SNAPSHOT SPECTRAL IMAGING USING IMAGE REPLICATION AND BIREFRINGENT INTERFEROMETRY: PRINCIPLES AND APPLICATIONS.

by

Alistair S. Gorman

Submitted for the Degree of
Doctor of Philosophy
at Heriot-Watt University
on completion of research in the
School of Engineering and Physical Sciences
October 2010

Heriot-Watt University

Abstract

SNAPSHOT SPECTRAL IMAGING USING IMAGE REPLICATION AND BIREFRINGENT INTERFEROMETRY: PRINCIPLES AND APPLICATIONS.

by Alistair S. Gorman

This thesis explores the image-replicating imaging spectrometer (IRIS). This relatively recent invention is a two-dimensional, snapshot spectral-imaging technology, capable of recording the spectral and spatial data from a scene instantaneously. Whereas conventional spectral-imaging technologies require multiple detector frames to record the entire data set, IRIS is able to record the data set in a single frame, a capability which is useful for highly dynamic scenes.

The IRIS concept and the design of IRIS systems are explained in detail, and constraints on the performance of IRIS are determined. Practical issue in the use of IRIS systems are identified and solutions are identified and appraised. Some applications of IRIS are also shown, demonstrating its viability as a spectral imaging technology.

Novel aspects of this work include the refinement of the IRIS design, demonstration of a registration algorithm for IRIS, designs for achromatic Wollaston prisms, a comparison of the IRIS technology with conventional spectral imaging technologies, and the application of IRIS to practical problems.

Acknowledgements

This work has been funded by EPSRC and Selex Galileo, I am grateful to these organisations for making this research possible.

I would like to thank Professor Andrew Harvey, my supervisor, for giving me the opportunity to conduct this research and for his patience and support in helping me complete this work. I would also like to thank my past and present colleagues in the Imaging Concepts Group; in particular David Fletcher-Holmes for his previous work on the topic and for all his help at the start of my research, and Gonzalo Muyo and Eid Alabboud for their friendship and help. The late Sonny Ramachandran was also a good friend from the beginning of my time at Heriot-Watt University and is fondly remembered. I'd also like to thank Peter Sinclair and my other collaborators at Selex for their help and support.

I am grateful to my parents Marion and Peter, and the rest of my family, especially Mike and Li. I am looking forward to having more time to spend with my family, including my brother Stuart, my sister Amy, and my nephew Sean.

I am also extremely grateful to Laura Valkonen for her patience and support during the difficult years of the PhD, and for her excellent proof reading skills.

Alistair Gorman
October 2010

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Summary: This chapter provides an review of existing spectral-imaging technologies in order to put the IRIS technology into context. An overview of spectral-imaging applications and processing techniques is also given.

1.1 The need for snapshot spectral imaging

The use of electromagnetic radiation to relay signals is well known, be it by sending radio waves through space, laser light through optical fibers or even semaphore. Messages are encoded into electromagnetic radiation by modulating its properties. By combining measurements of these properties with appropriate knowledge, a remote observer may be able to recover the intended message. Nature encodes information in radiation in a similar manner; imprinted on a flux of photons are details of the distribution of matter and energy involved in the emission and transport of these photons. Physical models of these processes represent the knowledge required for decoding this information from the measurements.

To obtain the most complete radiometric observation possible requires that the momentum and spin of each photon, together with the time and position of detection, be measured with resolution up to whatever limit nature imposes. This is impractical for a number of reasons and so it is the macroscopic properties that are often measured, namely the instantaneous power delivered at each frequency, in each polarisation state, in each direction, at each point in space. Even at this macro level it is only feasible, and for most purposes sufficient, to measure a subset of these properties at any time.

In terms of information content, the so-called optical wavelengths, stretching from around 300 nm to 15 μm, are of interest for a number of reasons. Firstly, the Earth's atmosphere is highly transparent to much of this wavelength range, making it possible to detect radiation at long distances from the source. Secondly, as this region also includes the peak thermal emission from the Sun and Earth there is a large amount of radiation available for detection, many artificial sources of radiation also exist for these

wavelengths. Finally, the photon energies in the optical regime also correspond to many atomic and molecular transitions which leave characteristic features in the spectrum of radiation emitted, absorbed and scattered by matter. If it is supposed that the set of transitions is unique for each chemical compound then it is possible to use spectral measurements (i.e. measurements of the distribution of photon energies) to discriminate between, and identify, different types of matter.

Spectral imaging involves the acquisition of both spatial and spectral information from a scene. The human visual system can be considered a type of spectral-imaging system. The image formed at the retina by the cornea and lens is sampled by an arrangement of photosensitive cells. In moderate and high levels of light, three distinct types of photoreceptor cells are active, each type with a different spectral sensitivity. Comparing firing rates for these cells helps the brain to create the perception of colour. The majority of imaging technologies in the optical region work in a similar manner to the eye, using a system of lenses or mirrors to create an image over a 2D array of detectors. As with the eye, these detectors are effective only over a finite region of the spectrum, though unlike the eye these arrays cannot differentiate detected photons by energy.

With thousands of detectors working concurrently, 2D panchromatic arrays are the most efficient way to collect for spectral imaging. Due to the inability of these detectors to differentiate between different photon frequencies, the spectral distribution of radiation must be modulated in some known way prior to detection, for instance by blocking all but a narrow range of frequencies from reaching the detector. To date, the majority of spectral-imaging systems have not been able to acquire the data for contiguous 2D areas in a single frame, instead requiring multiple acquisitions to record the entire data set. The need to sequentially acquire the data causes difficulties if the object space changes more rapidly than the required scan time. This challenge has motivated the development of a number of snapshot spectral-imaging instruments, able to collect the entire data set for a contiguous area in a single frame.

The subject of this thesis is a recently developed snapshot technology, the image-replicating imaging spectrometer (IRIS) [10, 11], which uses birefringent materials to create an array of spectrally distinct images at the detector. To put IRIS into context within the field of spectral imaging, the remainder of this chapter will provide an overview of spectral-imaging technologies and applications.

1.2 A review of spectral imaging

Spectral imaging is the convergence of the ability to record images and the study of spectroscopy, the goal being to record a stack of spectrally distinct images (the 'data cube') of a given scene. This chapter offers a survey of historical advances and current

Figure 1.1: Early (1905-1915) colour photograph by Sergei Mikhailovich Prokudin-Gorskii. Colour composite and separate blue, green, and red frames are shown. [1].

technology, beginning with the first attempts to record colour information. Current detector technology is then summarised, followed by a discussion of the immature technology of energy-resolving detectors, which if developed further may prove to be useful for spectral imaging in future. Finally, current techniques for spectral imaging are reviewed, including the small number of snapshot spectral-imaging instruments that currently exist.

1.2.1 Colour imaging

The earliest recorded success in spectral imaging dates to James Clerk Maxwell's slightly fortuitous experiments in colour photography [12]. Maxwell intended to record three separate images from the spectrum; one each in the red, green, and blue spectral bands of the three-colour model proposed by Thomas Young. To do this, Maxwell used three photographic plates with separate coloured-glass filters[1]. The images from the developed plates were then superimposed by projection onto a white surface using lanterns and similar coloured filters. The mixture of light reflected from this surface into the eye resulted in a colour image approximating the scene for a human observer. This technique continued to be refined and used to great effect in still photography and motion pictures. The expense and inconvenience of multiple emulsions was reduced somewhat with the introduction of Autochrome Lumiére [2], which consisted of a random array of microscopic filters in front of a single panchromatic emulsion (Figure 1.2(a)). Once de-

[1]Maxwell's fortune was that although the photographic emulsion he was using was insensitive to red light, the red parts of his object, a tartan ribbon, reflected ultraviolet light which was also passed by the red filter, exposing the film well enough to reproduce the colour image [13]

veloped, the photograph was viewed through the same filter screen to decode the colour information. Chemical photography continued to improve over the 20th century, with the introduction of monopack colour films consisting of several layers of dye-coupled emulsions, each absorbing a portion of the incident spectrum while transmitting the remaining light to subsequent layers.

One notable colour-photographic technique that goes someway toward recreating the distal stimulus of radiation at the image plane, and not just an appropriate proximal stimulus in humans, is the Lippmann plate. Developed in 1891 by Gabriel Lippmann, this process uses the interference of radiation to set up standing waves in a photographic emulsion at half the wavelength of the light [14]. After processing, the changes in the emulsion due to the standing waves become refractive index differences in a gelatin sheet; Bragg diffraction from these structures gives rise to colour images when viewed in diffuse light. The invention of this true snapshot spectral-imaging system earned Lippmann a Nobel prize in Physics, but was not practical due to limitations in the materials, and because of difficultly viewing and copying the images. At the time of its invention, it would not have been possible to access or use the vast amount of information in a Lippmann plate. Advances in telecommunications, electronics, and computing would begin to provide the ability to cope with large data sets, and spectral imaging would gradually emerge as a viable scientific tool.

Although in some senses, particularly aesthetics, chemical film has still not been surpassed, the inconvenience and expense of processing make it unattractive for many applications. Calibration of emulsion for scientific measurements is also difficult. In the early 20th century electrical and electronic technologies led to more convenient ways to capture, store and transmit images, aiding the rise of television, video, and digital photography. Today electronic cameras are commonplace and computers are routinely used to manipulate and display images. These sensors cannot resolve photon energy and so digital colour photography is performed in a similar manner to early colour photography. High-quality digital colour cameras use a separate detector array for each RGB component, while lower-quality systems use a single array, with multiplexed bands, or with an array of microscopic filters known as a Bayer filter (Figure 1.2(b)), which is analogous to the Autochrome system. These techniques work well for a small number of colour channels, but can become impractical as the desired numbers of spectral bands rise. There are a variety of spectral-imaging methods, these will be discussed in Sections 1.2.5 and 1.2.6 after a review of current detector technologies.

1.2.2 Optical detectors

At radio wavelengths, the phase and intensity information of radiation can be directly sensed and measured using an antenna connected to a receiver and an amplifier. Imaging is carried out either by scanning a single detector with a small field of view, or by using

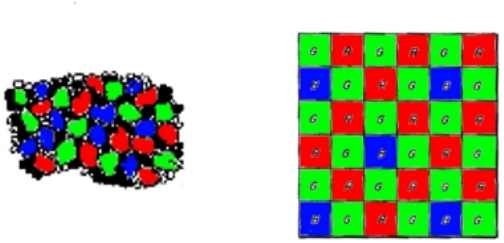

Figure 1.2: Illustrations, from the original patents, of two examples of colour filter arrays, with colour added: (a) small section of an Autochrome Lumiére screen [2], (b) a section of a Bayer filter matrix pattern as used in many colour digital cameras, each individual filter element is aligned with a single pixel; the relative proportions of colours are chosen to approximate the response of the human eye [3].

a large array of such detectors. At X-ray energies it becomes possible to measure the energy of detected photons. At optical wavelengths it is all but impossible to detect phase or energy using current sensors alone; however, the detector remains a important part of any spectral imaging system and an understanding of the technologies is therefore vital.

Single-colour detectors

Modern imaging detectors for the infrared (IR), visible and ultraviolet (UV) frequencies are sensitive over some finite band and are unable to resolve wavelength. These sensors are single-colour insomuch that for each image element a single intensity value is returned. The detectors function via the generation of photoelectrons within a wafer of semiconductor material. The photoelectrons generated by incoming light are captured and stored in thousands of microscopic potential wells which are formed by depositing electrodes on the surface of the wafer. The probability of a photon with wavenumber k being detected having entered the bulk of the semiconductor is known as the quantum efficiency $\eta(k)$. If the number of photons entering the detector at time t is $I(k,t)$, the expectation value of the electron count resulting from these photons is given by

$$\langle N \rangle = \int_{k_1}^{k_2} \eta(k)\, I(k,t)\, dk + \epsilon, \tag{1.1}$$

where ϵ is the average bias on the measurement. The uncertainty on the measurement is due to sampling error and noise introduced by the detection process.

The most common type of detector array for use with visible radiation is the charge coupled device (CCD), a two-dimensional regular array of quantum wells (pixels) within a wafer of silicon; the charge generated by photons near or in each well is stored until shift registering is used to read out the frame. CCDs offer much higher quantum efficiencies

than film and are sensitive to radiation up to around 1 μm; they are often cooled to reduce noise due to thermal excitation within the detector and to allow longer integration times. The quantum efficiency of CCD arrays can be increased by thinning the silicon and exposing the circuitry-free side of the chip to the incoming radiation; these are known as back-illuminated arrays. As higher energy photons are absorbed near the surface of the silicon, thinning greatly increases the sensitivity to shorter wavelengths; longer wavelengths are absorbed over a wider range of depths so there are smaller gains in the IR region. While CCDs offer the best quality images [15], CMOS sensors are also gaining in popularity, due in part to improvements in the shuttering technology that had previously limited them. CMOS detectors are also silicon-based but consume less power than CCDs and offer some on-chip processing capability and random access of pixels, allowing independent acquisition of multiple regions of interest, high frame rates are also possible. Standard silicon detectors are sensitive from around 300 nm to 1.1 μm. To sensitise silicon detectors to UV photons, phosphorescent coatings are used.

Arrays of photoelectric detectors are also common for imaging at wavelengths from 1 μm to 12 μm; InSb, InGaAs, PbS, PtSi and HgCdTe are popular photosensitive materials for this region [7,8]. A separate silicon circuit, in contact with the detector material, is commonly used to read the charge from each pixel. The typical width of spectral sensitivity for these photoelectric materials is around 3 μm [7], meaning more than one array is needed to cover the entire infrared region. Many IR cameras operate in one of two atmospheric windows, occurring at approximately 3 μm to 5 μm and 8 μm to 14 μm. At temperatures typically encountered in the Earth's atmosphere, the thermal emission of matter is strongest in the infrared; photoelectric detectors therefore tend to be cooled, and cold stops or more sophisticated cryogenic systems are used to reduce unwanted radiation. For thermal infrared detection, micro-bolometer arrays made from V_3O_5 or silicon are often used as an alternative to photoelectric devices [7]. Quantum-well infrared photo-detectors (QWIP) are another type of photon detector. These devices are constructed from thin layers of semiconductors with different bandgaps. Through careful choice of the materials and their thicknesses, a potential well with two discrete energy levels can be formed. The higher of these energy levels can be made to align with the edge of the potential well by applying a bias voltage. When a photon with the appropriate energy is incident upon the well structure, an electron is excited to the conduction band giving rise to a photocurrent. QWIP are an active area of research, as the materials can be grown with high uniformity and the spectral sensitivity is easily controlled; however, quantum efficiencies are currently low. Table 1.1 summarises the important properties of the different detector array types.

Detector type	Wavelength range	Maximum array size	Pixel size	η
Si	0.3 μm - 1.1 μm	9216×9216 [16]	7 μm	65%
InSb	1 μm - 5.5 μm	1024×1024 [17]	30 μm	45%
HgCdTe	2 μm - 15 μm	2048×2048 [18]	30 μm	65%
InGaAs	1 μm -2.6 μm	640×512 [19]	30 μm	85%
PbS	1 μm -3 μm	320×240 [20]	30 μm	50%
PtSi	1 μm -5 μm	640×480 [20]	30 μm	10%
Bolometer	1 μm - 20 μm	640×480 [6]	30 μm	80%
QWIP	2 μm - 15 μm	640×480 [6]	20 μm	20%

Table 1.1: Properties of single-colour detector arrays [6–9].

Two-colour detectors

For measurement over a wider spectral range, two-colour detectors have been produced; these output two intensity values for each image element. The simplest of these devices consist of a stack of two detectors. The top layer detects a portion of the incident spectrum but is transparent to wavelengths to which the bottom layer is sensitive [7,21]. The dual outputs from each pixel can be read by multiplexing or by using two read circuits for true simultaneous measurement. As well as offering an extended spectral bandwidth these detectors can be used for crude spectral discrimination.

Voltage-tunable arrays have also been developed for use as two-colour detectors in the long-wave infrared region [22]. Again these sensors can detect radiation from separate bands, the peaks of which can be tuned through an applied voltage bias, allowing the spectral sensitivity to be adjusted.

Tunable detectors are manufactured from stacks of quantum-well infrared photodetectors (QWIP) [23]. Electron transfer between ground states of adjacent quantum wells under a voltage bias is used to shift the peak sensitivity. Some tunable instruments offer only a single channel at a time, the bias voltage being used to switch the sensitivity between two singularly-peaked functions, with the intermediate states exhibiting a double peak. Other devices exhibit simultaneous, tunable dual-bands [24]. These detectors therefore enable sequential or simultaneous two-colour imaging with peak wavelengths that can be shifted by anything up to several microns. As an example, a 640×512 pixel tunable array operating between 3 μm and 5 μm, has a peak spectral sensitivity that can shift from 4.1 μm to 4.7 μm [25].

Four-colour elemental detector arrays based on QWIP have also been reported in the literature [26]. The current state of the art is a 640×512 pixel array with simultaneous readouts in four bands with peaks of 4 μm, 5 μm, 8 μm and 9.7 μm [27].

Although the intended use of these detectors is to cover a wider spectral range, the bands and tunability can be used for crude spectral discrimination.

1.2.3 Energy-resolving detectors

Energy-resolving CCDs currently exist for use at X-ray wavelengths. Like proportional counters, these detectors utilise the ability of high-energy photons to create electron cascades within the detector. The cascade effect is proportional to the energy of the incoming photon.

Superconductors are also utilised in energy-resolving detectors; three different types of superconducting detector are currently in development: superconducting tunnel-junctions (STJ) [28], transition-edge sensors (TES) [29], and kinetic-inductance detectors (KID) [30]. STJ and KID devices rely on the breaking of Cooper pairs by photons. In STJ devices this particle breaking results in electrons tunneling across a voltage-biased junction, which in turn leads to a photocurrent. In KID devices particle breaking results in measurable changes in the the complex conductivity of the superconductor. TES exploit the strong temperature dependence of the electrical resistance of the superconductor when it is cooled to below its critical temperature and an electrical bias is applied; incoming photons increase the temperature of the material and measurements of the resistance can be used to determine the energy falling on the detector. These devices are of considerable interest as the band gap in superconductors can be of similar magnitude to visible and infrared photon energies, making them suitable for use over large spectral ranges and out to IR wavelengths. Such detectors offer moderate spectral resolution with high quantum efficiencies over a broad bandwidth.

A 64-element array has been constructed using 20 μm square, TES detectors giving a spectral resolution of 40 nm (FWHM) in the visible, and a temporal resolution of 100ns [29]. The signal readout is achieved by using an array of superconducting quantum interference devices (SQUID); in order to achieve array sizes larger than a few hundred pixels a multiplexing scheme [29] is required. KID and STJ based devices may become favored over TES devices which require extremely low temperatures for good spectral resolution. A multiplexed readout of KID detectors by means of a microwave probe [31] may solve some of the engineering difficulties that currently restrict array size. A 36-element array of STJ detectors has also been constructed [32] with spectral resolution of around 30 nm in the visible, and temporal resolution of 100 μs. A multiplex technique based on coincidence measurements across rows and columns of an STJ array has also been demonstrated, although this strategy currently reduces the total count-rate of the array [33].

The fundamental drawback of these superconductor-based techniques is the extremely low operating temperature required (<1K). Work is being carried out to improve the spectral resolution and increase detection rates, and engineering issues surrounding large array sizes may be overcome for each technology. It remains to be seen what the limits of these technologies are.

If practical energy-resolving detector arrays can be developed, they will undoubtedly become the instrument of choice for spectral imaging. At present two-dimensional single-colour detector arrays offer the most efficient means of making radiometric measurements. If this efficiency is to be utilised, some method for decoding the spectral information from raw photon counts is required; the remainder of this section gives a review of these techniques.

1.2.4 Spectral imaging using multiple detector arrays

Electronic detector arrays are now inexpensive and it is feasible to use several arrays, each equipped with a separate colour filter, for snapshot spectral imaging. This technique is used in three-CCD cameras which use dichroic beam-splitters to split a beam into the three RGB components and direct each colour to a separate detector [34]. Although conceptually simple, such a system requires multiples of the power and space required for the single detector, making it unattractive for large numbers of bands. Calibration of multiple detector arrays is also more complicated, making it harder to meaningfully combine the data for scientific use. If the detectors do not share a common optical system, as is likely in the case of large numbers of bands, there may be difficulties in retrofitting such devices to existing imaging systems. Non-boresighted multi-camera systems are only suitable when the scene can be approximated as a two-dimensional space so that differences in viewing geometry can be neglected; this is often the case for airborne or satellite-based sensors and for other measurements made at distances far from the scene.

In the following review of technology, the discussion will be limited to techniques that use a single two-dimensional detector array.

1.2.5 Acquiring the data cube sequentially

As current detector arrays are not energy-resolving, the radiation must be modulated in some known way prior to detection. At present, spectral-imaging systems have largely been adaptations of spectroscopic techniques initially intended to measure radiation from a spatially homogeneous source. The most convenient adaptations of these techniques to imaging have resulted in instruments that require some form of temporal scanning in order to collect the data set for a contiguous two-dimensional area.

The majority of current spectral-imaging techniques work by recording the spectral

information independently for each spatial element. Regardless of how the sampling is performed, the set of spectral measurements for a single spatial element can be represented by a discretised Fredholm equation. With k the wavenumber and $x(k)$ the flux of incident radiation, the expectation value y^j, of the jth measurement, recorded with a detector with quantum efficiency $\eta(k)$, can be written as

$$y^j = \int_{k_1}^{k_2} \eta(k)\, \tau^j(k)\, x(k)\, dk. \tag{1.2}$$

Where $\tau^j(k)$ is the spectral transmission of jth spectral band, which is defined as the fraction of incident light reaching the detector after passing through the optical system. The set of j measurements are often represented as a measurement vector $\mathbf{y}$. Spectroscopic techniques are often classified as being either direct or transform [35] according to the general form of $\tau^j(k)$. For direct techniques each band records only a small range of wavelengths within the detector spectral range, whereas for transform techniques each measurement typically records radiation from across the entire spectral bandwidth of the system. The most common type of transform instruments are Fourier-transform spectrometers (FTS) for which the transmission functions are cosinusoidal in form. The remainder of this section will review current and past spectral-imaging technology. A summary of the important properties of each technique is given at the end of this section in Table 1.2.

Direct spectral-imaging technologies

Dispersive instruments

Some of the earliest investigations of the electromagnetic spectrum of radiation were carried out using dispersive prisms, which cause collinear beams of different wavelengths to diverge. If the collimated radiation from a point source is passed through a prism and then an imaging lens, a dispersed image of the source is formed at the image plane, which can be sampled using a linear detector array. With an area array, a line of point sources produced using a slit at a conjugate image plane can be measured simultaneously as in Figure 1.3. By scanning the image across the slit in the direction of dispersion, a two-dimensional area can be sampled [36]. The spectral bandpass of prism based dispersive systems varies with wavelength and can be calculated from the focal length of the imaging lens, the pixel area, and the angular dispersion of the prism. The prism dispersion can be predicted from the prism geometry and the dispersion relation for the prism material, the dispersion is not in general a linear function of wavelength. The f-number is constrained by the lens and prism apertures and the transmission of prism-based dispersive systems is typically $> 95\%$.

Diffraction gratings can be used in an almost identical manner to prisms [37] and are often preferred due to the linear dispersion of radiation and higher spectral resolutions

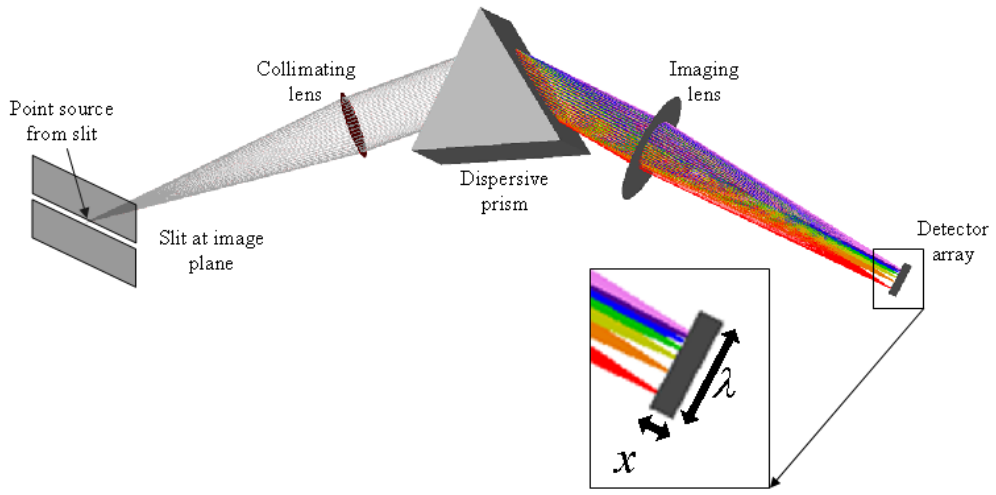

Figure 1.3: Ray-trace of a prism based dispersive system.

achievable. Collimated monochromatic radiation incident upon a grating is diffracted by the individual elements; when brought to focus at the detector, the radiation interferes to create a number of sharp maxima (or orders) at the image plane, as shown in Figure 1.4(a). A portion of an order not overlapping with radiation from another order can be isolated with a detector array (1.4(b)). The bandpass of the grating-based system again depends on the imaging lens, the pixel size, and the angular dispersion of the grating. According to diffraction theory, the dispersion can be calculated by computing the Fourier transform of the convolution of the individual slit shape with a comb of delta functions, with pitch equal to that of the slits. Blazed reflection gratings are often used to concentrate the radiation into a more convenient order. Grating efficiencies are often a complicated and rapidly changing function of wavelength. Over a narrow range of wavelengths peak efficiency can be as high as 90% for ruled gratings, and 60% for holographic gratings. In general, the efficiency also depends on the polarisation state.

Dispersive spectrometers are often used in moving platforms, where the motion can be exploited to scan the object space as shown in Figure 1.5; this is referred to as pushbroom scanning [38].

Filter based systems

Dispersive spectrometers collect a contiguous region of the spectrum. If only certain sections of the spectrum are of interest, then it may be preferable to extend the three-colour technique of recording multiple images using filters with different passbands. This is usually achieved by multiplexing the desired wavelengths onto a single detector array to record a series of monochromatic images. Instead of coloured glass, as used by Maxwell to record his colour photograph, modern filters usually work via interference. These filters tend to have higher peak transmissions and lower nulls and are more versatile than absorption filters. Interference filters are made using layers of thin films

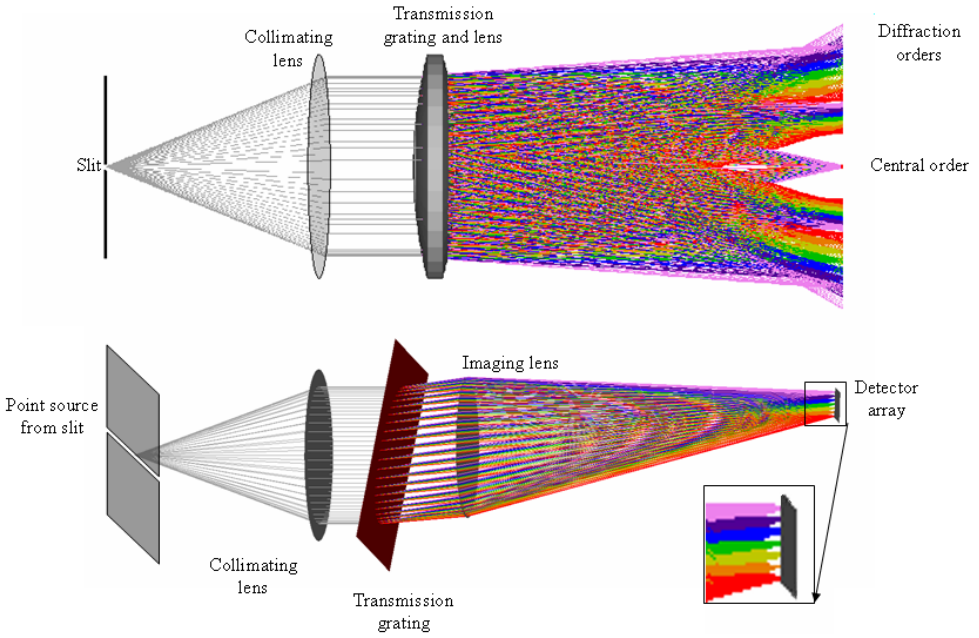

Figure 1.4: Grating based dispersive systems: (a) multiple diffraction orders, (b) isolating a single order with a detector array.

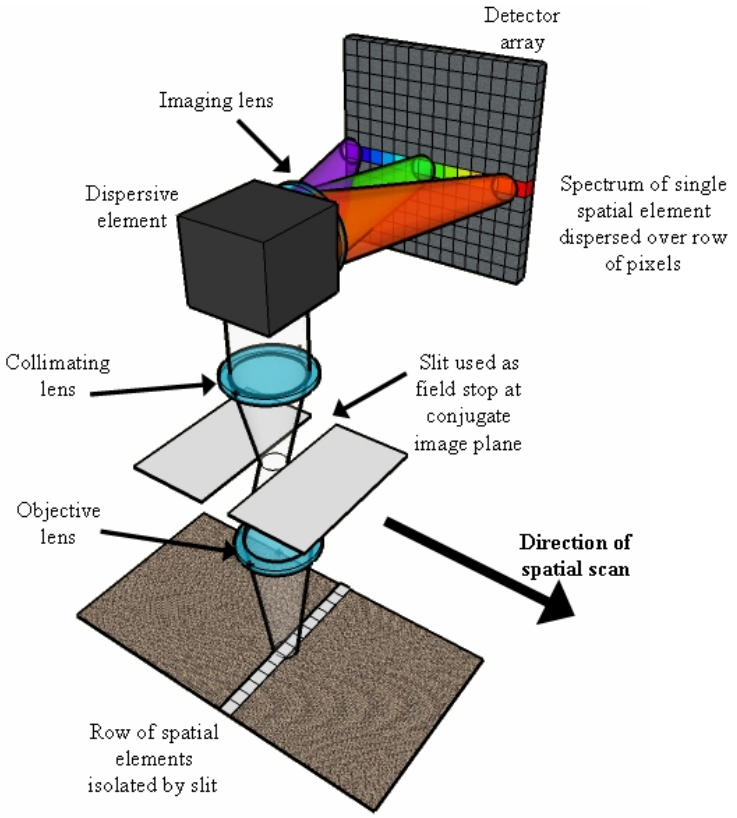

Figure 1.5: Pushbroom scanning with a dispersive system.

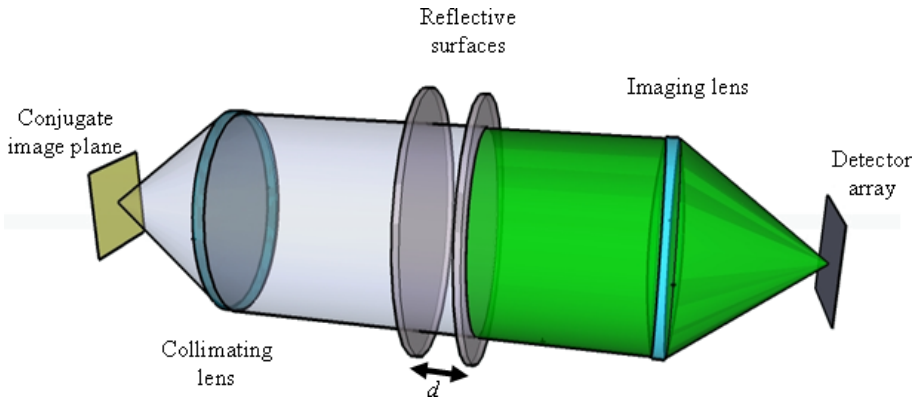

Figure 1.6: Schematic of an etalon used in a collimated beam.

chosen to give the desired transmission characteristics [39].

The most basic, single-film, interference-filters work in an identical manner to an etalon, which consists of two parallel surfaces, each with intensity reflectance R, and separated by a distance d (Figure 1.6). When radiation is incident upon an etalon, it is reflected multiple times by the surfaces and a portion exits the etalon in the forward direction. Due to the path difference between the different beams, the radiation interferes such that for wavenumber k the transmission can be written as

$$\tau\left(k\right) = \frac{\left(1-R\right)^2}{1+R^2-2R\cos 2kd}. \tag{1.3}$$

With the appropriate choice of parameters this function will take the form of a series of sharp maxima in the transmission function (Figure 1.7). A single peak of this function can be isolated using short- and long-pass filters. The transmission functions for multi-layer filters can be calculated from the single-layer equation and are generally complicated in form, though they are usually designed to pass only a narrow region of the spectrum. In this case the transmission functions can be approximated by a rectangular or Gaussian function with little error. As the interference of radiation depends on the path differences in the beam, the transmission of interference filters varies with incidence angle. As the filter is tilted, the peak transmission shifts towards shorter wavelengths until at some angle multiple peaks arise. This sensitivity to the angle of incidence means that interference filters should not be used in strongly convergent beams, and should preferably be used in collimated beams containing a narrow range of angles. The variation of transmission with angle is also true of dispersive instruments, but is less of an issue in operation, as the use of a slit as a spatial filter constrains the range of angles incident upon the prism or grating. Peak transmissions of etalons and interference filters can be $> 95\%$.

The Lyot filter [40] is a type of birefringent, polarising interference-filter, which consists of $M+1$ co-aligned linear polarisers interspersed by M retarders, with the

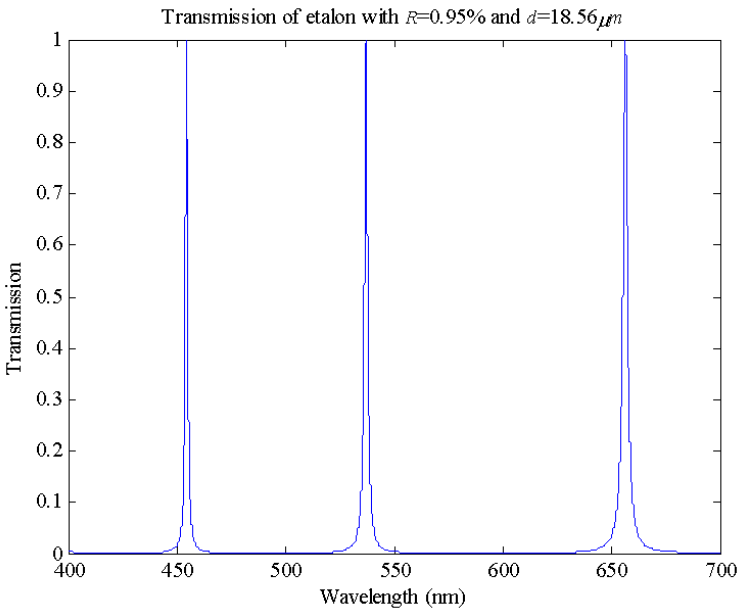

Figure 1.7: Typical transmission of an etalon.

transmission axis of the polarisers at 45° to the retarders' optic axes (Figure 1.8). The birefringence of the retarder material is wavelength-dependent, therefore the polarisation state of linearly-polarised radiation is coupled to the frequency and so is transmitted with varying efficiency by the polariser following the retarder. If the thickness of the mth retarder is d_m and the birefringence is B_m, the transmission function can be written

$$\tau(k) = \prod_{m=1}^{M} \cos^2(\pi B_m(k)\, d_m k). \tag{1.4}$$

With the path differences increasing by a factor of two at each stage the transmission becomes a series of sharp peaks. As for the etalon, a single peak is isolated with high- and low-pass filters. The peak transmission of a Lyot filter can be greater than 90% in linearly polarised light. The closely related Solc filter [41] consists of an even number of identical retarders of thickness d, between two co-aligned linear polarisers, and exists in either the fan or folded configuration. For the folded Solc filter the mth retarder optic axis is at an angle of $(-1)^{m-1}\phi$ to the transmission axes of the polarisers, where ϕ is an arbitrary angle. In the fan configuration, the angle between the polariser axis and the mth retarder optic axis is $(2m-1)\phi$. The transmission of a folded Solc filter is given by Equation 1.5; the transmission of the fan configuration is similar, but the phase retardation is shifted by π [42].

$$\tau(k) = \left[\tan(2\varphi)\cos(\chi)\frac{\sin(N\chi)}{\sin(\chi)}\right]^2, \tag{1.5}$$

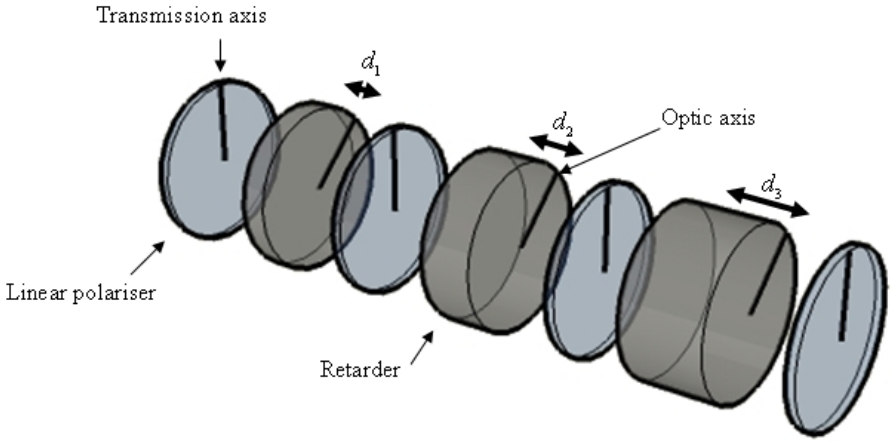

Figure 1.8: Schematic of a Lyot filter.

with

$$\cos(\chi) = \cos(2\varphi)\sin(B(k)\,dk/2).$$

Using separate filters to record each spectral band requires some form of switching to collect the entire data set. As a way of avoiding the need for a large number of filters and some of the associated mechanical issues, a single variable filter can be used. Variable filters are interference filters and are usually either circularly or linearly variable, such that the transmission changes with angle around or distance along the filter. Circularly variable filters offer fast random access, while linearly variable filters can be used in moving platforms in a similar manner to dispersive instruments. No slit is required for this type of operation, but the filter must be used directly in front of the detector or at a conjugate image plane. A wedge filter, an etalon with non-parallel surfaces, can be used in a similar pushbroom arrangement.

Another way of reducing the number of filters required is to use a tunable filter. One such filter exploits the change in transmission of interference filters with angle of incidence [39]; for this a mechanism for tilting the filter by a known amount is required. Etalons can also be tunable [43]: altering the separation between the reflective surfaces changes the path lengths and therefore the transmission. There are also two common types of electronically-switchable tunable filters [44]: the liquid-crystal tunable filter (LCTF) and the acousto-optic filter, both of these designs offer the benefit of no mechanical parts and fast random access.

The LCTF is a Lyot filter, with electronically switchable liquid crystal retarders following each solid retarder, the ability to alter the retardation at each stage allows for selection of transmission [45]. Peak transmission varies with wavelength but is typically less than 70% in polarised light and can be as low as 10%.

Acousto-optic filters rely on the interaction of radiation with sound waves in an anisotropic material, and can be used for spectral imaging in two configurations: collinear

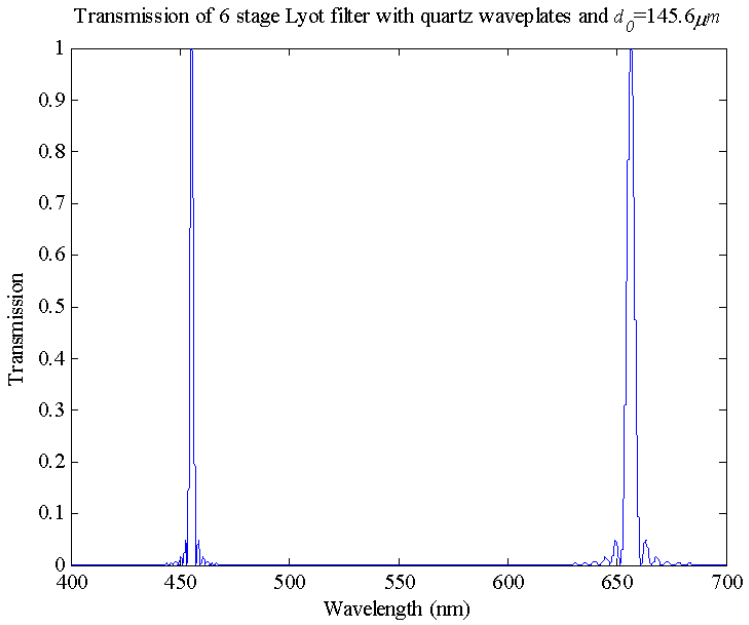

Figure 1.9: Transmission of a Lyot filter.

and non-collinear [46]. In the collinear mode, linearly polarised radiation interacts with the changes in refractive index caused by the acoustic waves, which propagate in the same direction as the radiation. The result of this acousto-optic interaction is that the wavelength of radiation is coupled to the polarisation state of the beam, in a similar manner to the Lyot filter. A polariser is again used to select portions of the spectrum. In the non-collinear configuration, the acoustic wave propagates in a different direction to the radiation. In this case the refractive index pattern caused by the sound wave in crystalline material acts like a 3D diffraction grating, causing certain wavelengths to be refracted through a different angle to the others. With the appropriate acoustic wave the desired parts of the spectrum can be isolated using stops. This configuration can be used in unpolarised light but the orthogonal components diverge upon exiting the prism, so only a single component is selected. Efficiencies for both arrangements can be as high as 90% in polarised light.

Transform instruments

A transform instrument is most generally defined as one in which each spectral band records wavelengths from over the entire bandwidth. Any of the filter-based techniques discussed in the previous section can in theory be used as transform techniques with the appropriate choice of transmission function. Most often the term refers to imaging Fourier transform spectrometers (IFTS) or the closely related tomographic spectral-imaging techniques.

IFTS come in a variety of designs but like direct instruments, can be separated into

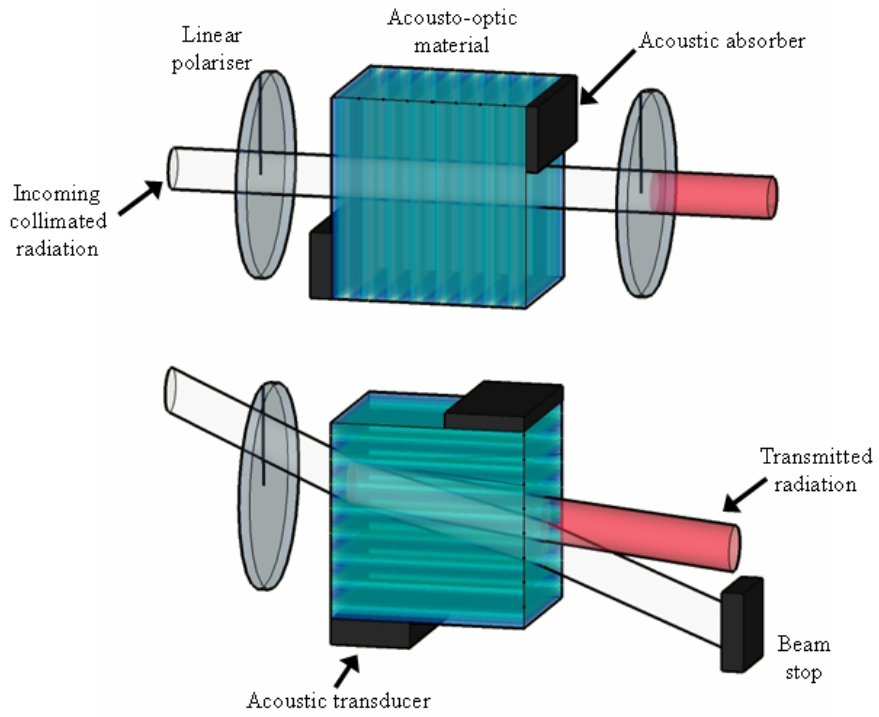

Figure 1.10: Acousto-optic tunable filter in (a) collinear configuration and (b) non-collinear configuration.

those collecting the spectral samples from a single image element in an single frame (in a similar manner to dispersive spectrometers) and those that require multiple exposures to do so (as with filter-based systems). These are usually termed as spatial (SIFTS) and temporal (TIFTS) Fourier transform spectrometers respectively. The temporal IFTS can be understood in terms of a Michelson interferometer [47]. If one of the mirrors of the Michelson is translated with constant speed, the path difference between wavefronts changes. The wavefronts interfere at the detector such that the recorded power is the cosine transform of the incident spectrum, plus a constant term. The transmission for an arm displacement of d^j is then given by

$$\tau^j(k) = \frac{1 + \cos(2\pi k d^j)}{2}. \tag{1.6}$$

Figure 1.11 shows the first 16 samples for a series of mirror displacements with magnitude equal to the smallest wavelength being sampled. Figure 1.12 shows the Michelson design adapted for imaging [48]. As with the other interferometric techniques, the system *étendue* is limited by the sensitivity to angle of incidence. The maximum transmission of temporal IFTS is 50%.

Spatial IFTS possess the same transmission functions as TIFTS, but the interferogram for a spatial element is created instantaneously at a plane in space instead of sequentially [49]. The interferogram is created by causing a collimated beam to split in

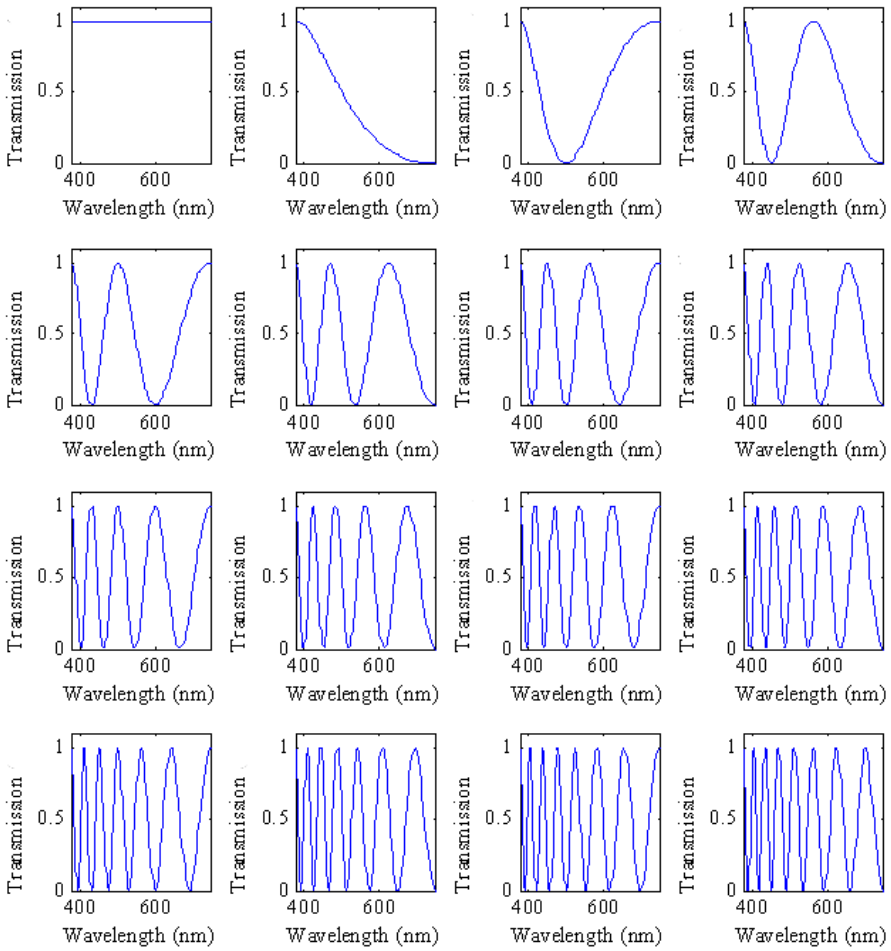

Figure 1.11: The first sixteen bands for a Fourier transform spectrometer.

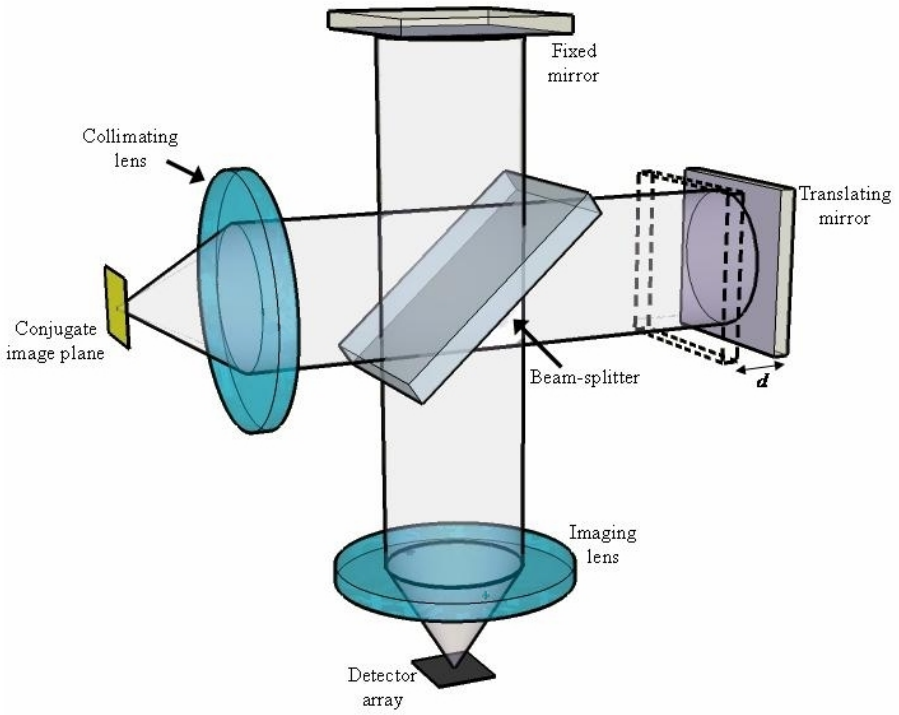

Figure 1.12: Schematic of a Michelson-based temporal IFTS.

two and recombine with the wavefronts of each beam sheared relative to one another. Figure 1.13 shows a Sagnac interferometer adapted for use as a SIFTS. Like dispersive systems, a two-dimensional detector array can be used to collect the spectral information for a row of spatial elements in a single integration. SIFTS are therefore another type of instrument suitable for use in the pushbroom arrangement; this together with the reduced sensitivity to vibration (due to the common path of both beams) mean that SIFTS are commonly employed in the field. Spatial IFTS have a maximum transmission of 50%.

Several types of FTS have been developed to further reduce the effects of vibration and shock that make field implementation of IFTS problematic. These include a Sagnac interferometer employing a pentaprism [50] to replace the beam splitter and mirrors, and a number of temporal [10] and spatial [51] IFTS instruments using Wollaston prisms to split the wavefronts and introduce a common path difference or shear at the detector.

Active instruments

Active spectral imaging refers to spectral imaging in which the illumination is supplied by the observer. This could mean supplying a broadband source and then filtering the return signals using a conventional spectrometer [52]. Alternatively this could refer to using a narrow-band source created using filters [53] or using by using narrow-band

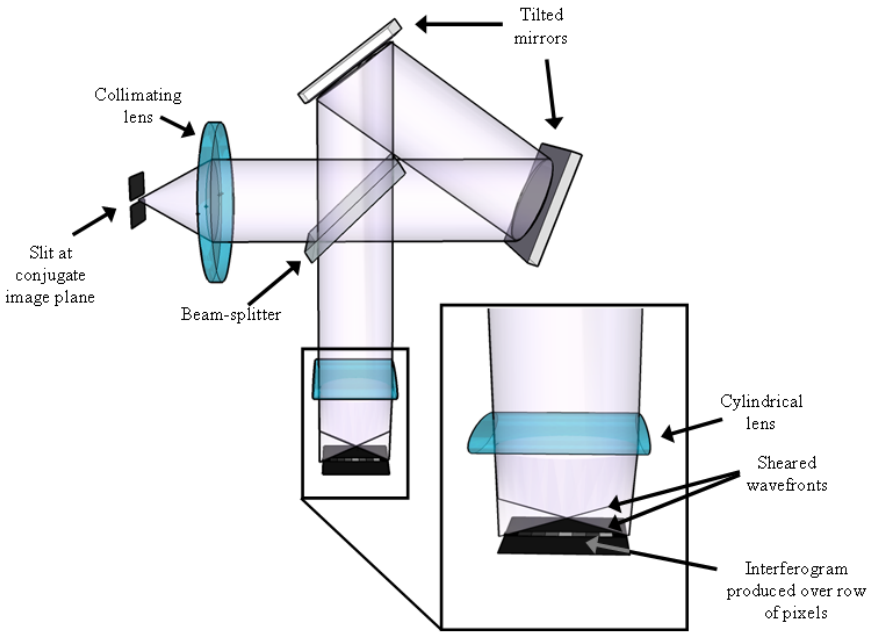

Figure 1.13: Schematic of a Sagnac-based spatial IFTS.

Instrument type	Number of bands	Random access	Suitable for bands	Typical average efficiency
Prism	100s	No	Vis to Far-IR	>90%
Grating	100s	No	Vis to Far-IR	<80%
Filter wheel	10s	Yes	Vis to Far-IR	>90%
Liquid-crystal filter	100s	Yes	Vis to Near-IR	<40%
Acousto-optic filter	100s	Yes	Vis to Far-IR	<80%
Temporal FTIS	100s	No	Vis to Far-IR	<50%
Spatial FTIS	100s	No	Vis to Far-IR	<50%

Table 1.2: Properties of sequential spectral imaging instruments.

lasers [54]. A grating-based spectral-imaging system using a broadband laser source, also capable of measuring range, has been demonstrated in a dark laboratory setting over a distance of 10m [55]. It is likely that the use of active spectral imaging will continue to expand given the rise of efficient low-cost illumination sources such as LED; the high power required for efficient spectral imaging of an extended area from distance is likely to remain an obstacle for remote-sensing applications.

1.2.6 Acquiring the data cube instantaneously

The technologies described thus far all rely on capturing many sequential detector frames in order to build up the spectral data for an extended two-dimensional scene. The need to scan in this manner is an issue for highly dynamic scenes has motivated research into capturing the spectral information for an extended scene in a single frame. Snapshot techniques have existed for some time in astronomy, where the sparsity of the scene can be exploited so that a slitless dispersive spectrometer [56] can be used to resolve multiple spatially separated objects in a single frame. This idea is extended by coupling the output from an intermediate image plane to the input slit of a dispersive spectrometer using fibre optics [57], allowing for the spectral imaging of crowded areas of the sky to be imaged. This method is quite general and fiber-reformatting has also been employed outside of astronomy to spectrally image a contiguous two-dimensional area [58, 59]. The principle is illustrated in Figure 1.14.

Another approach to spectral imaging using image reformatting is used in the spectral slicing imaging spectrometer [60]. This device uses a bulk reformatter known as an image slicer, consisting of an arrangement of thin strip mirrors tilted relative to one another. The slicer is used at an intermediate image plane, and in combination with an array of lenses, acts to split and project a lower-resolution image into spatially separated 1D strips on a high-resolution detector array in such a way that there are large voids at the detector between these image strips. By inserting a dispersive element into the collimated beam between the image slicer and the lens array, the spectra from each image strip are dispersed across the void area on the detector. With careful choice of the dispersive power and optics, the entire image can be spectrally imaged without overlap of the dispersed slices. The major disadvantage of these reformatting instruments at present is the cost of engineering the image slicer to achieve high spatial resolution. The lens array is also likely to require custom manufacture.

The reformatting techniques are all direct techniques; there also exist two noteworthy transform snapshot techniques, in which each detector pixel generally records spectral and spatial information from a number of spatial elements, using computer processing to recover a spectral data cube from the detected image. These snapshot transform techniques are the computed tomographic imaging spectrometer (CTIS) and the coded aperture snapshot spectral imager (CASSI).

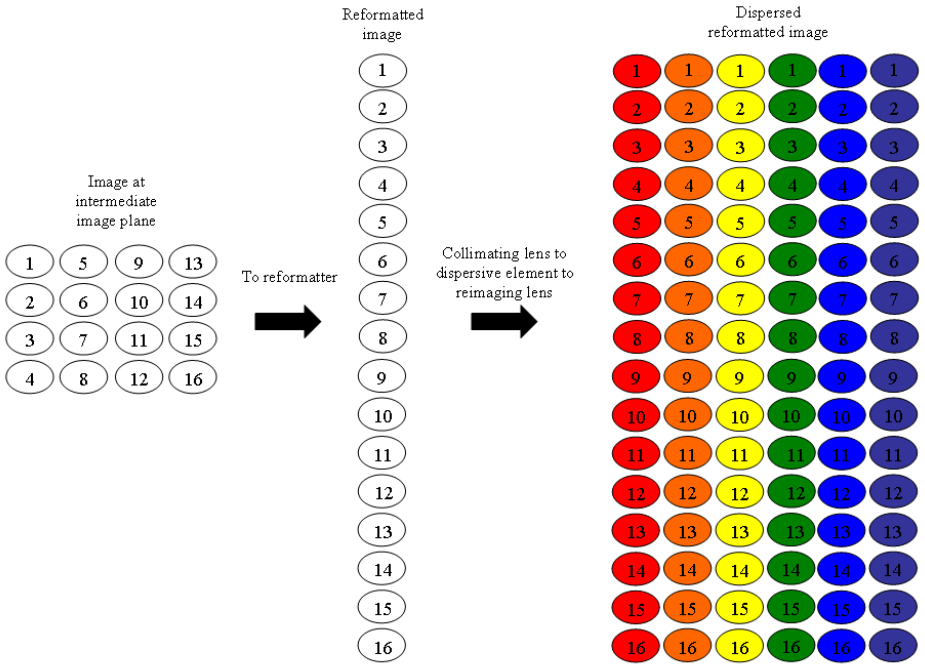

Figure 1.14: Reformatter process of operation. A two-dimensional scene is reformatted onto the entrance slit of a dispersive instrument, the dispersed output image can be rearranged to form the data cube using software.

CTIS [61] operates in an identical manner to a conventional grating-based imaging spectrometer, as described in Section 1.2.5, with the slit replaced by a square field stop and the standard grating replaced by a two-dimensional dispersive grating (Figure 1.15). This two-dimensional grating produces multiple diffraction orders in a two-dimensional array at the detector. A CTIS image (simulated as part of the work of this thesis) in which the object is an image of the letter F with a gray spectrum, is shown in Figure 1.16. The central (zero) order is a broadband image of the scene and the orders surrounding it are dispersed by greater amounts with distance from the centre of the detector.

The spectral data cube is retrieved from the detected CTIS image by means of algorithms common to emission tomography; this is why the instrument is known as the computed tomography imaging spectrometer. These algorithms are explained in more detail in Section 4.5, to summarise it is assumed that the three-dimensional object space $f_{x,y,\lambda}$ (two spatial and one spectral dimension) maps to the array of pixels at the detector $g_{u,v}$ according to the linear relation

$$g_{u,v} = \sum_{x} \sum_{y} \sum_{\lambda} H_{u,v,x,y,\lambda} f_{x,y,\lambda},$$

with $H_{u,v,x,y,\lambda}$ describing the instrument response. By arranging the detected image and

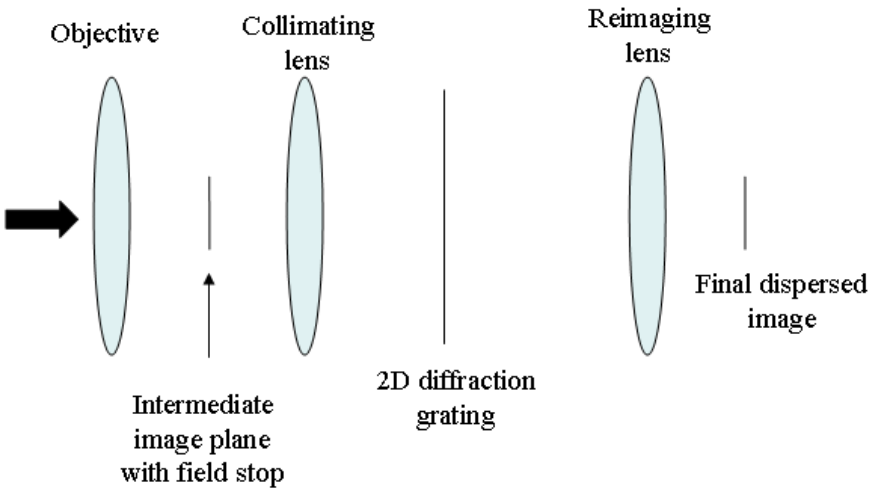

Figure 1.15: CTIS layout.

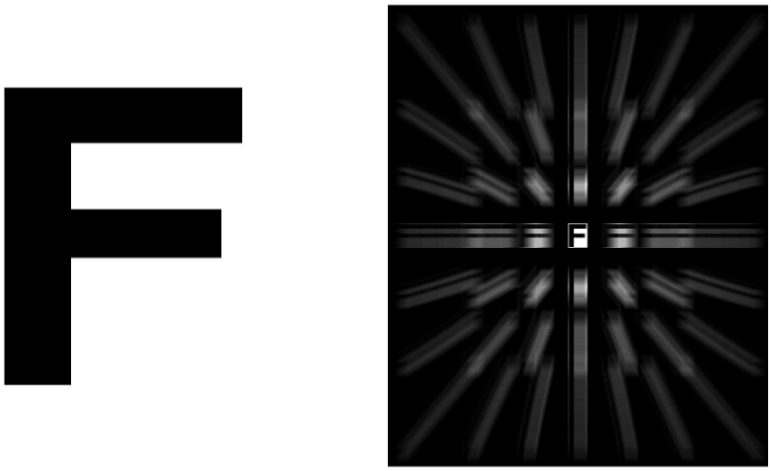

Figure 1.16: (a) Input image and (b) image at the detector for a CTIS instrument.

discretised object space as vectors $\mathbf{g}$ and $\mathbf{f}$, and the $H_{u,v,x,y,\lambda}$ as a matrix $\mathbf{H}$ this can be written as

$$\mathbf{g} = \mathbf{H}\mathbf{f}.$$

Depending on how well-posed this linear problem is, there are a number of techniques for inverting this equation to recover $\mathbf{f}$ from $\mathbf{g}$ [62, 63]

The disadvantages of CTIS include computational complexity in recovering the data cube, relatively low pixel resolution and the need for involved calibration. The advantages are the snapshot capability, the small size and the low cost. CTIS also rejects very little light, which may lead to a high signal-to-noise ratio in the final spectral data cube; however a thorough error analysis is needed to determine how any measurement errors in the detected image transform to the recovered spectral data cube.

Coded aperture snapshot spectral imaging devices (CASSI) currently come in two different designs: the dual-dispersive architecture (CASSIDD) [64] and the single-dispersive archetecture (CASSISD) [65]. The layout of CASSIDD is shown in Figure 1.17; a slitless prism-based dispersive spectrometer is employed, similar to those used in astronomy for sparse scenes. In the general case, that is when the scene is not sparse, each pixel in the image formed after the prism contains mixtures of the spatial and spectral components from nearby object points in the scene, making determination of the spectral data cube from this image a severely under constrained problem. By placing a coded aperture at the plane of the dispersed image, small portions of this image can be blocked. The coded aperture in this case is a specific type of patterned binary mask made from a replicated 3×15 unit cell; an example, reproduced here using information from [64], is shown in Figure 1.18. When the dispersion in this coded dispersed image is reversed using an identical prism and then re-imaged, the final coded image is better constrained, albeit for a smaller number of object points. Figure 1.19 shows the simulated sequence of operation of CASSIDD. The broadband input image in this example is a MacBeth ColorChecker® [2]. Also shown are the dispersed image, the coded dispersed image, and the final coded image. Figure 1.20 shows a closer view of a section of the input image and the same section of the detected image.

The single-dispersive architecture operates along similar principles and is identical to the conventional dispersive imaging spectrometer except a coded aperture replaces the slit so that the input image is coded. For CASSISD the coded aperture is a random binary pattern; an example pattern is shown in Figure 1.18(b). A simulated image is shown in 1.22(b). Whereas in the CASSIDD design, incomplete spectra are measured for all spatial elements, for CASSISD complete spectra are measured for a reduced set of spatial elements.

The techniques for reconstructing the spectral data cube for CASSI are similar to

[2]The spectral reflectances for the MacBeth ColorChecker® have been obtained from Munsell Color Science Laboratory.

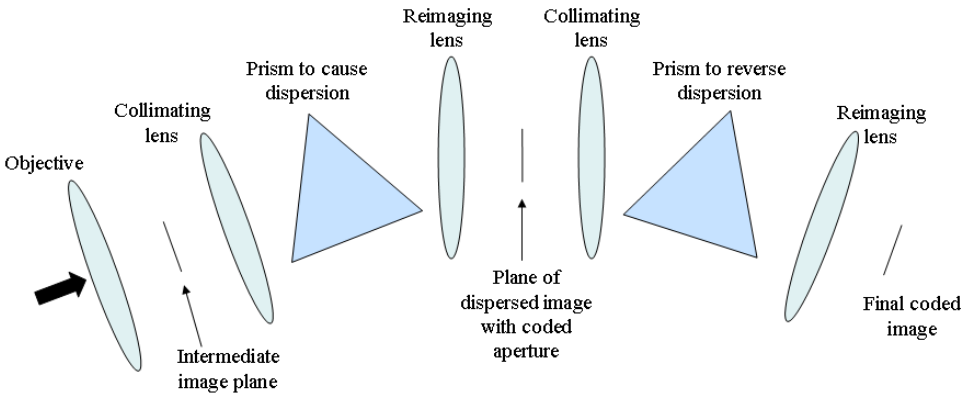

Figure 1.17: Layout of a dual-dispersive CASSI

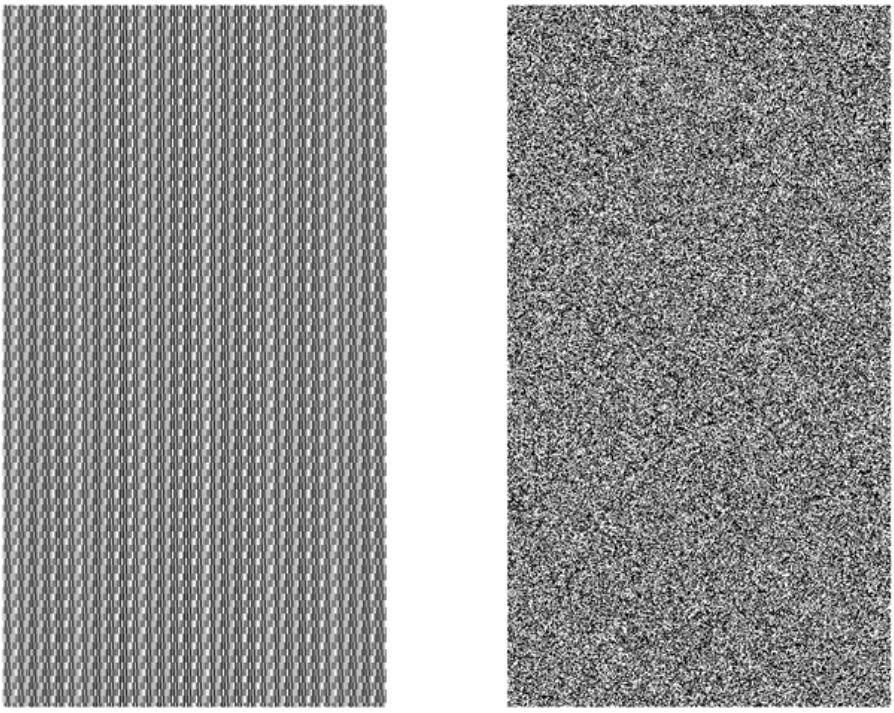

Figure 1.18: Coded apertures for CASSI systems. Shown are example masks for (a) CASSIDD and (b) CASSISD

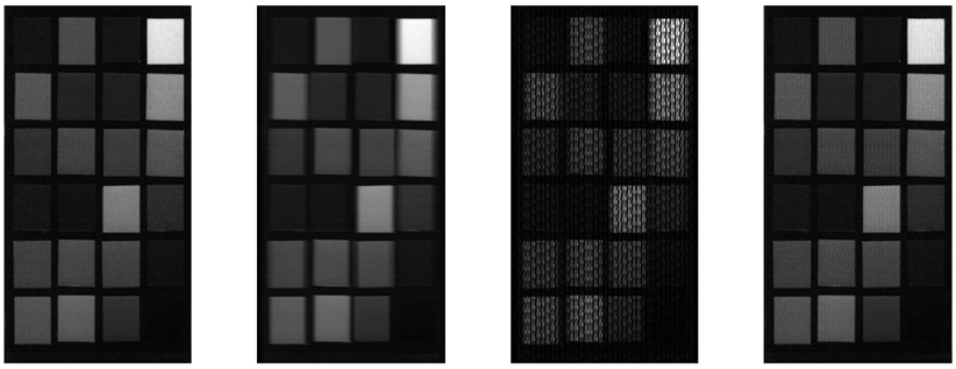

Figure 1.19: Sequence of operation for CASSIDD: (a) Broadband input image, (b) dispersed image, (c) coded version of (b), and (d) coded image (c) with dispersion reversed.

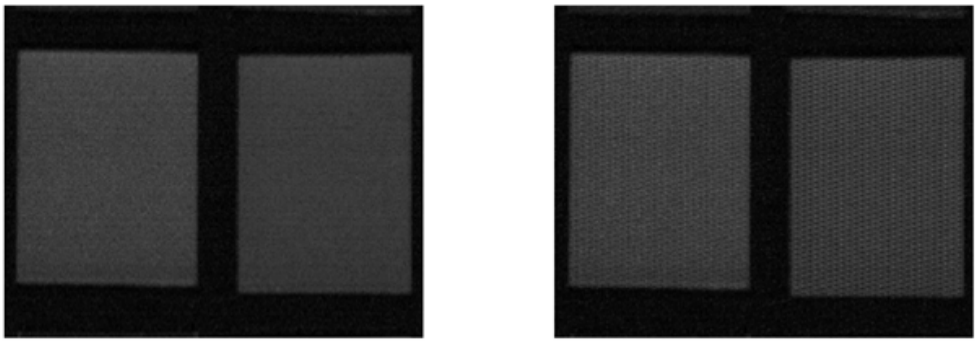

Figure 1.20: (a) Close up view of broadband input image in Figure 1.19(a). (b) Close up view of coded image in Figure 1.19(d)

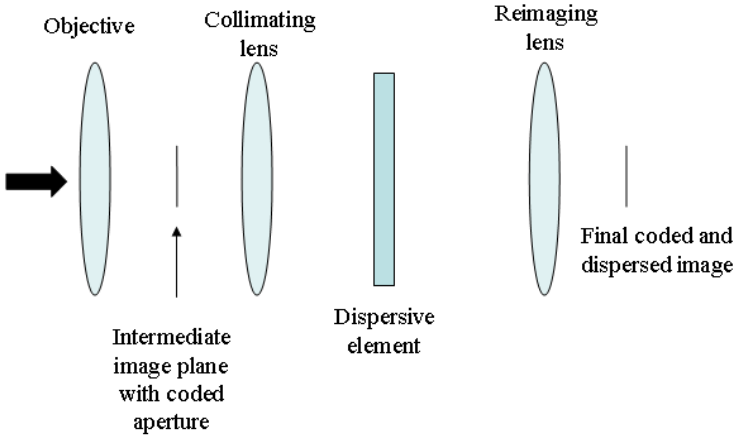

Figure 1.21: Single-dispersive CASSI layout.

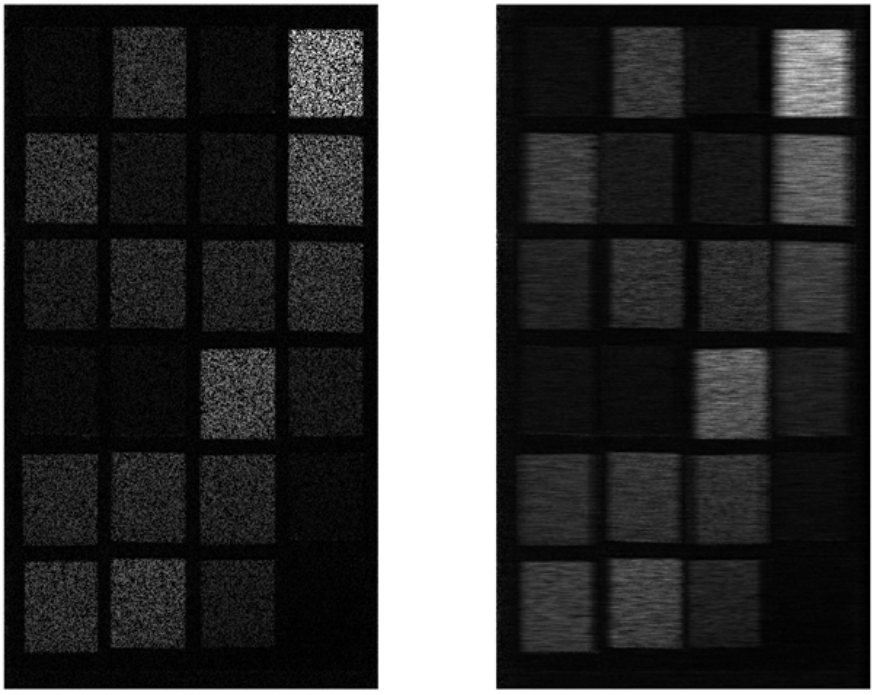

Figure 1.22: (a) Input image coded with binary coded-aperture. (b) Dispersed coded image at the detector.

those used for CTIS data and a number of algorithms have been suggested [64, 65].

CASSI is relatively low cost; apart from the coded apertures , the required components are available as commercial off-the-shelf (COTS) products. Like CTIS, CASSI requires computationally intensive inversion to recover the spectral data cube from the measurements. The compressive nature of CASSI may also result in large inaccuracies in the retrieved data cube.

The subject of this thesis is the image-replicating imaging spectrometer. This is a snapshot system which creates an array of spectrally distinct images simultaneously at the detector. IRIS uses an arrangement of retarders and Wollaston prisms to achieve this and has no moving parts. Unlike CTIS and CASSI the IRIS images are captured without the need for inversion, making it easier to utilise the information in real time and removing uncertainties associated with the inversion process. The main issues with IRIS are the expense of Wollaston prisms and also chromatic aberrations associated with the optical system. However, solutions to these problems are discussed in later Chapters.

1.3 Applications of spectral imaging

Spectral imaging is used in a number of fields to identify, discriminate between, and calculate abundances of materials in a scene. An overview of a representative sample of

applications is given in this section. Some of the most common techniques for processing spectral data are also described.

1.3.1 Spectral-imaging missions

One of the main application areas for spectral imaging has been in remote sensing. An example of such a system is the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). This is a satellite-based dispersive system, which uses multiple gratings and detector arrays to cover the range 400-2500 nm with 224 channels [66] to image the Earth's surface. AVIRIS has been in use since 1987 on a variety of platforms, collecting terabytes of data from North America and Europe. The data from AVIRIS is used in many remote sensing based studies, including:

- Differentiating between different forms of snow and ice [67] [68].
- Studies of the geological composition of the Earth's surface [69].
- Determination of the concentration of chlorophyll and other biological species in water [70] [71].
- Monitoring atmospheric carbon dioxide [72].

AVIRIS is also useful for studying vegetation as it samples the 680-750 nm region of the spectrum where the reflectance of chlorophyll increases substantially. Spectral measurements in this region can be used to determine chlorophyll concentration which can help monitor plant stress and other properties of the biomass. Examples of this include studies of forest composition [73], and monitoring agricultural crops [74] and other vegetation [75].

Other important examples of remote sensing missions include the Landsat program and Envisat. The first five Landsat satellites carried the 4 band multi-spectral scanner [76] a dispersive system operating in the visible and infrared. The two most recent Landsat satellites carry the thematic mapper, a 7-band filter-based system using multiple detectors to acquire data in the visible as well as the near-, mid- and far-infrared. Landsat data has been used for agriculture, forestry, geology, mapping, and oceanography. Envisat carries the medium resolution imaging spectrometer (MERIS), a dispersive system capturing a large number of bands in the visible and near-infrared. The primary use of MERIS is for oceanography [77].

Another major application of spectral imaging is in medical imaging and the life sciences. Spectral imaging can be used in microscopy to discriminate between fluorescent dyes as an alternative or compliment to fluorescence lifetime imaging. A system using a Sagnac-based Fourier-transform system [78] is able to successfully identify all human chromosomes simultaneously. Another example application in microscopy is the use

of an acousto-optic tunable filter in Forster resonance energy-transfer measurements [79]. Medical examples include an acousto-optic tunable-filter-based system for in-vivo studies of skin and brain [80] and a system for detecting cancer in vivo using a liquid-crystal tunable filter [81]. Narrow-band endoscopy has been studied for use in diagnosing gastro-intestinal disease and multi-spectral retinal imaging for oximetry is an active area of research [82] [83].

Spectral imaging also has applications in surveillance and defense related fields. Spectral information from surveillance and reconnaissance footage can be used to identify camouflage and materials such as explosives [84]. Spectral imaging of rocket plumes is also of interest and has been attempted using a fiber-reformated dispersive instrument [85]. The increased contrast offered by spectral imaging compared to monochromatic or colour cameras can also be used to detect sub-pixel targets [86] and to improve automated tracking and detection.

Spectral imaging also finds use in industry. Accurate colour reproduction in print [87] is one application. Multi-spectral imaging of artwork using a filter wheel has been successful in producing minimally metameric prints of the original [88]. Other industrial examples include quality control in manufacturing pharmaceuticals [89] using a near-infrared LCTF, and in food production [90] using a grating-based system in visible light.

1.3.2 Spectral-image processing

Point-by-point inspection of spectral data or separate examination of each monochrome image is laborious and prone to human error. Instead the data are usually processed using a computer prior to human interpretation.

Registration, false colour images and unmixing

Standard spectral imaging algorithms require that samples are arranged into sets according to spatial coordinate, a pre-processing step known as registration. This is equivalent to constructing a set of images for which the mapping of the spatial coordinate system to the image coordinate system is the same in all images. Each set of pixels sharing the same image-coordinate in the registered set can be identified with a measurement vector $\mathbf{y}$ of the J spectral samples y^j for that spatial element.

Registration is achieved by searching the transformation space to maximise some objective such as correlation or mutual information for the entire image. Alternatively, for a single pair of images, the transformation can be found by matching pixels corresponding to the same spatial element, these pixels are known as control points. The control-point method is often preferred when the transformation space is large enough to make search methods prohibitive.

With the images registered, two or three out of the set can be overlaid to create a false-colour image for visualisation. To exploit more of the available information, the data can be projected onto some smaller number of dimensions containing a large amount of the relevant information. Common projections include the principal component analysis (PCA) and the independent component analysis (ICA). ICA is conceptually very close to spectral-unmixing algorithms, these are source separation techniques in which it is assumed that each spatial element is composed of a finite number of distinct materials, the relative proportions of which can be found for each spatial element by decomposing the J samples in some manner. The spectral signatures of the pure materials within the scene are known as end-members and can be found from the statistics of the data set or chosen from an existing library of spectral signatures.

Contrast and clustering

One of the major benefits of spectral imaging compared to monochromatic and colour imaging is the potential increase in contrast offered by multiple independent samples of the radiation from each spatial element. To make use of this contrast, some metric of dissimilarity is needed. Commonly used metrics are the norm of the difference between two measurement vectors $|\mathbf{y}_1 - \mathbf{y}_2|$, the angle between the vectors $\cos^{-1}\left(\mathbf{y}_1\mathbf{y}_2^T/|\mathbf{y}_1|\,|\mathbf{y}_2|\right)$, and the Mahalanobis distance

$$\sqrt{(\mathbf{y}_1 - \mathbf{y}_2)\,\Sigma^{-1}\,(\mathbf{y}_1 - \mathbf{y}_2)^T}, \tag{1.7}$$

where Σ is the covariance of the distribution from which the measurement vectors are drawn.

A similarity metric can be used with a decision rule to group the spatial elements into multiple sets. Depending on the amount of *a priori* information required this process is known as supervised or unsupervised classification; unsupervised classification is also known as clustering.

Inversion of data to spectrum

In many instruments, particularly transform instruments, the data for each point needs to be inverted if the spectrum of radiation from the spatial element is to be obtained. As the spectrum is a continuous function, the inverse transforms used to recover the spectrum from the finite number of samples are not unique and can be found from the transmission functions of the instrument. The Backus-Gilbert (maximum resolution) inverse is used in this thesis.

Spectral and spatial processing

There are only a small number of reported algorithms that use both the spatial and spectral information. Anomaly detection algorithms, such as the RX-algorithm, make use of windowing to compare the spectral samples from a target sized area of the image with the statistics of a larger area. Most similarity metrics can also be extended to include the two spatial dimensions as well as the spectral dimension so that the signatures are volumes of measurement space. Morphological and topological analyses can also be extended to these higher dimensional spaces.

1.4 Thesis outline

This chapter has described the context and motivation for snapshot spectral imaging techniques, in particular the image-replicating imaging spectrometer. Chapter 2 will introduce the IRIS concept and design principles. Detailed modeling of these systems is described in Chapter 3 along with an appraisal of IRIS systems, comparing experimental results with predicted performance. Optical aberrations are a serious issue for spectral imaging technologies; Chapter 4 discusses aberrations encountered in IRIS systems and presents methods for dealing with them. Chapter 5 compares IRIS to existing spectral-imaging techniques, such a comparison may useful when determining the suitability of IRIS for a given spectral-imaging application. Results from several applications of IRIS systems are presented in Chapter 6. Conclusions and recommendations for future work are given in Chapter 7

Chapter 2

IRIS Concept and Design

Summary: This chapter introduces the IRIS concept and details the optical design of IRIS systems, providing refinements to the initial concept. The optical design principles are then demonstrated in the first-order design of an 8-channel IRIS system. The specific challenges in implementing IRIS systems in the thermal infrared are illustrated for a prototype IRIS system operating from 3-5 μm. The IRIS system optical parameters resulting in the optimal throughput for each field of view are calculated, and a survey of suitable birefringent materials is carried out. The chapter concludes with a discussion of potential image aberrations in IRIS systems and a description of the issues surrounding the registration of the IRIS images.

2.1 Introduction

The Image Replicating Imaging Spectrometer (IRIS) is the primary focus of this thesis. IRIS is the invention of Andrew Harvey[1] and David Fletcher-Holmes[2], the concept is examined and refined in this thesis.

As well as theoretical work, four different IRIS instruments are used in this thesis to illustrate different aspects of the design and a number of issues that arise:

- A mid-wave infrared system manufactured by *Selex Galileo* (formally *BAE systems*, based upon a design and concept outlined in a feasibility report produced by David Fletcher-Holmes and Andrew Harvey [91]. This system is designed for general surveillance related tasks in the wavelength range 3-5 μm. Hereby referred to as the MWIR IRIS system.

- An 8-channel, narrow-band system designed by the author and manufactured by Gonzalo Muyo[3]. This system is used for retinal imaging over the wavelength range 560-600 nm. Hereby referred to as the retinal-imaging IRIS system.

[1]Professor, Heriot Watt University
[2]Research associate, Heriot Watt University (2000-2005). Presently a Senior Scientist at CRi, Inc.
[3]Research associate, Heriot Watt University

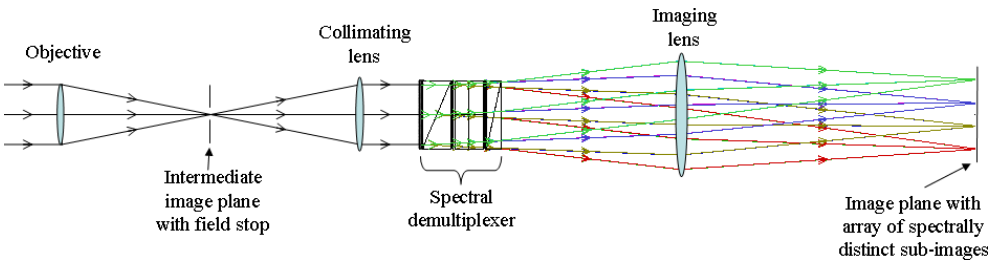

Figure 2.1: Layout of IRIS systems.

- A 16-channel, broadband system designed and manufactured by the author. This system was designed with a wide field of view for surveillance related tasks in the wavelength range 420-830 nm. Hereby referred to as the 16-channel broadband IRIS system.

- An 8-channel, broadband system designed and manufactured by David Fletcher-Holmes. This system is a general use system for the wavelength range 500-820 nm. Hereby referred to as the 8-channel broadband IRIS system.

This chapter introduces the IRIS concept and discusses the implementation of real systems. After the IRIS concept is introduced, the transmission functions of IRIS systems are discussed in detail. The optical design of IRIS systems is then explained, including the separate issues faced when designing systems for use in the infrared; choice of suitable birefringent materials is also discussed. The chapter concludes by discussing potential aberrations and the co-registration of the sub-images.

Two pieces of ray tracing software are also used in the course of this thesis. Zemax from *ZEMAX Development Corporation* is used for geometric ray tracing of IRIS systems using compound lenses, and an interferometric ray trace written by the author is used to calculate the transmission functions of IRIS, and for easier automation of certain ray tracing tasks. Where necessary, it shall be stated in the text which is program is used.

As the IRIS concept is the work of Andrew Harvey and David Fletcher-Holmes some of the design principles were determined prior to the work of this thesis. Where it is applicable, credit is given in the text, the remainder of the work in this thesis is the author's own unless explicitly stated otherwise.

2.2 A generalisation of the Lyot filter

Figure 2.1 shows the layout of an IRIS system [91]. IRIS operates to simultaneously project an array of spectrally distinct images of the scene onto two-dimensional detector array. An example of the image formed at the detector of an IRIS instrument is shown in Figure 2.2. This operation is achieved using a spectral demultiplexer to split incoming

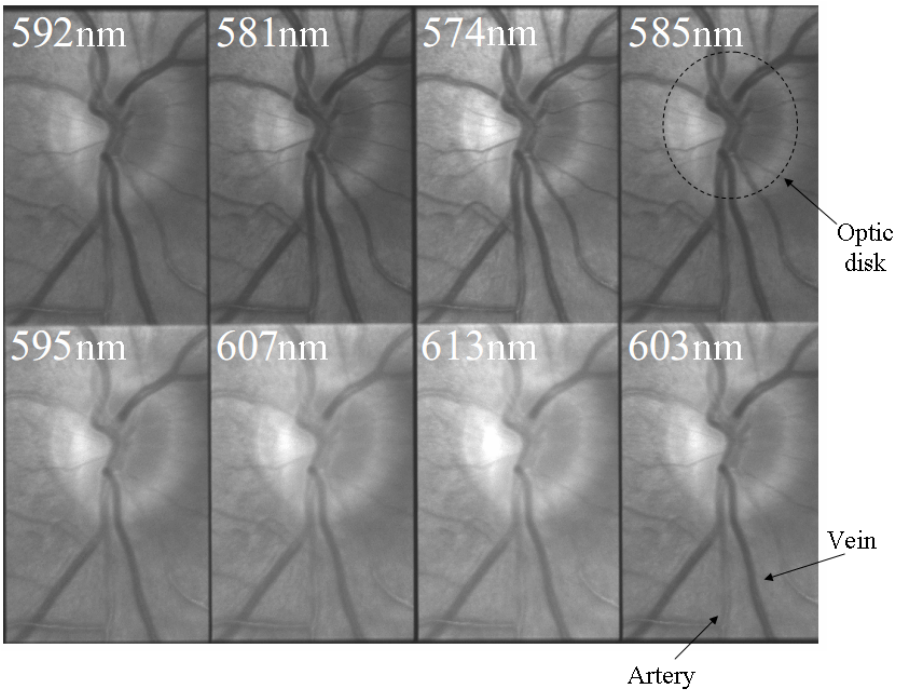

Figure 2.2: Example IRIS frame from the retinal imaging IRIS system, consisting of eight spectrally-distinct replicated images of the optic disk and major retinal blood vessels for a healthy human retina, peak wavelengths are shown in the corner of each sub-image. Due to the difference in spectral absorption of oxygenated and deoxygenated blood the arterial vessels appear lighter than the veins at longer wavelengths. See Chapter 6.3 for further details of the application of IRIS to retinal oximetry.

collimated radiation into multiple, mutually-divergent, spectrally-distinct, collimated beams. These beams can then be brought to focus at different areas on the detector using an imaging lens.

Figure 2.3 shows a schematic of a spectral demultiplexer, which consists of a linear polariser followed by M retarders and M Wollaston prisms, with the transmission-axes of the prisms at 45° to the optic-axes of the retarders. The demultiplexer can be considered an extension of the Lyot filter (Section 1.2.5); after the first retarder the linear polarisers in the Lyot filter are replaced with Wollaston prisms.

Each Wollaston prism analyses the radiation in a collimated beam into two orthogonal, linearly-polarised beams. The polarisation of the input beam is, in general, different for each wavelength, due to the retarder preceding the prism. As a result of this wavelength-dependent polarisation state, the spectral transmissions of the two output beams are different from one another. The angle between the two beams produced by the Wollaston prism is known as the splitting angle of the Wollaston prism and is a function of the birefringence of the prism material, $B(\lambda)$, and the wedge angle of the prism, θ.

Figure 2.4 shows a prism being used to produce two replicated images. Using M prisms in series, the image will be replicated $J = 2^M$ times. To prevent the replicated images from overlapping, a field stop must be used at the conjugate image plane.

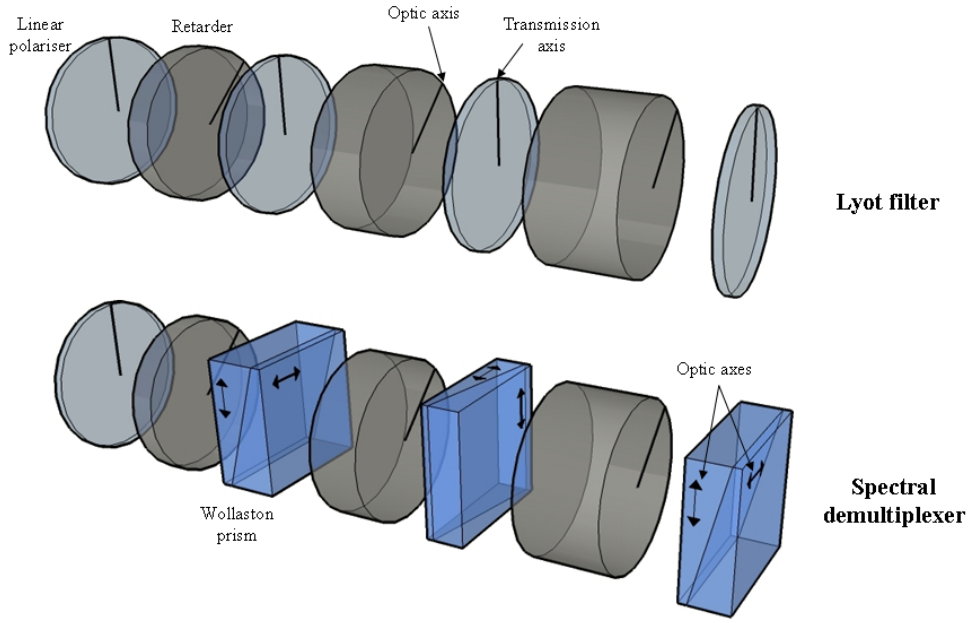

Figure 2.3: The spectral demultiplexer compared to a Lyot filter.

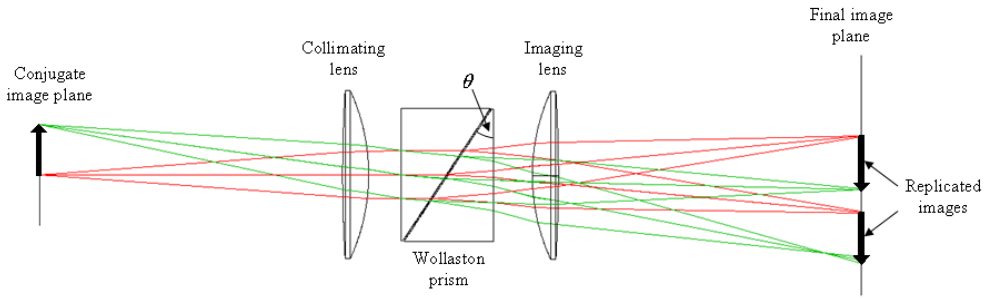

Figure 2.4: Replicating an image using a Wollaston prism

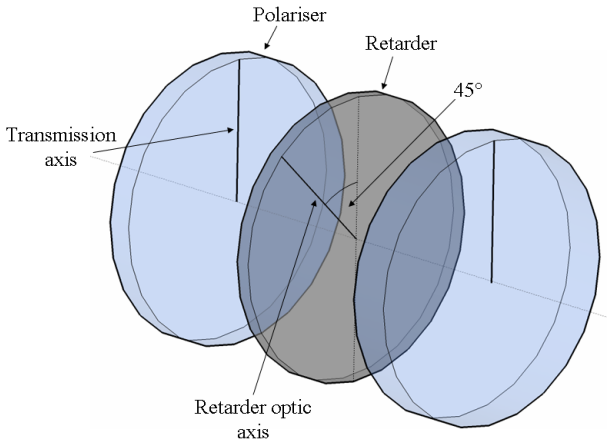

Figure 2.5: Arrangement of a retarder between two coaligned polarisers with the retarder optic axis at 45° to the transmission axes of the polarisers.

2.3 Transmission functions

The transmission functions of a demultiplexer can be understood by considering the arrangement in Figure 2.5. The phase difference introduced between orthogonal polarisation components by a single retarder of thickness d, can be written as

$$\delta(\lambda) = \frac{2\pi B(\lambda)\, d}{\lambda} = 2\pi k B(k)\, d, \tag{2.1}$$

where $B(\lambda)$ is the birefringence at wavelength λ (wavenumber $k = 1/\lambda$).

The action of the arrangement in Figure 2.5 on normally-incident radiation described by an electric field vector $\mathbf{E}$, is therefore given by the Jones matrix $\mathbf{W}$,

$$\mathbf{W}(\lambda)\mathbf{E}(\lambda) = \mathbf{PR}(\lambda)\mathbf{PE}(\lambda), \tag{2.2}$$

where $\mathbf{R}$ is the Jones matrix for the retarder and $\mathbf{P}$ is the Jones matrix for the polarisers,

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{2.3}$$

$$\mathbf{R}(\lambda) = \begin{bmatrix} \frac{1}{2}e^{i\delta(\lambda)} + \frac{1}{2} & \frac{1}{2}e^{i\delta(\lambda)} - \frac{1}{2} \\ \frac{1}{2}e^{i\delta(\lambda)} - \frac{1}{2} & \frac{1}{2}e^{i\delta(\lambda)} + \frac{1}{2} \end{bmatrix}. \tag{2.4}$$

Assuming absorption is negligible, the transmission of this system in vertically polarised light, represented by the vector

$$E = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{2.5}$$

is given by

$$\tau(k) = \mathbf{E}^T\mathbf{W}^T\mathbf{W}\mathbf{E} = \cos^2(kl), \tag{2.6}$$

where $l = \pi Bd$. The transmission of a Lyot filter with M stages is therefore given by

$$\tau(k) = \prod_{m=1}^{M} \cos^2(l_m k). \tag{2.7}$$

At each Wollaston prism in the demultiplexer both the cosine and sine squared components now pass as mutually divergent, collimated beams. M retarder-prism pairs give rise to $J = 2^M$ beams with transmissions given in the first approximation by the 2^M distinct products of the successive cosine- and sine-squared terms. For example, with $M = 3$, the first-order approximations for the 8 transmission functions are as follows [91]:

$$\begin{aligned}
\tau^1 &= \cos^2(l_1k)\cos^2(l_2k)\cos^2(l_3k)\\
\tau^2 &= \sin^2(l_1k)\sin^2(l_2k)\sin^2(l_3k)\\
\tau^3 &= \cos^2(l_1k)\sin^2(l_2k)\cos^2(l_3k)\\
\tau^4 &= \sin^2(l_1k)\cos^2(l_2k)\sin^2(l_3k)\\
\tau^5 &= \cos^2(l_1k)\cos^2(l_2k)\sin^2(l_3k)\\
\tau^6 &= \sin^2(l_1k)\sin^2(l_2k)\cos^2(l_3k)\\
\tau^7 &= \cos^2(l_1k)\sin^2(l_2k)\sin^2(l_3k)\\
\tau^8 &= \sin^2(l_1k)\cos^2(l_2k)\cos^2(l_3k),
\end{aligned} \tag{2.8}$$

where the subscripts refer to the retarder index. As with the Lyot filter, the functions are periodic and the passband is limited using bandpass filters at some appropriate point in the optical path.

As with other interferometric techniques the transmission functions for IRIS (and for the bandpass filters) are slow functions of the angle of incidence. In the case of IRIS this is because the optical-path-length difference introduced by a retarder depends on the angle of incidence, which changes both the distance traveled through the material and the extraordinary refractive index. Optimisation and calibration of the IRIS bands is therefore simpler when the radiation incident upon the demultiplexer and bandpass filters is collimated and contains a narrow range of angles.

Chapter 3 discusses the accurate determination of the transmission of IRIS systems across the entire field of view using ray tracing. For a given angle of incidence θ, the approximate transmission for a single retarder between two coaligned polarisers can be found by calculating the distances do_θ and de_θ, that the ordinary and extraordinary rays travel through the retarder using Snell's law, with the refractive index of the retarder material set to n_o and n_e, the ordinary and extraordinary refractive indices of the material respectively. This is an approximation as the extraordinary ray refractive

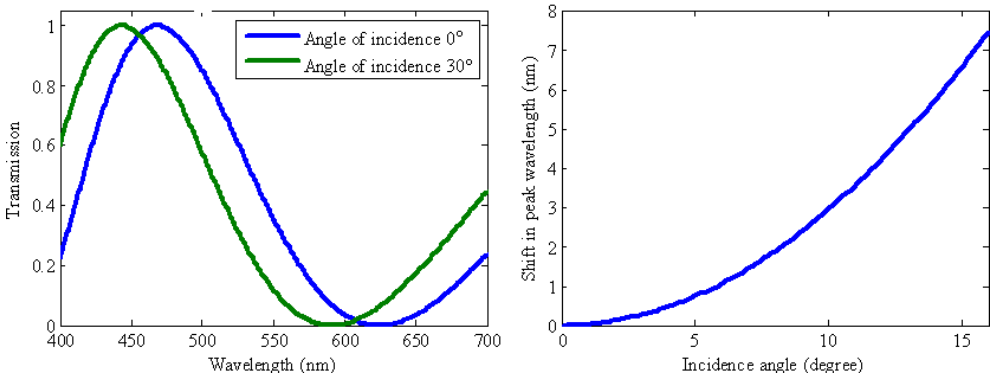

Figure 2.6: (a) Calculated change in transmission through a 102 μm-thick quartz retarder between coaligned polarisers for an angle of incidence of 30°, between 400 nm and 700 nm. (b) Shift in peak wavelength as a function of incidence angle for the same arrangement as in (a).

index actually changes with incidence angle. The transmission is then calculated using

$$\tau\left(k\right) = \cos^2\left(kB\pi do_\theta - kB\pi de_\theta\right). \tag{2.9}$$

Figure 2.6(a) demonstrates the change in transmission through a 102 μm-thick quartz retarder between coaligned polarisers for an angle of incidence of 30°, for the wavelength range 400-700 nm. Figure 2.6(b) shows the shift in peak wavelength as a function of incidence angle for the same arrangement. For example, in this case, if the tolerance on peak wavelength shift is < 1 nm, the maximum angle through the retarder should be $< 6°$.

2.3.1 Optimisation of bands

Figure 2.7 shows an example of the transmission functions for the 16-channel broadband IRIS system using retarder thicknesses that double consecutively, as in the Lyot filter, with the seed thickness (that of the thinest retarder) given by $d_{min} = 1/B\Delta k$. There is no reason to suppose that a set of bands chosen in this way will be useful for a specific task. Instead the thicknesses are chosen by maximising a suitable objective function.

Metrics for selecting retarders

The transmission functions of the demultiplexer depend only on the retarder thicknesses and on the wavelength-dependent ordinary and extraordinary refractive indices of the retarder material. The refractive indices can be found for each wavelength using empirically determined dispersion equations, which are listed, for a large number of materials, in [92]. Choice of retarder material will be discussed in more detail in the next section.

With the materials chosen it remains to pick the thicknesses of the retarders to optimise the objective function. The most suitable metric will depend on the intended

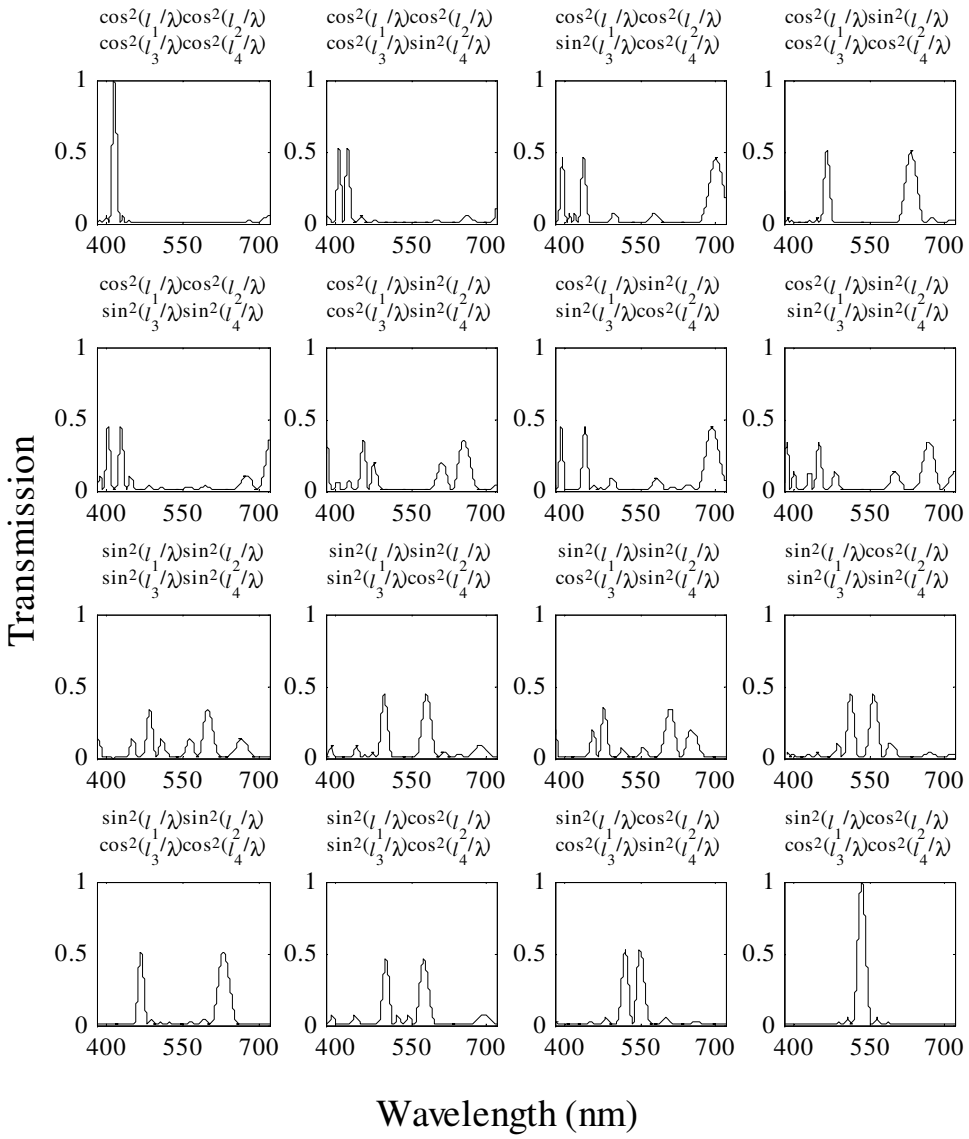

Figure 2.7: Example of a set of unoptimised bands a 16-channel IRIS system. The mathematical form of the corresponding transmission function is shown above each plot.

application and the amount of a priori information available.

As discussed in Section 1.3.2, the main types of spectral processing are discrimination and sub space projections such as linear unmixing.

If there is no prior knowledge of the state space, then the separation between the bands themselves can be maximised via the orthogonality or some other measure. Figure 2.8 shows the result of optimisation for an 8-channel system with quartz retarders, operating between 375-750 nm. The optmisation strategy is to minimise the average squared difference between the IRIS transmission functions and a set of target transmission functions. The target in this case is the set of idealised contiguous, orthogonal, direct bands shown in Figure 2.9. The optimisation was carried out using an exhaustive search of the three retarder thicknesses between 50 and 2000 μm at a resolution of 0.1 μm.

If the range of spectra to be recorded and the intended processing algorithm are known in advance then this information can be used in the optimisation. The set of transmission functions shown in Figure 2.10 have been optimised for spectral angle discrimination (Section 1.3.2) of the spectral reflectances of the eighteen coloured squares on the MacBeth ColorChecker® shown in Figure 2.11. The spectral angle between two spectral measurement vectors $\mathbf{u}$ and $\mathbf{v}$ (Section 1.2.5) is

$$\cos^{-1}\mathbf{u} \cdot \mathbf{v} \tag{2.10}$$

The goal for this optimisation was to maximise the average spectral angle between the set of eighteen measurement vectors corresponding to the coloured squares. The optimisation was carried out using an exhaustive search of the three retarder thicknesses between 50 and 2000 μm at a resolution of 0.1 μm. The reflectance spectra of the squares in the absence of noise is 38° compared to 42° for the set of direct bands shown in Figure 2.9.

There are also other factors that it may be desirable to account for in the objective function. As will be explained in Section 2.5 the birefringence of prism materials generally decreases with wavelength, causing each point in the scene to be dispersed over a number of pixels in the detector plane. As the extent of this smear increases with the spectral width of the band, the effects of the birefringent dispersion can be reduced by minimising the sidelobes in each band. The dynamic range of the detector is another factor to consider; a band with a low amount of expected energy may not be measured to acceptable accuracy in the same exposure as a band with a high amount of energy. To avoid this, equalisation of the expected power across the bands can also be included in the objective function.

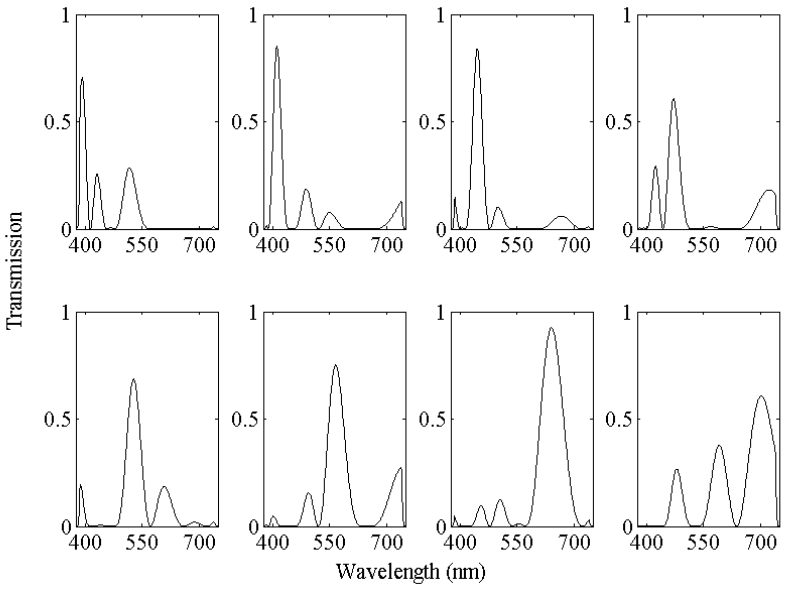

Figure 2.8: Transmission functions for an 8-channel IRIS system optimised for similarity (in the least squares sense) to the target transmissions shown in Figure 2.9.

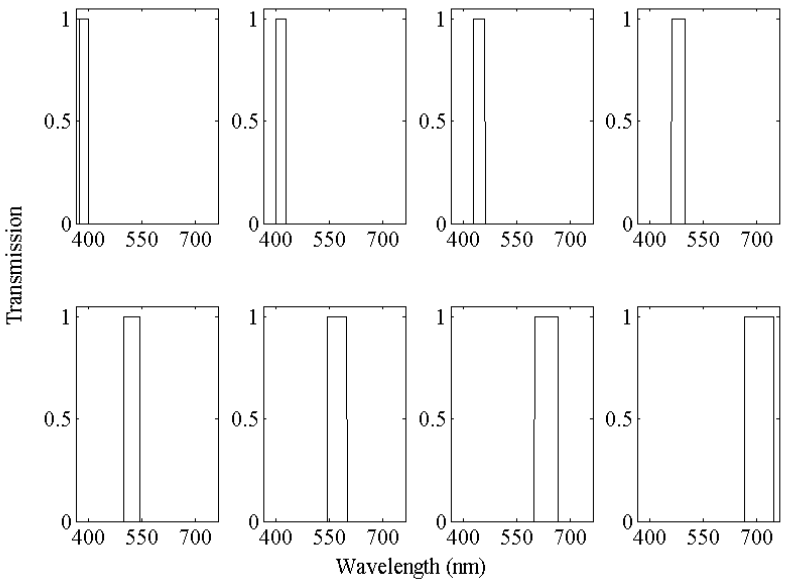

Figure 2.9: Target transmission functions for optmisation of bands shown in Figure 2.8.

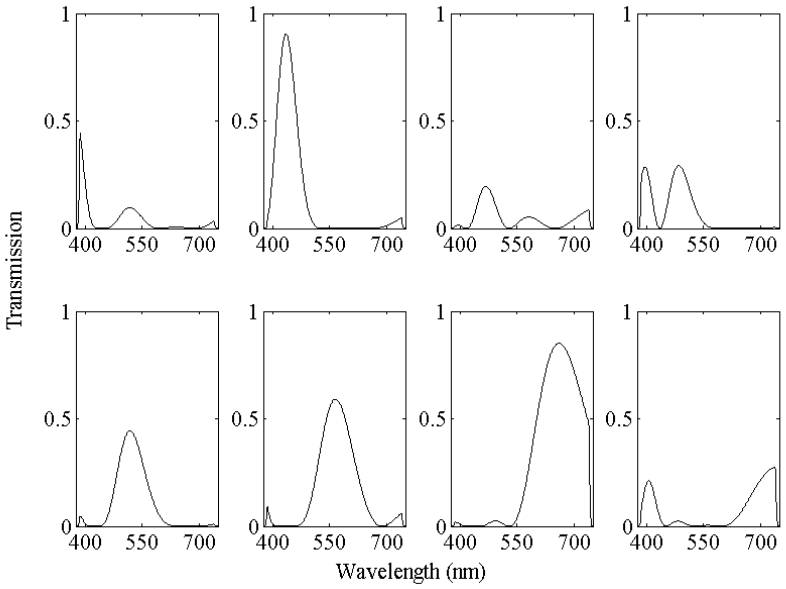

Figure 2.10: A set of transmission functions for an 8-channel IRIS system optimised for spectral discrimination of the coloured tiles of the MacBeth ColorChecker® shown in Figure 2.11. The optimisation was carried out using an exhaustive search.

Figure 2.11: MacBeth colour checker®.

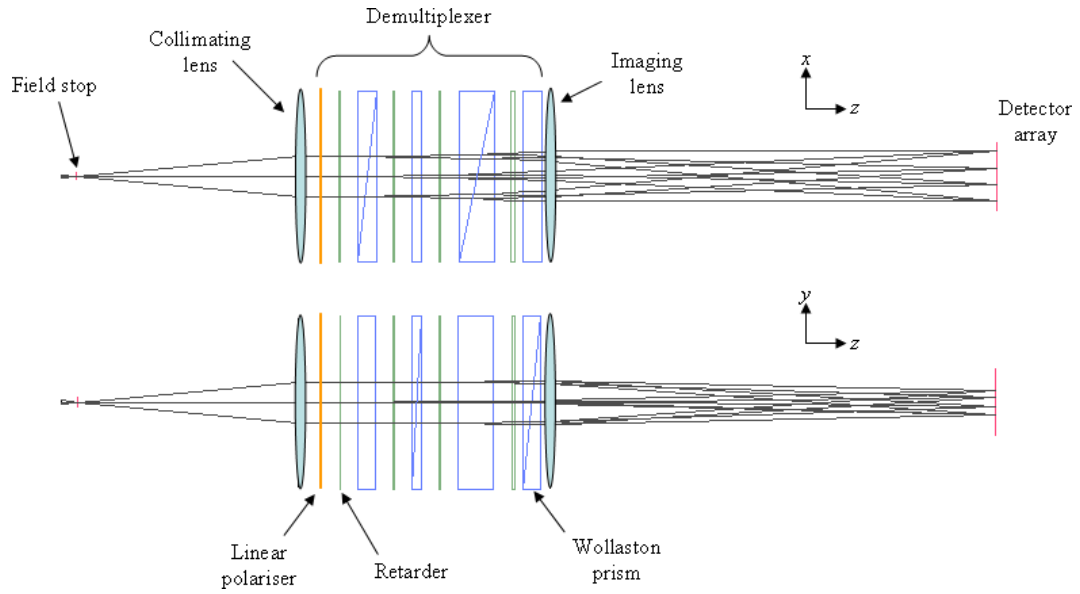

Figure 2.12: Top and side orthographic projection of 16-channel IRIS systems, shown with three rays traced from the centre of field stop to the detector.

2.4 Optical design

Figure 2.12 shows the 16-channel broadband IRIS system, for clarity the demultiplexer is shown as an exploded view; the beam is split twice in each of the $x-z$ and $y-z$ planes, resulting in 16 sub-images arranged 4×4 at the detector.

To ensure that the detector area is used efficiently, certain dependencies between optical parameters exist. Knowledge of these relations facilitates first-order optical design, which can be further refined using geometric ray tracing. What follows is an explanation of the optical design, including a survey of candidate optical materials, and a discussion of practical issues arising during the construction and use of IRIS systems.

2.4.1 Optical relations

In Figure 2.12 an image of the scene is formed at the plane of the field stop by an objective lens or other imaging apparatus of focal length f_O. A lens of focal length f_C collimates rays from the field stop; these rays then pass through the demultiplexer before being brought to focus at the detector by an imaging lens of focal length f_I. This arrangement, when correctly designed, will result in an array of sub-images at the focal plane. The $J = 2^M$ sub-images will be arranged $2^K \times 2^L$, where K and L are the number of times the beam is split in the $x-z$ and $y-z$ planes. The aspect ratio of a detector array is usually between 1:1 and 2:1, and the desired aspect ratio of the spectral images will often lie within this range. It will therefore be common to choose $K = L = M/2$ if M is even, or $K = ((M+1)/2)$ and $L = (((M-1)/2)$, if M is odd.

The required sets of K and L Wollaston-prism splitting angles are given by the

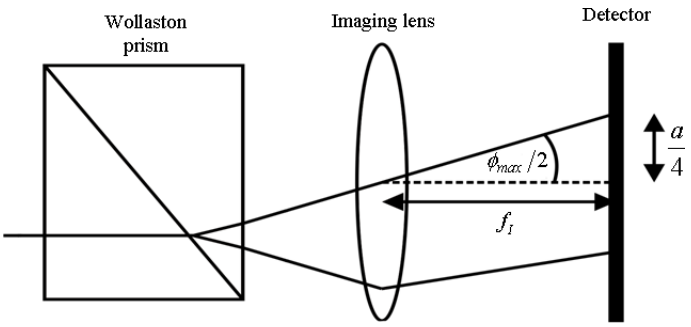

Figure 2.13: Illustration of the relation between the imaging lens focal length f_I, the maximum Wollaston prism splitting angle ϕ_{max} and the detector size a.

geometric progressions $\phi_k = [\phi_{max}]_x / 2^k$ and $\phi_l = [\phi_{max}]_y \Big/ 2^l$, where k and l take values from 0 to $K-1$ and $L-1$ respectively, and where ϕ_{max} is the maximum splitting angle in that plane. The ordering of prisms within the demultiplexer is arbitrary. As can be appreciated from Figure 2.13, the focal length f_I, of the imaging lens is related to ϕ_{max} and the dimensions of the detector $a_x \times a_y$ by [91]

$$f_I = \frac{a_x}{4\tan\left((\phi_{max})_x/2\right)} = \frac{a_y}{4\tan\left((\phi_{max})_y \Big/ 2\right)}. \tag{2.11}$$

The magnification between the plane of the field stop and the detector plane is f_I/f_C; this relates the dimensions of the detector $a_x \times a_y$ to those of the field stop aperture $S_x \times S_y$ through [91]

$$\frac{f_I}{f_C} = \frac{a_x}{2^K S_x} = \frac{a_y}{2^L S_y}. \tag{2.12}$$

As discussed in Section 2.3 the angle of incidence of rays entering the demultiplexer should be limited if the transmission is to stay reasonably constant across the field of view. The required tolerance on the transmissions is highly task-specific but, once known, can be related to a, and f_I. The maximum angle leaving the demultiplexer corresponds to a ray striking the corner of the detector:

$$\alpha_{max} = \tan^{-1}\left(\frac{\sqrt{(a_x/2)^2 + (a_y/2)^2}}{f_I}\right). \tag{2.13}$$

However the final retarder in a demultiplexer precedes the final Wollaston prism; therefore assuming that the final Wollaston prism gives the largest splitting angle, the maximum angle leaving the final retarder is given by

$$\beta_{max} = \tan^{-1}\left(\frac{\sqrt{(a_x/2)^2 + (a_y/4)^2}}{f_I}\right). \tag{2.14}$$

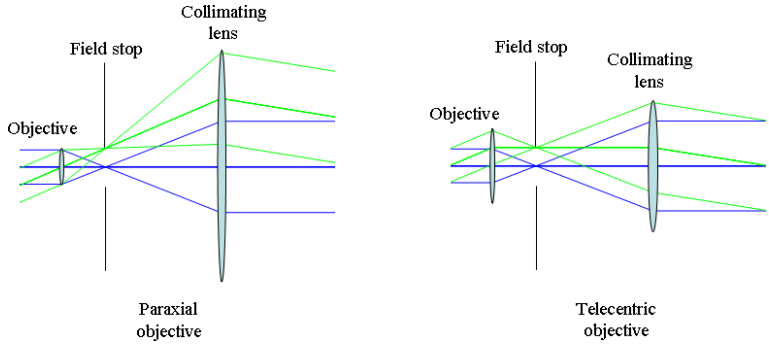

Figure 2.14: Paraxial objective lens vs. telecentric objective lens.

This is approximately equal to the maximum incidence angle (in air) on the final retarder.

The relationships thus far, if satisfied, ensure the correct arrangement of images at the detector. To avoid vignetting it is also required that the demultiplexer and lenses are sufficiently large. Calculation of the required apertures is difficult without accurate ray tracing, as the radial extent of the ray bundles through the system depend on the exact optical design. There are, however, two common approximations: when a lens is treated as a simple paraxial lens, for which the exit pupil is coincident with the lens itself, and when a lens is considered to be telecentric, for which the exit pupil is at infinity. Figure 2.14 shows the geometric ray trace for a paraxial objective and for a telecentric objective respectively, this ray trace was carried out using Zemax. If the objective lens can be approximated by a single paraxial lens of f-number F_O, so that the aperture of the objective lens is given by $D_O = f_O/F_O$, then the minimum required aperture of the collimating lens D_C is related to f_O, D_O, f_C, and the field stop size S, by

$$\frac{D_O + S}{f_O} = \frac{D_C - S}{f_C}. \tag{2.15}$$

If the objective lens is image-space telecentric then the required collimating lens aperture is given by

$$D_C = S + \frac{f_C}{F_O}. \tag{2.16}$$

The required sizes of the demultiplexer aperture and imaging lens entrance pupil are again difficult to state precisely without ray tracing, as they depend on the type of lenses used and the separation of components. Certainly these apertures have to be at least as large as the collimated beam diameter, which is given by $A = f_C/F_O$. In any case, the apertures should not have to be much larger than D_C and this is a reasonable starting value. The system f-number F_S is given by the ratio of f_I to the diameter of the collimated beam at the aperture of the imaging lens. If there is no vignetting by the collimating lens, then the diameter of the collimated beam A is determined by the f-number of the objective and the focal length of the collimating lens such that

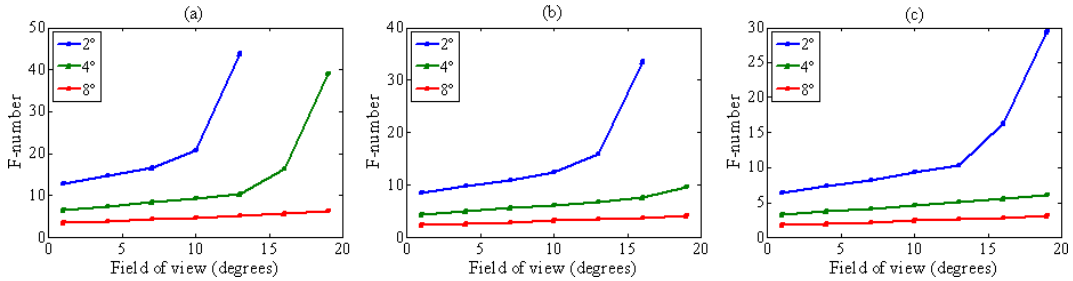

Figure 2.15: Minimum f-number against field of view for a 8 or 16 channel IRIS system with a paraxially approximated objective lens, with the Wollaston prism square aperture equal to (a) 20 mm, (b) 30 mm, and (c) 40 mm for different incidence angles, with $a_x = a_y = 15$ mm.

$A = f_C/F_O$. If vignetting is avoided then F_S is given by

$$F_S = F_O \frac{f_I}{f_C}. \tag{2.17}$$

A sub-image will be Nyquist-sampled if the pixel sampling frequency is equal to twice the cutoff frequency of the optical system which for a diffraction limited system with a circular aperture is approximately equal to $1/\lambda F_S$. It is not uncommon for the image produced by an optical imaging system to be undersampled.

The field of view of the system, as a half angle ψ is given by

$$\psi = \tan^{-1} \frac{S}{2f_O}. \tag{2.18}$$

2.4.2 Maximum field of view versus f-number

It is instructive to consider the maximum field of view as a function of system f-number F_S to understand the limitations of IRIS systems. This can be found using Equations 2.12 to 2.18.

For the purposes of this analysis the f-number of the lenses will be assumed to be greater than 1.4; this is a typical value for high-speed lenses in the visible region. Faster lens designs exist particularly in the thermal infrared region where materials with large refractive indices are available, making control of aberrations easier. Birefringent components with apertures larger than 40 mm in diameter are likely to be difficult to source. This limits D_C and D_I to around 40 mm. With these constraints, Figure 2.15 shows the minimum system f-number attainable for each field of view (as a half angle) for an 8 or 16 channel system with a paraxial objective, with the Wollaston prism square aperture equal to (a) 20 mm, (b) 30 mm, and (c) 40 mm. The legend in the plots refers to the maximum allowable angle (in air) incident on the final retarder. Figure 2.16 shows the same plots for a telecentric objective. Figure 2.17 shows the difference in minimum f-number with field of view with different numbers of channels, with Wollaston prism apertures of 30 mm, maximum incidence angle of 4°, and $a_x = a_y = 15$ mm.

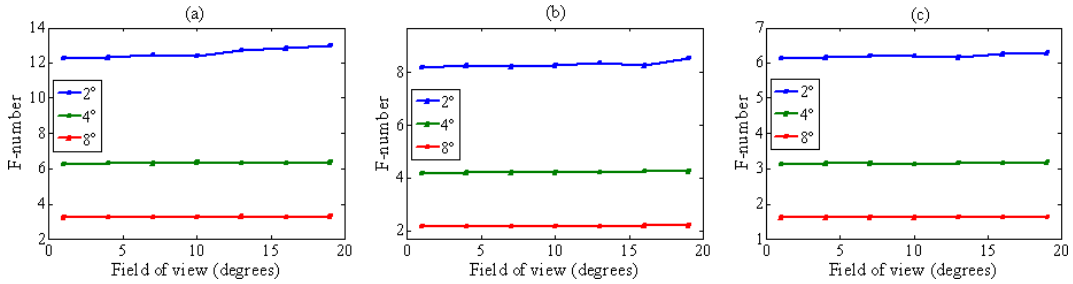

Figure 2.16: Minimum f-number against field of view for a 8 or 16 channel IRIS system with a telecentric approximated objective lens, with the Wollaston prism square aperture equal to (a) 20 mm, (b) 30 mm, and (c) 40 mm for different incidence angles, with $a_x = a_y = 15$ mm.

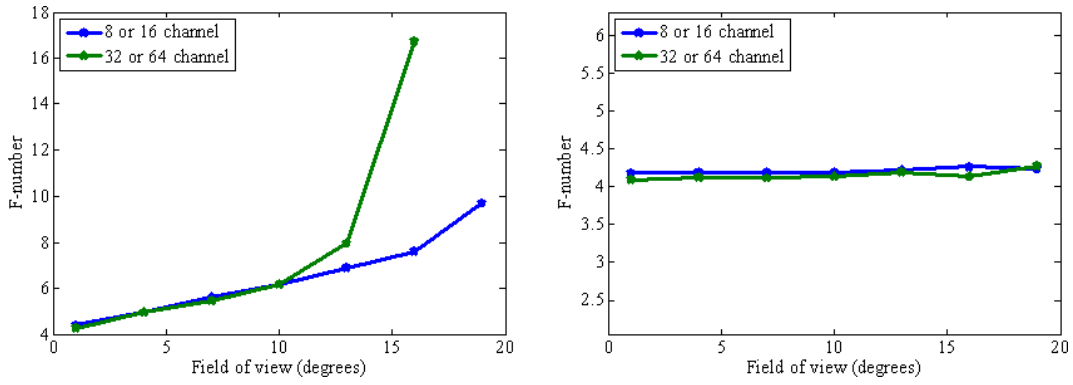

Figure 2.17: Minimum f-number against field of view for 8 or 16 channel IRIS systems compared to 32- or 64-channel systems, with Wollaston prism apertures of 30 mm, maximum incidence angle of 4°, and $a/2^M = 3.5$ mm, for (a) a paraxial objective, and (b) a telecentric objective.

FOV (°)	F_S	F_O	f_O(mm)	S(mm)	f_C(mm)	D_C(mm)	f_I(mm)
1	3.28	1.4	45.9	1.6	51.2	40	120
4	3.74	1.4	10.1	1.41	45	39.8	120
7	4.19	1.4	5.1	1.25	40.2	39.8	120
10	4.65	1.4	3.2	1.13	36.1	39.6	120
13	5.17	1.4	2.2	1.02	32.5	39.2	120
16	5.55	1.4	1.65	0.946	30.3	40	120
19	6.1	1.4	1.25	0.861	27.6	39.6	120

Table 2.1: Optical design parameters for 8 or 16 channel systems, for minimum f-number at each field of view (paraxial objective).

FOV (°)	F_S	F_O	f_O(mm)	S(mm)	f_C(mm)	D_C(mm)	f_I(mm)
1	3.13	1.4	48	1.68	53.6	40	120
4	3.14	1.4	12	1.67	53.5	39.9	120
7	3.14	1.4	6.8	1.67	53.4	39.8	120
10	3.13	1.4	4.75	1.68	53.6	40	120
13	3.16	1.4	3.6	1.66	53.2	39.7	120
16	3.16	1.4	2.9	1.66	53.2	39.7	120
19	3.18	1.4	2.4	1.65	52.9	39.4	120

Table 2.2: Optical design parameters for 8 or 16 channel systems, for minimum f-number at each field of view (telecentric objective).

Tables 2.1 and 2.2 show the optical parameters required for 8 or 16- channel systems, for minimum f-number for several values of the field of view, when the Wollaston prism apertures are 40 mm and the angle of incidence tolerance is $\pm 4°$.

2.4.3 Example IRIS design

To illustrate the design principles, consider the 8-channel retinal IRIS system introduced in 2.1, manufactured for retrofitting to a modified DisCam ophthalmoscope (Figure 2.18). The instrument is used for spectral imaging of the retina. The opthalmoscope serves as the objective lens in this example; it has a f-number of $F_O = 5.4$ and is highly telecentric in the image space. The field of view of the final system should be as large as possible, implying that the field stop and thus f_I/f_C be as large as possible. Calcite Wollaston prisms with square aperture 30 mm $\times$ 30 mm are used for this system, which limits f_I to a minimum of 40 mm for an f-number of 1.4; a 50 mm lens is used, as these are more commonly available commercially. The detector is 15 mm $\times$ 15 mm and the sub-images are arranged in a 2×4 configuration. With $f_C = 50$ mm, the three required splitting angles are (from Equation 2.11) 8.58° in the x-plane, and 4.29° and 8.58° in the y-plane.

For a telecentric objective, $D_C = S + f_C/F_O$ and $S_x = f_C a_x / f_I K$. With D_C lim-

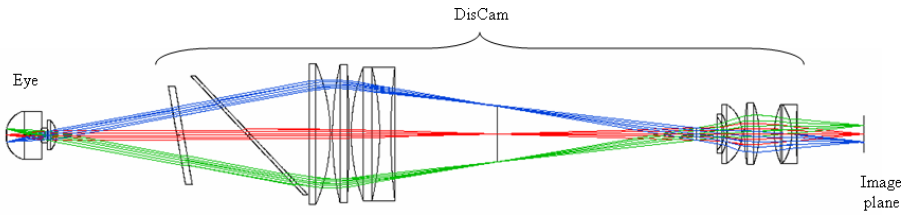

Figure 2.18: Zemax ray trace of a modified DisCam opthalmoscope.

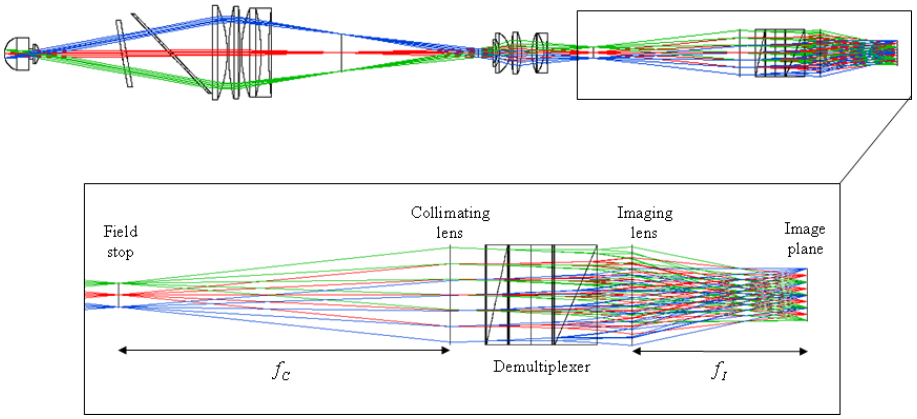

Figure 2.19: Zemax ray trace of a modified DisCam ophthalmoscope with the retinal IRIS system.

ited to 30 mm and $K = 2$ this limits f_C to a maximum of 90 mm, making $S =$ 13.5 mm $\times$ 6.75 mm. In actual fact, the output image size of the DisCam is limited to 9 mm $\times$ 9 mm and thus $f_C < 110$ mm. A 100 mm lens will be used, again due to availability, giving the size of S to be 9 mm$\times$ 7mm, and a sub-image size of 4.5 mm$\times$3.5 mm, leaving some of the detector area unused. The system f-number is 2.8; this increase in speed over the DisCam decreases the spatial sampling when compared to the same detector placed at the image plane of the ophthalmoscope .

The optical relations in this section facilitate the first order design of IRIS systems and ensure that the the area of interest is correctly imaged onto the detector, however care must also be taken to ensure that unwanted sources of radiation do not reach the detector. A discussion of the control of this stray light follows.

2.4.4 Stray light

There are a number of potential sources of unwanted light in an IRIS system and the optical design should attempt to remove or minimise the contribution of these sources. This subsection describes steps that should be taken to achieve this for visible/near-infrared systems and for thermal-infrared systems, in which stray light control is inherently more challenging.

Figure 2.20 illustrates the sources of stray light in IRIS systems by tracing rays outward from the detector through the optical system. The field stop in IRIS systems

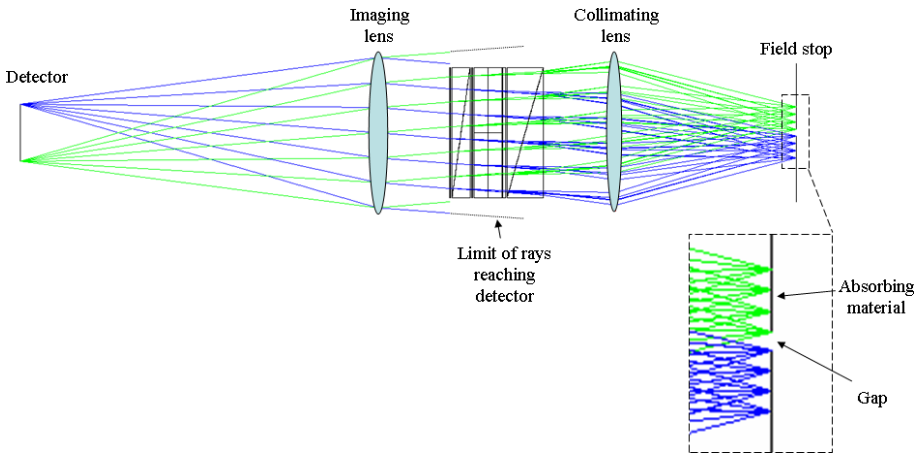

Figure 2.20: Zemax ray trace illustrating of the potential sources of stray light in IRIS systems. Note that the detector is considered the source in this case.

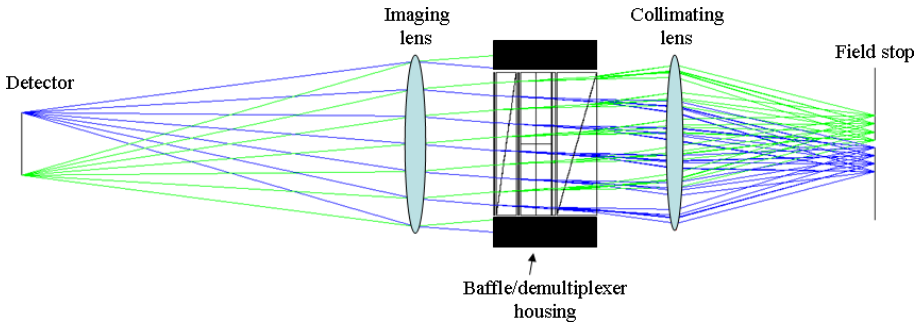

Figure 2.21: Zemax ray trace showing the use of a baffle to reduce stray light in IRIS systems.

is required to prevent overlap of the sub-images at the detector, and this results in the material surrounding the gap also being imaged onto the detector. The field-stop surface should therefore not emit or reflect appreciable amounts of radiation over the range of wavelengths to which the detector is sensitive. Depending on the optical design, the imaging lens may also accept radiation not originating from the plane of the field stop, including radiation originating from the demultiplexer housing. This potential source of stray light can be eliminated by placing an absorbing baffle (Figure 2.21) around the exit aperture of the demultiplexer and by ensuring that the demultiplexer housing has high absorptance over the detector spectral range.

In the thermal infrared these objectives become more difficult as efficient absorbers of radiation are also efficient emitters of radiation, meaning that any stops or baffles must also be cooled. The need to cool the housing and baffles can be removed by ensuring that the system is designed so that all radiation reaching the detector originates from the plane of the field stop. This is aided by moving the components as close together as possible and may require the entrance aperture of the camera to be stopped down. In this case, stray light can be controlled by cooling the field stop. The cooled area of the field stop should be at least $S_x \left(2^{K+1}+1\right) \times S_y \left(2^{L+1}+1\right)$ if all non-signal rays reaching

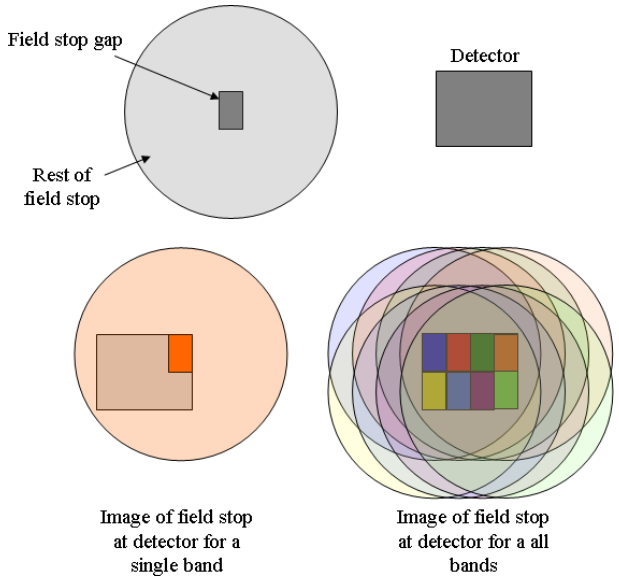

Figure 2.22: Illustration of the replicated images of the entire field stop at the detector plane. The cooled area of the field stop should be large enough that its image in each band covers the entire detector.

the detector are to originate from a cooled surface (Figure 2.22).

If the field stop can be considered a blackbody of temperature T, the spectral photon radiance from this surface is given by the Plank function $B(k,T)$. If the spectral photon radiance of the source is $x(k)$, and the solid angle of radiation arriving the detector is written as $\pi/4F_S^2$ with F_S the system f-number, then for sub-image i, the number of photons detected by a single pixel of area b^2, in time t, is given by

$$y^i = \frac{\pi b^2 t}{4F_S^2}\int_{k_1}^{k_2}\eta(k)\,\tau^i(k)\,x(k)\,dk + \frac{\pi b^2 t}{4F_I^2}\sum_{i\neq j}^{J}\int_{k_1}^{k_2}\eta(k)\,\tau^j(k)\,B(k,T)\,dk \tag{2.19}$$

where the first term in the sum corresponds to radiation from the scene and the second term is radiation from the field stop. The radiation from the field stop constitutes unwanted background radiation which will decrease the signal to noise ratio for a given integration time. The photon-counting process is expected to have Poisson statistics, meaning that the uncertainty associated with a count of N photons is $\sqrt{N}$. In the absence of other sources of noise, the signal-to-noise ratio (SNR) is given by

$$SNR = \frac{\frac{\pi d^2 t}{4F_S^2}\int_{k_1}^{k_2}\eta(k)\,\tau^i(k)\,x(k)\,dk}{\sqrt{y^i}}. \tag{2.20}$$

To demonstrate the problems faced when designing infrared systems, consider the MWIR IRIS system shown in Figure 2.23. This prototype system is based on a design concept outlined in a report produced by David Fletcher Homes and Andrew Harvey [91]; the system was manufactured by SELEX-Galileo (then BAE systems) using existing

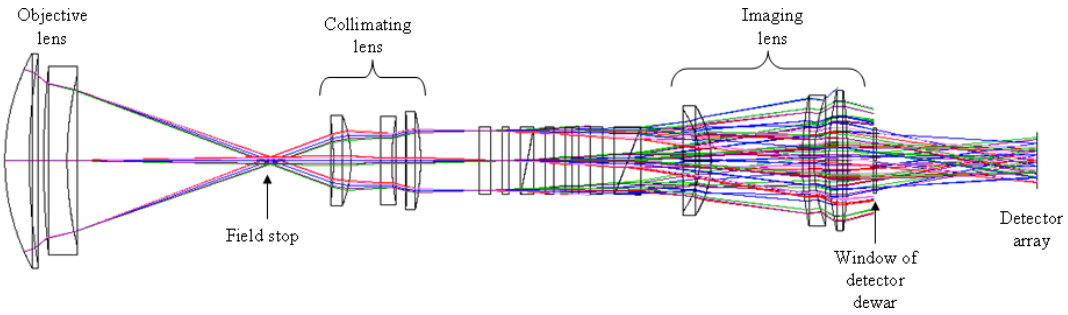

Figure 2.23: Zemax ray trace of the MWIR IRIS system.

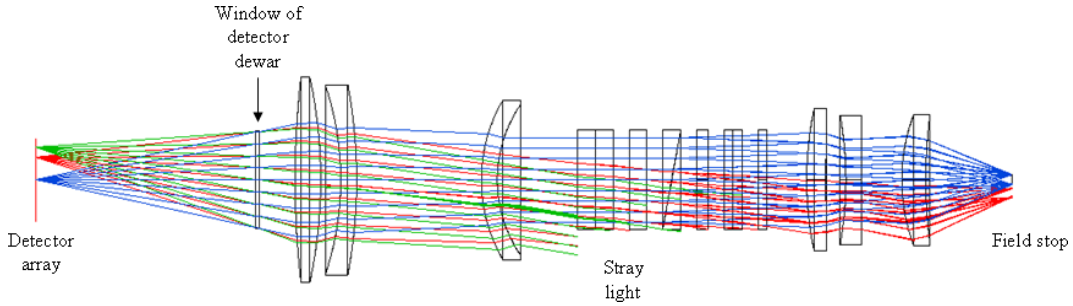

Figure 2.24: Zemax ray trace of the MWIR IRIS system with rays traced outwards from detector.

lenses, 16 mm diameter circular MgF_2 retarders and square YVO_4 Wollaston prisms with side 11.5 mm. Upon construction of this system it became apparent that there were problems with stray light. This system has a field of view of 1.5° and operates in the 3-5 μm range, projecting 8 sub-images, arranged in a 4×2 configuration onto 320×160 square pixels, each of area 900 μm^2. The system has an f-number $F_S = 2.41$, although due to vignetting of between 30% and 80% the true speed is between $f/4.4$ and $f/2.7$ depending on the sub-image. The field-stop aperture is circular with a diameter of 1.17 mm; the field stop surface is cooled using a thermoelectric cooler and has diameter 5 mm. Tracing rays from the detector to the field stop, as in Figure 2.24, reveals the problems with stray light from the imaging lens and instrument housing. Figure 2.25 shows that the fraction of rays leaving the detector which intersect with the plane of the field stop decreases with distance from the centre of the detector.

To validate this stray-light analysis, a measured image can be compared to an image simulated using the commercial ray tracing program Zemax. The camera used is a 14 bit Phoenix camera from Indigo Systems using an InSb detector cooled to 77 K; the well capacity is 18 million electrons and the quantum efficiency is 80%. Figure 2.26 shows the detected image when the object is a homogeneous blackbody at temperature 293 K. The integration time is 0.4 ms and there is a background count of 4900 grey-levels. The temperature of the optics and housing for these measurements is 293 K and the cooled area of the field stop is at a temperature of 273 K. With these same parameters and treating all surfaces as blackbodies, the simulated image is shown Figure 2.27, and is

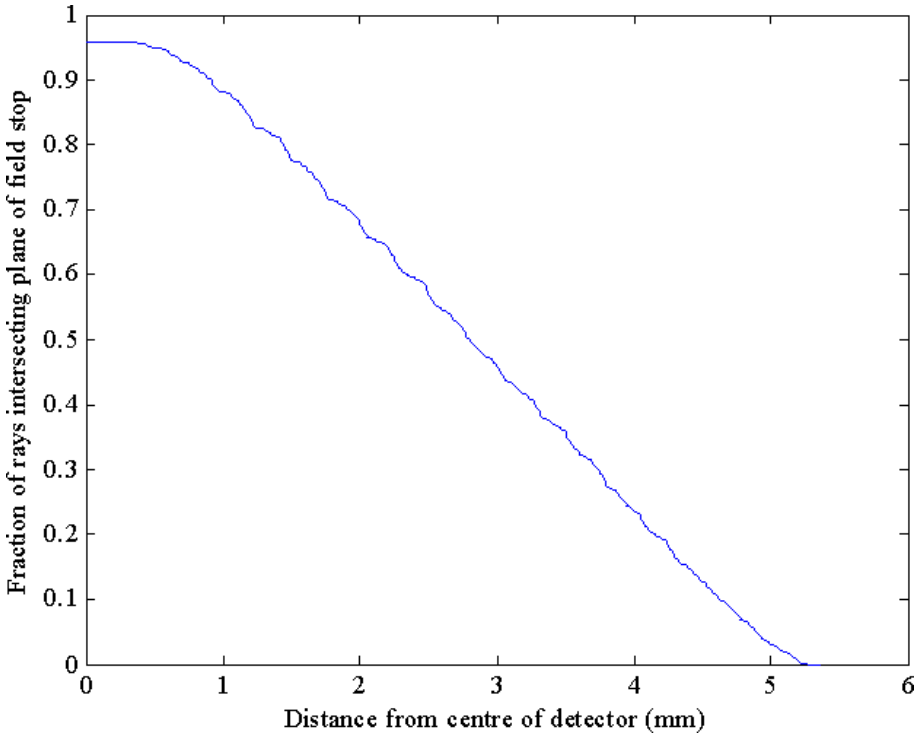

Figure 2.25: Fraction of rays leaving detector that intersect with the plane of the field stop.

in good agreement with the measured image. The assumption that the surfaces are perfect emitters is likely to be poor for the housing and field stop, as these surfaces are partially reflective. This is not likely to change the expectation value of the detected signal but may introduce additional, possibly signal-dependent, uncertainty into the measurements.

The signal-to-noise ratio (SNR) for the IRIS system is estimated by aquiring multiple images of the blackbody as shown in Figure 2.26 and finding the variance for a pixel at the centre of the field stop. The square root of this value is the denominator in Equation 2.20 and the numerator is estimated from the simulated data. The SNR calculated using this method is shown in Figure 2.28; the SNR calculated using the ray trace is shown in Figure 2.29. The measurements have a standard deviation roughly twice as large as the the ray-trace analysis suggests. This difference is most likely due to the weak assumption that the housing and field stop are perfect emitters, and to a lesser extent because other sources of noise have been ignored. For this system and an object of 293K, the maximum integration time before saturation of the sub-image area occurs is 4.38 ms. In this case the SNR, as calculated from the stray-light analysis, is shown in Figure 2.30. If the stray light is reduced or removed, then the integration time can be increased to improve the SNR. A possible reconfiguration to achieve this is shown in Figure 2.31. The effective speed of this system is $f/3.8$, and all radiation reaching the detector comes from the plane of the field stop. This is achieved by using a reflective 'warm shield' to stop down the camera at the dewar window, as well as moving the

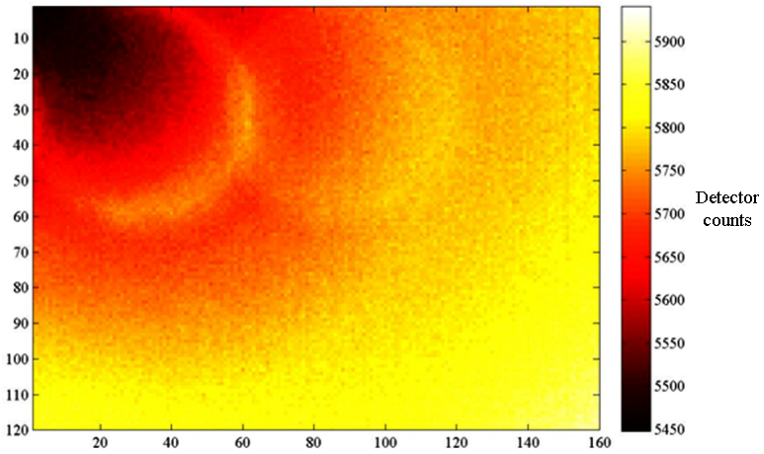

Figure 2.26: Lower right corner of measured image from the MWIR IRIS system, when the object is a 293 K blackbody.

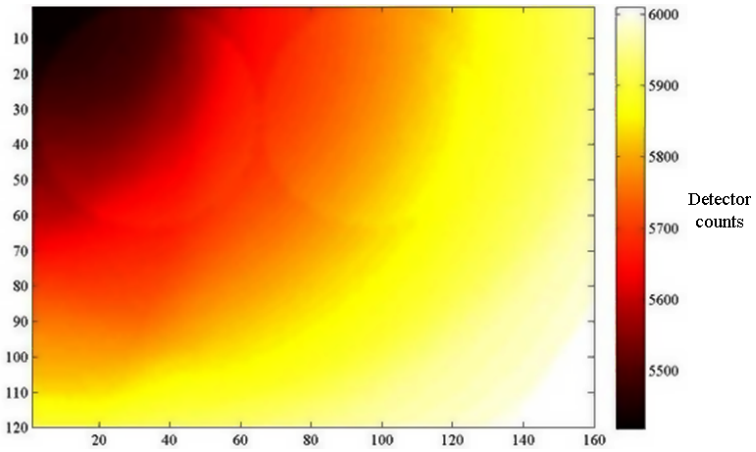

Figure 2.27: Lower right corner of simulated image from the MWIR IRIS system, when the object is a 293 K blackbody.

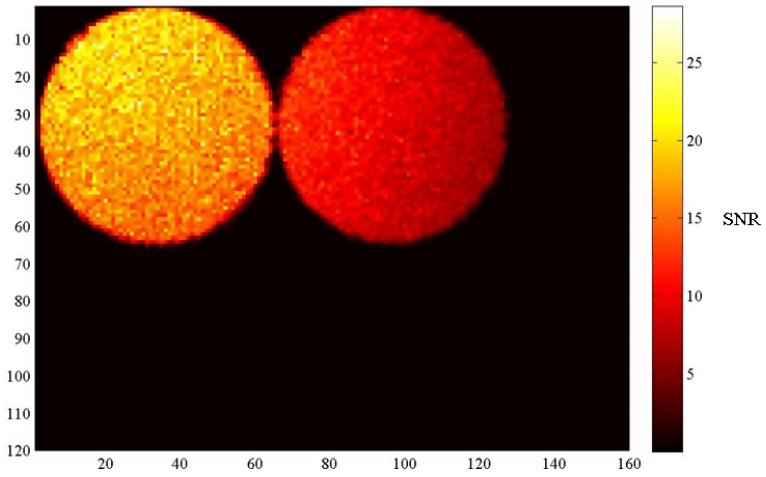

Figure 2.28: Measured SNR for lower right corner of simulated image from the 8-channel 3-5 μm IRIS system, when the object is a 293 K blackbody.

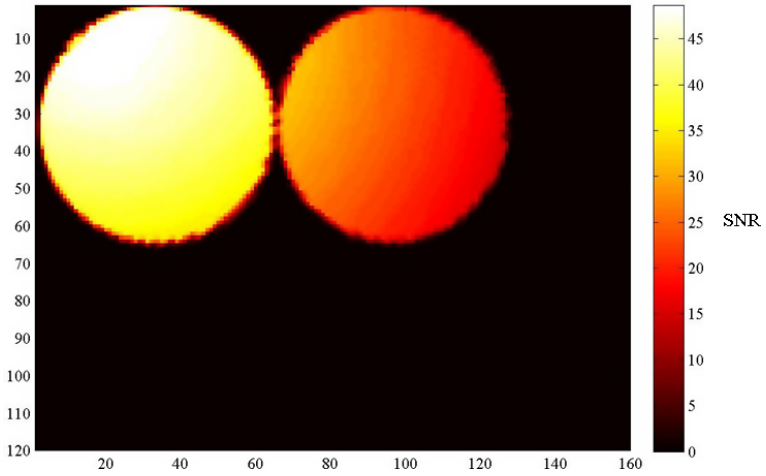

Figure 2.29: SNR for lower right corner of simulated image from the MWIR IRIS system, when the object is a 293 K blackbody.

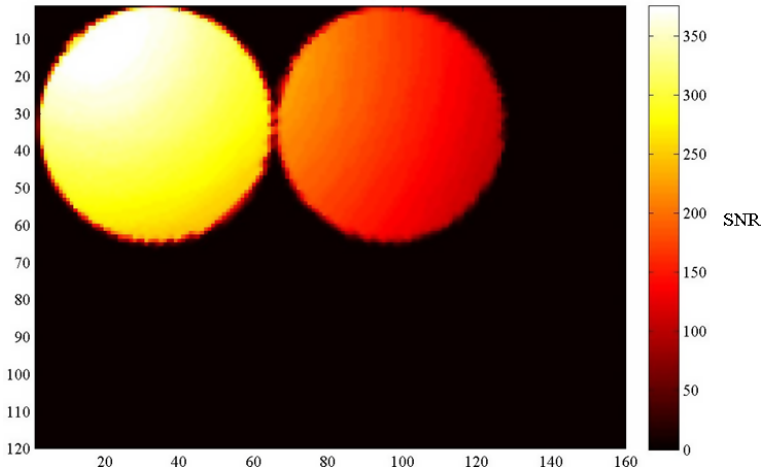

Figure 2.30: Maximum possible SNR for lower right corner of simulated image from the MWIR IRIS system, when the object is a 293 K blackbody.

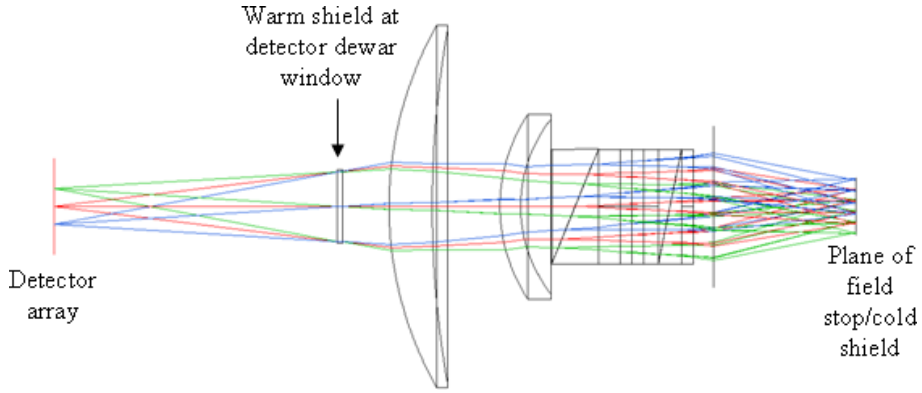

Figure 2.31: Redesign of imaging lens and demultiplexer for the MWIR IRIS system to ensure all rays originate from the plane of the field stop.

components together and redesigning the imaging lenses. The field stop aperture is also made square. As this design is restricted to using existing components, it is unlikely to be optimal in terms of speed and imaging performance. Figure 2.32 shows the expected SNR for an object at 293 K with an integration time of 4.38 ms. It can be seen that for the same integration time, this design offers roughly half the SNR of the previous system, with the advantage that the potentially time-dependent stray light is removed. With the stray light removed the maximum integration time without saturation of the image area can be increased to 23 ms. Figure 2.33 shows the expected SNR for the 293 K object with this integration time, for the 293 K object. The maximum SNR is roughly twice as large as can be obtained with the original system. Unfortunately, implementation of this reconfiguration, or any other improvement, has yet to be carried out due to restrictions on the availability of the equipment.

This section has described the optical design of IRIS systems, once the first order design is in place ray-tracing software can be used to refine the design. One further

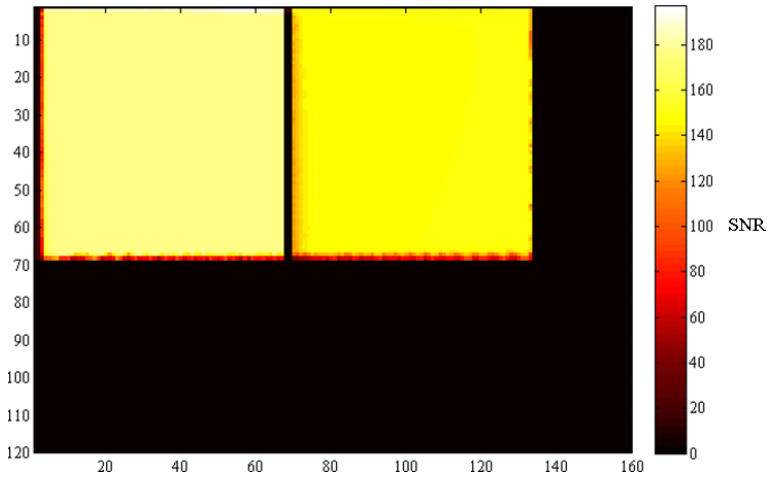

Figure 2.32: Maximum possible SNR for the redesigned MWIR IRIS system, when the object is a 293 K blackbody, with integration time of 4.38 ms.

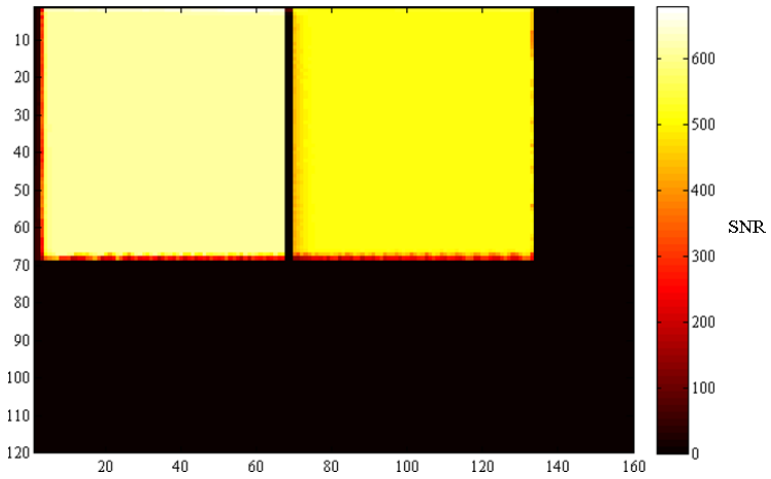

Figure 2.33: Maximum possible SNR for the redesigned MWIR IRIS system, when the object is a 293 K blackbody, with integration time of 23 ms.

important consideration in the design and manufacture of IRIS systems is the choice of optical materials, this is discussed in the next section.

2.5 Optical materials

Imaging components

The IRIS optical system consists of the demultiplexer, the field stop, and components required for collimation and imaging. A fundamental requirement for the system is that the throughput is large. This requires the efficiency of each component to be as high as possible across the wavelength range of interest, and also requires that they are available in the suitable sizes. For the retarders and Wollaston prisms the materials must also exhibit a degree of birefringence such that the components are of reasonable size. The materials must also be easily machined and robust enough to withstand operating conditions.

Lens technology is very mature for the optical regime and there exist a large number of glasses suitable for the UV, visible and near-infrared. At longer infrared wavelengths, silicon, germanium, and zinc selenide are commonly used. Mirrors are also a mature technology and materials with high reflectivity exist across the entire optical regime.

Retarder materials

Figure 2.34 shows the transmission ranges of many birefringent materials over the optical regime [92–95]. Quartz and MgF_2 are commonly used for retarders in the visible and near infrared. If the number of channels is 2^M and the system bandwidth is $\Delta k \mathrm{cm}^{-1}$ then, from Equation 2.6, a retarder thickness of $2^M/B\Delta k$ will give a resolution of approximately $\Delta k/2^M$. If this is the thickest retarder in the demultiplexer, then the thinnest retarder will have a thickness of approximately $1/B\Delta k$.

If high spectral resolution is necessary then a material with larger birefringence, such as calcite, will reduce the required thickness of retarder. Magnesium fluoride remains a useful retarder material across the mid-infrared, as does sapphire. CdSe and CdS are also commonly used for retarders at far-infrared wavelengths.

Prism materials

The precise action of a Wollaston prism made from a material with birefringence $B(k)$, and wedge angle θ, on a collimated beam can be found via ray tracing using Snell's law and the expression for the extraordinary-ray refractive index as a function of incidence angle (Appendix A). For small splitting angles it can be assumed that the splitting of the beam is nearly symmetric about the optical axis, in this case Figure 2.35 shows that the splitting angle introduced by a prism can be approximated by $\psi = 2B(k)\tan\theta$.

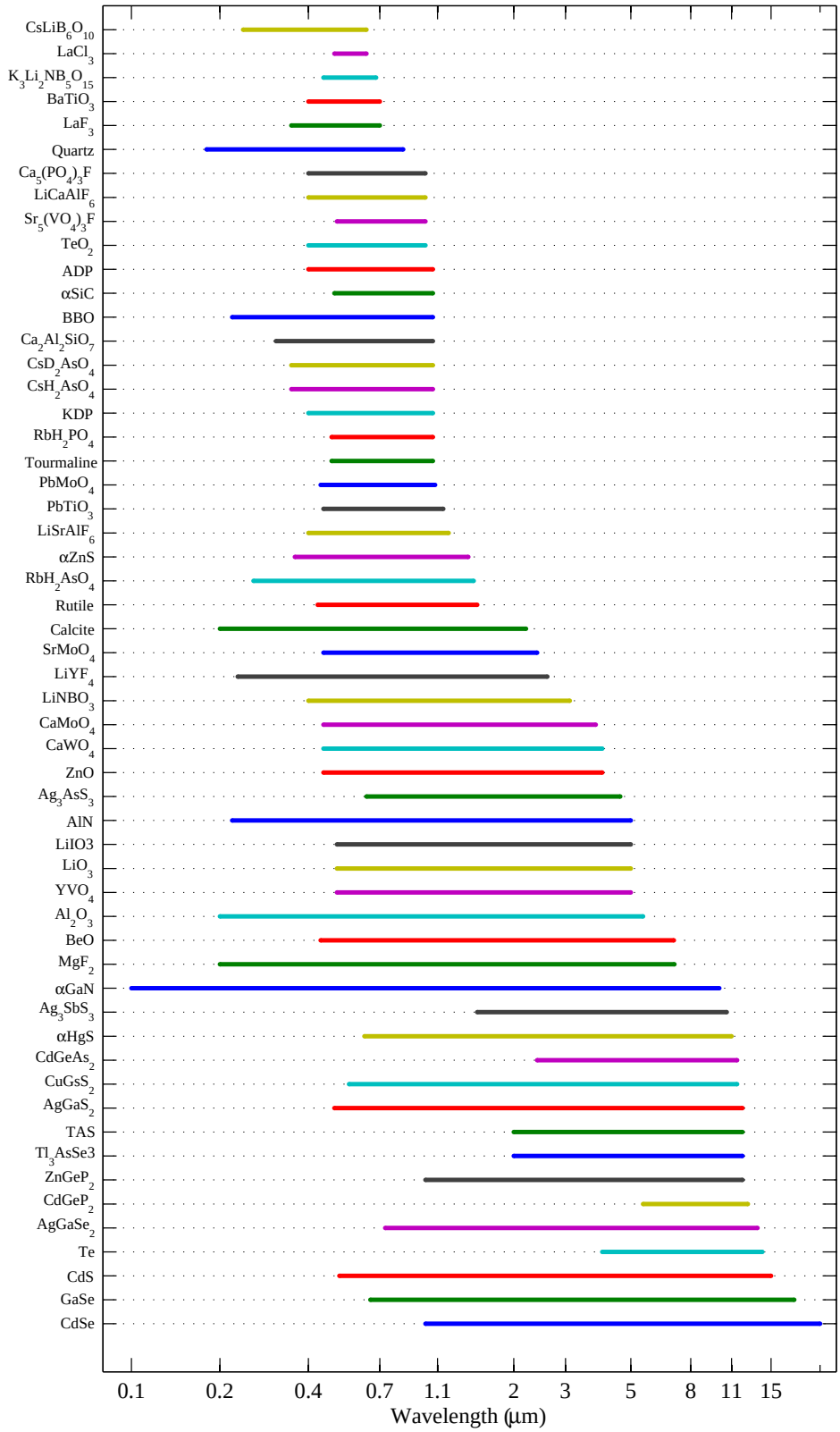

Figure 2.34: Wavelength range over which birefringent materials are known to be transmissive. A material is regarded as transmissive at a given wavelength if the transmission of a 1 mm thick sample is greater than 10%.

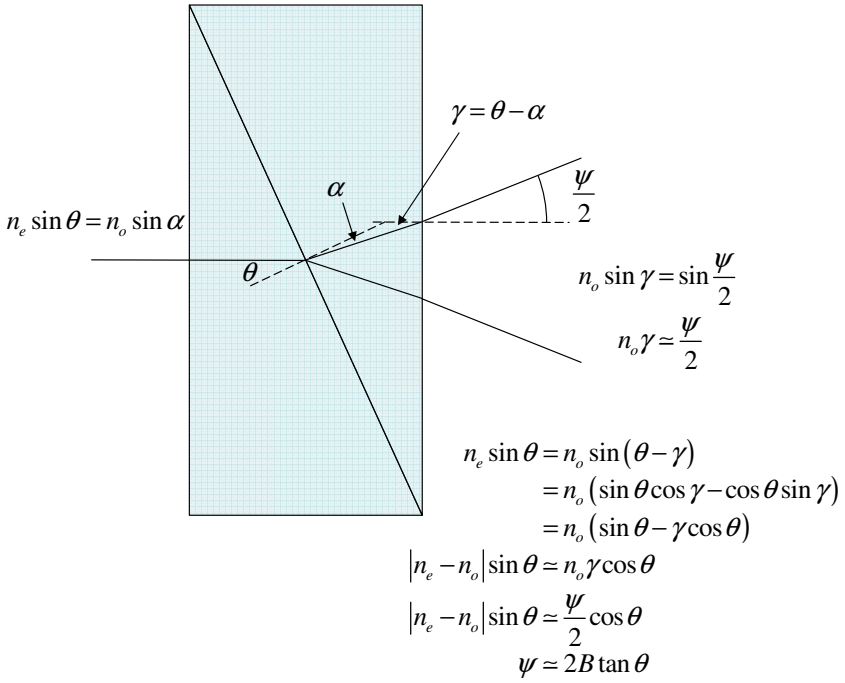

Figure 2.35: Approximate splitting angle for Wollaston prisms with small ($< 30°$) splitting angles, derived using Snell's law [4].

Figure 2.36 shows the maximum splitting angle required as a function of f_I for a detector size of $a = 15$ mm. Figure 2.37 shows the required wedge angle as function of birefringence for several splitting angles. To reduce the prism volume, materials with larger birefringence are preferred.

A second important consideration when choosing a prism material is dispersion in the birefringence. This dispersion causes radiation from a point source to be spread over a line of pixels in each sub-image. As will be described in more detail in Section 2.6.1, the maximum extent of dispersion in a sub-image is approximately proportional to the maximum change in birefringence divided by the average birefringence, $\Delta B/\bar{B}$. Figures 2.38 to 2.44 show the birefringence and $\Delta B/\bar{B}$ for birefringent materials over several wavelength ranges [92–95]. Note that this is not an exhaustive list of materials. Calcite is commonly used to make Wollaston prisms in the visible region, due to its high birefringence. Rutile, BBO, and YVO_4 also have relatively high birefringence and offer less dispersion than calcite, though rutile is expensive [96]. For very small splitting angles magnesium fluoride may be an attractive option.

In the mid-infrared YVO_4 has a relatively low dispersion and suitably large birefringence. At longer infrared wavelengths choice becomes limited; CdSe, CdS, and $AgGaS_2$ components are available but the birefringence is low. TAS is already used for acousto-optic filters in the far-infrared and has moderately large birefringence. Tellurium is a very promising material, as it has high birefringence, low dispersion, and is available in

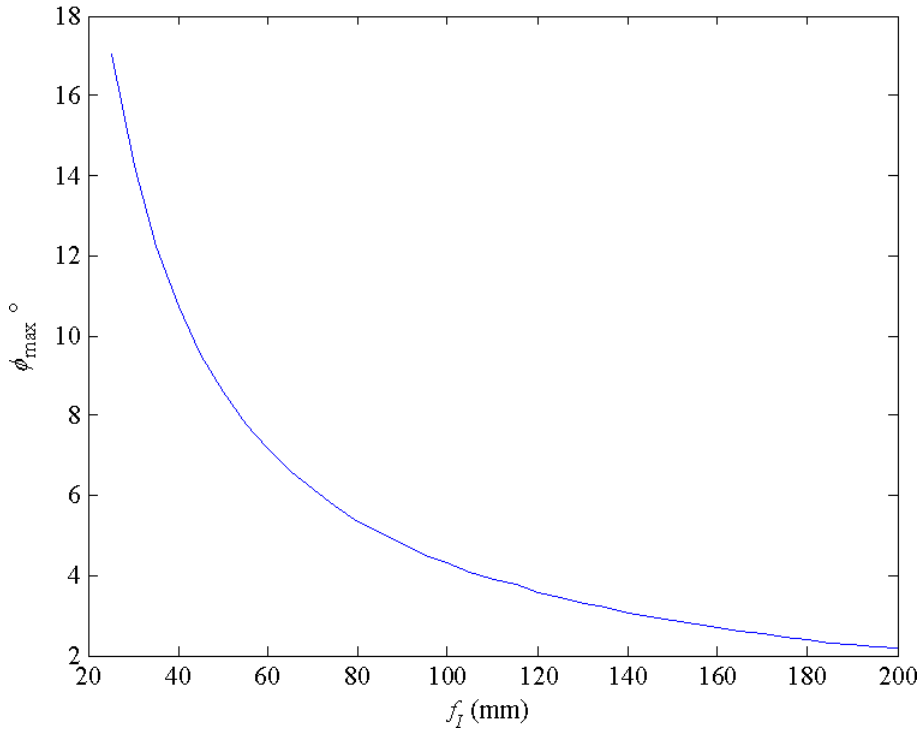

Figure 2.36: Splitting angle as a function of imaging lens focal length.

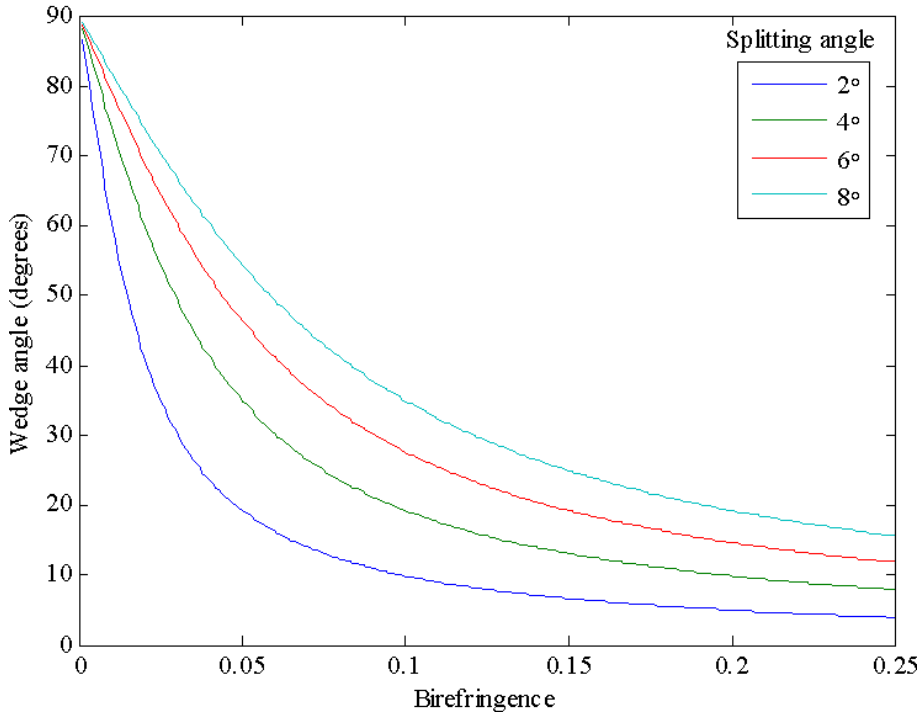

Figure 2.37: The required Wollaston prism wedge angle as function of birefringence for several splitting angles.

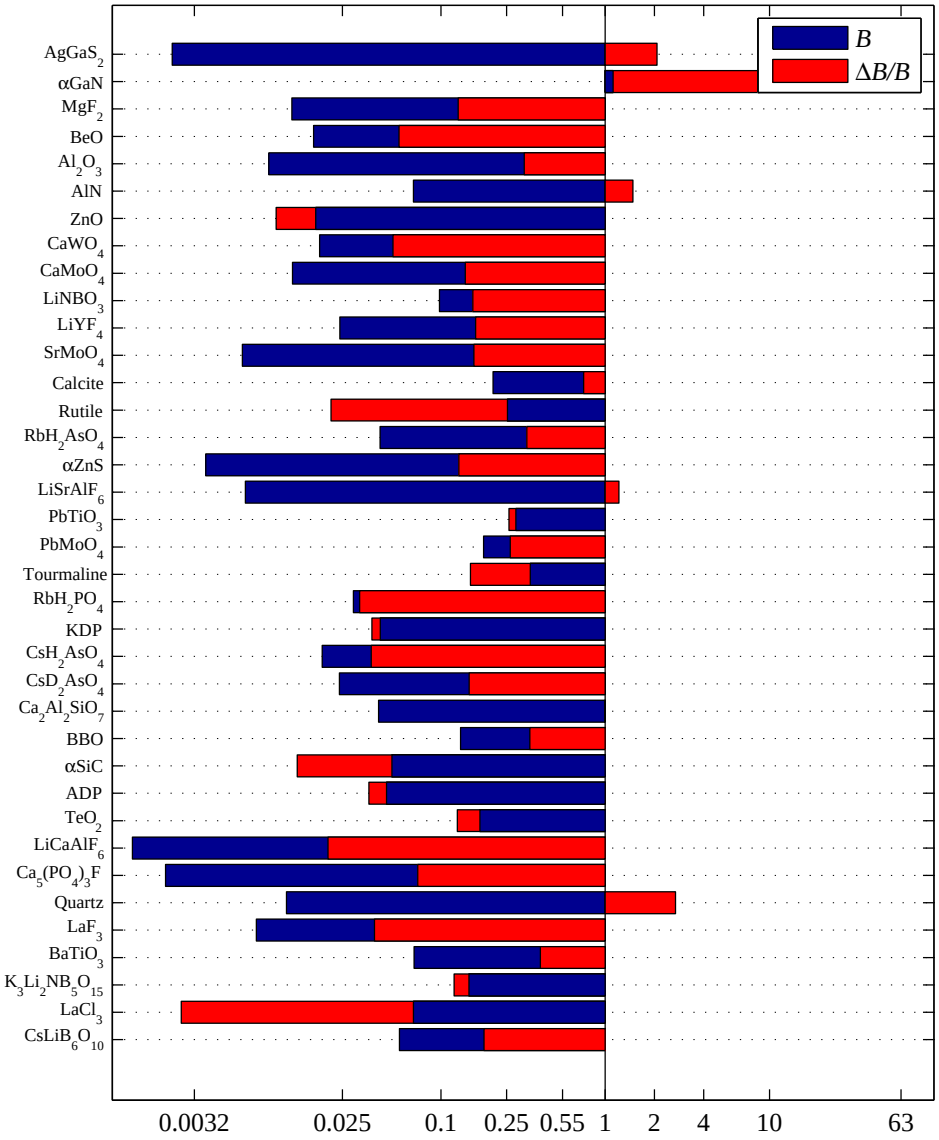

Figure 2.38: Values for birefringence and $\Delta B/B$ for materials transparent in the range 0.1 nm to 0.5 nm, larger values for the birefringence (blue bars) and smaller values for the dispersion (red bars) are desirable for prism materials. Both quantities are unitless.

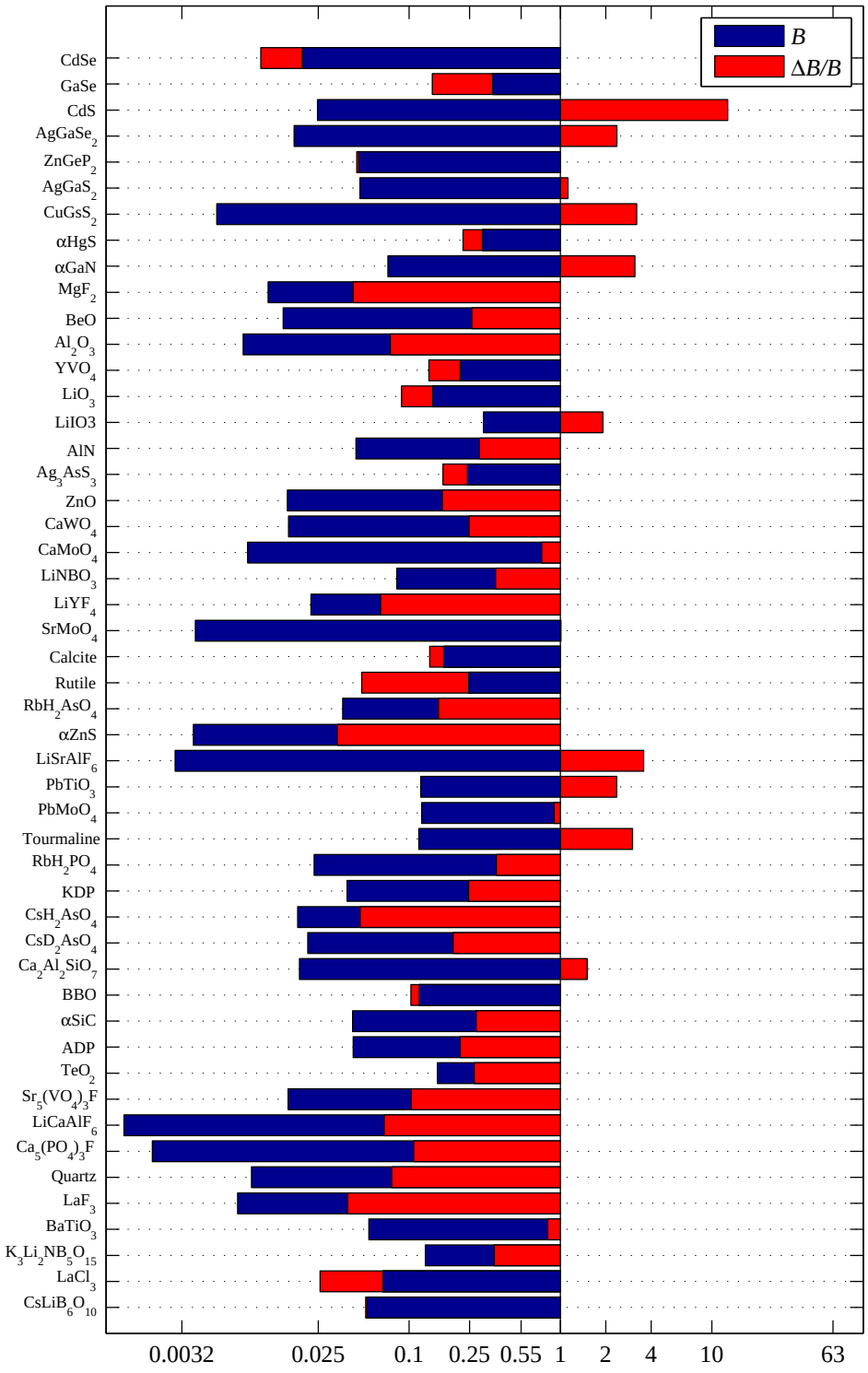

Figure 2.39: Values for birefringence and $\Delta B/B$ for materials transparent in the range 440 nm to 1.1 μm. Both quantities are unitless.

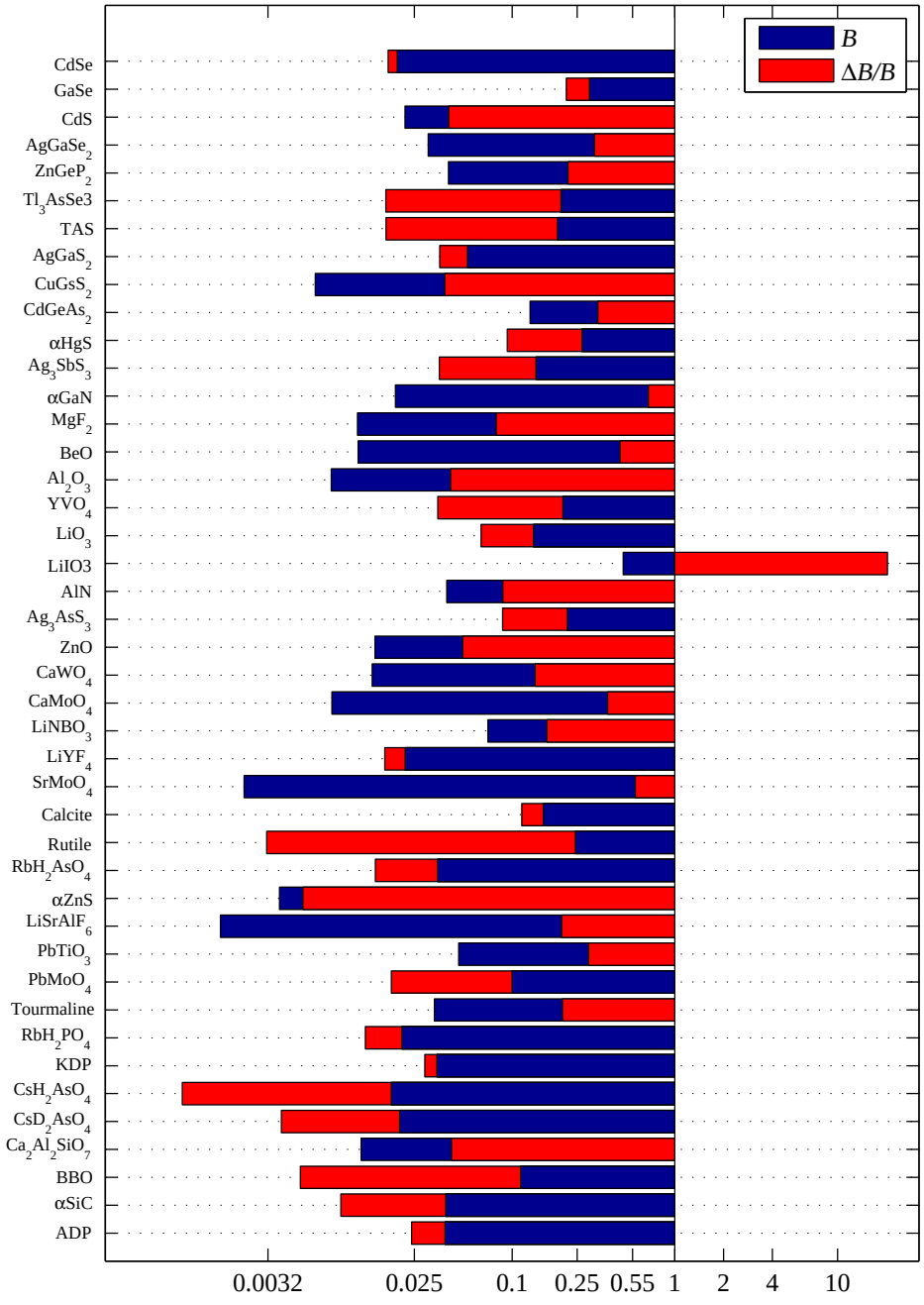

Figure 2.40: Values for birefringence and $\Delta B/B$ for materials transparent in the range 1 μm to 3 μm. Both quantities are unitless.

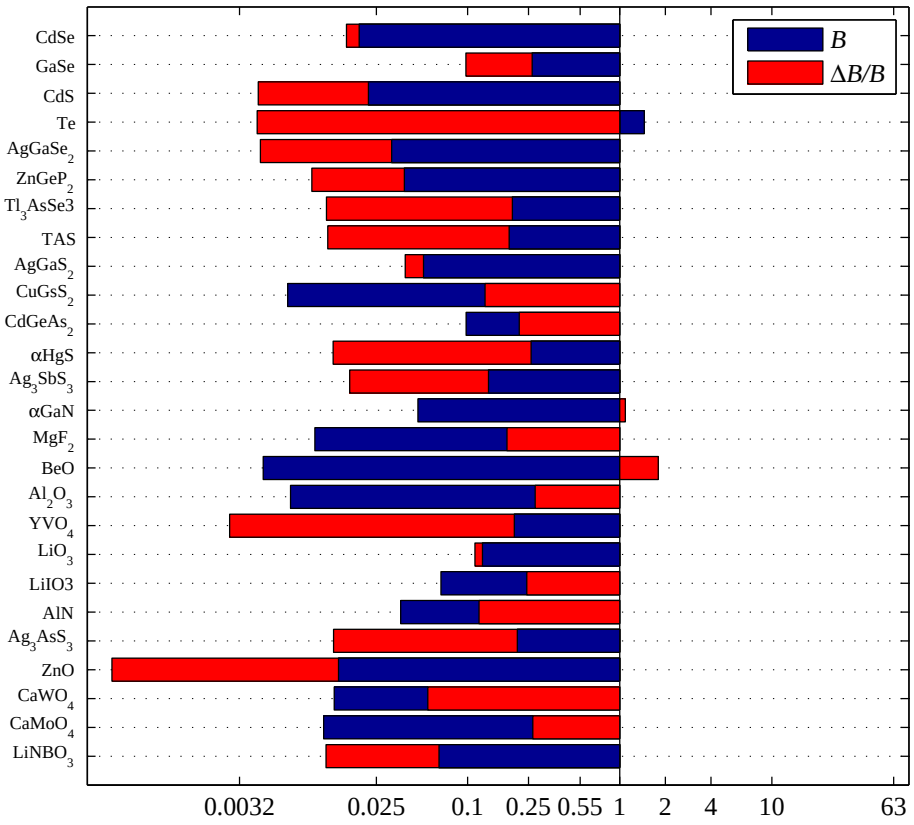

Figure 2.41: Values for birefringence and $\Delta B/B$ for materials transparent in the range 3 μm to 5 μm. Both quantities are unitless.

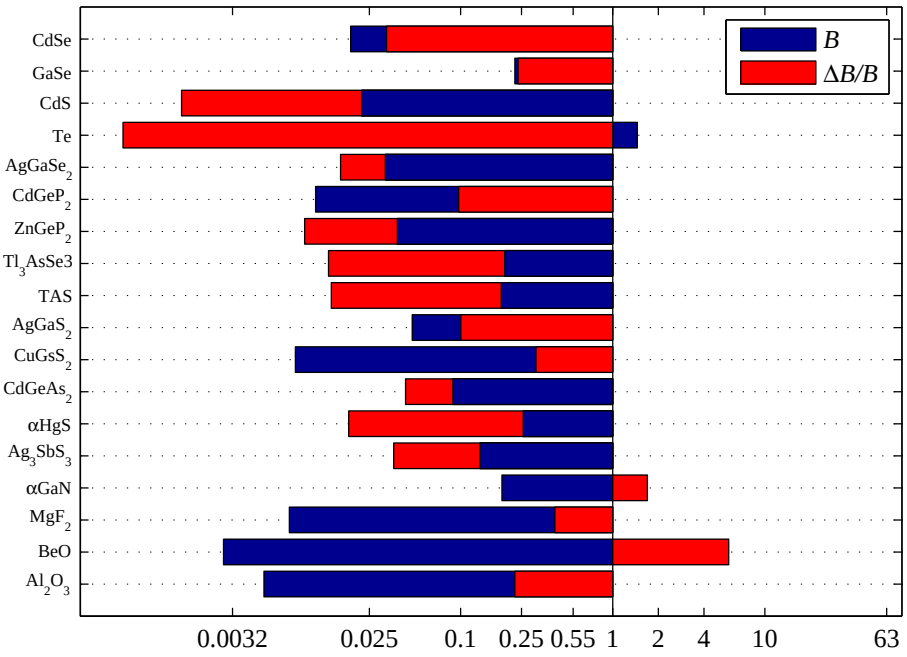

Figure 2.42: Values for birefringence and $\Delta B/B$ for materials transparent in the range 5 μm to 8 μm. Both quantities are unitless.

large single crystals [97]; the suitability of Tellurium as an acousto-optic material has also been investigated [98].

Minimising Fresnel losses

Losses will occur through bulk absorption and Fresnel losses for transmissive components. To minimise Fresnel losses, anti-reflection (AR) coatings can be used on optical surfaces. Single-layer MgF_2 coatings are inexpensive and (with a refractive index of 1.38) suitable across a wide range of wavelengths and materials, reducing reflectivity for many commonly used glasses. For materials with refractive index < 1.8, better performance may be achieved using fluoropolymer coatings. Multilayer coatings can also be tailored to give high transmission over a wide range.

For the demultiplexer, better performance is obtained for lower cost by using index-matching material in between components rather than AR coating all surfaces. This approach has been used for all visible-band systems described in this thesis. The optimal refractive index n_m for matching two materials with indices n_1 and n_2 is found by differentiating with respect to n_m the expression for the total Fresnel reflectance and

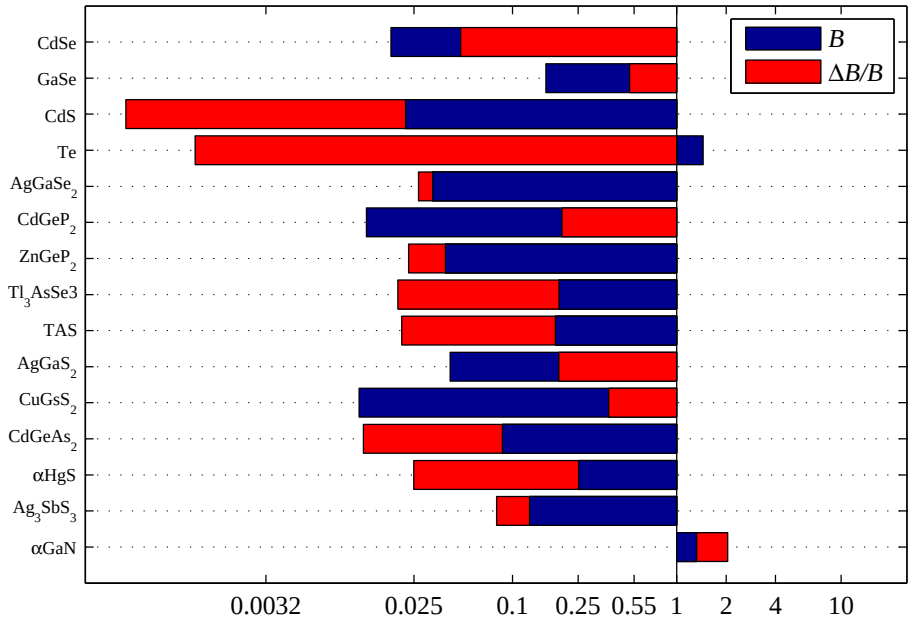

Figure 2.43: Values for birefringence and $\Delta B/B$ for materials transparent in the range 8 μm to 11 μm. Both quantities are unitless.

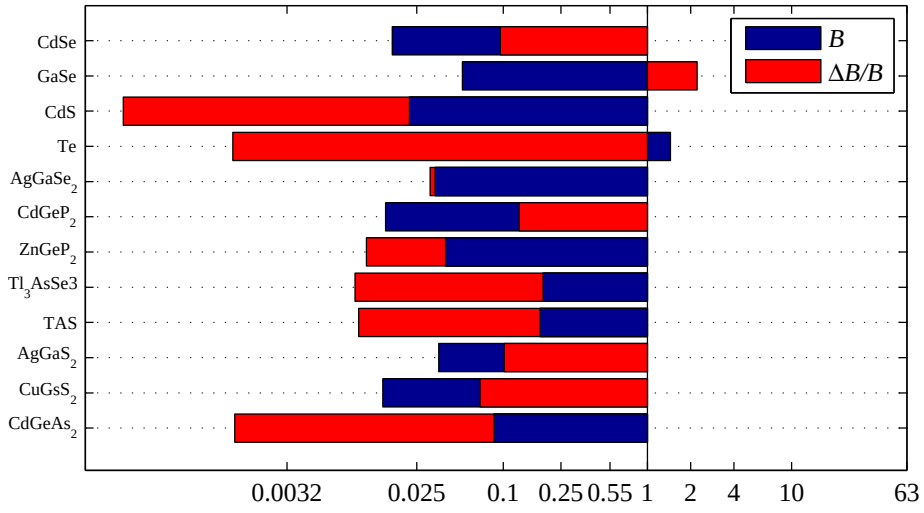

Figure 2.44: Values for birefringence and $\Delta B/B$ for materials transparent in the range 11 μm to 15 μm. Both quantities are unitless.

setting the result equal to 0.

$$\begin{aligned}\frac{dR}{dn_m} &= \frac{d}{dn_m}\left(\frac{n_1 - n_m}{n_1 + n_m}\right)\left(\frac{n_m - n_2}{n_m + n_2}\right) \\ &= \frac{d}{dn_m}\frac{n_1 n_m - n_m^2 + n_2 n_m - n_1 n_2}{n_1 n_m + n_m^2 + n_2 n_m + n_1 n_2} = 0.\end{aligned}$$

The physically possible solution to this is $n_m = \sqrt{n_1 n_2}$.

Another option for reducing reflection losses within the retarder is to optically contact surfaces [99]. This is attractive for the mid- and far-infrared as index matching fluids are not available in this region. The YVO_4 Wollaston prisms used in the MWIR IRIS system are optically contacted.

2.6 Practical issues

The previous sections flagged some of the potential practical problems encountered when designing and operating IRIS systems. These are now examined in more detail in this section.

2.6.1 Potential aberrations

Monochromatic aberrations

A thorough discussion of the aberrations introduced by the focusing and collimating components falls outside the scope of this work. Needless to say, the major aberrations, such as the five third-order aberrations, should be minimised during the optical design. Optical design techniques are reasonably mature and there are many tools available to aid with the reduction of image aberrations, not to mention a wide variety of commercially available lenses and mirrors, optimised to varying degrees and for various tasks. One additional complication for IRIS is that the optimisation should ideally be performed for all J sub-images simultaneously.

The surfaces in the demultiplexer can also distort the collimated wave fronts. Consider a plane wave front which is described by a surface w, with a standard deviation σ_w, propagating in a medium with refractive index n_1. If this wavefront is incident upon a surface R, described by a standard deviation σ_R, which is the boundary to a second medium with refractive index n_2, then the aberrations of the wave-front in the second medium are described by

$$\sigma_X^2 = \left(\frac{n_1}{n_2}\right)^2 \sigma_w^2 + \left(1 - \frac{n_1}{n_2}\right)^2 \sigma_R^2 + 2\mathrm{cov}\left(\left(\frac{n_1}{n_2}\right) w, \left(1 - \frac{n_1}{n_2}\right) R\right). \tag{2.21}$$

This equation allows the aberrations through the demultiplexer to be tracked and allows

one to calculate the tolerances for the surface flatness of the Wollaston prisms and retarders. The Strehl ratio is commonly used to relate the wavefront aberrations to imaging performance, and is defined as the ratio of aberrated to unaberrated intensities at the Gaussian image point. This is equal to the normalised zero-order component of the Fourier transform of the aberrated aperture function. Assuming the field amplitude does not vary over the wavefront, the aperture function is described purely by the phase variation. By expanding the zero-order component of the Fourier transform as Taylor series, the Strehl ratio can be shown to be, for small aberrations, $1 - (2\pi\sigma)^2$. A Strehl ratio of 0.8 is commonly considered acceptable; this is equivalent to $\sigma_X \approx \lambda/14$. For calcite prisms and quartz retarders, with uncorrelated surface flatnesses of $\sigma_R = \lambda/28$ and no index-matching, the $\lambda/14$ limit is breached with a $M = 3$ demultiplexer. With the same value for σ_R, and with index-matching chosen to minimise Fresnel losses, the limit is not breached for upwards of $M = 200$, a far larger number of stages than is feasible. In the context of wavefront aberrations, surface flatnesses of $\sigma_R = \lambda/12$ are likely to be sufficient for the Wollaston prisms and retarders in any index-matched demultiplexer with $M < 10$.

Chromatic aberrations

In the discussion of image aberrations so far, chromatic effects have not been mentioned. Due to the wavelength-dependence of refractive indices, the imaging properties of a lens system are also wavelength-dependent. This complicates the minimisation of aberrations and can result in lateral chromatic aberration (a difference in the image magnification between wavelengths) and longitudinal chromatic aberration (a shift in the focal plane between wavelengths). Commercially available lenses are usually sufficiently well corrected for lateral chromatic aberrations but less so for longitudinal aberrations. As IRIS systems require three lenses, there is more sensitivity to chromatic aberrations.

Longitudinal chromatic aberrations in an otherwise well-corrected imaging system will result in varying degrees of defocus relative to a properly focused reference wavelength. If such defocus cannot be corrected in the imaging components then it may instead be possible to use an array of J glass elements of varying thicknesses at, or close to, the detector plane, one element in front of each sub-image. For this to be successful the secondary peaks of the filter functions must be minimised or be made close to the main peak in spectral space. Figure 2.45 illustrates this principle.

For the paraxial case, the thickness d, of glass with refractive index n, required to correct a defocus of Δz, in a beam of f-number F, can be found from simple trigonometry (Figure 2.46) and is given by

$$d = \frac{\Delta z}{(1 - 2F\tan\phi)}, \tag{2.22}$$

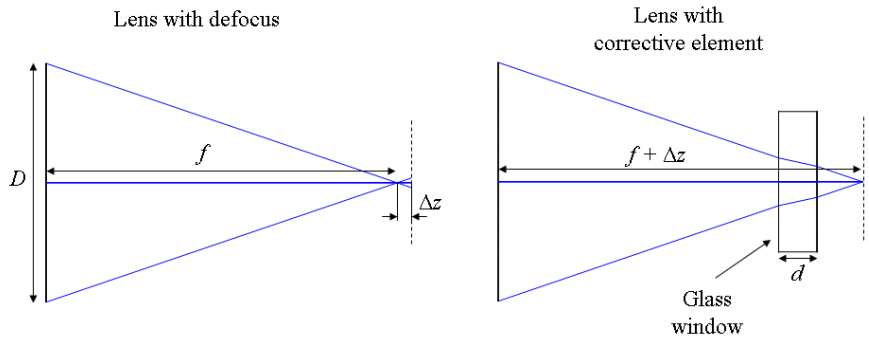

Figure 2.45: Ray trace showing the use a window of glass to correct defocus.

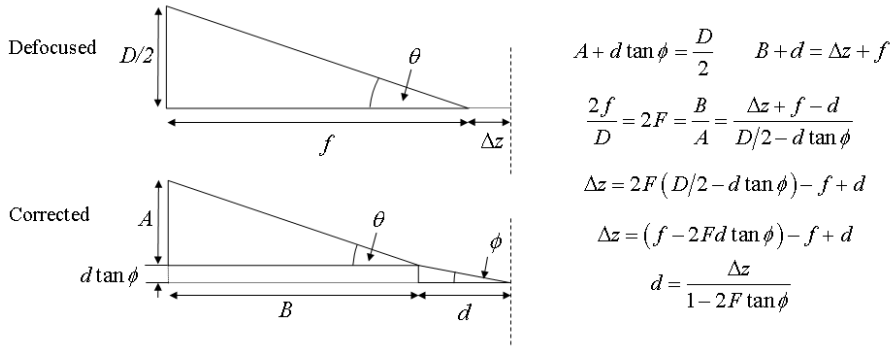

Figure 2.46: Calculation of required glass thickness to correct a defocus of Δz.

where ϕ can be found using Snell's law

$$\phi = \sin^{-1}\left(\frac{\sin\left(\tan^{-1}\left(1/2F\right)\right)}{n}\right). \tag{2.23}$$

A flat piece of glass in a convergent beam will introduce negative longitudinal-spherical aberration (Figure 2.47), as demonstrated by the dependence of d on F in Equation 2.22. For rays close to the optical axis, the small-angle approximation for Equation 2.22 gives the shift in focus due to a thickness d to be $\Delta z' = d\,(1 - 1/n)$. The extent of the longitudinal-spherical aberration is therefore given by $\Delta z - \Delta z' = \Delta z - d\,(1 - 1/n)$. The circle of least confusion is one quarter of this distance away from the marginal focus, giving the minimum spot size to be $(\Delta z - d\,(1 - 1/n))/4F$. For example, for a $F = 2$ system with a defocus of $\Delta z = 1.5$ mm, Equation 2.22 gives a required thickness of 4.35 mm of $n = 1.5$ glass to correct the defocus. The circle of least confusion in this case has diameter 6.3 μm.

As always, real lenses will depart from paraxial behavior and an accurate value for the thickness of glass required to correct any longitudinal aberration is best determined using ray tracing, or empirically if the lens design is unknown.

The corrective glass elements will also cause vignetting for fields at the edge of each

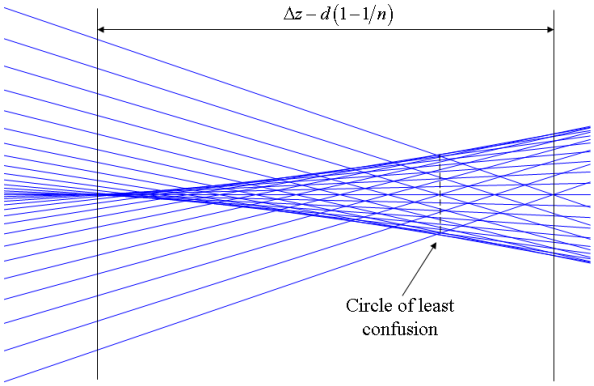

Figure 2.47: Spherical aberration introduced by a glass window in a convergent beam.

sub-image, the distance a from the edge of the sub-image which is affected by vignetting, depends on the distance d, of the front surface of the corrective element from the detector and F, the f-number of the beam. If the lens can be considered telecentric then

$$a = \frac{d}{2F} \tag{2.24}$$

If a is larger than the smallest extent of the sub-image then vignetting is present across the entire sub-image. To ensure that the vignetted light does not contribute to stray light, the edges of the corrective elements need to absorb light.

Chromatic aberration is not an issue for mirrors and these can be used for collimating and imaging. The presence of the demultiplexer requires that off-axis or annular mirrors are used. The feasibility of off-axis parabolic mirrors was investigated using ray-tracing software: chromatic aberrations are removed as expected; however only relatively simple designs using a small number of mirrors were considered and monochromatic aberration control is problematic. A full attempt at solutions using more complex optical designs is beyond the scope of this work; however, it is likely that a fully achromatic optical design can be found using reflective components.

A form of chromatic aberration specific to IRIS is caused by the dispersion in the birefringence of the prism material. As mentioned previously, the splitting angle of a Wollaston prism is approximately $\psi = 2B(k)\tan\theta$. Due to the dependence of B on wavenumber, the splitting angle varies with photon energy, causing a point source to be dispersed over a line of pixels in each sub-image. The intensity at each wavelength is modulated by the transmission of the band, thus minimising the secondary peaks in the transmission will help reduce the extent of the effects of dispersion. For a system with K Wollaston prisms splitting in the x-plane, the number of pixels of smear for a sub-image q places from the centre of the detector in the x-direction, is

$$p = R_x \Delta B\,(2q-1)/\left(2^{K+1}\bar{B}\right), \tag{2.25}$$

where R_x is the resolution of the detector in the x-direction.

There will always be some secondary peaks in the transmission functions even after optimisation of the retarder thicknesses. There are a number of options for dealing with the chromatic aberration due to birefringent dispersion: the secondary peaks can be further reduced by using an array of M interference filters prior to the image plane to reject radiation outside the main peak of each band, or one can attempt to find a design for a polarising beamsplitter with minimal dispersion; a final option is to use image processing algorithms to attempt to remove the effects of the dispersion from the the image. These techniques will be investigated in Chapter 4.

2.6.2 Distortion and co-registration

Registration of spectral-imaging data is a key pre-processing step. The required accuracy of registration for a specific task can be determined via analysis of the propagation of measurement error to the retrieved variables. For instance, for the Landsat Thematic Mapper mentioned in Section 1.3.1, the required accuracy is less than a fifth of a pixel registration error for a 10% error rate in a Mahalanobis-distance-based classification algorithm [100].

The effect of registration error on the measurement error can be estimated by simulating the sampling of the spatial information and investigating the effect of misregistration. The image formed at the detector can be regarded as the convolution of the two-dimensional projection of the scene, and the instrument point-spread function. This image is then sampled by the detector array.

It will be assumed that the PSF can be approximated by an Airy disk, and also that each pixel is square with sensitivity constant across its entire area. Critical sampling of the PSF occurs approximately when the first zeros of the Airy disk are separated by twice the pixel pitch. In this analysis the critically sampled case will be considered, as well as an under sampled case with the zeros separated by a single pixel width, and an over sampled case with a separation of four-pixels (Figure 2.48). These PSF are convolved with the source and the resulting image is binned into pixels. Figure 2.49 shows the binned values for a point source against a black background.

By applying an offset in the position of the source, the effect of misregistration on the measurement value can be found. This calculation is performed for 5 different square sources with sizes 0.05, 0.25, 0.5, 1 and 2 pixels. The sources are initially concentric with the pixel and a maximum offset of 1 pixel is applied, the results are shown in Figures 2.50 to 2.52. For a point source in the centre of the pixel, to achieve less than 1% error, the registration accuracy should be better than a 16th of a pixel in the under-sampled case, and less than 11th of a pixel for the critically-sampled case. As can be seen from these plots, the sensitivity to registration error increases if the source is not located at the centre of the pixel. If the source is located towards the edge of the pixel, for 1%

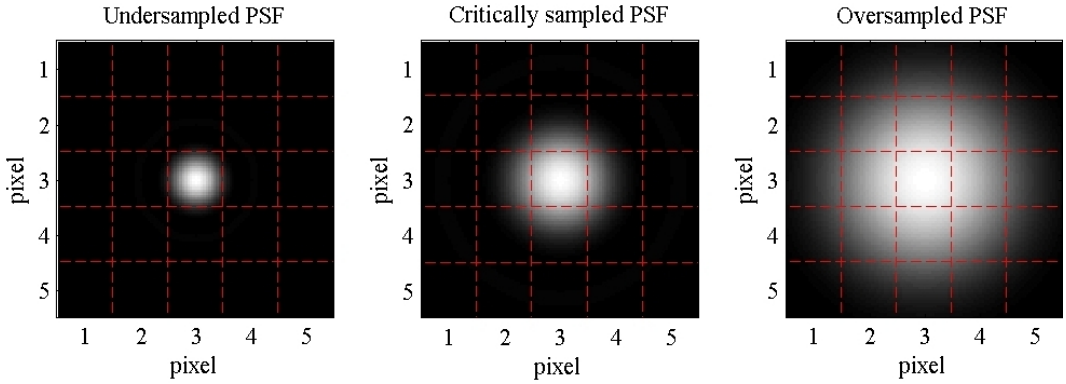

Figure 2.48: PSF for registration error analysis

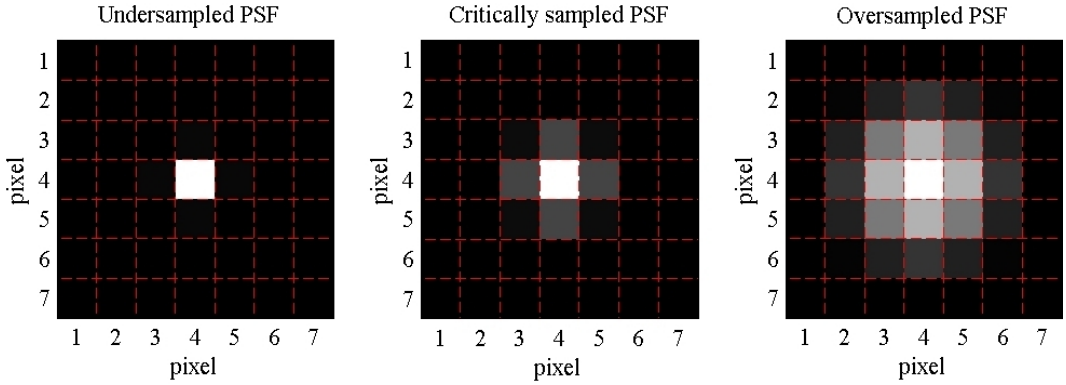

Figure 2.49: PSF for registration error analysis, binned into pixels.

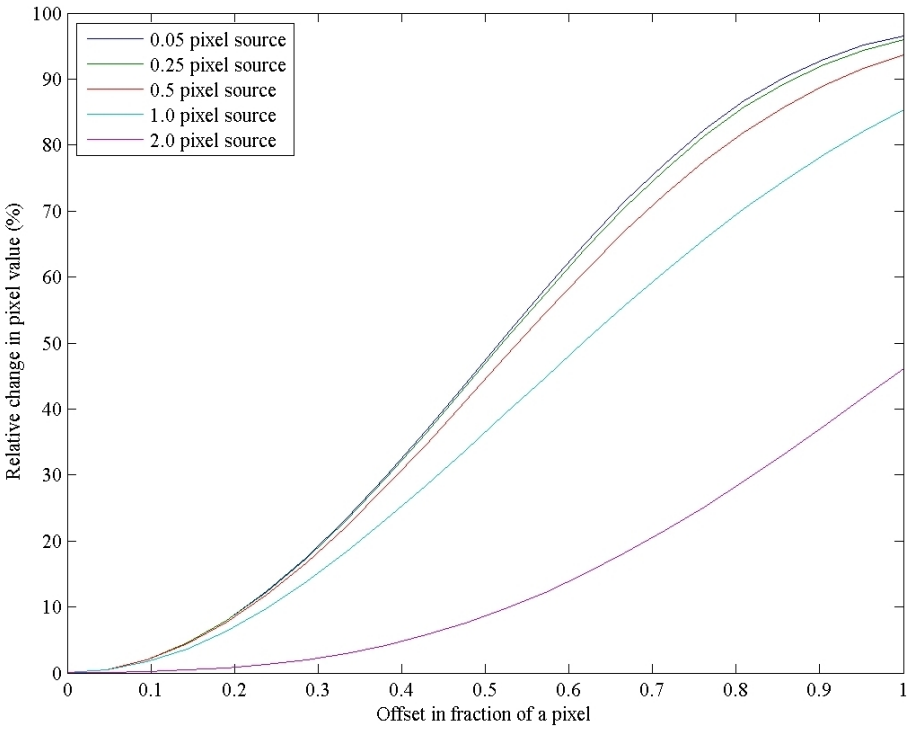

Figure 2.50: Change in pixel value as a function of registration error for the under-sampled case.

error, 100th-of-a-pixel registration accuracy is required in the critically-sampled case, and 200th-of-a-pixel accuracy in the under-sampled case. The example of a point source against a black background is the most sensitive to misregistration so these estimates of required accuracy are upper limits.

The registration of IRIS sub-images can be regarded as a two-step process: the sub-images are cropped from the full frame and then a transformation applied to each sub-image to bring the entire set into a common coordinate system. The type of transformation required can be found by examining the imaging process.

The wavefronts from an individual spatial element in the IRIS system follow a single, common path until the demultiplexer is reached. If the demultiplexer and imaging lens do not introduce any aberrations in the final image then the transformations required for registration are simple translations. In this case the required amount of translation can be found by calculating the normalised cross-correlation or some similar metric. Assuming that the normalised cross-correlation is a smooth function, its maximum can be determined with sub-pixel accuracy. Formally, sinc interpolation is optimal; however, due to the finite bandwidth of images, spline interpolation often performs better [101] in practice.

If the surface-flatness of the constituent components is adequate, then the demultiplexer will introduce negligible distortion. The imaging lens is likely to introduce non-trivial aberrations, most significantly geometric distortion [102], meaning that a polynomial transformation is necessary for registration. A two-dimensional polynomial

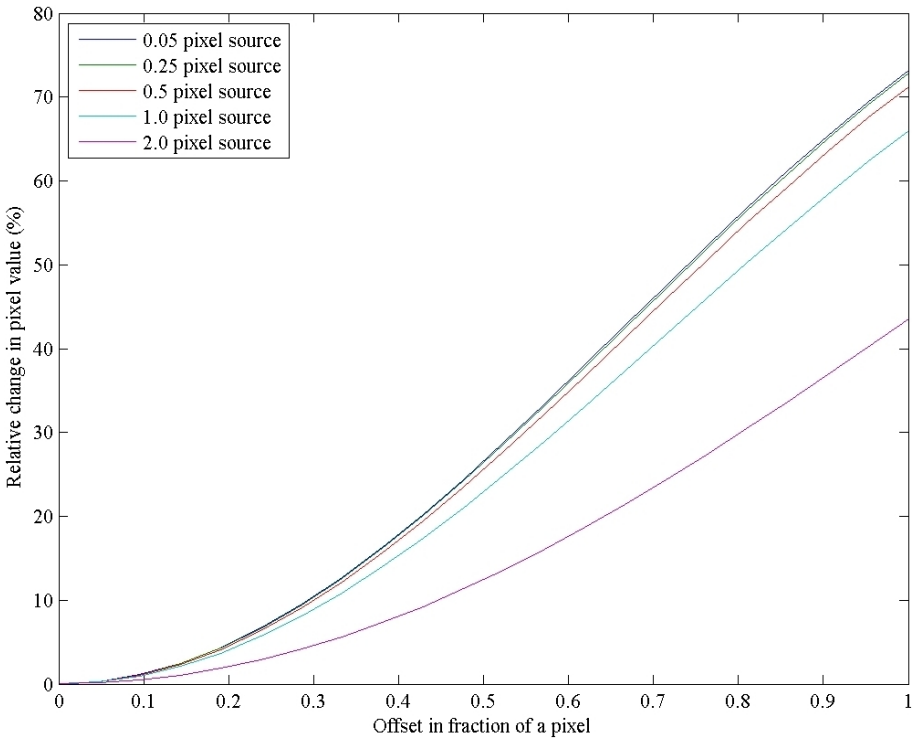

Figure 2.51: Change in pixel value as a function of registration error for critically-sampled case.

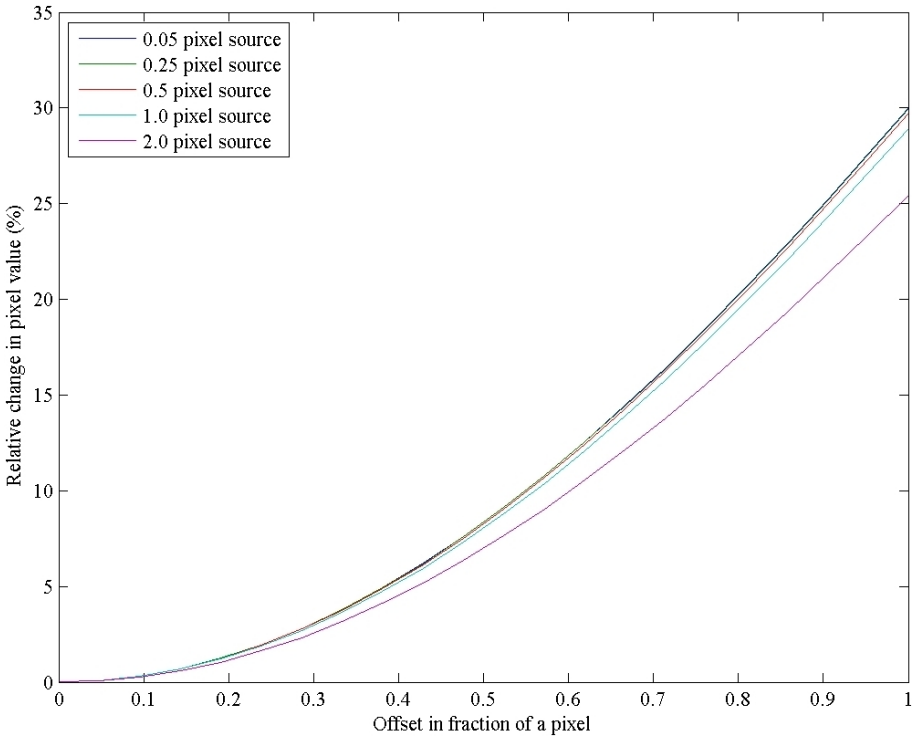

Figure 2.52: Change in pixel value as a function of registration error for over-sampled case.

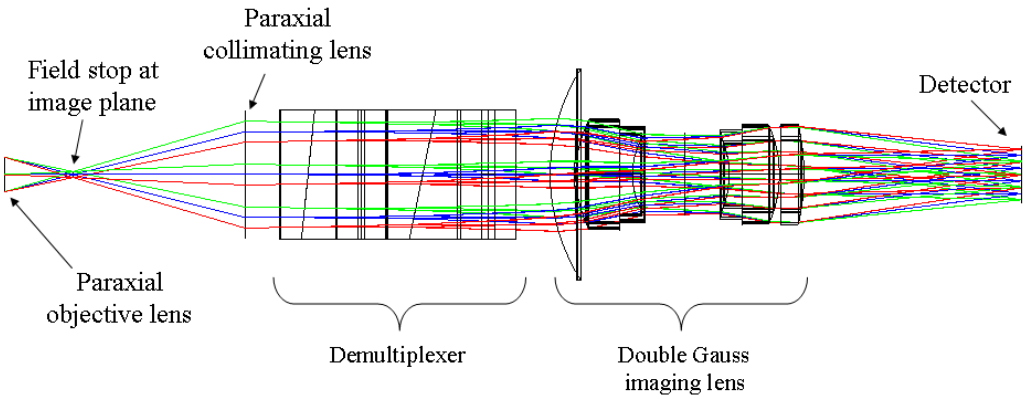

Figure 2.53: Zemax ray trace of the 16 channel IRIS system with a Double Gauss imaging lens.

transformation of degree N, from coordinate system (u, v) to (p, q), can be written as

$$\begin{aligned} p &= \sum_{s=0}^{N} \sum_{t=0}^{N-s} a_{st} u^{N-s-t} v^t \\ q &= \sum_{s=0}^{N} \sum_{t=0}^{N-s} b_{st} u^{N-s-t} v^t. \end{aligned} \tag{2.26}$$

In order to estimate the required degree of polynomial a ray trace is carried out using Zemax for the 16-channel IRIS system as shown in Figure 2.53. The sub-images are 3.75 mm $\times$ 1.875 mm and are arranged in a 4×4 matrix. This model uses paraxial lenses for the objective and collimating lenses, while a 100 mm double-Gauss lens is used to simulate a realistic imaging lens. There is no distortion introduced by the demultiplexer.

By tracing rays from a regular grid of point sources, the corresponding coordinates in each sub-image can be found. Figure 2.54 illustrates the distortion present in four of the images; the average displacement of the points from the grid is approximately 1 pixel and the standard deviation of these displacements is around 0.7 pixels. For the ith source, the polynomial transform from image coordinates (u_i, v_i) to object space coordinates (p_i, q_i) can be written as the inner products $p_i = \mathbf{w}_i \cdot \mathbf{a}$ and $q_i = \mathbf{w}_i \cdot \mathbf{b}$, where

$$\begin{aligned} \mathbf{w}_i &= \begin{bmatrix} 1 & u_i & v_i & u_i^2 & v_i^2 & v_i u_i & \dots \end{bmatrix}^T \\ \mathbf{a} &= \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & \dots \end{bmatrix}^T \\ \mathbf{b} &= \begin{bmatrix} b_0 & b_1 & b_2 & b_3 & b_4 & b_5 & \dots \end{bmatrix}^T. \end{aligned} \tag{2.27}$$

For the entire set of sources, the transformation can be written as the linear equations

$$\begin{aligned} \mathbf{p} &= \mathbf{W}\mathbf{a} \\ \mathbf{q} &= \mathbf{W}\mathbf{b} \end{aligned} \tag{2.28}$$

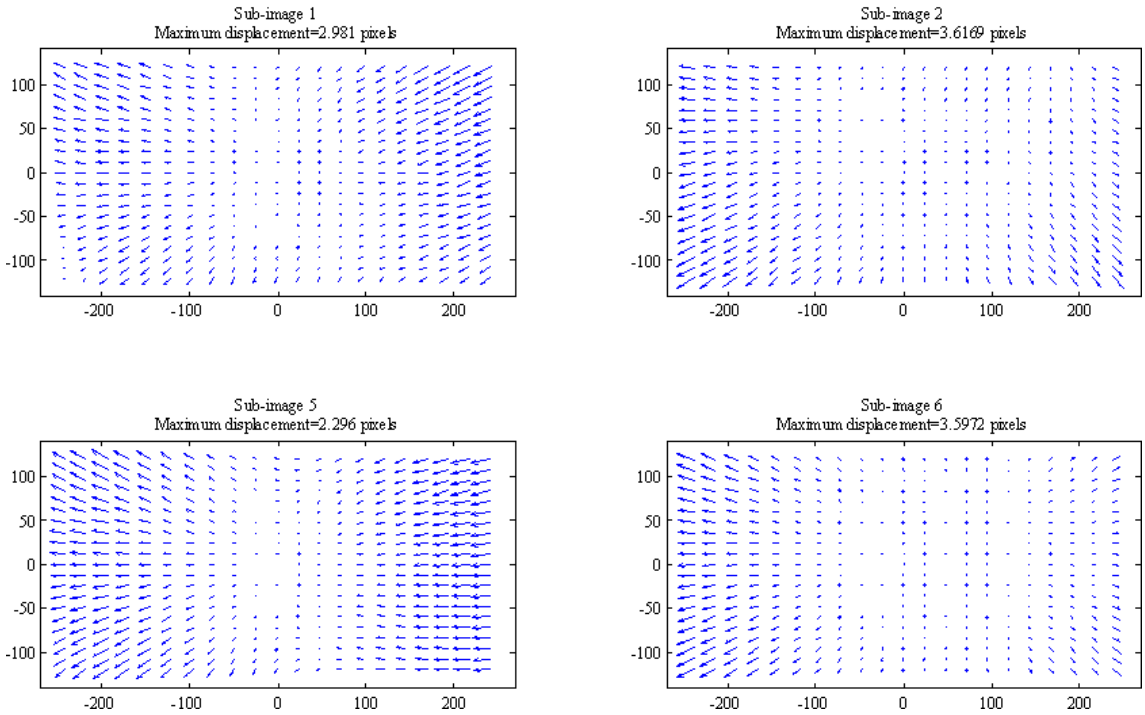

Figure 2.54: Distortion present in each sub-image. The position of these sub-images at the detector is shown in Figure 2.55.

where $\mathbf{p}$ and $\mathbf{q}$ are vectors of the object space coordinates and $\mathbf{W}$ is a matrix whose ith row is the polynomial vector $\mathbf{w}_i$. $\mathbf{a}$ and $\mathbf{b}$ are the vectors of $N\,(N+1)/2$ coefficients for a polynomial of order N. Provided that the number of point sources used is greater than or equal to the number of coefficients, the coefficients $\mathbf{a}$ and $\mathbf{b}$ can be found exactly, or in the least squares sense, by calculating the inverse, or pseudo-inverse, of $\mathbf{W}$.

Figure 2.55 shows the accuracy of the registration process for the 16-channel system as a function of the order of polynomial transformation used. The order required for a given registration accuracy increases with distance from the detector, but using 4th-order polynomials will ensure 10th-of-a-pixel accuracy in all cases, while 5th-order polynomials will give 100th-of-a-pixel accuracy.

Assuming that the optical performance of the system does not change over time, registration is only a calibration procedure. This means that once the sub-images from a single frame have been registered, the same transformations can be applied to subsequent frames to register the entire sequence of data.

Another benefit of the registration process is that, by choosing to register to the object-space coordinates, geometric distortion caused by all components can be removed. The transformation used reveals the type and magnitude of distortion present in each sub-image.

A discussion of the registration of real data can be found in Chapter 3.

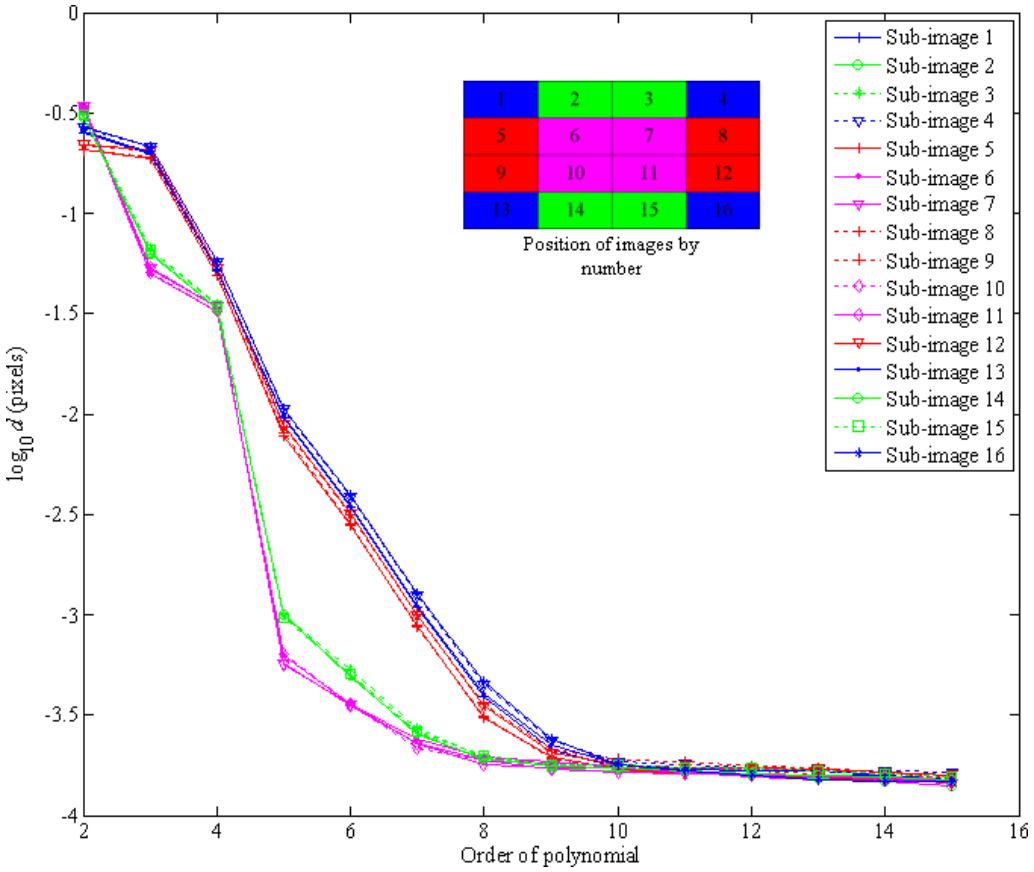

Figure 2.55: Registration error as a function of polynomial order.

2.7 Conclusions

This chapter has explained the IRIS concept and the design principles, which were then demonstrated in the first order design of the 8-channel retinal IRIS system. Potential issues with stray light have been discussed and the specific challenges in implementing IRIS systems in the infrared have been illustrated in the optical redesign of the MWIR IRIS prototype system. The IRIS system optical parameters resulting in the optimal throughput for each field of view have been calculated, and a survey of suitable birefringent materials presented. A discussion of potential image aberrations in IRIS systems was also given as well as a description of the issues surrounding registration of the IRIS images.

The next chapter shows the use of the design principles given in this chapter in the design of the 16-channel broadband IRIS system. This initial design is then modelled using polarisation ray tracing which is used to refine the design and characterise the system. The manufactured 16-channel broadband system is also characterised through measurements which are compared with the simulated data.

Chapter 3

IRIS Appraisal and modelling

Summary: This chapter details the design and construction of the 16-channel broadband system. Interferometric ray racing is used to accurately model the performance of the system, and measurements on the manufactured system are compared with the output of the ray tracing program. The chapter concludes by demonstrating the registration of IRIS images for the retinal-imaging IRIS system.

3.1 Introduction

The previous chapter described the first order design of IRIS systems and touched upon the effects of the varying angle of incidence and the effects of dispersion on the performance of IRIS systems. To model the behaviour of IRIS more accurately, polarisation ray tracing is required. The outputs of the ray trace can then be compared with measurements to validate the model. The measurements also serve to characterise the instrument properly for future use and are an essential part of the construction process.

3.2 Ray tracing through birefringent media

An explanation of the propagation of radiation through birefringent media is given in Appendix A. To summarise: for each direction of propagation within a birefringent material there can exist two mutually orthogonal waves. For one wave, known as the ordinary wave, the wave vector is perpendicular to an axis of rotational symmetry within the crystal known as the optic axis; this wave moves with a speed c/n_o, where n_o is the ordinary refractive index of the material. The rays corresponding to this field are parallel with the wave vector and propagate at the same speed. The second wave is the extraordinary wave and is coplanar with the optic axis. The speed of this wave is $c/n(\phi)$ where the refractive index $n(\phi)$ depends on the angle ϕ between the optic axis and the wave vector. $n(\phi)$ reaches an extremum of $n(\phi) = n_e$ at $\phi = 0°$ and n_e is known as the extraordinary refractive index of the material. The rays corresponding to

the extraordinary wave are not in general parallel to the extraordinary wave vector, but are coplanar with the incident and transmitted wave vectors.

As the refractive indices are different for the ordinary and extraordinary waves corresponding to a given wave direction, these waves must also correspond to different incident wave directions. Conversely, a single wave incident upon a birefringent surface is split into two orthogonally polarised waves with different angles of refraction. The angle of refraction θ_o for the ordinary wave, for a given angle of incidence θ_i can be found from Snell's law:

$$n_o \sin\theta_o = n_i \sin\theta_i, \tag{3.1}$$

where n_i is the refractive index of the medium in which the incident wave propagates. Likewise the angle of refraction of the extraordinary wave-vector can be found from

$$n\left(\phi\right) \sin\theta_e = n_i \sin\theta_i, \tag{3.2}$$

where the extraordinary refractive wave index is as given by Equation A.3.

Equation A.3 together with Equation 3.2 result in a quartic equation in ϕ with one physically possible root corresponding to the correct value of ϕ. A closed-form solution can be found for this Equation [103] or it can be solved using an iterative approach, such as Newton's method.

The optical path length through the medium for the ordinary ray is given by the product of the ray physical path length and the ordinary refractive-index, n_o. The optical path length for the extraordinary ray is given by the product of the ray physical path length and ray refractive index, which is given by $n_r = n_e\left(\phi\right)\cos\rho$ [104], where ρ is the angle between the extraordinary ray and wave vectors and is given by [103]

$$\tan\rho = \frac{\left({n_e}^2 - {n_o}^2\right)\tan\left(\phi\right)}{{n_e}^2 + {n_o}^2\tan^2\left(\phi\right)}. \tag{3.3}$$

As with isotropic materials, the ordinary and extraordinary refractive indices depend on wavelength. Also, in general $dn_e/d\lambda \neq dn_o/d\lambda$, which leads to dispersion in the birefringence.

Using the above relations, a polarisation ray-tracing program for the simulation of IRIS systems has been implemented in Matlab.

3.3 16-channel broadband visible IRIS system

The ray-tracing program implemented in Matlab has been used to aid the design of the 16-channel broadband IRIS system introduced in Section 2.1. The resulting design has also been manufactured, and measurements made with this system are compared to results from the ray trace.

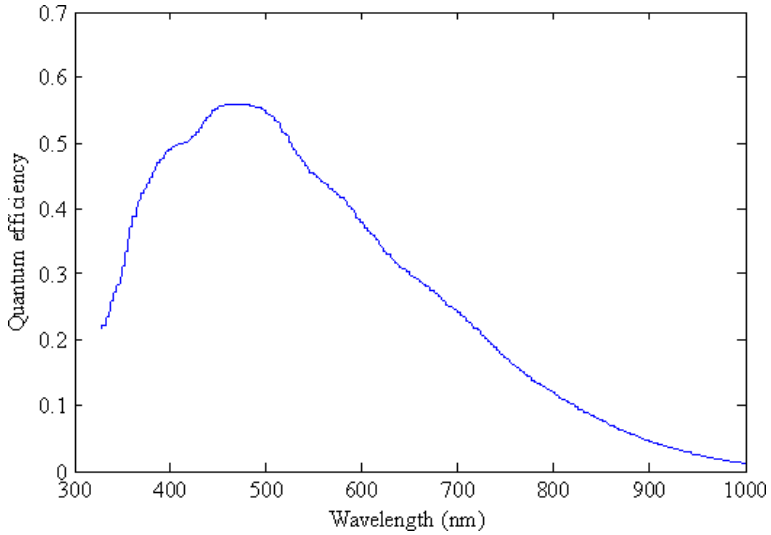

Figure 3.1: Quantum efficiency of the Retiga 4000R CCD detector.

The requirement for the 16-channel system is that the field of view is as large as possible while remaining fast enough to record at video rates in a relatively high light-level environment.

The 16-channel system is also designed to be used with an array of interference filters at the detector to reduce the chromatic aberrations caused by spectral dispersion in the splitting angle of the prisms. This dispersion causes each point in the scene to be imaged to a line of pixels at the detector, the extent of which increases with bandwidth. Each filter in the filter array covers a single sub-image which blocks radiation from the secondary peaks of the IRIS transmission functions, this in turn reduces the extent of the smear. The matrix filter will not be used here but is described in more detail in Section 4.4.

This system is originally designed for use with an *AVT* Pike-210 B/C camera; the detector array resolution is 1920×1080 and the pixels are square with length of side 7.4 μm. However, the camera used to take the measurements described in this chapter is a Retiga 4000R produced by *QImaging*®. This uses a cooled CCD detector array with 2048×2048 pixels. The pixels are square with length of side 7.4 μm. This means that around 51% of the pixels on the Retiga 4000R remain unused. The detector is sensitive to wavelengths from 350 nm to around 1 μm. The well capacity for each pixel is 40000e^-, digitised at up to 12 bits. The read noise is 12e^- which is around 1 grey level. The quantum efficiency of the detector array is plotted in Figure 3.1 [1].

[1]The quantum-efficiency data has been obtained from the manufacturer.

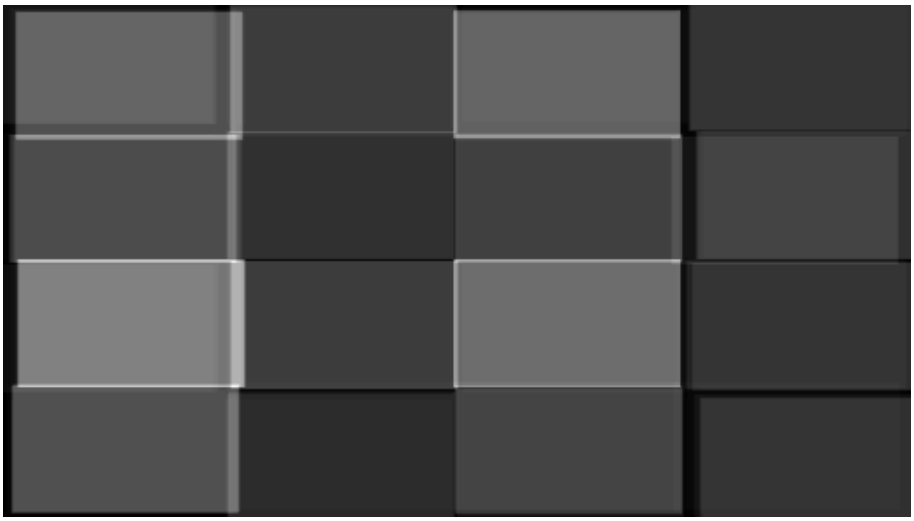

Figure 3.2: Simulated image for 16-channel IRIS broadband system showing overlap of the sub-images due to dispersion

3.4 Optical design relations

For the 16 band system, the objective lens is a $f/2.8$, 3.5 mm video lens, which is image-space telecentric. The chosen imaging lens is a $f/2$, 100 mm SLR lens manufactured by Canon. The AVT Pike detector dimensions are $a_x = 14.2$ mm by $a_y = 8.0$ mm. From Equation 2.11, the maximum prism splitting angles required are $[\phi_{max}]_x = 4.07°$ and $[\phi_{max}]_y = 2.29°$. The remaining two splitting angles are $\phi_x = 2.03°$ and $\phi_y = 1.14°$. The chosen prism material is calcite, due to the large sizes in which these prisms can be obtained. The mean refractive indices of calcite over the wavelength range 420-830 nm are $\bar{n_o} = 1.658$ and $\bar{n_o} = 1.485$; the mean birefringence is $\bar{B} = 0.171$. The required wedge angles for calcite prisms are $\theta_{x1} = 11.55°$ and $\theta_{x2} = 5.82°$, and $\theta_{y1} = 6.56°$ and $\theta_{y2} = 3.28°$.

A 50 mm lens is used as the collimating lens. As the magnification between the field stop and the detector array is 2, the field stop dimensions are (from Equation 2.12) $S_x = 1.85$ mm and $S_y = 0.92$ mm. The system f-number is 5.6.

Figure 3.2 shows an the image at the detector simulated using the Matlab interferometric ray trace, for a spatially-uniform white source for these parameters. As can be seen from this image, there is overlap between the sub-images, which is due to dispersion.

The design was refined using the interferometric ray tracing software, reducing the size of the field stop to $S_x = 1.54$mm and $S_y = 0.76$mm while increasing the prism splitting angles to ensure there is no overlap of the sub-images. The fields of view for a 3.5 mm objective lens with this size of field stop are (from Equation 2.18) $\psi_x = 14.8°$ and $\psi_y = 6.2°$. Figure 3.3 shows two orthographic projections of the ray trace through this 16-channel system.

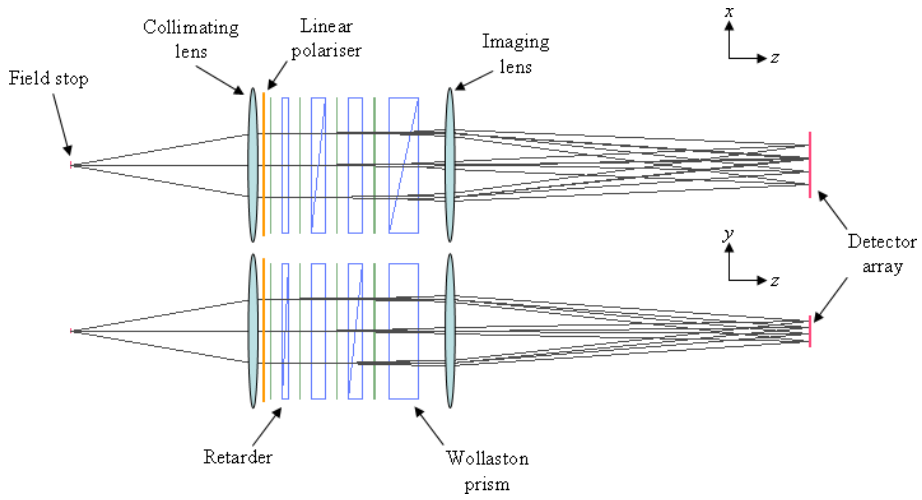

Figure 3.3: Orthographic projections of a ray trace of the 16-channel system.

The demultiplexer was constructed using quartz retarders and calcite Wollaston prisms, both supplied by *Halbo Optics*. The average refractive indices of quartz are $n_o = 1.54$ and $n_e = 1.55$; from Section 2.5, the optimal refractive index to match calcite and quartz is between 1.5 and 1.6; a *Norland* index-matching fluid with a refractive index of 1.52 is used here. The first and last surfaces of the demultiplexer are coated with a single quarter-wavelength (at 530nm) layer of magnesium fluoride, reducing Fresnel losses from each of these surfaces to 2% or less across the entire spectrum ($< 99.5\%$ at 530nm).

Figures 3.4 and 3.5 show the simulated and measured images for the manufactured system at 520 nm when the source is spatially uniform across the field of view. The source for the measured image is an integrating sphere, with the input illumination filtered using a interference filter. The measured image shows that for the real system there is a rotation of the field stop of approximately 36 arcminutes clockwise compared to the modelled system, this is due to misalignment during construction; however, this rotation will not impact negatively on the system performance. Otherwise, the images compare well.

3.5 Transmission

The four quartz retarders in the 16-channel system have thicknesses of 79.4 μm, 94.2 μm, 189.7 μm, and 415.8 μm. These were chosen in order to maximise total throughput when used with the array of interference filters at the detector, as will be described in Section 4.4. As mentioned the purpose of the filter array is to block radiation from the secondary peaks in the IRIS transmission functions to reduce the effects of dispersion. The transmission measurements have been performed without the filter array in place.

Using the ray trace, the transmissions at the centre of each sub-image have been cal-

Figure 3.4: Simulated 16-channel IRIS image at 520 nm with no overlap of the sub-images

Figure 3.5: Measured image, recorded at 520 nm with the 16-channel IRIS system. The image has been dark calibrated.

culated. For the real system the transmissions of the demultiplexer have been measured using an *Ocean Optics* USB4000-VIS-NIR spectrometer for which the input is a single strand of optical fibre whose head can be scanned over the image plane of the IRIS. Light accepted by the fibre feeds a diffraction grating, which disperses the radiation over an uncooled, linear CCD array. To measure the transmission a reference spectrum is first measured at the centre of the single sub-image produced by the system when the demultiplexer is removed from the optical path. The spectrum at the centre of each of the sub-images is then measured with the demultiplexer in the path, and each spectrum is divided by the reference spectrum to produce a transmission. The spectral resolution of the USB4000-VIS-NIR is 1.5 nm, spline interpolation is used to increase the resolution to 0.1 nm. The measured and simulated transmissions are plotted in Figure 3.6 along with the transmissions as predicted using the approximate expressions for the transmission functions from Section 2.3. The measured, simulated, and approximate transmissions are all in good agreement.

The transmissions around the peaks of each band are shown in Figure 3.7. The second measured transmission in the top row appears to be the least similar to the predicted transmission; this is due to poor signal to noise in the spectrometer measurements above 800 nm. The maximum predicted difference between the peak wavelength, as calculated using the approximate transmission functions, and the simulated transmissions is 0.4 ± 0.05 nm and occurs in the top-left sub-image; the measured difference is 0.6 ± 0.05 nm. The total transmission of the demultiplexer is shown in Figure 3.8. The average transmission across the band is 96.8%.

As the beams corresponding to different field angles follow different optical paths through the demultiplexer, there is a small change in transmission across the field of view. For example, the predicted shift in peak wavelength between the inner and outer corner of the bottom-right sub-image is 0.2 ± 0.05 nm (Figure 3.9), the measured shift (Figure 3.10) is 0.2 ± 0.05 nm.

3.6 Image quality

To assess the imaging performance of the IRIS system a spoke target (Figure 3.11) has been imaged. This target is challenging for an imaging system to reproduce accurately as it contains a large number of spatial frequencies, including very high frequencies. Visual inspection of the resulting images can be used to assess image quality.

Figure 3.12 shows the IRIS image of the spoke target when illuminated by a broadband tungsten source. It should be noticed that there is dispersion due to the Wollaston prisms; this is more evident in the logarithm of this image which is shown in Figure 3.13. The dispersion will be characterised in the next section.

To assess the image quality in the absence of dispersion, the spoke target can be

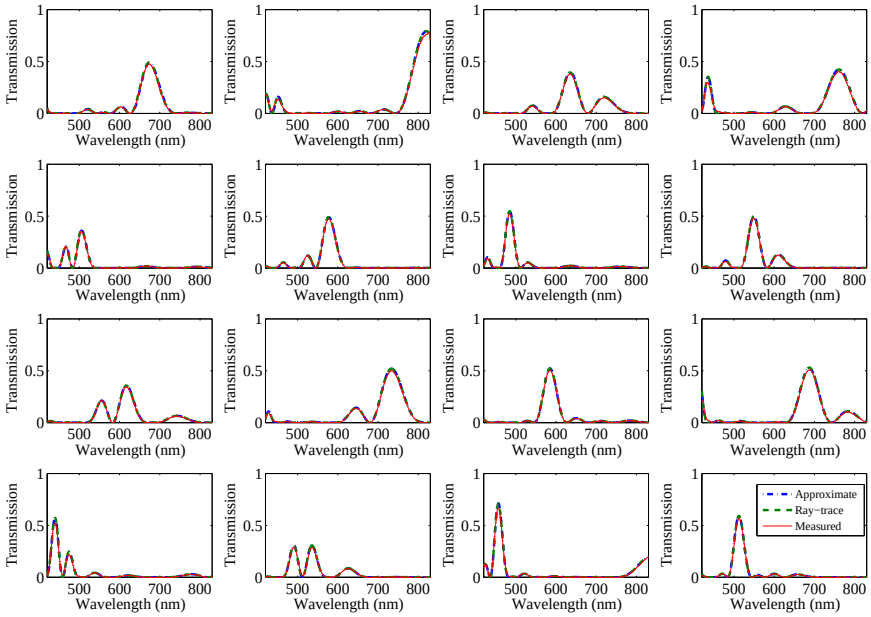

Figure 3.6: Measured, simulated and approximate transmissions for a 16-channel IRIS system.

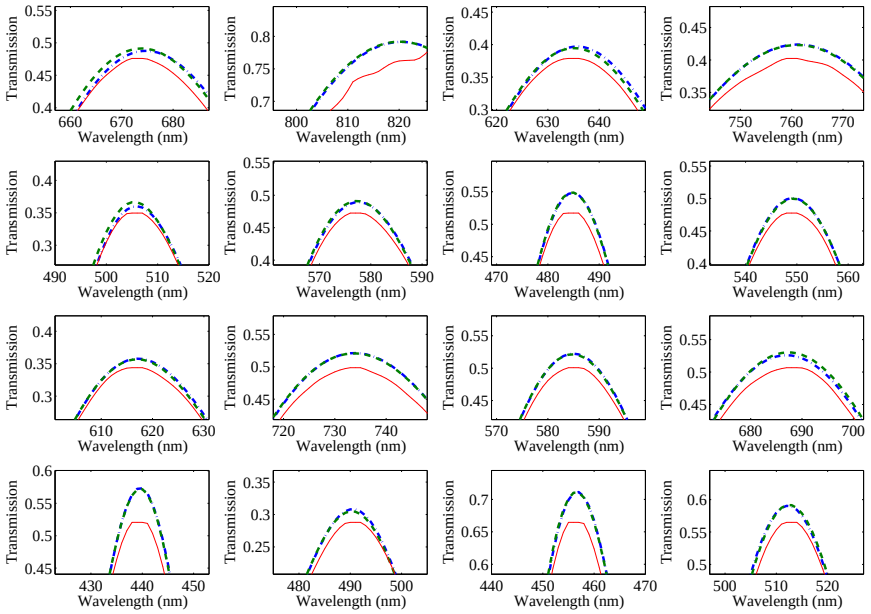

Figure 3.7: Measured, simulated and approximate transmissions around the peak of the bands.

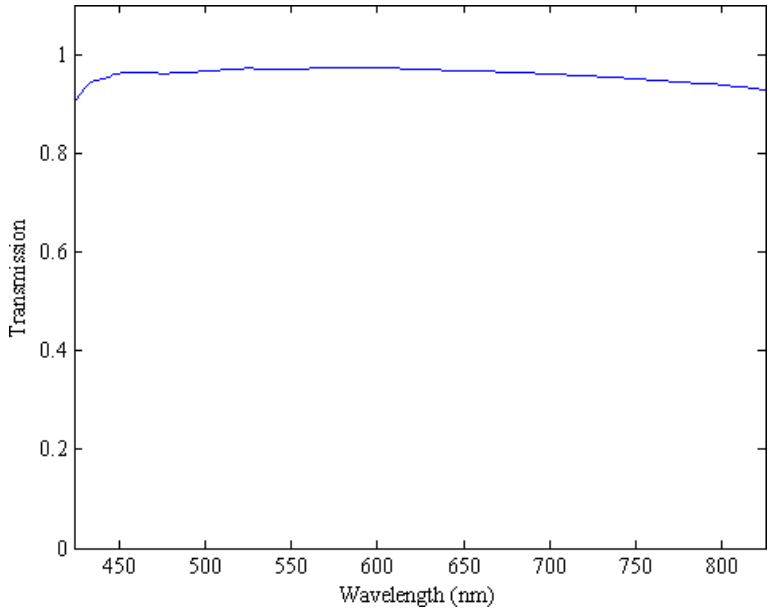

Figure 3.8: Total transmission of the 16-channel demultiplexer.

illuminated with quasi-monochromatic light created by filtering the tungsten source using an interference filter with a narrow passband. Figure 3.14 shows the best-focus image of the centre of the spoke target taken at 580 nm without the demultiplexer in place. Figure 3.15 shows the best focus at five wavelengths across the pass band of the system with the demultiplexer in place; there appears to be no significant loss of image quality due to the demultiplexer.

To achieve best focus, the imaging lens requires refocusing at each wavelength due to longitudinal chromatic aberration within the lenses. Figure 3.16 shows the best focus image at 850 nm, and the image at 500 nm with the same focus setting; there is severe defocus for the 500 nm image.

To characterise the longitudinal chromatic aberration, the illumination is filtered at 20 nm intervals from 440-720 nm using a liquid-crystal tunable filter, and at 20 nm intervals from 740-840 nm using interference filters. For each wavelength the detector position is scanned along the optical axis using a computer-controlled linear positioning stage. At each wavelength and position, an image of the spoke target is captured; the metric used for image-sharpness is the total variance of the image. By interpolating the variance as a function of distance along the optical axis, the position of best focus can be determined for each wavelength. The results are shown in Figure 3.18; the minimum focus distance from the imaging lens occurs at 510 nm and the maximum shift in focus across the passband is 1.5 mm.

This system uses commercial off-the-shelf (COTS) lenses which cannot be redesigned to correct the longitudinal chromatic aberration. For broadband systems in general, lenses that are sufficiently corrected for all wavelengths are not likely to be available

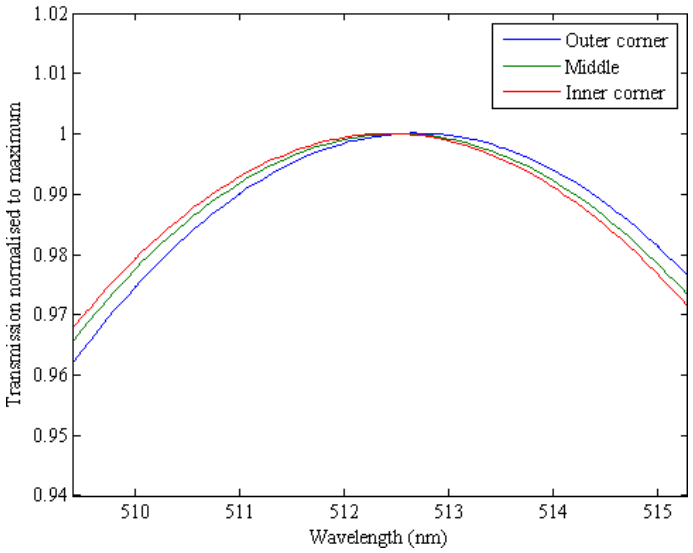

Figure 3.9: Simulated variation in transmission across image for the bottom-right sub-image.

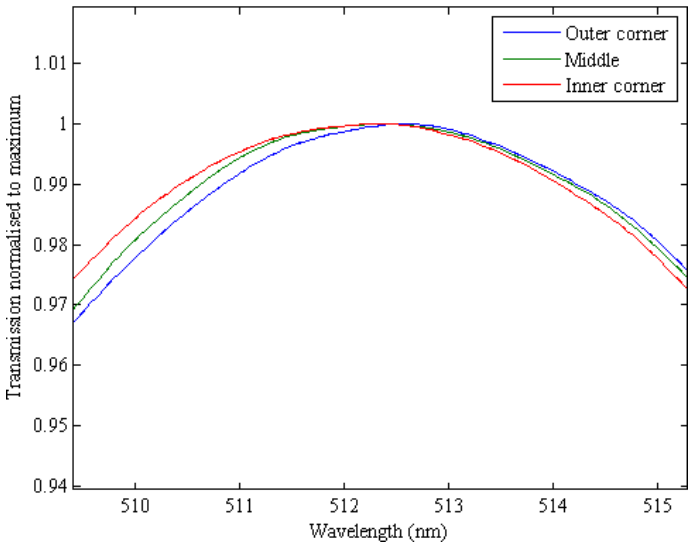

Figure 3.10: Measured variation in transmission across image for the bottom-right sub-image.

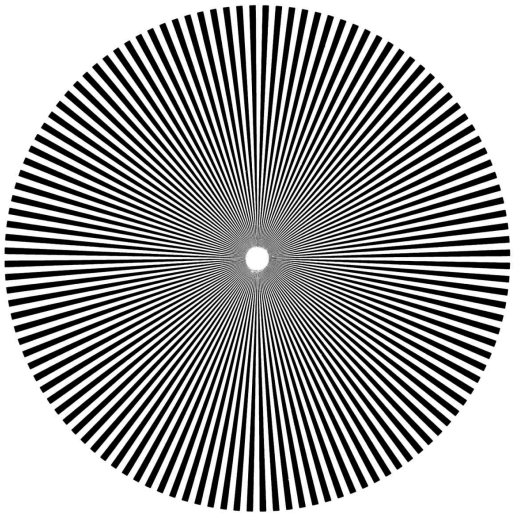

Figure 3.11: Spoke target used to assess image quality.

off-the-shelf, and instead require sophisticated custom designs at significantly higher cost. As described in Section 2.6.1, the aberration may instead be corrected at the peak wavelength in each sub-image by placing an array of corrective glass elements of different thicknesses, one in front of each sub-image image at the detector. The use of corrective glass elements comes at the expense of spherical aberration and a loss of signal through vignetting. Using Equation 2.22 the required corrective thickness of glass can be determined; a plot of the required thicknesses against wavelength is shown in Figure 3.19 for a glass with a refractive index of 1.5. The effects of spherical aberration from these glass elements are difficult to predict without prescription for the lenses but from Equation 2.22 the circle of least confusion for an otherwise well corrected system is calculated to be ≈ 7 μm for a 4.5 mm thick glass element; this is smaller than the pixel width. Using Equation 2.24, vignetting will be begin 0.5 mm from the edge of the sub-image for a piece of corrective glass 5.5 mm from the detector, the centre of each sub-image is unaffected by vignetting.

Figure 3.20 shows the image at 500 nm when focused for 850 nm as in Figure 3.16, but with a 4.4 mm thicknesss of glass inserted into the convergent beam near the detector. Although the optimal thickness to correct the aberration is actually 4.5mm, the defocus is almost entirely removed although there are differences in the contrast across the image. Section 4.4 describes the correction of the longitudinal chromatic aberration in this system with an array of glass elements; these elements are used in conjunction with an array of interference filters which reduce the dispersion due to the Wollaston prisms in each sub-image.

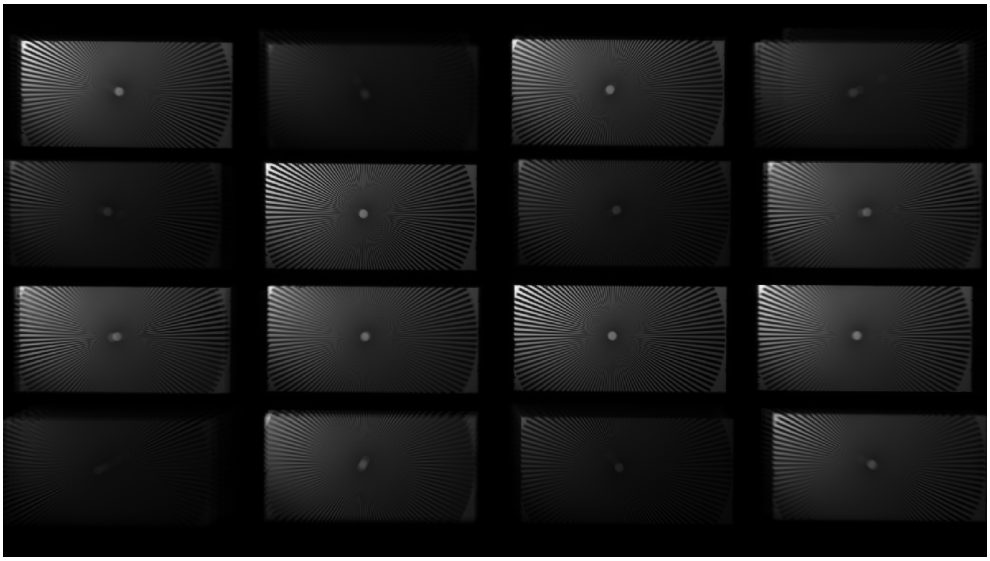

Figure 3.12: Image of a spoke target in broadband illumination.

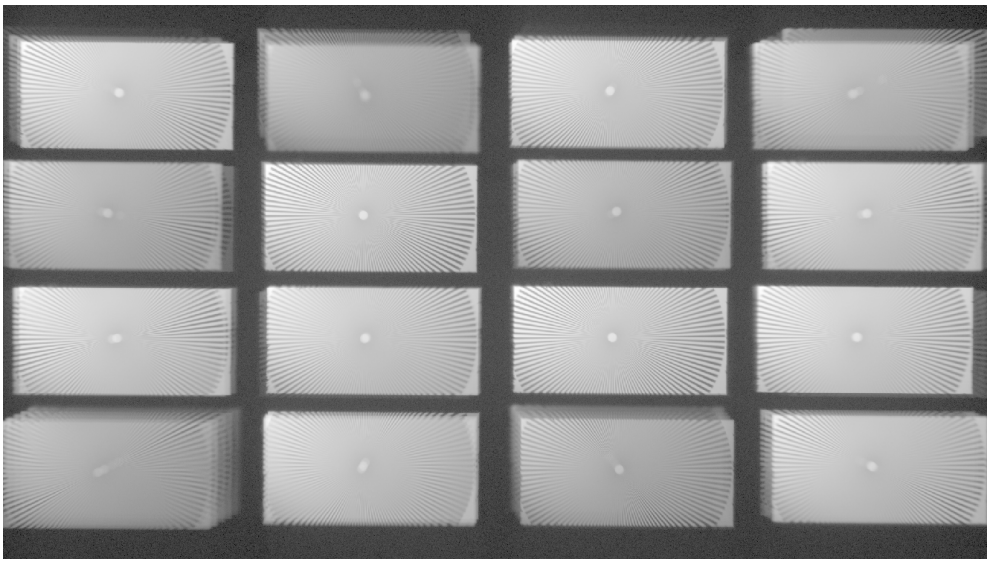

Figure 3.13: Logarithm of the image in Figure 3.12.

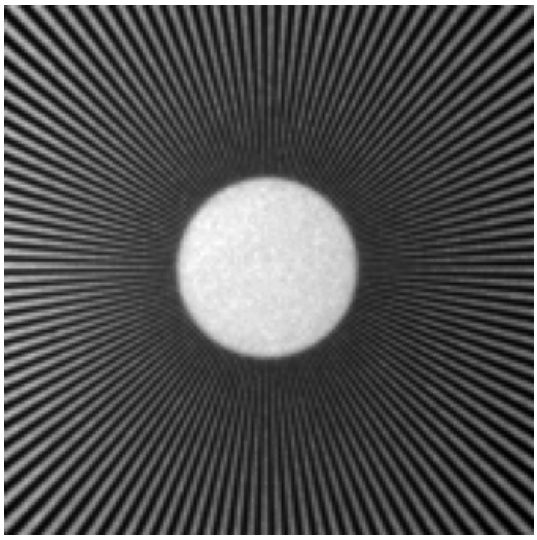

Figure 3.14: Image of a spoke target at 580 nm taken with the IRIS system without the demultiplexer in the optical path.

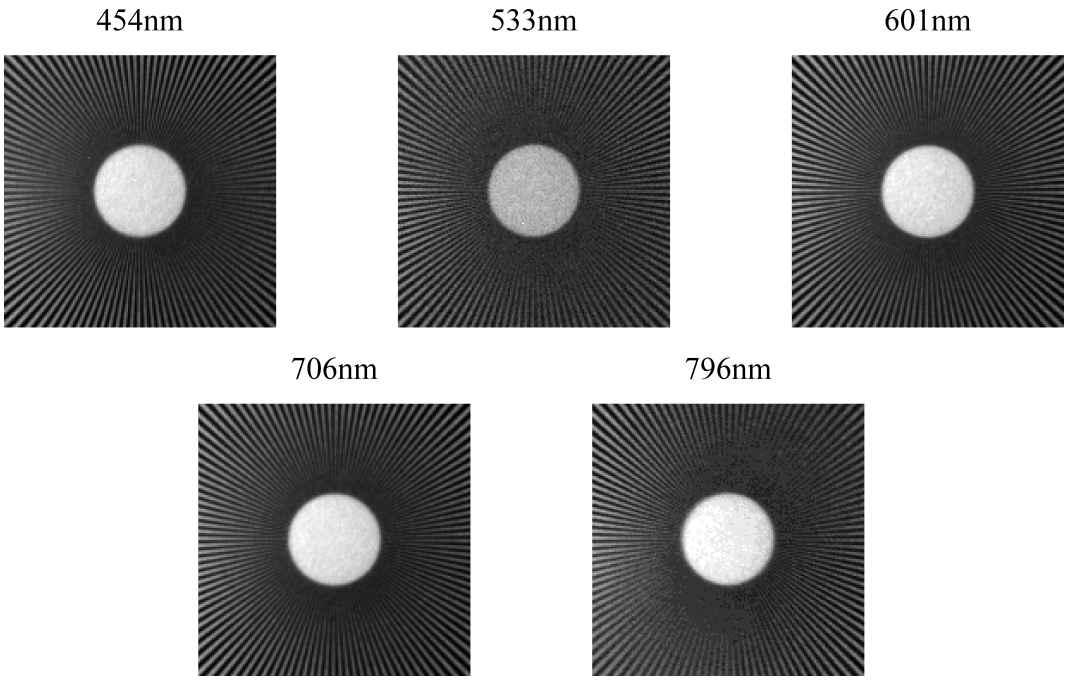

Figure 3.15: Image of a spoke target at 5 wavelengths, taken with the IRIS system with the demultiplexer in the optical path. Demonstrating that the demultiplexer does not noticeably decrease image quality

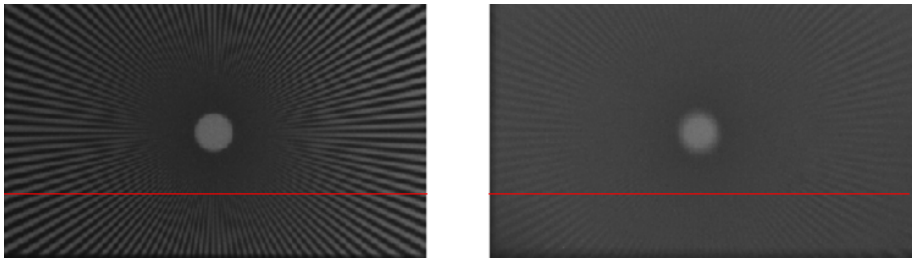

Figure 3.16: (a) best focus image of spoke target at 850 nm and (b) the image at 500 nm with the same focus setting. Figure 3.17 shows the intensity profiles along the red line in each image.

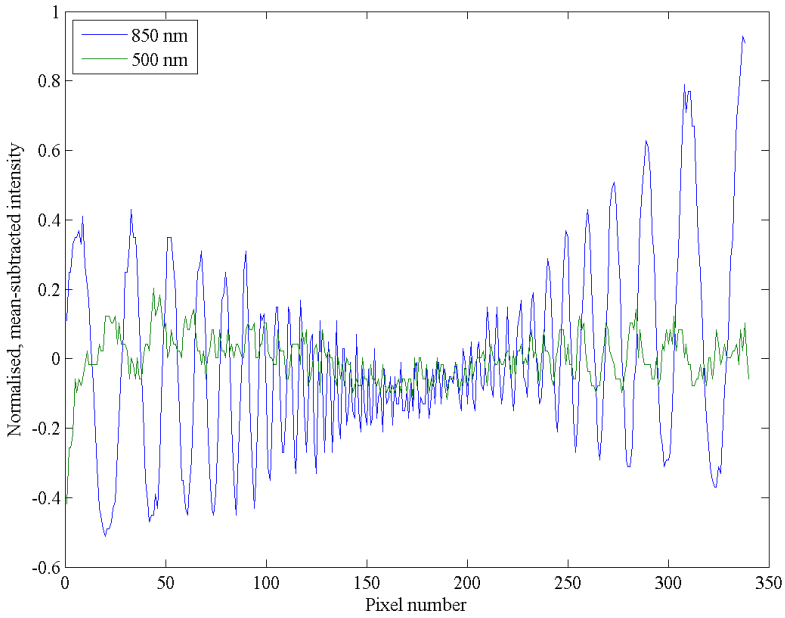

Figure 3.17: Normalised, mean-subtracted intensity profiles corresponding to the red lines in Figure 3.16.

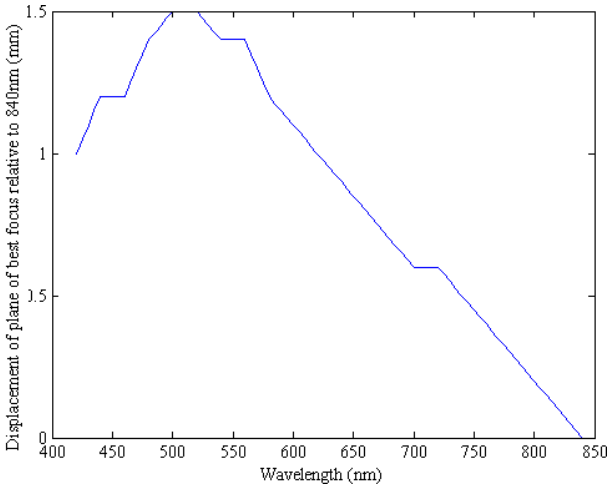

Figure 3.18: Variation in focal distance with wavelength for 16-channel system.

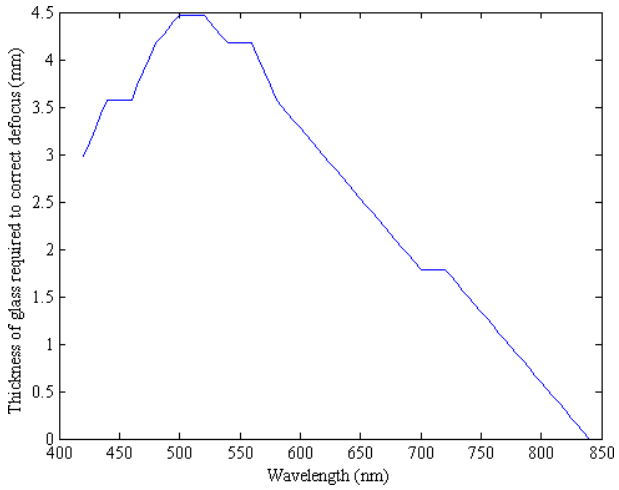

Figure 3.19: Thickness of glass required to correct the longitudinal chromatic aberration in the 16-channel IRIS system.

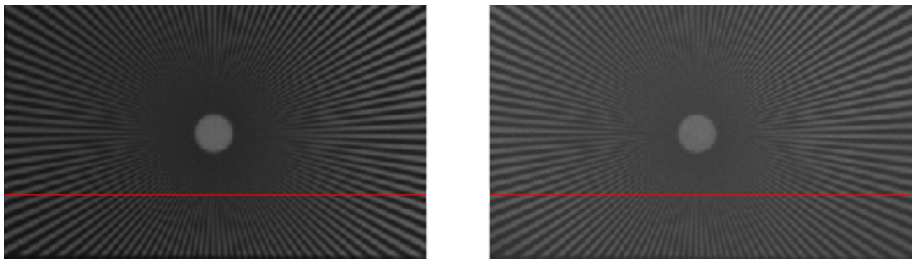

Figure 3.20: (a) best focus image of spoke target at 850 nm and (b) the image at 500 nm with the same focus setting but now with 4.4 mm of glass inserted into the convergent beam. Figure 3.21 shows the intensity profiles along the red line in each image.

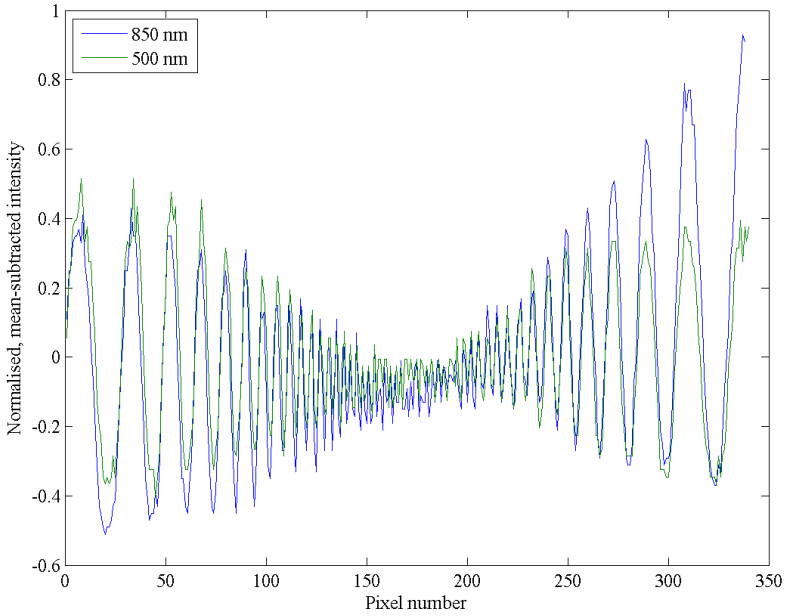

Figure 3.21: Normalised, mean-subtracted intensity profiles corresponding to the red lines in Figure 3.20.

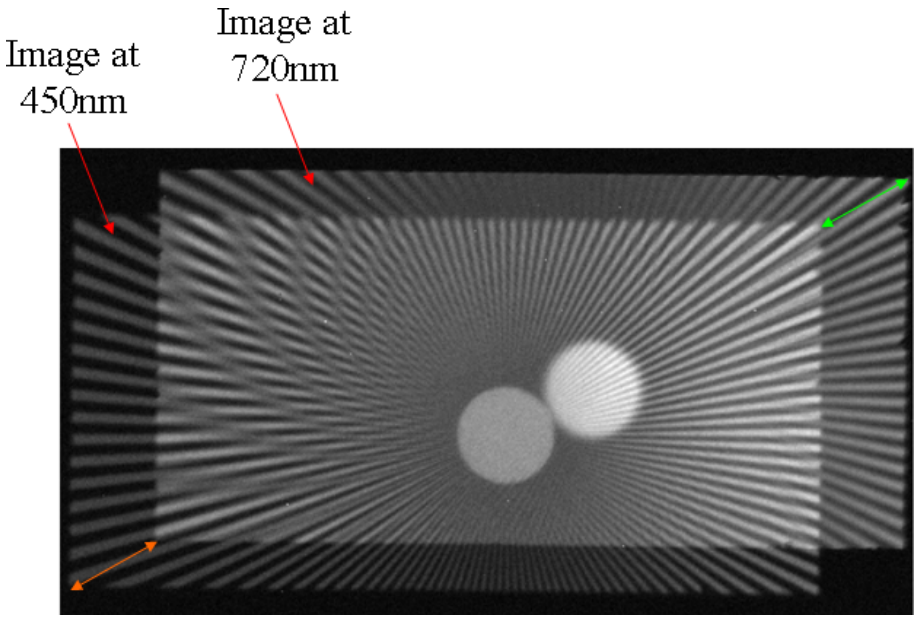

Figure 3.22: Bottom-left corner of superposed IRIS images at 450 nm and 720 nm, the amount of dispersion is equal for the inner (green arrow) and outer corners (orange arrow).

3.7 Dispersion from Wollaston prisms

To measure the spectral dispersion in the images due to the Wollaston prism, the change in position of the image of the field stop with wavelength can be calculated. Figure 3.22 illustrates the displacement in position for the bottom-left sub-image at 450 nm and 720 nm. The change in position of the inner and outer corners is measured from this image to be 47 pixels horizontally and 26 pixels vertically; this is equal for both the inner and outer corners meaning that the dispersion of the Wollaston prism does not change appreciably over the sub-image. The simulated change in position of these same points is 49 pixels horizontally and 25 pixels vertically. Again the simulation is in good agreement with the measurements.

3.8 System f-number and transmission losses from lenses

The system f-number is determined by measuring the convergence angle of the beam along the optical axis near the detector. This measurement is illustrated in Figure 3.23; the convergent beam is produced by imaging a point source produced by placing a pinhole in front of an extended source. Light from the pinhole is collimated using a lens so that the objective lens may be focused at infinity. The collimated beam must also be larger than the entrance aperture of the objective lens. The detector is moved along the optic axis using a linear positioning stage, and by measuring the diameter of the

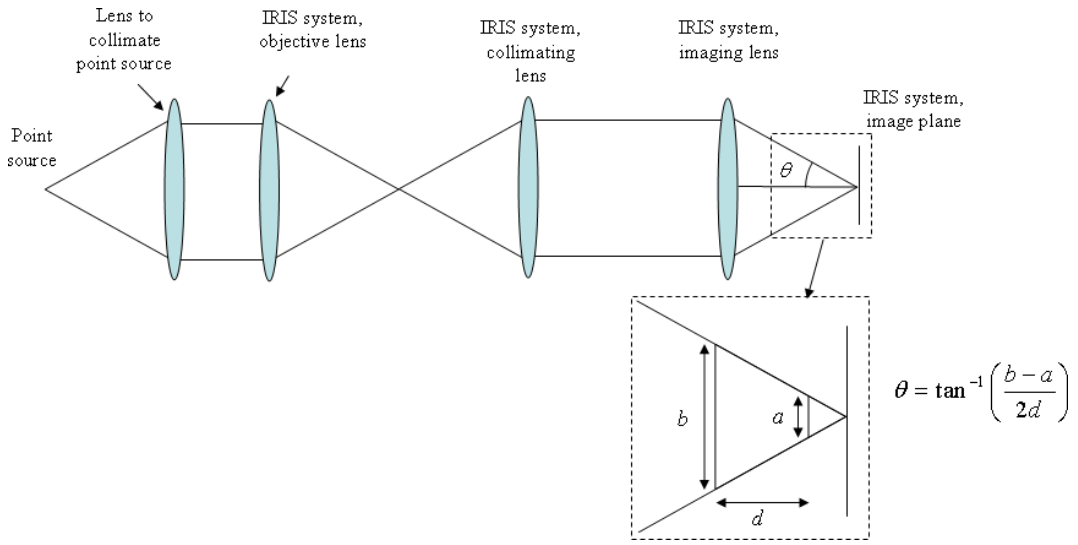

Figure 3.23: An arrangement to measure the convergence angle of the IRIS lens system.

beam at two position separated by a distance d, the convergence angle of the beam and thus system f-number can be calculated. For convenience, the measurement is made without the demultiplexer in place; the f-number is found to be $f/5.9$ compared to the predicted f-number of $f/5.6$.

The transmission losses due to the lenses in the IRIS system are estimated by comparing the detector intensity counts for an image of a spatially uniform white source with an image of the same source taken with a *Nikon* AF-D SLR lens for which the f-number is known and the transmission is estimated to be 90%. Relative to this lens the total transmission of the lenses in the IRIS-system is measured to be 77%.

3.9 Co-registration

As explained in Section 2.6.2, in general each IRIS sub-image is subject to a different non-rectilinear geometric distortion due to the different beam paths through the imaging lens. Registration of the sub-images therefore requires that a polynomial spatial-transformation is applied to each sub-image to bring the sub-images into a common spatial coordinate system. In this section one of the sub-images is chosen as a reference image and the remaining 'target' sub-images are registered to the reference coordinate system.

As the distortion caused by the imaging lens is expected to remain constant, the required spatial transformations for a given IRIS system need only be determined once as a calibration step.

To assess the suitability of this approach for real images, registration has been carried out for the retinal IRIS system operating over 560-600 nm (Section 2.1). A narrow bandwidth system is used so that the registration effectiveness can be assessed in the

absence of severe chromatic aberrations. This system is designed for retinal imaging, and the required registration accuracy for this application is one fifth of a pixel.

As described in Section 2.6.2, the polynomial transform from the ith coordinate in the target image (u_i, v_i) to the correct reference coordinate (p_i, q_i) can be written as the inner products $p_i = \mathbf{w}_i \cdot \mathbf{a}$ and $q_i = \mathbf{w}_i \cdot \mathbf{b}$, where

$$\begin{aligned}
\mathbf{w}_i &= \begin{bmatrix} 1 & u_i & v_i & u_i^2 & v_i^2 & v_i u_i & \dots \end{bmatrix}^T \\
\mathbf{a} &= \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & \dots \end{bmatrix}^T \\
\mathbf{b} &= \begin{bmatrix} b_0 & b_1 & b_2 & b_3 & b_4 & b_5 & \dots \end{bmatrix}^T .
\end{aligned}$$

For a set of coordinates, the transformation can be written as the linear equations

$$\begin{aligned}
\mathbf{p} &= \mathbf{W}\mathbf{a} \\
\mathbf{q} &= \mathbf{W}\mathbf{b},
\end{aligned}$$

where $\mathbf{p}$ and $\mathbf{q}$ are vectors of the target space coordinates and $\mathbf{W}$ is a matrix whose ith row is the polynomial vector $\mathbf{w}_i$. $\mathbf{a}$ and $\mathbf{b}$ are the vectors of $N(N+1)/2$ coefficients for a polynomial of order N. The polynomial coefficients can be found by solving this set of simultaneous linear equations for the set of image features whose coordinates (u_i, v_i) and (p_i, q_i), are known in both the target and reference images; these coordinates are known as control points. For a least-squares solution to the simultaneous equations, the number of control points should be greater than the number of coefficients.

To determine control points for all IRIS sub-images, an array of 78 black dots against a white background has been imaged as shown in Figure 3.24; the coordinates of the dots are used as the control points. To extract these control points from the image, the reference sub-image (sub-image 6 in this case) is first cropped from the IRIS frame and the normalised cross correlation of the reference with the entire IRIS frame is calculated and used to locate the remaining target sub-images. The result of this operation is a set of approximately registered sub-images as shown in Figure 3.25.

Some lateral chromatic aberration due to the prisms is evident in the 2nd sub-image (Figure 3.26). Towards the corner of the detector in sub-images 1, 4 , 5, and 8, there are additional image aberrations due to the lenses (Figure 3.27). There is also a very small amount of longitudinal chromatic aberration in the system leading to a slight change in focus between the sub-images. These aberrations will remain after coregistration.

Each sub-image is divided by a low-pass filtered copy of itself to remove intensity variations across the sub-image. The magnitude of the difference between the 1st and 6th of these flattened sub-images is shown in Figure 3.28, it is clear that due to lens distortion the dots are not well aligned between the sub-images.

The coordinates of the dots are extracted by cropping a single dot from the reference

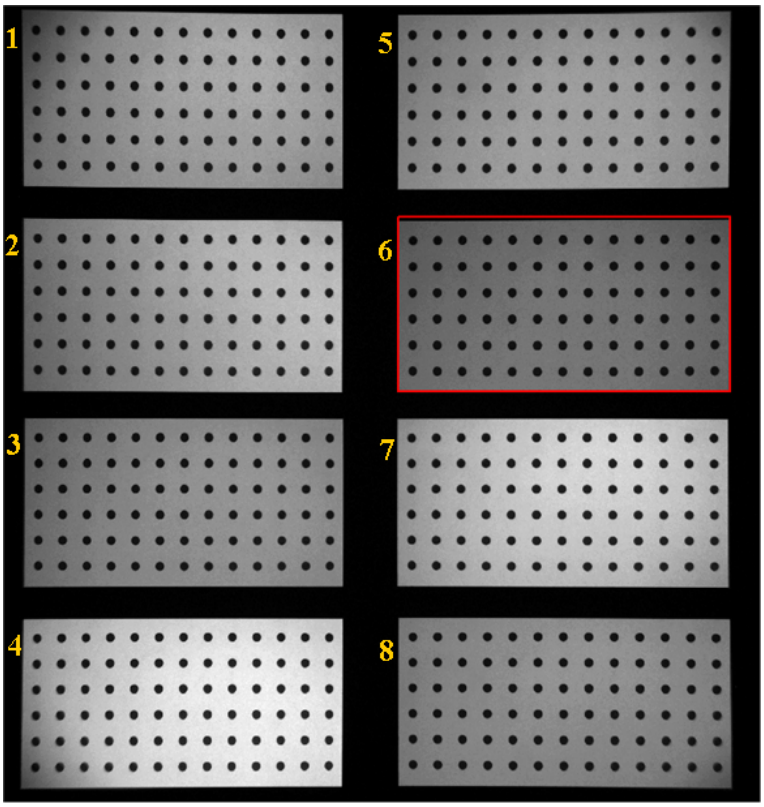

Figure 3.24: An 8-channel IRIS image of an array of dots used as features for coregistration.

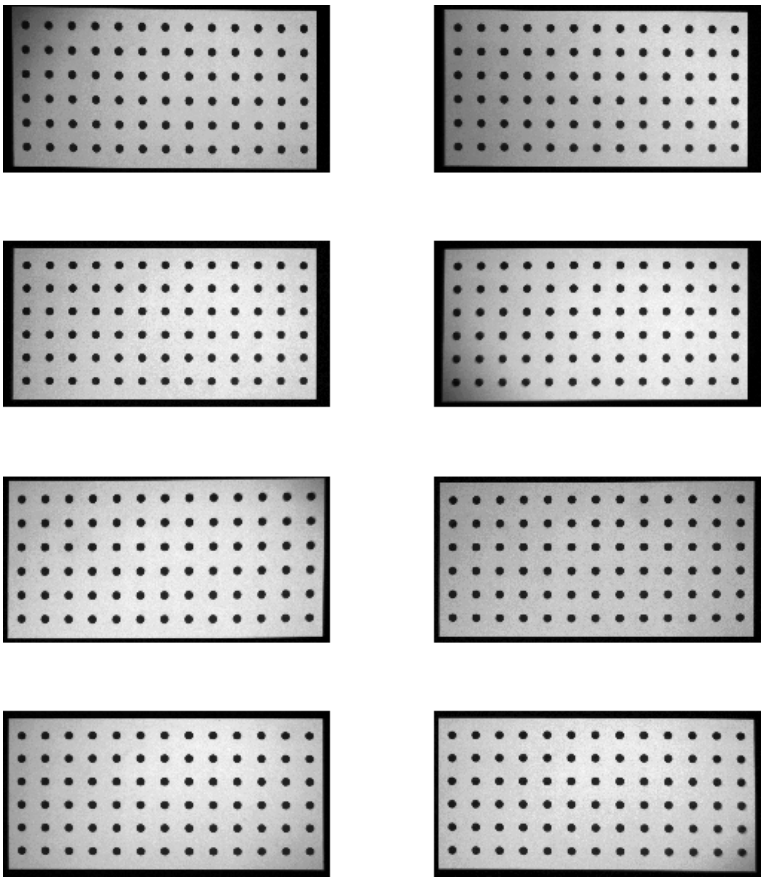

Figure 3.25: Cropped and approximately registered sub-images.

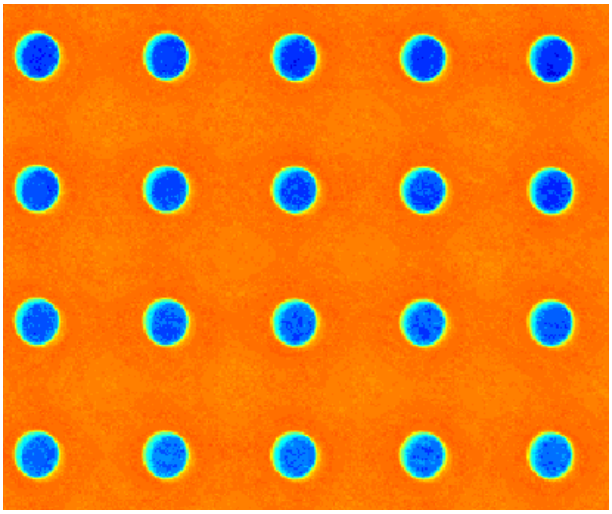

Figure 3.26: False colour image of a section of the 2nd sub-image. The lateral chromatic aberration due to the Wollaston prisms is evident as the light-blue crescent in the top left of each dot.

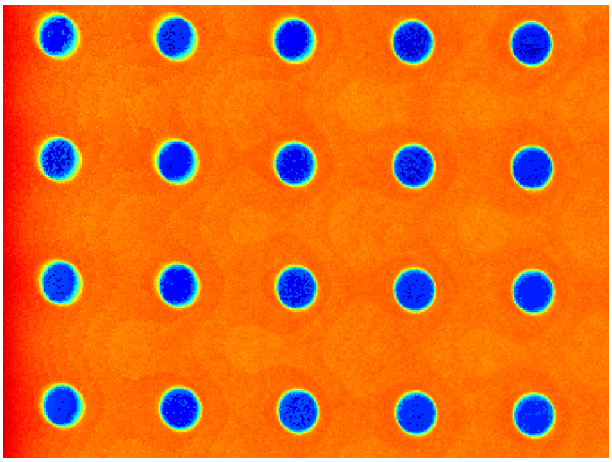

Figure 3.27: False colour image of the top-left corner of the 1st sub-image, demonstrating the aberrations near the corner of the detector due to the lenses.

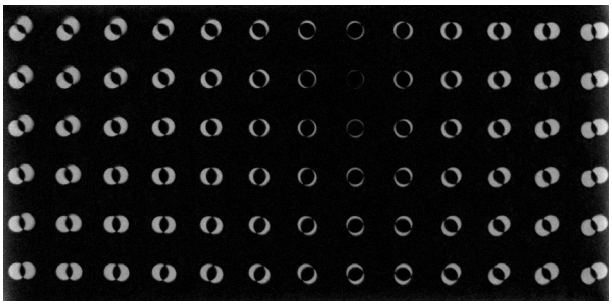

Figure 3.28: Difference between the 1st and 6th approximately registered sub-images.

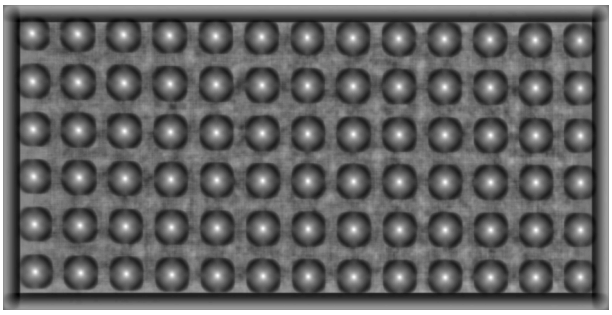

Figure 3.29: Normalised cross-correlation of a single dot with the 1st sub-image.

image and calculating the normalised cross correlation of this dot with every sub-image; an example of the cross correlation of the dot with sub-image 1 is shown in Figure 3.29. The coordinates of the 78 maxima are extracted from this correlation image using 2-dimensional spline interpolation to achieve 100th of a pixel precision. Using these control points the average registration error in the untransformed images is found to be 6.35 pixels.

The control points are separated into two sets; a random selection of 58 of the control points from each sub-image is used as training data to determine the transformation coefficients; ten of the remaining control points are used as validation data to determine the correct polynomial fit, and the final ten control points are used as test data to assess the accuracy of the registration. To prevent over fitting, the order of polynomial is determined by applying successive third-order polynomial transformations until the validation control points are registered to less than 100th of a pixel accuracy. In this case three third-order transformations are applied, which is equivalent to a 27th order polynomial in which only terms of order 1-9, 12, 18, and 27 are nonzero.

Figure 3.30 shows the magnitude of the difference between the 1st and 6th transformed images; as before, the images have been flattened before differencing. The alignment of the dots is much improved although the effects of the image aberrations present in the 1st sub-image can be seen in the top-left corner. The slight difference in focus between the images manifests as the annuli in the middle and right of this image; these result from differencing a sharp image of the dot with a slightly defocused image.

The average registration error as determined from the 10 test control points is better than 300th of one pixel, which is more than sufficient for the intended application.

3.10 Conclusions

This chapter has demonstrated the manufacture of the 16-channel broadband IRIS system. The performance of the manufactured system has also been characterised through measurements of the relevant system parameters. An interferometric ray-tracing program has also been constructed to aid with the design of IRIS systems, and to accurately

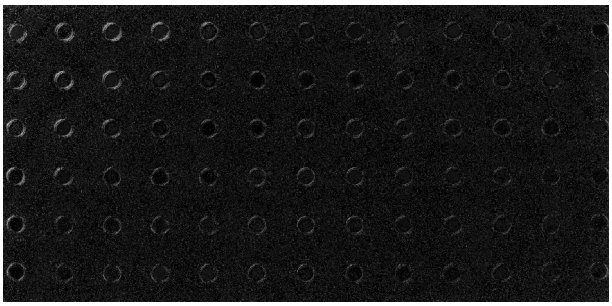

Figure 3.30: Difference between the 1st and 6th transformed sub-images.

predict their performance prior to construction. There is good agreement between this model and the measurements. The registration of the IRIS images was also demonstrated.

It is clear from the preceding work that chromatic aberrations are a major issue in IRIS systems, particularly those with broad bandwidths. The following chapter investigates methods for dealing with these aberrations.

Chapter 4

Countering Birefringent dispersion

Summary: Dispersion in the birefringence of Wollaston prism materials causes a form of lateral chromatic aberration in the output image of IRIS. The extent of this aberration increases with wavelength making it an issue for broad-band systems. In this chapter, several methods for removing or reducing the effects of this aberration are investigated: minimisation of the spectral width of the bands during optimisation, achromatisation of Wollaston prisms, using an array of interference filters at the image plane to reduce the spectral width of each band, and attempting to remove the aberration using image processing.

4.1 The effects of dispersion

The refractive indices and birefringence of birefringent materials exhibit a dependence on wavelength. For a conventional Wollaston prism the splitting angle will therefore also depend on wavelength.

As described in Sections 2.6.1 and 3.7, this birefringent dispersion leads to a type of lateral chromatic aberration in the sub-images in which the radiation from a point source is spread over a straight line at the detector. If the detector has R_x pixels in the x direction, and the image is replicated 2^K times in this direction, then in a sub-image q_x places from the centre of the detector, the number of pixels of smear in the x direction is given approximately by

$$p = R_x \Delta B \, (2q - 1) / \left(2^{K+1} \bar{B}\right), \tag{4.1}$$

where ΔB is the maximum difference in birefringence over the wavelength range and $\bar{B}$ is the average birefringence over this range. Figure 4.1 shows a plot of the number of pixels of smear as a function of bandwidth (centered on 550nm), for calcite and BBO prisms, with $R_x = 2000$, with the image replicated four times in the x-direction ($K = 2$), for the 2nd sub-image from the centre of the detector ($q_x = 2$).

At a single point on the detector, this dispersion mixes radiation from different spa-

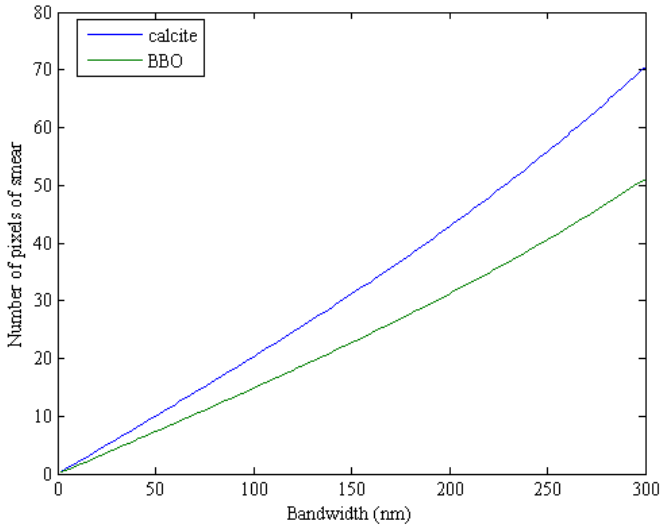

Figure 4.1: Amount of smear as a function of bandwidth for calcite and BBO prisms, with the image replicated 4 times in the x-direction over 2000 pixels, for the 2nd image from the centre of the detector, the central wavelength is 550nm.

tial elements at different wavelengths. This complicates interpretation and processing of the data using standard spectral processing algorithms, making it desirable to reduce or remove the effects of dispersion prior to processing. This chapter discusses options for achieving this goal.

4.2 Minimising the effects of smear in the transmission functions

One method that has already been discussed for reducing dispersion is to minimise the secondary peaks in the transmission functions during optimisation of the bands, thereby allowing only a narrow range of wavelengths in each sub-image. In theory this will reduce the full extent of dispersion p, by a factor of $1/J$, where J is the number of bands. From Figure 4.1, for a system using calcite prisms this will be effective over a bandwidth of 40nm for an 8 channel system and 80nm for a 16-channel system. Even for these narrow band-widths, it will not generally be possible to remove all of the effects of dispersion in this manner. Other constraints in the objective function used to optimise the bands may further limit the effectiveness of this approach. As an example, Figure 4.2 shows a set of bands for an 8-channel system with passband 530-570nm, optimised to minimise the power in the secondary peaks of the bands. The point-spread functions (PSF) for an 8-channel system with these bands and calcite prisms, has been simulated via ray tracing and is shown in Figure 4.3, the base-10 logarithm of these PSF are shown in Figure 4.4. As can be seen, the effects of dispersion are still present in the system.

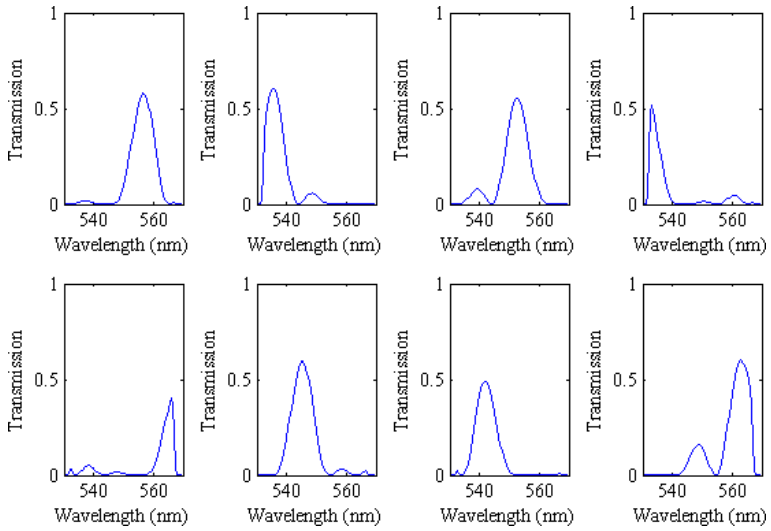

Figure 4.2: Set of bands for an 8-channel system with passband 530-570nm, optimised using the simplex method to minimise the power in the secondary peaks of the bands.

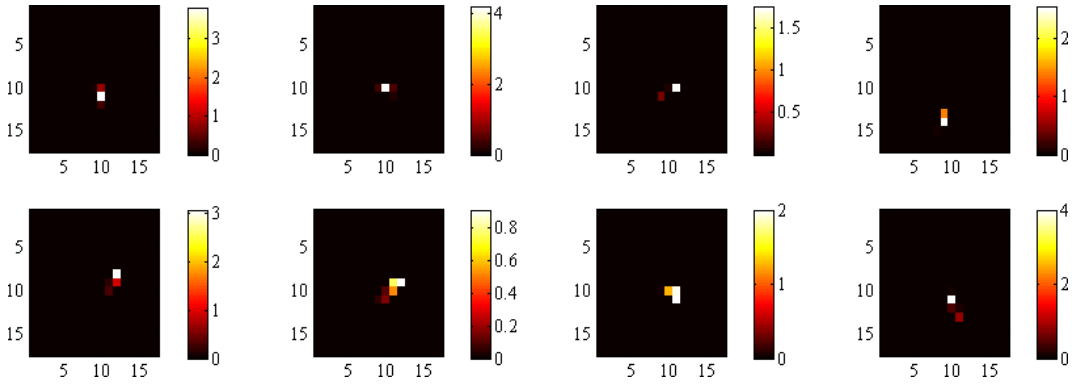

Figure 4.3: The point-spread functions (PSF) for an 8-channel system with bands as shown in Figure 4.2 and calcite prisms.

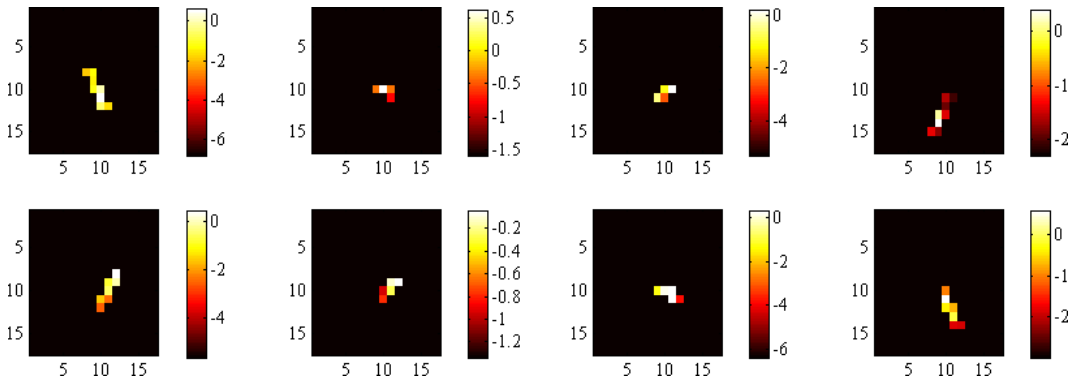

Figure 4.4: Logarithm of the PSF shown in Figure 4.3.

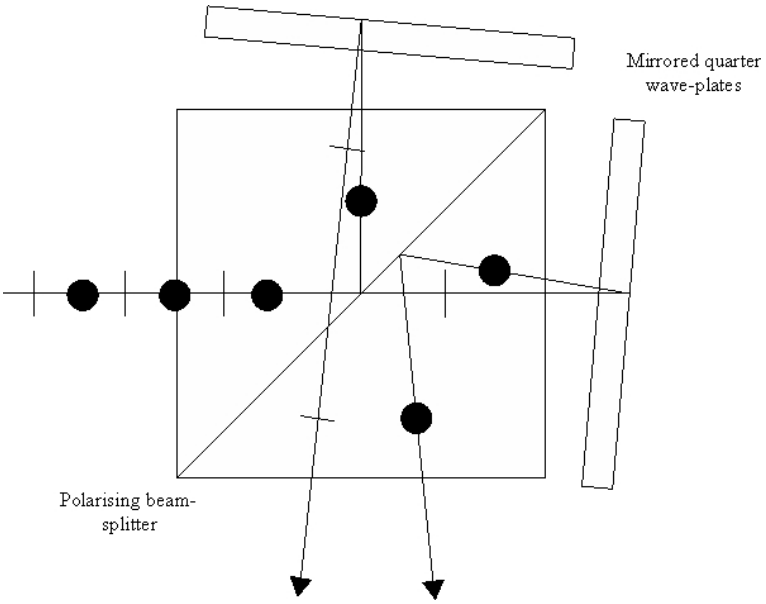

Figure 4.5: Arrangement of components for a Wollaston-like prism.

4.3 Achromatising Wollaston prisms

4.3.1 Reflective, Wollaston-like prisms

Little has been published on producing achromatic Wollaston prisms. One alternative to using conventional Wollaston prisms uses reflection instead of refraction to split the beams [105]. This tunable, 'Wollaston-like' prism makes use of a polarising beam-splitter, and two achromatic quarter-wave plates with a reflective coating on their back-surfaces. When arranged as in Figure 4.5 this arrangement acts like a Wollaston prism, although with the output beams folded 90°. By changing the angle that the mirrored waveplates make with the beam splitter, the splitting angle can be 'tuned' to the desired value. As the splitting of the beams occurs via reflection rather than refraction, the splitting angle is independent of wavelength. Reflection losses and wavefront distortion at the prism and wave plate surfaces may be an issue, but as the tunability is unnecessary for IRIS, the spaces between these components may be filled with an index matched achromatised prism. Also, changes in the performance of the beam-splitter coating and waveplates, with both angle and wavelength, may introduce reflection losses and stray light. The suitability of this device for use in IRIS systems is still to be investigated in detail, but it may prove additionally attractive due to the likelihood of lower production costs, as isotropic materials are more commonly available in large sizes across the entire optical region.

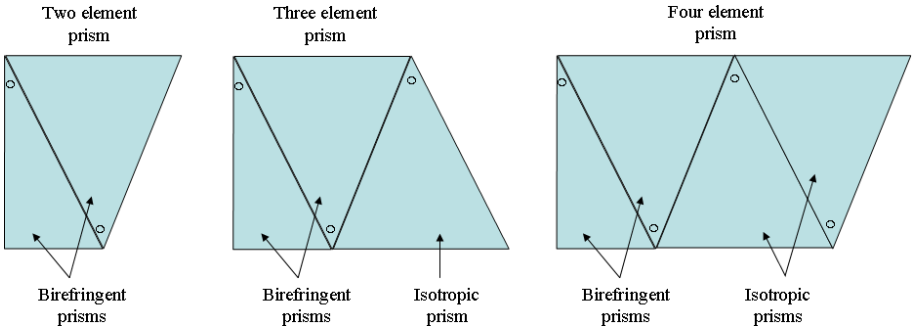

Figure 4.6: Illustration of the geometry of the (a) two-element, (b) three element, and (c) four element prisms in the achromatising optimisation. Angles marked with a ∘ are allowed to vary as part of the optimisation.

4.3.2 Achromatising Wollaston prisms using combinations of materials

In optical design it is common to produce achromatic and apochromatic components by using combinations of high-dispersion and low-dispersion materials. It is reasonable to suppose that, using similar techniques, achromatic Wollaston prisms may be produced.

To determine the possibility of constructing an achromatic Wollaston prism by employing combinations of different materials, the database of material transmissions and dispersion formulae collated in Section 2.5 has been used with the ray tracing software developed in Section 3.2, in a computer optimisation to determine the lowest dispersion combinations. If different materials are used for each half of the prism, with indices n_{e1} and n_{o1}, and n_{e2} and n_{o2} respectively, the optic axes of the prisms can be parallel to one another if $n_{e1} \neq n_{e2}$ and $n_{o1} \neq n_{o2}$, or perpendicular if $n_{o1} \neq n_{e2}$ and $n_{e1} \neq n_{o2}$. It should be noted that if the two half-prisms are made from different materials, the splitting of the beam will, in general, no longer be symmetric about the direction defined by the input ray. This may make index matching of the demultiplexer components problematic but is otherwise not an issue. Because of this probable asymmetry, the exit surface normal is also allowed to vary in the following optimisation (Figure 4.6 (a)).

To simplify the comparison between combinations, the average splitting angle across the range of wavelengths should be the same for each combination. The objective function in the optimisation has been chosen heuristically to be $\left(\left(\bar{\phi} - \phi_t\right) + 1\right)\sigma_\phi$, where $\bar{\phi}$ is the average splitting angle, ϕ_t is the target splitting angle, and σ_ϕ is the average standard deviation of the angle across the wavelength range.

As an example, there are 21 birefringent materials in the database that are transmissive over the entire range $0.4\text{nm} - 0.7\text{nm}$, giving 420 ordered combinations. With the wedge angles constrained to $\leq 45°$, and $\phi_t = 5°$, Table 4.1 shows the 10 combinations of materials resulting in the lowest dispersions when the optic axes of the two prism halves are perpendicular to one another. Table 4.2 shows the best combinations when the optic

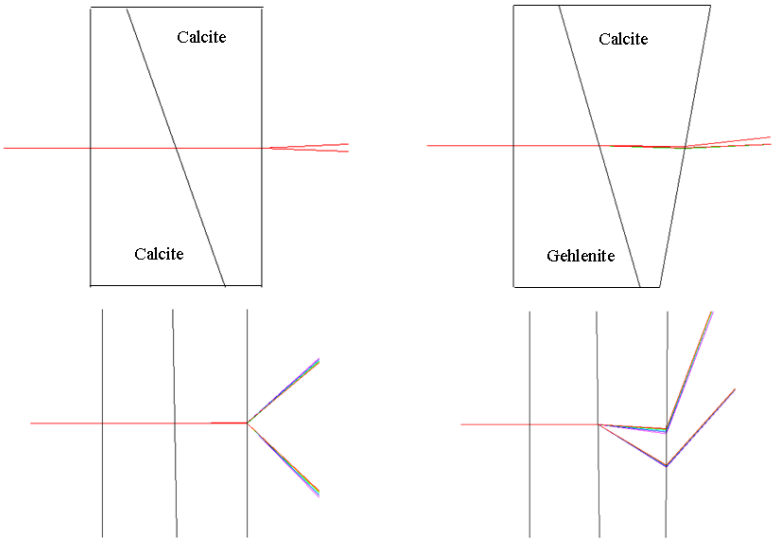

Figure 4.7: Ray trace of a calcite-calcite prism and a gehlenite-calcite prism.

axes are parallel. When the optic axes are perpendicular, there is little improvement to be had over a BBO-BBO prism. With the axes parallel, the combination of $Ca_2Al_2SiO_7$ (gehlenite) and calcite results in dispersion of around 7% that of a conventional calcite prism. Figure 4.7 shows a trace of the rays through this combination compared to a conventional Wollaston prism made from calcite, also shown are the same traces with the splitting angles exaggerated for clarity. Figure 4.8 shows a plot of the change in the position with wavelength, of the spot produced at the detector by an intially on-axis ray when a imaging lens with 50mm focal length is used with this prism in the arrangement shown in Figure 2.13. The axis of symmetry for rays produced the gehlenite-calcite prism makes an angle of around 4.5° with the optical axis of the system. Calcite occurs naturally in large single crystals and prisms of up to 4cm aperture are available. Transparent gehlenite also occurs naturally, and it appears that single crystals of 10mm in dimension can be grown using the Czochralski method [106]. With more care it may be possible to grow larger crystals.

It is possible to extend the attempts at constructing an achromatic Wollaston prism by using additional prisms, made from isotropic materials, following the first two birefringent prisms (Figure 4.6 (b) and (c)). The Schott Glass catalogue [107] lists 106 glasses for use over the visible range; this is not an exhaustive list of available materials but is reasonably comprehensive. With each extra isotropic prism, the number of possible combinations of materials grows by a factor of 106, so that there are over 40000 combinations of materials when a single additional prism is used, and over 4 million combinations for 4 prisms. There are also $2P$ surface normals where P is the number of

Materials	**Wedge angles (degrees)**	**Splitting angle (degrees)**	**Dispersion relative to calcite prism**
$LiCaAlF_6$ BBO	0 39.75 9.446	5	0.6613
BBO $CsLiB_6O_{10}$	0 -31.28 5.149	5	0.6751
$CsLiB_6O_{10}$ BBO	0 -35.65 -5.12	5	0.6759
Al_2O_3 BBO	0 -33.65 -1.293	5	0.7034
$Ca_5\,(PO_4)_3\,F$ BBO	0 36.65 2.053	5	0.7173
BBO $LiCaAlF_6$	0 33.35 -9.749	5	0.7216
BBO BBO	0 -20 -0.008441	5	0.7234
BBO Al_2O_3	0 -35.27 -0.3991	5	0.7239
BBO $Ca_5\,(PO_4)_3\,F$	0 -35.71 1.395	5	0.7385
Quartz BBO	0 44.3 8.554	5	0.7389

Table 4.1: Combinations of two materials resulting in lowest dispersion Wollaston prisms over the 400-700nm region. The prism optic axes are perpendicular.

Materials	**Wedge angles (degrees)**	**Splitting angle (degrees)**	**Dispersion relative to calcite prism**
$Ca_2Al_2SiO_7$ Calcite	0 -15.52 9.952	3.668	0.06766
Calcite $Ca_2Al_2SiO_7$	0 29.79 9.983	5	0.295
CsH_2AsO_4 BBO	0 43.32 1.282	5	0.474
BBO CsH_2AsO_4	0 -40.67 -2.475	5	0.5179
ADP BBO	0 45 6.606	3.81	0.6387
$LaCl_3$ BBO	0 45 10	3.687	0.6715
$Ca_5\,(PO_4)_3\,F$ BBO	0 -44.3 -9.675	5	0.7025
BBO ADP	0 -45 10	4.476	0.7044
Quartz BBO	0 38.3 6.161	5	0.706
Al_2O_3 BBO	0 30.65 -0.7108	5	0.7083

Table 4.2: Combinations of two materials resulting in lowest dispersion Wollaston prisms over the 400-700nm region. The prism optic axes are parallel.

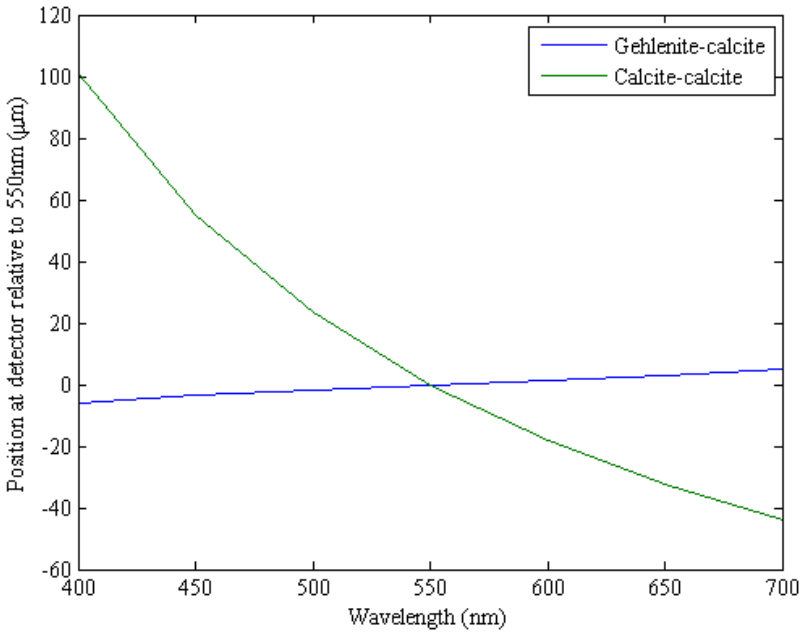

Figure 4.8: Change in position with wavelength, of the spot produced by initially on axis ray when a imaging lens with a focal length of 50mm is used, for a calcite-calcite prism and a gehlenite-calcite prism.

prisms, birefringent and isotropic, in the system. To reduce the number of parameters in the optimisation, the front surface of the first prism is kept normal to the incident ray and the normals of adjacent internal surfaces are parallel to one another; the direction of the final surface normal is also allowed to vary. As well as reducing the number of surface-normal variables to P, this allows the components of the Wollaston prism to be index matched, reducing Fresnel losses and distortion. As before, the splitting of the beam by these arrangements of prisms is no longer likely to be symmetric around the input ray direction, meaning that the whole demultiplexer will not be easily index matched. Anti-reflection coatings may then be necessary on external surfaces, although immersion of the entire demultiplexer in index-matching fluid is also a possibility.

It is computationally expensive to search the entire range of combinations, so a random set of 10000 were generated and optimised for both the three- and four-material prisms. The ten lowest-dispersion, three-material combinations, when the optic axes are perpendicular, are shown in Table 4.3. The results for parallel optic axes are shown in Table 4.4. Ray traces for the lowest disperison three- and four-material combinations are shown in Figure 4.9; in both cases the axis of symmetry of the exit rays makes an angle of around 4.7° with the optical axis. Combinations containing calcite and gehlenite, or BBO and gehlenite, offer low dispersion in both cases. $LiNbO_3$ and $BaTiO_3$ also appear to be potentially useful materials. $BaTiO_3$ is a photoreactive crystal and single crystals are available in at least 5mm dimensions. $LiNbO_3$ is commonly used in non-linear optics and is available with dimensions of 15mm. Other materials that appear in some of the

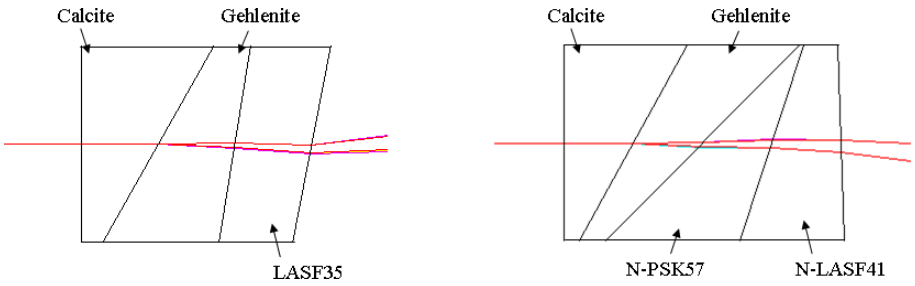

Figure 4.9: Ray trace of the lowest dispersion (a) three- (first entry in Table 4.3) and (b) four-material (first entry in Table 4.5) combinations.

lower dispersion combination are $LiSrAlF_6$ and $\alpha - GaN$, $\alpha - GaN$, neither appears to exist in large single crystals, although doped $LiSrAlF_6$ crystals have been grown with 7mm dimensions [108].

4.4 Using an array of spectral filters at the image plane to remove secondary peaks in the IRIS transmission functions

One obvious solution to the problem of dispersion is to use a small interference filter in front of each sub-image to create narrower transmission functions by removing the secondary peaks in the transmission functions. The filters can be placed directly in front of the detector if possible, or otherwise at a conjugate image plane. This requires the filters to be arranged in a 2D array. If a filter array is used, the bands will become closer to direct bands at the expense of a loss in total throughput. The objective function in the optimisation of the bands should be to maximise the power in the filter peaks.

The 16-channel system described in Section 3.3 is designed for use with a filter array. The 16 individual filter tiles for this array were manufactured by *Envin Scientific*; the measured transmissions of these filters are shown in Figure 4.10. The bands of the 16-channel system, which were shown in Section 3.5, were optimised to maximise the total power through this filter array. Figure 4.11 shows the predicted point spread functions due to dispersion without the filter array in place, the dispersion in the corner images is around 80 pixels in the horizontal and 40 in the vertical. Figure 4.12 shows the point spread functions with the filter array in place; the dispersion in the corner sub-images is reduced to around 10 pixels in the horizontal and 5 in the vertical.

Initially the filter array was to be placed 12mm in front of the detector in the plane at the which the waist of the convergent beams was expected to be a minimum, as shown in Figure 4.13. This plane was chosen as it allowed for the most relaxed tolerances on cutting and positioning the filter tiles. The imaging lens used in this ray trace is a double Gauss lens which was assumed to be a reasonable approximation to the actual

Materials	**Wedge angles (degrees)**	**Splitting angle (degrees)**	**Dispersion relative to calcite prism**
Calcite $Ca_2Al_2SiO_7$ LASF35	0 27.4 8.103 10	4.694	0.02897
BBO $Ca_2Al_2SiO_7$ N-SK18	0 38.11 26.72 10	5	0.03844
$Ca_2Al_2SiO_7$ Calcite N-LAF2	0 -16.61 6.482 10	3.795	0.04016
$Ca_2Al_2SiO_7$ BBO K10	0 -25.42 38.29 10	8.581	0.07926
$LiNbO_3$ $BaTiO_3$ SF5	0 43.63 22.67 10	3.604	0.08288
BBO $\alpha-$GaN N-SSK8	0 37.32 35.98 10	5	0.08605
Calcite $\alpha-$GaN SK51	0 29.14 26.82 10	5.186	0.08662
$\alpha-$GaN Calcite N-LAF2	0 -1.591 16.29 10	3.034	0.1013
$\alpha-$GaN BBO N-BAF3	0 -1.7 35.53 10	5	0.1173
$BaTiO_3$ $LiNbO_3$ N-SK14	0 -33.4 27.43 10	5	0.1254

Table 4.3: Combinations of three materials (two birefringent and one isotropic) resulting in lowest dispersion Wollaston prisms over the 400-700nm region. The prism optic axes are perpendicular.

Materials	**Wedge angles (degrees)**	**Splitting angle (degrees)**	**Dispersion relative to calcite prism**
BBO $Ca_2Al_2SiO_7$ N-BAK4	0 -28.02 -36.3 10	3.496	0.03532
$LiNbO_3$ BBO N-SF5	0 -14.74 -42.83 10	3.736	0.1386
Calcite $Ca_2Al_2SiO_7$ N-LAK21	0 -28.16 -42.04 10	5	0.1525
$LiSrAlF_6$ BBO N-LASF44	0 33.35 -6.6 10	5	0.2233
BBO $\alpha - GaN$ N-LAF2	0 27.89 28.67 10	3.531	0.2267
BBO $LiSrAlF_6$ N-LAF21	0 -34.69 -6.994 10	5	0.2593
$BaTiO_3$ BBO N-SF10	0 5.34 40.37 10	5	0.2915
BBO $BaTiO_3$ N-F2	0 -31.77 -35.24 10	4.245	0.3142
$Ca_2Al_2SiO_7$ BBO N-SK10	0 -12.46 -44.3 10	5	0.3367
Calcite $BaTiO_3$ N-SF15	0 -23.54 -29.5 10	4.209	0.3584

Table 4.4: Combinations of three materials (two birefringent and one isotropic) resulting in lowest dispersion Wollaston prisms over the 400-700nm region. The prism optic axes are parallel.

Materials	**Wedge angles (degrees)**	**Splitting angle (degrees)**	**Dispersion relative to calcite prism**
Calcite $Ca_2Al_2SiO_7$ N-PSK57 N-LASF41	0 28.86 44.75 18.3 -1.88	5	0.02498
BBO $Ca_2Al_2SiO_7$ LAKL12 SF57	0 33.96 42.77 9 5.851	4.19	0.03619
Calcite $\alpha - GaN$ SF11 N-LAK7	0 31.46 33.87 31.7 0.8165	5.766	0.08309
BBO $BaTiO_3$ BASF51 N-BAK4	0 -24.72 -30 31.39 4.809	3.03	0.1576
$\alpha - GaN$ Calcite N-SSK8 N-BK7	0 -3 -30.14 -6.645 6.471	5	0.2025
AlN BBO N-BALF4 SF1	0 30.65 -18.06 25.76 -5.042	5	0.2517
BBO $\alpha - GaN$ N-BAK2 N-ZK7	0 41.32 42.8 -45 -9.988	6.371	0.4601
Calcite $LiSrAlF_6$ N-SK18 N-LAK10	0 28 -19.16 40.97 -7.135	5	0.4681
Quartz $LiYF_4$ N-BK10 N-F2	0 45 -45 9.788 -10	3.175	0.5673
$LiSrAlF_6$ BBO LF5 LAFN7	0 28.97 -11.84 15 8.64	5	0.6108

Table 4.5: Combinations of four materials (two birefringent and two isotropic) resulting in lowest dispersion Wollaston prisms over the 400-700nm region. The prism optic axes are perpendicular.

Materials	**Wedge angles (degrees)**	**Splitting angle (degrees)**	**Dispersion relative to calcite prism**
$\alpha - GaN$ BBO N-KZFS4 LLF1	0 1.776 -35.34 -32.69 -9.701	5	0.0969
Calcite $\alpha - GaN$ N-SK5 SFL57	0 -28.71 -26.28 14.84 8.937	5.09	0.1365
$BaTiO_3$ $LiNbO_3$ N-BALF5 KZFSN4	0 37.3 -34.32 -16.65 4.526	6.665	0.3113
BBO $Ca_2Al_2SiO_7$ N-SF1 K7	0 36.3 32.36 31.71 -1.078	5	0.463
$Ca_2Al_2SiO_7$ Calcite N-PSK53 N-BAF3	0 -9.419 19.32 6.8 9	5	0.5875
ADP BBO N-SF1 N-BAF4	0 34.95 -16.42 -21.65 6.206	5	0.6293
CsH_2AsO_4 ADP N-KZFS11 SF66	0 -45 45 -27.43 -10	3.591	0.6336
CsH_2AsO_4 BBO N-SK15 N-BAF3	0 -21.75 22 2.464 -0.8532	5	0.6357
BBO $CsLiB_6O_{10}$ LAKL12 LASF35	0 -36.3 -39.55 5 2.257	5	0.7072

Table 4.6: Combinations of four materials (two birefringent and two isotropic) resulting in the lowest dispersion Wollaston prisms over the 400-700nm region. The prism optic axes are parallel.

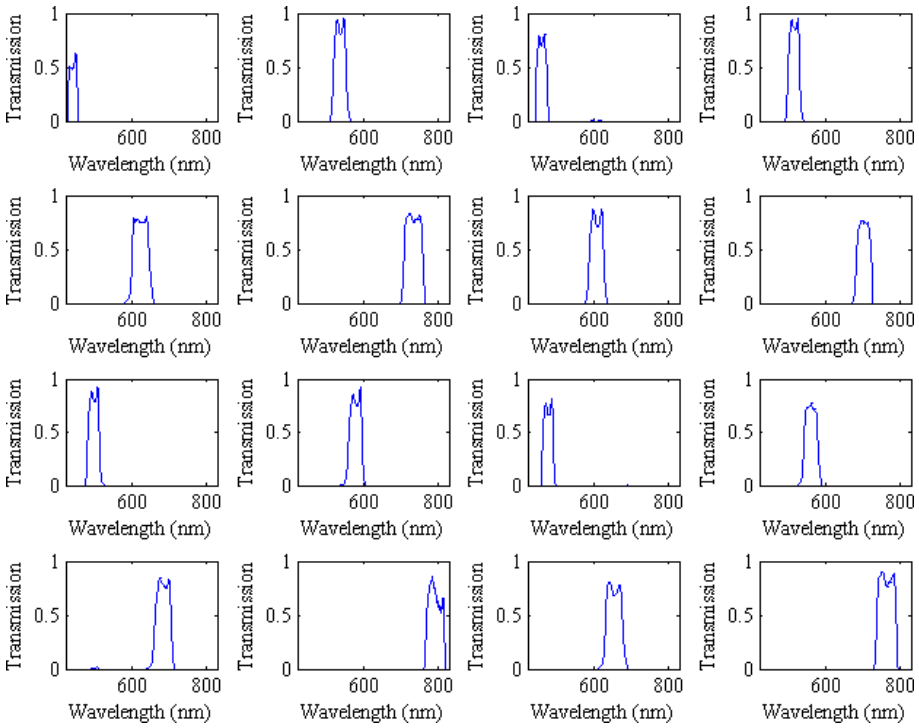

Figure 4.10: Transmission functions of the individual filters in the filter array.

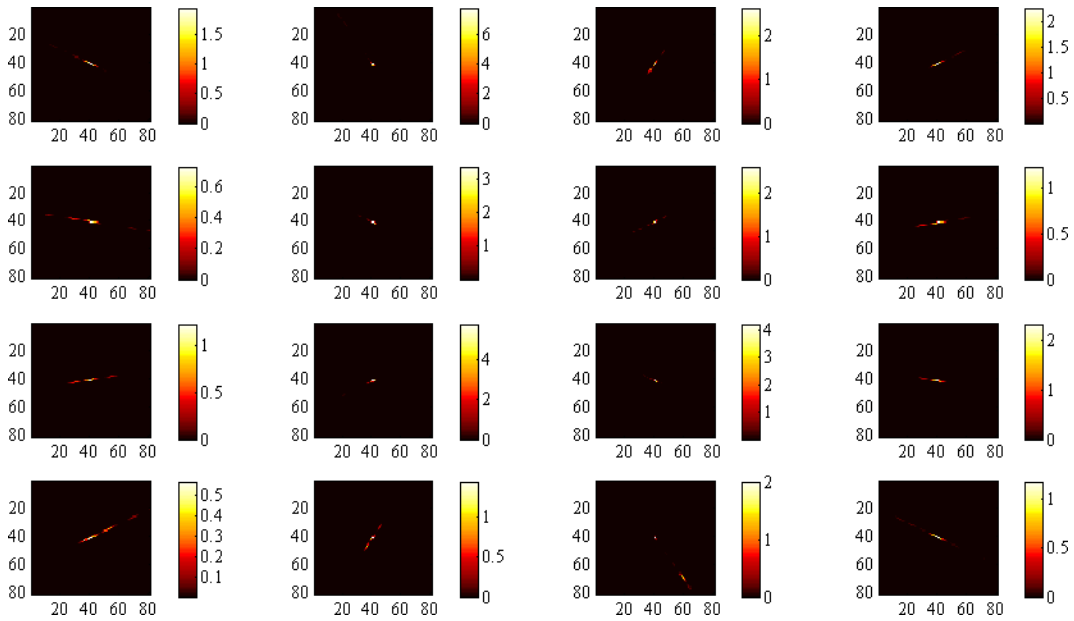

Figure 4.11: PSFs of the 16-channel system without the filter array.

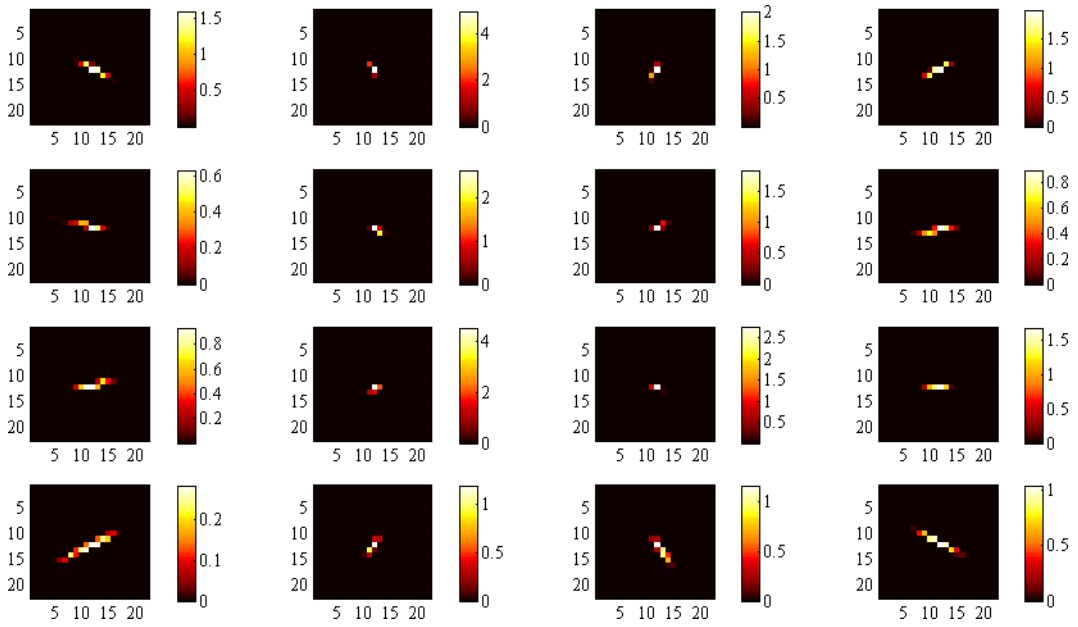

Figure 4.12: PSFs of the 16-channel system with the filter array.

imaging lens, which is a Canon SLR lens for which the prescription is unknown. A first prototype matrix filter for use in this plane was manufactured by Envin Scientific. However, it was found upon construction of the IRIS system that the beam waists in this plane are not a minimum and are in fact larger than the filter tiles. As the prescriptions of the lenses used are not available, it is not known why these differences in the beam shapes occur. It was instead found that the minimum beam diameter occurs at the detector. Unfortunately the filter array constructed by Envin Scientific was not suitable for use at the detector due to the positioning of the filter tiles. Further attempts at constructing a filter array for use directly in front of the detector were made as part of the work of this thesis. In both the initial prototype and later attempts, the filter array was produced by affixing the filter tiles to a substrate of N-BK7 with an optical adhesive. Figure shows the final filter array 4.14. Although still not optimal, the filter positions are as well aligned as is possible with the existing filter tiles. The performance of the array could be improved in future with more accurate cutting of the tiles.

As mentioned in Section 3.6, the filter array has been combined with an array of glass tiles of varying thickness in order to correct the longitudinal chromatic aberrations present in the system. These corrective tiles were affixed to the reverse side of the filter array substrate as illustrated in Figure 4.15. The required thicknesses of these tiles were determined from the plot in Figure 3.19. Unfortunately at the time of construction, the glass tiles were available only with thicknesses in multiples of 1.1mm. From Equation 2.24 the

The constructed filter array was placed approximately 2mm in front of the detector. The measured transmissions with the filter array in place are shown in Figure 4.16.

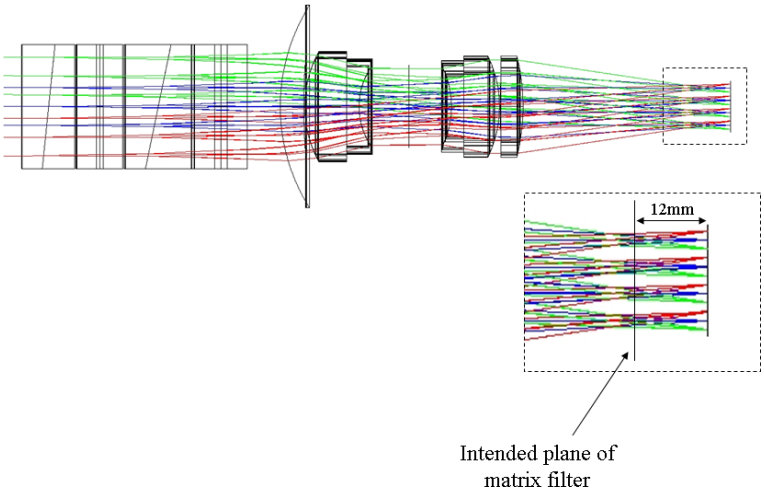

Figure 4.13: Initial position of the filter array.

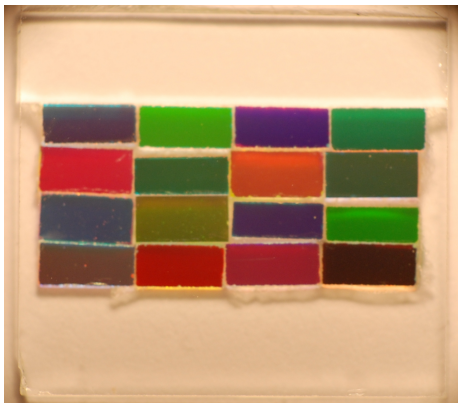

Figure 4.14: Photograph of the filter array.

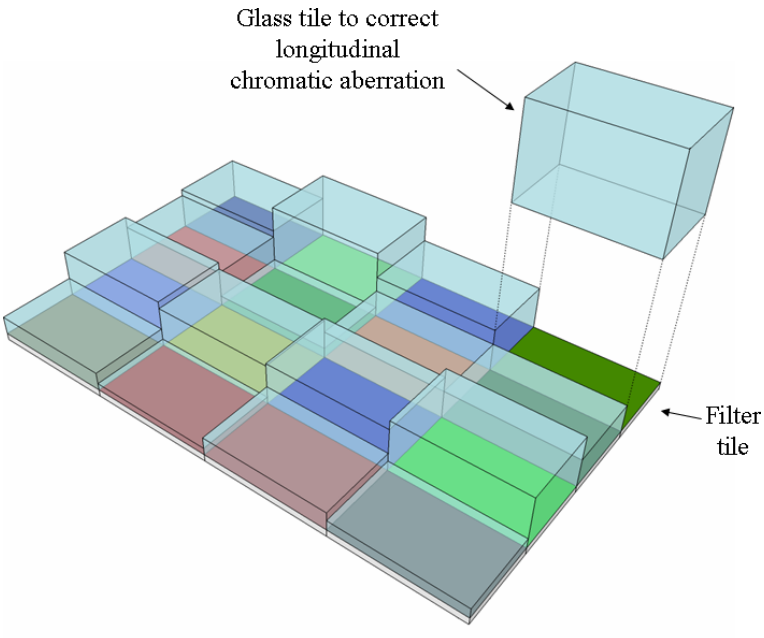

Figure 4.15: Filter array with glass tiles to correct longitudinal chromatic aberration.

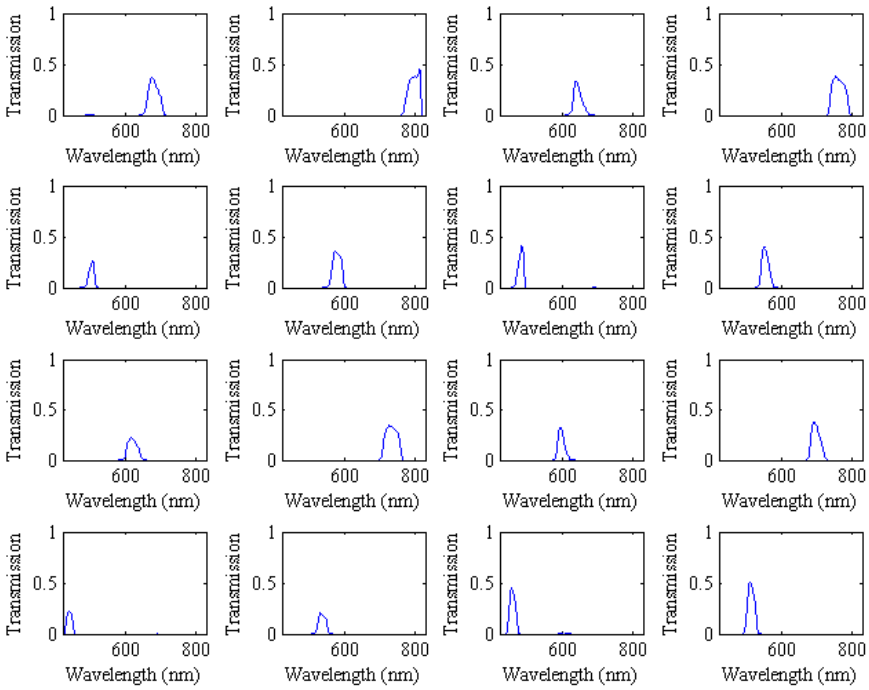

Figure 4.16: Measured transmission functions of the 16-channel system with the filter array.

Figure 4.17 shows an image taken without the filter array, the effects of dispersion are particularly obvious in sub-images 4, 5 and 13. Defocus due to longitudinal chromatic aberration is obvious in sub-images 5, 13, and 15. Figure 4.18 shows a similar image with the filter array in place. The suboptimal positioning of the filter tiles is apparent around the edge of each sub-image, however, the effects of dispersion are greatly reduced. The difference in focus between sub-images is also reduced by the glass tiles. However, some residual defocus is particularly apparent in sub-images 6 and 15. The longitudinal chromatic aberration correction could be improved by using glass tiles of optimal thickness. Figure 4.19 shows a false colour image of a spectral angle map (Section 1.3.2) produced from the frame in Figure 4.18, the portion of the MacBeth ColorChecker in the spectral image is also shown alongside for reference. Note the discrimination between the background vegetation and the green tiles on the ColorChecker.

So far only hardware based solutions to the problem of birefringent dispersion have been considered, the next section discusses the use of image processing to reverse the effects of this dispersion.

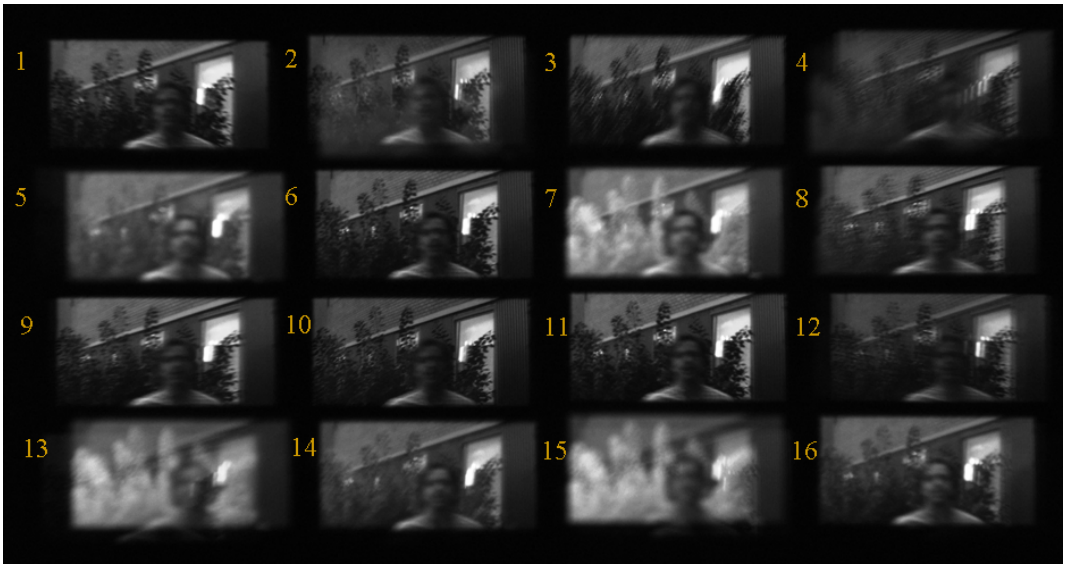

Figure 4.17: Image from 16 channel system without filter array or longitudinal chromatic aberration correction.

Figure 4.18: Image from 16-channel system with filter array and array of corrective glass tiles.

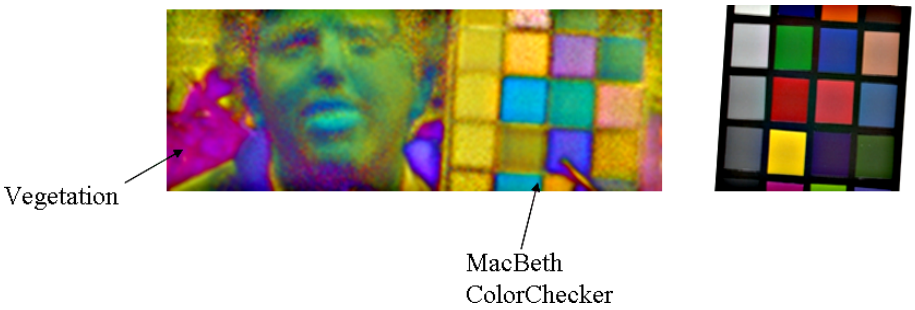

Figure 4.19: A image of a spectral-angle map produced using the 16 sub-images shown in Figure 4.18.

4.5 Software based correction of the birefringent dispersion

It is common practice to model the effects of imaging aberrations as the convolution of the system PSF and the unaberrated image. Image processing, using image-deconvolution techniques such as maximum entropy deconvolution or the Richardson-Lucy algorithm, is often employed to attempt to reverse the effects of these aberrations. In the most simple case, the PSF is known and considered constant across the field of view. The problem becomes more complicated if the PSF varies across the field of view and more difficult still if the PSF is poorly known. These problems are usually known as shift-variant deconvolution and blind deconvolution respectively. Much work has been published on blind-deconvolution: [109] and [110] offer a review of algorithms for solving this class of problems.

Ignoring aberrations due to the imaging and collimating components, the PSF for a point source, in each IRIS sub-image, depends not only on the transmission function and birefringent dispersion but also on the spectrum of the source, which is unknown. The type of correction needed for IRIS is therefore in the class of shift-variant blind-deconvolution algorithms; unfortunately many of the existing algorithms for solving these problems are not suitable for IRIS, as they make the assumption that the PSF is smoothly varying, or can be parameterised as in the case of motion blur. Also, as the deconvolution problem is in general ill-posed, it is preferable to use the spectral and spatial information from the entire set of sub-images to better constrain the solution.

With this in mind, one option that is particularly attractive for reducing the effects of smear with the help of software, is to note that the measurement made by each pixel is equivalent to a weighted line integral at some angle through the object cube, which has two spatial dimensions x and y, and one spectral dimension k. As recognised by the creators of CTIS (described in Section 1.2.6), such spectral measurements are similar to emission-tomography measurements, and therefore, inversion of the data to recover the object-cube may be possible using reconstruction techniques commonly used in computer-tomography. A commonly cited reconstruction technique is the expectation-maximisation algorithm detailed in [62]. This maximum-likelihood method is used to invert the CTIS data, and assumes that at a point in the object cube, the number of photons that are detected in a given time follows a Poisson distribution. The object cube can be discretised into B boxes; the number of detected photons emitted from box $b_{x,y,k}$ is $m(b)$, which is distributed as

$$P\left(m\left(b\right)=M\right)=e^{-\lambda(b)}\frac{\lambda\left(b\right)^{M}}{M\left(b\right)!}, \tag{4.2}$$

where $\lambda\left(b\right)$ is the expectation value of $m\left(b\right)$. The expectation value of the number of

photons emitted in b and detected by detector d is then given by $\langle Q(b,d)\rangle = \lambda(b)\,p(b,d)$, where $p(b,d)$ is the probability of an emission in b being detected by detector d. $p(b,d)$ describes the mapping of each point in the object cube to the detector, and can be determined using accurate ray-tracing or via calibration. The expectation value of the number of total counts in detector d is $\langle n(d)\rangle = \sum_{b=1}^{B} \lambda(b)\,p(b,d)$ and the likelihood of the measurements is

$$L(\mathbf{m}) = P(\mathbf{n}\,|\mathbf{m}) = \prod_{d=1}^{D} e^{-\langle n(d)\rangle} \frac{\langle n(d)\rangle^{n(d)}}{n(d)!}. \tag{4.3}$$

For this probabilistic model, [62] shows that the log-likelihood is a concave function and that the Khun-Tucker conditions lead to a single-step iterative expectation-maximisation scheme:

$$\lambda_{r+1}(b) = \lambda_r(b) \sum_{d=1}^{D} \frac{n(d)\,p(b,d)}{\sum_{b'=1}^{B} \lambda_n(b')\,p(b',d)} \tag{4.4}$$

In this scheme, the likelihood of the retrieved object cube given the measured data increases with each iteration.

To assess the feasibility of this algorithm for correcting the dispersion present in IRIS systems, simulations of the retrieval process have been performed. The probabilities $p(b,d)$ is the transmission of the IRIS system for detector d for the point b in the object cube, for an 8-channel system operating between 400nm and 700nm. These transmissions are calculated using the interferometric ray tracing program, by simulating an impulse at each point b and recording the impulse response in each d. BBO has been assumed as the material for the Wollaston prisms as it offers lower dispersion than calcite. The simulated system uses a 1000×1000 pixel detector with 5μm pixels. The bands for this system have been optimised to provide an impulse response close to a direct system, when the Backus-Gilbert method is used to invert the data to wave-number space. The bands are shown, as ordered at the detector, in Figure 4.20. This is not likely to be the optimal set of bands for ensuring success at removing the dispersion with software, but should serve as a useful starting position.

The object-cube consists of a set of spectral-reflectance images of the MacBeth ColorChecker®. A colour image of this chart is shown in Figure 4.21 together with the image at 550nm. This spectral data is discretised into the desired number of boxes; in this case the data is represented by a stack of 32 monochromatic images of size 434×201 pixels. As long as correct order is preserved, the transformation from the set of emissions $m(b)$ to the detector $n(d)$ can be written as the matrix equation $\mathbf{n} = \mathbf{P}\mathbf{m}$, where $\mathbf{P}$ is the transfer matrix as calculated using the ray trace. The simulated image is shown in Figure 4.22. Dispersion is most apparent in the two leftmost images in the top row and the leftmost image in the bottom row, there is little dispersion in images

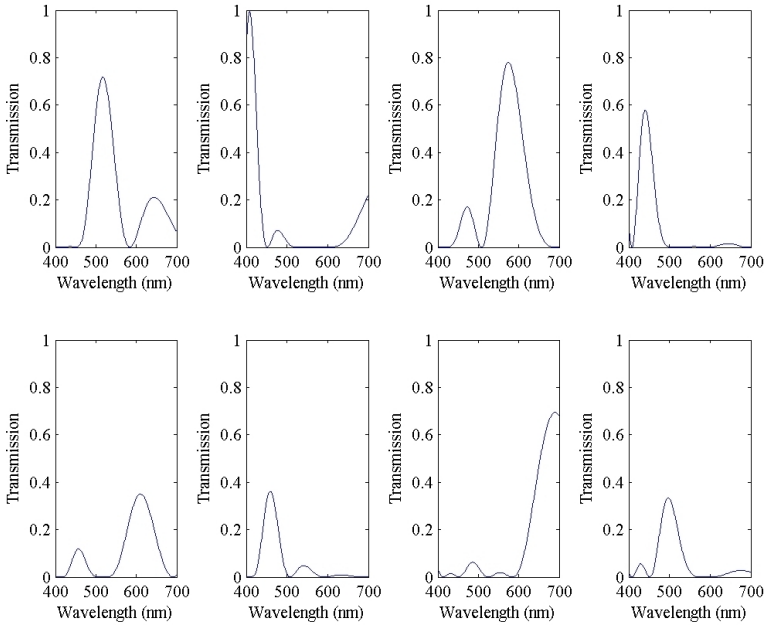

Figure 4.20: Bands for the 8-channel system used in the dispersion simulation, ordered according to position at the detector.

Figure 4.21: MacBeth colour chart used in the dispersion simulation. Colour image and gray-scale image at 550nm.

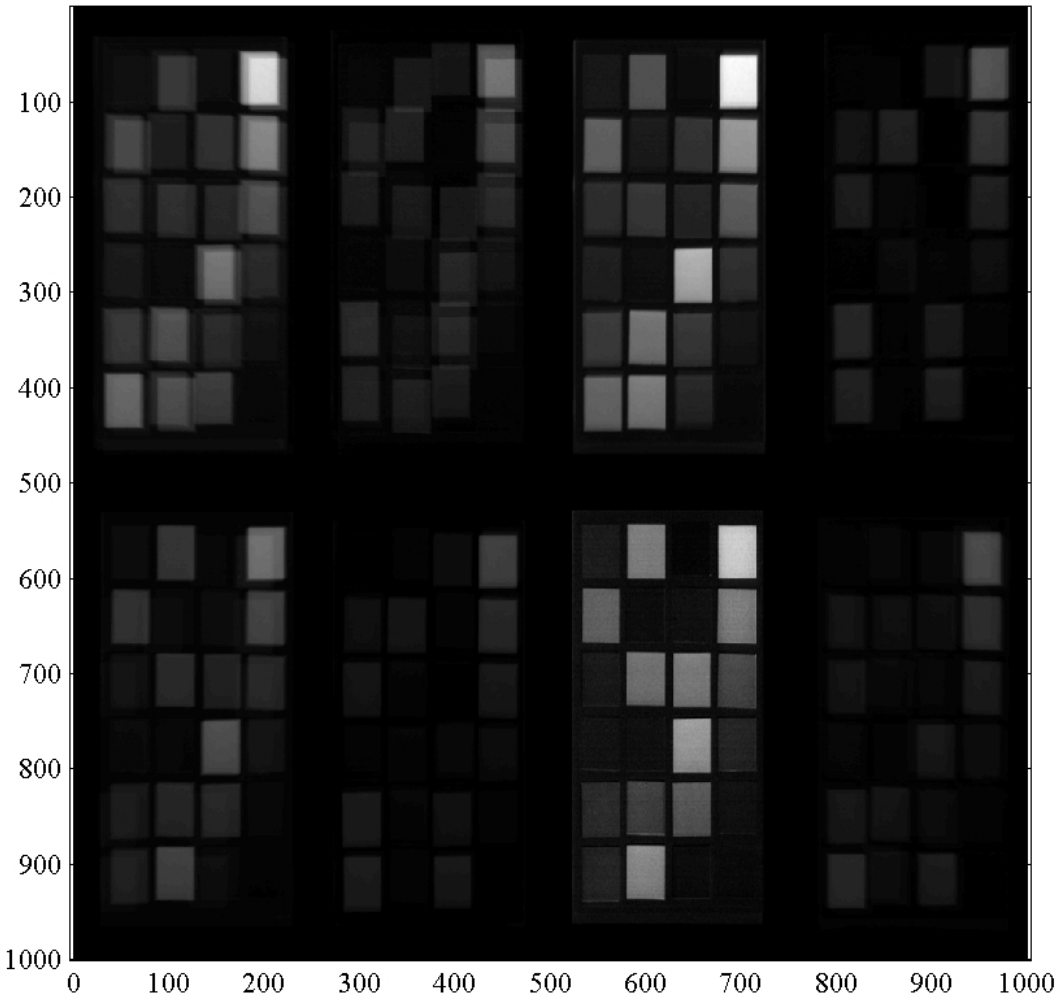

Figure 4.22: Simulated dispersion image of MacBeth colour chart.

third from the left in the top and bottom rows.

It is common to use the filtered back projection of the data as the initial estimate of $\lambda(b)$. The filtered back projection is given by low-pass filtering the back projection, $\mathbf{P^T n}$ of the data. With this starting point, the EM-algorithm is applied to the simulated data for 100 iterations. Binning the 32 spectral-samples into eight bins gives eight spectral images corresponding to 38nm spectral windows. The original set of images is shown in Figure 4.23, while the recovered images are shown in Figure 4.24. The absolute difference between these sets of images are shown in Figure 4.25.

The highest error values appear at edges within the image. The average relative error for the retrieval across the whole data set is shown in Figure 4.26; although the likelihood increases at each step, it can be seen from the plot that this is no guarantee that the quality of retrieval is increasing, and so some method of halting the iteration is desirable. The lowest average retrieval error is around 9.5%.

There are a number of techniques for improving on this standard EM algorithm by some form of regularisation. One potentially attractive route is the Bayesian method-

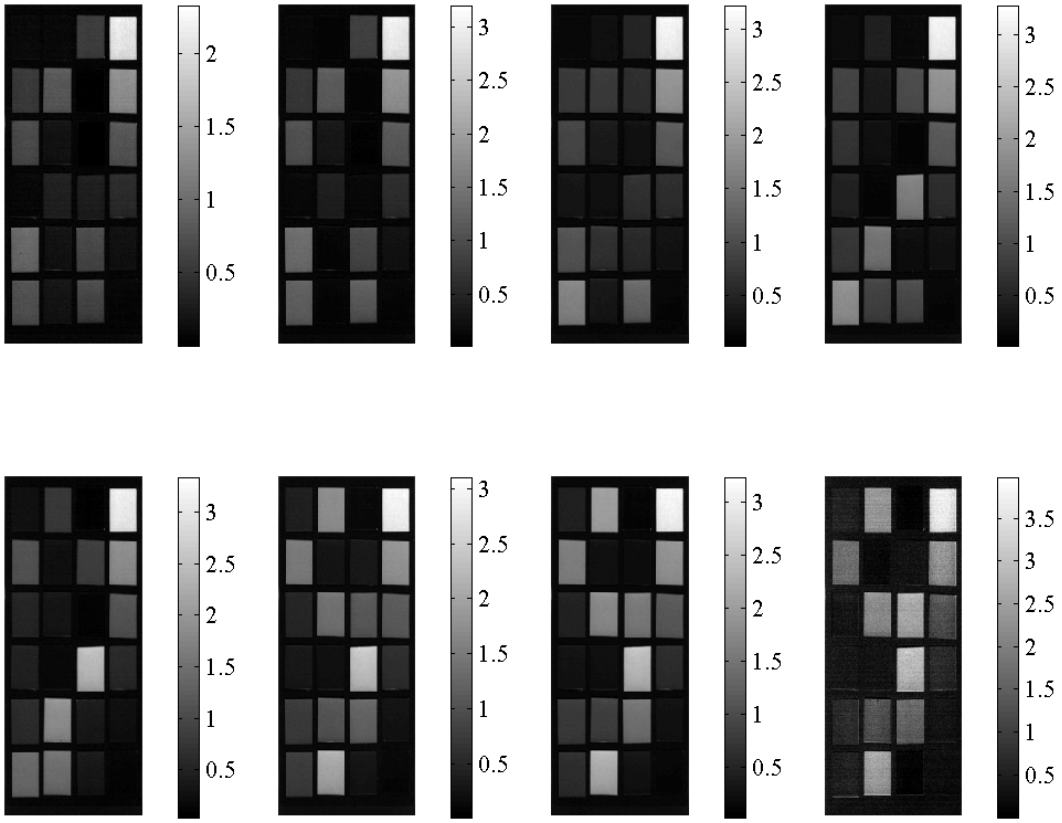

Figure 4.23: Actual spectral images.

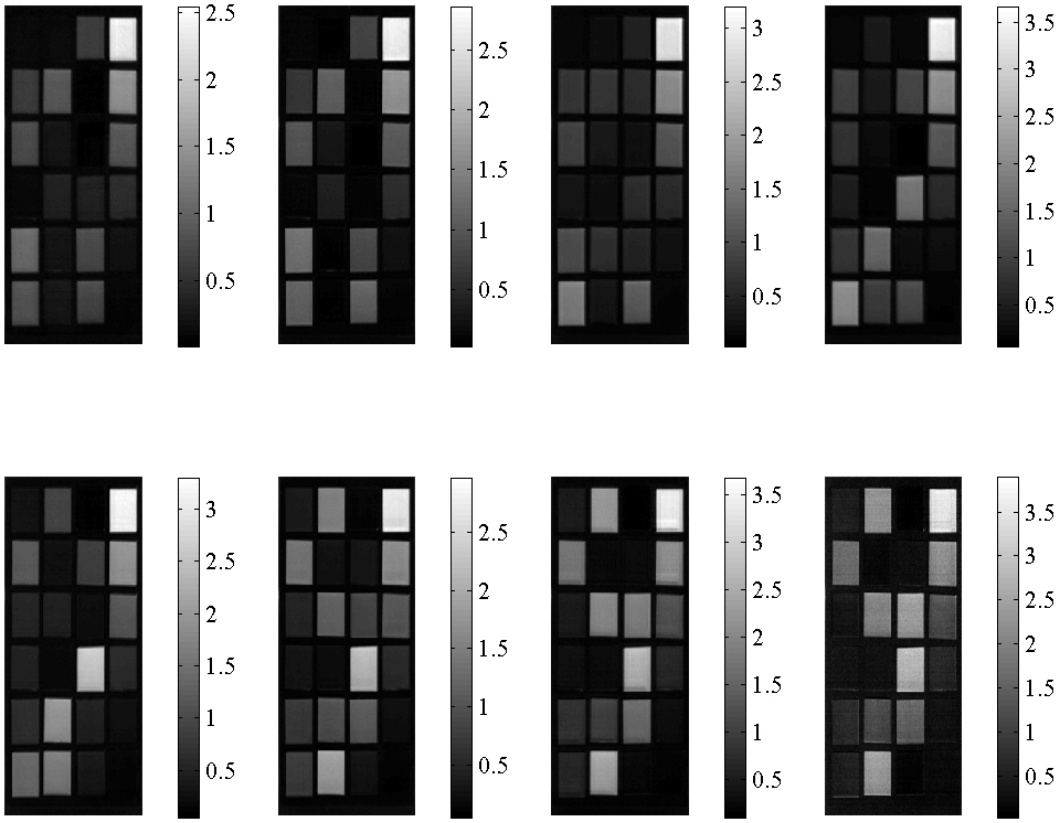

Figure 4.24: Recovered spectral images.

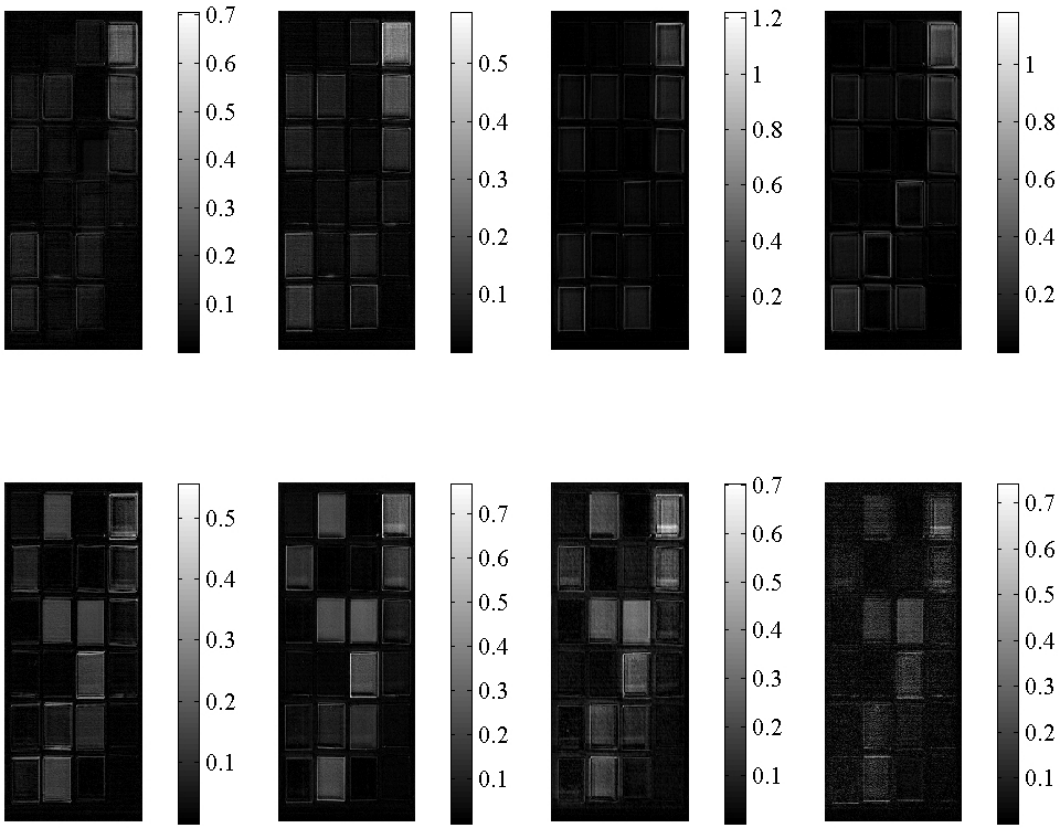

Figure 4.25: Absolute difference between recovered and original spectral images.

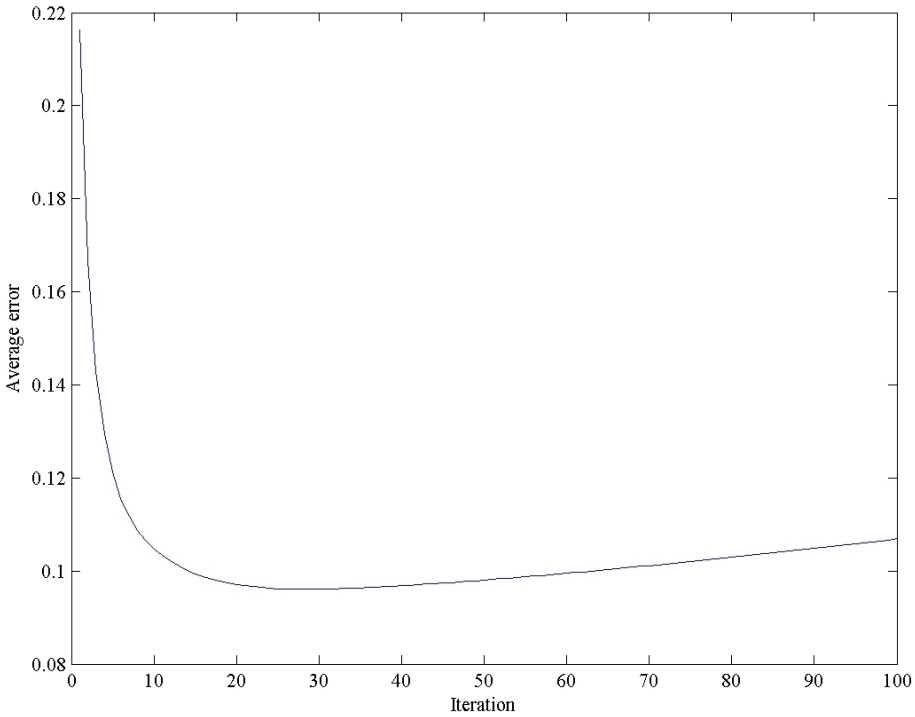

Figure 4.26: Error at each iteration.

ology which facilitates the use of *a priori* data. For IRIS the *a priori* data might be in the form of constraints on the smoothness of the spectra, or the spatial information in the sub-images with minimal dispersion may be used to constrain the spatial components of the retrieved data. This would suggest that attempting to find the maximum *a posteriori* (MAP) solution could be a suitable approach. A MAP solution using an expectation-maximisation algorithm has been proposed in [63]. The *a priori* information exists in the form of λ_A (an estimate of the mean discretised intensity distribution) and $\mathbf{S}_A$ (the covariance matrix which describes the estimated uncertainties and covariances on λ_A). Assuming, for simplicity, that the *a priori* probability of the measurements can be described by a Gaussian distribution leads to the *a posteriori* probability:

$$P\left(\mathbf{n}\left|\mathbf{m}\right.\right) = c\left[\prod_{d=1}^{D} e^{-\langle n(d)\rangle}\frac{\langle n\left(d\right)\rangle^{n(d)}}{n\left(d\right)!}\right]\cdot e^{-(\lambda-\lambda_A)^T\mathbf{S}_A(\lambda-\lambda_A)},$$

which is the conditional probability in Equation 4.3, multiplied by the *a priori* distribution; c is the normalisation constant. Derivation of the iterative MAP algorithm in [63] leads to maximisation of the log-expectation:

$$E = \left[\sum_{d=1}^{D}\sum_{b=1}^{B}\left(-p_{bd}\lambda_b^{(r+1)} + \frac{n_d p_{bd}\lambda_b^{(r)}}{\sum_{b=1}^{B} p_{bd}\lambda_b^{(r)}}\ln\left(p_{bd}\lambda_b^{(r+1)}\right)\right)\right] - \left(\lambda-\lambda_A\right)^T\mathbf{S}_A\left(\lambda-\lambda_A\right)\Big/2.$$

Differentiating this expression with respect to λ and setting equal to zero gives

$$-1 + \frac{\lambda_b^{(r)}}{\lambda_b^{(r+1)}}\sum_{d=1}^{D}\frac{n_d p_{bd}\lambda_b^{(r)}}{\sum_{b=1}^{B} p_{bd}\lambda_b^{(r)}} - \sum_{b=1}^{B} S_{bd}\left(\lambda_b - \left(\lambda_A\right)_b\right) = 0. \tag{4.5}$$

Unfortunately, an iterative EM-algorithm is only determined for the simple case of a diagonal *a priori* covariance matrix, meaning that the expected correlations in the spectral and spatial components cannot be easily encoded using this algorithm. In this simple case [63] shows that the solution of Equation 4.5 leads to the iterative scheme:

$$\lambda_b^{(r+1)} = \frac{1}{2s_b}\left\{s_b\left[\lambda_A\right]_b - 1 + \sqrt{\left(s_b\left[\lambda_A\right]_b - 1\right)^2 + 4s_b\lambda_b^{(r)}\sum_{d=1}^{D}\frac{n_d p_{bd}}{\sum_{b=1}^{B}\lambda_b^{(r)}p_{bd}}}\right\},$$

with λ_A chosen so that every element is equal to the mean emission from each box, which could be estimated from the measurements in the case of real data. The covariance matrix S_A is the identity matrix multiplied by σ_A. A plot of the minimum mean relative error as a function of σ_A is shown in Figure 4.27. The point, $\sigma_A = 0$, corresponds to a zero prior, which is equivalent to the maximum likelihood estimate. The minimum retrieval error occurs for $\sigma_A = 0.055$, giving an improvement of around 0.5% over the

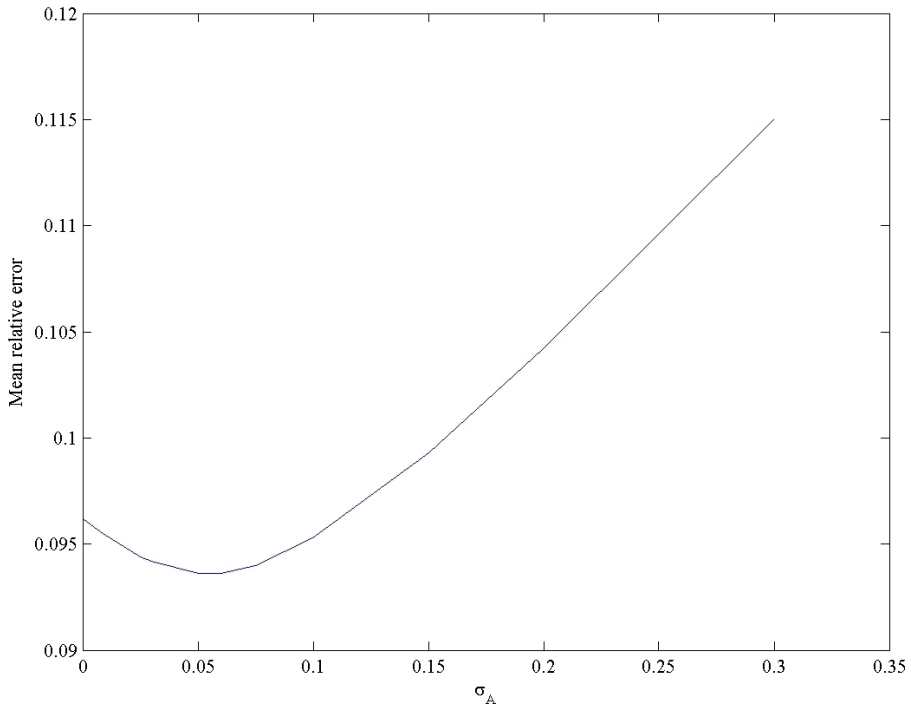

Figure 4.27: Mean relative error as a function of σ_A

maximum-likelihood estimate, a plot of the error as a function of iteration number for this value is shown in Figure 4.28. It can be seen that the additional regularisation term also reduces the penalty for not halting the iteration quickly enough.

It is likely that the retrieval can be further improved with the incorporation of more detailed *a priori* data via a MAP approach or using some other method. Optimisation of the bands, in terms of transmission and position at the detector, may also reduce the errors in the retrieval. As this offers a low-cost, general approach to countering birefringent dispersion, further work is warranted in future.

4.6 Conclusions

Chromatic aberration due to spectral birefringent dispersion in the Wollaston prism material is currently a significant issue in implementing IRIS systems. Out of the methods described in this chapter, the reflective Wollaston prism reported by Feldman et al. and described in Section 4.3.1 is perhaps the most attractive option: as well as being free from dispersion, it offers the potential for significant cost reductions in manufacture, and further opportunity for tuning the IRIS transmission functions. This option should be considered in detail, with full modelling of the change in the optical performance of the beam splitter with angle and wavelength. The designs determined in Section 4.3.2, for achromatic prisms which use combinations of materials, are another promising option for removing dispersion, although cost may again be prohibitive. The

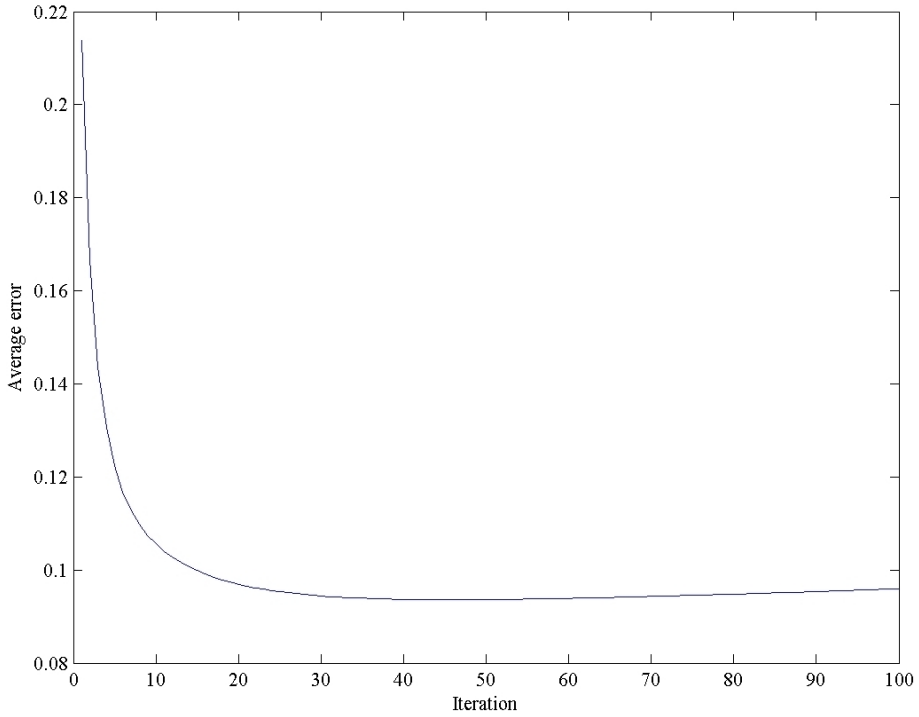

Figure 4.28: Error at each iteration for $\sigma_A = 0.055$

filter array used with the 16-channel broadband IRIS system, as described in Section 4.4 reduces but does not entirely remove the birefringent dispersion in that system; however, the technique may be still be useful when combined with other methods of reducing dispersion. A software based correction of the chromatic aberration also appears to be feasible, and as a low cost option for countering dispersion is another potentially promising area for future work.

Chapter 5

Comparison of spectral imaging techniques

Summary: This chapter introduces three figures of merit (FOM) for comparing spectral-imaging techniques. The first of these FOM is the number of degrees of freedom of signal; this indicates the number of independent quantities for which measurement decreases the uncertainty and is the upper limit on the number of unknowns that can be solved for in a linear system. The second FOM is the information content, which is related to the number of distinct states that measurement can detect; this is a useful FOM in applications for which spectral imaging is used for discrimination. The final FOM is the signal-to-noise ratio (SNR), which is commonly used to assess the performance of radiometric instruments. The difficulty in comparing the SNR in spectral imaging is that the measurement spaces are different for each type of system, so the measurement noise alone cannot be used. This problem is dealt with by transforming the measurements into a common space before comparing SNR. These FOM are used to compare IRIS to direct imaging spectrometers (DIS) and Fourier-transform imaging spectrometer (FTIS) in the general case.

5.1 Motivation for the comparison

It is not immediately obvious what type of spectral-imaging system is most suitable for a given task and there are many factors to consider in choosing a technology. Foremost, the instrument must be able to detect the signals of interest from the scene with some specified degree of accuracy; this will depend on the measurement errors but also on the efficiency of the optics and the form of the transmission functions. The transmission functions of IRIS are unconventional compared to direct imaging spectrometers (DIS) and Fourier-transform imaging spectrometer (FTIS) and so a comparison based purely on the transmission functions is useful to determine those scenarios in which IRIS may offer improved performance over conventional instruments. Aside from IRIS there are three other significant snapshot techniques: the computed tomography imag-

ing spectrometer (CTIS), coded aperture snapshot spectral imaging (CASSI), and direct reformating techniques (which are a class of DIS). CTIS and coded aperture spectral imaging are transform techniques in which each pixel records spectral and spatial information from a number of spatial elements; computational techniques are then used to recover the spectrum for each spatial element. Due to the scene-dependent and iterative nature of the reconstruction process for CTIS and CASSI, these instrument classes are not included in the analyses presented in this chapter.

If there is little foreknowledge of the scene then some general measure of the information content of the measurements may be used to compare IRIS with DIS and FTIS. The first two figures of merit (FOM) used in this chapter are such general measures: the number of degrees of freedom associated with the measurements, and the Shannon information content of the measurements. The analyses for these two FOM follow closely from [111] and have in the past been used to compare remote sounding instruments used to measure atmospheric carbon monoxide [112], and have also been applied to the optimisation of high-resolution spectral measurements [113].

The final general method of comparison is to determine the expected signal-to-noise of IRIS compared to that of DIS and FTIS. The signal-to-noise ratio is the ratio of the component of the measurement which is due to the signal compared to the component due to the measurement noise, and is commonly used to compare the performance of radiometric instruments. As DIS, FTIS and IRIS measurements all represent transformations of the spectrum to different lower dimensional spaces, how the SNR for these different instruments should be compared is not immediately clear. The approach used here is to transform the measurements into a common space; in this case wavenumber space is chosen as this is the form in which spectroscopic data is usually quoted and directly illustrates the transition energies associated with the photon emissions. A further justification for this choice is that if the spectrum is recovered accurately at each wavenumber then there has been no loss of information.

The chapter concludes with a comparison of the performance of the 8-channel retinal IRIS system with that of an 8-channel liquid crystal tunable filter, for the specific task of determining the oxygen saturation of blood vessels.

5.2 Spectral measurements

As described in Section 1.2, spectroscopic measurements are most often made using detectors that cannot resolve frequency on any scale smaller than the spectral bandwidth of the detector. To obtain finer measurements of the spectrum with these detectors, the spectral distribution of incident radiation must be modulated in some manner before the photons are integrated at the detector. The set of J spectroscopic measurements corresponding to a single spatial element can be written as a vector $\mathbf{y}$. The jth element

of this vector can be written as a discretised Fredholm equation, with k the wavenumber, and $x(k)$ the spectrum of incident radiation:

$$y^j = \int_{k_1}^{k_2} \tau^j(k)\, x(k)\, dk. \tag{5.1}$$

$\tau^j(k)$ is known as the transmission function of the jth channel.

Spectroscopic techniques are classified as direct or transform according to the general form of $\tau^j(k)$ for the device. For direct techniques each band records only a narrow contiguous range of wavelengths within the bandwidth of the instrument, whereas for transform techniques each band typically records modulated radiation from across the entire bandwidth. The most common type of transform instruments are the Fourier transform spectrometers (FTS) for which the transmission functions are cosinusoidal in form. IRIS is, in general, a transform instrument, although with suitable optimisation the bands can be made to approach those of direct spectrometers.

It is presumed that the spatial information will be equally well sampled across all classes of instrument and therefore a comparison can be made purely on the nature of the spectral sampling.

5.2.1 Band types

For the sake of simplicity the comparison will be carried out over the same fixed bandwidth for all systems, with the same number of bands. The detector characteristics are assumed to be the same across all instruments and all bands, and the quantum efficiency is assumed constant for all wavelengths. The sets of idealised bands listed in Table 5.1 will be used to represent the different classes of instrument.

The analyses presented in this chapter are carried out for the best case scenario for IRIS, that is, when the measurements are made in polarised light and one is only interested in the field of view offered by the IRIS system. In this case the exposure time for IRIS is J times longer than for each frame captured by a filter based DIS or TFTIS, for the same total acquisition time. IRIS must be scanned J times to cover the same field of view offered by these systems, in this case the snapshot advantage is diminished. For SFTIS and dispersive DIS the entrance aperture is usually a single slit so that each frame records only a single column (or row) of the spatial data. For lower spectral resolutions a low power dispersive element can be used in conjunction with multiple slits to record the entire data cube in the same number of frames as a filter based DIS.

The efficiency for all instrument types in polarised light will be assumed to be 98%; in theory, this is achievable for IRIS, for filter based temporal DIS, and for polarising FTIS such as the Wollaston prism based system reported in [10]. For Michelson- or Sagnac-

Class	Transmission funtions
Direct	$\tau^i(k) = \tau_{max} \sqcap \left(\frac{M\left(k_o^i - k\right)}{2\Delta k}\right)$, $k_o^i = k_1 + (2i-1)\frac{\Delta k}{2M}$
Fourier Transform [a]	$\tau^i = \tau_{max} \cos^2\left(\frac{\pi(i-1)k}{2k_2}\right)$, $k_2 = 2k_1$
IRIS [b]	$\tau^i = \tau_{max} \prod_{n=1}^{\log_2 M} \cos^2\left(\frac{kB_n d_n - g_n^i \pi}{2}\right)$, $g_n^i \in \{1,0\}$, $g_{\cdot}^p = g_{\cdot}^q \Leftrightarrow p = q$

[a] Fourier transform is step and lock system with the step size chosen to be half the smallest wavelength

[b] B_n and d_n are the birefringence and thickness of the nth retarder in the IRIS assembly, these are chosen so that the bands approach those of a direct instrument

Table 5.1: Band types

based FTIS the transmission in polarised light is limited to 50%; for some grating-based dispersive DIS and filter based DIS such as liquid crystal tunable filters, the efficiency in polarised light may be significantly lower than this. In unpolarised light IRIS and FTIS are limited to 50% transmission, whereas the efficiency of DIS are largely unaffected (with the exception of Lyot- and Solc-type filters, such as the LCTF, which are limited to 50%).

Figures 5.1 to 5.3 show plots of the transmission functions used in the analyses of this chapter for DIS, IRIS and FTIS with $J = 16$, with a passband from 375nm to 750nm. The IRIS bands have been chosen so that the impulse response is as close to the equivalent direct instrument as possible.

5.2.2 Spectral imaging as linear problem

Often the integral equations in the forward problem are well approximated as a set of linear equations,

$$\mathbf{y} \approx \mathbf{T}\mathbf{x} + \epsilon. \tag{5.2}$$

with the matrix $\mathbf{T}$ having the normalised transmission functions τ as its rows (Figure 5.4), discretised with high enough resolution that the error due to linearisation is negligible so that Equation 5.2 can be regarded as an equality. The spectrum is represented as a vector $\mathbf{x}$, and ϵ is a vector representing the measurement error.

Before the measurement is made, the knowledge of the spectrum can be described by a vector $\mathbf{x_p}$ representing the a priori best estimate of x, and a covariance matrix $\mathbf{S_{x_p}}$ representing the uncertainty on this estimate. The a priori measurement vector is thus $\mathbf{y_p} = \mathbf{T}\mathbf{x_p} + \bar{\epsilon}$, where $\bar{\epsilon}$ is the expectation value of the noise; the covariance of the noise is $\mathbf{S}_\epsilon$. The a priori covariance matrix corresponding to $\mathbf{y_p}$, is $\mathbf{S_{y_p}} = \mathbf{T}\mathbf{S_{x_p}}\mathbf{T^T} + \mathbf{S}_\epsilon$.

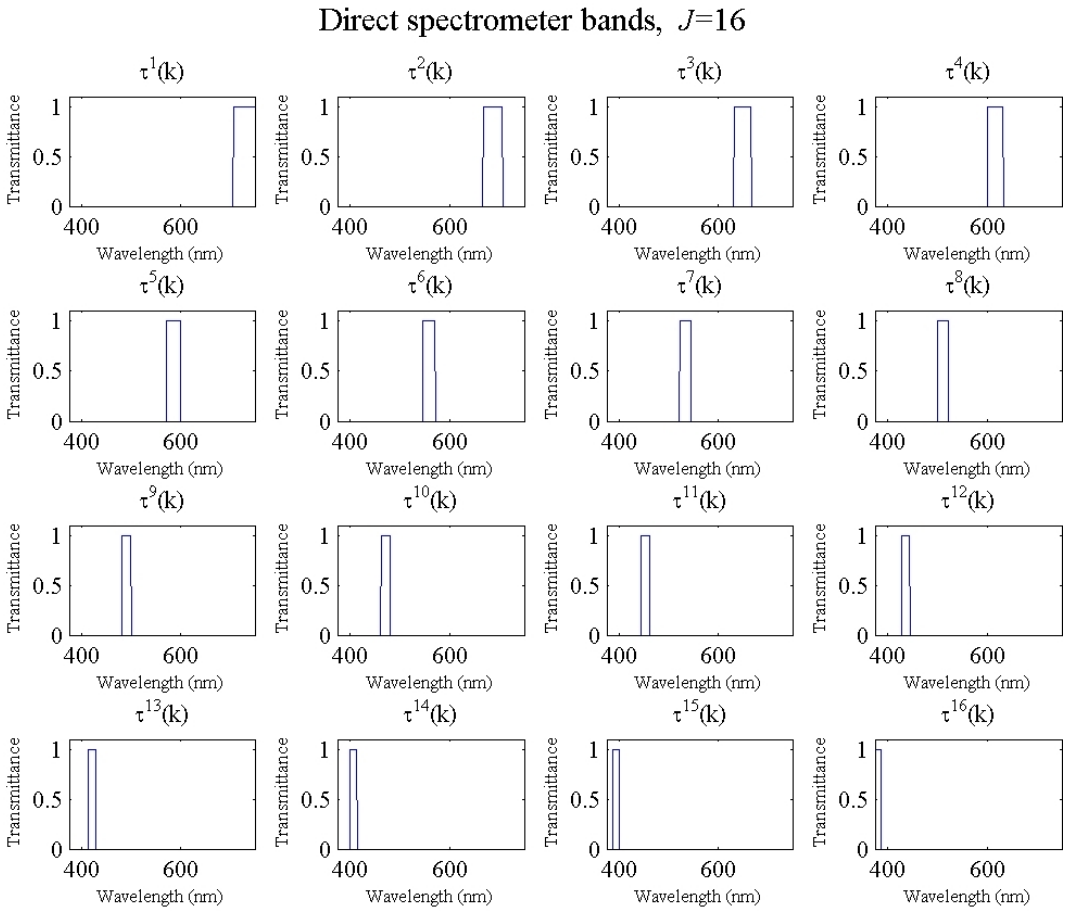

Figure 5.1: Direct spectrometer bands as a function of wavelength, J=16

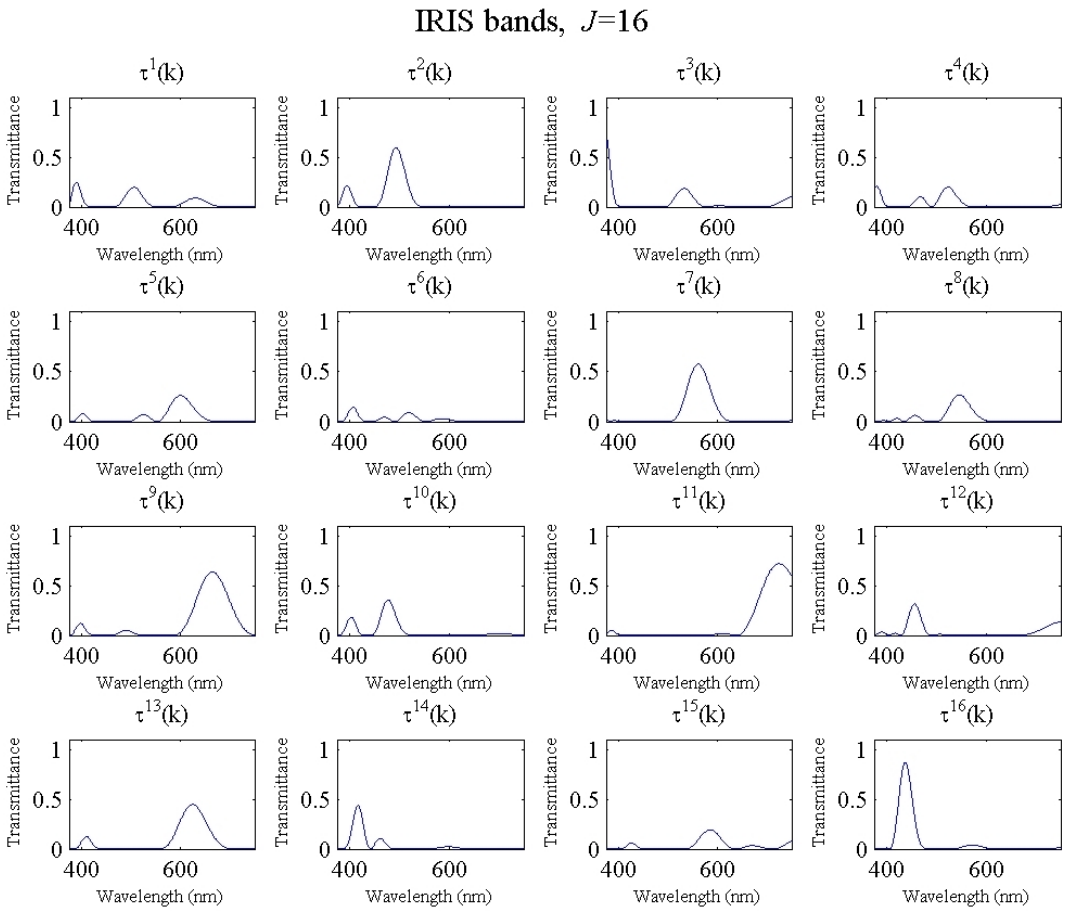

Figure 5.2: IRIS bands as a function of wavelength, J=16

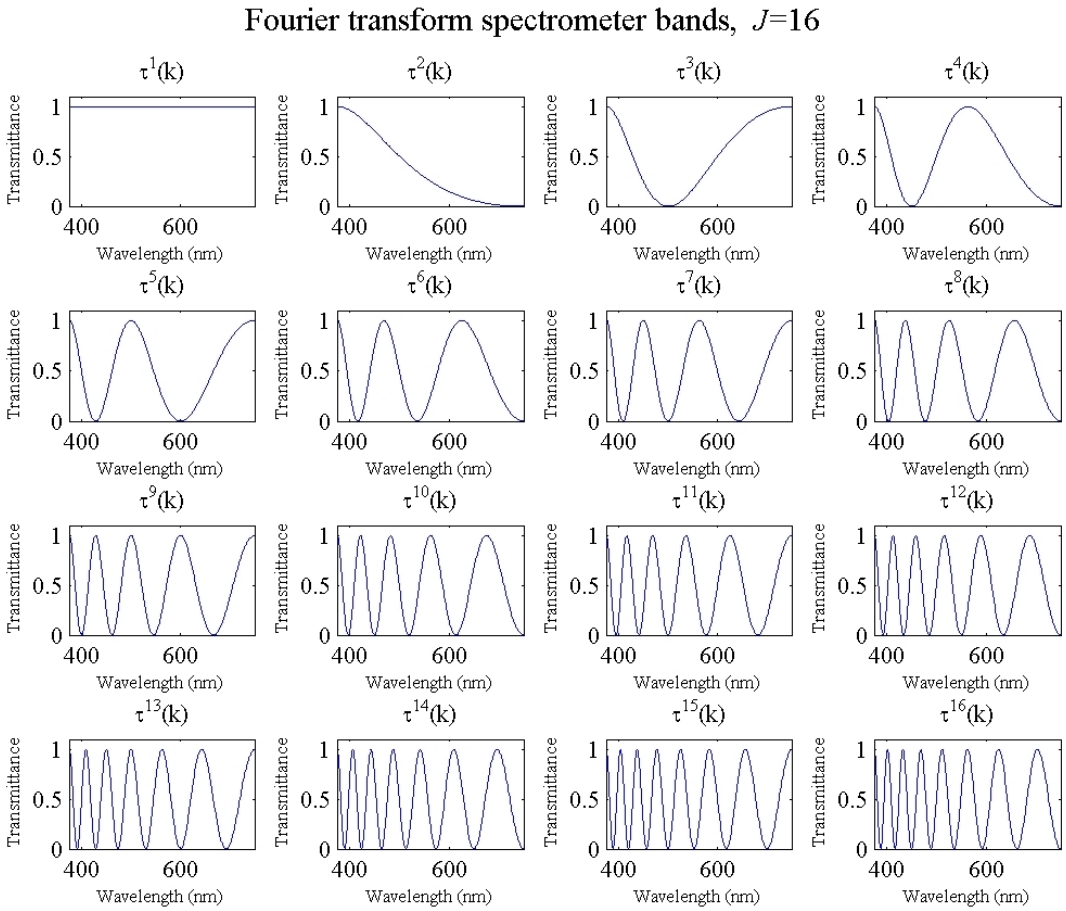

Figure 5.3: Fourier transform spectrometer bands as a function of wavelength, J=16

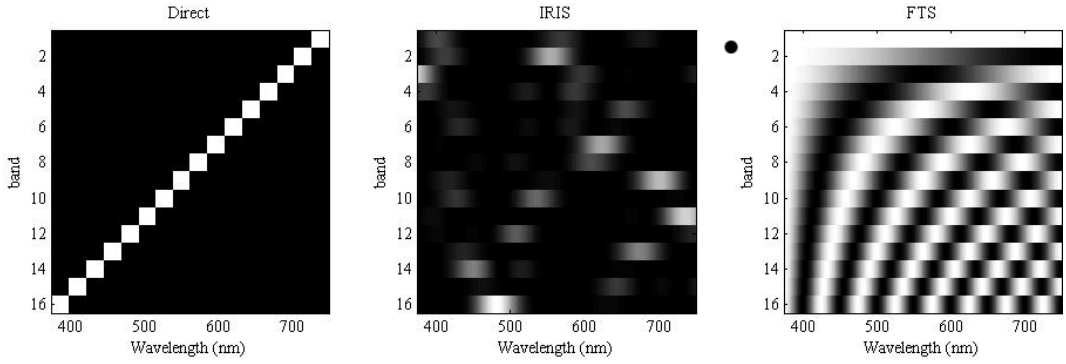

Figure 5.4: Visualisation of the transmittance matrices - 16 x 480 images

Each spectral measurement is each likely to be the sum of many quasi-monochromatic sources, each detected with Poisson statistics or according to some other finite variance distribution. From the central limit theorem the measurement variable is therefore likely to be well described by a Gaussian distribution. It is also common to model the noise ϵ as a Gaussian distributed variable with mean $\mathbf{0}$ and a covariance matrix $\mathbf{S}_\epsilon$. It is assumed in this chapter that the a priori and a posteriori distributions of x are also Gaussian.

5.2.3 Noise and uncertainty

Photon-counting detector arrays function through electronic measurement of the number of electrons generated in each detector element by source photons, over the exposure time for that element. The currents, or voltages, generated for each element are then digitised and arranged to form an output image. Photon detection is probabilistic in nature and therefore there is an inherent source of error in the detection process. Photon counting is often assumed to be well modeled as a Poisson process, meaning that the measured number of photons emanating from the object will be distributed around some unknown expectation value with standard deviation equal to the square root of this expectation value. Sources of photons other than the object, such as stray light or thermal emission from the optical system, will add to the measured number of photons. In many cases the expectation value of the number of electrons due to these unwanted sources can be measured and then subtracted from subsequent measurements. However, the error associated with the photon-counting process will remain.

There are additional sources of error in the detector and readout system. A primary source of noise is the accumulation of electrons generated via thermal motion in the bulk of the detector rather than via the photoelectric effect. Again it is common to measure the expectation value of these thermally generated electrons and subtract this average from any measurements, leaving the error associated with electron counting. The read noise is the term given to the total error introduced in the electronic measurement and digitisation process.

Assuming the signal and noise sources outlined above generate electrons at a constant rate over the exposure time of the detector, and that they can be approximated as Poisson processes, the measurement signal-to-noise ratio for a single detector can be written as

$$SNR = \frac{N_s t}{\sqrt{(N_s + N_o + N_d)\, t}}, \tag{5.3}$$

where N_s is the number of electrons per second generated by source photons, N_o is the number of electrons per second generated by other sources of photons, and N_d is the number of electrons per second generated by the detector and readout system. In a

spectral imaging system, N_s for the jth band is given by

$$N_s^j = E \int_{k_a}^{k_b} \eta(k)\, \tau^j(k) x(k)\, dk, \tag{5.4}$$

where $x(k)$ is the number of photons from the source, which is assumed to be constant across E, the system *étendue* for a single pixel. N_o is in general a complicated term, which depends on the specific optical and mechanical design of the system. N_d is also difficult to characterise in general, although thermal noise is a Poisson process.

Two limiting cases for the denominator in Equation 5.3 are of interest: measurements whose accuracy is limited by the photon shot noise of the measurements $N_s t \gg ((N_o + N_d))$, and measurements for which the photon shot noise is negligible compared to other noise sources $N_s t \ll ((N_o + N_d)\, t)$. Both of these limits shall be considered in this chapter with the noise levels chosen in both cases to give an effective measurement signal-to-noise ratio of 4 at $t = 1$.

5.3 Degrees of freedom and information content

5.3.1 Degrees of freedom

When J independent measurements are made, it should in theory be possible to determine J independent quantities from the data; however, in the presence of noise this number may be less than J. The actual number of independent quantities that can be recovered in a set of measurements is known as the number of degrees of freedom (DOF) for signal and is important in applications such as spectral unmixing which involve solving a set of simultaneous equations.

One indication of the DOF is the rank of the matrix $\mathbf{T}$ in Equation 5.2, which gives an upper limit on the number of free parameters in the system. However, this does not take into account the effects of measurement error. Determination of the DOF including measurement error can be made by analysis of the covariance matrix $\mathbf{S_y}$ of the measurements $\mathbf{y}$. This comparison is complicated by the fact that $\mathbf{S_y}$ is not in general diagonal. Determination of the DOF in the presence of noise is made simpler by first transforming to a basis in which the covariance matrices $\mathbf{S}_{\tilde{x}_p}$, and $\mathbf{S}_{\tilde{\varepsilon}}$ are identity matrices;

$$\tilde{\mathbf{T}} = \mathbf{S}_{\varepsilon}^{-1/2} \mathbf{T} \mathbf{S}_x^{1/2}, \tag{5.5}$$

where

$$\mathbf{S}^{1/2} = \mathbf{L} \Lambda^{1/2} \mathbf{L}^T \tag{5.6}$$

is the positive definite matrix square root of the matrix $\mathbf{S}$, with $\mathbf{L}$ a matrix of it eigen-

vectors, and $\mathbf{\Lambda}$ a diagonal matrix with the corresponding eigenvalues along the diagonal. $\mathbf{S}^{-1/2}$ is the matrix-inverse of $\mathbf{S}^{1/2}$. Defining $\tilde{\mathbf{x}} = \mathbf{S}_x^{-1/2}\mathbf{x}$, $\tilde{\varepsilon} = \mathbf{S}_\varepsilon^{-1/2}\varepsilon$, and $\tilde{\mathbf{y}} = \mathbf{S}_\varepsilon^{-1/2}\mathbf{y}$, Equation 5.2 can be written in this transformed basis as

$$\begin{aligned}\tilde{\mathbf{y}} = \mathbf{S}_\varepsilon^{-1/2}\mathbf{y} &= \left(\mathbf{S}_\varepsilon^{-1/2}\mathbf{T}\mathbf{S}_x^{1/2}\right)\mathbf{S}_x^{-1/2}\mathbf{x} + \mathbf{S}_\varepsilon^{-1/2}\varepsilon \\ &= \tilde{\mathbf{T}}\mathbf{S}_x^{-1/2}\mathbf{x} + \mathbf{S}_\varepsilon^{-1/2}\varepsilon \\ &= \tilde{\mathbf{T}}\tilde{\mathbf{x}} + \tilde{\varepsilon}.\end{aligned}$$

Although $\mathbf{S}_{\tilde{x}a}$, and $\mathbf{S}_{\tilde{\varepsilon}}$ are identity matrices, $\mathbf{S}_{\tilde{y}}$, is not diagonal in this basis, but can be made so by transforming $\tilde{\mathbf{y}}$ using the matrix of left singular vectors $\mathbf{U}$ from the singular value decomposition of $\tilde{\mathbf{T}} = \mathbf{U}\Lambda\mathbf{V}^T$:

$$\mathbf{y}' = \mathbf{U}^T\tilde{\mathbf{y}} = \mathbf{U}^T\tilde{\mathbf{T}}\tilde{\mathbf{x}} + \mathbf{U}^T\tilde{\varepsilon}. \tag{5.7}$$

Defining $\mathbf{x}' = \mathbf{V}^T\tilde{\mathbf{x}}$ (so that $\tilde{\mathbf{x}} = \mathbf{V}\mathbf{x}'$) and $\varepsilon' = \mathbf{U}^{\mathbf{T}}\tilde{\varepsilon}$, this can be written as

$$\mathbf{y}' = \mathbf{U}^T\tilde{\mathbf{T}}\mathbf{V}\mathbf{x}' + \varepsilon' = \Lambda\mathbf{x}' + \varepsilon'. \tag{5.8}$$

As $\mathbf{U}^T\mathbf{U}$ and $\mathbf{V}^T\mathbf{V}$ are identity matrices, in this second transformed basis $\mathbf{S}_{\epsilon'}$ and $\mathbf{S}_{\mathbf{x}'}$ are identity matrices, while the covariance matrix of $\mathbf{y}'$ is $\mathbf{S}_{y'} = \mathbf{\Lambda}^2 + \mathbf{I}_J$, with $\mathbf{I}_J$ a $J \times J$ identity matrix. The eigenvalues of $\tilde{\mathrm{T}}\tilde{\mathrm{T}}^{\mathbf{T}}$ which are greater than unity therefore correspond to elements of $\mathbf{y}'$ which carry more information about the signal than they do about the noise. The fractional number of degrees of freedom for signal is given by

$$\mathrm{tr}\left(\mathbf{\Lambda}^2\left(\mathbf{\Lambda}^2 + \mathbf{I}_J\right)^{-1}\right). \tag{5.9}$$

Figure 5.5 shows the number of DOF for signal as a function of the exposure time for the three classes of instrument with $\tau_{\max} = 1$. $\mathbf{S}_\epsilon$ is diagonal and the mean SNR of the measurements defined as $\mathrm{tr}\left(\mathrm{T}\mathbf{S}_x\mathrm{T}^T\mathbf{S}_{\tilde{\varepsilon}}^{-1}\right)/M$. For the shot-noise limit it can be seen that direct instruments offer a greater number of DOF for signal than FTIS and IRIS at all SNR, while IRIS measurements offer a greater number of DOF for signal than FTIS. If shot noise is not dominant then IRIS offer a greater number of DOF for signal than direct systems and FTIS at all SNR, while FTIS outperform direct systems.

Figure 5.6 shows the number of DOF for signal for different numbers of bands, with a total exposure time of $t = 100$. In all cases, the number of DOF is close to the number of bands, although IRIS begins to under perform the other systems when the number of bands is greater than 64, particularly in the case where shot noise is not dominant.

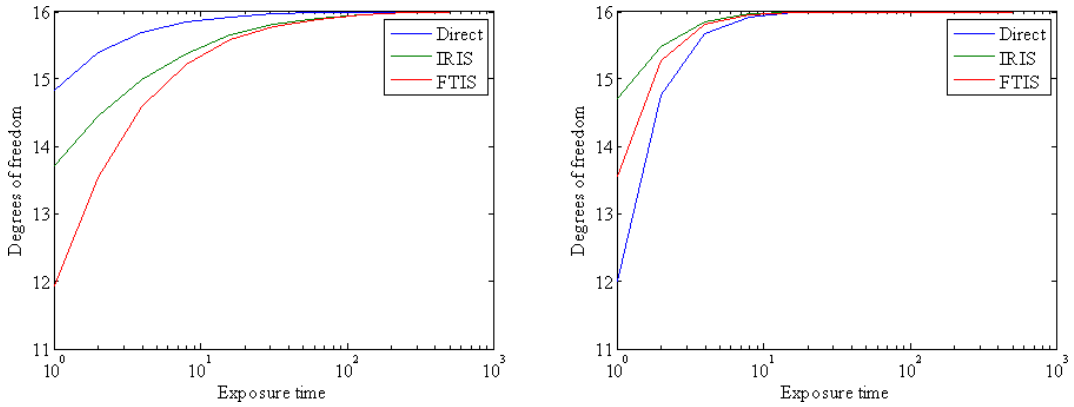

Figure 5.5: Degrees of freedom for signal for $J = 16$ as a function of total exposure time for the case that (a) the noise is dominated by shot noise (b) when the noise is dominated by sources other than shot noise.

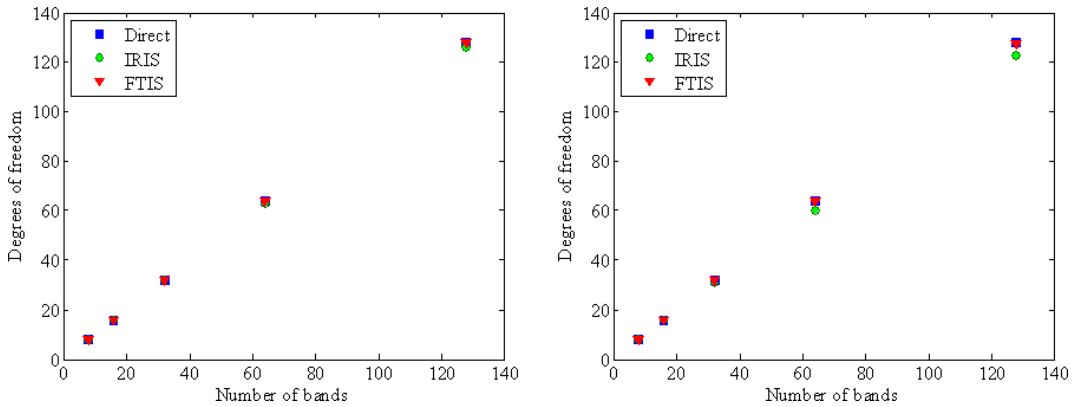

Figure 5.6: Degrees of freedom for signal for different numbers of bands with a total exposure time of $t = 100$.

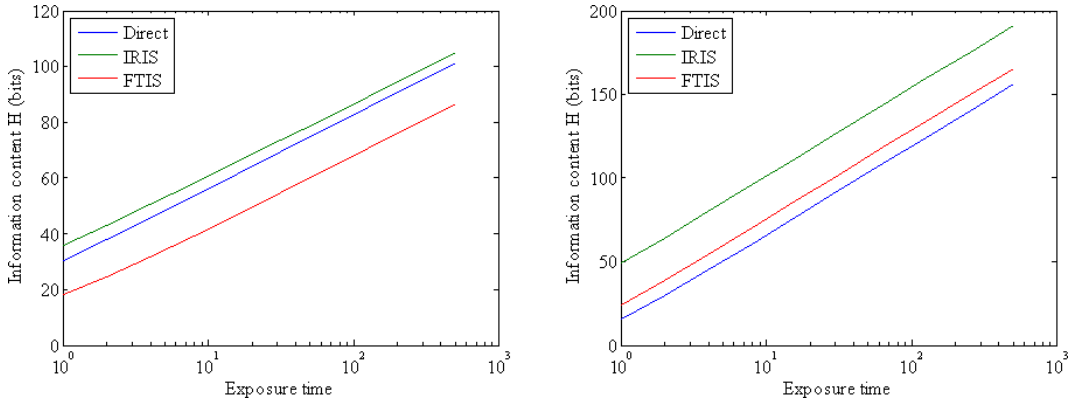

Figure 5.7: Information content for $J = 16$ as a function of total exposure time for the case that (a) the noise is dominated by shot noise (b) when the noise is dominated by sources other than shot noise.

5.3.2 Shannon information content

Discrimination is a major application for spectral imaging and is possible if the instrument response is consistently different for dissimilar inputs. A measure of the ability of a set of measurements to distinguish between different states is the Shannon information content of the measurement. The information content can also be considered qualitatively as the factor by which knowledge of the spectrum is improved by the measurements. Information content is the difference in the entropy of the prior and posterior distributions of $\mathbf{y}$; both distributions are assumed to be Gaussian.

The covariance prior to the measurement is given by $\mathbf{S}_{y_p} = \mathrm{T}\mathbf{S}_{x_p}\mathrm{T}^T + \mathbf{S}_\epsilon$ and afterward is $\mathbf{S}_y = \mathbf{S}_\epsilon$. The information content of the measurements is therefore given by

$$H = \log_2 \sqrt{\left|\mathbf{S}_y^{-1}\mathbf{S}_{y_p}\right|} = \log_2 \sqrt{\left|\tilde{\mathrm{T}}\tilde{\mathrm{T}}^T + \mathbf{I}_M\right|}. \tag{5.10}$$

Figure 5.7 illustrates the difference in information content between the different types of bands for $J = 16$, with IRIS giving higher information content in both noise limits. Again the direct measurements outperform FTIS in the shot-noise limit, while the opposite is true if other sources of noise dominate. It can be seen that the difference in information content between the band types stays reasonably constant. The actual amount of information available in the measurements may be limited by the precision of the detector.

Figure 5.8 shows entropy for different numbers of bands for a total exposure time of $t = 100$. In both noise limits the IRIS measurements have a higher information content that the DIS and FTIS measurements for all numbers of bands, with DIS offering a greater information content than FTIS in the shot-noise limit, while the converse is true in the non shot-noise limit.

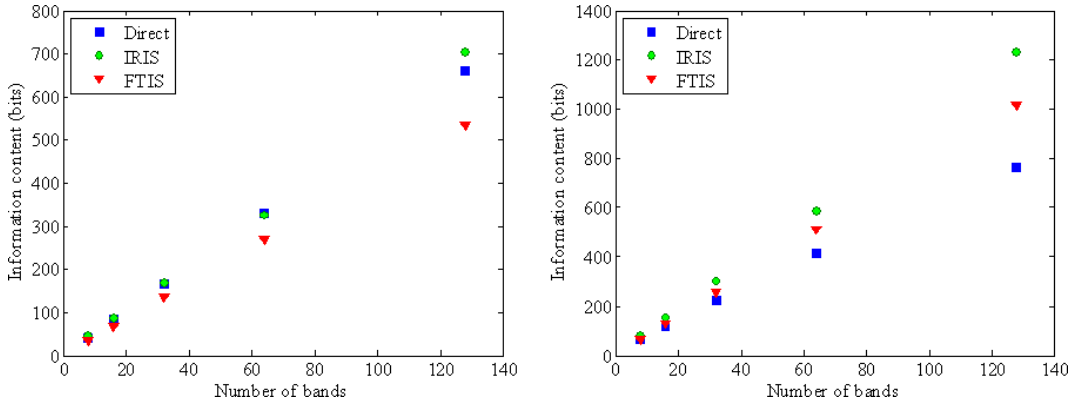

Figure 5.8: Information content for different numbers of bands with a total exposure time of $t = 100$.

5.4 Comparison of signal-to-noise ratios

A comparison of the signal-to-noise ratio of direct and FTIS was made in [35]. This analysis showed the expected signal-to-noise ratio advantage of FTIS over DIS, when the measurement noise is limited by the detector (Fellgett's advantage). It was further shown that, when the noise is instead limited by the shot-noise associated with photon counting, DIS offer a greater signal-to-noise ratio than FTIS. In this section, IRIS is compared to direct imaging spectrometers and FTIS by transforming the measurement error ϵ to wavenumber space; to do this the measurement process in treated as a linear system and the recovery of the spectrum as an inverse problem.

Measurements of the continuous spectrum $x(k)$ can be regarded as transforms to lower dimensional spaces. The comparison of signal-to-noise ratio is therefore made easier by transforming each set of measurements to a common space; an obvious choice, of space for spectral measurements is wavenumber space. There is no unique transformation to achieve this and in this case the Backus-Gilbert method is used [111]; this is the optimal-resolution transformation and is an intuitive choice, as the goal is to accurately recover the spectrum. The Backus-Gilbert transformation can be represented as a multiplication of the vector of measurements $\mathbf{y}$, by a matrix $\mathbf{D}$ so that the recovered spectrum is given by

$$\hat{\mathbf{x}} = \mathbf{D}\mathbf{y} = \mathbf{D}\mathrm{T}\mathbf{x} + \mathbf{D}\epsilon. \tag{5.11}$$

The matrix $\mathbf{A} = \mathbf{DT}$ is known as the averaging kernel matrix and in the absence of noise represents the mapping of the input spectrum to the recovered spectrum. The fidelity of the system depends on the averaging kernels and the domain of spectra being measured. Assuming the fidelity is adequate for all band-types, the comparison can be made purely by calculating how the measurement error propagates to the retrieved spectrum. From Equation 5.11, this error is $\mathbf{D}\epsilon$, so that $\mathbf{S}_{\hat{x}} = \mathbf{D}\mathbf{S}_{\epsilon}\mathbf{D}^{T}$. Figures 5.9 to 5.11 show the averaging kernel matrices for $J = 16$.

If the *étendue* for a single pixel is assumed to be the same for all classes of instrument,

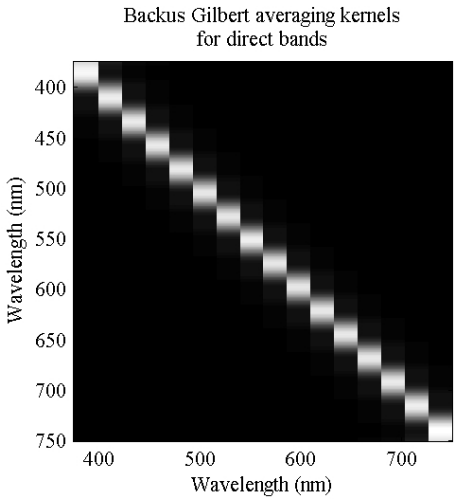

Figure 5.9: Averaging kernels for direct bands, with $J = 16$.

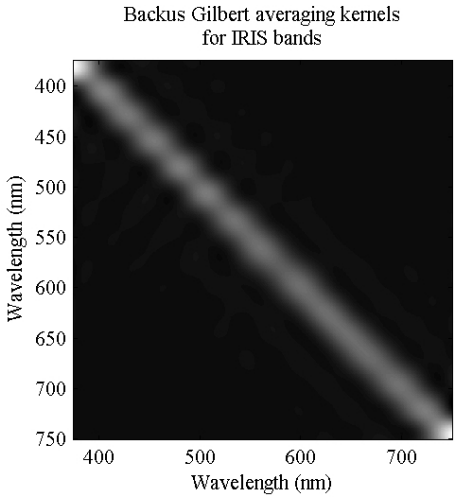

Figure 5.10: Averaging kernels for IRIS bands, with $J = 16$.

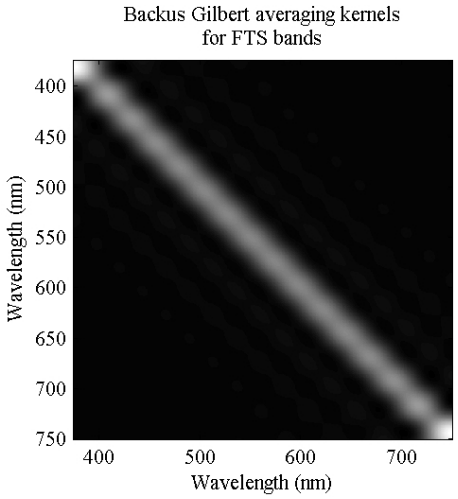

Figure 5.11: Averaging kernels for Fourier transform spectrometer bands, with $J = 16$.

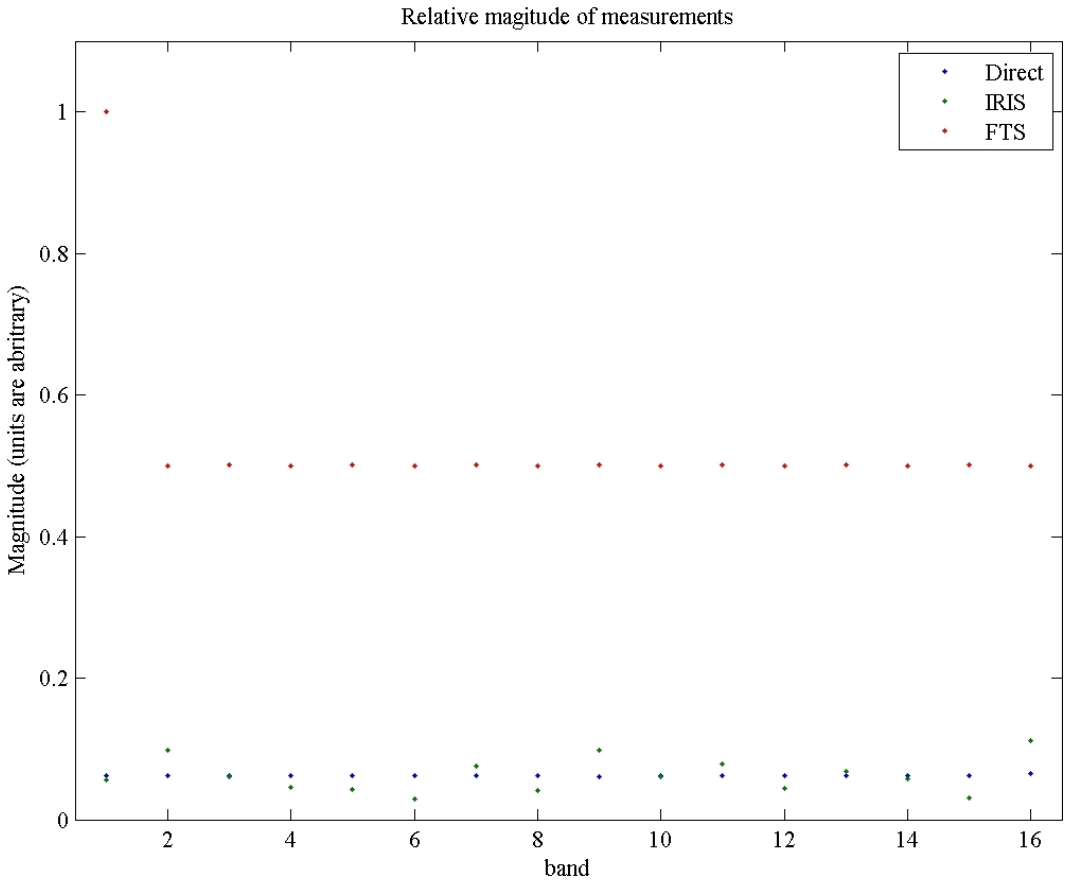

Figure 5.12: Magnitude of measurements for grey source, with $J = 16$.

the source spectrum is taken as grey and the quantum efficiency of the detector constant, then $N_s^i = c\int_{k_a}^{k_b} \tau^i(k)dk$, with the same constant of proportionality c across all bands and all instrument types. Figure 5.12 shows the relative magnitude of N_s^i; the magnitude of IRIS and the direct measurements will on average be $2/J$ times smaller than the FTS measurments ($1/J$ smaller than the zero path difference FTS measurement), all other things being equal.

For the case that the shot-noise of the source limits the SNR, the standard deviation of the error on each measurement is $\sqrt{N_s^i}$. The error on the retrieved spectra can be therefore be compared by comparing the diagonal elements of $\mathbf{DS}_\epsilon\mathbf{D}^T$, where S_ϵ is a diagonal matrix with $S_\epsilon^{ii} = \int_{k_b}^{k_b} \tau^i(k)\,dk$. Figure 5.13 shows plots of the relative retrieval error for $J = 16$.

In the case that the other terms, N_d and N_o, limit the measurement SNR, the error is presumed equal in all channels and uncorrelated between measurements. The retrieval error in this limit may be compared by comparing diagonal elements of $\mathbf{DD}^T$ to determine the relative error on the retrieval. Figure 5.13 (b) shows the relative retrieval error for $J = 16$. Figures 5.14 show the rms error as a function of the number of bands. As can be seen, IRIS outperforms DIS and FTIS when the number of bands is ≤ 16. For larger numbers of bands the highest SNR is offered by DIS in the shot-noise limited

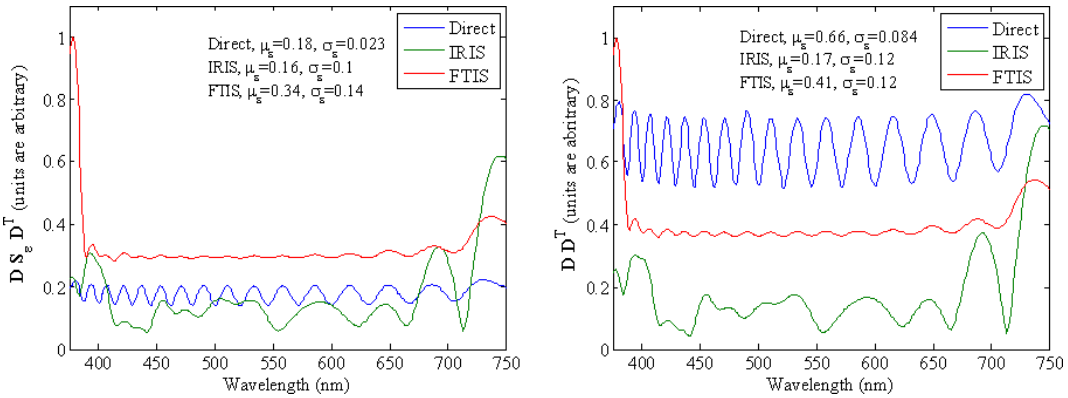

Figure 5.13: Relative error on retrieved spectrum for Backus-Gilbert method, in (a) the shot-noise limited case, and (b) the non shot-noise limited case.

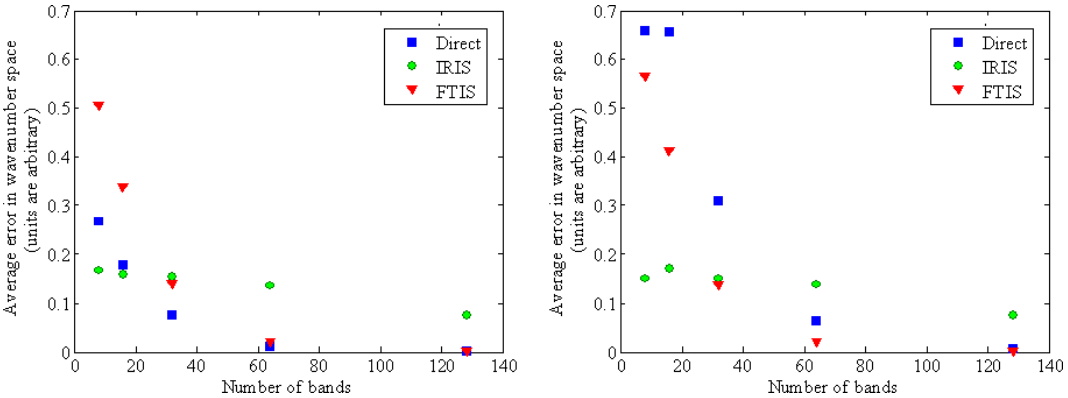

Figure 5.14: Relative error on retrieved spectrum for Backus-Gilbert method, in (a) the shot-noise limited case, and (b) the non shot-noise limited case.

case, and FTIS when other sources of noise dominate. This suggests that, even with its snapshot capability, for larger numbers of bands IRIS systems will not outperform other spectral imaging technologies over the same total integration time.

It should be noted that the chances of the set of IRIS bands being the optimal set diminishes as the number of bands increases, due to the exponentially increasing size of search space in the optimisation of the IRIS bands. The merit function used in the optimisation might also be suboptimal. It should also be noted that inversion of the measurements to obtain the spectrum in wavenumber space is not a necessary step in many spectral imaging applications. Often one is only interested in the abundances of different chemical species or discrimination between different materials and the preceding analysis is not proof of the performance of IRIS in these applications. Accurate recovery of the spectrum would, however, indicate that there is little loss of information in the measurement process.

5.5 Comparison of LCTF and retinal IRIS for the purpose of blood oximetry

As a practical example, the 8-channel retinal IRIS system (for which the bands are shown in Figure 5.15) has been compared to 8 bands of a *CRi* VariSpec LCTF (for which the bands are shown in Figure 5.16) for the purposes of retinal oximetry. The application of IRIS to retinal oximetry is discussed in more detail in Section 6.3. Both of these instruments can be fitted to a fundus camera, enabling imaging of the retina using flash illumination via the pupil. This illumination scatters within the retina and exits the eye through the pupil. Figure 5.17 illustrates the assumed physical model, in which the light emanating from the retina behind a blood vessel is considered to be attenuated by the blood within the vessel according to the modified Beer-Lambert law. Due to the different spectral absorptivity of oxygenated and deoxygenated blood, the amount of attenuation can be used to determine the oxygen saturation of the blood. To determine the oxygen saturation of the blood the transmission T^j of the vessel in each band is calculated as the ratio of the number of detector counts at the centre of the vessel N_V^j, to the number of detector counts adjacent to the vessel $\bar{N}_R^j$. To this set of measured vessel transmissions the following equation is fitted:

$$T^j = \frac{S\int_{k_1}^{k_2} \tau^j(k)\,\eta(k)\,I_{\bar{R}}(k)\,10^{-lc\left(SO_2\mu_{HbO_2}(k)+(1-SO_2)\mu_{Hb}(k)\right)}dk}{\int_{k_1}^{k_2} \tau^j(k)\,\eta(k)\,I_{\bar{R}}(k)\,dk}, \tag{5.12}$$

where k is the wavenumber, τ^j is the transmission function of the jth band, SO_2 the oxygen saturation, S is a parameter representing attenuation due to scattering, l is the path length of light through the vessel, c is the total concentration of haemoglobin in

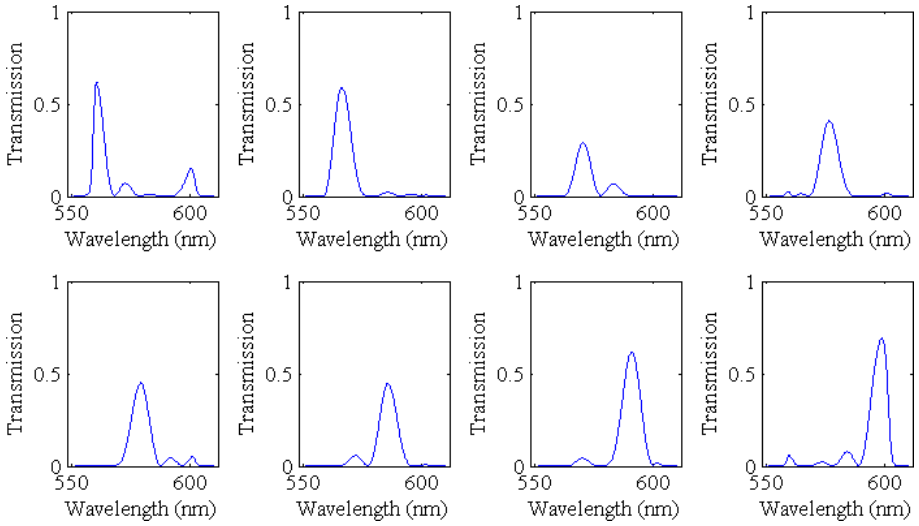

Figure 5.15: The eight bands of the retinal IRIS system.

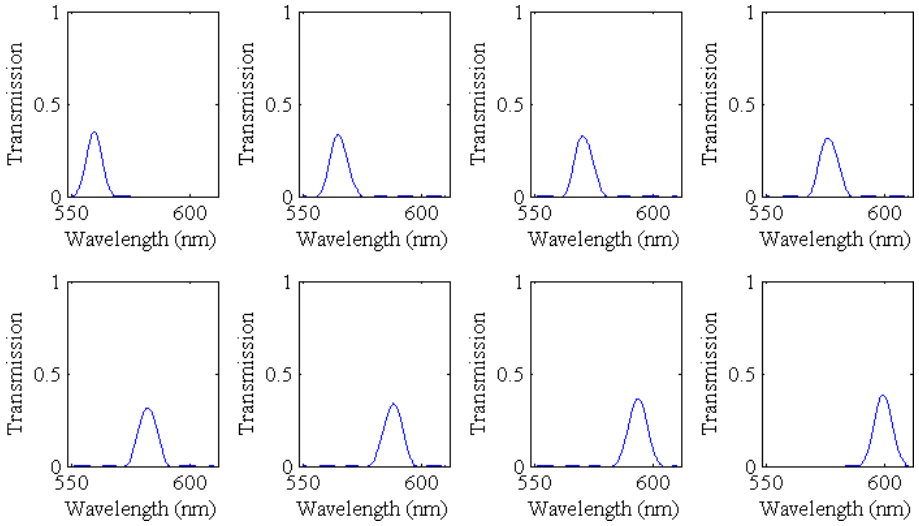

Figure 5.16: The eight LCTF bands for comparison the retrieval of oxygen saturation.

the blood, and μ_{HbO_2} and μ_{Hb} are the absorption coefficients of oxygenated and deoxygenated haemoglobin respectively. $I_{\bar{R}}(k)$ is the number of photons reaching the detector from the retina adjacent to the vessel and $\eta(k)$ is the spectral quantum efficiency of the detector.

To assess the performance of IRIS and the LCTF for the retrieval of oxygen saturation, the maximum value of the $\bar{N}_R^j$ is kept constant while 100 values of N_V^j are simulated, each corresponding to different randomly chosen values for SO_2, l and S. For the sake of simplicity, $I_{\bar{R}}(k)$ and η are assumed to be constant with wave number. The dominant source of noise in these measurments is expected to be shot noise; to simulate this, random Gaussian noise with standard deviations of $\sqrt{\bar{N}_R^j}$ and $\sqrt{N_V^j}$ are added to each of the $\bar{N}_R^j$ and N_V^j before calculating the transmissions $T^j = N_V^j/\bar{N}_R^j$. An attempt is then made to recover SO_2 using a non-linear fit of Equation 5.12 to the

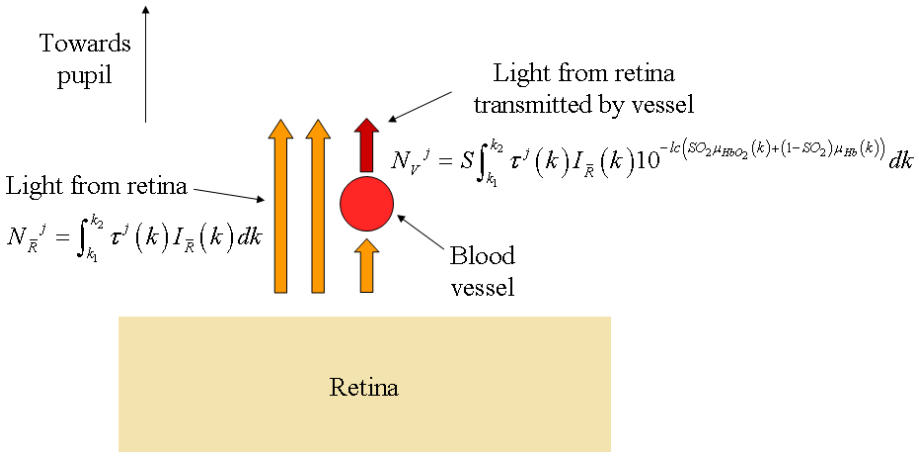

Figure 5.17: The assumed physical model for the oximetry of retinal blood vessels.

noisy transmissions using the simplex method. The standard deviation on the error in the retrieval of SO_2 over the 100 different retrievals is then calculated. This is repeated for a number of values of $\bar{N}_R^j$; Figure 5.18 shows a plot of the standard deviation of the error on the retrieved oxygen saturation against the maximum number of counts from the retina. It is clear that the LCTF offers higher accuracy than the retinal IRIS system for the same number of counts, although to achieve this the flash needs to be twice as intense for the LCTF than for IRIS and 8 sequential flashes are needed compared to a single flash for IRIS. If an error of 2% on the retieved oxygen saturation is acceptable, a maximum count of at least 30000 photons is needed for the LCTF and 60000 photons for the retinal IRIS system. Well depths for commonly used scientific grade CCD are often $<80000e^-$ and in some areas of the image the count may be far less than this; smoothing of the image data and averaging of the transmissions along the vessel may therefore be required to increase the accuracy to acceptable levels.

5.6 Conclusions

This chapter has compared IRIS with conventional spectral-imaging technologies. The general methods of comparison suggest that for 8 or 16 bands IRIS will outperform DIS and FTIS for the same total integration time. Above 16 bands, IRIS offers fewer degrees of freedom and the error on the recovered spectrum is larger than for DIS or FTIS. IRIS does, however, offer a greater information content than either DIS or FTIS, suggesting that it may outperform the other instruments in discrimination tasks. The practical comparison of the retinal IRIS system and the LCTF for retinal oximetry shows that IRIS compares well to an LCTF despite the total integration time being eight times shorter for IRIS than for the LCTF.

There are other factors that shape the choice of instrument for a specific task, es-

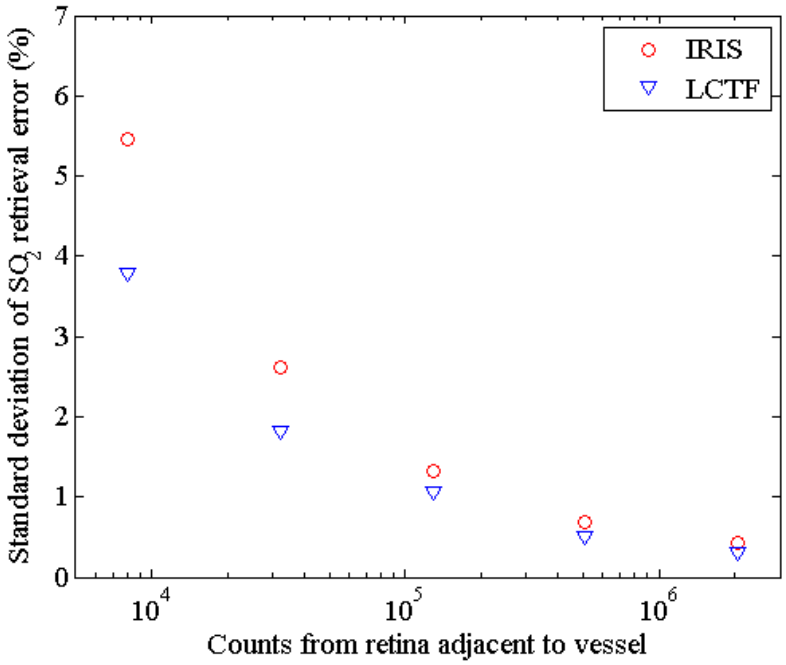

Figure 5.18: Error on retrieved oxygen saturation against the count level from the retina.

pecially when all classes of instrument perform well within the acceptable error limits. Again these as there are many variables to consider, such as cost of production, ease of maintenance, type of platform, detector performance and power requirements. A comprehensive analysis is outside the scope of this work, and it will suffice to mention some of the qualities of IRIS that may be relevant to a systems designer. Certainly, as IRIS has no moving parts, it satisfies the low maintenance requirement; if mechanical scanning is not required then this may also mean that the power required is lower than that ofthe other systems. However, in randomly polarised light IRIS is at most 50% efficient, and in polarised light the transmission is between 0 and 100% depending on the polarisation state. Cost of birefringent materials for the prisms may also be an issue but as described in Section 4.3 the Wollaston-like prism could potentially lower costs while also eliminating dispersion.

Chapter 6

Applications of IRIS

Summary: An example of a spectral-angle map produced from a video frame acquired using the 16-channel system has already been shown in Figure 4.19. In order to demonstrate further the spectral-imaging and real-time capabilities of IRIS, this chapter presents some example data acquired using IRIS systems. A practical application in which IRIS is used to measure the oxygen content of retinal blood vessels is also shown.

6.1 Retrieval of spectra from linearly variable filter

As an example of the spectral imaging capability of IRIS, the 8-channel retinal-imaging IRIS system (Section 2.1) has been used to image a back-illuminated, linearly-variable interference filter; a colour photograph of this filter is shown in Figure 6.1. The illumination is provided by a tungsten source. This filter is transmissive over the wavelength range 400-700nm, whereas the IRIS system is only able to measure the the small section outlined in the figure corresponding to the passband of the retinal-imaging IRIS system. A plot of several transmission profiles of the filter over this range is shown in Figure 6.2, the transmission measurements were made using an *Ocean Optics* USB 4000 spectrometer.

Figure **??** shows the eight registered IRIS sub-images arranged by the peak wavelength of the transmission functions; the measured transmission function for each sub-image is overlaid. The secondary peaks of the transmission functions are particularly

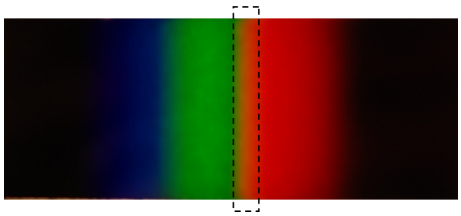

Figure 6.1: Colour photo of a linearly variable filter.

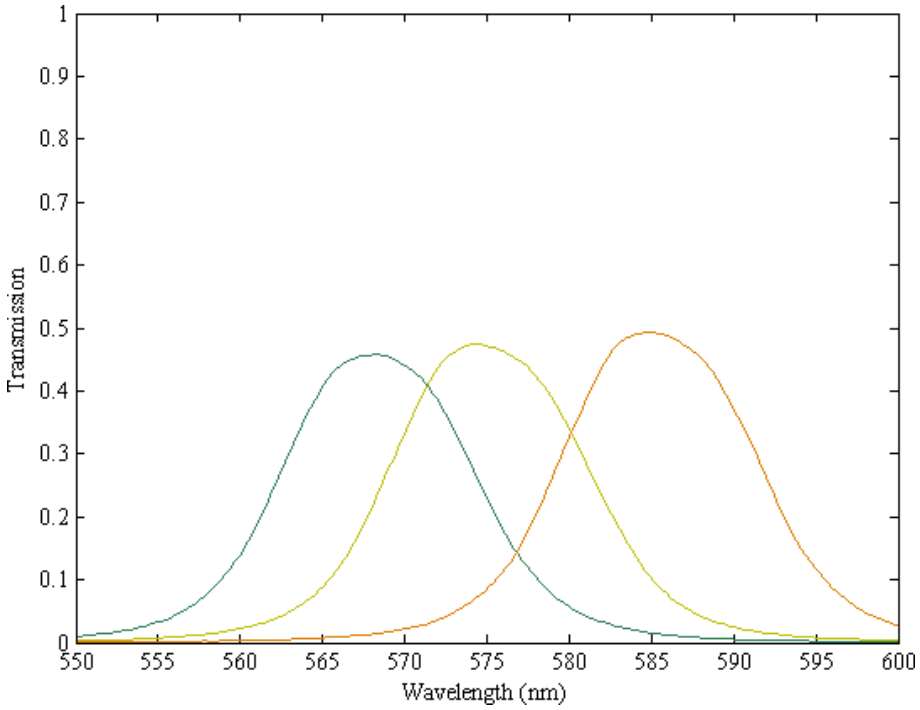

Figure 6.2: Transmission profiles of the linearly variable filter between 550 nm and 600 nm.

apparent in sub-images 1, 7, and 8. To gauge the effectiveness of IRIS at measuring spectral information, the Backus Gilbert inverse is used to recover the spectra from the measurements for the random selection of image points shown in Figure. 6.4. A plot of 10 of these recovered spectra are shown in Figure 6.5, as can be seen, the spectra in the centre of the passband are similar in form to the measured spectra. A plot of the peak wavelength of the recovered spectra against the peak wavelength for the corresponding points on the filter is shown in Figure 6.4. The retrieved peak wavelengths correspond well to the actual wavelengths, with an average absolute error of 1.8 nm and a standard deviation of 1.6 nm. The peak wavelengths of the IRIS bands are shown on the plot as circles; the average absolute error at these points is 1.0 nm.

6.2 Chemical flame test imaging

As an example of the real-time capability of IRIS, the 8-channel broadband system (Section 2.1) has been used to record spectral video from a series of chemical flame tests. Combustion and other highly dynamic scenes are challenging to image using sequential spectral-imaging techniques [114]; this is one of the main motivations behind snapshot spectral imaging techniques such as IRIS.

The 8-channel broadband system records wavelengths over the range 500-830 nm. The transmission functions for this system are shown in Figure 6.7.

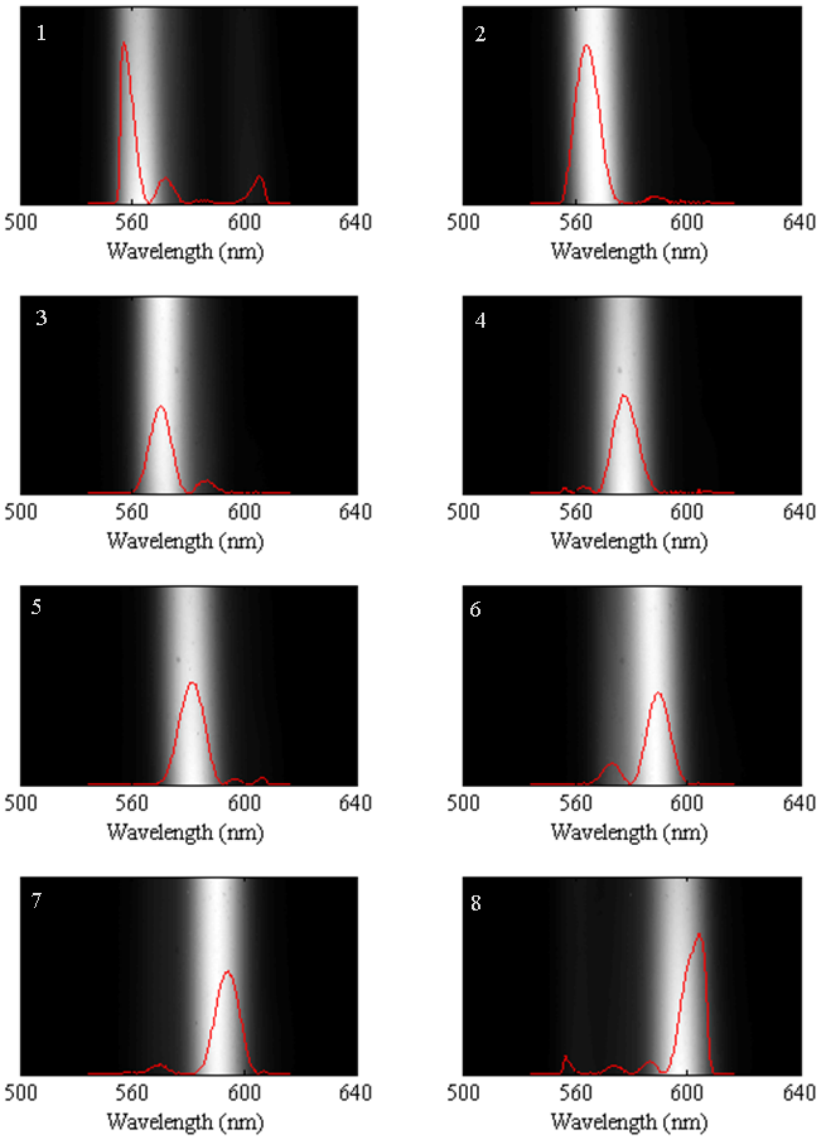

Figure 6.3: 8 registered IRIS sub-images of the linearly variable filter between 560 nm and 600 nm. Plots of the transmission functions as measured with the *Ocean Optics* USB4000 spectrometer are shown overlayed on the sub-image.

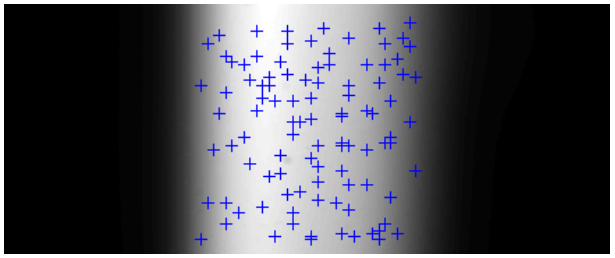

Figure 6.4: Location of points retrieved in Figure 6.6.

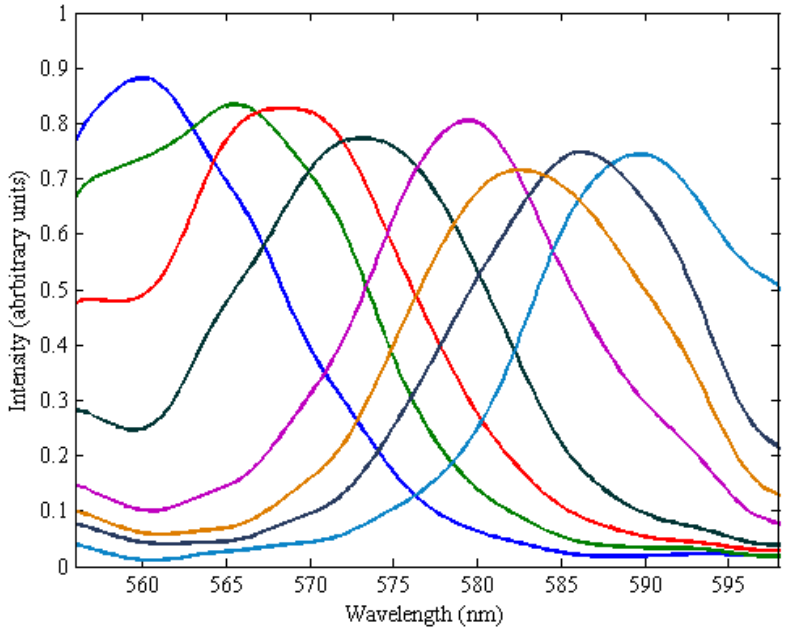

Figure 6.5: Plots of the transmitted intensity of the interference filter at different points, retrieved from the IRIS data using the Backus Gilbert inverse.

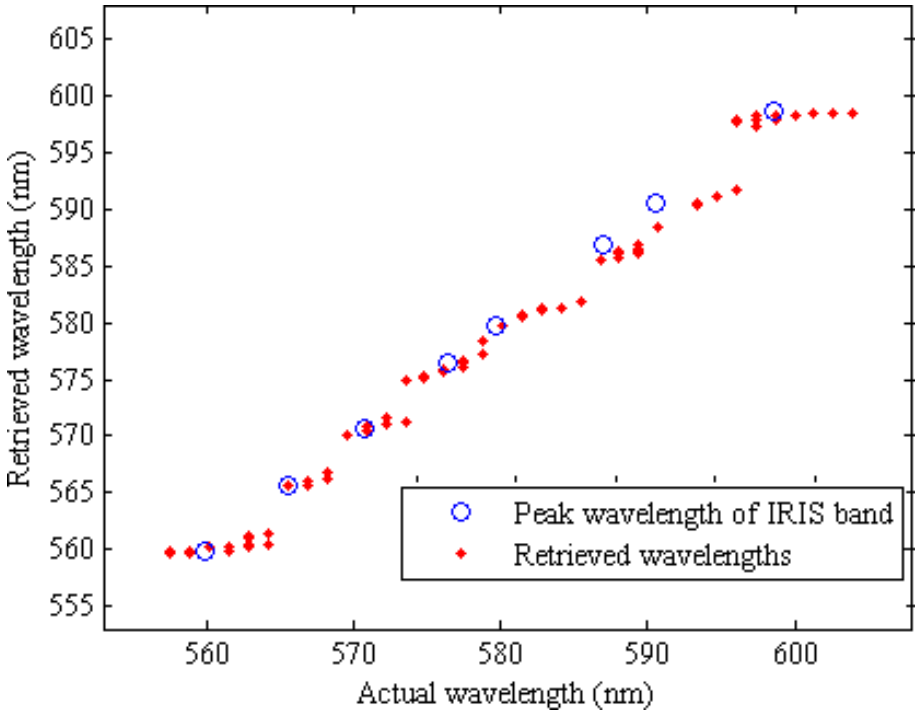

Figure 6.6: Location of peak wavelengths retrieved using the the Backus-Gilbert method versus actual wavelength.

Figure 6.8 shows colour photos of flame tests for (a) Cu_2O[1], (b) $CuSO_4$[2], (c) Li_2SO_4[3], (d) $Na_2B_4O_7$[3], and (e) Na_2CO_3[2]. Figure 6.9 shows frames from the spectral video of the flame tests for these same chemicals. As can be appreciated from Figures 6.8 (a) and (e), Cu_2O and Na_2CO_3[2] appear similar in colour images but in the IRIS frames Cu_2O can be differentiated by its near-infrared emission between approximately 760 nm and 810 nm [115]. Likewise, Figures 6.8 (b) and (d), $CuSO_4$ and $Na_2B_4O_7$ both impart a green hue when introduced to the flame, despite having markedly dissimilar spectra when viewed with IRIS. In the case of $Na_2B_4O_7$ the green hue is due to combination of emissions around 580 nm (yellow) as seen in the IRIS frame, and emissions below 500 nm (blue) which are not detected in the passband of this instrument. The emission of $CuSO_4$ around 540 nm is responsible for the perception of green in the flame. Conversely, $Na_2B_4O_7$ and Na_2CO_3[2] appear similar in the IRIS frames despite giving different rise to different colours in the flame test, although emission from $Na_2B_4O_7$ in the 780 nm sub-image means these elements can be distinguished.

6.3 Retinal imaging for blood oximetry using IRIS

An example of a practical application of IRIS systems is spectral imaging of the retina for the measurement of oxygen saturation in retinal blood vessels. Figure 6.10 shows a diagram of the human eye: the retinal images are measured through the pupil of the eye using a fundus camera. A colour fundus-camera image of a retina, taken by David Mordant[4] is shown in Figure 6.11. Although the vessels all appear very similar in colour, the difference in the spectral properties of oxygenated and deoxygenated blood makes oximetry using spectral measurements possible.

Oximetry of the retinal vessels is a promising tool for the diagnosis and monitoring of diseases such as glaucoma and diabetic retinopathy, which result in damage to the retinal tissues and in doing so effect a change in the metabolism of the eye. This change in metabolic activity causes the oxygen saturation of the haemoglobin in the blood vessels to differ from that expected in a healthy eye. Due to the different spectral characteristics of oxygenated and deoxygenated haemoglobin (Hb And HbO_2), such changes in oxygenation can be detected and quantified in a non-invasive manner using spectral measurements. The number of spectral-bands required for successful retinal oximetry is an unanswered question at present. Although there are only two quantities (the relative abundances of Hb and HbO_2) to recover, the presence of other chromophores and the scattering of light within the tissue, mean that at least three wavelengths are required for absolute oximetry. As such, instruments that use two wavelengths can only be assumed

[1]Source: David Gessel
[2]Source: Søren Wedel Nielsen
[3]Source: Unknown
[4]Research Fellow in Ophthalmology, Gloucestershire Hospitals NHS Foundation Trust

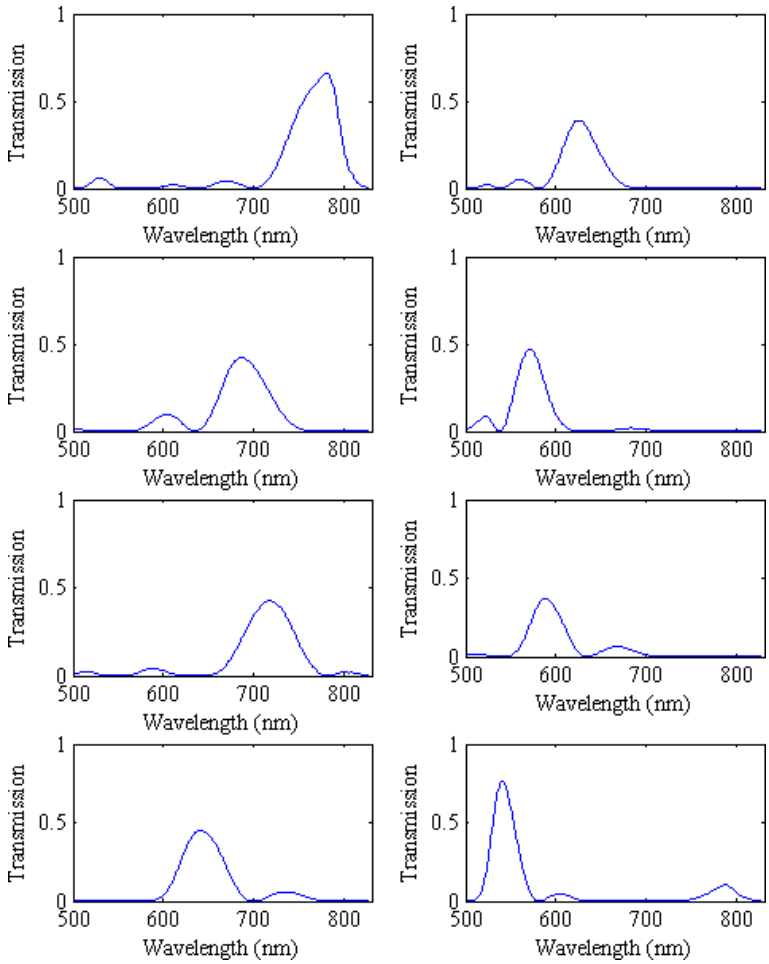

Figure 6.7: Transmission functions for 8-channel broadband system.

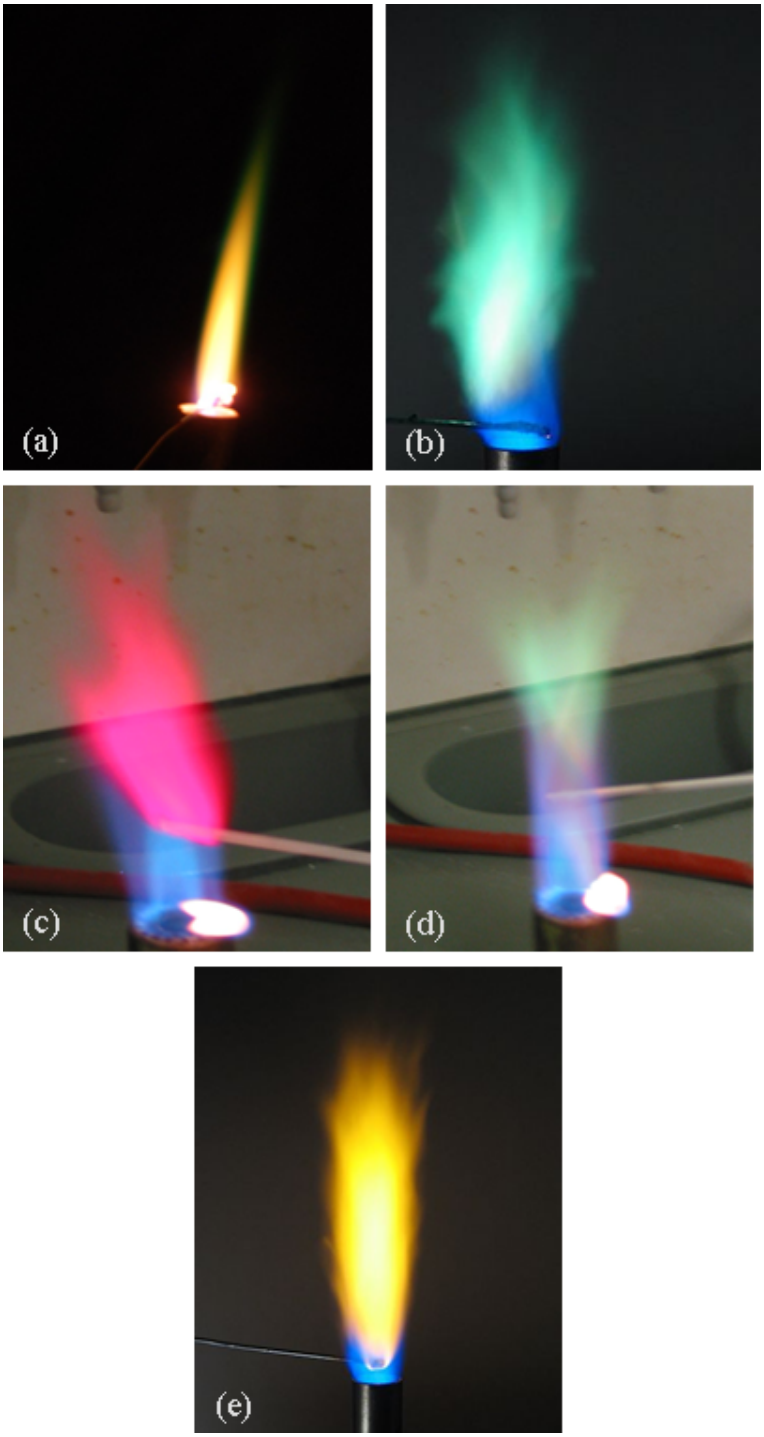

Figure 6.8: Colour photos of flame tests for (a) Cu_2O , (b) $CuSO_4$, (c) Li_2SO_4, (d) $Na_2B_4O_7$, and (e) Na_2CO_3.

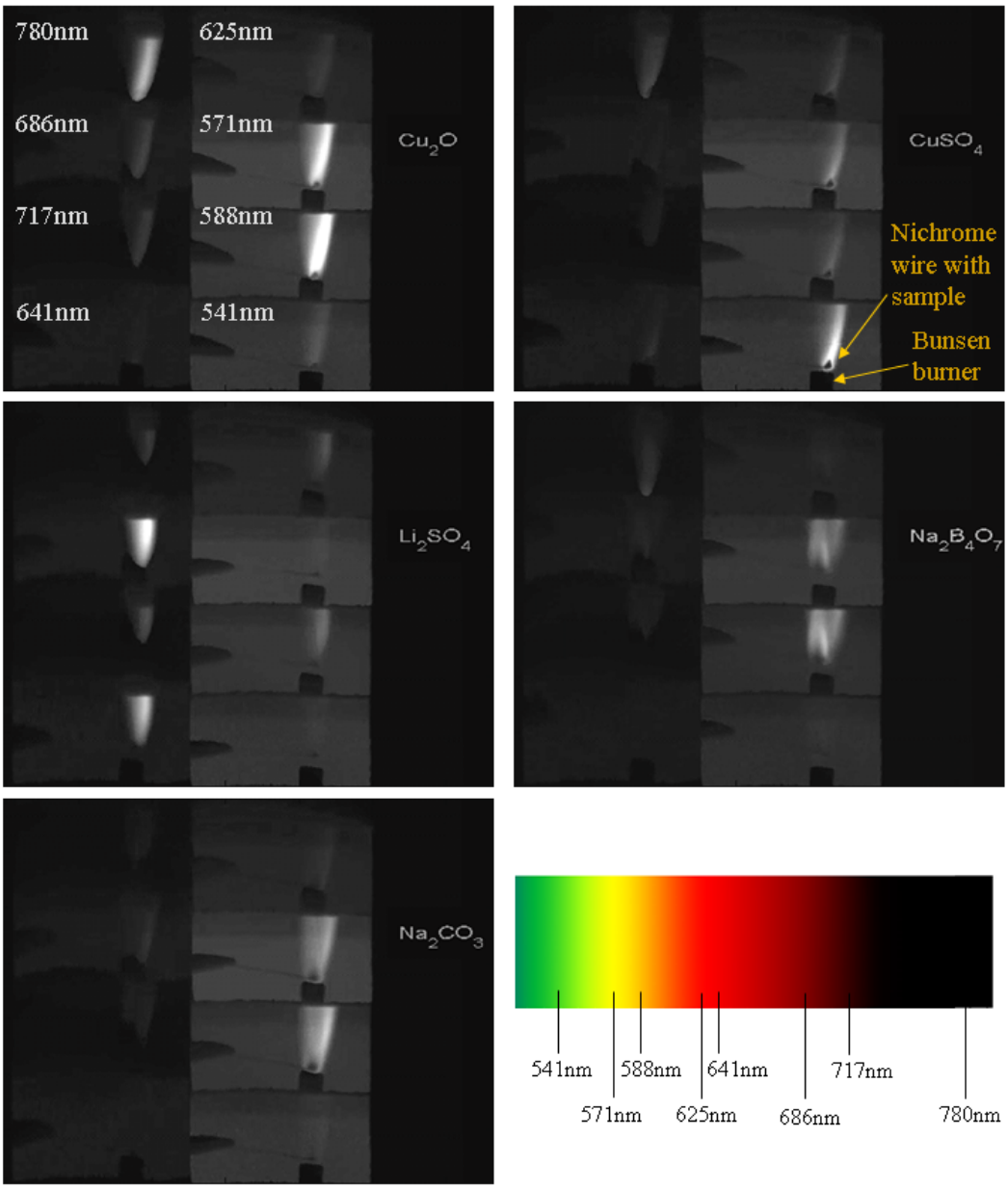

Figure 6.9: Frames from a series of videos of flame tests, the peak wavelengths for each band are shown in the top-left figure. Also shown is a diagram approximating the apparent colour seen for a narrow spectral line at a given wavelength.

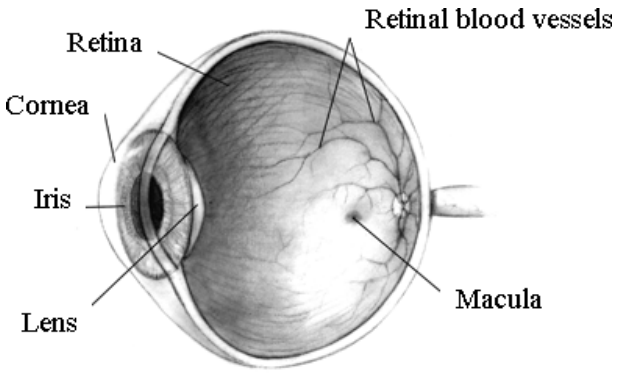

Figure 6.10: Structure of the eye, the body of the eye is filled with the transparent vitreous humour.

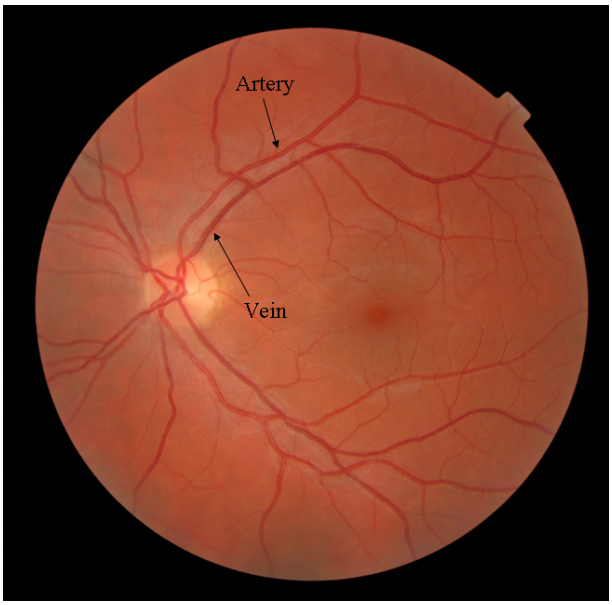

Figure 6.11: Colour image of a retina.

to return relative values for the oxygen saturation. In this section the 8-channel IRIS system designed for retinal oximetry (Section 2.1) has been fitted to a *TopCon* TRC 50-IA fundus camera (Figure 6.12) and used to take images for a phantom eye constructed for the purposes of validation. A demonstration of the feasibility of oximetry using this eye-phantom data is presented in this section.

6.3.1 Background

Previous retinal spectral imaging studies have used both dispersive devices, for which only a narrow profile across the retina can be measured in a single frame, and filter based systems, which acquire a number of narrow-band spectral images sequentially. Xenon flashes with exposure times on the order of 10 ms are currently the most common light source for fundus cameras. To acquire an image using either of these techniques therefore

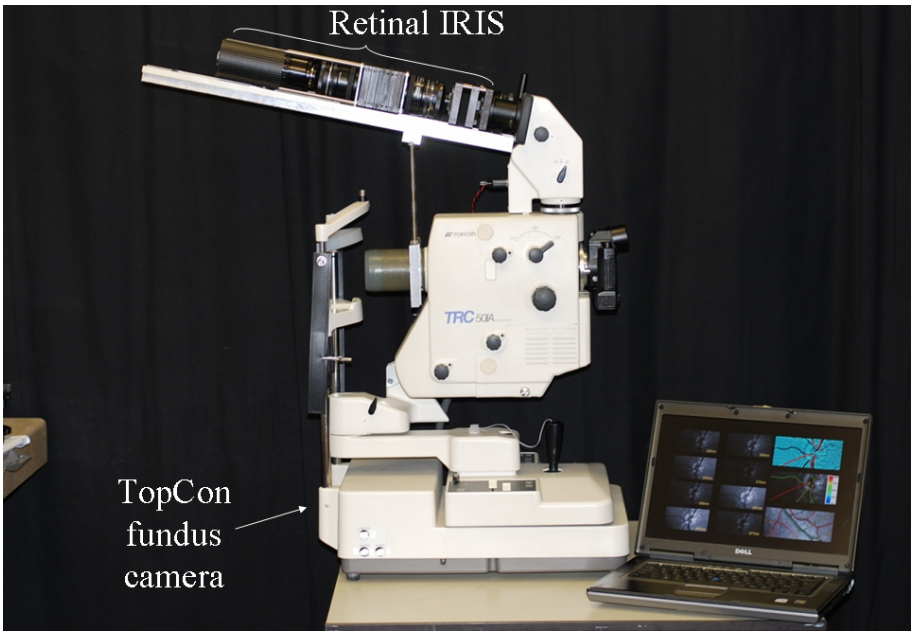

Figure 6.12: Retinal IRIS system fitted to a *TopCon* fundus camera.

requires a number of repeated flashes which can be uncomfortable for patients. The eye is also a rapidly moving environment due to saccades and slower movements; these movements translate and rotate the image of the retina at the detector and change the distribution of light incident on retina, making co-registration and combination of the data non-trivial for any sequential system. The biochemistry of the eye is also dynamic; blood volume flowing within the vessels and tissues of the eye changes with time and in response to light intensity, and the relative abundances of chromophores in the eye may also change.

A number of snapshot techniques have also been applied to the oximetry in the eye including, two-, three- and four-wavelength systems using dichroic beam splitters to split the input beam in a manner similar to a three-CCD camera [116]. CTIS (described in Section 1.2) has also been applied to retinal imaging [117], and a snapshot system using an array of lenslets with different spectral transmissions to re-image a broadband image of the retina projected onto a viewing screen [118].

IRIS is particularly suitable for retinal oximetry as it offers the potential for a wide field of view and high light throughput. The snapshot capability removes a number of the problems associated with the dynamic nature of the eye and offers improved patient comfort compared to sequential techniques.

6.3.2 Retrieval of oxygen saturation using a physical model

Figure 6.13 illustrates the relavent layers of the retina. The outer layer is the sclera, the white material which forms the majority of the exterior of the eye. At the back of

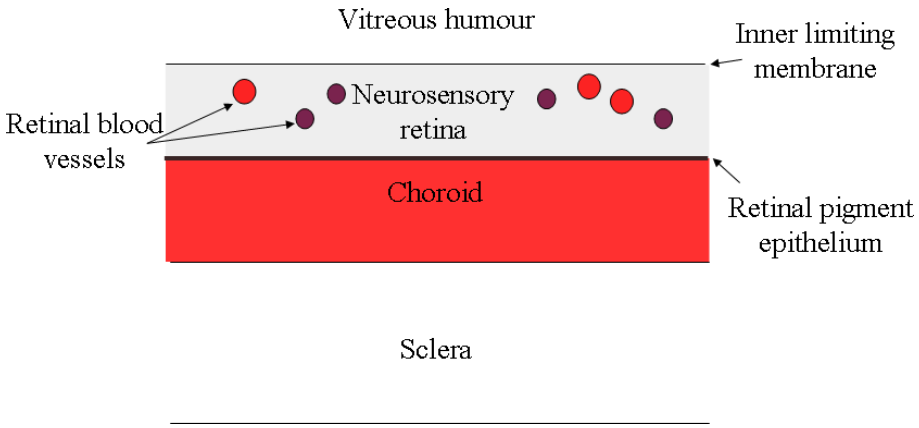

Figure 6.13: A simple schematic of the retina.

the eye the sclera is 1 mm thick. The choroid is a 300 μm layer of highly oxygenated blood vessels; blood flow in the choroid is high and helps to nourish the highly metabolically active outer neurosensory retina, including the photo-receptors. The retinal blood vessels run through the largely-transparent, neurosensory retina (around 100 μm thick) and provide nourishment mainly to the inner parts of this layer [119]. The retina is separated from the transparent vitreous by the inner limiting membrane.

The propagation of light in the eye is complicated due to scattering and absorption by the tissues, and retrieval of the oxygen saturation from spectral imaging measurements requires an accurate radiative-transfer model. A Monte-Carlo simulation of the propagation of light in a blood vessel in front of a scattering background (representing a simplified retina) has been implemented by the author using the Zemax ray-tracing software. This simulation is based upon a basic Zemax model of a blood vessel created by Gonzalo Muyo. Monte-Carlo methods are a common technique for simulating photon transport in turbid media such as tissue [120], and are used to aid understanding and provide validation of physical models where in vivo validation is difficult. The Monte-Carlo simulation works by simulating the paths of photons packets as a random walk through a three-dimensional structure. For each step in the walk the photon scattering probability with distance is an exponential cumulative distrubution parameterised by the 'scattering coefficient'. The angle of a scattered photon follows the Henyey-Greenstein probability distribution [121] which is characterised by a parameter g. The scattering parameters used in the Monte-Carlo simulation, for oxygenated and deoxygenated blood have been taken from [122, 123]. Light is also absorbed as it moves through a medium, exhibiting an exponential decrease in intensity with path length, the amount of attenuation per unit path length is characterised by the absorption coefficient of the medium; the absorption coefficients of oxygenated and deoxygenated haemoglobin are shown in Figure 6.14.

Cross sections of the three-dimensional model and some typical photon paths are

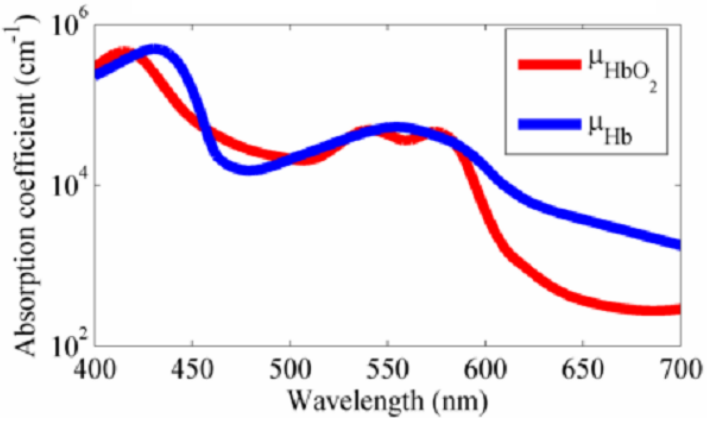

Figure 6.14: Absorption coefficients for oxygenated and de-oxygenated haemoglobin in the typical concentrations found in humans [5].

shown in Figure 6.15. The intensity of a vessel in an image depends on the relative proportions of these different photon paths.

Most reported oximetry algorithms for application to the retinal vasculature assume that the total amount of light backscattered from the vessel as in Figure 6.15 (c), is negligible. The majority of photons emanating from the vessel are therefore considered to be incident from the adjacent retinal layers as in Figures 6.15 (b) and (c). As illustrated in Figure 6.16, absorption by haemoglobin in the blood attenuates a portion of this light on its passage through the vessel toward the detector, making the vessel appear darker than the adjacent retina. Further light is lost due to scattering from the blood cells. If a measurement is made in monochromatic light, then for this scenario, the intensity of a vessel compared to the background is described by the modified Beer-Lambert law, for which the transmission of a vessel is

$$T = Se^{-lc\left(SO_2\mu_{HbO_2}+(1-SO_2)\mu_{Hb}\right)}, \tag{6.1}$$

where μ_{Hb} and μ_{HbO_2} the absorption coefficients of deoxygenated and oxygenated haemoglobin respectively, l is the path length through the blood column, c is the total concentration of haemoglobin in the blood, SO_2 is the average oxygen saturation of the haemoglobin, and S is a parameter to account for scattering within the blood column.

In the IRIS system the vessel transmission is sampled by the IRIS bands $\tau^j(k)$, meaning that the intensity of the vessel compared to the background in the IRIS system is not given by the Beer-Lambert law but is instead

$$T^j = q\frac{S\int_{k_1}^{k_2}\tau^j(k)\,\eta(k)\,I_{\bar{R}}(k)\,10^{-lc\left(SO_2\mu_{HbO_2}(k)+(1-SO_2)\mu_{Hb}(k)\right)}dk}{\int_{k_1}^{k_2}\tau^j(k)\,\eta(k)\,I_{\bar{R}}(k)\,dk}. \tag{6.2}$$

$I_{\bar{R}}(k)$ is the number of photons reaching the detector from the retina adjacent to the

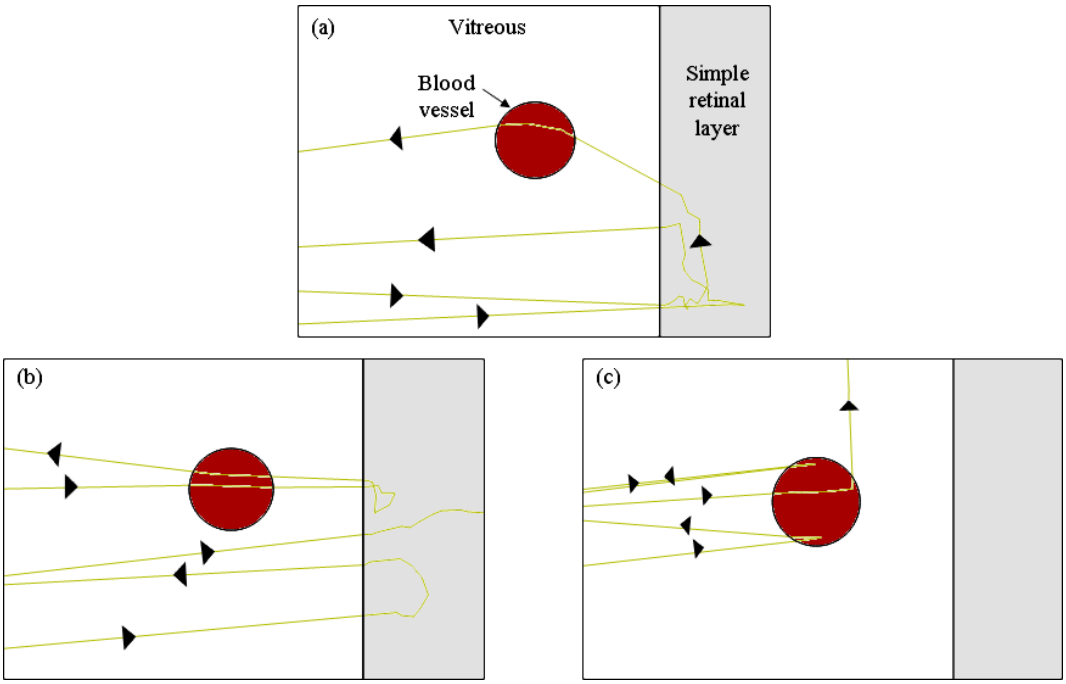

Figure 6.15: Cross sections from three-dimensional Monte-Carlo model of light propagating through a blood vessel and simple scattering background. Some typical paths in the vessel of detected photons are shown: (a) photons which scatter through the retina then pass once through the vessel, (b) photons which pass through the vessel onto the retinal layer and then return through the vessel, (c) photons which backscatter from the vessel without interacting with the retinal layer.

vessel which depends on the source and the reflectivity of the retina, $\eta(k)$ is the spectral quantum efficiency of the detector.

For in vivo measurements, S , lc , and SO_2, are unknown, and so a minimum of three wavelengths must be used to fit Equation 6.2 to the measured transmission of the vessel and solve for SO_2. $I_{\bar{R}}(k)$ must be measured or estimated from existing empirical data. The quantum efficiency of the detector must be measured.

In each IRIS sub-image, the transmission for each point along the vessel is calculated as the ratio $T = I_V / \bar{I}_F$, with I_V the gray level number at the centre of the vessel, and $\bar{I}_F$ the average gray level number from the fundus adjacent to the vessel, as illustrated in Figure 6.17. This results in eight different transmissions for each point along the vessel. Equation 6.2 is then fitted to these transmissions to determine the correct value for SO_2.

6.3.3 Method

In-vivo validation of retinal vessel oximetry is challenging; as alternative methods of measuring the *in vivo* oxygen saturations of the blood vessels are not practical; for instance, the vessels cannot easily be probed using electrodes to measure the partial pressure of oxygen in the blood. A phantom eye has been constructed to provide some validation of the physical models and the oximetry algorithm. A schematic of the

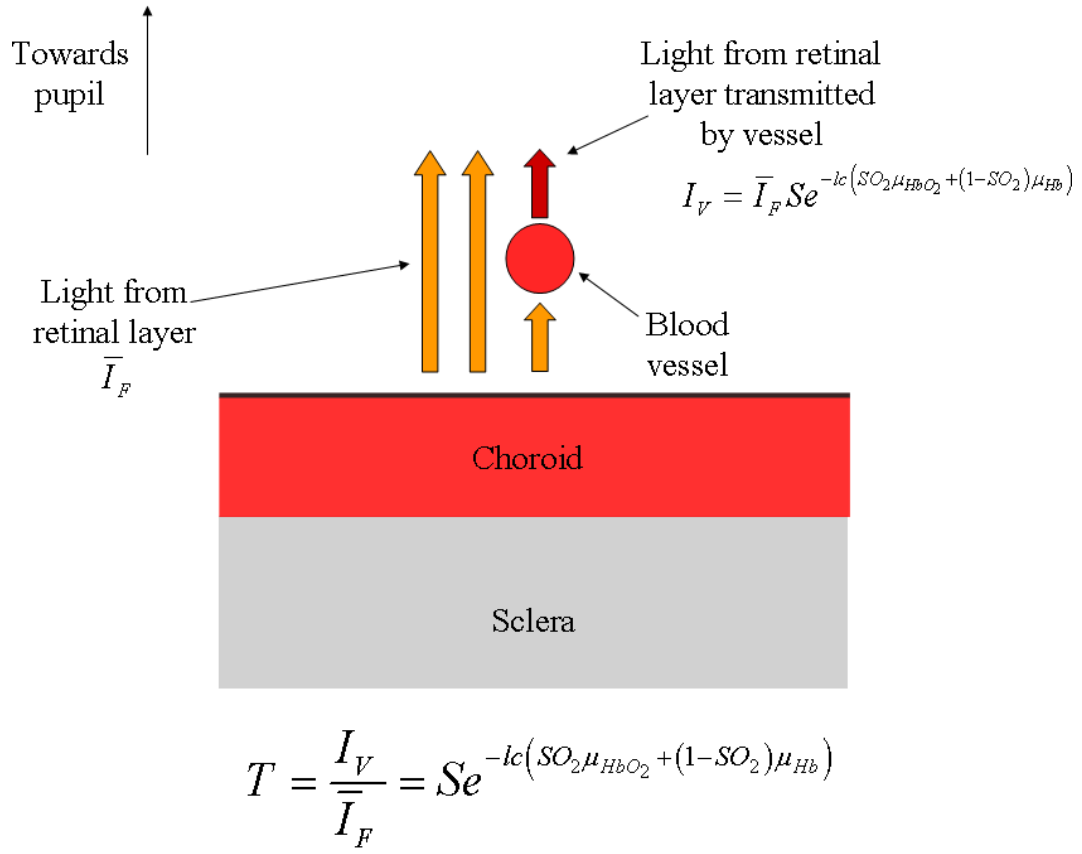

Figure 6.16: Illustration of the assumed physical model of light propagation in retinal imaging. Light is transmitted by the vessel according to the Beer-Lambert law.

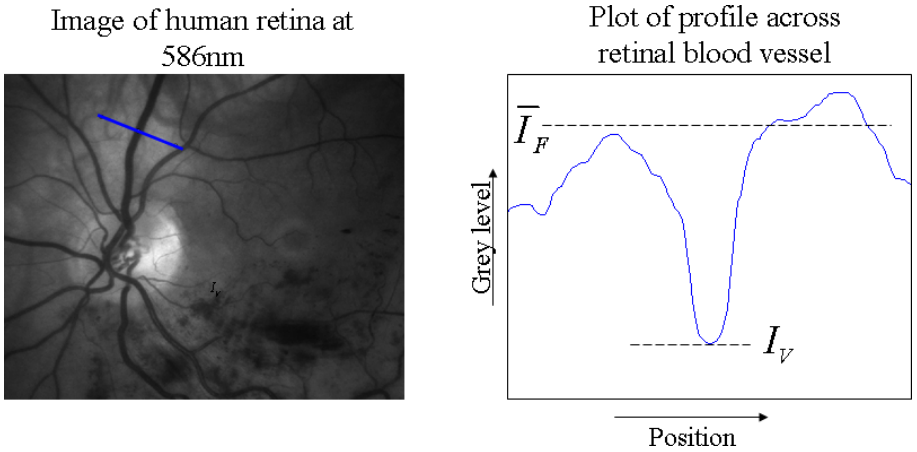

Figure 6.17: The transmission for a point midway across a vessel is determined by taking the ratio of the number of detector counts at the centre of the vessel I_V , and the average number of counts from the fundus adjacent to the vessel $\bar{I}_F$.

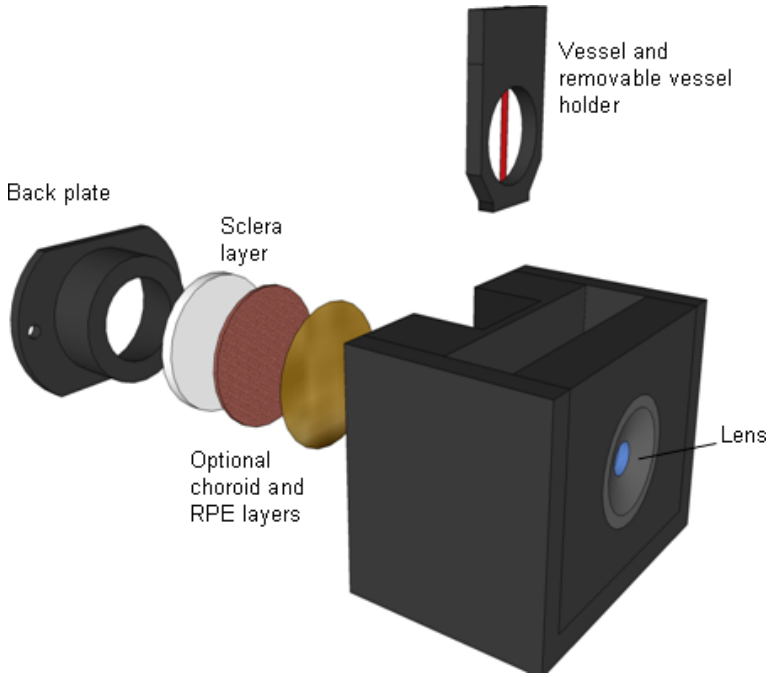

Figure 6.18: Schematic of the eye phantom used in the oximetry experiment.

phantom is shown in Figure 6.18. The phantom consists of a waterproof body with a 22 mm focal-length lens to simulate the combined power of the lens and cornea. The properties of the retina can be changed by placing different materials between the vessel and the sclera layer. The blood vessels are simulated using fluorinated ethylene propylene (FEP) capillaries, the body of the eye is filled with a saline solution which has the same refractive index as the FEP. The phantom was designed by Gonzalo Muyo and the author, and constructed at Heriot-Watt university.

The retinal IRIS system was used to record data from the eye phantom. The eye phantom was used with 150 μm diameter capillaries filled with horse blood to represent large blood vessels; for the retinal layer a 3mm thick layer of 99% reflective *Spectralon* (a commonly used reflectance standard) was used. This scenario has also been modelled using the Zemax Monte-Carlo simulation, which suggests that intensity of the backscattered component can be regarded as negligible (less than 2% of the intensity of the *Spectralon* background), and that the modified Beer-Lambert law describes the relative intensities of the vessel and retina adjacent to the vessel.

The horse blood used in the eye-phantom was deoxygenated in a controlled manner using sodium-dithionite [124], which is a reducing agent commonly used for this purpose. The concentration of sodium-dithionite in blood has been determined to show a linear relationship with the oxygen saturation of the blood. Eight different concentrations have been used here to produce eight blood samples with known oxygenation saturations between 100% and 10% with an error of $\pm$3%. The IRIS system was used to aquire images of the eye-phantom retina for capillaries filled with each sample.

6.3.4 Results

Figure 6.19 shows the eight registered and dark calibrated IRIS images of a vessel in the eye phantom. To recover the oxygenation from the IRIS measurements the transmission of the vessel is first calculate by dividing the number of counts at the centre of the vessel by the average number number of counts adjacent to the vessel. Equation 6.1 is then fitted to these transmissions The reflectivity of the retinal layer in the eye phantom is constant with wavelength and the illumination is known. The camera used is a *QImaging* Retiga 4000R for which the quantum efficiency data is supplied by the manufacturer. This equation is modified further to account for a loss of image contrast in the optical system. This loss of contrast was identified by the author during calibration of the system and appears to be due to veiling glare. To account for this contrast loss, the following equation is fitted to the measured transmissions

$$T^j = (1-K)\,\frac{S\int_{k_1}^{k_2}\tau^j(k)\,\eta(k)\,I_{\bar{R}}(k)\,10^{-lc\left(SO_2\mu_{HbO_2}(k)+(1-SO_2)\mu_{Hb}(k)\right)}dk}{\int_{k_1}^{k_2}\tau^j(k)\,\eta(k)\,I_{\bar{R}}(k)\,dk}+K, \qquad (6.3)$$

where K is a parameter describing the contrast loss.

Figure 6.20 shows the result of fit of Equation 6.3, to IRIS measurements on a section of a capillary containing blood with an oxygen saturation of 80%. The coefficient of determination for this fit is $R^2 = 0.9934$ and the retrieved oxygenation is 82%. Figure 6.21 shows the results of the oxygen saturation retreival along a length of vessel, the mean retrieved oxygenation is 83% and standard deviation of the retrieved oxygenation along the vessel is 2%.

Figure 6.22 shows a plot of the average retrieved oxygen saturation against the known oxygen saturation for each blood sample, the average difference between the known and retrieved values is 4% the standard deivation on this error is 3%. The coefficient of determination of the fit is 97%.

Much work remains on improving the physical model to determine the effects of the different textures and pigments that are present in the retina, on the retrieval of oxygen saturation. However, these results suggest that IRIS is a potentially useful tool for retinal oximetry.

6.4 Conclusions

This chapter has demonstrated that IRIS is a viable snapshot imaging technology. The example of flame tests imaging shows that IRIS systems are capable of imaging highly dynamic scenes. As a practical example, the application of IRIS to oximetry measurements has also been demonstrated.

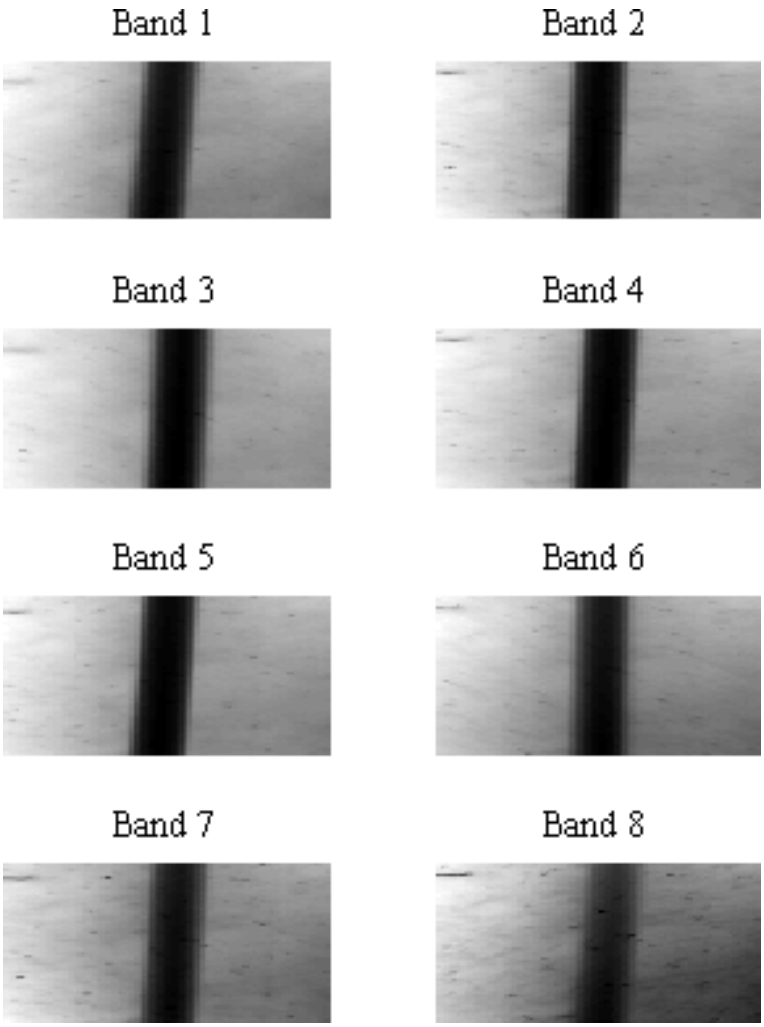

Figure 6.19: The eight registered and dark-calibrated IRIS images of a capillary in the eye phantom, the known oxygen saturation of the blood in the capillary is 80%.

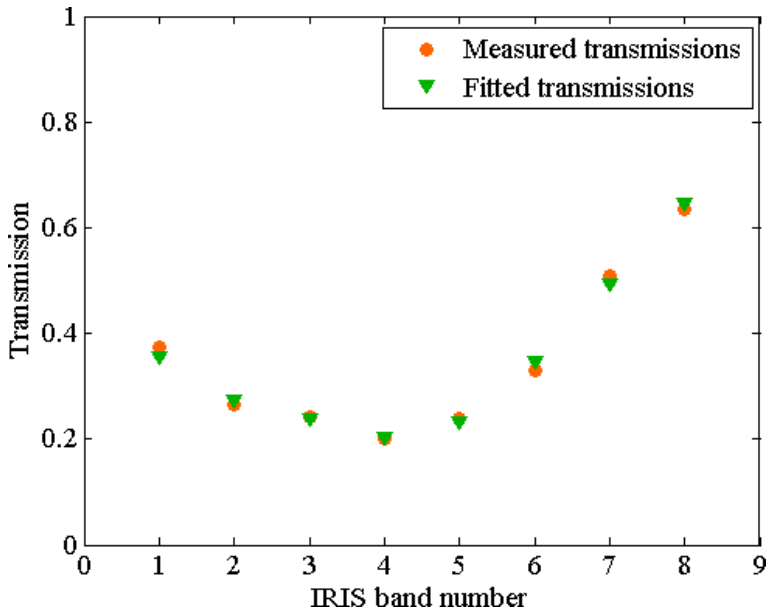

Figure 6.20: Nonlinear fit of Equation 6.3 to IRIS measurements on the eye phantom. The retrieved oxygenation of the blood is 83% and the known oxygenation is 80%.

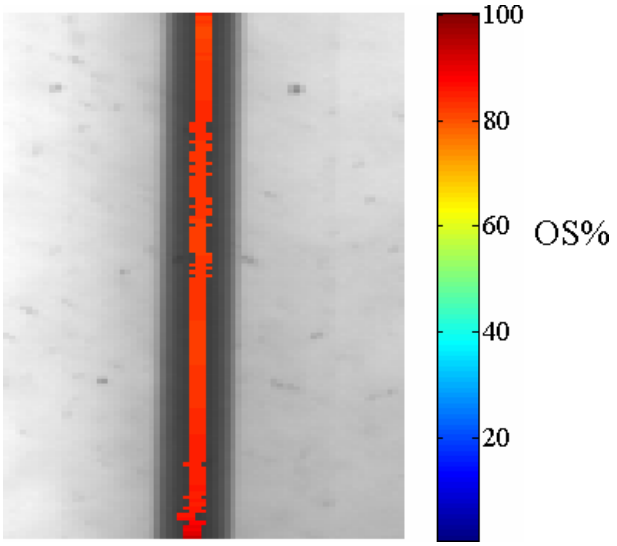

Figure 6.21: The retrieved oxygen saturations along a the vessel shown in 6.19.

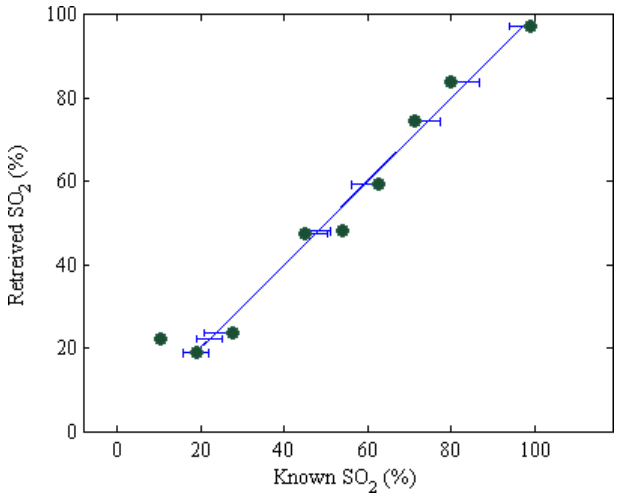

Figure 6.22: A plot of the oxygen saturation of blood retrieved from IRIS measurements on the eye-phantom, against the known oxygen saturations.

Chapter 7

Conclusions

7.1 Summary of thesis

This work has explored the IRIS technology in detail. Chapter 1 provided a summary of existing spectral imaging technologies in order to put the IRIS technology in context. Chapter 2 introduced the existing IRIS concept and detailed the optical design of IRIS systems, providing refinements to the initial concept. The optical design principles were then demonstrated in the first order design of the 8-channel retinal IRIS system. The specific challenges in implementing IRIS systems in the infrared were illustrated in the optical redesign of the MWIR IRIS prototype system to remove stray light. The IRIS system optical parameters resulting in the optimal throughput for each field of view were calculated, and a survey of suitable birefringent materials was presented. Chapter 2 concluded with a discussion of potential image aberrations in IRIS systems and a description of the issues surrounding registration of the IRIS images.

Chapter 3 discussed the implementation of an interferometric ray tracing program. This program was constructed to aid the IRIS design process, which was demonstrated by its use in the refinement of the first-order design of the 16-channel broadband IRIS system; the first order design was produced using the optical relations presented in Chapter 2. The interferometric ray tracing program was also used to predict the performance of the 16-channel broadband IRIS system. Manufacture of the 16-channel broadband system was described and measurements of the spectral transmission of this system were shown to be in excellent agreement with the results of the ray trace. The measurements also revealed the extent of the longitudinal chromatic aberration present in the lens system of the 16-channel broadband IRIS system and discussed a potential solution to this problem. A co-registration algorithm for IRIS was also devised and successfully applied to the retinal IRIS system.

Chapter 4 examined the chromatic aberrations present in IRIS systems, exemplified by the measurements on the 16-channel broadband IRIS system presented in Chapter 3. Several methods for dealing with the chromatic aberrations due to birefringent spectral

dispersion in Wollaston prism materials were discussed. These included an existing design for an achromatic, reflective Wollaston prism which also has the potential to reduce the cost of manufacturing IRIS systems. An attempt was also made at finding a novel design of achromatic Wollaston prisms using combinations of different materials to reduce dispersion; several combinations resulting in greatly reduced dispersion compared to conventional Wollaston prisms were found, and the feasibility of the required materials discussed. A practical demonstration was given of the use of an array of interference filters to remove the secondary peaks from the IRIS bands and thus reduce the extent of the dispersion present in the 16-channel broadband IRIS system. This filter array was combined with an array of corrective glass elements to remove the longitudinal chromatic aberration due to the lenses. Chapter 4 concluded with a demonstration of the feasibility of using image processing to reverse the effects of dispersion in an 8-channel system, simulated using the interferometric ray tracing program.

Chapter 5 has compared IRIS to conventional spectral imaging technologies using several general figures of merit; it was found that, in terms of these figures of merit, IRIS systems are expected to outperform these conventional systems when the number of bands is 16 or lower. If the number of bands is greater than 32 then the benefits provided by the snapshot capabilities IRIS are diminished, although for the same total integration time, the per-pixel Shannon information content of IRIS instruments is expected to be larger than for conventional instruments with the same number of bands regardless of the characteristics of the noise. A specific practical example, comparing the 8-channel retinal IRIS system with an LCTF for the purposes of retinal oximetry, showed that in a single exposure IRIS performs as well as the LCTF does with the eight successive exposures required to record the spectral images. Although the LCTF can record a spatial area eight times larger than IRIS in these exposures.

Chapter 6 presented results from some applications of IRIS, including a demonstration of video rate imaging of chemical flame tests, an example of a highly dynamic scene that is challenging to image at these rates using conventional spectral imaging technologies. An example of the practical application of IRIS for determining the oxygen saturation of blood within blood vessels was also presented.

7.2 Conclusions and recommendations for future work

It has been shown that the IRIS technology is a viable snapshot spectral imaging technology offering many potential benefits in the imaging of dynamic scenes. The technology offers relatively high throughput and the potential for instantaneous spectral imaging with large field of views. IRIS is also a compact instrument with no moving parts. The field of view and/or pixel resolution of IRIS is, however, reduced compared to conventional spectral imaging instruments of similar spectral resolution. This means that the

advantages due to the snapshot capability of IRIS may be diminished if spatial scanning is required to acquire a larger field of view. If scanning is required the snapshot capability may still be an advantage for scenes in which small extended areas of the image change too rapidly to be acquired accurately using a sequentially scanned system.

It has also been shown that the transmission functions of IRIS are in general less optimal than those of conventional spectral imaging technologies, meaning that all else being equal, a snapshot spectral imaging system with direct bands would be expected to outperform IRIS if the noise is dominated by shot noise, and a snapshot Fourier transform imaging spectrometer would be expected to outperform IRIS if the noise is dominated by sources other than shot noise.

Compared to other snapshot spectral imaging technologies, all else being equal, IRIS would be expected to underperform direct reformatting technologies such as the image slicing spectrometer (ISS) and optical fibre reformatting techniques. The comparison of IRIS with snapshot transform instruments such as the computed tomography imaging spectrometer (CTIS) is less clear, as the success of the CTIS reconstruction depends on the scene, the design of the dispersive element, and the reconstruction algorithm used. There has been no published analysis of the signal-to-noise ratio of CTIS measurements; the total efficiency is moderately high which suggests that the signal-to-noise ratio of the reconstructed spectral data cube may also be high. However, there are many unused detector pixels in CTIS systems and the resolution is low and/or field of view small. It is possible that a direct comparison is unwarranted as IRIS and CTIS fulfill slightly different roles, with CTIS useful for measurements at higher spectral resolutions and smaller field of view, and IRIS for measurements with lower spectral resolution and larger fields of view. For similar reasons to CTIS, the coded-aperture snapshot spectral imaging (CASSI) devices cannot easily be compared to IRIS, although again the efficiency of CASSI is high and in the best case the measurement should be equivalent to a direct system. Future work could include a more thorough analysis of these transform instruments.

The cost of manufacture of the demultiplexer in IRIS systems may be prohibitive to their wide-spread use; although retarders are inexpensive, Wollaston prisms are an expensive and potentially fragile component. The reformatting elements in ISS and fibre reformatters are challenging to engineer and also likely to be high cost. The hardware required for CTIS and CASSI is relatively low cost compared to IRIS systems.

If some of the engineering problems involved in the ISS can be solved, and the cost of manufacture reduced, then the ISS is possibly a more attractive option than IRIS as these systems can in theory be designed to offer the same field of view with more optimal transmissions functions (in the shot-noise limit). ISS are also likely to outperform transform instruments. A thorough analysis of the efficiency of ISS and limitations on the field of view is, however, required before drawing any firm conclusions.

Currently the major difficulty in implementing IRIS systems is the chromatic aberration present in the system due to the Wollaston prisms. Achromatising of the system or otherwise removing the effects of dispersion is therefore the most important aspect of any future research. The reflective Wollaston prism reported by Feldman et al. and described in Section 4.2.1 is perhaps the most attractive option: as well as being free from dispersion, it offers the potential for significant cost reductions in manufacture, and further opportunity for tuning the IRIS transmission functions. This option should be considered in detail, with full modelling of the change in the optical performance of the beam splitter with angle and wavelength. The designs determined in Section 4.2.2, for achromatic prisms which use combinations of materials, are also a promising option for removing dispersion although cost may again be prohibitive. The filter array used with the 16-channel broadband IRIS system in Section 4.3 does not entirely remove the birefringent dispersion in that system; however, the technique may be still be useful when combined with other methods of reducing dispersion. A software based correction of the chromatic aberration also appears possible, and as a low cost option for countering dispersion is another potentially promising area for future work. A final suggestion for future research is the investigation of optical designs for lens systems free from longitudinal chromatic aberration for use with IRIS systems, although this aberration can be corrected for potentially lower cost by using corrective glass elements as demonstrated in Section 4.3.

Appendix A

Propagation of light in birefringent media

Key to the accurate modeling of the propagation of light through an IRIS system is polarisation ray tracing through the birefringent wave plates and Wollaston prisms. Birefringent materials are uniaxial crystals, a class of electrically anisotropic crystals in which there exists a single axis of symmetry, it will be assumed that the materials are non-magnetic.

While it is straightforward to trace rays in linear isotropic dielectrics using Snell's law, the process in linear anistropic dielectrics is more complicated. Within a dielectric an electric field $\mathbf{E}$ acts to produce a dipole moment density per unit volume in the material $\mathbf{P}$ [125]. In a linear anisotropic medium P is related to the electric field through $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$ where

$$\chi = \begin{bmatrix} \chi_{xx} & \chi_{yx} & \chi_{zx} \\ \chi_{yx} & \chi_{yy} & \chi_{zy} \\ \chi_{zx} & \chi_{yz} & \chi_{zz} \end{bmatrix}$$

is the susceptability tensor.

Maxwell's equations in matter in the absence of currents and free charges are

$$\nabla \cdot \mathbf{D} = 0 \qquad \nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\dot{\mathbf{B}} \qquad \nabla \times \mathbf{H} = \dot{\mathbf{D}}$$

with $\mathbf{B} = \mu \mathbf{H}$ where μ is the permeability of the material. The electric displacement is given by $\mathbf{D} = \varepsilon_\mathbf{0} \mathbf{E} + \mathbf{P} = \varepsilon \mathbf{E}$ where $\varepsilon = \varepsilon_0 (1 + \chi)$ is the dielectric tensor.

Plane wave solutions to Maxwell's equations within the dielectric can be represented by

$$\mathbf{E} = \mathbf{E_0} e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

$$\mathbf{H} = \mathbf{H_0} e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

where $\mathbf{k} = n\omega\hat{\mathbf{k}}\Big/c$ is the wave vector and ω the angular frequency of the wave. For these wave solutions

$$\nabla \times \mathbf{E} = i\mathbf{k} \times \mathbf{E} = c^{-1}in\omega\hat{\mathbf{k}} \times \mathbf{E}, \quad \nabla \times \mathbf{H} = i\mathbf{k} \times \mathbf{H} = c^{-1}in\omega\hat{\mathbf{k}} \times \mathbf{H} \ ,$$

and

$$\dot{\mathbf{E}} = i\omega\mathbf{E}, \quad \dot{\mathbf{H}} = i\omega\mathbf{H}.$$

Inserting the wave solutions into the vector Maxwell-relations therefore gives

$$c^{-1}n\hat{\mathbf{k}} \times \mathbf{E} = \mu\mathbf{H}, \qquad c^{-1}n\hat{\mathbf{k}} \times \mathbf{H} = -\mathbf{D},$$

eliminating $\mathbf{H}$,

$$\mathbf{D} = \frac{-n^2}{\mu c^2}\hat{\mathbf{k}} \times \left(\hat{\mathbf{k}} \times \mathbf{E}\right) = \frac{-n^2}{\mu c^2}\left[\mathbf{E} - \hat{\mathbf{k}}\left(\hat{\mathbf{k}} \cdot \mathbf{E}\right)\right]. \tag{A.1}$$

$\mathbf{D}$, $\mathbf{E}$ and $\hat{\mathbf{k}}$ are therefore perpendicular to $\mathbf{H}$, and $\mathbf{D}$ is perpendicular to $\hat{\mathbf{k}}$ but in general $\mathbf{E}$ is not. From Poynting's theorem energy transport is in the direction of $\mathbf{E} \times \mathbf{H}$, this is also the direction of the light rays. This means the wave vector and the ray vector are not in general collinear in an anisotropic dielectric.

The dielectric tensor is symmetric [126] and can thus be diagonalised, the axes of a coordinate system in which this tensor is diagonal are known as principle dielectric axes. The diagonal elements, ε_i, ε_j, ε_k of the tensor in such a system are known as the principle dielectric constants of the material. In a uniaxial crystal two of these constants are equal, e.g., $\varepsilon_j = \varepsilon_k = \varepsilon_o$ and $\varepsilon_i = \varepsilon_e$. In a principle coordinate system Equation A.1 becomes

$$\varepsilon_i E_i = -\frac{n^2}{\mu c^2}\left[E_i - \hat{k}_i\left(\mathbf{E} \cdot \hat{\mathbf{k}}\right)\right]$$

$$E_i = \frac{n^2\hat{k}_i\left(\mathbf{E} \cdot \hat{\mathbf{k}}\right)}{n^2 - \mu\varepsilon_i c^2} = \frac{n^2\hat{k}_i\left(\mathbf{E} \cdot \hat{\mathbf{k}}\right)}{n^2 - \varepsilon_i/\varepsilon_0}$$

$$\mathbf{E} \cdot \hat{\mathbf{k}} = \sum_i \frac{n^2\hat{k}_i\left(\mathbf{E} \cdot \hat{\mathbf{k}}\right)}{n^2 - \varepsilon_i/\varepsilon_0}$$

$$\frac{1}{n^2} = \sum_i \frac{\hat{k}_i}{n^2 - \varepsilon_i/\varepsilon_0}$$

This final equation is quadratic in n so there are two refractive indices and therefore two sets of fields for each propagation direction.

The directions of propagation of the the two electric displacements $\mathbf{D_1}$ and $\mathbf{D_2}$ can be found [127] by inserting

$$\mathbf{E} = \varepsilon^{-1}\mathbf{D} = \frac{1}{\varepsilon_0}\eta\mathbf{D},$$

with $\eta = \varepsilon_0 \varepsilon^{-1}$, into Equation A.1,

$$\frac{1}{n^2}\mathbf{D} = \frac{-1}{\mu c^2}\hat{\mathbf{k}} \times \left(\hat{\mathbf{k}} \times \mathbf{E}\right) = \frac{-1}{\mu c^2 \varepsilon_0}\hat{\mathbf{k}} \times \left(\hat{\mathbf{k}} \times \eta\mathbf{D}\right)$$

$$\mathbf{k} \times (\mathbf{k} \times \eta\mathbf{D}) + \frac{1}{n^2}\mathbf{D} = 0$$

Working in a Cartesian coordinate system in which $\mathbf{k}$ is parallel to the z-axis, thereby making $\mathbf{D_1}$ and $\mathbf{D_2}$ perpendicular to this axis, the above can be written as

$$\begin{pmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{pmatrix} \mathbf{D} = \eta_t \mathbf{D} = \frac{1}{n^2}\mathbf{D}$$

$\mathbf{D_1}$ and $\mathbf{D_2}$ are thus eigenvectors of η_t with eigenvalues $1/{n_1}^2$ and $1/{n_2}^2$ respectively. $\mathbf{D_1}$ and $\mathbf{D_2}$ are therefore orthogonal to one another.

The matrix η describes a spheroid known as the index ellipsoid, in a principle coordinate system of the crystal

$$\eta = \varepsilon_0 \begin{bmatrix} \frac{1}{\varepsilon_e} & 0 & 0 \\ 0 & \frac{1}{\varepsilon_o} & 0 \\ 0 & 0 & \frac{1}{\varepsilon_o} \end{bmatrix} = \begin{bmatrix} \frac{1}{{n_e}^2} & 0 & 0 \\ 0 & \frac{1}{{n_o}^2} & 0 \\ 0 & 0 & \frac{1}{{n_o}^2} \end{bmatrix}$$

η_t describes an elliptical cross section through the center of this spheroid. As $\mathbf{D_1}$ and $\mathbf{D_2}$ are eigenvectors of η_t, they lie along the principle axes of this ellipse, and therefore in two of the principle planes of the spheroid. Figure A.1 illustrates the geometry for the case that $\varepsilon_e > \varepsilon_o$ in which case the spheroid is oblate, if $\varepsilon_o > \varepsilon_e$ the ellipsoid is prolate. The direction corresponding to the rotational axis of symmetry is known as the optic axis of the crystal. $\mathbf{D}_1 = \mathbf{D}_o$ lies perpendicular to the optic axis in the circular principle-plane, and is termed an ordinary wave, $\mathbf{D}_2 = \mathbf{D}_e$ is perpendicular to $\mathbf{D}_o$ and $\mathbf{k}$ and is known as an extraordinary wave. The ordinary refractive index $n_1 = n_o$ corresponds to the ordinary wave and is constant for all angles of propagation. The refractive index for the extraordinary wave on the other hand depends on the angle of propagation within the crystal and reaches an extremum of $n_1 = n_e$ when $\mathbf{k}$ is perpendicular to the optic axes, n_e is known as the extraordinary refractive index of the material. To find $n_1 = n(\phi)$ when the angle between the wave vector and the optic axis is ϕ consider the construction in Figure A.2, from this $n(\phi)$ is given by

$$\frac{1}{n_e^2(\phi)} = \frac{\sin^2\phi}{n_e^2} + \frac{\cos^2\phi}{n_o^2}.$$

$$n_e(\phi) = \frac{n_e n_o}{\left(n_o^2 \sin^2\phi + n_e^2 \cos^2\phi\right)^{1/2}}. \tag{A.2}$$

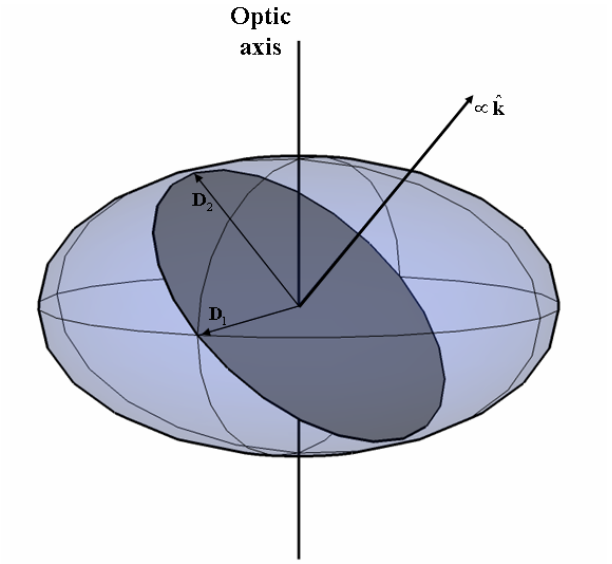

Figure A.1: Spheroid illustrating the directions of the electric displacements

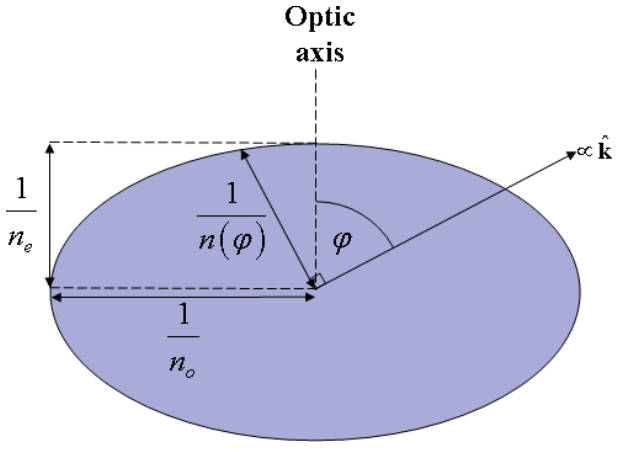

Figure A.2: Construction to determine refractive index for extraordinary wave.

With the wave vectors known, the ray vectors $\mathbf{s_o}$ and $\mathbf{s_e}$ can be found from the relation $\mathbf{D} = \varepsilon\mathbf{E}$. Choosing a coordinate system in which the optic axis is parallel to the x-axis and the extraordinary wave $\mathbf{D_e}$ is in the $x - y$ plane (so that the ordinary wave $\mathbf{D_o}$ is $y - z$ plane), the relationship between $\mathbf{E}$ and $\mathbf{D}$ can be written as

$$\mathbf{E} = \begin{bmatrix} \frac{1}{\varepsilon_e} & 0 & 0 \\ 0 & \frac{1}{\varepsilon_o} & 0 \\ 0 & 0 & \frac{1}{\varepsilon_o} \end{bmatrix} \mathbf{D}. \tag{A.3}$$

$\mathbf{E_o}$ (and hence $\mathbf{s_o}$) is rotated with respect to $\mathbf{D_o}$ ($\mathbf{k_o}$) in the $y - z$ plane by

$$\begin{bmatrix} \frac{1}{\varepsilon_o} & 0 \\ 0 & \frac{1}{\varepsilon_o} \end{bmatrix}, \tag{A.4}$$

so $\mathbf{s_o}$ is collinear with $\mathbf{k_o}$. $\mathbf{E_e}$ (and hence $\mathbf{k_e}$) is rotated with respect to $\mathbf{D_e}$ ($\mathbf{k_e}$) in the $y - z$ plane by

$$\begin{bmatrix} \frac{1}{\varepsilon_e} & 0 \\ 0 & \frac{1}{\varepsilon_o} \end{bmatrix}. \tag{A.5}$$

Equations A.3 and A.5, together with Snell's law, enable geometric ray tracing in birefringent media.

Bibliography

[1] S. M. Prokudin-Gorskii. Po r. Karolitskhali ("On the Karolitskhali River"). *Library of Congress Prints and Photographs Division Washington, D.C. 20540 USA*, Created between 1905 and 1915.

[2] A. Lumière and L. Lumière. Photographic plate for color imaging. June 1906. US Patent 822.532.

[3] B.E. Bayer. Color imaging array. July 1976. US Patent 3971065.

[4] M. Françon and S. Mallick. *Polarization Interferometers: Applications in Microscopy and Macroscopy*. Wiley-Interscience, New York, 1971.

[5] S. Prahl. Optical absorption of hemoglobin. *Oregon Medical Laser Center*, 2002.

[6] S. Tohyama, M. Miyoshi, S. Kurashina, N. Ito, T. Sasaki, A. Ajisawa, Y. Tanaka, A. Kawahara, K. Iida, and N. Oda. New thermally isolated pixel structure for high-resolution (640 × 480) uncooled infrared focal plane arrays. *Optical Engineering*, 45(1):014001, 2006.

[7] A. Rogalski. Infrared detectors: status and trends. *Progress in Quantum Electronics*, 27(2-3):59–210, 2003.

[8] E.L. Dereniak and G.D. Boreman. *Infrared detectors and systems*. Wiley-Interscience, New York, 1996.

[9] A. Rogalski. Infrared detectors: an overview. *Infrared physics and technology*, 43(3-5):187–210, 2002.

[10] A. Harvey and D. Fletcher-Holmes. Birefringent Fourier-transform imaging spectrometer. *Optics Express*, 12(22):5368–5374, 2004.

[11] A. Gorman, D.W. Fletcher-Holmes, and A.R. Harvey. Generalization of the Lyot filter and its application to snapshot spectral imaging. *Optics Express*, 18(6):5602–5608, 2010.

[12] R.C. Dougal, C.A. Greated, and A.E. Marson. Then and now: James Clerk Maxwell and colour. *Optics and Laser Technology*, 38(4-6):210–218, 2006.

[13] R.M. Evans. Maxwell's color photograph. *Scientific American*, 205(5):118–127, 1961.

[14] G. Lippmann. La photographie des couleurs. *Comptes Rendus de L' Académie des Sciences*, 112:274–275, 1891.

[15] R. Hain, C.J. Kahler, and C. Tropea. Comparison of CCD, CMOS and intensified cameras. *Experiments in Fluids*, 42(3):403–411, 2007.

[16] B. Matthews. Ultrahigh-resolution electro-optical framing camera for reconnaissance and other applications using a 9216 x 9216 pixel wafer-scale focal plane array. *Proceedings of SPIE*, 3431:144, 1998.

[17] J. Fischer, F. Vrba, D. Toomey, R. Lucke, S. Wang, A. Henden, J. Robichaud, P. Onaka, B. Hicks, F. Harris, et al. ASTROCAM: An Offner Re-imaging 1024x1024 InSb Camera for Near-Infrared Astrometry on the USNO 1.55-m Telescope. *Proceedings of SPIE*, 4841:564–577, 2003.

[18] K. Vural, L.J. Kozlowski, D.E. Cooper, C.A. Chen, G.L. Bostrup, C.A. Cabelli, J.M. Arias, J. Bajaj, K.W. Hodapp, D.N.B. Hall, et al. 2048x2048 HgCdTe focal plane arrays for astronomy applications. *Proceedings of SPIE*, 3698:24, 1999.

[19] M.H. Ettenberg, M.J. Lange, M.T. O'Grady, J.S. Vermaak, M.J. Cohen, and G.H. Olsen. A room temperature 640x512 pixel near-infrared InGaAs focal plane array. *Proceedings of SPIE*, 4028:201–207, 2000.

[20] E.N. Lewis, L.H. Kidder, E. Lee, and K.S. Haber. *Spectrochemical Analysis Using Infrared Multichannel Detectors*, chapter Near-infrared spectral imaging with focal plane array detectors, page 25. Blackwell Publishing, Oxford, UK, 2007.

[21] S.V. Bandara, S.D. Gunapala, D.Z. Ting, J.K. Liu, C.J. Hill, J.M. Mumolo, and S. Keo. Monolithically integrated near-infrared and mid-infrared detector array for spectral imaging. *Infrared Physics and Technology*, 50(2-3):211–216, 2007.

[22] I. Grave, A. Shakouri, N. Kuze, A. Yariv, et al. Voltage-controlled tunable GaAs/AlGaAs multistack quantum well infrared detector. *Applied Physics Letters*, 60(19):2362–2364, 1992.

[23] A. Majumdar, K.K. Choi, J.L. Reno, L.P. Rokhinson, and D.C. Tsui. Two-color quantum-well infrared photodetector with voltage tunable peaks. *Applied Physics Letters*, 80(5):707–709, 2002.

[24] C.J. Chen, K.K. Choi, W.H. Chang, and D.C. Tsui. Two-color corrugated quantum-well infrared photodetector for remote temperature sensing. *Applied Physics Letters*, 72(1):7–9, 1998.

[25] M. Kaldirim, S.U. Eker, Y. Arslan, U. Tumkaya, and C. Besikci. Large-Format Voltage-Tunable Dual-Color Midwavelength Infrared Quantum-Well Infrared Photodetector Focal Plane Array. *IEEE Photonics Technology Letters*, 20(9-12):709–711, 2008.

[26] S.V. Bandara, S.D. Gunapala, J.K. Liu, S.B. Rafol, D.Z. Ting, J.M. Mumolo, R.W. Chuang, T.Q. Trinh, J.H. Liu, K.K. Choi, et al. Four-band quantum well infrared photodetector array. *Infrared Physics and Technology*, 44(5-6):369–375, 2003.

[27] A. Soibel, S.V. Bandara, D.Z. Ting, J.K. Liu, J.M. Mumolo, S.B. Rafol, W.R. Johnson, D.W. Wilson, and S.D. Gunapala. A super-pixel QWIP focal plane array for imaging multiple waveband temperature sensor. *Infrared Physics and Technology*, 56(6):403–407, 2009.

[28] P. Verhoeve, N. Rando, A. Peacock, A. Van Dordrecht, A. Poelaert, and D.J. Goldie. Superconducting tunnel junctions as photon counting detectors in the infrared to the ultraviolet. *IEEE Transactions on Applied Superconductivity*, 7(2):3359–3362, 1997.

[29] A.J. Miller, B. Cabrera, R.W. Romani, E. Figueroa-Feliciano, S.W. Nam, and R.M. Clarke. Development of wide-band, time and energy resolving, optical photon detectors with application to imaging astronomy. *Nuclear Instruments and Methods in Physics Research Section A*, 444(1-2):445–448, 2000.

[30] B.A. Mazin, P.K. Day, H.G. LeDuc, A. Vayonakis, and J. Zmuidzinas. Superconducting kinetic inductance photon detectors. *Proceedings of SPIE*, 4849:203, 2002.

[31] J. Zmuidzinas, A. Vayonakis, P.K. Day, H.G. LeDuc, and B.A. Mazin. A broadband superconducting detector suitable for use in large arrays. *Nature*, 425:817–821, 2003.

[32] P. Verhoeve, N. Rando, A. Peacock, D. Martin, and R. den Hartog. Superconducting tunnel junctions as photon-counting imaging spectrometers from the optical to the X-ray band. *Optical Engineering*, 41:1170, 2002.

[33] A.M. Kadin, D.K. Brock, and D. Gupta. Superconducting Digital Multiplexers for Sensor Arrays. Technical report, NASA, 2004.

[34] M. Saffian. Optical system for a color television camera, 1965. US Patent 3202039.

[35] A.R. Harvey, J. Beale, A.H. Greenaway, T.J. Hanlon, and J. Williams. Technology options for imaging spectrometry. *Proceedings of SPIE*, 4132:14–24, 2000.

[36] L.J. Rickard, R.W. Basedow, E.F. Zalewski, P.R. Silverglate, and M. Landers. HYDICE: An airborne system for hyperspectral imaging. *Proceedings of SPIE*, 173(1937), 1993.

[37] R.S. Balaban, I. Kurtz, H.E. Cascio, and P.D. Smith. Microscopic spectral imaging using a video camera. *Journal of Microscopy*, 141(Pt 1):31, 1986.

[38] A.A. Sterk. Multispectral pushbroom array imager. *Sensor Systems for the 80's Conference, Colorado Springs, Colorado, December 2-4, 1980, Technical Papers.*, pages 109–112, 1980.

[39] L.N. Hadley and D.M. Dennison. Reflection and transmission interference filters. *Journal of the Optical Society of America*, 37(6):451–453, 1947.

[40] B. Lyot. Le filtre monochromatique polarisant et ses applications en physique solaire. *Annales d'Astrophysique*, 7:31–79, 1944.

[41] J.W. Evans. Solc birefringent filter. *Journal of the Optical Society of America*, 48(3):142–143, 1958.

[42] G. Shabtay, E. Eidinger, Z. Zalevsky, D. Mendlovic, and E. Marom. Tunable birefringent filters-optimal iterative design. *Optics Express*, 10(26):1534–1541, 2002.

[43] W.J. Marinelli, C.M. Gittins, A.H. Gelb, and B.D. Green. Tunable Fabry–Perot etalon-based long-wavelength infrared imaging spectroradiometer. *Applied Optics*, 38(12):2594–2604, 1999.

[44] N. Gat. Imaging spectroscopy using tunable filters: a review. *Proceedings of SPIE*, 4056:50–64, 2000.

[45] R.W. Slawson, Z. Ninkov, and E.P. Horch. Hyperspectral imaging: wide-area spectrophotometry using a liquid-crystal tunable filter. *Publications of the Astronomical Society of the Pacific*, 111(759):621–626, 1999.

[46] L. Bei, G.I. Dennis, H.M. Miller, T.W. Spaine, and J.W. Carnahan. Acousto-optic tunable filters: fundamentals and applications as applied to chemical analysis techniques. *Progress in Quantum Electronics*, 28(2):67–87, 2004.

[47] A.A. Michelson. Application of Interference Methods to Spectroscopic Measurement. *Publications of the Astronomical Society of the Pacific*, 4(25):190–192, 1892.

[48] C.W. Wells, A.E. Potter, and T.H. Morgan. Near-infrared spectral imaging Michelson interferometer for astronomical applications. In *Infrared imaging systems technology; Proceedings of the Seminar, Washington, DC, April 10, 11, 1980*,

pages 61–64, Bellingham, WA. Society of Photo-Optical Instrumentation Engineers.

[49] T. Okamoto, S. Kawata, and S. Minami. Fourier transform spectrometer with a self-scanning photodiode array. *Applied Optics*, 23(2):269–273, 1984.

[50] J.G. Hirschberg and E. Kohen. Pentaferometer: a solid Sagnac interferometer. *Applied Optics*, 38(1), 1999.

[51] M.J. Padgett and A.R. Harvey. A static Fourier-transform spectrometer based on Wollaston prisms. *Review of Scientific Instruments*, 66(4):2807, 1995.

[52] M.L. Nischan, R.M. Joseph, J.C. Libby, and J.P. Kerekes. Active spectral imaging. *Lincoln Laboratory Journal*, 14(1):131–144, 2003.

[53] I. Alabboud, A. McNaught, D. Mordant, and AR Harvey. Quantitative spectral imaging of the retina. In *ARVO Annual Meeting Fort Lauderdale, Florida*, 2007.

[54] K. Grieve, P. Tiruveedhula, Y. Zhang, and A. Roorda. Multi-wavelength imaging with the adaptive optics scanning laser ophthalmoscope. *Optics Express*, 14(25):12230–12242, 2006.

[55] B. Johnson, R. Joseph, M. Nischan, A. Newbury, J. Kerekes, H. Barclay, B. Willard, and J. Zayhowski. A compact, active hyperspectral imaging system for the detection of concealed targets. *Proceedings of SPIE*, 3710:144–153, 1999.

[56] J.B. Oke. Absolute spectral energy distributions in stars. *Annual Review of Astronomy and Astrophysics*, 3(1):23–46, 1965.

[57] C.A. Clayton. The implications of image scrambling and focal ratio degradation in fibre optics on the design of astronomical instrumentation. *Astronomy and Astrophysics (ISSN 0004-6361)*, 213(1-2), 1989.

[58] D.W. Fletcher-Holmes and A.R. Harvey. Real-time imaging with a hyperspectral fovea. *Journal of Optics A: Pure and Applied Optics*, 7(6):S298–S302, 2005.

[59] D. Rock, M. Roth, and S.F. Sagan. Broadband spectrometer with fiber optic reformattor, July 7 1987. US Patent 4678332.

[60] L. Gao, R.T. Kester, and T.S. Tkaczyk. Compact Image Slicing Spectrometer (ISS) for hyperspectral fluorescence microscopy. *Optics Express*, 17(15):12293–12308, 2009.

[61] M. Descour and E. Dereniak. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Applied Optics*, 34(22):4817–4826, 1995.

[62] L. A. Shepp and Y. Vardi. Maximum Likelihood Reconstruction for Emission Tomography. *IEEE Transactions on Medical Imaging*, 1(2):113–122, 1982.

[63] E. Levitan and G.T. Herman. A Maximum A Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography. *IEEE Transactions on Medical Imaging*, MI-6:185–192, 1987.

[64] M.E. Gehm, R. John, D.J. Brady, R.M. Willett, and T.J. Schulz. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics Express*, 15(21):14013–14027, 2007.

[65] A. Wagadarikar, R. John, R. Willett, and D. Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied Optics*, 47(10):44–51, 2008.

[66] R.O. Green, M.L. Eastwood, C.M. Sarture, T.G. Chrien, M. Aronsson, B.J. Chippendale, J.A. Faust, B.E. Pavri, C.J. Chovit, M. Solis, et al. Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer(AVIRIS). *Remote Sensing of Environment*, 65(3):227–248, 1998.

[67] A.W. Nolin and J. Dozier. Estimating snow grain size using AVIRIS data. *Remote Sensing of Environment*, 44(2):231–238, 1993.

[68] R.O. Green and J. Dozier. Retrieval of Surface Snow Grainsize and Melt Water from AVIRIS Spectra. Technical report, JPL, Pasadena, California, 1996.

[69] N.A. Drake, S. Mackin, and J.J. Settle. Mapping vegetation, soils, and geology in semiarid shrublands using spectral matching and mixture modeling of SWIR AVIRIS imagery. *Remote Sensing of Environment*, 68(1):12–25, 1999.

[70] M.K. Hamilton, C.O. Davis, W.J. Rhea, S.H. Pilorz, and K.L. Carder. Estimating chlorophyll content and bathymetry of Lake Tahoe using AVIRIS data. *Remote sensing of environment*, 44(2-3):217–230, 1993.

[71] R.F. Kokaly, R.N. Clark, and K.E. Livo. Mapping the biology and mineralogy of Yellowstone National Park using imaging spectroscopy. In *JPL Airborne Earth Science Workshop, 7: 97–21*, 1998.

[72] R.O. Green. Measuring the spectral expression of carbon dioxide in the solar reflected spectrum with AVIRIS. In *AVIRIS Workshop, JPL/NASA, Pasadena, California*, 2001.

[73] M.E. Martin, S.D. Newman, J.D. Aber, and R.G. Congalton. Determining forest species composition using high spectral resolution remote sensing data. *Remote Sensing of Environment*, 65(3):249–254, 1998.

[74] R. Green, B. Pavri, D. Roberts, and S. Ustin. Mapping Agricultural Crops with AVIRIS Specta in Washington State. Technical report, Earth-Science Airborne Workshop, Pasadena, California, USA, 1999.

[75] R.F. Kokaly, D.G. Despain, R.N. Clark, and K.E. Livo. Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data. *Remote Sensing of Environment*, 84(3):437–456, 2003.

[76] J.C. Lansing Jr and R.W. Cline. The four-and five-band multispectral scanners for Landsat. *Optical Engineering*, 14:312–322, 1975.

[77] J.L. Bezy, S. Delwart, and M. Rast. MERIS-A new generation of ocean-colour sensor onboard Envisat. *ESA bulletin*, 103:48–56, 2000.

[78] E. Schrock, S. Du Manoir, T. Veldman, B. Schoell, J. Wienberg, MA Ferguson-Smith, Y. Ning, DH Ledbetter, I. Bar-Am, D. Soenksen, et al. Multicolor spectral karyotyping of human chromosomes. *Science*, 273(5274):494–497, 1996.

[79] C. Thaler, S.V. Koushik, P.S. Blank, and S.S. Vogel. Quantitative multiphoton spectral imaging and its use for measuring resonance energy transfer. *Biophysical journal*, 89(4):2736–2749, 2005.

[80] T. Vo-Dinh, B. Cullum, and P. Kasili. Development of a multi-spectral imaging system for medical applications. *Journal of Physics D: Applied Physics*, 36(14):1663, 2003.

[81] M.E. Martin, M.B. Wabuyele, K. Chen, P. Kasili, M. Panjehpour, M. Phan, B. Overholt, G. Cunningham, D. Wilson, R.C. DeNovo, et al. Development of an advanced hyperspectral imaging (HSI) system with applications for cancer detection. *Annals of biomedical engineering*, 34(6):1061–1068, 2006.

[82] I.B. Styles, A. Calcagni, E. Claridge, F. Orihuela-Espina, and J.M. Gibson. Quantitative analysis of multispectral fundus images. *Medical Image Analysis*, 10(4):578–597, 2006.

[83] D.J. Mordant, I. Al-Abboud, A.R. Harvey, and A.I. McNaught. Hyperspectral Imaging of the Human Retina–Oximetric Studies. In *Investigative Ophthalmology & Visual Science*, volume 48, pages E–148, 2007.

[84] M.S. Dombrowski, P.D. Willson, and C.C. LaBaw. Defeating camouflage and finding explosives through spectral matched filtering of hyperspectral imagery. *Proceedings of SPIE*, 2933:109, 1997.

[85] N. Gat, G. Scriven, J. Garman, M. De Li, and J. Zhang. Development of four-dimensional imaging spectrometers (4D-IS). *Proceedings of SPIE*, 6302, 2006.

[86] H. Ren and C.I. Chang. Target-constrained interference-minimized approach to subpixel target detection for hyperspectral images. *Optical Engineering*, 39(12):3138–3145, 2000.

[87] R.S. Berns, F.H. Imai, P.D. Burns, and D.Y. Tzeng. Multi-spectral-based color reproduction research at the Munsell Color Science Laboratory. *Proceedings of SPIE*, 3049:14–25, 1998.

[88] F.H. Imai and R.S. Berns. Spectral estimation of artist oil paints using multifilter trichromatic imaging. *Proceedings of SPIE*, pages 504–507, 2002.

[89] R.C. Lyon, D.S. Lester, E.N. Lewis, E. Lee, L.X. Yu, E.H. Jefferson, and A.S. Hussain. Near-infrared spectral imaging for quality assurance of pharmaceutical products: analysis of tablets to assess powder blend homogeneity. *AAPS PharmSciTech*, 3(3):1–15, 2002.

[90] K. Chao, P.M. Mehl, and Y.R. Chen. Use of hyper-and multi-spectral imaging for detection of chicken skin tumors. *Applied Engineering in Agriculture*, 18(1):113–120, 2002.

[91] D.W. Fletcher-Holmes and Harvey A.R. The feasibility of a LWIR IRIS II system.

[92] Marvin J. Weber. *Handbook of Optical Materials*. CRC Press, Boca Raton, FL, 2003.

[93] D.B. Chenault and R.A. Chipman. Infrared birefringence spectra for cadmium sulfide and cadmium selenide. *Applied Optics*, 32(22):4223–4227, 1993.

[94] D.A. Roberts. Dispersion equations for nonlinear optical crystals: KDP, AgGaSe_2, and AgGaS_2. *Applied Optics*, 35(24):4677–4688, 1996.

[95] D.E. Zelmon, J.J. Lee, K.M. Currin, J.M. Northridge, and D. Perlov. Revisiting the optical properties of Nd doped yttrium orthovanadate. *Applied Optics*, 49(4):644–647, 2010.

[96] E. Oliva, S. Gennari, L. Vanzi, A. Caruso, and M. Ciofini. Optical materials for near infrared Wollaston prisms. *Astronomy and Astrophysics Supplement Series*, 123(1):179–182, 1997.

[97] J.J. Loferski. Infrared optical properties of single crystals of tellurium. *Physical Review*, 93(4):707–716, 1954.

[98] G.A. Knyazev and V.B. Voloshinov. Diffraction of IR radiation by ultrasound in tellurium single crystals. *Bulletin of the Russian Academy of Sciences: Physics*, 72(12):1643–1647, 2008.

[99] R.N. Smartt and J.V. Ramsay. On the production and use of the optical contact bond. *Journal of Scientific Instruments*, 41(8):514, 1964.

[100] X. Dai and S. Khorram. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):1566–1577, September 1998.

[101] J.A. Parker, R.V. Kenyon, and D.E. Troxel. Comparison of interpolating methods for image resampling. *IEEE Transactions on Medical Imaging*, 2(1):31–39, 1983.

[102] M. Bass, E. W. V. Stryland, D. R. Williams, and W. L. Wolfe. *Handbook of Optics Volume I: Fundamentals, Techniques, and Design*. McGraw-Hill, New York, 2nd edition, 1995.

[103] Q.T. Liang. Simple ray tracing formulas for uniaxial optical crystals. *Applied Optics*, 29(7):1008–1010, 1990.

[104] M.C. Simon and K.V. Gottschalk. Optical path in birefringent media and Fermat's principle. *Pure and Applied Optics: Journal of the European Optical Society Part A*, 7(6):1403, 1998.

[105] M. Feldman, A. El-Amawy, A. Srivastava, and R. Vaidyanathan. Adjustable Wollaston-like prisms. *Review of Scientific Instruments*, 77(6):066109, June 2006.

[106] Z. Burshtein, Y. Shimony, and I. Levy. Refractive-index anisotropy and dispersion in gehlenite, Ca2Al2SiO7 between 308 and 1064 nm. *Journal of the Optical Society of America A*, 10(10):2246–2247, October 1993.

[107] Schott Glaswerke. *Schott catalogue of optical glass*.

[108] D. Klimm, G. Lacayo, and P. Reiche. Growth of Cr: LiCaAlF6 and Cr: LiSrAlF6 by the Czochralski method. *Journal of Crystal Growth*, 210(4):683–693, 2000.

[109] D. Kundur and D. Hatzinakos. Blind Image Deconvolution. *IEEE Signal Processing Magazine*, 13(3):43–64, 1996.

[110] D. Kundur and D. Hatzinakos. Blind image deconvolution revisited. *IEEE Signal Processing Magazine*, 13(6):61–62, 1996.

[111] C.D. Rodgers. *Inverse methods for atmospheric sounding: Theory and practice*. World Scientific Press, Singapore, 2000.

[112] C.D. Rodgers and B.J. Connor. Intercomparison of remote sounding instruments. *Journal of Geophysical Research*, 108(D3):4116–4229, 2003.

[113] C.D. Rodgers. Information content and optimisation of high spectral resolution remote measurements. *Advances in Space Research*, 21(3):361–367, 1998.

[114] K.C. Gross, C. Borel, A. White, S. Sakai, R. DeVasher, and G.P. Perram. First imaging Fourier-transform spectral measurements of detonation in an internal combustion engine. *Proceedings of SPIE*, 7812:78120J, 2010.

[115] N. Bellakhal, K. Draou, and JL Brisset. Electrochemical investigation of copper oxide films formed by oxygen plasma treatment. *Journal of Applied Electrochemistry*, 27(4):414–421, 1997.

[116] S.H. Hardarson, A. Harris, R.A. Karlsson, G.H. Halldorsson, L. Kagemann, E. Rechtman, G.M. Zoega, T. Eysteinsson, J.A. Benediktsson, A. Thorsteinsson, et al. Automatic retinal oximetry. *Investigative Ophthalmology & Visual Science*, 47(11):5011, 2006.

[117] W.R. Johnson, D.W. Wilson, W. Fink, M. Humayun, and G. Bearman. Snapshot hyperspectral imaging in ophthalmology. *Journal of Biomedical Optics*, 12(1):014036, 2007.

[118] J.C. Ramella-Roman, S.A. Mathews, H. Kandimalla, A. Nabili, D.D. Duncan, S.A. D'Anna, S.M. Shah, and Q.D. Nguyen. Measurement of oxygen saturation in the retina with a spectroscopic sensitive multi aperture camera. *Optics Express*, 16(9):6170–6182, 2008.

[119] S. Standring, H. Ellis, J. Healy, D. Johnson, and A. Williams. *Grays anatomy.*, volume 1112. London, 39th edition, 2005.

[120] T. Vo-Dinh and Masters B.M. Biomedical photonics handbook. *Journal of Biomedical Optics*, 9:1110, 2004.

[121] L.G. Henyey and J.L. Greenstein. Diffuse radiation in the galaxy. *Annales d'Astrophysique*, 3:117–137, 1940.

[122] M. Friebel, A. Roggan, G. M
"uller, and M. Meinke. Determination of optical properties of human blood in the spectral range 250 to 1100 nm using Monte Carlo simulations with hematocrit-dependent effective scattering phase functions. *Journal of Biomedical Optics*, 11(3):034021, 2006.

[123] M. Friebel, J. Helfmann, U. Netz, and M. Meinke. Influence of oxygen saturation on the optical scattering properties of human red blood cells in the spectral range 250 to 2000 nm. *Journal of Biomedical Optics*, 14(3):034001, 2009.

[124] K. Briely-Sæbø and A. Bjørnerud. Accurate de-oxygenation of ex vivo whole blood using sodium dithionite. *Proceedings of the 8th Annual Meeting, ISMRM, Denver*, 2000.

[125] D.J. Griffiths and C. Inglefield. *Introduction to electrodynamics*. Prentice Hall, New Jersey, 1999.

[126] M. Born and E. Wolf. *Principles of optics*. Pergamon press, Oxford, 7th edition, 1975.

[127] A. Yariv and P. Yeh. *Optical waves in crystals*. Wiley, New York, 1984.