A circle with two intersecting tangents

where (x_c, y_c) is the centre of the circle, i.e. the axis of rotation. The circle can also be represented by

$$x = x_c + r\cos\theta \tag{3.4}$$

$$y = y_c + r\sin\theta \tag{3.5}$$

where θ is the angle between (x, y), the centre of the circle and the horizontal.

Suppose there is a point outside this circle, J, at (x_j, y_j). There must be two lines which pass through this point and are tangential to the circle. See Figure 3.8.

If the implicit equations of the two tangents are in the form

$$ax + by + c = 0 \tag{3.6}$$

then according to [72], a, b, and c can be found as follows:

$$a = \frac{\mp r(x_c - x_j) - (y_c - y_j)\sqrt{(x_c - x_j)^2 + (y_c - y_j)^2 - r^2}}{(x_c - x_j)^2 + (y_c - y_j)^2} \tag{3.7}$$

$$b = \frac{\mp r(y_c - y_j) - (x_c - x_j)\sqrt{(x_c - x_j)^2 + (y_c - y_j)^2 - r^2}}{(x_c - x_j)^2 + (y_c - y_j)^2} \tag{3.8}$$

$$c = -ax_j - by_j \tag{3.9}$$

The equations of the tangents can then be used to determine the two points of intersection with the circle (P and Q on Figure 3.8):

$$\begin{aligned} ax + by + c &= 0 \\ \Rightarrow y &= -\frac{a}{b}x - \frac{c}{b} \end{aligned}$$

Let $m = -\frac{a}{b}$ and $C = -\frac{c}{b}$

$$\begin{aligned} \Rightarrow y &= mx + C \\ (x - x_c)^2 + (y - y_c)^2 &= r^2 \\ \Rightarrow (x - x_c)^2 + (mx + C - y_c)^2 &= r^2 \\ \Rightarrow x^2 - 2x_c x + x_c^2 + (mx + C)^2 - 2y_c(mx + C) + y_c^2 &= r^2 \end{aligned}$$

$$\begin{aligned} \Rightarrow x^2 - 2x_c x + x_c^2 + m^2x^2 + 2mxC + C^2 - 2y_c mx & \\ +2y_c C + y_c^2 &= r^2 \\ \Rightarrow x^2(m^2 + 1) + x(2mC - 2x_c - 2my_c) + x_c^2 + y_c^2 & \\ -2y_c C + C^2 &= r^2 \end{aligned}$$

$$\begin{aligned} \Rightarrow x^2(m^2 + 1) + x(2mC - 2x_c - 2my_c) + x_c^2 + y_c^2 & \\ -2y_c C + C^2 - r^2 &= 0 \end{aligned} \tag{3.10}$$

Equation 3.10 can then be solved for x using the quadratic formula giving the x co-ordinates of the points of intersection with the circle. The equations of the tangents can be used to compute the corresponding y co-ordinates. Equation 3.4 or 3.5 can then be used to calculate the two θs. The Euclidean distance between J and P gives one tangent length and the Euclidean distance between J and Q gives the other. Thus, point J can be converted from Cartesian representation to two angles and two tangent lengths. These values can then be used to extract a pixel from the output scan image. Repeating this process for every (x, y) point allows a Cartesian representation of the scanned object to be generated.

3.5.2 Compensating for a horizontal offset

As mentioned previously, the horizontal offset is easier to deal with. A horizontal offset of zero would mean that the axis of rotation is in the middle of the line scanned and so each of the tangents determined in Section 3.5.1 would be of equal length. If the two parts of each scanned line are called 'Left Scan' and 'Right Scan' respectively, then their lengths can be evaluated using the horizontal offset, HO, as follows:

$$LeftScanLength = \frac{TotalScanLength}{2} + HO \tag{3.11}$$

$$RightScanLength = \frac{TotalScanLength}{2} - HO \tag{3.12}$$

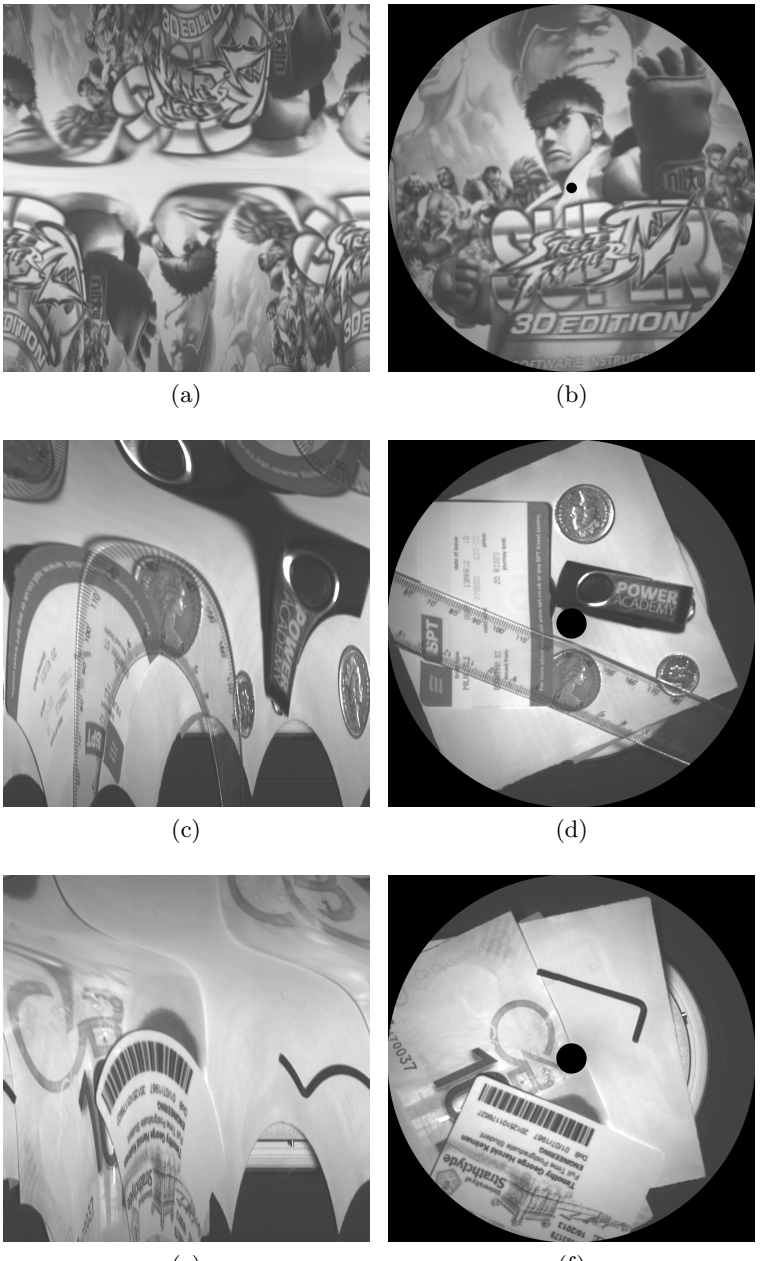
The output of the conversion to Cartesian representation is an angle and a tangent length. If this tangent length is less than or equal to the Left Scan Length, then a value for the Left Scan exists. The same is true for the Right Scan. There are four possible combinations: there is a value for just the left scan; there is a value for just the right scan; there is a value for both scans; or there is a value for neither scan. This explains why part of the image is scanned twice as discussed in Section 3.4.

3.5.3 Example Scans

Figure 3.9 shows some original and converted scans. Figures 3.9(a), 3.9(c) and 3.9(e) are the original offset scans and Figures 3.9(b), 3.9(d) and 3.9(f) are the converted scans. Note that Figures 3.9(a) and 3.9(b) have a smaller vertical offset, hence the smaller black circle in the centre of the converted scan.

3.6 Determining the offsets

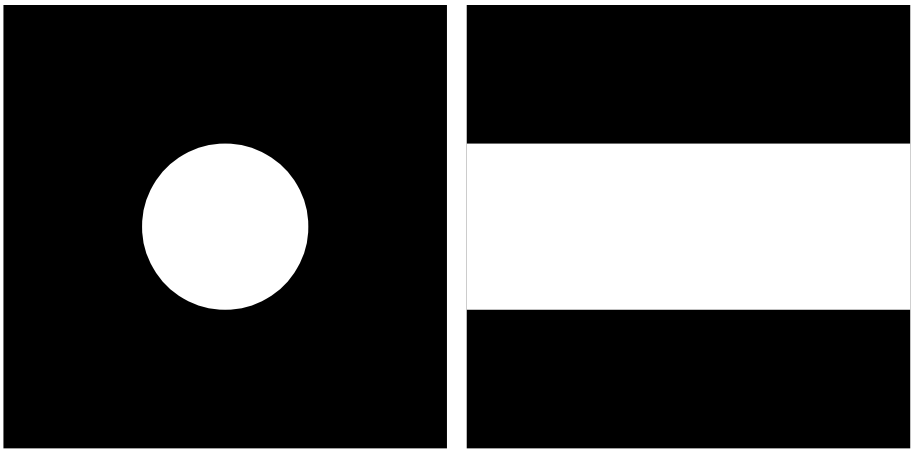
As explained in Section 3.5, the offsets must be known to allow the conversion algorithm to operate. Initially, these offsets were determined through trial and error, but this was both time consuming and inaccurate. This section explains two automatic methods of determining the offset parameters using a calibration object. In essence, an object of known dimensions is scanned and the output is compared to a theoretical perfect scan of



(a) (b) (c) (d) (e) (f)

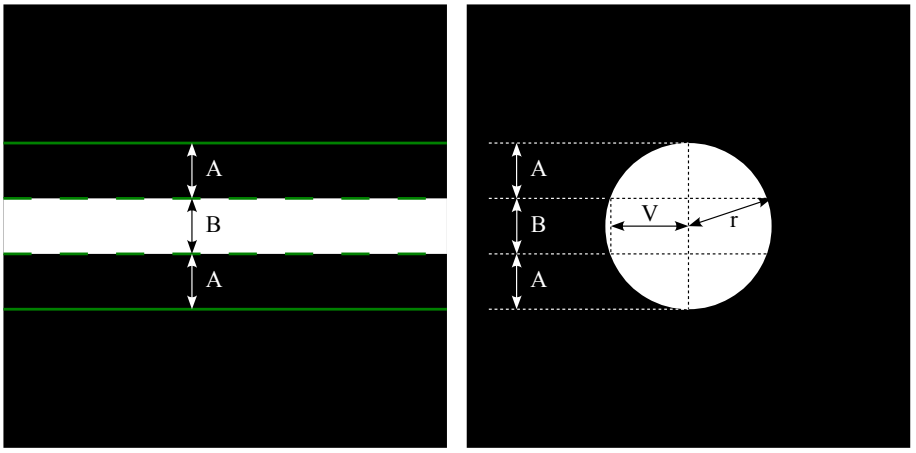
Figure 3.9: Example rotational scans converted to Cartesian form. (a) Computer game manual original scan. (b) Computer game manual converted scan. (c) Collection of objects original scan. (d) Collection of objects converted scan. (e) ID, money and an Allen key original scan. (f) ID, money and an Allen key converted scan.

that object. The difference between the actual output scan and the theoretical output scan should reveal the offset parameters.



(a) (b)

Figure 3.10: An ideal scan of the calibration object. (a) The calibration object (Cartesian space). (b) The output of an ideal scan (*θ-tangent* space).



(a) (b)

Figure 3.11: A vertically offset scan of the calibration object. (a) The output of the scan (*θ-tangent* space). (b) Calibration object showing the difference between an ideal scan and a vertically offset scan (Cartesian space).

3.6.1 Using a calibration object with perfect alignment

The calibration object used is shown in Figure 3.10(a). For the offset determination technique explained in this section, the calibration object must be aligned such that the centre of the circle coincides with the axis of rotation. The theoretical output of this scan (no offsets) is a horizontal bar vertically in the centre of the image with a height equal to the diameter of the white circle, as shown in Figure 3.10(b).

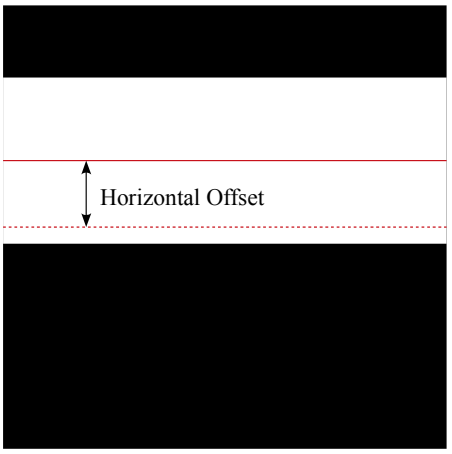


Figure 3.12: Horizontally offset calibration object output scan (θ-*tangent* space).

3.6.1.1 Calculating the Vertical Offset

Figure 3.11(a) shows the output of a vertically offset scan. Using the knowledge of what an ideal scan should look like, we can calculate the vertical offset V using Pythagoras' Theorem as follows:

$$\begin{aligned} r^2 &= V^2 + (\frac{B}{2})^2 \\ \Rightarrow V^2 &= r^2 - (\frac{B}{2})^2 \\ \Rightarrow V &= \sqrt{r^2 - (\frac{B}{2})^2} \end{aligned} \tag{3.13}$$

3.6.1.2 Calculating the Horizontal Offset

Figure 3.12 shows a horizontally offset output scan of the calibration object. The difference between the centre of the white bar (solid red line) and the centre of the image (dotted red line) is equal to the horizontal offset.

3.6.2 Using a calibration object with imperfect alignment

In the previous section, the centre of the calibration object had to be aligned perfectly with the axis of rotation. This is relatively simple to achieve when the imaging set-up is a stationary camera and moving object. If, for example, the rotational movement is provided by a motorised rotation stage, the centre of the axis of rotation is easy to determine. If the set-up involves a stationary object and rotating camera (for example,

a camera attached to a robotic arm), the exact location of the axis of rotation is hard to determine. If the axis of rotation cannot be accurately determined, then the offset determination technique explained in Section 3.6.1 will fail. This section explains a technique for determining the offsets using a calibration circle that does not need to be accurately aligned.

3.6.2.1 Calibration circle requirements

Although the calibration circle does not need to be accurately aligned, there are several requirements that must still be met:

1. The calibration circle must be small enough that it is fully scanned (apart from the blind spot caused by the offset).
2. The calibration circle must overlap with the offset circle.
3. The calibration circle must be larger than the offset circle.

Figure 3.13 shows four calibration circles (white) with a different offset circle (red) overlaid on each one. Figure 3.13(a) and Figure 3.13(b) are acceptable. Figure 3.13(c) would fail since the offset circle is larger than the calibration circle and Figure 3.13(d) would fail since the offset circle is entirely outwith the calibration circle.

Figure 3.14 shows what the output scan in each of these calibrations situations would look like.

By looking at these output scans, we can determine if the calibration requirements have been met.

Since the calibration circle does not appear at the edge (top or bottom) of the output scan in any of the scenarios, then requirement 1 has been met.

In the Figure 3.14(d), there are two distinct regions, therefore the offset circle is entirely outwith the calibration circle and requirement 2 has not been met. The other three scans all only have one distinct region so meet requirement 2. Note that for this requirement, regions are allowed to wrap around the edges (left and right) of the scan since the original starting angle of the scan does not matter. For example, the scan in Figure 3.14(b)

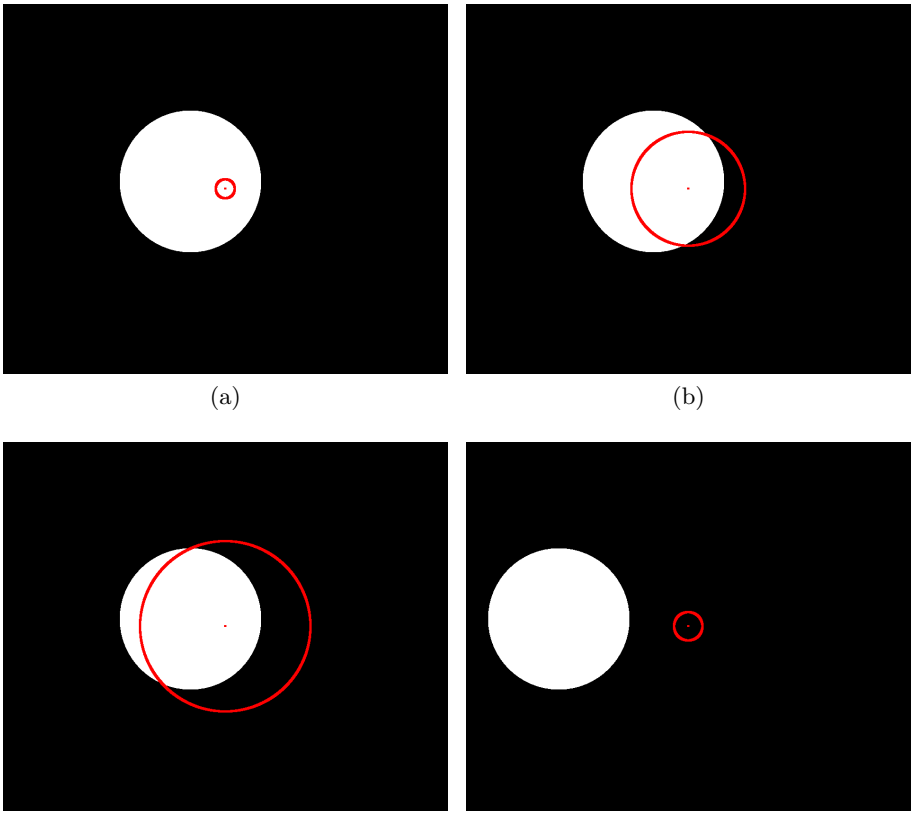


Figure 3.13: Examples of calibration circles (white) with offset circles (red). Each figure is in Cartesian Space. (a) Valid. (b) Valid. (c) Invalid. (d) Invalid.

could be shifted by π radians resulting in two separate regions. These regions would wrap around, however, resulting in only one distinct region.

Requirement 3 is the hardest to determine. Since the calibration circle in Figure 3.14(a) spans the entire width of the scan, the calibration circle must be larger than the offset circle. Figure 3.14(d) has already been determined to be invalid. Figure 3.14(b) and Figure 3.14(c) require further analysis. First, the location of the tangent whose intersection with the offset circle occurs closest to the centre of the calibration circle needs to be found. This location is an angle. This can be found looking at the plot of the sum of the output scan (in the vertical direction i.e., the number of white pixels at any x location). If the offset circle overlaps with the centre of the radius of the calibration circle, like in Figure 3.13(b), then this location will be at the maximum value in the plot of the sum. See Figure 3.15(a). If the offset circle radius does not overlap with the centre of the calibration circle, like in Figure 3.13(c), then this location will be at the local minimum in the plot of the sum. See Figure 3.15(b).

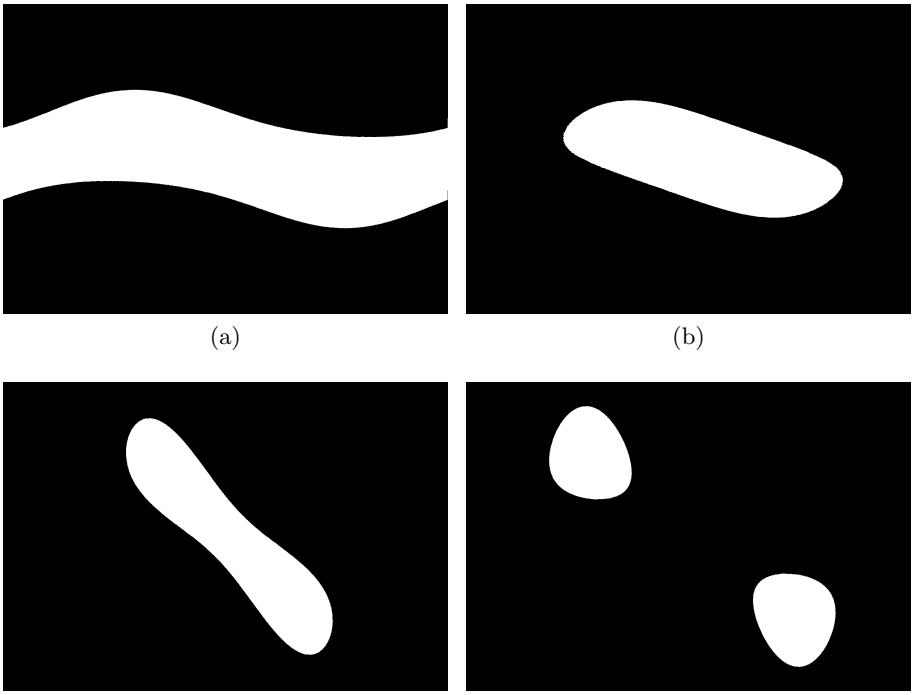


Figure 3.14: Examples of the scans produced from Figure 3.13. Each figure is in *θ-tangent* space. (a) Valid. (b) Valid. (c) Invalid. (d) Invalid.

It can be seen that the angle of the maximum in Figure 3.15(a) is 3.37 radians and the angle of the local minimum in Figure 3.15(b) is 3.19 radians.

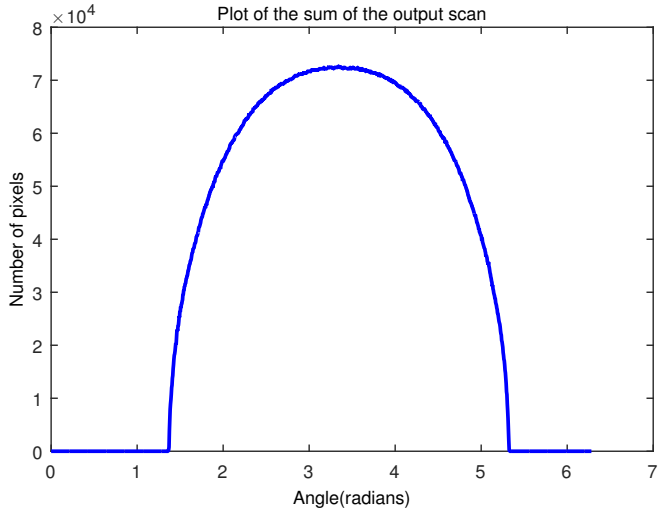
After this angle has been found, the values in the sum plot $\frac{\pi}{2}$ radians either side of this angle need to be noted. This will give the locations (angles) of the tangents perpendicular to the one found previously. If this tangent intersects with the calibration circle, then the offset circle must be smaller than the calibration circle and vice versa.

From Figure 3.15(a), these two tangents are at 1.80 radians and 4.05 radians and have corresponding non-zero values in the sum plot meaning that they intersect with the calibration circle.

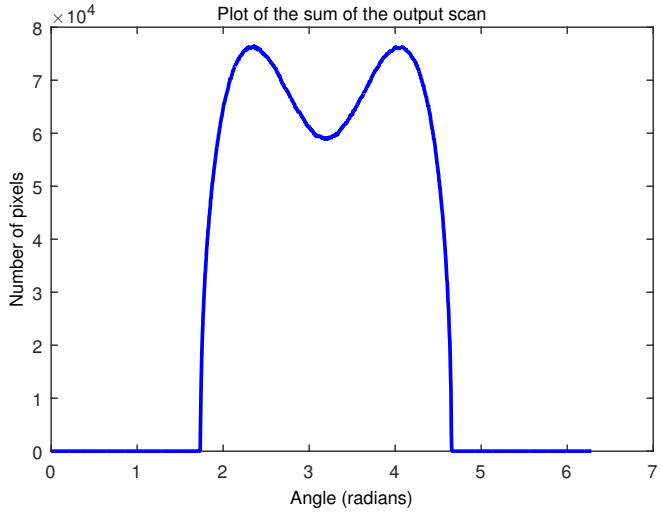
In Figure 3.15(b), these values are at 1.62 radians and 4.76 radians, but in this case, their corresponding values in the sum plots are zero meaning that they do not intersect with the calibration circle.

Therefore Figure 3.13(b) meets requirement 3 and Figure 3.13(c) does not.

If all the requirements are met, then the scanned image can be used to calculate the vertical and horizontal offsets.



(a)



(b)

Figure 3.15: Plots of the sums of two output scans (a) Maximum (b) Local minimum

3.6.2.2 Calculating the Vertical Offset

The method used to calculate the vertical offset (i.e., the radius of the offset circle), depends up the location of the calibration image in the scan. In this section, four examples will be used to show how an output scan can be used to calculate the vertical offset. For each example, three figures will be shown: the offset circle overlaid on the calibration circle, the output of the scan and a plot of the sum of the output scan. The output scan and the plot of the sum will be used to calculate the vertical offset for each example. These four examples are all valid scans and do not correspond to the four examples in the previous section.

Offset circle fits entirely within the calibration circle and overlaps the centre of the calibration circle

Figure 3.16(a) shows a calibration circle (white) with an offset circle (red) overlaid. This offset circle fits entirely within the calibration circle and it also overlaps the centre of the calibration circle. Figure 3.16(b) shows what the output of this scan would look like. Figure 3.16(c) shows a plot of the sum of each column in the output scan.

The first step is to find the length of tangent of the offset circle that is closest to the centre of the calibration circle, the green line labelled Ta on Figure 3.16(a). This can be found by examining the sum of plots of the output scan shown in Figure 3.16(c). Since the offset circle overlaps with the centre of the calibration circle, the tangent closest to the centre of the calibration circle will also be the longest. From the sum of plots, it can be seen that the highest value is 195 and occurs at 3.1416 (This happens to be π since the offset circle is in line horizontally with the calibration circle and the scan started at the right hand side of the offset circle). Therefore $Ta = 195$.

Next, the tangent furthest from the centre of the calibration circle, Tb, must be found. This is exactly $\pm\pi$ radians away from Ta. Looking at the plot of sums, Tb is therefore 167 pixels.

CCr on Figure 3.16(a) is the calibration circle radius. This needs to be known for this process to work, and in this case is 100 pixels.

A and B can then be calculated as follows:

$$CCr^2 = A^2 + (\frac{Ta}{2})^2$$

$$\Rightarrow A^2 = CCr^2 - (\frac{Ta}{2})^2$$

$$\Rightarrow A = \sqrt{CCr^2 - (\frac{Ta}{2})^2}$$

$$\Rightarrow A = \sqrt{100^2 - (\frac{195}{2})^2}$$

$$\Rightarrow A = 22.2205$$

$$CCr^2 = B^2 + (\frac{Tb}{2})^2$$

$$\Rightarrow B^2 = CCr^2 - (\frac{Tb}{2})^2$$

$$\Rightarrow B = \sqrt{CCr^2 - (\frac{Tb}{2})^2}$$

$$\Rightarrow B = \sqrt{100^2 - (\frac{167}{2})^2}$$

$$\Rightarrow B = 55.025$$

Looking at Figure 3.16(a), the radius of the offset circle (i.e. the vertical offset) is found as follows:

$$VO = \frac{A + B}{2}$$

$$\Rightarrow VO = \frac{22.2205 + 55.0250}{2}$$

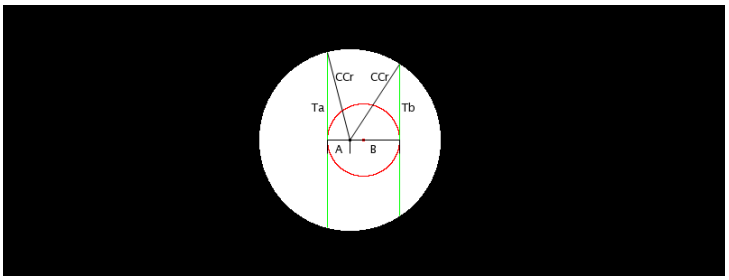
$$\Rightarrow VO = 38.6227$$

The actual vertical offset used in this example was 40 pixels, so while 38.6 is close, it would still produce some errors in a reconstructed scan. The reasons for this inaccuracy are discussed later in the chapter.

Offset circle fits entirely within the calibration circle and does not overlap the centre of the calibration circle

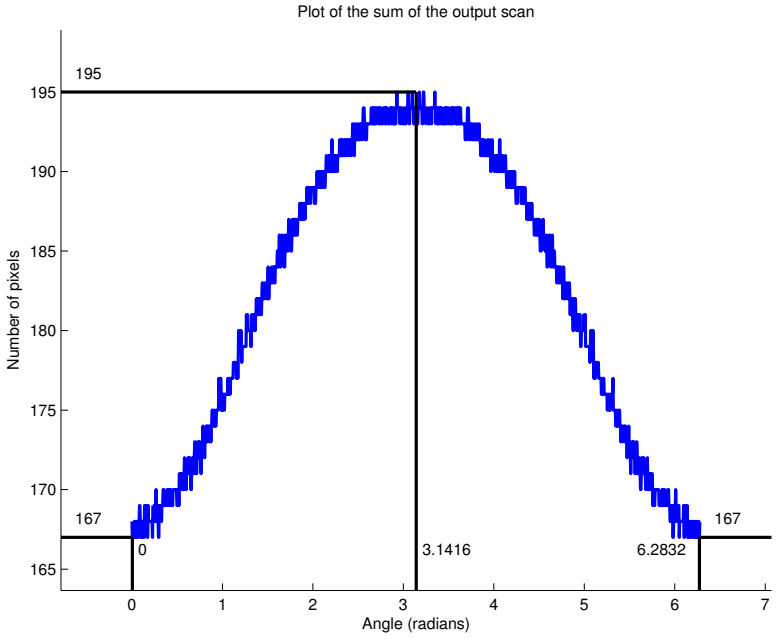
Figure 3.17(a) shows a calibration circle (white) with an offset circle (red) overlaid. This offset circle fits entirely within the calibration circle but, in this case, it does not overlap the centre of the calibration circle. Figure 3.17(b) shows what the output of this scan would look like. Figure 3.17(c) shows a plot of the sum of each column in the output scan. The main difference between this scan and the previous one, is that the tangent increases to the maximum, decreases to a local minimum, returns to the maximum before decreasing to the global minimum again. This is not very apparent in Figure 3.17(b), but is very clear in the plot of sums, Figure 3.17(c).

Again, the first step in determining the vertical offset is to find the length of the tangent to the offset circle closest to the centre of the calibration circle. This is the green line labelled Ta on Figure 3.17(a). Since the offset circle does not overlap the centre of the calibration circle, this will be a local minimum on the plot of sums, Figure 3.17(c). From the plot, the length of Ta is 197 pixels.



(a)

(b)



(c)

Figure 3.16: Example Scan 1. (a) Annotated calibration circle and offset circle (Cartesian space). (b) Output of the scan (θ-*tangent* space). (c) Plot of the sum of the output scan.

The other tangent length required, Tb, is $\pm\pi$ radians away from Tb. Again from the plot of sums, this is 64 pixels.

As before, A and B can be calculated as follows:

$$CCr^2 = A^2 + (\frac{Ta}{2})^2$$

$$\Rightarrow A^2 = CCr^2 - (\frac{Ta}{2})^2$$

$$\Rightarrow A = \sqrt{CCr^2 - (\frac{Ta}{2})^2}$$

$$\Rightarrow A = \sqrt{100^2 - (\frac{197}{2})^2}$$

$$\Rightarrow A = 17.2554$$

$$CCr^2 = B^2 + (\frac{Tb}{2})^2$$

$$\Rightarrow B^2 = CCr^2 - (\frac{Tb}{2})^2$$

$$\Rightarrow B = \sqrt{CCr^2 - (\frac{Tb}{2})^2}$$

$$\Rightarrow B = \sqrt{100^2 - (\frac{64}{2})^2}$$

$$\Rightarrow B = 94.7418$$

Looking at Figure 3.17(a), the radius of the offset circle (i.e. the vertical offset) is found slightly differently to before:

$$VO = \frac{B - A}{2}$$

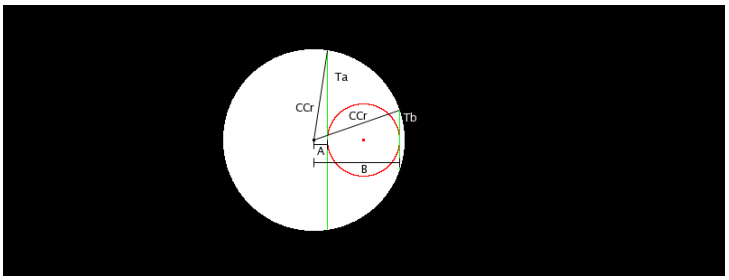
$$\Rightarrow VO = \frac{94.7418 - 17.2554}{2}$$

$$\Rightarrow VO = 38.7432$$

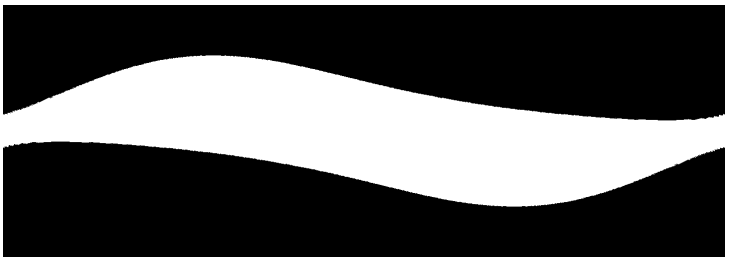
Again, the actual vertical offset was 40 pixels.

Offset circle does not fit entirely within the calibration circle and does overlap the centre of the calibration circle

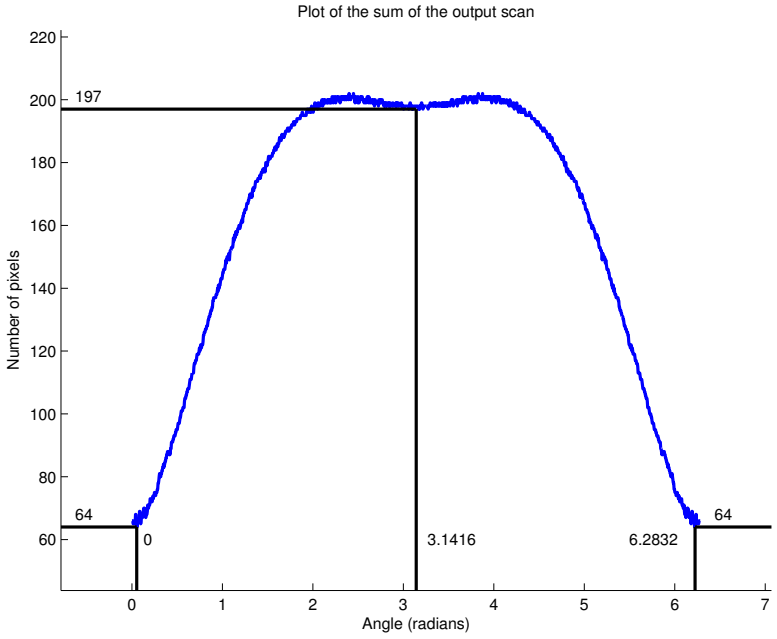
Figure 3.18(a) shows a calibration circle (white) with an offset circle (red) overlaid. This offset circle does not fit entirely within the calibration circle but it does overlap the centre of the calibration circle. Figure 3.18(b) shows what the output of this scan would look like. Figure 3.18(c) shows a plot of the sum of each column in the output scan. Since the offset circle overlaps the centre of the calibration circle, the plot of sums has no local minimum.



(a)



(b)



(c)

Figure 3.17: Example Scan 2. (a) Annotated calibration circle and offset circle (Cartesian space). (b) Output of the scan (θ-*tangent* space). (c) Plot of the sum of the output scan.

The tangent closest to the centre of the calibration circle, Ta, is found the same way as in Figure 3.16. The maximum value in the plot of sums is 195 pixels at 3.1416. In this case, the other tangent length required is perpendicular (rather than parallel) to Ta. This tangent, Tb, is therefore $\pm\frac{\pi}{2}$ radians away from Ta. Again from the plot of sums, this is 122 pixels.

Once Tb has been found, B can be calculated as follows:

$$CCr^2 = B^2 + (\frac{Tb}{2})^2$$

$$\Rightarrow B^2 = CCr^2 - (\frac{Tb}{2})^2$$

$$\Rightarrow B = \sqrt{CCr^2 - (\frac{Tb}{2})^2}$$

$$\Rightarrow B = \sqrt{100^2 - (\frac{122}{2})^2}$$

$$\Rightarrow B = 79.2401$$

In this scenario, B is the vertical offset and no further calculation is required. The actual value for the vertical offset in this case was 80.

Offset circle does not fit entirely within the calibration circle and does not overlap the centre of the calibration circle

Figure 3.19(a) shows a calibration circle (white) with an offset circle (red) overlaid. This offset circle does not fit entirely within the calibration circle and it does not overlap the centre of the calibration circle. Figure 3.19(b) shows what the output of this scan would look like. Figure 3.19(c) shows a plot of the sum of each column in the output scan. Since the offset circle does not overlap the centre of the calibration circle, the plot of sums has a local minimum, as in Figure 3.17.

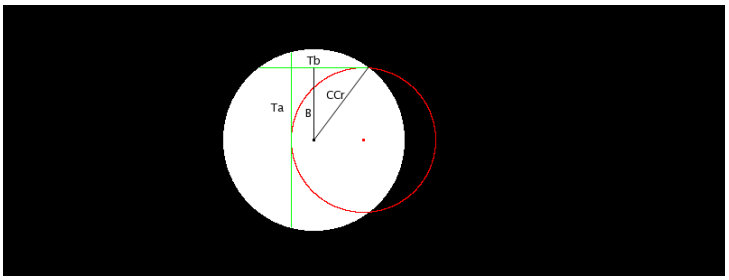
The tangent closest to the centre of the calibration circle, Ta, is found the same way as in Figure 3.17. The local minimum value in the plot of sums is 183 pixels at 3.1416. As in Figure 3.18, the other tangent length required is perpendicular (rather than parallel) to Ta. This tangent, Tb, is therefore $\pm\frac{\pi}{2}$ radians away from Ta. From the plot of sums, this is 123 pixels.

Once Tb has been found, B can be calculated as follows:

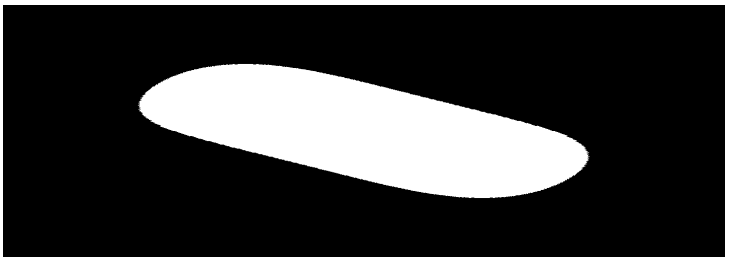
$$CCr^2 = B^2 + (\frac{Tb}{2})^2$$

$$\Rightarrow B^2 = CCr^2 - (\frac{Tb}{2})^2$$

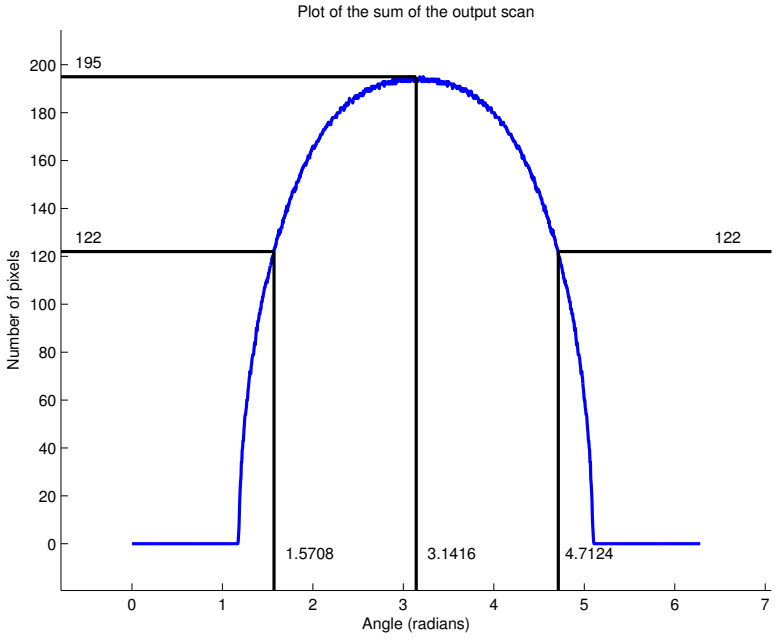
$$\Rightarrow B = \sqrt{CCr^2 - (\frac{Tb}{2})^2}$$



(a)



(b)



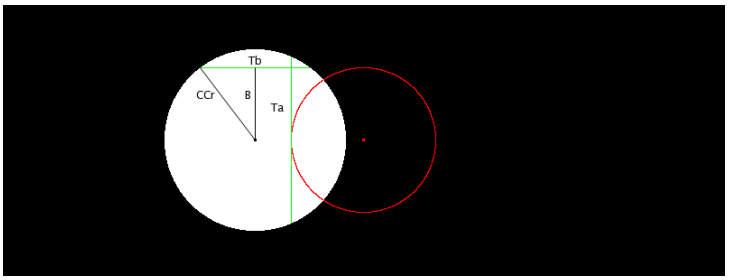
(c)

Figure 3.18: Example Scan 3. (a) Annotated calibration circle and offset circle (Cartesian space). (b) Output of the scan (θ-*tangent* space). (c) Plot of the sum of the output scan.

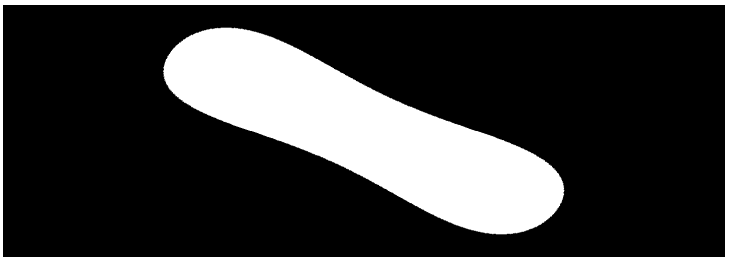
$$\Rightarrow B = \sqrt{100^2 - (\frac{123}{2})^2}$$

$$\Rightarrow B = 78.8527$$

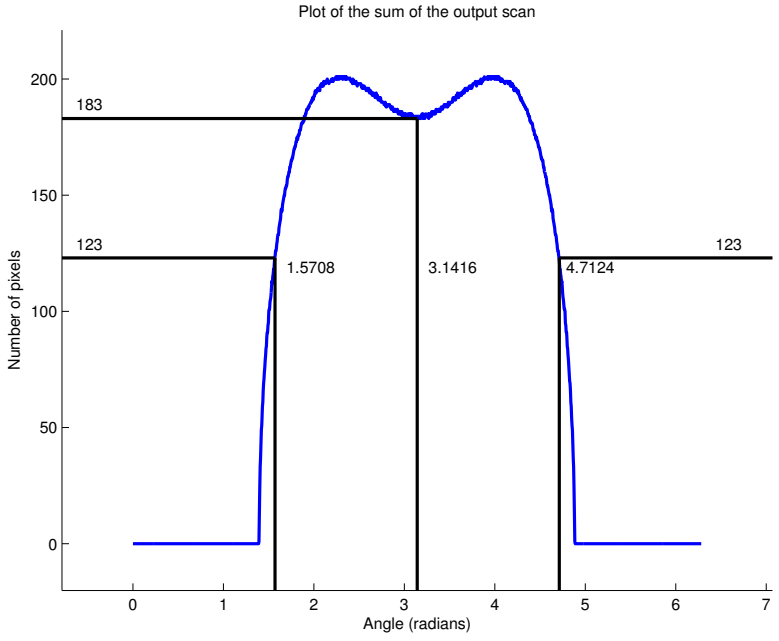
In this scenario, B is the vertical offset and no further calculation is required. The actual value for the vertical offset in this case was 80.



(a)



(b)



(c)

Figure 3.19: Example Scan 4. (a) Annotated calibration circle and offset circle (Cartesian space). (b) Output of the scan (θ-*tangent* space). (c) Plot of the sum of the output scan.

3.6.2.3 Calculating the Horizontal Offset

Calculating the horizontal offset follows a similar process to the one described in Section 3.6.1; the shift in the tangent from the centre is measured. The difference with an unaligned calibration circle, however, is that the tangent method length changes throughout the scan. To ensure that the horizontal offset is calculated correctly, the

tangent closest to the centre of the calibration circle must be used. If the vertical offset has been found, then this has already been determined (Ta in the examples in this chapter).

3.6.3 Errors in the Calculated Offsets

As was noted from the results of the offset calculations earlier in this section, the calculated offsets had varying degrees of accuracy. This was mainly caused by the fact that the calibration objects used were synthetic and not truly circular. Figure 3.20 shows how the exact radius of the circle varies at different points on the circumference if the shape is not a perfect circle. As a result of this, the plot of sums measurements in the previous section were effectively quantised. Using real calibration objects that are much closer to a perfect circle would result in significantly higher accuracies.

3.7 Interpolation

When converting to Cartesian representation, some interpolation is required as none of the pixels in the Cartesian representation will have an exact corresponding pixel in the scanned image. This is illustrated in Figure 3.21 where it can be seen that not only do the pixels from each representation not align, they are not even the same shape. This section explains the different interpolation techniques used to extract a Cartesian point from a rotationally scanned image.

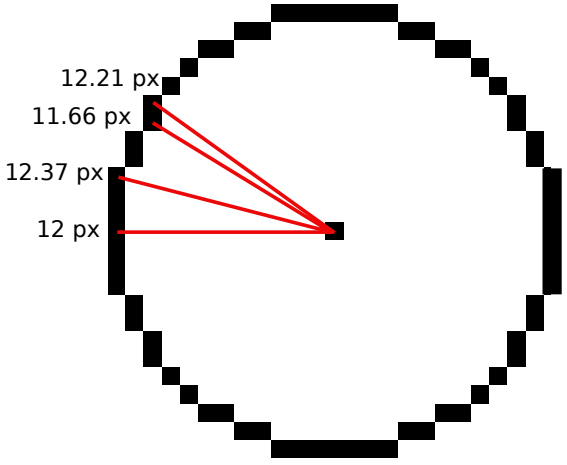


Figure 3.20: Imperfect circle showing different radius values

3.7.1 Nearest-Neighbour Interpolation

Nearest-neighbour interpolation can be used as a very simple method of determining the value of a point that lies between other known points. As the name implies, nearest-neighbour interpolation selects the closest known value. In the examples in the Section 3.5, this is achieved by simply rounding to the nearest whole pixel value. Figure 3.22(a) shows part of a scan reconstructed using nearest-neighbour interpolation. From the figure, it is clear that the checkerboard pattern has not been reconstructed accurately - the straight edges show signs of severe aliasing.

3.7.2 Bilinear interpolation

Bilinear interpolation differs from nearest-neighbour interpolation in that it uses the distances from the unknown point to the 4 nearest points to perform a weighted mean. Figure 3.22(b) shows part of a scan reconstructed using bilinear interpolation. By comparing Figures 3.22(a) and 3.22(b) it can be seen that bilinear interpolation produces smoother edges. Further information on bilinear interpolation can be found in [73].

3.7.3 Bicubic interpolation

Whilst bilinear interpolation considers 4 surrounding pixels to perform a weighted mean, bicubic interpolation extends this by considering a further 12 pixels (16 pixels in total). Figure 3.22(c) shows part of a scan reconstructed using bicubic interpolation. Again, the

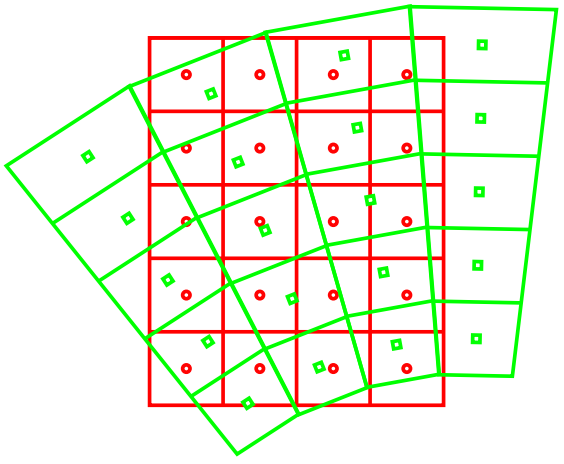
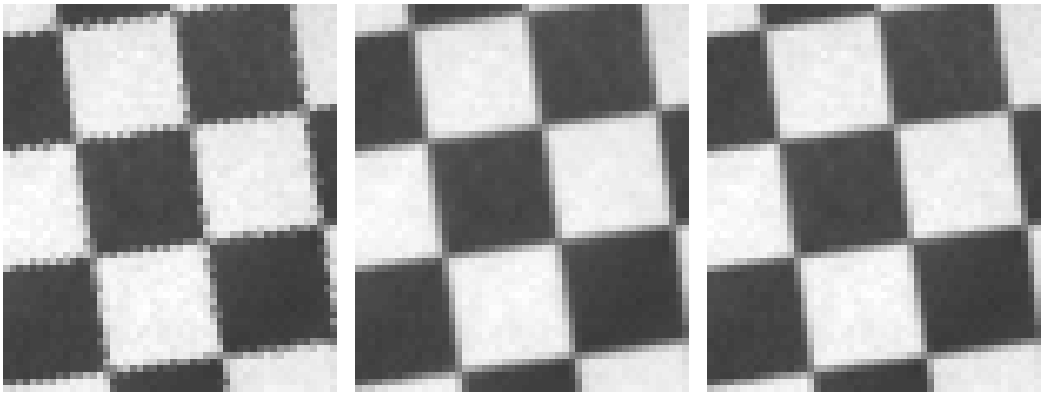


Figure 3.21: Cartesian (red) and scanned image (green) coordinate grids



(a) (b) (c)

Figure 3.22: Examples of three different types of interpolation on a scan of a chequerboard. (a) Nearest-neighbour interpolation. (b) Bilinear interpolation. (c) Bicubic interpolation.

result is smoother than when using nearest-neighbour interpolation, but offers almost no improvement over bilinear interpolation. It is also significantly more computationally intensive than both.

3.8 Quantitative Evaluation

To compare the effectiveness of using rotational scanning as an alternative method to linear scanning, three randomly generated binary images were printed out and imaged using both techniques. By comparing these scans with the originals, scores can be produced which indicate their similarity. The three random images used are shown in Figures 3.23(a)-(c). These images were generated by creating three random images where each pixel in each lies in the range [0,1]. A binary image was then created from each of these by thresholding at 0.5. Each of these binary images was then filtered three times using a median filter. Filter sizes of 35 x 35, 20 x 20 and 5 x 5 were used. These images are ideal for quantitative evaluations such as this one as they contain many different features of varying sizes.

Although the random images are binary, the camera used for the linear and rotational scans captures a greyscale image for each wavelength. As in other parts of this chapter, only a single wavelength band needed to be considered, as the effectiveness of the technique is not wavelength dependent. To convert a single wavelength band from greyscale to binary, a threshold must be set. Since the ratio of white pixels to dark pixels is

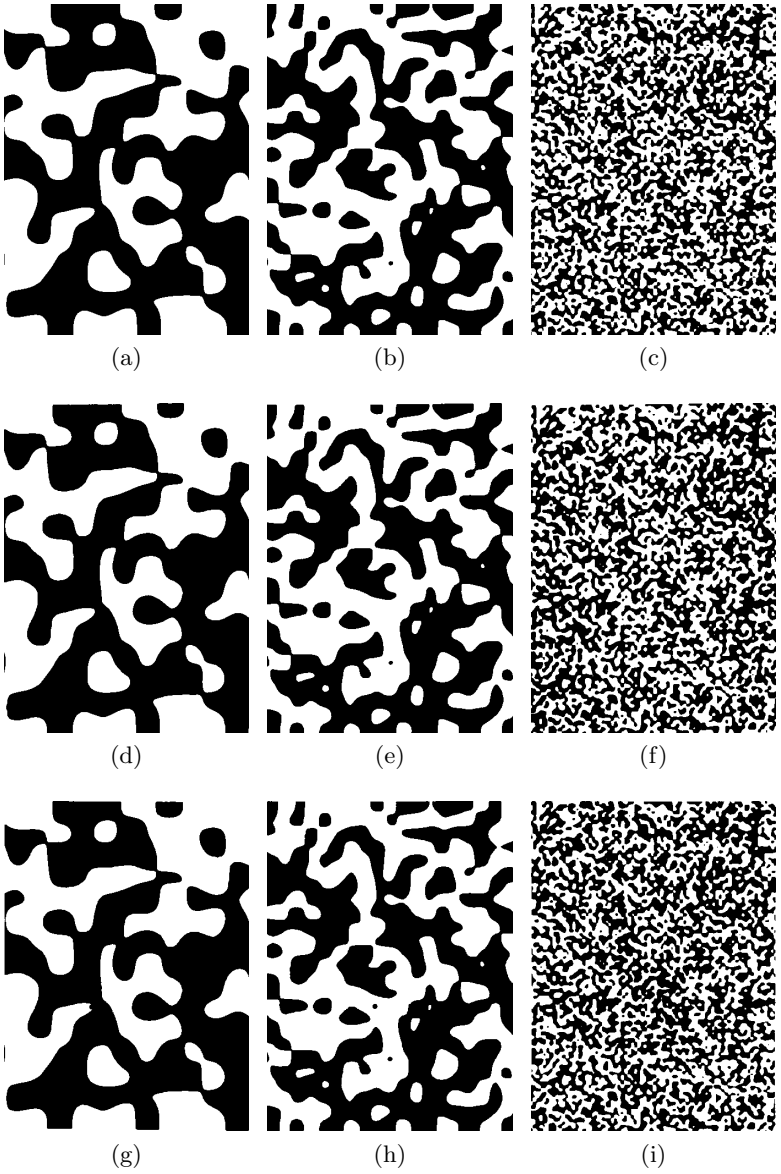


Figure 3.23: Random images for quantitative evaluation. (a) Random Image 1 (filter size 35 x 35). (b) Random Image 2 (filter size 20 x 20). (c) Random Image 3 (filter size 5 x 5). (d) Random Image 1 Linear Scan. (e) Random Image 2 Linear Scan. (f) Random Image 3 Linear Scan. (g) Random Image 1 Rotational Scan. (h) Random Image 2 Rotational Scan. (i) Random Image 3 Rotational Scan.

approximately 1:1 in each of the random images, setting the threshold to the mean of the greyscale image was appropriate.

Before the images could be compared, they had to be scaled to the same size and rotated to the same orientation. This was required as the ground truth images were in a digital format and the scale of the scanned print outs depended on both the printer and camera resolutions. The linear scans only had to be scaled as the orientation is simple to fix before acquisition. The scaling was performed using bicubic interpolation. The

rotational images had to be rotated before scaling. Since a rotation is simply a horizontal shift in the tangent length/angle coordinate space, no interpolation was required for the rotation. The same scaling technique used for the linear scans was then applied resulting in all 6 scanned images sharing the same size and orientation.

This was performed for each scanned image and can be seen in Figures 3.23(d)-(i).

3.8.1 Image Similarity Measures

Image similarity is a subjective measure, so the only 'correct' method of determining the similarity between two images is through human subjective evaluation [74]. Human subjective evaluation, however, is rarely an option, so there is a need for an objective measure of image similarity. A common technique is to use the mean squared error (MSE) where the dissimilarity is computed as the average of the squared intensity between the source image and the distorted image [75]. Another common technique for determining image similarity is the Structural SIMilarity Index (SSIM) which compares images based on their structural information and gives a score based on how much they differ [74]. SSIM operates on the assumption that human visual perception is suited to extracting a scene's structural information and therefore a good basis for determining image similarity. Correlation has also been shown to be a measure of how similar two images are to each other [76]. Other image similarity measures include techniques based on the discrete cosine transform [77], multi-scale wavelet decompositions [78] and feature-based sparse representation [79].

For this study, MSE, SSIM and correlation were chosen as three measures of image similarity. These techniques were chosen as they are simple to implement and are also effective on binary image problems. For MSE, a dissimilarity score is produced, i.e., the lower the better. For SSIM and correlation, a high score indicates high image similarity. For each technique, three measures were generated comparing the ground truth to the linear scan image, the ground truth to the rotational scan image and the linear scan image to the rational scan image. This was repeated for each of the three random images. The results are shown in Table 3.1.

From Table 3.1 the performance of the rotational scan is generally comparable with that of the linear scan. For Images 1 and 2, the rotational scan outperforms the linear

Table 3.1: Image similarity measures between synthetic images and rotational and linear scans

Image Reference	Image Check	MSE	Correlation	SSIM
Image 1 Original	Image 1 Linear Scan	0.045	0.910	0.811
Image 1 Original	Image 1 Rotational Scan	0.034	0.932	0.837
Image 2 Original	Image 2 Linear Scan	0.040	0.921	0.791
Image 2 Original	Image 2 Rotational Scan	0.032	0.937	0.807
Image 3 Original	Image 3 Linear Scan	0.163	0.675	0.375
Image 3 Original	Image 3 Rotational Scan	0.183	0.634	0.343
Image 1 Linear Scan	Image 1 Rotational Scan	0.055	0.889	0.803
Image 2 Linear Scan	Image 2 Rotational Scan	0.049	0.903	0.759
Image 3 Linear Scan	Image 3 Rotational Scan	0.185	0.629	0.310

scan in each similarity measure. For Image 3, the linear scan scores slightly better in each case. Since Image 3 contained smaller features, this might indicate that rotational scanning is better suited to scenes with large spatial features. To determine exactly how the size of the spatial features affects the performance, multiple images could be created for each filter size and many more filter sizes could be used (e.g. every second integer between 3 and 35). This would allow the results to be cross-validated allowing more robust conclusions to be drawn. The fact that each similarity measure ranked each test image in the same order, however, gives more credibility to the results presented here.

Another consideration when comparing linear and rotational scanning is the blind spot caused by VO in the rotational scan. For these similarity measures, the blind spot was ignored which will naturally degrade the performance somewhat. By only considering an area of the scan without the blind spot, it is expected that the results would further improve.

The angular resolution of the scan will also have an effect on the accuracy of the output image. Further tests could be performed with different angular resolutions to determine how much of a factor this is.

Each of the rotational images used in this quantitative evaluation were reconstructed using bilinear interpolation. Using a different interpolation method would have an impact on the results.

3.9 Other Considerations

There are a number of other factors which must be considered when using a rotational scan over a linear one.

3.9.1 Angular Resolution

The angular resolution of the scan will also have an effect on the quality of the reconstructed image. This angular resolution is determined by the scan speed; the slower the scan speed, the higher the angular resolution. Like linear scanning, changing the scan speed changes the pixel shape. However, the algorithm for reconstructing a rotational scan described in this chapter is capable of using any scan speed, therefore the only considerations when changing the scan speed are acquisition time and angular resolution. The width of a pixel in the scan can be approximated with

$$s = \theta r \tag{3.14}$$

where s is the pixel width, θ is the angle in radians and r is the distance from the axis of rotation. The angle θ can be determined by

$$\theta = \frac{2\pi}{n} \tag{3.15}$$

where n is the number of lines in the scan.

$$\Rightarrow s = \frac{2\pi r}{n} \tag{3.16}$$

Equation 3.16 is only an approximation as the distance s is the length of the arc that subtends angle θ. The pixel width can not be exactly represented by an arc length in this way. The size of the VO determines the accuracy of this approximation; a smaller VO gives a more accurate approximation for the pixel width. If a specific maximum pixel size is required, then Equation 3.16 can be used to approximate the minimum n needed.

3.9.2 Image Storage

As explained in Section 2.2, hypercubes captured using a linescan tend to be captured and saved in BIL format. This is still the case with a rotational line scan except the coordinate space is $(tangent\ length, \theta, \lambda)$ rather than (x, y, λ). The algorithm described in this chapter could either be used to convert the whole hypercube and save it in (x, y, λ) format or software could be developed to convert each wavelength band on-the-fly when viewing. All spectral (rather than spatial) processing techniques could be performed on the $(tangent\ length, \theta, \lambda)$ hypercube with no adverse effects. Spatial techniques would have different results, however there may be some applications where this could be beneficial such as in [80] where morphological operations are performed on images polar-logarithmic form.

3.9.3 Conversion

Using the conversion algorithm in this chapter has a non-negligible computation time. The solving of the Equation 3.10 is the most computationally expensive part. Rather than performing this for every pixel at every wavelength, a lookup table can be constructed for a single wavelength and then rapidly repeated for all others. Furthermore, this lookup table will show some rotational symmetry which could possibly be exploited, drastically reducing the overall computation time.

3.9.4 Regions of Interest

Due to the nature of a rotational scan, the resulting resolution of the converted image is much higher nearer the axis of rotation. Unfortunately, this is also where the blind spot is. If care is taken when aligning the camera, then specific areas of interest could be located towards (but not obscured by) the blind spot, improving the quality of the signal captured. The resolution at different points in the image can be approximated using Equation 3.16.

3.9.5 The Radon Transform

On first look, it appears that the rotational scans in this chapter could be manipulated in a similar way as the Radon transform [81]. The Radon transform, however, is used with a scan where the axis of rotation is perpendicular to the axis of the detector, similar to the other systems referred to in the introduction of this chapter.

3.10 Summary

In this chapter, an alternative to linear pushbroom scanning using rotational motion has been presented. An algorithm has been created to allow the reconstruction of a Cartesian image from a rotational line-scan. Two methods for measuring and compensating for the two sources of error in a rotational line-scan have also been explained. Bilinear interpolation was selected as the best interpolation method as it offered the best trade off between performance and computational complexity. This interpolation method could also be used in future to improve accuracy in the calculation of the offsets. Rotational scanning was compared to traditional linear pushbroom scanning using MSE, SSIM and correlation, and the results indicate that the two methods have a similar performance, demonstrating that rotational scanning is a viable alternative hyperspectral imaging line-scan method. Finally, some other important factors related to rotational scanning were discussed with some possible suggestions for future developments of a rotational scanning hyperspectral system.

Chapter 4

Supervised Classification Techniques in Hyperspectral Imaging

4.1 Introduction

Supervised classification is the process of assigning a new observation to one of a previously defined set of classes, based on training data whose class is already known [82]. Unsupervised classification, such as clustering, involves assigning a new observation to a class which is hitherto unknown. In this chapter, only supervised classification will be considered. Further information on unsupervised classification of hyperspectral data can be found in [33]. In a classification context, an observation is a set of quantifiable properties. With regards to hyperspectral data, an observation is usually a pixel, and the properties are normally spectral responses or principal components. Data with known classes (i.e. ground truth) is usually split into training and testing sets. The training set it used to train the classifier and the testing set is used to analyse its performance. This means that the testing set is not 'unknown' but its class is hidden from the classifier until performance evaluation. Adjustments can then be made before the classifier is applied to truly unknown data.

In this chapter, five different classification techniques are discussed. The operation of each classifier is explained with some synthetic data used where appropriate followed by

a discussion of some important factors that need to be considered during implementation. A classification problem using Chinese tea samples is then used to compare the implementation and performance of each. For each classifier, the adjustment of any parameters is discussed and evaluated where appropriate before the results are summarised.

The five classification techniques used are as follows:

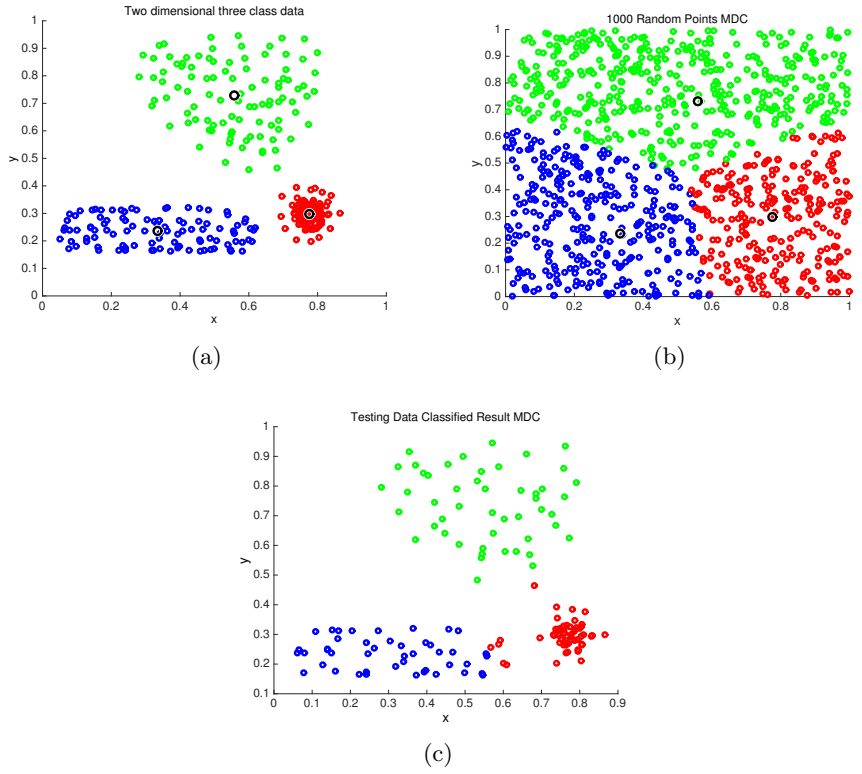
- Minimum Distance Classifier (MDC)
- Maximum Likelihood Classifier (MLC)
- Artificial Neural Network (ANN)
- Support Vector Machine (SVM)
- Adaptive Boosting (AdaBoost)

While only a small subset of the available classification techniques available, these five provide a good spread in terms of complexity and performance. They also form the basis for many other classification techniques.

4.2 Classifiers

4.2.1 Minimum Distance Classifier

The Minimum Distance Classifier (MDC) is one of the most simple classifiers available. Training data is used to determine the mean value of each class. Each testing/unknown point is assigned a class based on the Euclidean distance between it and each class mean [33]. The minimum distance (hence the name) determines the class to which it is assigned. Consider the red, green, and blue circles plotted in Figure 4.1(a). Each of these circles represents an observation where the quantifiable properties are x and y and each colour represents a different class; the data is two-dimensional and three class. The thicker black circles represent the mean of each class. 1000 random samples are then classified based on the class whose mean they are closest to (based on the Euclidean distance), shown in Figure 4.1(b). From this figure, it can be seen that



(a) (b)

(c)

Figure 4.1: MDC Classification (a) 2D 3 Class data set and class means (b) Classification of 1000 random points (c) Testing data classified

the classified points are distributed evenly, with each class having a similar shape and classification boundaries being generally straight lines. To evaluate the performance of this classifier, the data from Figure 4.1(a) is split into two equal sections; training data and testing data. The mean of each class is calculated from the training data, then used to classify the testing data. Figure 4.1(c) shows the result of this classification. Overall, the performance appears high, with only around 6 points classified incorrectly. The misclassification is caused by the classifier being unable to take account of the distribution of the data; the long thin nature of the blue class, the compactness of the red class and the spread of the green class. This example is only two-dimensional so not comparable to hyperspectral imaging, however, Euclidean distance can be calculated with multiple dimensions easily, so extending this to hyperspectral data is a trivial task.

Performance Factors

Speed and simplicity are the two main benefits of MDC. Shortly after data collection, classification can be performed. The simple nature of the classification means that no complex mathematical transforms are required, reducing the computational requirements. It can easily be used in a multi-class problem (such as the one above) without

modification, unlike binary classifiers. A simple threshold may also be set by requiring the minimum distance to be below a certain level before a sample can be assigned to a class. This is a quick way to remove unwanted (and spectrally different) background pixels. For example, consider a simple classification problem where a farmer wishes to know which parts of a field have diseased crops. The classifier could be trained on healthy and diseased crops and then a whole field could be classified. Other parts of the image, such as paths, roads or adjacent fields do not need to be classified, so using a threshold may prevent the classifier from labelling these regions. Without setting a threshold, every pixel in the image will be classified, producing potentially confusing results.

As can be seen from the synthetic example, MDC is unable to take account of the distribution of data. If the data is biased in any way, then the performance of MDC will suffer [33]. As a result, MLC is often only used as a first test before other more sophisticated classification techniques are employed.

4.2.2 Maximum Likelihood Classifier

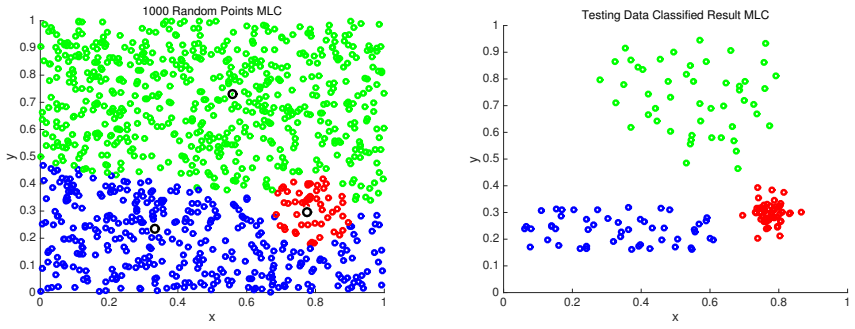
The Maximum Likelihood Classifier (MLC) is similar to MDC except that it uses the covariance of each class as training data, as well as the mean. Each unknown observation is given a likelihood score of belonging to each class based on the following equation [33]:

$$g_i(\mathbf{x}) = -\ln|\mathbf{C}_i| - (\mathbf{x}-\mathbf{m}_i)^T\mathbf{C}_i^{-1}(\mathbf{x}-\mathbf{m}_i) \tag{4.1}$$

Where g is the likelihood result, $_i$ is the class being tested, $\mathbf{C}_i$ is the covariance matrix for the $_ith$ class, $\mathbf{x}$ is the observation being tested and $\mathbf{m}_i$ is the mean of the $_ith$ class.

The class with the highest likelihood is used as the label for the unknown observation. Similar to the MDC, a threshold can be set so that only likelihoods above the threshold value are classified. Anything below the threshold value is left unclassified.

Using the same synthetic data as in the MDC example, 1000 random points are classified with an MLC, shown in Figure 4.2(a). Unlike in the MDC example (Figure 4.1(b)), each class has a different shape and the classification boundaries are not all straight lines. Using the same training and testing data as the MDC example, the testing data is



(a) (b)

Figure 4.2: MLC Classification (a) Classification of 1000 random points (b) Testing data classified

classified and the result is shown in Figure 4.2(b). Clearly, the MLC has performed well, correctly classifying every point.

If class probabilities are known before classification, then a more complex form of maximum likelihood classification can be performed [33]:

$$g_i(\mathbf{x}) = \ln p(\omega_i) - \frac{1}{2}\ln|\mathbf{C}_i| - \frac{1}{2}(\mathbf{x}-\mathbf{m}_i)^T\mathbf{C}_i^{-1}(\mathbf{x}-\mathbf{m}_i) \tag{4.2}$$

where $p(\omega_i)$ is the prior the probability of pixel $\mathbf{x}$ belonging to the $_ith$ class.

Performance Factors

Like MDC, speed and simplicity are significant advantages when compared to other more complex classifiers. By using the covariance matrix from each training class, MLC is able to take account of the distribution of the data, leading to improved classification results. As with MDC, multi-class classification is easy to implement.

Although MLC takes into account the distribution of the data, its classification accuracy is still lower than some of the other more popular (and more complex) classifiers discussed later in this section. Its simple nature, however, makes it an ideal candidate for use as a weak learner in a boosting situation, also described later in this section.

4.2.3 Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models based on the operation of the human central nervous system. The effectiveness of ANNs for solving classification problems has been demonstrated repeatedly since the early 1980s and their use in hyperspectral (specifically remote sensing) problems has steadily increased since 1989 [83].

Simply, ANNs consist of interconnected nodes (known as 'neurons') which compute a value based on some inputs. These neurons are arranged in layers consisting of an input layer, some hidden layers and an output layer, see Figure 4.3. The number of hidden layers is a parameter that must be set in advance. The neurons are joined together by connections, with each connection having a weighting. The values computed by neurons in one layer affect the ones in the next and so forth. When the network is trained, training data is presented to the input layer and the result at the output layer is compared to the desired result (ground truth). The weightings of the connections are then iteratively updated to improve the result at the output layer. This trained network can then be used to classify unknown input data.

An implementation of ANN would be as follows. For a given input vector $\mathbf{x} = (x_1, x_2, \ldots, x_d)^T$, the output of a single neuron, z, is determined as

$$z = g(\mathbf{w}^T\mathbf{x} - b) = g(\sum_{i=1}^{d} w_i x_i - b)$$

where $\mathbf{w} = (w_1, w_2, \ldots, w_d)^T$ denotes a weight vector and $g(.)$ refers to an activation function to decide whether the perceptron should fire or not. The sigmoid function $Sig(x) = (1 + e^{-x})^{-1}$ is the most popular activation function [84].

The same process used to compute the output of a single neuron can be used to calculate the output of the whole network in a topological manner. This means that for each neuron, its inputs from other neurons must be computed before determining its output. Consequently, the weight vector and the bias associated to each connection and each node will influence the results, and they are determined in training as follows.

Before training the ANN, the topology of the ANN must be specified, and in this case the feed-forward ANN will be used. A feed-forward ANN is a multi-layer perceptron (MLPP)

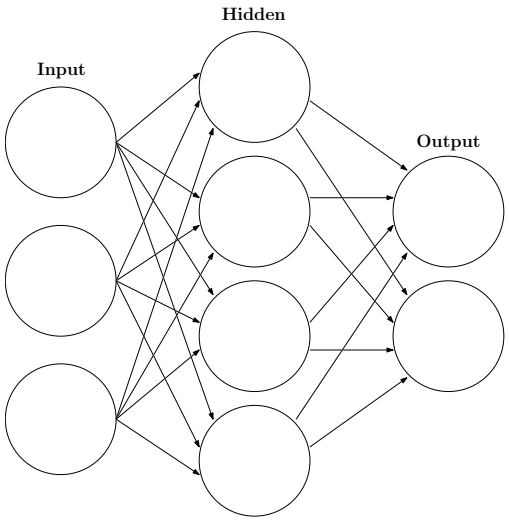


Figure 4.3: Diagrammatic representation of ANN.

which contains three or more layers of neurons, i.e. one input layer, one output layer and at least one hidden layer. With a given training set, a specified activation function and a learning ratio γ where $\gamma \in (0, 1)$, the learning process for supervised training using the well-known back-propagation algorithm can be described in the following three stages.

First, the initial weights are set randomly between $[-1, 1]$ to attain a group of output $\mathbf{z}^{(t)}$ at $t = 1$ referring to the first round of iteration. Then, an error function is decided as $\varepsilon(t) = \sum_{i=1}^{M}(y_i - z_i^{(t)})^2/2$ using the sum of squared error between the estimated output z and the target output y. Finally, the error signal at the output units is propagated backwards through the whole network to update the weights using the gradient descent rule

$$\Delta w_{ij}(t) = -\gamma \frac{\partial \varepsilon(t)}{\partial w_{ij}}$$

where w_{ij} refers to a weight between the jth node in a given layer and the ith node in the following layer. With updated weights, $t = t + 1$ can be set to start a new iteration until the network converges. This can be measured by using a small change ratio of $\varepsilon(.)$ or a given number of iterations.

Performance Factors

ANNs are popular due to their ability to generalise well in noisy environments and their high accuracy even when the feature subspace is complex or when the source data is comprised of different statistical distributions [85].

The main disadvantage of ANNs is the time taken to train the network. Once the network is trained, however, classification of more data is comparatively quick compared to other competing techniques. Another notable drawback is their reliance on random initial weights. This means that two implementations with the same parameters on the same data may produce different results, making it hard to compare with other techniques without several repeats.

4.2.4 Support Vector Machines

Based on an algorithm developed in 1963, Support Vector Machines (SVMs) as they are used today were developed in 1995 [86]. SVMs aim to classify data by finding the optimal hyperplane which separates the data into the correct classes, whilst maximising the margin between the hyperplane and the data. Figure 4.4(a) shows three hyperplanes (lines in this case) separating two-dimensional two-class data. H_1 fails to separate the data correctly while both H_2 and H_3 are successful. Looking at the margins, the distances from each hyperplane to the closest point in each class, it can be seen that H_3 separates the data with much larger margins. In this case, it is largest margin possible, therefore H_3 is the optimal hyperplane.

If the data is not linearly separable, a technique known as the 'kernel trick' can be used to map the data into a new feature space where the data is linearly separable [86]. Consider the data shown in Figure 4.4(b). Splitting the data correctly is not possible using a linear function. By transforming the data from one dimension to two dimensions, as shown in Figure 4.4(c), a straight line can now be used to separate the data. Figures 4.4(d) and 4.4(e) show two-dimensional data mapped into three dimensions, allowing a plane to separate the data. In both of these cases, a simple quadratic function is used to map the data into a higher dimension space where the data is linearly separable. Rarely is a problem this simple, however, and so a Gaussian radial basis function (RBF) is commonly used [33].

When using the RBF kernel, two parameters C and γ must be set for the SVM to operate. C stands for *cost* and determines how strongly to penalise misclassification. γ is a parameter for the RBF kernel which determines the width of the peaks in the higher dimensional space. The width of these peaks is inversely proportional to γ. Although

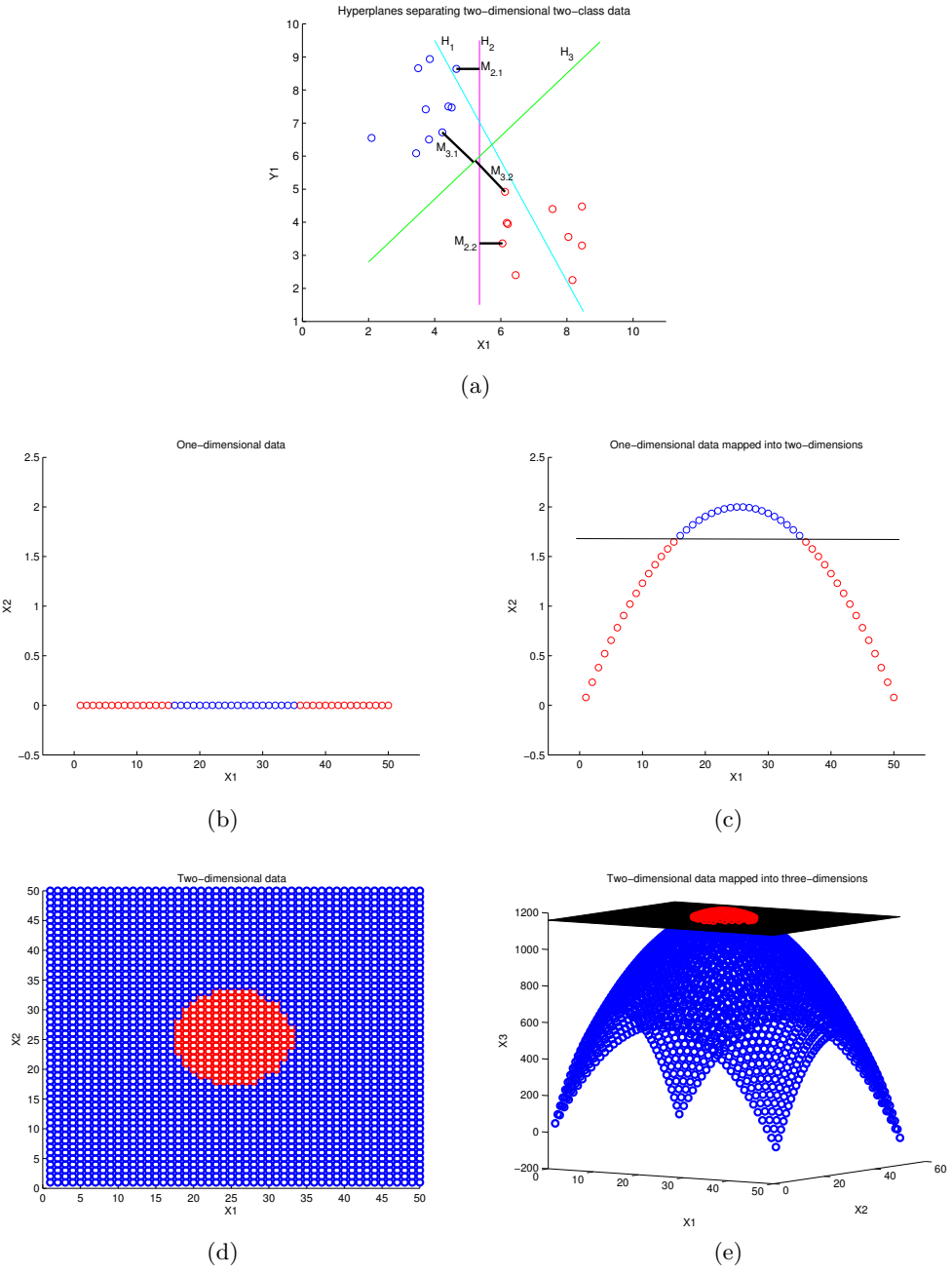


Figure 4.4: SVM classification (a) Hyperplanes separating the data (b) Non-linearly separable one-dimensional data (c) One-dimensional data linearly separable when mapped in two dimensions (d) Non-linearly separable two-dimensional data (e) Two-dimensional data linearly separable when mapped in three dimensions

these parameters are critical to the performance [87], they can be set accurately using a cross-validated grid-search technique removing the need for the user input.

Performance Factors

SVMs tend to achieve robust and accurate results with multidimensional capabilities and low requirements for training data. It is because of these features that SVMs are

one of the most popular classification methods [88]. They are by nature, however, binary classifiers. This is obviously a problem for multi-class data, but there are several methods to classify multi-class data with a binary classifier and these work well with SVMs [89]. Although the parameter selection can be performed automatically, this can significantly increase the time it takes to train the SVM.

4.2.5 AdaBoost

AdaBoost (short for 'Adaptive Boosting', [90]) is part of a subset of methods known as boosting, which is part of a larger subset known as ensemble methods. An ensemble method is a process where multiple learning algorithms are combined to improve the prediction accuracy to a rate higher than what any of the single algorithms could achieve on their own. AdaBoost does this by combining multiple weak learners (classifiers) using a weighted sum. The term 'adaptive' comes from the way the weak learners are adjusted with each iteration to favour the points which have been previously misclassified. The classification accuracy of each weak learner needs to be only slightly better than random guessing, so for a binary classifier this would be anything greater than 50%. In fact, since it is a binary classifier, anything less than 50% would also be acceptable as this result would just be inverted. Pseudocode for an implementation of AdaBoost is shown below in Algorithm 2.

Algorithm 2: AdaBoost Algorithm

Construct a training set $(x_1, y_1), \ldots, (x_m, y_m)$ where $x_i \in X, y_i \in Y = \{-1, +1\}$;

Set a number of iterations (number of weak learners), N;

Initialise the weight W_i to $\frac{1}{m}, i = 1, \ldots, m$;

for $n = 1$ **to** N **do**

- Select from the available weak classifiers, the classifier with the largest absolute difference between the weighted error rate, e_n, and 0.5, h_n;
- Find the weight of the classifier h_n, $\alpha_n = \frac{1}{2} \ln \frac{1-e_n}{e_n}$;
- Update the weight of each sample in the training set. If a sample has been correctly classified, decrease its weight. $W_i = W_i \frac{e_n}{1-e_n}$, where i is each correctly classified point;
- Normalise W such that it is a probability distribution (i.e., $\sum_{i=1}^{m} W_i = 1$);

The final classifier is then $H(x) = \sum_{n=1}^{N} \alpha_n h_n(x)$

In this study, AdaBoost is used with multiple instances of MLC as the weak learners, where each MLC is trained on a different subset of data. To use MLC as the weak learner, it needs to be capable of using weighted inputs. To achieve this, a weighted mean and weight covariance is used during the training process. The number of weak learners used is a parameter that must be set in advance.

Performance Factors

The main advantage of AdaBoost over many other classification techniques is its ability to use simple weak learners. Techniques such as SVM and ANN may be able to achieve higher accuracies in some instances, but they are much more complex algorithms. Since any weak learner can be used, AdaBoost is very versatile and can therefore be used in a wide range of problems. Despite its simplicity, it can be very computationally intensive compared to SVM and ANN, especially if a large number of weak learners are to be used.

4.3 Tea Classification

In this section, each of the techniques described earlier is implemented in a classification problem: the classification of five different kinds of Chinese Tea. The reasons tea samples were chosen for classification are three fold: firstly, this topic has rarely been investigated [44]; secondly, consumption of tea is a world-wide popular activity; and thirdly, tea price varies significantly in terms of brand and associated grades. Since conventional methods for tea analysis and classification need subjective input from domain experts, hyperspectral imaging provides a unique means for objective and consistent quality control even when the samples under examination appear similar to each other. This research shows how hyperspectral imaging could become an important tool for tea quality analysis and control, potentially identifying inferior or even fake products from established quality brands. It also serves as an ideal test to measure the performance of each of the classifiers described in this chapter.

4.3.1 Data Preparation and Pre-processing

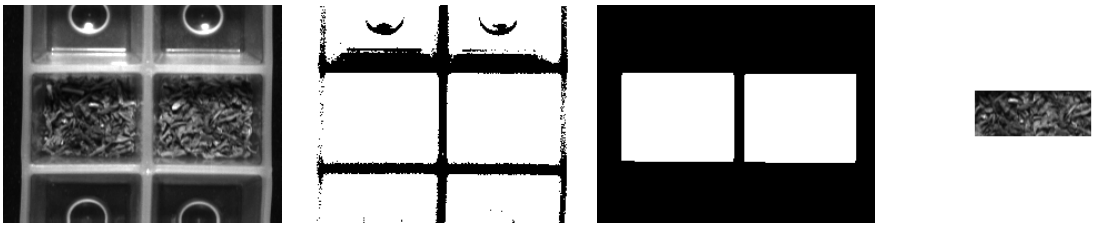
In total, five different types of tea were used in the experiments: China Black (Tea 1), Lung Ching (Tea 2), Tikuanyin (Tea 3), Yunnan (Tea 4) and Jasmine (Tea 5). All

the tea samples were imaged using an Andor Luca EMCCD camera with a Specim V8E spectrograph attached. Samples were placed on a Zolix KSA 11-200S4N motorised stage and illuminated with an Armley 150W Halogen lamp. This is a pushbroom system, similar to the one described in Section 2.2.8.

For each type of tea, two samples were placed into the compartments of an ice cube tray and imaged simultaneously to generate a hypercube. Since each tea was imaged twice, 10 hypercubes were produced from the five types of tea. Each hypercube was then scaled using a white and dark reference image so each pixel has a value in the range 0 to 1 (see Section 2.2.9). For each hypercube, there were 256 spectral bands, however not all of these bands contained information; some were just noise due to camera and the spectrograph not having the same spectral range. This redundant data in each hypercube was removed leaving 170 bands in the range 400 to 800 nm.

There were a substantial number of non-tea pixels in each of the acquired hypercubes, which were mainly from the container holding the tea samples whilst the data was captured, i.e. the ice cube tray. Therefore, these pixels were excluded so that only valid tea samples remained for classification. To this end, the hypercubes were cropped so that only the tea data was processed.

Rather than using manual image cropping, each hypercube was converted to a binary image and a morphological opening was performed with a rectangular structuring element of size 70 by 80 pixels, leaving just two white boxes of tea. The sizes and the locations of these white boxes were then used to crop the hypercube and valid tea samples were extracted as a region of interest in the centre of the cropped cube. The size of each region of interest used was 39 by 51 pixels. This process is shown in Figure 4.5.



(a) (b) (c) (d)

Figure 4.5: Pre-processing to extract regions of interest. (a) Original tea sample. (b)After thresholding. (c) Extracted masks. (d) Determined ROIs.

After applying the above cropping process to all 10 hypercubes where each contained two samples from the same tea, 20 different samples were produced. This resulted in 4 samples for each of the 5 teas. Images of these samples at the spectral band of 790 nm are shown in Figure 4.6. As can be seen at this spectral band, some of the tea samples appear quite similar to each other, such as Tea 3 (in the third row), Tea 4 (in the fourth row) and also Tea 1 (in the first row). This shows how difficult it is to classify these samples if only one grey level intensity is available, i.e. using a conventional method.

4.3.2 Feature Extraction and Selection

As explained in Section 2.3, a hypercube contains data derived from a range of contiguous spectral bands, resulting in a significant volume of highly correlated data, especially in adjacent bands. Processing this high volume of data is computationally expensive and due to the high correlation, potentially inefficient. Therefore, it is worthwhile to extract uncorrelated components as features from this highly correlated data for both efficiency and effectiveness. To achieve this, PCA was employed to reduce the data from 170 bands to 10 principal components for most of the analysis. The effect of the number of components retained is also explored. More information on PCA can be found in Section 2.3.1.

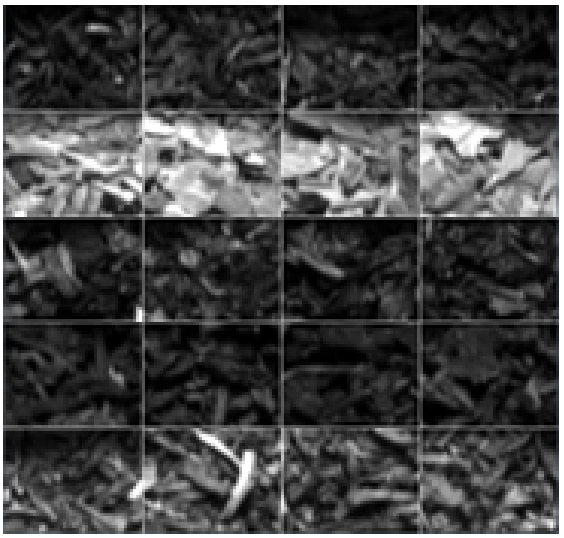


Figure 4.6: All 20 tea samples, each row contains 4 samples for one tea

4.3.3 Classification

Out of the five techniques described earlier, ANNs and AdaBoost have a parameter that must be chosen before classification. As explained earlier, the number of principal components retained is also a parameter that can be altered. Another parameter to be considered is the size (in pixels) of a square modal filter that will be used to post-process the results. The final parameter is the percentage of the data used for training (and therefore the percentage used for testing). If every combination of every parameter was tested, then the number of classification attempts would be huge. Instead, all but one parameter will be fixed while the other is altered to see its effects. This will then be repeated for all parameters. This forms the basis for all the classification attempts. The parameter settings are shown below in Table 4.1. MATLAB notation is used so 1:2:15 is every second integer between 1 and 15 (inclusive) and 25:25:75 is 25, 50 and 75, etc.

Table 4.1: Parameter Selection Settings

Test	Components	Filter Size	Hidden Layers	Weak Learners	Training Data
PCA	3:1:15	5	10	20	50%
Filter Size	10	1:2:15	10	20	50%
Hidden Layers	10	5	3:1:15	20	50%
Weak Learners	10	5	10	3:3:27	50%
Training Data	10	5	10	20	25:25:75%

For each of these tests, a plot will be produced showing how changing the parameter has an effect on the overall classification accuracy for each tea or classifier, as appropriate.

All of the results are cross-validated by taking different cubes of tea for training and testing data. For example, for a training percentage of 50%, two columns from the hypercube are selected as training data. The remaining two become the testing data. This is repeated for all combinations and the results are averaged together.

A cross-validated confusion matrix will then be produced for each classifier for a specific set of parameters. Confusion matrices are ideal for displaying data such as this as they show exactly which classes were misclassified and which class they were 'confused' with. Some pixel-based visual results are also produced where each class is assigned a colour. This shows exactly where the classifier has classified incorrectly.

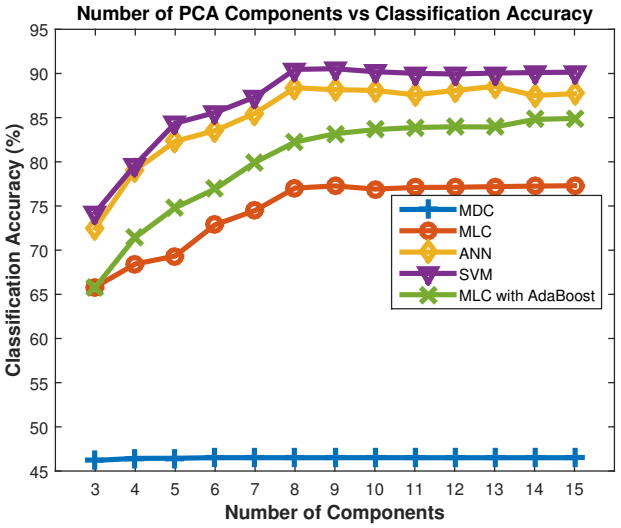


Figure 4.7: Number of PCA Components vs Classification Accuracy for all 5 Classifiers.

4.4 Results

This section presents the results of the experiments outlined in the previous section. The term 'accuracy' is used to describe the percentage of pixels which have been classified correctly, either for all the testing data pixels, or for the set of testing pixels for each individual class.

4.4.1 PCA

The overall effect of changing the number of principal components retained for each classifier is shown in Figure 4.7. From the plot, it is clear that increasing the number of principal components retained tends to improve the accuracy of the classifier. It is interesting to note that the improvement offered is not consistent between methods. MDC, for example, very quickly reaches a maximum classification accuracy using just 4 components. The accuracy of MLC with AdaBoost, however, continues to increase with number of components even when going from 14 to 15 components. Despite MLC being the weak learner used, MLC on its own reaches a peak accuracy around 8 components. This shows that the boosting algorithm is able to take advantage of the extra information provided by the less significant components whereas MLC cannot.

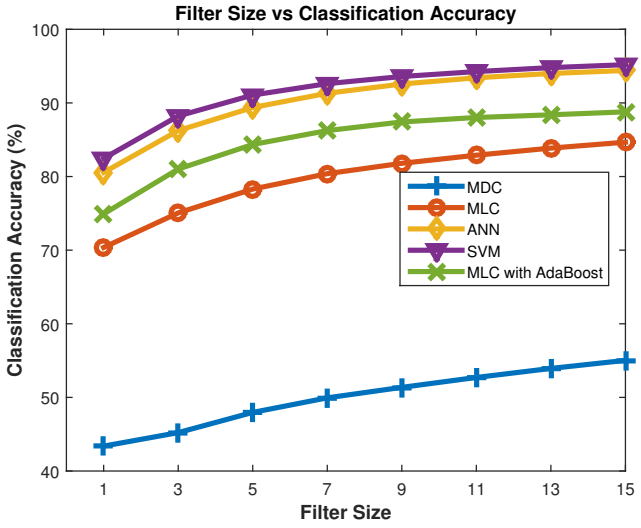


Figure 4.8: Filter Size vs Classification Accuracy Accuracy for all 5 Classifiers.

4.4.2 Filter Size

The overall effect of changing the modal filter size used in post-processing is shown in Figure 4.8. As might be expected, increasing the size of the modal filter increases the classification accuracy. Filter sizes bigger than 5 by 5 pixels continue to improve the overall accuracy, however, this results in a significant increase in computation time.

4.4.3 Hidden Layers

For the ANN classifier, changing the number of hidden layers had an effect on the classification accuracy. This is shown in Figure 4.9. For a hidden layer size of less than 5, the classification accuracy tends to suffer. The results, however, fluctuate significantly. This is likely caused by the fact that the initial random weighting which affects the final classification result. To achieve a more statistically significant result, the results would need to be averaged over a large number of classification attempts.

4.4.4 Weak Learners

The number of weak learners used for the MLC with AdaBoost classifier had a noticeable impact on the results as shown in Figure 4.10. In general, the classification accuracy increases for each class as the number of weak learners used increases. This improved accuracy, however, is relatively small when compared to the increase in computational

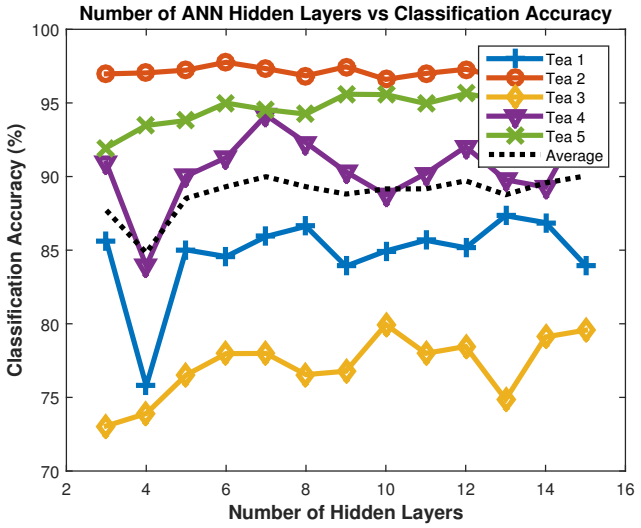


Figure 4.9: ANN Hidden Layer Size vs Classification Accuracy.

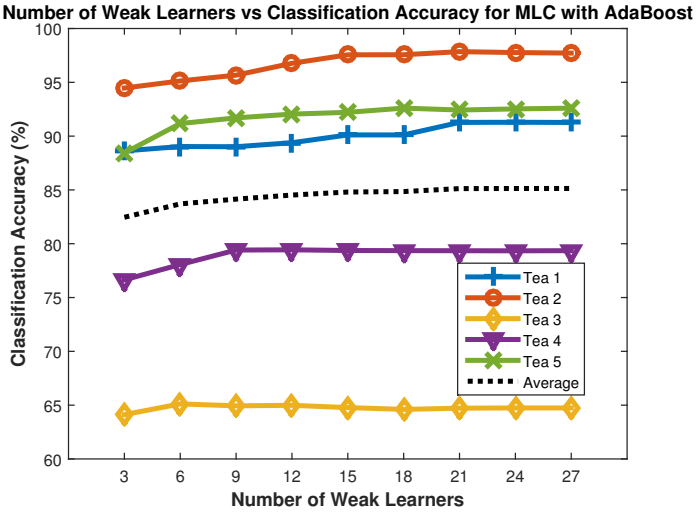


Figure 4.10: Number of Weak Learners vs Classification Accuracy for MLC with AdaBoost.

complexity. For example, the training and classification process took 77 seconds for 3 weak learners while it took 399 seconds for 15 weak learners and 639 seconds for 25 weak learners.

4.4.5 Training Data

The effects of altering the percentage of training data used are shown in Figure 4.11. Interestingly, increasing the percentage of data used for training did not result in a consistent increase in the classification accuracy. In fact, for SVM, ANN, and AdaBoost

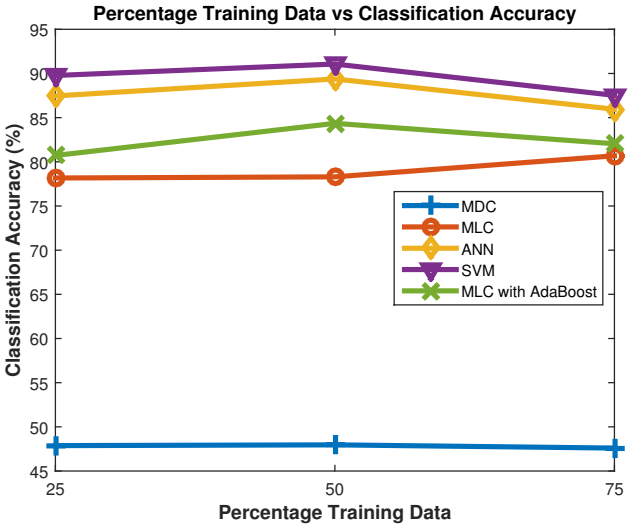


Figure 4.11: Training Data Percentage vs Classification Accuracy for all 5 Classifiers.

with MLC, 50% training data provided a higher accuracy than 75%. This demonstrates that the data is already well generalised by only 25% of the data.

4.4.6 Confusion Matrices and Visual Results

As explained earlier, confusions matrices are an effective way to show how the classifier is performing, specifically how accurate each individual class is. The confusions matrices in this section are generated using each classifier with the parameters set as the fixed ones in Table 4.1, i.e., 10 PCA components, a filter size of 5 by 5 pixels and a training percentage of 50%. For the ANN, 10 hidden layers are used and 20 weak learners are used for the MLC with AdaBoost.

Table 4.2 shows a confusion matrix for each classifier. From these it can be seen that Tea 3 tends to be the worst performer and is often confused with Tea 4. They also show that although the overall accuracy for MLC with AdaBoost is lower than both ANN and SVM, it is higher for Tea 1 and Tea 2.

Pixel-based visual results are generated by assigning each pixel a colour from a single classification result for each classifier without cross-validation. These visual results are shown in Figure 4.12. Each image here is a single classification result that contributed to the confusion matrix results earlier.

		Predicted				
		Tea 1	**Tea 2**	**Tea 3**	**Tea 4**	**Tea 5**
Actual	**Tea 1**	23.63 %	1.12 %	3.32 %	44.55 %	27.38 %
	Tea 2	3.14 %	78.44 %	1.52 %	1.28 %	15.63 %
	Tea 3	9.59 %	6.26 %	10.59 %	56.81 %	16.76 %
	Tea 4	1.71 %	0.00 %	14.48 %	78.03 %	5.77 %
	Tea 5	5.12 %	22.66 %	4.90 %	18.19 %	49.12 %

(a)

		Predicted				
		Tea 1	**Tea 2**	**Tea 3**	**Tea 4**	**Tea 5**
Actual	**Tea 1**	85.36 %	0.08 %	1.59 %	10.85 %	2.12 %
	Tea 2	0.00 %	95.82 %	0.00 %	0.00 %	4.18 %
	Tea 3	19.19 %	0.81 %	46.67 %	27.40 %	5.93 %
	Tea 4	18.82 %	0.02 %	2.66 %	77.16 %	1.33 %
	Tea 5	5.77 %	7.62 %	0.13 %	0.00 %	86.48 %

(b)

		Predicted				
		Tea 1	**Tea 2**	**Tea 3**	**Tea 4**	**Tea 5**
Actual	**Tea 1**	86.16 %	0.01 %	10.24 %	2.31 %	1.27 %
	Tea 2	0.01 %	97.07 %	0.00 %	0.06 %	2.85 %
	Tea 3	10.30 %	0.05 %	77.17 %	12.08 %	0.40 %
	Tea 4	2.23 %	0.00 %	7.36 %	90.41 %	0.00 %
	Tea 5	0.99 %	3.20 %	0.23 %	0.13 %	95.44 %

(c)

		Predicted				
		Tea 1	**Tea 2**	**Tea 3**	**Tea 4**	**Tea 5**
Actual	**Tea 1**	90.51 %	0.01 %	7.48 %	1.44 %	0.56 %
	Tea 2	0.00 %	97.45 %	0.06 %	0.00 %	2.48 %
	Tea 3	11.28 %	0.05 %	80.20 %	8.33 %	0.13 %
	Tea 4	1.93 %	0.00 %	6.04 %	92.02 %	0.00 %
	Tea 5	0.52 %	3.17 %	0.23 %	0.00 %	96.08 %

(d)

		Predicted				
		Tea 1	**Tea 2**	**Tea 3**	**Tea 4**	**Tea 5**
Actual	**Tea 1**	91.24 %	0.00 %	3.17 %	5.18 %	0.41 %
	Tea 2	0.10 %	97.68 %	0.10 %	0.00 %	2.12 %
	Tea 3	21.18 %	0.13 %	64.74 %	13.42 %	0.54 %
	Tea 4	19.19 %	0.00 %	1.44 %	79.34 %	0.03 %
	Tea 5	3.52 %	3.82 %	0.18 %	0.03 %	92.45 %

(e)

Table 4.2: Confusion Matrices (a) MDC (b) MLC (c) ANN (d) SVM (e) MLC with AdaBoost

Looking at these pixel-based visual results, it confirms that Tea 3 (the red class in these figures) is the most the most likely to be classified incorrectly and the cyan regions in the third row show where it has been misclassified as Tea 4.

4.4.7 Statistical Reliability

So far, only the classification percentage has been used as a measure of performance. This measure does not compensate for the fact that some pixels could be correctly classified due to random chance. Cohen's Kappa, κ, takes into account the probability of chance agreement and is calculated as follows:

$$\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$$

where $Pr(a)$ is the observed agreement and $Pr(e)$ is the probability of random agreement [91]. Table 4.3 summarises Cohen's Kappa for the classified results used to generate the confusion matrices in Table 4.2.

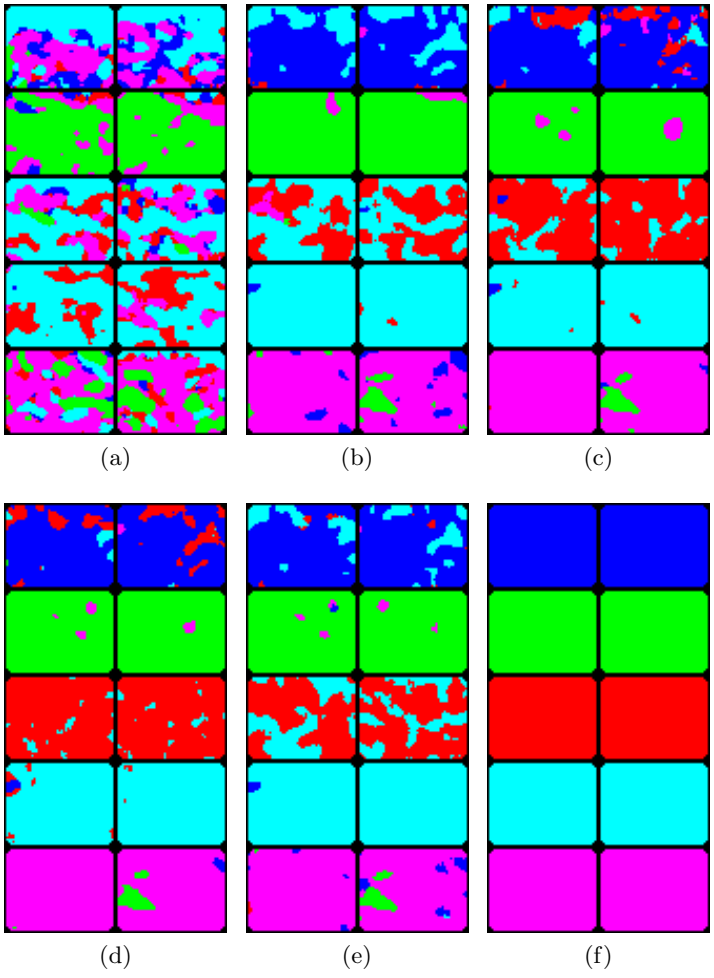


Figure 4.12: Pixel-based Classification Results. (a) MDC (b) MLC (c) ANN (d) SVM (e) MLC with AdaBoost (f) Ground Truth

Table 4.3: Cohen's Kappa for each Classifier

Classifier	Cohen's Kappa
MDC	0.35
MLC	0.73
ANN	0.87
SVM	0.89
MLC with AdaBoost	0.81

From the results, SVM exhibits the highest Kappa coefficient. This is unsurprising as it had the highest overall accuracy in all of the tests in this study. In fact, the ordering of the coefficients is identical to that of the overall classification accuracies. The poorest performer, MDC, has a significantly lower coefficient than the others. This suggests that many of the classified pixels could have been classified by random chance rather than any distinguishing features identified by the classifier.

Another factor influencing the results is the number of tea samples available for analysis. Although cross-validation was used, the statistical reliability is limited by there being only four samples of each tea available. If this experiment was to be repeated, many more samples could be obtained simply by shaking the ice cube tray to redistribute the tea leaves before image acquisition. With this extra data, more reliable results would be produced and error bars could be added to the plots showing the variance of the results.

4.5 Summary

Classification is an important tool used extensively in hyperspectral image analysis. In this chapter, five popular classification techniques were introduced and their operation explained. Each technique was then used in a classification problem where five different types of Chinese tea had to be identified.

The results demonstrate that hyperspectral imaging can be used to identify each of the five teas, but the success of the process is highly dependent on the classification algorithm used. Both post and pre-processing of the data is also crucial in extracting the best results.

Increasing the number of principal components retained tended to improve the classification accuracy, up until around 10 components. Retaining more than 10 components offered no improvement to the overall accuracy.

Performing modal filtering of the data after classification improved the results. Increasing the filter size improved accuracy up until around 9 by 9 pixels, after which it offered little improvement.

The number of hidden layers when using an ANN classifier had an effect on the classification result. Using fewer than 5 layers resulted in a poor performance. The peak performance of the classifier was when 7 hidden layers were used. It should be noted, however, that the initial random weightings of an ANN mean that these results are not necessarily fair comparisons. To overcome this, future experiments using ANN should be performed multiple times for each hidden layer size so the results can be averaged together.

The performance of the MLC with AdaBoost was improved as more weak learners were used. Adding these weak learners, however, was at the expense of increased computation time. Using 15 weak learners offered a reasonable trade-off between performance and complexity.

Increasing the percentage of training data used did not improve the classification results. This suggests that the data was already well generalised. In fact, in some cases, increasing the percentage of training data decreased the accuracy.

Chapter 5

Hyperspectral Analysis of Baked Sponges

5.1 Introduction

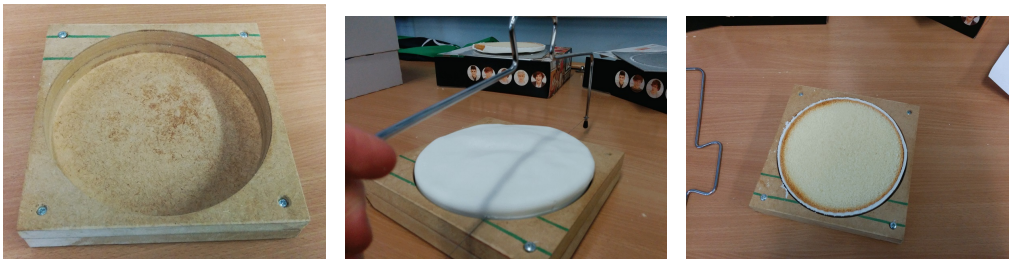
This chapter presents the results of an analysis of baked sponges using hyperspectral imaging. As noted in Section 2.4, food and drink analysis is a major area for hyperspectral imaging. The baking area in particular has seen a number of developments with hyperspectral imaging used to analyse bread, biscuits, cakes and other goods for the purposes of product innovation, reformulation and for making healthier alternatives [92].

The work presented in this chapter looks at how hyperspectral imaging can be utilised within the baking industry with a focus on the problem of assessing the shelf life of baked sponges and determining the distribution of moisture within them. Currently, organoleptic testing is used to predict the shelf life of baked sponges. The organoleptic tests are performed by a tasting panel, generally consisting of trained bakery staff, who eat ageing sponges on a daily basis and score them based on factors such as taste, texture and appearance. This process is repeated regularly using ageing sponges from the same batch until the panel determine that the cake is no longer suitable for consumption. The time taken for the sponge to reach this stage of deterioration can then be used to predict the shelf life of that product and to help ensure that when sold to consumers, sponges maintain their expected quality until the end of their shelf life.

Organoleptic testing as described above is used across the baking industry for the purposes of evaluating food quality. One drawback with this approach is the subjective nature of the test itself. Many external factors can affect a taster's perception of taste at the time of eating, drastically affecting the results. This makes it difficult to obtain an objective measure of how a sponge actually tastes at any point in time after it has been baked. Another major drawback is the limitation on the number of samples a single individual can taste in a single day. The result of this is that it is difficult to obtain statistically relevant sample sizes as increasing the number of samples tasted would decrease the accuracy of each tasting result.

In this study, an alternative and objective technique for analysing the eating quality of baked sponges of different ages with the aim of accurately predicting the shelf life of these products is presented. The technique uses a VNIR hyperspectral imaging system to acquire images of sponges ageing over a period of 40 days. By analysing the changes in measured reflectance as a function of time at a known water absorption band at 970 nm [93], it is possible to estimate the reduction in water content (drying out) of the sponge over time. The results of this analysis are compared with and shown to correlate with the tasting panel results obtained when tasting a different set of ageing sponges from the same batch. Furthermore, due the the nature of hyperspectral imaging, the distribution of moisture can be visualised. This moisture distribution information is used to train a classifier which can classify a sponge of unknown moisture distribution showing exactly where the sponge is drying out.

The remainder of this chapter is set out as follows: Section 5.2 describes a robust and repeatable technique for imaging sponges of different ages to ensure that the imaging process is performed consistently over the duration of the experiment. In Section 5.3 the techniques used to process and analyse the hyperspectral images to estimate the shelf life of the sponges are described and explained in detail. Also described is the process involved in using the spatial information from each hyperspectral image to give an indication of the distribution of moisture content throughout each sponge. Section 5.4 presents the results of analysing the data for one batch of cakes captured over a 40 day period and provides a comparison of the scores computed using the hyperspectral data with the scores allocated by a tasting panel. In Section 5.5, an explanation is provided as to how the results from this chapter could be used to implement a hyperspectral



(a) (b) (c)

Figure 5.1: The fixture used to ensure consistent sponge cut and location. (a) The empty fixture. (b) A sponge being cut. (c) A cut sponge.

sponge quality assessment system on a bakery production line. Section 5.6 provides some summarising remarks.

5.2 Data Acquisition

For the experiments described in this chapter, one vanilla (white) and one chocolate (brown) sponge were imaged each day (excluding weekends) over a 40 day period using a visible hyperspectral system resulting in 30 images of each type of sponge. Each sponge was taken from the same batch produced on the first day. The system consists of a Hamamatsu ORCA-05G camera with a Specim V10E spectrograph attached giving a spectral range of 400-1000 nm and resolution of 2.34 nm/pixel. This system captures hyperspectral image cubes using the popular pushbroom method explained in Section 2.2.4.3.

Sponges to be imaged were placed in a bespoke fixture and the top of the sponge was removed prior to imaging as shown in Figure 5.1.

The fixture was designed to match the diameter of the sponges with a depth of 30 mm. This was used to ensure that the distance between the camera and the surface of each sponge was consistent in every image captured. After removing the top of each sponge, the fixture containing the remaining section of the sponge to be imaged was positioned on a Zolix KSA 11-200S4N motorised stage and the area under the camera was illuminated with two 150 W halogen lamps. The bespoke fixture was constructed with locating screws on the underside to ensure that the position of the fixture, and hence the sponge contained within it, were always accurately positioned in the same location on the translational stage for each imaging session. It is known that the illumination

(a) (b)

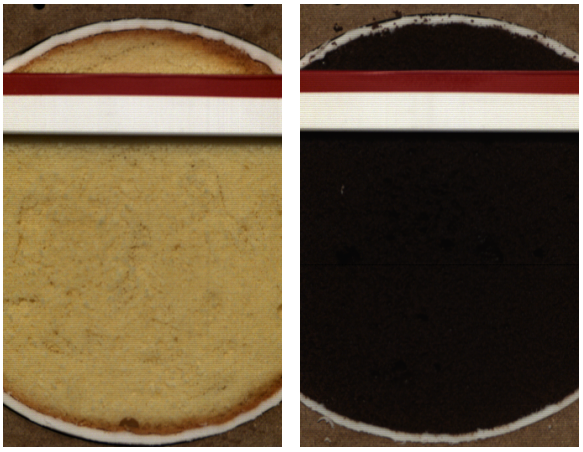
Figure 5.2: HSI system used for imaging the sponges. (a) Entire setup. (b) Close up of a sponge being imaged on the translation stage.

from the halogen lamp can be transient for around 30 minutes after switching on until it becomes stable. To ensure stability in the illumination source during data capture, the lamps were switched on one hour before imaging the sponges each day. The halogen lamp itself can cause the sponges to dry out due to the radiation which it emits. For this reason the system was set to capture an image of each sponge in 15 seconds. This resulted in captured hypercubes of dimensions $336 \times 765 \times 256$ pixels ($x \times y \times \lambda$). This time was optimal to ensure a high quality image of the scene was acquired in the shortest possible time thus minimising the effects of the illumination source on the experiment. An image of the entire set up while imaging a sponge is shown in Figure 5.2.

5.3 Data Analysis

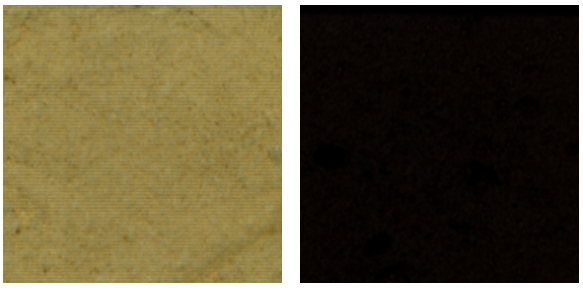
5.3.1 Pre-processing

Although care was taken to ensure illumination was as uniform as possible, the two lamps created a light source that was neither spatially nor spectrally uniform. Furthermore, the spectral response of the sensor was also non-uniform. To overcome this, 'white' and 'dark' references were used to correct each image. Each image captured contained a calibration tile, as shown in Figure 5.3. The images in Figure 5.3 are false colour images generated from three bands, as explained in Section 2.2.2. Cropping the image to just the white part of tile provided the white reference. The dark reference was acquired by capturing a small image while the lens was covered with a lens cap. The red tape on the tile was used as an extra way to check consistency between images.



(a) (b)

Figure 5.3: False colour images of the sponges on Day 1 before cropping. (a) Vanilla. (b) Chocolate.



(a) (b)

Figure 5.4: False colour images of the sponges on Day 1 after cropping. (a) Vanilla. (b) Chocolate.

The correction was performed as explained in Section 2.2.9.1.

After correction, each imaged was cropped to a 275×275 section in the centre of the sponge. An example image of each type sponge after pre-processing is shown in Figure 5.4 again, in false colour.

5.3.2 Observing Water Content

As stated earlier, a known water absorption band is at 970 nm. A sponge with a high moisture content is therefore expected to have a low reflectance (high absorbance) value at 970 nm compared to a sponge with a lower moisture content. Figure 5.5 shows a section of a vanilla sponge at 970 nm with a small amount of water placed on the left hand side. The sponge is clearly darker and thus more moist where the water has been added. Figure 5.6 shows the average spectral reflectance of this vanilla sponge across

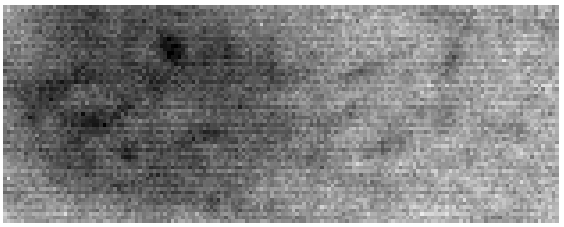


Figure 5.5: Vanilla sponge. Left - wet, right - dry.

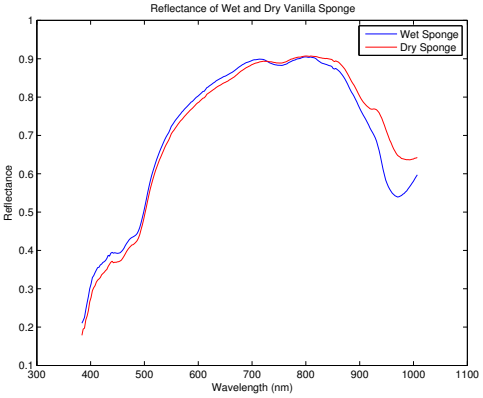


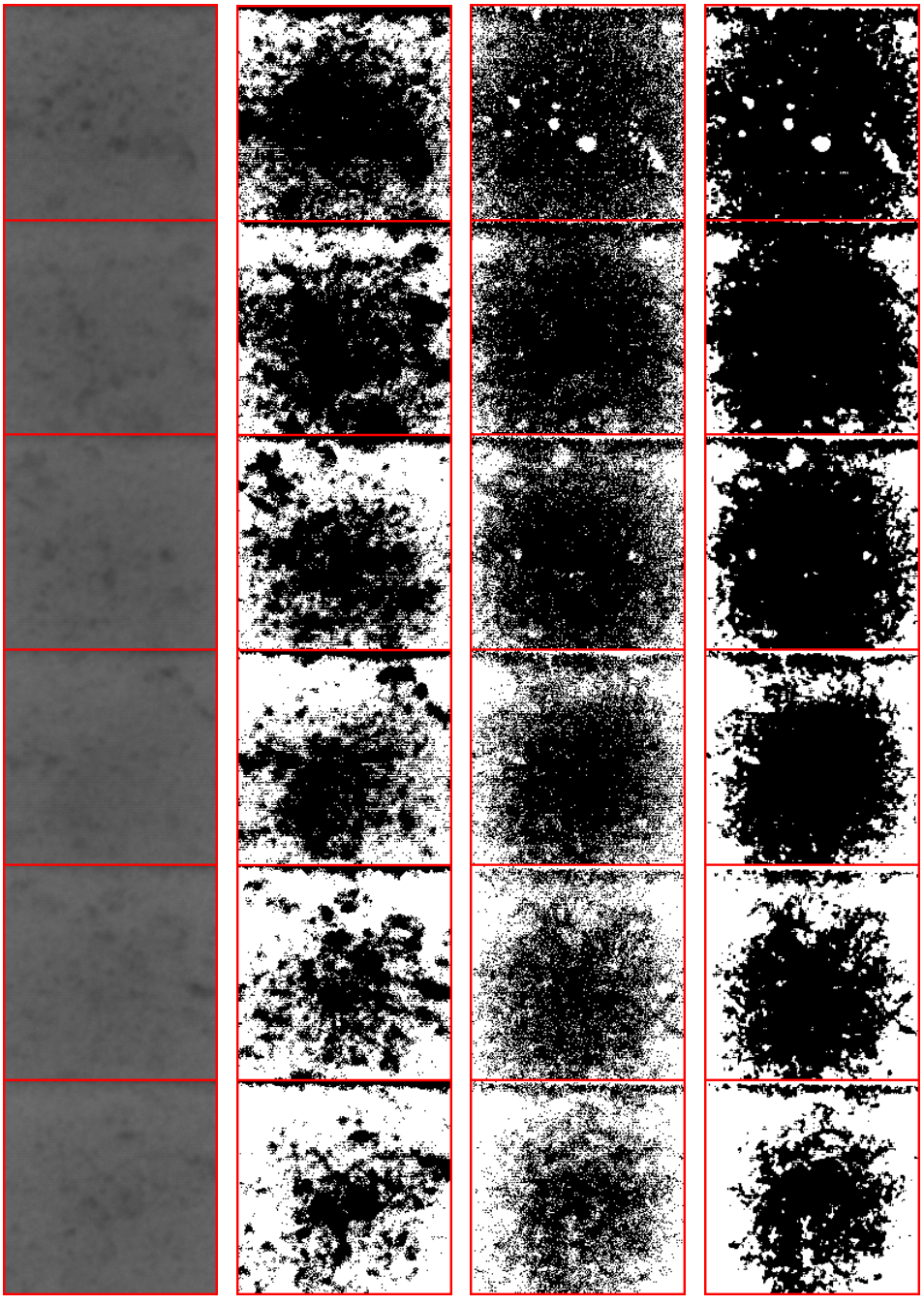
Figure 5.6: Reflectance values for wet and dry vanilla sponge.

the dry section and the wet section. Looking at reflectance values around 970 nm, it is easy to see that the reflectance is higher for the dry sponge, confirming that 970 nm is indeed a water absorption band and therefore suitable for measuring moisture content.

By inverting the reflectance at 970 nm and normalising for all sponges in an experiment, a score can be produced within the range 0 to 1, where 0 corresponds to the driest sponge and 1 corresponds to the most moist. Since sponges dry out with age, it is expected that this score will give an indication of the age of this sponge.

5.3.3 Spatial Information

Using the moisture content information explained in the previous section, it is possible to estimate the distribution of moisture throughout a spatial region of a hyperspectral image. Since the day-to-day variation between sponges can be significant, an average sponge image is created for each week, shown at 970 nm in Figure 5.7(a). Since the moisture content at each pixel is a function of pixel intensity, the freshest (most moist) average sponge (top) is the darkest and the oldest (bottom) average sponge is the brightest.



(a) (b) (c) (d)

Figure 5.7: Average sponge for each week. From top to bottom: Week 1 to Week 6. (a) Reflectance at 970 nm. (b) Reflectance at 970 nm thresholded. (c) MLC classified result. (d) MLC classified and filtered result.

The variance in pixel intensity in Figure 5.7(a) is relatively low, however, resulting in low contrast images. If each pixel can be considered as either 'dry' or 'moist', it should be possible to set a threshold which separates these classes. Since no other information about the moisture content of the sponges is known, a thresholded image of these average sponges should contain an equal number of black pixels as white pixels, i.e. the mean

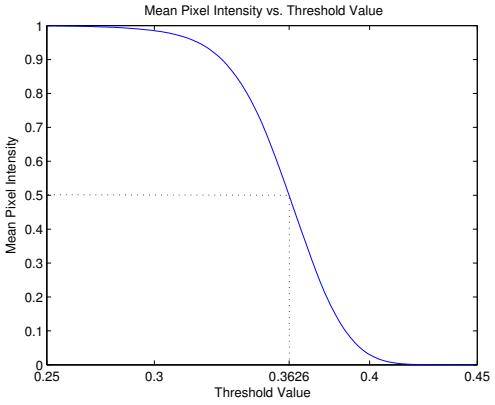


Figure 5.8: Mean Pixel Intensity at different threshold values

pixel value in the image is 0.5 (where 0 is black and 1 is white). To find this threshold value, the mean pixel intensity is plotted against threshold values from 0 to 1. See Figure 5.8. From the figure, the threshold value which achieves a mean pixel intensity of 0.5 is 0.3626. Applying this threshold to the average sponges for each week produces the image shown in Figure 5.7(b). The thresholded result for all sponge images at this value is shown in Figure 5.9 where the top left is the freshest sponge and the bottom right is the oldest sponge.

From the results in Figures 5.7 and 5.9, it is clear that each sponge has dried from the edge inwards. This information can be used to train a binary classifier by selecting regions of dry and moist sponge for training data, see Figure 5.10. The centre of the freshest sponge has been selected as 'moist' (blue) and the edges of the oldest sponge have been selected as 'dry' (green).

The thresholded result shown in Figure 5.9 only uses information from one spectral band (970 nm). It is expected that a more robust result could be obtained using more of the spectral information available. Rather than trying to process every spectral band, Principal Component Analysis (see Section 2.3.1) was employed to reduce the data from 256 wavelength bands to just 20 principal components.

The Maximum Likelihood Classifier (see Section 4.2.2) was chosen to classify the data as it is quick to implement whilst still providing a robust classification result. The speed of classification is important, especially if a system such as this one is to be used on a production line where results are required almost instantly. So while SVMs were shown to have a higher performance in Chapter 4, the higher speed of MLC made it most suitable for the classification presented here. For this study, two classes were used:

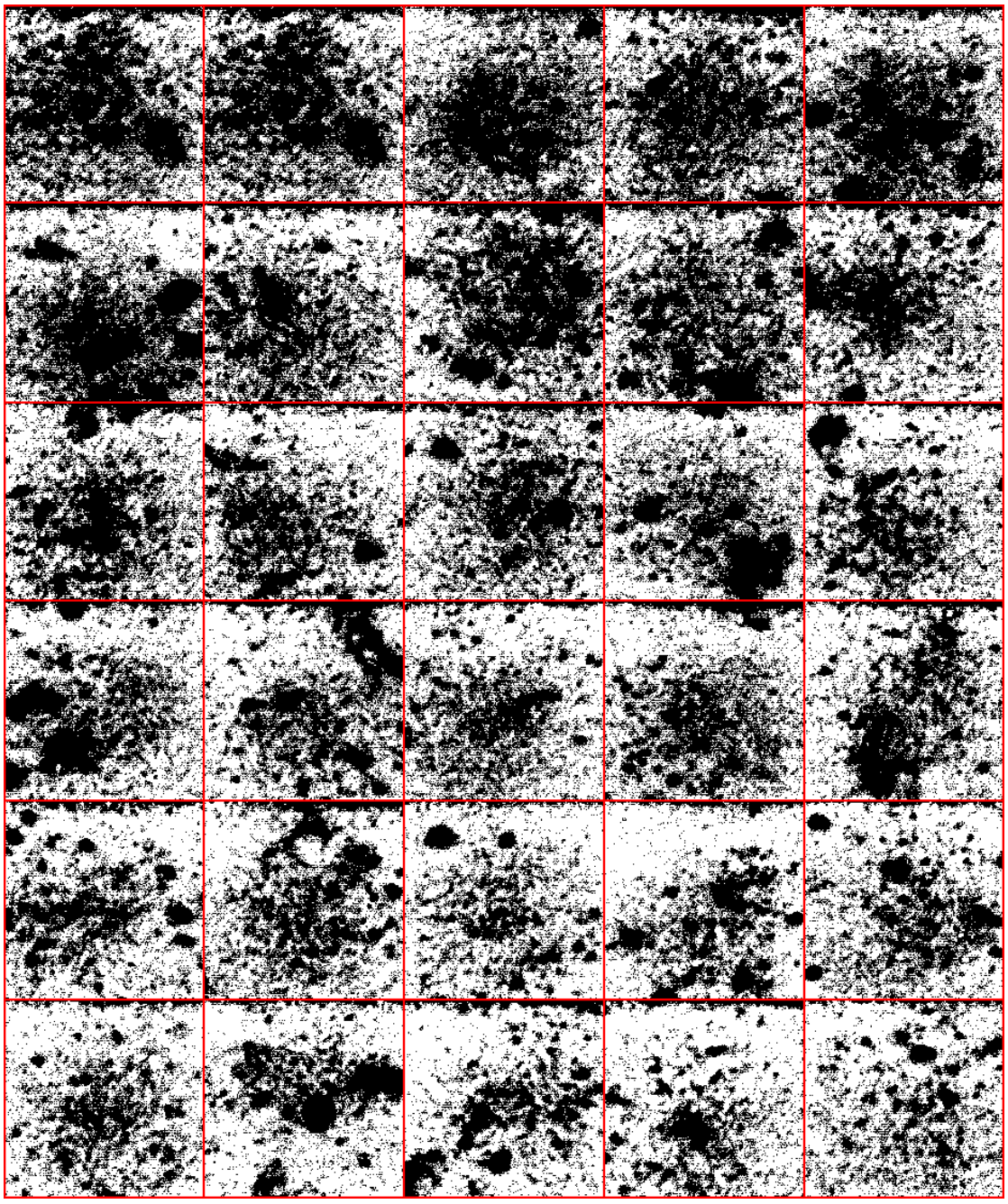


Figure 5.9: All sponges at 970 nm thresholded. Each row corresponds to a week and each column corresponds to a day of the week.

moist sponge and dry sponge. Training data was selected as described earlier. The results of the classification are discussed in Section 5.4.

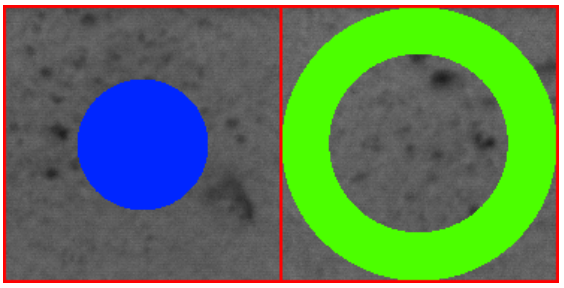
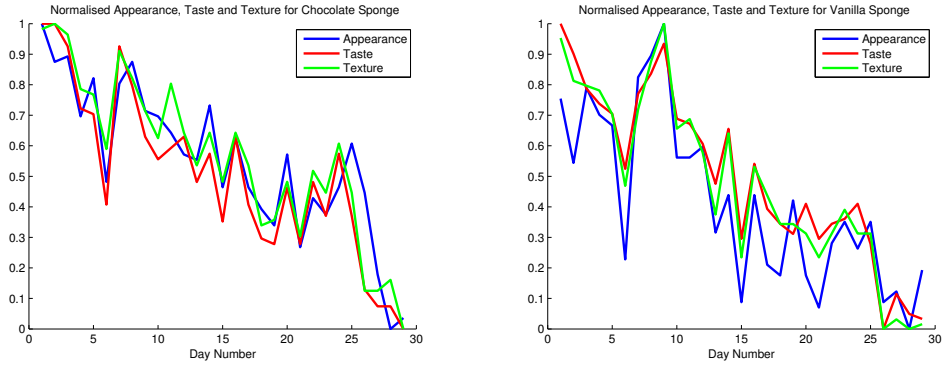


Figure 5.10: Training data selected from freshest sponge (left) and oldest sponge (right). Blue corresponds to 'moist' and green corresponds to 'dry'.



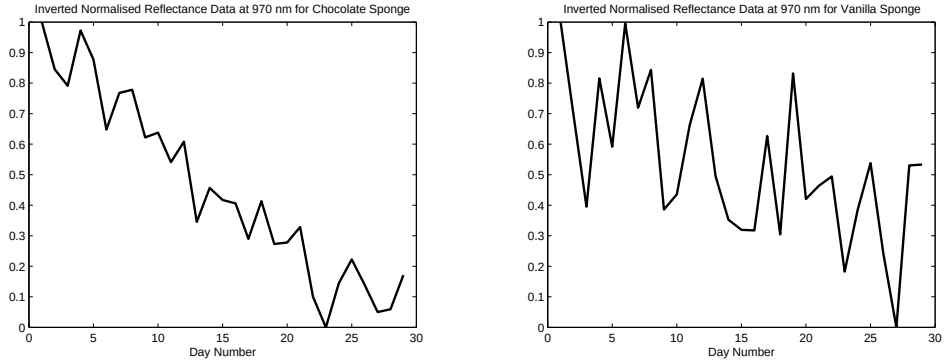
(a) (b)

Figure 5.11: Normalised mean organoleptic scores for Appearance, Taste and Texture. (a) Chocolate. (b) Vanilla.

5.4 Experiments and Results

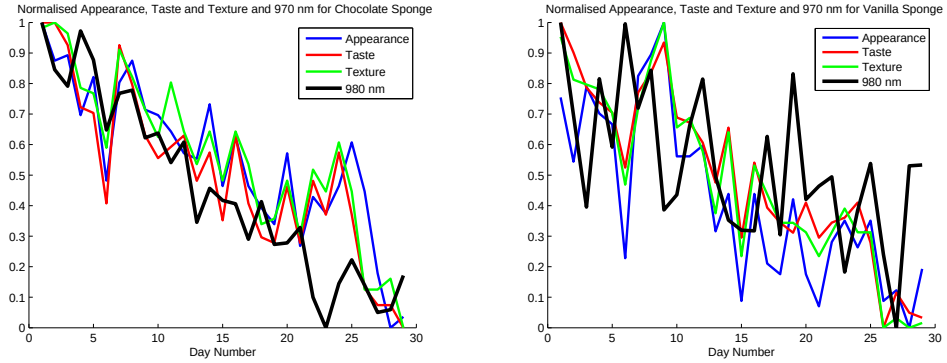
5.4.1 Organoleptic Results

A team of five tasting experts at the bakery analysed one chocolate sponge and one vanilla sponge each day over a 40 day period from the same batch as used in the imaging process explained in Section 5.2. Each sponge was scored out of 10 by each panel member for taste, texture and appearance. In order to make comparisons with the scores generated from the hyperspectral data, these scores were averaged and normalised, resulting in 3 scores (taste, texture, appearance) for each sponge on each day, with each score being within the range 0 to 1. This normalised mean response is shown in Figure 5.11. As expected, the scores tend to decrease over time.



(a) (b)

Figure 5.12: Normalised and inverted response at 970 nm. (a) Chocolate. (b) Vanilla.



(a) (b)

Figure 5.13: Normalised Taste, Texture and Appearance and 970 nm. (a) Chocolate. (b) Vanilla.

5.4.2 Hyperspectral Results

The inverted and normalised scores produced from the hyperspectral data for each type of sponge are shown in Figure 5.12. The chocolate sponge clearly exhibits a downward trend over time. The vanilla results are less conclusive, although the overall trend of the data is still slightly downwards. Figure 5.13 shows the hyperspectral results on the same axes as the tasting panel results. From here it can be seen that there is a strong correlation between results. The chocolate sponge produces more consistent scores both organoleptically and hyperspectrally compared to the vanilla results. This indicates that there may be more variation within a single batch of vanilla sponges than in a single batch of chocolate sponges. Indeed, further discussion with the bakery confirmed that they find there is more variation between vanilla sponges during the baking process.

5.4.3 Classification Results

As explained in Section 5.3, a classifier needs to be trained using training data before its performance can be evaluated using testing data. Using the results of the thresholding at 970 nm, the edge of the oldest sponge was labelled as dry (green) and the centre of the freshest sponge was labelled as moist (blue), as in Figure 5.10. All sponges were used as testing data and an average sponge was produced for each week. These average sponge classification results are shown in Figure 5.7(c). To further improve the classification result, each average sponge was filtered using a median filter with window size 3x3 pixels, as seen in Figure 5.7(d). Comparing with the other results in Figure 5.7, the improvement offered by classification over simple thresholding and also the improvement offered by median filtering can be seen.

5.5 Developing a System

The results from the previous section clearly indicate that hyperspectral information closely correlates both with the organoleptic quality scores and the expected moisture content distribution of a baked sponge. If a sponge of unknown quality and moisture distribution were to be imaged using a similar set up to the one described in this chapter, a score could be produced giving an estimate of its eating quality. Furthermore, this sponge image could be classified using the already trained MLC and the predicted distribution of moisture content could be visualised. Currently, the bakery uses a human tasting panel to analyse a very small subset of all sponges produced. This is subject to a number of drawbacks, identified in Section 5.1. With the proposed hyperspectral system, significantly more sponges could be analysed. In fact, due to the production line nature of the baking process, every sponge produced could be imaged and analysed, providing an unprecedented level of information about the baking process. The effects of any changes in the baking process (e.g. new flour supplier) may not be apparent under existing methods, so hyperspectral imaging could provide a means of objectively quantifying any differences making it easier test new methods which may improve the quality of sponges produced.

The bakery has recently acquired a moisture meter which provides an accurate percentage moisture content for a small sponge sample ($<10\ g$) by weighing it before and after

drying it out. Like the organoleptic testing, this test is also destructive so cannot be used to analyse any more than a small subset of all sponges produced. In future work, however, this moisture data could be used to verify and improve the performance of the hyperspectral analysis.

Since using the full spectral data with the MLC produced better results than simple thresholding at 970 nm, it is likely that more accurate scores could be produced for the analysis in Section 5.4.2. If only 970 nm data is used, however, then a hyperspectral system is not required. A single 970 nm band pass filter could be used to achieve the same results. Obviously, this would limit the spatial information to the simple thresholded result.

5.6 Summary

In this chapter, a method for using hyperspectral imaging to determine the quality of a baked sponge is presented. Two batches of sponges (chocolate and vanilla flavour) were used to analyse how the age of the sponge affects its quality. The ground truth measurement was taken from an expert panel of tasters, who tested one of each sponge each day. From the same batches of sponges, one of each sponge was imaged using an VNIR hyperspectral system. By calibrating and correcting every image, an average score relating to sponge moisture content was produced for each sponge. The results show that these scores are strongly correlated with the organoleptic measurements from the human tasters. Furthermore, a classifier was trained to identify moist and dry regions of a sponge image. This allows an unknown sponge to be imaged and an estimation of its moisture distribution to be visualised. With the techniques proposed in this chapter, a system could be developed to allow every sponge produced to be imaged and analysed, allowing the bakery to see the measured quality of each sponge, not a just subset analysed by human tasters. With this information, the baking process can be modified and any changes in the quality of the sponges can be immediately quantified.

Chapter 6

Spatial, Spectral and Colour Information - A Quantitative Analysis

6.1 Introduction

The aim of the research presented in this chapter is to quantify the improvement offered by HSI over standard imaging techniques (RGB and greyscale) and standard spectroscopic techniques. A simple task of classifying four different kinds of rice is used to measure the performance of each technique. Rice was chosen as there were several different kinds readily available and each variety showed some differences in terms of shape, size, colour and, one would expect, spectral response.

SWIR HSI offers clear advantages over conventional imaging due the rich chemical information available, but the high equipment costs often prohibit its use. VNIR HSI systems, however, costs significantly less allowing the potential of HSI to be realised in many more situations. The advantage of VNIR HSI over RGB or greyscale imaging is less apparent. For this reason, VNIR HSI is the considered technology of interest in this chapter.

To show the improvement HSI offers over standard imaging, hyperspectral data is compared with true colour images generated from the same dataset. The process to produce

these true colour images is the same as explained in Section 2.2.3. The comparison with spectroscopic techniques is achieved by simply considering only the spectral response of each pixel and ignoring any spatial information.

Spatial information is obtained from the images using standard image processing techniques, explained later in this chapter. Thus, from a single hyperspectral image, we can produce five datasets for comparison:

- Spatial information only
- Colour information only
- Spectral information only
- Spatial and colour information
- Spatial and spectral information

To compare these five data sets, an SVM classifier is trained using a sub-sample of each kind of rice. While any of the classification techniques explained in Chapter 4 could have been used, SVMs were used as they consistently had the highest accuracy. Since the datasets are relatively small (compared to the tea images in Chapter 4 and the sponge images in Chapter 5), the classification speed is not so critical. The performance of the SVM is evaluated against a different sub-sample, giving a classification accuracy as a percentage. This is repeated for each dataset, and the best performance will be indicated by the highest classification accuracy. To ensure reliable conclusions can be drawn, this classification process is repeated over a number of different sub-samples for each dataset, cross-validating the results.

In [30, 94], several techniques are described for extracting spatial features and combining them with spectral information for classification. These techniques, however, focussed on remote sensing data so are not applicable to the rice grain analysis presented in this chapter. Colour is determined to be an important attribute for consumers with regards to food in [95] and in [96], dates are classified using an SVM based on colour, texture and shape. The four different types of date used in [96] are clearly visually similar, unlike the rice grains used in this chapter.

The remainder of this chapter is laid out as follows: firstly, the different types of rice are described along with justification for their inclusion, the imaging process is then explained in detail along with some example images, followed by the steps involved in pre-processing all the captured data. Then the classification process is described. The classification results are then presented and summarised before some conclusions are drawn.

6.2 Rice Types

As outlined in the introduction to this chapter, rice was chosen as the subject for this experiment for a number of reasons:

- Several kinds readily available at low cost
- Grain shape variance between types
- Grain size variance between types
- Grain colour variance between types
- Expected spectral variances between types

Using the tea leaves from Chapter 4 or the sponges from Chapter 5 was considered for this study, but the spatial information in each of these cases was difficult to determine (tea leaves) or almost non-existent (sponges).

The four types chosen, shown in Figure 6.1, are Long Grain and Wild Rice, Basmati Rice, Pudding Rice and Value Long Grain Rice. For the rest of this chapter, these will be referred to as Rice 1, Rice 2, Rice 3, and Rice 4, respectively.

From visual inspection of each type, the following conclusions could be immediately drawn:

- Rice 1 is made up of black and white grains
- Rice 3 has generally more rounded shape than the others
- Rice 4 has a more inconsistent size due to the number of broken grains

Figure 6.1: The four kinds of rice. From left to right: Basmati, Long Grain and Wild, Pudding Rice, Value Long Grain

Since Rice 1 consists of grains which are one of two different colours, all of the black colour grains were discarded. If they had been left in, then the problem would essentially have become five-class, with one class being clearly different from the rest offering little improvement to any conclusions which could be drawn.

6.3 Imaging

Since accurate spatial information is critical in this study, the grains of rice were imaged in a grid formation with no two grains touching each other. A well-plate was used to hold each grain in a specific place. By isolating each grain of rice like this, the required image processing task of identifying each grain and its spatial properties should become significantly easier since it is much easier to distinguish a grain of rice from the background compared to distinguishing one grain of rice from another when they are touching.

Each type of rice was imaged separately in a grid of 72 formation. Then a final image consisting of 18 grains of each type was imaged. Each rice type in this case corresponded to three rows in the grid. This resulted in a total of five hypercubes and a total of 360 individual grains imaged.

The imaging set-up was the one described in Section 2.2.8, and the imaging was performed in 4x4 binning mode, resulting in 336 pixels of 256 wavelength bands for each line in each image. Using a white calibration tile, each image was corrected, as in Section 2.2.9.1. A mercury calibration lamp was also used to calibrate each pixel to a specific wavelength, as in Section 2.2.9.2.

Figure 6.2 shows a true colour representation (explained in Section 2.2.3) of each image captured after correction and calibration, but before any other pre-processing. The white bar at the top of each image is the calibration tile. Since the true colour image consists of three channels - red, green and blue - the non-spatial data therefore consists of all 256 spectral features and 3 colour features.

6.4 Data pre-processing

6.4.1 Mask Creation

From looking at the images in Figure 6.2, it is clear that much of each image consists of pixels which are not rice. The first step in pre-processing is therefore to remove all non-rice pixels. To achieve this, a mask must be created for each image which consists of a binary image where a one represents a rice pixel and a zero represents a non-rice pixel. Creation of these masks was not as straightforward as first expected and a number of techniques needed to be attempted before a satisfactory result was obtained. This section details each of these techniques.

6.4.1.1 Mask Creation By Thresholding

Since the rice grains were imaged on a black, highly absorbent background, it was expected that there would be a single band at which all rice grain pixels were significantly brighter than the background pixels. Some unwanted bright pixels caused by reflections were anticipated, but it was thought that a morphological opening on a thresholded result would be adequate, as in Section 4.3.1. The reflections (unwanted bright pixels), however were more problematic than anticipated, preventing accurate masks from being generated in this way. Figure 6.3(b) shows some regions of interest selected from the rice mixture hypercube. Two of these regions are of rice pixels, and four are of background

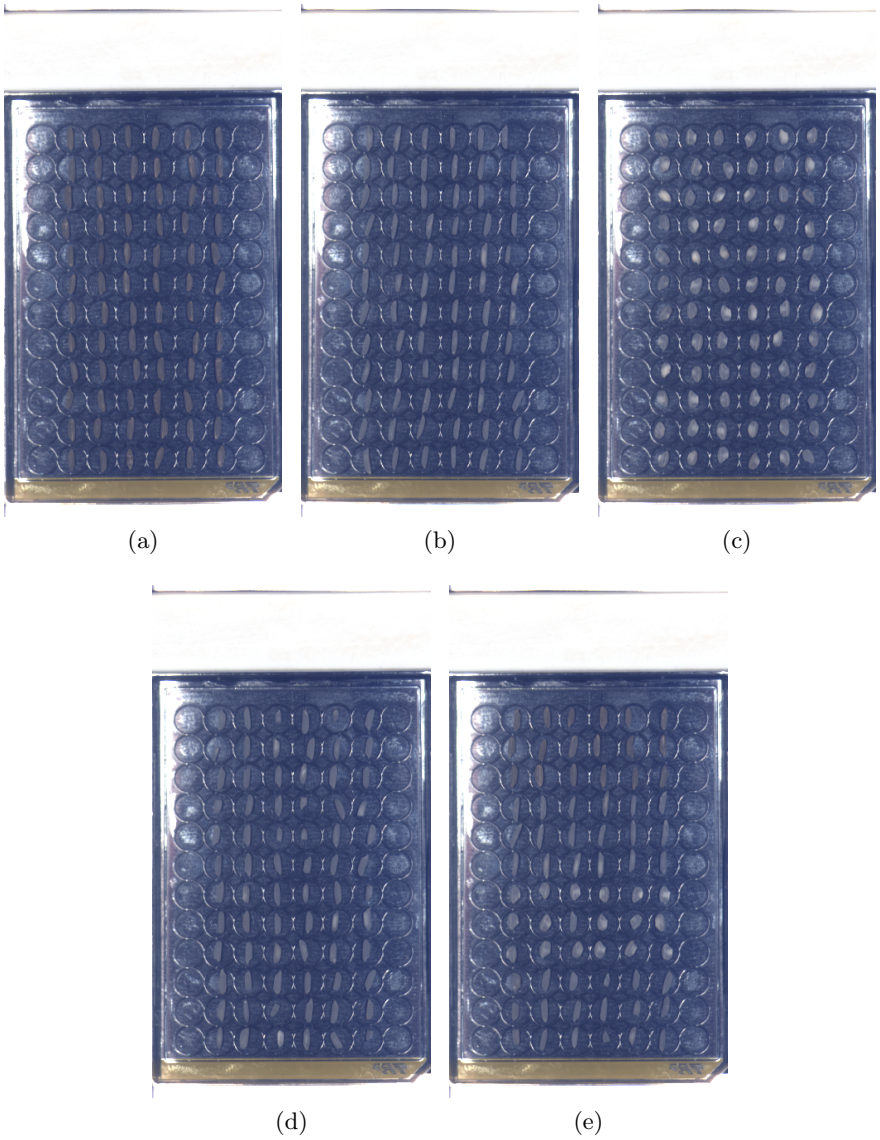


Figure 6.2: False colour representations of the five captured hypercubes (a) Rice 1 (b) Rice 2 (c) Rice 3 (d) Rice 4 (e) Mix of Rice 1, 2, 3, and 4

pixels, i.e., the well-plate. Figure 6.3(c) shows the mean spectrum for each of these regions. From this plot, it is clear that the background is sometimes brighter than the rice pixels and sometimes darker, regardless of the wavelength chosen. In cases such as this, separating objects based on a threshold value is impossible.

6.4.1.2 Mask Creation By Spectral Trough Steepness

From this plot, however, it can be seen that the background pixels all exhibit distinct troughs at 543 nm at 608 nm while the rice pixels maintain a relatively flat response. These two troughs are caused by the combination of the fluorescent room lighting that

was used during the imaging and the highly absorbent tape used behind the well plate. In the uncorrected image, there are peaks at these wavelengths on both the white calibration tile and the rice, but not on the well-plate, resulting in troughs produced in the corrected spectra.

The steepness of these troughs can be calculated by taking a simple ratio of the trough wavelength and another wavelength around 10 nm away. Figure 6.3(d) shows the image at 543 nm divided by the image at 533 nm. The rice grains are clearly brighter than the background pixels. To completely remove all the non-rice pixels, the image was thresholded and any pixels much larger than a rice grain (400 pixels) or much smaller (70 pixels) were masked out. The resultant mask is shown in Figure 6.3(e).

The masks for all rice images are shown in Figure 6.4. Each hypercube at each wavelength band was then ANDed with its corresponding mask. This produced hypercubes of identical size to the originals, but with the majority of pixels in each hypercube consisting of zeros.

6.4.2 Sparse Representation

Since a masked hypercube consists mainly of zeros, there is huge redundancy in storing every pixel at every wavelength band. To remove this redundancy, each hypercube was reshaped from a 3-D matrix (x, y, λ), to a 2-D matrix $(x \times y, \lambda)$. Two extra columns were then added to this data corresponding to the x and y values of each pixel resulting in a dataset of dimensions $(x \times y, \lambda + 2)$. All rows with zeros values for all wavelengths (the masked pixels) were then removed, leaving a dataset of much smaller size. For example, for the hypercube in Figure 6.2(a), the storage requirements can be calculated as $336(x) \times 620(y) \times 256(\lambda) \times 32\ bits = 203.438\ MB$. The storage requirements for the sparse representation of this hypercube are $12611 \times (256 + 2) \times 32\ bits = 12.412\ MB$, a huge decrease. This decrease not only reduces storage requirements, but results in significantly faster processing times as much less data has to be classified.

6.4.3 Spatial Feature Extraction

After considering a number of possible spatial features that could be extracted, the following three were chosen:

- Major axis length
- Minor axis length
- Eccentricity

Each of these features refer to properties of an ellipse. Clearly, not every rice grain has the shape of an ellipse, so a method is needed for determining the size and shape of an ellipse which represents a single grain of rice. This is done by isolating each grain of rice in the mask images, and calculating its second-moments. Since the second-moments define an ellipse [97], they can be used to determine the corresponding ellipse for each grain of rice in all of the mask images.

Major Axis Length

The major axis length of an ellipse is its longest diameter. Since the ellipse is determined by the second-moments of the rice grain, this will give an indication of each grain's length.

Minor Axis Length

The minor axis length is the shortest diameter of an ellipse. This is perpendicular to the major axis explained above and gives an indication of each grain's width.

Eccentricity

The eccentricity of an ellipse is another measure which is determined by the major and minor axis lengths. From [98] it is defined as

$$E = \frac{\sqrt{L^2 - l^2}}{L} \tag{6.1}$$

where E is the eccentricity, L is the major axis length, and l is the minor axis length. The eccentricity of an ellipse is always in the range 0 to 1, where 0 represents a circle and 1 represents a line.

Each of these spatial features was calculated for each grain of rice in each mask image. The three features were the incorporated into the hyperspectral data by creating an extra band for each feature, similar to the extra two bands created for the x and y locations of each pixels.

Other spatial features such as textures were considered, but manual examination of the rice images made it clear that any texture measures would not have helped separate any of the rice classes.

Another factor that needed to be considered, was the removal of noisy bands that offered no information. These were caused the sensor not being sensitive enough in the wavelength regions towards the ultraviolet end of the spectrum. The first ten spectral bands (306 nm - 408 nm) were removed from each sparsely represented hypercube. The λ dimension for each pixel therefore consisted of 246 spectral bands + 2 spatial coordinates + 3 spatial features = 251 variables.

6.5 Classification

As explained in Section 4.2.4, SVMs are consistently top performers in classification problems, especially when the sample features are hyperspectral data, so they make a good choice for this study. Although SVMs are binary classifiers, one-vs-all classification is used to allow multi-class classification.

Since the data is arranged into sparse form with 251 variables per pixel, selecting subsets for the data is simply an indexing problem. Selecting the training and testing data is also an indexing problem, but it is not as simple to solve. Perhaps the most intuitive method is to use the data from the hypercubes of each rice on its own as the training data and the hypercube of the mixture of rice as the testing data. The benefit of this method is that since the testing data contains spatial information from a single image, a visual result can be produced showing exactly where each classified point lies. The drawback however, is that it reduces the possibility of cross-validation, since there is only one way to split the data. An alternative method is to gather all the data together (individual rice classes and the mix of rice) into a large matrix where each row represents a pixel and each column is one of the 251 variables. Training and testing sets can then be chosen by selecting rows at random from this matrix.

For example, to perform ten-fold cross-validation, the large matrix would be split into ten sections and in turn each of the sections would be used as testing data with the rest used as training data. The results would then be averaged together. For this study, however, this method would not be appropriate due to the nature of the spatial

information. During the feature extraction process, three spatial features were calculated for each pixel. These spatial features were related to the shape of the rice grain to which it belonged, not to anything specific about this pixel. Therefore, for each pixel with specific spatial feature values, there exists several other pixels with the exact same spatial properties. With the cross-validation technique used above this means that some of the testing pixels could have identical spatial features to the ones in the training set since they could both originate from the same rice grain. This could not happen in a real-life classification situation and it would unfairly skew the results of the classification in favour of the spatial information.

To overcome this problem, the entire data set was split into individual rice grain sections. Since there are 5 images with 72 grains in each one, the data is now subset into 360 sections. Cross validation can then be performed by choosing different groups of these rice sections for training and testing sets without any pixels from a single grain appearing in both.

6.6 Results and Discussions

This section shows the results of the classification and aims to quantify the advantage offered by HSI over conventional imaging techniques. Different sized sets of training and testing data were used for quantitative evaluation of each imaging technique and the effect this had on the results are discussed.

6.6.1 Classification with Visual Result

As explained in the previous section, an intuitive way to split the data was to simply use the individual rice images as training data and the mixed rice image as testing data. This produced a visual result, shown in Figure 6.5 alongside the ground truth image. This ground truth image was used to calculate the accuracies shown in Table 6.1. Obviously, only rice pixels contributed to the accuracy, not any of the black background. Looking at these classification results, the highest classification result is with spectral and spatial information, as would be expected since this is the subset containing the most information. Without cross-validation, however, there can be no confidence in these results. Perhaps the split of training and testing this way happened to suit the SVM

particularly well and the classification performance would suffer in other circumstances. The effect of changing the percentage of training data also may have a great influence.

Table 6.1: Classification Accuracies for SVM Classification with Rice Mix as Testing Data

Dataset	**Accuracy**
Spatial information	69.91%
Colour information	67.17%
Spectral information	74.27%
Colour and spatial information	81.67%
Spectral and spatial information	90.20%

6.6.2 Cross-validation Results

By splitting the entire dataset into 360 individual rice grain sections, a number of different levels of cross-validation were performed. The technique used was repeated random sub-sampling, where the data is randomly split into multiple training and testing sets. The classification accuracy is then calculated for each pair of sets and an average is calculated. This entire process can then be repeated for various percentages of training (and testing) data. In this study, 20 different pairs were randomly selected for each training percentage. The results of this are shown in Table 6.2.

From Table 6.2, the results appear broadly similar to the non-cross-validated results, suggesting that the data generalises well and that the initial split based on the individual rice images and the mixed rice image was a good approximation of random selection. As before, combining spectral and spatial information produced the highest classification accuracies. It is interesting to note that the spatial, colour, and spectral subsets had relatively low accuracies individually, but combined together well. This shows that the spatial differences were to an extent unrelated to the spectral differences and that both were required for accurate classification. Adding colour information to the spatial information improved the classification result, but not as much as adding spectral information. This was expected since the colour information is a subset of the spectral information.

Table 6.2: Cross-validated Classification Accuracies for SVM Classification

Training Percentage	**Spatial**	**Colour**	**Spectral**	**Colour + spatial**	**Spectral + spatial**
5%	57.80%	61.04%	58.73%	58.72%	70.44%
10%	64.76%	62.96%	63.01%	65.86%	84.69%
15%	65.74%	64.97%	64.07%	67.61%	88.98%
20%	66.63%	65.26%	65.49%	73.90%	91.19%
25%	68.70%	66.74%	66.22%	76.72%	91.48%
30%	66.84%	66.88%	66.33%	79.27%	92.12%
35%	68.31%	67.75%	67.75%	82.99%	92.35%
40%	67.05%	66.87%	68.02%	80.33%	92.28%
45%	67.30%	67.41%	67.40%	83.66%	92.75%
50%	65.76%	67.51%	69.64%	81.76%	93.48%
55%	66.24%	67.20%	69.63%	84.63%	93.60%
60%	64.85%	67.85%	69.32%	84.07%	93.34%
65%	66.87%	68.50%	69.08%	81.87%	93.70%
70%	68.88%	68.16%	67.94%	84.93%	94.33%
75%	67.85%	68.55%	69.43%	85.96%	93.07%
80%	66.26%	69.15%	69.59%	84.68%	94.13%
85%	68.99%	67.18%	70.86%	86.65%	94.33%
90%	67.79%	69.97%	69.58%	84.82%	93.65%
95%	69.20%	66.24%	69.91%	89.78%	94.32%
Average	**66.62%**	**66.85%**	**67.47%**	**79.91%**	**91.27%**

As the training percentage increased, the individual classification accuracies also increased, as would be expected. The combinations of colour+spatial and spectral+spatial showed the steepest increases.

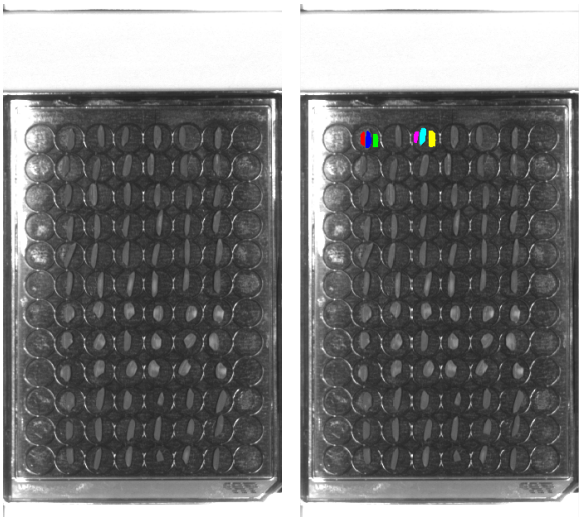
6.7 Summary

In this chapter, hyperspectral images of rice were classified using an SVM classifier. To evaluate the performance improvement offered by hyperspectral data over conventional imaging techniques, the classifier was implemented with five different data subsets: spatial, colour, spectral, spatial + colour, and spectral + spatial. The varying steepness of a spectral trough in the data was exploited to create binary masks for each rice image to get rid of any background pixels. The second-moments of each rice grain in each of these masks was then used to determine the dimensions of corresponding ellipses. The major-axis length, minor-axis length and eccentricity were then measured for each of the ellipses and were used as the spatial information for classification. The colour

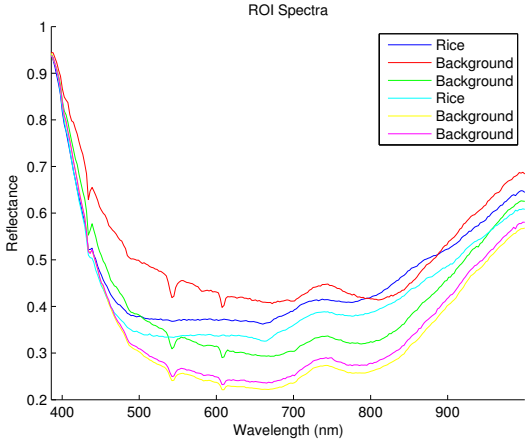
information was extracted from the hyperspectral data by means of a true colour image. By removing all rice pixels, each image reduced to a sparse representation, substantially reducing the storage and processing requirements.

The results show that the combination of spectral and spatial information produced the highest classification accuracy, averaging $> 90\%$ accuracy during each level of cross-validation. Using colour and spatial information together (such as would be available from a standard colour imaging system) produced the second most accurate results, averaging $>80\%$ accuracy. This was higher than the approximately 70% accuracies seen for spectral information on its own. From these results, it is shown that HSI in this context can improve classification, but only if spatial features are still considered.

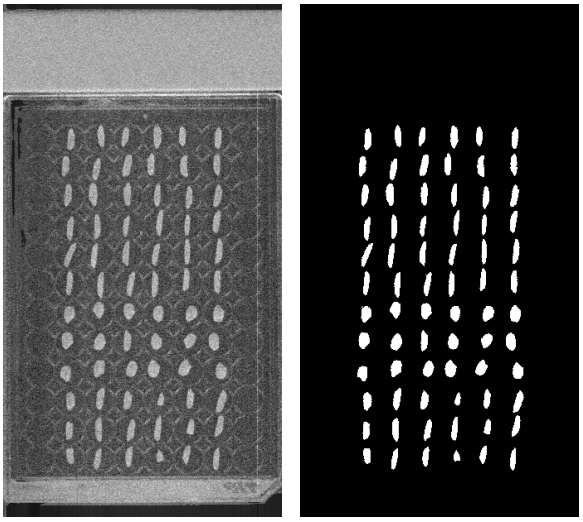
It has been demonstrated that VNIR HSI is a useful tool that when appropriately used can improve classification results. It does not, however, negate the need for standard image processing techniques. Since only four types of rice grain were analysed, there is no guarantee that these results would generalise well for other foodstuffs. For this reason, it would be worthwhile to repeat this experiment with new samples. This is further discussed in Chapter 7.



(a) (b)

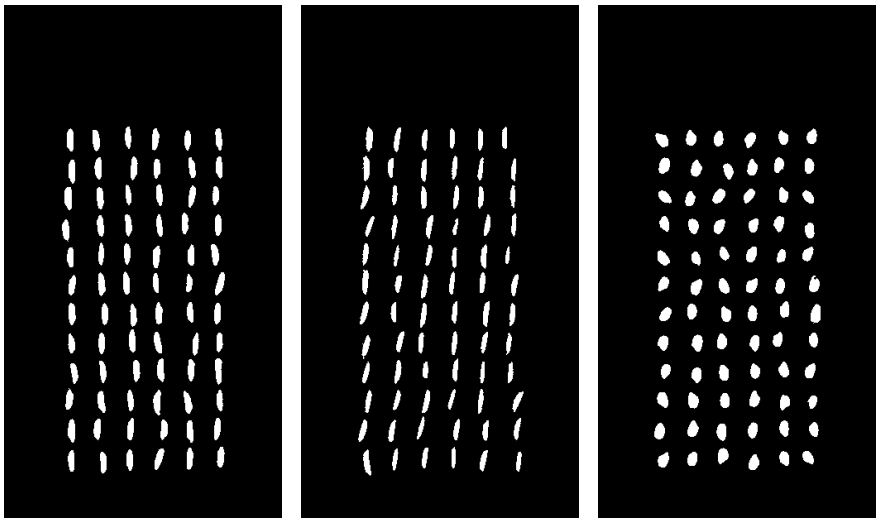


(c)

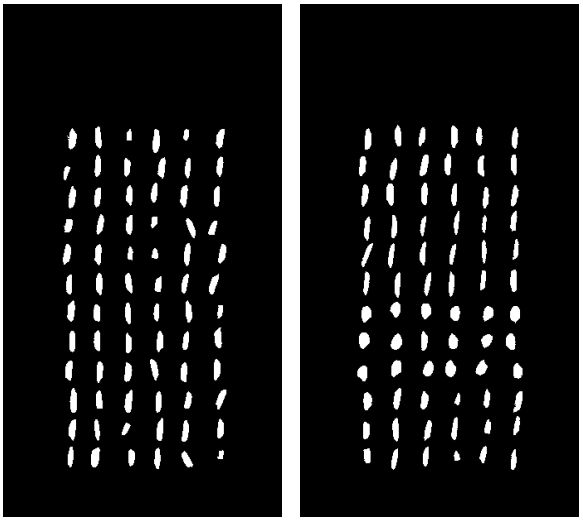


(d) (e)

Figure 6.3: Creation of the mask for the Mix of Rice 1, 2, 3 and 4 (a) At 624 nm (b) Regions of Interest (c) Plot of the Regions of Interest (d) The output of the trough steepness calculation (e) The final mask



(a) (b) (c)



(d) (e)

Figure 6.4: Masks for each hypercube to remove non-rice pixels (a) Rice 1 Mask (b) Rice 2 Mask (c) Rice 3 Mask (d) Rice 4 Mask (e) Mix of Rice 1, 2, 3, and 4 Mask

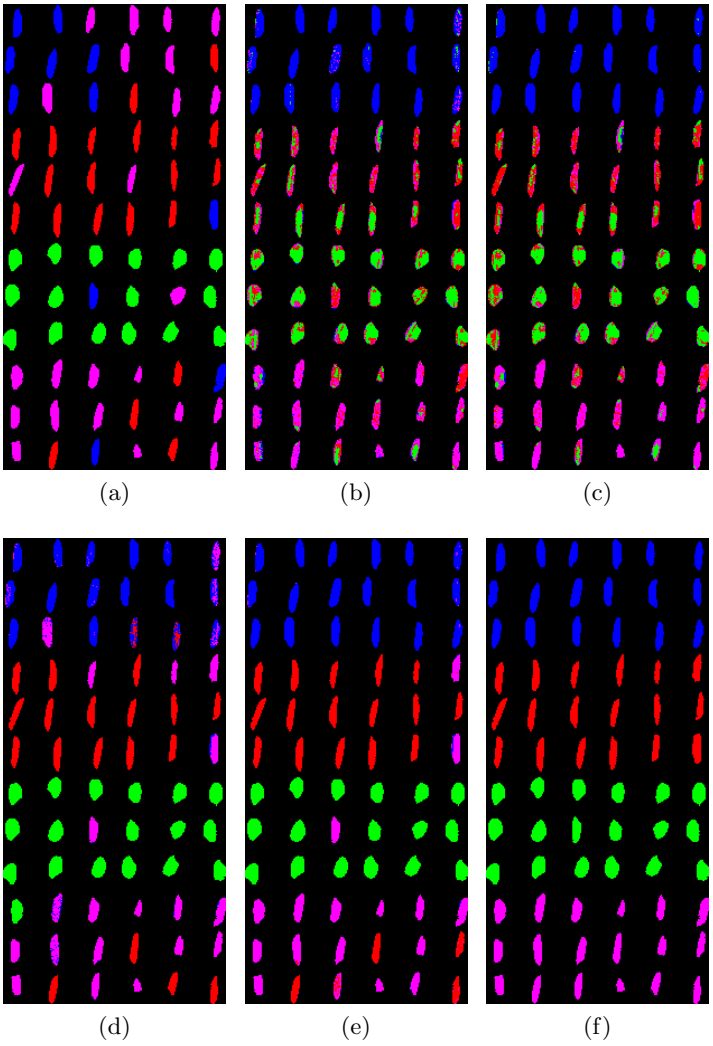


Figure 6.5: SVM Classification results with Rice Mix as testing data (a) Spatial information (b) Colour information (c) Spectral information (d) Colour and spatial information (e) Spectral and spatial information (f) Ground truth

Chapter 7

Conclusions and Further Work

The initial aims of the work in this project were to develop new methods of capture and analysis of hyperspectral data, and to conduct studies that could convince industrialists that hyperspectral imaging may be a financially viable improvement to an aspect of their business. This thesis demonstrates that these goals have clearly been achieved.

In terms of methods of capturing data, a major contribution has been the development of a novel algorithm which converts data captured using a rotational line-scan into Cartesian form making it an alternative to linear line-scan systems. This was enhanced by the addition of a calibration process which determined the inherent offsets in the scan using a calibration object that does not need to be perfectly aligned with the axis of rotation of the scan.

The Chinese tea classification, the analysis of baked sponges and the quantitative analysis of spatial, spectral and colour information have clearly demonstrated that hyperspectral imaging can offer many benefits. The classification of Chinese tea samples showed how hyperspectral imaging can be used to distinguish visually similar samples. The analysis of baked sponges clearly demonstrated the potential value of hyperspectral imaging as a quality analysis tool in a manufacturing environment and the quantitative analysis of spatial, spectral and colour information showed the improvements offered by hyperspectral imaging when compared to conventional imaging techniques.

In this chapter, the major conclusions of this work are presented and the impact of the demonstrated results is evaluated. Finally, a number of potential directions that future work in this field could take are discussed.

7.1 Rotational Scanning Techniques

In Chapter 2, the main image acquisition techniques for hyperspectral imaging were described. The most popular method was identified as pushbroom, where an entire line of pixels at all wavelengths is captured simultaneously, before the object or detector is moved and another line is captured. To ensure square pixels, an accurate way of moving the camera or the object is required. Chapter 3 proposed that since rotational movement is often naturally available or easier to achieve, a pushbroom acquisition technique that utilises a rotational line-scan would be beneficial.

First, the idea of a perfect rotational scan was explained where the centre of each line scanned lies exactly coincident with the axis of rotation of the scan. Since a perfect scan of this type is the equivalent of polar representation, conversion to Cartesian representation is trivial using publicly available toolboxes. It was then explained that a perfect scan of this nature is very hard to achieve, and that there will almost always be two inherent offsets related to each scan. These offsets were defined as a vertical offset and a horizontal offset. The horizontal offset represents a shift in the scan and as such is easy to account for. The vertical offset, however, represents the distance the line-scan is away from the axis of rotation. It was shown how this means that each line in the scan can be represented by a tangent of a circle whose radius is equal to the vertical offset. By defining the scan in this way, an algorithm was proposed which takes a rotationally scanned image and converts it into Cartesian form. Several examples of rotational scans converted to Cartesian form were shown.

To operate, this algorithm required that both the horizontal and vertical offsets were known. Rather than manually determining these offsets through trial and error, a calibration technique was proposed which used a calibration object consisting of a circle of known size and of a colour which highly contrasted with the background. By aligning the centre of this calibration object with the axis of rotation of the scan, it was shown how these two offsets could be easily calculated. This calibration process was then improved by relaxing the alignment requirements. Several examples were shown of how an imperfectly aligned calibration object could be used to calculate both of the offsets. Since nearly all of the pixels in Cartesian form do not have an exact corresponding pixel in the scanned image, some form of interpolation was required. Nearest-neighbour, bilinear and bicubic interpolation methods were compared, and bilinear was selected as

the most appropriate method as it offered the best trade off between performance and computational complexity.

Finally, three synthetic images were used to measure the performance of rotational scanning compared to linear scanning, by measuring the similarity between different pairs of images. In general, the results were very good with rotational scanning outperforming linear scanning in some cases, demonstrating that rotational scanning is a viable alternative hyperspectral imaging method to a linear line-scan.

7.2 Supervised Classification Techniques in Hyperspectral Imaging

In Chapter 4, five of the most popular classification techniques were explained before each one was implemented in a classification problem where five different types of Chinese tea leaves had to be distinguished. MDC, MLC, ANN, SVM and AdaBoost with MLC were selected as the five classifiers.

Due to the nature of the imaging of the tea leaves, each image consisted of many non-tea pixels, so some cropping was required. Rather than using manual cropping, each image was converted to a binary image and a rectangular structuring element was used to perform a morphological opening, resulting in a mask which was used to remove all non-tea pixels. Since hyperspectral images contain a significant amount of highly correlated data, PCA was used to reduce that data.

After splitting the data into multiple training and testing sets, each classifier was used to classify the data. Several parameters were investigated: the number of PCA components, the size of the modal filter used for post-processing, the number of hidden layers (ANN only), the number of weak learners (AdaBoost with MLC only) and the percentage of data used for training.

Plots were produced for each of these tests with the classification accuracy being the measure of performance. Cross-validation was used to ensure the results were statistically significant and Cohen's Kappa was calculated to confirm that the higher classification accuracies were not significantly effected by random chance.

Increasing the number of PCA components improved accuracy up until around 10 components where it offered little improvement. Modal filtering of the data improved the results up until around 9 by 9 pixels where the benefit was outweighed by increased computational complexity. Using more hidden layers in the ANN improved results, although more repeated results are needed to confirm this is not due to the random initial weighting. Adding more weak learners to the AdaBoost with MLC also improved results but significantly increased computation time - 15 weak learners offered a good compromise. Using more training data did not improve classification results and in some cases decreased the accuracy.

This study clearly demonstrated the potential for hyperspectral imaging in a classification context. Although initial classification results may be unsatisfactory, it was shown that appropriate pre and post-processing can greatly improve the classification performance of a system.

7.3 Analysis of Baked Sponges

An analysis of baked sponges using hyperspectral imaging was presented in Chapter 5. An expert panel of tasters scored a selection of sponges of various ages on taste, texture and appearance. A VNIR hyperspectral imaging system was used to image a similar selection of sponges taken from the same batch. Through calibration and correction of every image, a score relating to sponge moisture content for each sponge was produced. Comparisons of these scores with the tasting panel scores showed a strong correlation, especially for the chocolate sponges. By taking an average sponge image for each week, the moisture content information was used to train a classifier which identified moist and dry regions of each sponge image. This trained classifier could then be used to analyse an unknown sponge and give a visualisation of its estimated moisture distribution. It was then discussed how the techniques developed in this chapter could be implemented in the production line environment of a bakery, allowing every sponge that is produced to be imaged and scored, rather than the small subset the human tasters currently test. This information would allow any problems in the baking process to be identified earlier. Furthermore, alterations could be made to the baking process and any changes in the quality of sponges could be immediately quantified.

Although hyperspectral imaging will never completely replace the need for human tasters in the quality analysis of food products, this case study demonstrated the potential benefits a hyperspectral imaging system could provide. With ability to analyse many more samples than a human tasting panel ever could, and capture spatial information, hyperspectral imaging is destined to become a key technology in the baking industry and the wider food and drink sector in general.

7.4 A Quantitative Analysis of Spatial, Spectral and Colour Information

Chapter 6 provided a quantitative analysis of SVM classification using spatial, spectral and colour information. While it was expected that a SWIR system would provide the most accurate results, the benefits offered by a VNIR system tend to be less obvious. The much lower cost of a VNIR system means that it would be preferable to a SWIR one, if the benefits over standard colour imaging techniques could be demonstrated.

Using images of four different kinds of rice, four data subsets were produced: spatial, colour, spatial + colour, and spectral + spatial. Similar to the Chinese tea images, each image consisted of many unwanted background pixels.

By exploiting an unusual feature in the data caused by the fluorescent lighting in the room during the imaging, a mask was created to separate the rice from the background in each image. These mask images were then used to extract the second-moments of each rice grain. These second moments were then used to determine the dimensions of an ellipse for each grain, where the major-axis, minor-axis and eccentricity of each could be measured. These three properties were used as the spatial information. Spectral information was selected by ignoring any spatial features. Colour information was extracted as true colour images from the hyperspectral data. Since so much of the data consisted of the unwanted background pixels, each hypercube was reduced to a sparse representation where each rice pixel was stored with its spatial coordinates and all other pixels were discarded. This resulted in a significant reduction in storage and processing requirements.

As expected, using both spatial and spectral information produced the highest results from the SVM classification (>90% average accuracy after cross-validation). Spatial

information combined with colour information had the second best performance with nearly 80% accuracy, notably better than the approximately 67% accuracy seen for spectral information on its own. These results showed that VNIR hyperspectral imaging can significantly improve classification results compared to standard imaging techniques, but it does not negate the need for spatial information to be taken into account.

7.5 Further Work

From the results presented in this thesis, there are a number of areas in which further work could be carried out. These are discussed in this section.

7.5.1 Rotational Scanning Techniques

For the automatic determination of the rotational scan offsets, the requirements for the calibration object were relaxed to the point where perfect alignment was not required. The placement still had to satisfy some conditions, however. It is thought that these requirements could be relaxed further still perhaps with the use of a different shape of calibration object, or even by using multiple calibration objects.

Only three different images were used to measure the performance of the rotational scan and compare it with the traditional linear one. By creating multiple images using multiple filter sizes, a more comprehensive set of testing data could be produced. The size and shape of the features in a rotationally scanned image is also likely to influence the performance, so a more supervised approach to test pattern generation (rather than filtered random images) would be useful for exploring this.

As mentioned in Chapter 3, the research in [80] is based on applying image processing techniques to images in polar-logarithmic form. It would be interesting see how well these techniques would preform when used with images captured by a rotational scan.

A robotic arm would be an ideal way to provide the movement required for a rotational scan. As robotic arms are commonly used for non-destructive testing [99], combining the rotational scanning method proposed in this thesis with the non-destructive testing capabilities of hyperspectral imaging would be a perfect way to showcase both techniques.

7.5.2 Supervised Classification Techniques

With the Chinese tea classification, the modal filtering used during post-processing was carried out using a simple square structuring element. A potential improvement to the classification result could be to use a structuring element based on the shape and size of the tea leaves being classified.

Further work could be to improve the robustness of the results by performing more experiments. The ANN classification results in particular suffered from different initial random weightings, so more repeats of these classifications would be beneficial. Since only four different samples were available for each tea type, the number of possible cross-validations was limited was for all classifiers. A simple way to remedy this would be to image the tea leaves multiple times, but shaking the ice cube tray between repeat images.

Each of the five classification techniques could be applied to both the sponge data and the rice grain data. It would be interesting to see if the classifiers ranked the same as with the tea data or if certain classifiers were more suited to specific datasets.

7.5.3 Analysis of Baked Sponges

The analysis of baked sponges is perhaps where most of the potential for future work lies. Indeed, further research has already been carried out using a SWIR hyperspectral system and the initial results indicate an improvement upon the results presented in this thesis. The bakery has recently acquired a moisture meter which allows an accurate moisture percentage to be obtained for a small sponge sample ($<10\ g$). Further work has been planned to use this moisture meter to confirm the relationship between moisture content and hyperspectral data with the aim of accurately predicting moisture content using hyperspectral data alone. The end goal of this work is to implement a hyperspectral imaging system at the bakery where every single sponge produced could be imaged and analysed.

Although the main focus of this research has been the analysis of baked sponges, the bakery manufactures many other products such as chocolates, biscuits and other confectionery. Each of these products will have a similar method of quality assessment which

could well be improved by a similar type of hyperspectral analysis to the one described in this thesis.

7.5.3.1 Spatial, Spectral and Colour Information

The quantitative analysis carried out in Chapter 6 only looked at the classification of four kinds of rice grains. The robustness of this analysis would be greatly improved by considering other objects, perhaps with some which may exhibit a high spatial variation and low spectral variation such different pastas, and some which exhibit a high spectral variation and low spatial variation such as dry spices. By comparing these results, a score could be computed which could give an indication of spectral variance compared to spatial variance and be used to predict the benefit offered by hyperspectral imaging compared to conventional colour imaging.

Only three spatial features were used (major axis length, minor axis length and eccentricity of each rice grain). Although texture measures were dismissed due to the similarity of the rice grains, texture measures may be useful in other experiments outlined above. Other spatial features could also be extracted, possibly by using rotational scanning to find features in the proposed feature space.

Bibliography

[1] D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review part ii: Applications," *Innovative Food Science & Emerging Technologies*, vol. 19, no. 0, pp. 15 – 28, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1466856413000799

[2] G. Edelman, T. G. van Leeuwen, and M. C. Aalders, "Hyperspectral imaging for the age estimation of blood stains at the crime scene," *Forensic Science International*, vol. 223, no. 1?3, pp. 72 – 77, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0379073812003660

[3] K. Tiwari, M. Arora, and D. Singh, "An assessment of independent component analysis for detection of military targets from hyperspectral images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 5, pp. 730 – 740, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S030324341100050X

[4] J. K. Bowmaker and H. J. Dartnall, "Visual pigments of rods and cones in a human retina." *The Journal of Physiology*, vol. 298, no. 1, pp. 501–511, 1980. [Online]. Available: http://jp.physoc.org/content/298/1/501.abstract

[5] Matlab colon operator. [Online]. Available: http://www.mathworks.co.uk/help/matlab/ref/colon.html

[6] D. Akkaynak, T. Treibitz, B. Xiao, U. A. Gürkan, J. J. Allen, U. Demirci, and R. T. Hanlon, "Use of commercial off-the-shelf digital cameras for scientific data acquisition and scene-specific color calibration," *J. Opt. Soc. Am. A*, vol. 31, no. 2, pp. 312–321, Feb. 2014. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-31-2-312

[7] *Scyllarus MATLAB Toolbox*. [Online]. Available: http://scyllarus.research.nicta.com.au/

[8] D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review part i: Fundamentals," *Innovative Food Science & Emerging Technologies*, vol. 19, no. 0, pp. 1 – 14, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1466856413000775

[9] P. Schaare and D. Fraser, "Comparison of reflectance, interactance and transmission modes of visible-near infrared spectroscopy for measuring internal properties of kiwifruit (actinidia chinensis)," *Postharvest Biology and Technology*, vol. 20, no. 2, pp. 175 – 184, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925521400001307

[10] C. Starr, *Biology : concepts and applications*. Belmont, CA: Thomson, Brooks/-Cole, 2008.

[11] Z. Xiong, D.-W. Sun, X.-A. Zeng, and A. Xie, "Recent developments of hyperspectral imaging systems and their applications in detecting quality attributes of red meats: A review," *Journal of Food Engineering*, vol. 132, no. 0, pp. 1 – 13, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0260877414000648

[12] E. F. Schubert, *Light-Emitting Diodes*. Cambridge University Press, 2006.

[13] M. E. Klein, B. J. Aalderink, R. Padoan, G. De Bruin, and T. A. Steemers, "Quantitative hyperspectral reflectance imaging," *Sensors*, vol. 8, no. 9, pp. 5576–5618, 2008. [Online]. Available: http://www.mdpi.com/1424-8220/8/9/5576

[14] L. N. THIBOS, A. BRADLEY, and X. ZHANG, "Effect of ocular chromatic aberration on monocular visual performance." *Optometry & Vision Science*, vol. 68, no. 8, pp. –, 1991. [Online]. Available: http://journals.lww.com/optvissci/Fulltext/1991/08000/Effect_of_Ocular_Chromatic_Aberration_on_Monocular.5.aspx

[15] [Online]. Available: https://www.flickr.com/photos/nicholasjones/8529867715/in/photostream/

[16] D. Bannon and R. Thomas, "Harsh environments dictate design of imaging spectrometer," *Laser Focus World*, 2005.

[17] A. Korpel, "Acousto-optics - a review of fundamentals," *IEEE Proceedings*, vol. 69, pp. 48–53, Jan. 1981.

[18] S. A. Mathews, "Design and fabrication of a low-cost, multispectral imaging system," *Appl. Opt.*, vol. 47, no. 28, pp. F71–F76, Oct 2008. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-47-28-F71

[19] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–B51, Apr 2008. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-47-10-B44

[20] L. Gao, R. T. Kester, and T. S. Tkaczyk, "Compact image slicing spectrometer (iss) for hyperspectral fluorescence microscopy," *Opt. Express*, vol. 17, no. 15, pp. 12 293–12 308, Jul 2009. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-17-15-12293

[21] A. Gorman, D. W. Fletcher-Holmes, and A. R. Harvey, "Generalization of the lyot filter and its application to snapshot spectral imaging," *Opt. Express*, vol. 18, no. 6, pp. 5602–5608, Mar 2010. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-6-5602

[22] N. Gupta, P. R. Ashe, and S. Tan, "Miniature snapshot multispectral imager," *Optical Engineering*, vol. 50, no. 3, pp. 033 203–033 203–9, 2011. [Online]. Available: http://dx.doi.org/10.1117/1.3552665

[23] *Hamamatsu ORCA-05G Specifications*. [Online]. Available: http://www.hamamatsu.com/jp/en/product/alpha/O/C8484-05G02/index.html

[24] H. Yao and D. Lewis, "{CHAPTER} 2 - spectral preprocessing and calibration techniques," in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, pp. 45 – 78. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123747532100024

[25] "Mercury lamp response." [Online]. Available: http://www.zolix.com.cn/en/prodcon_370_378_369.html

[26] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *Information Theory, IEEE Transactions on*, vol. 14, no. 1, pp. 55–63, Jan 1968.

[27] N. Keshava and J. Mustard, "Spectral unmixing," *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 44–57, Jan 2002.

[28] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 8, pp. 1778–1790, Aug 2004.

[29] J. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 32, no. 4, pp. 779–785, Jul 1994.

[30] M. Fauvel, J. Benediktsson, J. Chanussot, and J. Sveinsson, "Spectral and spatial classification of hyperspectral data using svms and morphological profiles," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 11, pp. 3804–3814, Nov 2008.

[31] S. Kaewpijit, J. Le-Moigne, and T. El-Ghazawi, "Automatic reduction of hyperspectral imagery using wavelet spectral analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 4, pp. 863–871, April 2003.

[32] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 3, pp. 466–479, March 2005.

[33] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, 3rd ed., D. E. Ricken and W. Gessner, Eds. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1999.

[34] R. Dianat and S. Kasaei, "Dimension reduction of optical remote sensing images via minimum change rate deviation method," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 1, pp. 198–206, Jan 2010.

[35] J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, and J. Han, "Novel folded-pca for improved feature extraction and data reduction with hyperspectral imaging and {SAR} in remote sensing," {*ISPRS*} *Journal of*

Photogrammetry and Remote Sensing, vol. 93, pp. 112 – 122, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924271614000938

[36] L. Bagnasco, M. Zotti, N. Sitta, and P. Oliveri, "A pca-based hyperspectral approach to detect infections by mycophilic fungi on dried porcini mushrooms (boletus edulis and allied species)," *Talanta*, vol. 144, pp. 1225 – 1230, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0039914015301971

[37] M. Kim, A. Lefcourt, K. Chao, Y. Chen, I. Kim, and D. Chan, "Multispectral detection of fecal contamination on apples based on hyperspectral imagery: Part i. application of visible and near-infrared reflectance imaging," *Transactions of the American Society of Agricultural Engineers*, vol. 45, no. 6, pp. 2027–2037, 2002, cited By 136. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0242486677&partnerID=40&md5=4e16716df9d63814656b46bed2d77e6d

[38] J. Xing, W. Saeys, and J. De Baerdemaeker, "Combination of chemometric tools and image processing for bruise detection on apples," *Computers and Electronics in Agriculture*, vol. 56, no. 1, pp. 1–13, 2007, cited By 64. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-33846921806&partnerID=40&md5=cb88cfaa2ba91acfa8dcdfa48a927246

[39] C. wang DONG, Y. YE, J. qiang ZHANG, H. kai ZHU, and F. LIU, "Detection of thrips defect on green-peel citrus using hyperspectral imaging technology combining {PCA} and b-spline lighting correction method," *Journal of Integrative Agriculture*, vol. 13, no. 10, pp. 2229 – 2235, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2095311913606711

[40] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philosophical Magazine*, vol. 6, no. 2, pp. 559–572, 1901.

[41] H. Hotelling, "Analysis of a complex of statistical variables into principal components," 1933.

[42] I. Jolliffe, *Principal Component Analysis*, ser. Springer Series in Statistics. Springer, 2002. [Online]. Available: http://books.google.co.uk/books?id=_olByCrhjwIC

[43] J. Shan, "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Systems*, 2002.

[44] J. Zhao, Q. Chen, J. Cai, and Q. Ouyang, "Automated tea quality classification by hyperspectral imaging," *Appl. Opt.*, vol. 48, no. 19, pp. 3557–3564, Jul 2009. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-48-19-3557

[45] M. S. Andresen, B. S. Dissing, and H. L?je, "Quality assessment of butter cookies applying multispectral imaging," *Food Science & Nutrition*, vol. 1, no. 4, pp. 315–323, 2013. [Online]. Available: http://dx.doi.org/10.1002/fsn3.46

[46] Y.-Y. Pu and D.-W. Sun, "Visnir hyperspectral imaging in visualizing moisture distribution of mango slices during microwave-vacuum drying," *Food Chemistry*, vol. 188, pp. 271 – 278, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0308814615006779

[47] D. P. Ariana, R. Lu, and D. E. Guyer, "Near-infrared hyperspectral reflectance imaging for detection of bruises on pickling cucumbers," *Computers and Electronics in Agriculture*, vol. 53, no. 1, pp. 60 – 70, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169906000378

[48] J.-H. Cheng and D.-W. Sun, "Hyperspectral imaging as an effective tool for quality analysis and control of fish and other seafoods: Current research and potential applications," *Trends in Food Science & Technology*, vol. 37, no. 2, pp. 78 – 91, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924224414000648

[49] G. ElMasry, N. Wang, and C. Vigneault, "Detecting chilling injury in red delicious apple using hyperspectral imaging and neural networks," *Postharvest Biology and Technology*, vol. 52, no. 1, pp. 1 – 8, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925521408003220

[50] L. Pan, Q. Zhang, W. Zhang, Y. Sun, P. Hu, and K. Tu, "Detection of cold injury in peaches by hyperspectral reflectance imaging and artificial neural network," *Food Chemistry*, vol. 192, pp. 134 – 141, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S030881461500998X

[51] T. Qiao, J. Ren, C. Craigie, J. Zabalza, C. Maltin, and S. Marshall, "Singular spectrum analysis for improving hyperspectral imaging based beef eating quality evaluation," *Computers and Electronics in Agriculture*, vol. 115, pp. 21 –

25, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169915001295

[52] G. K. Naganathan, L. M. Grimes, J. Subbiah, C. R. Calkins, A. Samal, and G. E. Meyer, "Visible/near-infrared hyperspectral imaging for beef tenderness prediction," *Computers and Electronics in Agriculture*, vol. 64, no. 2, pp. 225 – 233, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016816990800152X

[53] T. Selige, J. Bohner, and U. Schmidhalter, "High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures," *Geoderma*, vol. 136, no. 1?2, pp. 235 – 244, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0016706106001030

[54] P. Goel, S. Prasher, J. Landry, R. Patel, R. Bonnell, A. Viau, and J. Miller, "Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn," *Computers and Electronics in Agriculture*, vol. 38, no. 2, pp. 99 – 124, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169902001382

[55] Z. yu LIU, J. feng HUANG, and R. xiang TAO, "Characterizing and estimating fungal disease severity of rice brown spot with hyperspectral reflectance data," *Rice Science*, vol. 15, no. 3, pp. 232 – 242, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1672630808600475

[56] J. K. . B. K. J. Shwetank, "Review of rice crop identification and classification using hyperspectral image processing system," *International Journal of Computer Science & Communication*, vol. 1, no. 1, pp. 253–258, 2010.

[57] M. A. F. de la Ossa, C. Garciaa-Ruiz, and J. M. Amigo, "Near infrared spectral imaging for the analysis of dynamite residues on human handprints," *Talanta*, vol. 130, no. 0, pp. 315 – 321, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0039914014005827

[58] G. Reed, K. Savage, D. Edwards, and N. N. Daeid, "Hyperspectral imaging of gel pen inks: An emerging tool in document analysis," *Science & Justice*, vol. 54, no. 1, pp. 71 – 80, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1355030613000907

[59] S. Karimkhan, K. Vongsy, A. K. Shaw, and D. Wicker, "Automated recognition and detection of dismounts and vehicles using close-in urban hyperspectral images," pp. 65 650X–65 650X–12, 2007. [Online]. Available: http://dx.doi.org/10.1117/12.723211

[60] J.-P. Ardouin, J. Levesque, and T. Rea, "A demonstration of hyperspectral image exploitation for military applications," in *Information Fusion, 2007 10th International Conference on*, July 2007, pp. 1–8.

[61] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of Biomedical Optics*, vol. 19, no. 1, p. 010901, 2014. [Online]. Available: http://dx.doi.org/10.1117/1.JBO.19.1.010901

[62] J. A. Chin, E. C. Wang, and M. R. Kibbe, "Evaluation of hyperspectral technology for assessing the presence and severity of peripheral artery disease," *Journal of Vascular Surgery*, vol. 54, no. 6, pp. 1679 – 1688, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S074152141101384X

[63] S. H. Hardarson, "Retinal oximetry," *Acta Ophthalmologica*, vol. 91, no. thesis2, pp. 1–47, 2013. [Online]. Available: http://dx.doi.org/10.1111/aos.12086

[64] D. J. Mordant, I. Al-Abboud, G. Muyo, A. Gorman, A. Sallam, P. Rodmell, J. Crowe, S. Morgan, P. Ritchie, A. R. Harvey, and A. I. McNaught, "Validation of human whole blood oximetry, using a hyperspectral fundus camera with a model eye," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 5, p. 2851, 2011. [Online]. Available: +http://dx.doi.org/10.1167/iovs.10-6217

[65] A. Gowen, C. O'Donnell, P. Cullen, and S. Bell, "Recent applications of chemical imaging to pharmaceutical process monitoring and quality control," *European Journal of Pharmaceutics and Biopharmaceutics*, vol. 69, no. 1, pp. 10 – 22, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0939641107003566

[66] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias, "Hyperspectral imaging an emerging process analytical tool for food quality and safety control," *Trends in Food Science and Technology*, vol. 18, no. 12, pp. 590 – 598, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924224407002026

[67] N. K. Boesche, C. Rogass, C. Mielke, S. Herrmann, F. Krting, A. Papenfu, C. Lubitz, M. Brell, S. Tonn, and U. Altenberger, "Chapter 16 - hyperspectral rare earth element mapping of three outcrops at the fen complex, norway: Calcitic, dolomitic, and ankeritic carbonatites," in *Rare Earths Industry*, I. B. D. L. L. Filho, Ed. Boston: Elsevier, 2016, pp. 235 – 265. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780128023280000164

[68] C. Yang, W. S. Lee, and P. Gader, "Hyperspectral band selection for detecting different blueberry fruit maturity stages," *Computers and Electronics in Agriculture*, vol. 109, pp. 23 – 31, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169914002087

[69] A. Al-Mallahi, T. Kataoka, and H. Okamoto, "Discrimination between potato tubers and clods by detecting the significant wavebands," *Biosystems Engineering*, vol. 100, no. 3, pp. 329 – 337, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1537511008001207

[70] H. Okamoto and W. S. Lee, "Green citrus detection using hyperspectral imaging," *Computers and Electronics in Agriculture*, vol. 66, no. 2, pp. 201 – 208, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169909000258

[71] P. Manandhar, "Polar to/from rectangular transform of images," *http://www.mathworks.com/matlabcentral/fileexchange/17933-polar-tofrom-rectangular-transform-of-images.*

[72] D. Marshall, "Proof of an implicit line equation through a point (not on the circle) that is tangent to a general circle," *http://www.cs.cf.ac.uk/Dave/CM0268/PDF/circle_tangent_intersect_proof.pdf.*

[73] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition).* Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.

[74] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, April 2004.

[75] H. L. Tan, Z. Li, Y. H. Tan, S. Rahardja, and C. Yeo, "A perceptually relevant mse-based image quality metric," *Image Processing, IEEE Transactions on*, vol. 22, no. 11, pp. 4447–4459, Nov 2013.

[76] B. Cramariuc, I. Shmulevich, M. Gabbouj, and A. Makela, "A new image similarity measure based on ordinal correlation," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 3, 2000, pp. 718–721 vol.3.

[77] J. Malo, A. Pons, and J. Artigas, "Subjective image fidelity metric based on bit allocation of the human visual system in the dct domain," *Image and Vision Computing*, vol. 15, no. 7, pp. 535 – 548, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0262885696000042

[78] Y.-K. Lai and C.-C. Kuo, "A haar wavelet approach to compressed image quality measurement," *Journal of Visual Communication and Image Representation*, vol. 11, no. 1, pp. 17 – 40, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1047320399904339

[79] L.-W. Kang, C.-Y. Hsu, H.-W. Chen, C.-S. Lu, C.-Y. Lin, and S.-C. Pei, "Feature-based sparse representation for image similarity assessment," *Multimedia, IEEE Transactions on*, vol. 13, no. 5, pp. 1019–1030, Oct 2011.

[80] M. Luengo-Oroz and J. Angulo, "Cyclic mathematical morphology in polar-logarithmic representation," *Image Processing, IEEE Transactions on*, vol. 18, no. 5, pp. 1090–1096, May 2009.

[81] A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Society for Industrial and Applied Mathematics, 2001. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/1.9780898719277

[82] S. Watanabe, *Pattern Recognition: Human and Mechanical*. New York, NY, USA: John Wiley & Sons, Inc., 1985.

[83] P. Goel, S. Prasher, R. Patel, J. Landry, R. Bonnell, and A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Computers and Electronics in Agriculture*, vol. 39, no. 2, pp. 67 – 93, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169903000206

[84] C. Ozkan and F. S. Erbek, "The comparison of activation functions for multispectral landsat tm image classification," *Photogrammetric Engineering & Remote Sensing*, vol. 69, no. 11, pp. 1225–1234, 2003.

[85] M. Awad, "Sea water chlorophyll-a estimation using hyperspectral images and supervised artificial neural network," *Ecological Informatics*, vol. 24, no. 0, pp. 60 – 68, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574954114000867

[86] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: http://dx.doi.org/10.1023/A:1022627411411

[87] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for support vector machines," *Machine Learning*, vol. 46, no. 1-3, pp. 131–159, 2002. [Online]. Available: http://dx.doi.org/10.1023/A%3A1012450327387

[88] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. McLachlan, A. Ng, B. Liu, P. Yu, Z.-H. Zhou, M. Steinbach, D. Hand, and D. Steinberg, "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008. [Online]. Available: http://dx.doi.org/10.1007/s10115-007-0114-2

[89] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *Neural Networks, IEEE Transactions on*, vol. 13, no. 2, pp. 415–425, Mar 2002.

[90] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S002200009791504X

[91] J. Carletta, "Assessing agreement on classification tasks: The kappa statistic," *Comput. Linguist.*, vol. 22, no. 2, pp. 249–254, Jun. 1996. [Online]. Available: http://dl.acm.org/citation.cfm?id=230386.230390

[92] C. BRI. (2013) Bakery product development. Available at: www.campdenbri.co.uk/services/bakery-product-development.php. [Online]. Available: http://www.campdenbri.co.uk/services/bakery-product-development.php

[93] P. J. Curran, "Remote sensing of foliar chemistry," *Remote Sensing of Environment*, vol. 30, no. 3, pp. 271 – 278, 1989. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0034425789900692

[94] M. Fauvel, Y. Tarabalka, J. Benediktsson, J. Chanussot, and J. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, March 2013.

[95] D. Wu and D.-W. Sun, "6 - food colour measurement using computer vision," in *Instrumental Assessment of Food Sensory Quality*, ser. Woodhead Publishing Series in Food Science, Technology and Nutrition, D. Kilcast, Ed. Woodhead Publishing, 2013, pp. 165 – 195e. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780857094391500067

[96] G. Muhammad, "Date fruits classification using texture descriptors and shape-size features," *Engineering Applications of Artificial Intelligence*, vol. 37, pp. 361 – 367, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0952197614002413

[97] A. Kavianpour, N. Bagherzadeh, and S. Shoari, "Finding elliptical shapes in an image using a pyramid architecture," in *Speech, Image Processing and Neural Networks, 1994. Proceedings, ISSIPNN '94., 1994 International Symposium on*, Apr 1994, pp. 445–448 vol.2.

[98] X. Ben, W. Meng, and R. Yan, "Dual-ellipse fitting approach for robust gait periodicity detection," *Neurocomputing*, vol. 79, no. 0, pp. 173 – 178, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092523121100628X

[99] C. Mineo, S. G. Pierce, P. I. Nicholson, and I. Cooper, "Robotic path planning for non-destructive testing a custom {MATLAB} toolbox approach," *Robotics and Computer-Integrated Manufacturing*, vol. 37, pp. 1 – 12, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0736584515000666

Publications by Author

[1] Timothy Kelman, Jinchang Ren, and Stephen Marshall. Effective classification of chinese tea samples in hyperspectral imaging. *Artificial Intelligence Research*, 2: 87–96, 2013.

[2] Timothy Kelman, Jinchang Ren, and Stephen Marshall. Rotational scanning techniques for hyperspectral imaging. In *Proceedings of Hyperspectral Imaging Conference*, 2012.

[3] Timothy Kelman, Jinchang Ren, and Stephen Marshall. Classification of chinese teas samples for food quality control using hyperspectral imaging. In *Proceedings of Hyperspectral Imaging Conference*, 2011.

[4] Stephen Marshall, Timothy Kelman, Tong Qiao, Paul Murray, and Jaime Zabalza. Hyperspectral imaging for food applications. In *Proceedings of European Signal Processing Conference (EUSIPCO)*, 2015.