Absorption spectral imaging in the mid-infrared and its application in cancer diagnosis

Laurence Alexander Drummond

Imperial College London, Department of Physics

Thesis submitted for examination for the PhD degree of Imperial College

Declaration of Originality

The work contained in this thesis is the author's own, and that of all else is appropriately referenced.

Copyright Declaration

Abstract

This thesis presents work on the development of two distinct systems for the purpose of spectral imaging in the mid infrared (MIR), with particular focus on a method than can be of use in cancer diagnosis. Wavelengths are used that correspond to positions of absorbance peaks from the vibrational modes in molecules commonly found in biological tissue. Images from different wavelengths are combined to show the relative concentration of nuclei in tissue, an accepted marker of malignancy that can be seen in cancerous cells. One system, we have termed the "*digistain*" system, uses a thermal source with narrow bandpass filters to select the wavelength, the other uses a pulsed tuneable quantum cascade laser (QCL) unit. Both systems show results that highlight a difference between healthy epithelium and epithelium that has entered a pre-cancerous stage. The *digistain* system has benefits in speed over other infrared spectral imaging systems, such as a scanning Raman system or a commercial multi-channel Fourier transform infrared (FTIR) system. Before a reliable direct comparison of the quantitative results of the *digistain* to histological grading of the same areas can be made, a method to identify the different tissue types that are present and mixed together must be created. The QCL system was also used for diffuse reflection imaging, with the conclusion that the differing surface roughness of the samples investigated played a larger role in creating contrast between areas of different chemical composition than did absorption.

Contents

List of figures and tables

1 Introduction

This thesis presents spectral imaging systems in the mid infrared (MIR) from 6–10µm wavelength, primarily for analysing the spectral absorption properties of biological tissue. The imaging takes place at discrete wavelengths, with the absorption attributed to molecular vibrational modes of functional groups and chemical bonds that are commonly found in biological tissue. There are two systems presented here – a laser-based system that uses a tuneable quantum cascade laser (QCL) to operate in both 'transmission' and 'reflection' mode, and the *digistain* system that uses a thermal source and bandpass filters. The *digistain* system is designed to look only at tissue, and in particular cancer samples, with the aim of being able to quantitatively measure the progress of the disease. The QCL system is more flexible in its choice of target sample, being able to tune to any wavelength in the 6–10µm region; the main samples looked at are the same cancer samples.

Many of the same operating principles and considerations are shared between the two systems, such as the optical elements to be used and methods to reduce the influence of environmental noise.

1.1 Motivation

Spectroscopic work in the MIR is often performed using a Fourier Transform Infrared (FTIR) setup. This has large advantages over other spectroscopic methods, as it allows the collection of an entire spectrum relatively quickly [1]. When we do not require an entire spectrum, it may be more beneficial to look at only a few discrete wavelengths. Using light sources such as tuneable laser sources can allow for higher radiance than a broadband thermal source at our chosen wavelengths, potentially allowing for a higher signal-to-noise ratio (SNR) [2]. Using a thermal source, selecting a small number of wavelengths can lead to faster data acquisition than taking a whole spectrum with an FTIR spectrometer (the exact time-saving will depend on factors such as the desired spectral resolution and number of wavelengths required) [3]. Using a quantum cascade laser (QCL) that outputs discrete wavelengths can similarly lead to improved data acquisitio times, with one estimate stating that imaging can be faster than an FTIR system by a factor of 1100/N, where N is the number of wavelengths to be collected [4].

A source that can tune across a wide spectral range can potentially look at any sample that has a characteristic absorption response in a given spectral range, but particular samples targeted in this work are oesophageal and breast cancer biopsies. This is in an attempt to account for shortcomings in the most common practice for diagnosing cancer – the inherent subjectivity of the histopathologist [5-7], the medical staff that carries out the ‘grading’ of tumours that contributes to assessing the risk to the patient.

1.1.1 Importance of a high SNR spectroscopic instrument

A key goal of this work is to address the shortcomings of current cancer diagnosis techniques by utilising the inherent benefits of working in the MIR when looking at biological molecules. As the data is all collected digitally, it can easily be processed. When taking an objective, quantitative approach, the accuracy and precision of the numbers is of great importance. A figure of merit that can help inform us how well we can trust any quantitative data is the SNR. The SNR can be increased by increasing acquisition time, but since a four times longer acquisition will only double SNR, increasing SNR by other means is preferable to keep acquisition times reasonable.

1.1.2 Cancer and its current diagnosis

The current process for the diagnosis of cancer is the H&E staining procedure, which allows medical staff to view a biopsy under a visible microscope and then use their expertise and some visual estimates to make a judgement on the state of disease, which forms a major role in deciding the best treatment option [8, 9]. The earlier the cancer is identified and diagnosed, the better the chances of survival for the patient [10, 11]. A biopsy for H&E staining is taken after the identification of tissue that may be cancerous. The way in which the H&E stained tissue is described in section 4.2.2.1.

1.1.2.1 H&E staining

Hematoxylin and eosin (H&E) staining is a staining technique used in histology, and is the current gold standard stain for histopatholgists to look at biopsy sections suspected of being cancerous. Two dyes are used to stain different parts of a cell – hemalum (derived from hematoxylin) primarily dyes the nuclei of cells a blue colour, while a solution of eosin Y dyes cellular proteins and most of the cytoplasm shades of red, pink and orange [12]. Additional dyes can be applied to highlight other cellular and tissue structures. The staining process is required as, under a visible microscope, a section thin enough to investigate at a cellular level would be transparent (without the presence of natural pigment such as melanin). H&E staining is not used exclusively for cancer diagnosis, and can be used to reveal cellular structure to help diagnose other diseases. An example of an H&E stained tissue biopsy in shown in Figure 1-1.

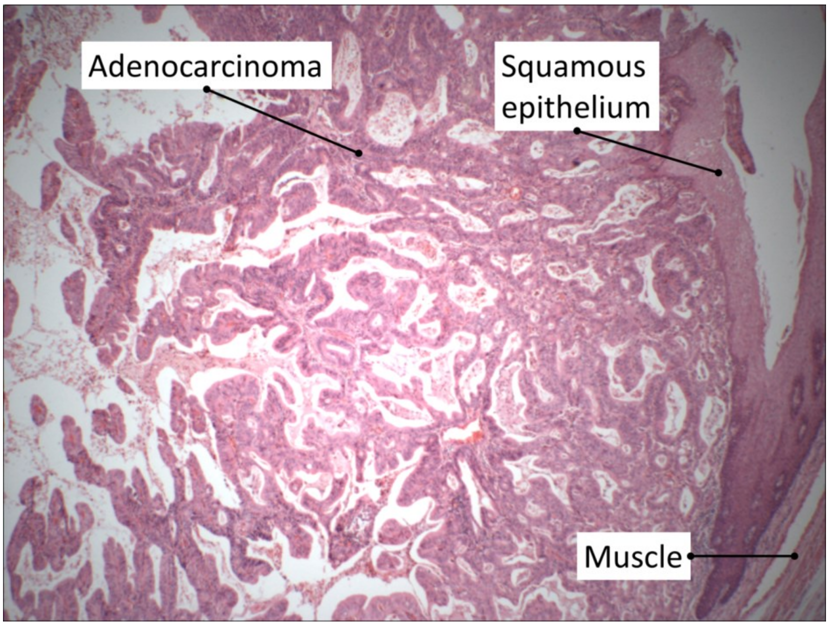


Figure 1-1. H&E stained oesophageal tissue biopsy. The size of the imaged area (height) is roughly 2mm. The majority of the tissue is adenocarcinoma (epithelial tissue of glandular origin that shows cancerous characteristics). The 'V' on the very right is squamous epithelium, and in the bottom right corner is shown muscle.

A dyed section is then looked at by an experienced histopathologist, who will identify the cells being looked at and reach a conclusion on which areas of the section may be in a cancerous stage, based on what they deem to be abnormal, largely through cell and tissue morphology.

There are other tests apart for visual inspection after H&E staining that can help in the diagnosis of cancer and the prognosis for the patient. Some of these tests are discussed in section 2.5.2.

1.1.2.2 Discrepancy in diagnoses from histopathologists

The main problem with the current H+E diagnosis technique is its subjective nature. Studies have shown that there can be disagreement between the diagnosis offered by different histopathologists [5-7]. One study of colorectal cancers showed that the agreement between three pathologists as to the grade of dysplasia was only 47.8% [5], whilst another study also looking at the intestinal area showed that for five pathologists, agreement over grade of dysplasia between any two pathologists ranged between 42% and 65% [7]. This second study showed that the best agreement was for biopsies that were said to contain no dysplasia. Additionally, there appeared to be a systematic bias within the pathologists' diagnoses, with some pathologists being more likely to diagnose a patient as having dysplasia than others, regardless of clinical outcome in developing carcinoma.

The grade of a cancer is an indicator that factors into a doctor's subsequent decision on what kind of treatment a patient should receive. If the cancer is thought to already be well developed, the patient may be recommended an aggressive and tolling treatment such as chemotherapy. It is best that a patient not be put through such a treatment if not necessary, and receives the appropriate treatment required. This is why the accurate diagnosis of disease is so important, to avoid unnecessary hardship or deaths.

Cancer diagnostics is an active field of research, and there are many emerging technologies and techniques for aiding in the diagnosis and prognosis of cancer [13-22]. Some of these are discussed in section 2.5.2.

1.1.2.3 Data on cancer deaths

Worldwide, cancer caused 8.2 million deaths in 2012, with 14.1 million new cases diagnosed [23]. 32.6 million people were living with cancer having been diagnosed within the previous 5 years [23]. Of the deaths due to cancer in 2012, the leading types were lung, liver, stomach, colorectal, breast and oesophageal cancer, in that order [23].

In the UK, 331,000 people were diagnosed with cancer in 2011 and it is estimated that over 1 in 3 people will develop a form of cancer during their lifetime [24]. In 2011 there were 159,000 deaths from cancer, more than 1 in 4 of all deaths [24]. However, half of people diagnosed with cancer now survive for over 5 years [24].

1.2 Previous work

The work presented in this thesis builds directly on previous work in this group, as presented in the theses of Hemmel Amrania [25] and Andrew McCrow [26], as well as one paper [27], using much of the same equipment as will be presented here.

Amrania's thesis made use of a tuneable MIR Optical Parametric Generator (OPG) used in conjunction with a modified commercial infrared microscope. Single cells were able to be imaged at MIR wavelengths with a temporal resolution of ~100ps. MIR spectra acquired with the OPG system was shown to have a correlation with those acquired using an FTIR spectrometer, and imaging at amide and phosphodiester peak wavelengths was achieved.

McCrow's thesis utilised the same OPG system, as well as a quantum cascade laser (QCL) to push further into the possibilities of cancer diagnosis, assessing the viability of getting numerical absorbance data with higher SNR values, with the conclusion that at the time, the QCL system with its higher SNR was better placed to give reliable readings.

The paper by Amrania *et al.* published in 2012 [27], to which the author of this thesis was a contributor, used a thermal source with bandpass filters to show the correlation between spectral features in the MIR and nucleus concentration within tissue.

1.3 Overview of *digistain* and QCL systems

The work performed on the *digistain* part of this thesis is a direct evolution of the work in the Amrania paper [27], with a focus on being able to distinguish and quantify the difference between healthy and cancerous tissue due to the chemical differences. As a result of the work reported here, the acquisition time to get an image that shows good contrast has been reduced by at least an order of magnitude (from over an hour to just a few minutes) and the data processing routine has been modified. Due to the introduction of germanium lenses, there are no longer shifts in magnification between wavelengths (germanium has relatively low dispersion in the MIR, see section 2.3). This improves overall image quality, as when images from different wavelengths are ratioed, the pixels correspond to the same areas on the sample. The system is now automated with a program written using LabVIEW software, with very little user interaction required. The SNR of the system has now reached a point where the acquisition times are faster than other infrared imaging systems (sometimes considerably so [28]). This is achieved despite a relatively 'low-tech' approach, with a broadband thermal source instead of a laser being used.

With the laser-based system, the QCL unit is used to produce the same type of images (a phosphodiester-to-amide ratio — "PA ratio" — image) that are possible with the *digistain* setup, for the first time. The main difficulties were in removing the coherence that introduces artefacts into the image and making sure that nonlinearities were not present that would affect the quantitative accuracy of the results. The QCL system is also used to make a first attempt at acquiring a diffuse reflection image. Although reflection images are obtained, it is questionable whether the contrast shown is due to the chemistry in the samples, and not just due to differences in surface roughness.

A generic setup for transmission imaging with the systems is shown in Figure 1-2. In both systems, the MIR source consists of more than one component – the laser-based system also has a diffuser to remove spatial coherence, whilst the *digistain* system has a shutter and a filter wheel with bandpass filters. A lens images the sample onto a MIR camera which sends the data to a computer for analysis. The sample can be moved in and out of the field of view by the data acquisition program for purposes of background correction, and is usually mounted on a substrate of barium fluoride or calcium fluoride which is transmissive in the MIR.

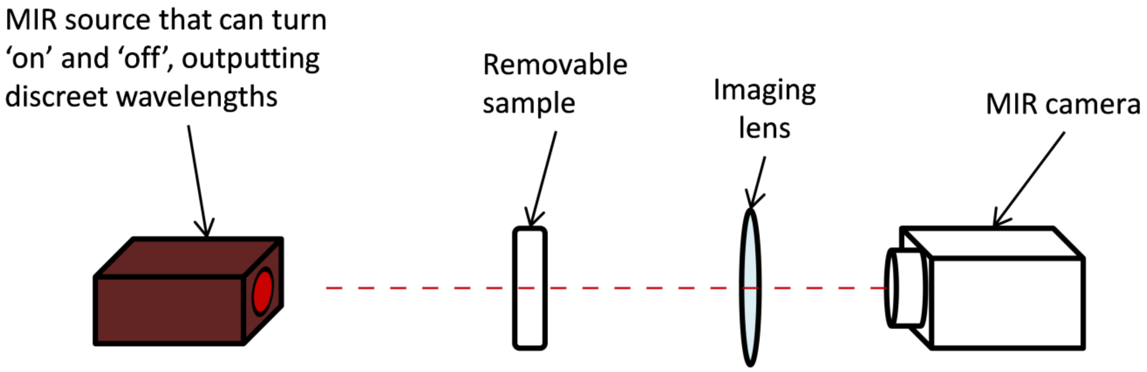


Figure 1-2. Setup for transmission imaging with the digistain and laser-based systems. In a simplified view, a source of MIR radiation is required that can provide light at discrete wavelengths is required. With the digistain, this is a thermal source with bandpass filters, whilst for the laser-based system it is a tuneable QCL. The MIR light is directed through the sample in which absorption occurs. For the purposes of background correction, the sample must be able to be moved in and out of the optical axis, and the MIR source must be to turn its output 'on' and 'off'. An imaging lens, in both systems made of germanium, is used to image the sample onto a MIR camera. Germanium is used to maintain the focus and position of the image over the MIR wavelengths – each pixel of the camera must be imaging the same part of the sample across the wavelengths to obtain reliable results when images at different wavelengths are combined.

With both systems, when imaging biological tissue samples, data is collected at wavelengths corresponding to MIR absorption peaks of a phosphodiester and amide molecular vibrations. The phosphodiester absorption peak gives information on the concentration of cell nuclei, the amide peak on the concentration of cellular proteins. This information is then combined in a ratio to give information on the relative concentration of cell nuclei, independent of sample thickness and cellular concentration. This is known as the phosphodiester-to-amide ratio.

This is of use, as it can be related to an aspect in which cancer cells differ from normal cells, the relative size and density of cell nuclei [27]. As cancerous cells divide more rapidly than normal cells, more cells within a tumour will be undergoing mitosis (cell division), or will have just completed division. Because of this, the nuclei in the cells will be large relative to the size of the cell. This difference within cancerous cells is commonly referred to by histopathologists as the 'NC ratio' (nuclear to cytoplasmic ratio). It is recognised that cancer cells will have a greater NC ratio [29], although the value is not currently used directly in the grading procedure. There are studies that have utilised algorithms to count the mean nuclear area of visible H&E stained images, a step towards calculating the NC ratio [21, 22]. By using the PA ratio acquired through infrared spectroscopy, the nuclear concentration can be digitised and quantified in a marker-free and non-subjective manner. The method of preparing samples for infrared spectroscopy is very similar to that used currently to prepare samples for viewing by a histopathologist, indicating that it would be a minimal change in this regard to use an infrared system.

1.4 Thesis Outline

Chapter 2 contains the bulk of the theory and background required to follow and understand this thesis as well as providing context to the methods used. It covers the physics of MIR absorption spectroscopy, spectroscopy techniques and considerations for spectral imaging as well as an introduction to the structure and makeup of tissue to aid in understanding the samples under study. The specific molecular vibrations being probed are detailed and a further look at the need for an instrument such as the *digistain* is taken and where it sits in relation to other techniques.

Chapter 3 looks at the specific hardware that comprises the entire systems, the data collection routines and processing, and some methods implemented in the software when displaying the results. It also characterises the systems, with a more in-depth look at how components work together. This chapter primarily gives insight into the design decisions taken as the system evolved.

Chapter 4 contains information on the particular samples being looked at as well as an explanation of more specific terminologies not covered in Chapter 2. Note that the samples looked at with the infrared systems are not the exact same samples as the H&E stained visible samples, but are instead samples taken as an adjacent slice from the patient's biopsy. The samples for the infrared have not been put through the staining procedure as the H&E samples. They are considered to be close enough spatially that the cell types and structures will be similar.

The results from the investigation of the tissue samples, and other samples investigated, are presented in Chapter 5, along with a discussion of these results.

Chapter 6 summarises the thesis and looks to the future continuation of this work.

2 Theory and background to research

This chapter's aims include providing an introduction to MIR absorption spectroscopy as well as a brief comparison of spectroscopy methods. Considerations for imaging in the MIR are discussed, as well as an overview of cellular and biological tissue structures, and molecular vibrational modes of interest. Lastly, the *digistain* technique and QCL system are put in context in regards to spectral imaging tools and methods of cancer diagnosis.

Different fields of science variably use wavenumber, wavelength, frequency or photon energy when discussing light that occupies different spectral regions. In this thesis, only wavenumber and wavelength will be used. The field of chemical spectroscopy tends to use wavenumber, defined as

$$wavenumber\ (units\ cm^{-1}) = \frac{10{,}000}{\lambda\ (units\ \mu m)} \qquad \text{2-1}$$

where λ is the wavelength of the light. This is also sometimes known as the 'spectroscopic wavenumber' to distinguish it from a definition of wavenumber commonly used in physics, the magnitude of the wavevector, $k = \frac{2\pi}{\lambda}$. Only the spectroscopic wavenumber will be used in this thesis. Wavelength is also used, as it is used more commonly in some relevant areas such as the sensitive range of a detector.

2.1 Introduction to spectroscopy

Spectroscopy is the wavelength-dependent study of light-matter interaction. Of particular interest in this thesis is absorption spectroscopy, whereby light is absorbed at a specific wavelength due to various mechanisms in the sample under study.

2.1.1 Absorption spectroscopy

Absorption peaks in the mid-infrared stem from molecular vibrations, such as those that can be found in biological tissue. Each absorption peak is due to a particular molecule being present, and is a result of a particular vibrational mode in a functional group being excited. The energy from the light drives the vibrational mode, putting its energy into the resonance [30].

A single molecule can have different vibrational modes, meaning that there can be multiple absorption peaks due to its presence. A molecule with N atoms has $(3N - 6)$ normal modes of vibration, whilst a linear molecule has $(3N-5)$ modes. The three modes of the 3-atom H_2O molecule are shown in Figure 2-1 as an example.

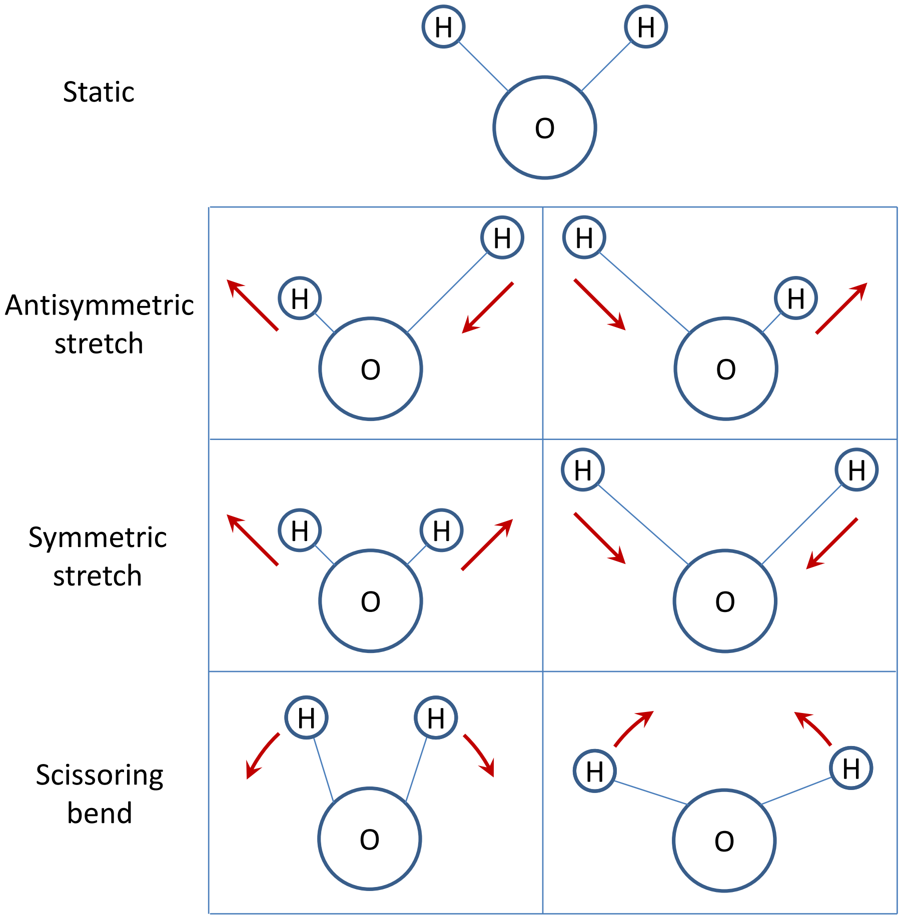


Figure 2-1. The vibrational modes of a 3-atom (nonlinear) molecule, H_2O, which has 3 distinct modes. Each image on the left and right shows a snapshot of the vibration of the molecule at their extreme positions, with arrows to indicate where the hydrogen atoms will next move. In the antisymmetric stretch mode, each 'arm' with a hydrogen atom oscillates out of phase. In the symmetric stretch mode, the hydrogen atoms change their distance from the oxygen at the same time. In the scissoring bend mode, the hydrogen atoms maintain their distance from the oxygen atom, but oscillate axially about their normal static position.

The strength of the absorption is proportional to the concentration of the molecule present, and the position of the peak (i.e. its vibrational frequency) can shift due to the structural properties, environments of the functional groups and isotropic effects. These peaks tend to be relatively wide (compared to laser linewidths or atomic transitions), and if close enough together, can overlap enough to create the appearance of a larger composite peak. In this case, the individual peaks can be separated through deconvolution methods [31].

Every molecule will have a unique series of absorption peaks across the MIR. Although the individual peak positions may overlap due to a shared common functional group, the combination of peaks will always differ, giving rise to 'spectral fingerprints' that can be used to identify the presence of certain molecules.

2.1.2 Spectral regions of interest for biological investigation

Table 2-1 shows the wavelengths that will be discussed in this section in the context of spectroscopy. Usually the division of the spectrum is split into ultraviolet, visible, infrared etc. but the infrared can also be subdivided. Bear in mind that the work in this thesis is only conducted in the mid infrared, and within this band, only between the wavelengths of 6 and 10µm.

Visible	**400-800nm**
Near Infrared	**0.8-2.5µm**
Mid Infrared	**2.5-25µm**
Far Infrared	**25-1000µm**

Table 2-1. Division of visible and IR spectral regions used in this thesis. The wavelengths used in this thesis fall in the middle of the mid infrared, between 6-10µm

Each region of the spectrum has its own advantages and suitable samples for investigation. The visible spectrum is useful for directly probing atomic transitions and plasmonic resonances in metals, and the instrumentation for experiments in the visible is relatively well developed. The use of fluorescent proteins has also proven adept at exploring cell mechanics [32, 33].

2.1.2.1 Infrared (IR)

The infrared region of the spectrum is of particular interest to biological spectral imaging not just because it can have higher penetration depths than visible (giving the possibility to study areas deeper below the skin in *in vivo* measurements), but also because the molecular vibrations typical to functional groups and chemical bonds in biological tissue lie within this region (see section 2.4.2). The absorption peaks from the resulting spectrum stem purely from the molecules present – unlike the visible spectrum, there is no need for the addition of markers. Raman spectroscopy operates on a different mechanism from straightforward absorption spectroscopy (although it probes the same energy states), and is most commonly detected at wavelengths in the infrared.

Near infrared (NIR) radiation can penetrate deeper than visible radiation due to reduced scattering and deeper than the mid infrared (MIR) due to reduced absorption – absorption peaks in the NIR are based on weaker molecular overtone and combination vibrations. However, the complicated overlapping of overtones can make it difficult to separate the spectrum into specific chemical components [34]. The MIR region of the spectrum is utilised far more than the FIR for spectroscopy purposes, as there is a greater presence of features from molecular bonds and functional groups in the MIR.

2.1.2.1.1 Mid infrared (MIR)

MIR spectral imaging is often known as chemical imaging, due to the presence of many fundamental vibrational modes of functional groups, though the name ‘chemical imaging’ is not restricted to only the MIR region – other IR bands and Raman spectroscopy can also reveal the presence of certain chemicals. MIR spectroscopy typically employs FTIR interferometric spectrometers instead of dispersive spectrometers, or tuneable sources. Water and CO_2 have strong absorption bands in the MIR, meaning at wavelengths where there is strong absorption (4.2µm, ~13.5-16.5µm for CO2, across much of the MIR for water), the data can be unreliable if the atmospheric content changes over the timescale of the data acquisition. Often the space in which the measurement is taking place is purged with nitrogen to allow spectra to be taken without atmospheric gases interfering.

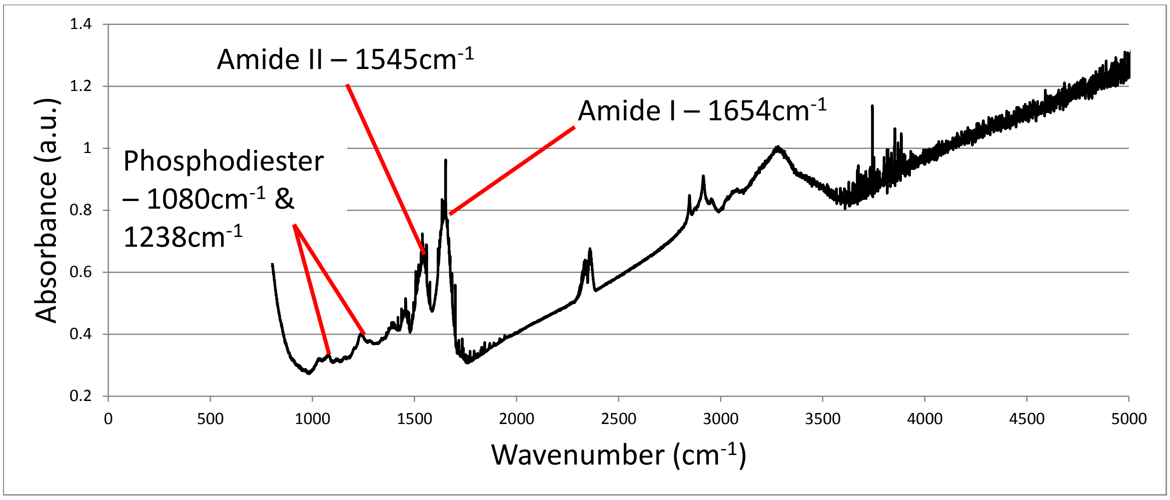


Figure 2-2. Typical absorption spectrum of tissue biopsy obtained with an FTIR. The large peaks at 1654cm^{-1} and 1545cm^{-1} are the amide I and amide II peaks of proteins respectively. Two further peaks at 1080cm^{-1} and 1238cm^{-1} are attributed to phosphodiester bonds from nucleic acid. This spectrum has not been baseline-corrected for scattering at higher wavenumbers. This spectrum also shows evidence of detector noise at higher wavenumbers, and contributions from absorption of atmospheric gas such as CO_2.

Many vibrational modes from biological molecules fall in the MIR, such as proteins, nucleic acid and lipids, making MIR spectroscopy a powerful technique for determining tissue composition without the need for markers or dyes. An absorption spectrum of a tissue biopsy is shown in Figure 2-2, highlighting some key peaks that can be attributed to proteins and nucleic acid.

2.1.2.1.2 Raman spectroscopy

Raman spectroscopy relies on inelastic scattering (Raman scattering), and can be observed through the FIR all the way to the near-UV. Light is much less likely to undergo Raman scattering than Rayleigh scattering, with only ~1 in every 10^8 photons Raman scattering [35]. The Raman-scattered light is then sent to some form of spectrometer for spectral analysis.

In the Raman process, monochromatic light interacts with the electron cloud and bonds of the molecule, exciting the molecule to a virtual energy state [36]. The molecule then returns to either the same state (Rayleigh) or a different state (Raman) [36]. In the case of Raman, the emitted photon will have a different energy from the incident light. If the photon is emitted with more energy, it is said to have undergone a Stokes shift, while a photon with less energy has an anti-Stokes shift [36]. A diagram demonstrating the Raman process is shown in Figure 2-3.

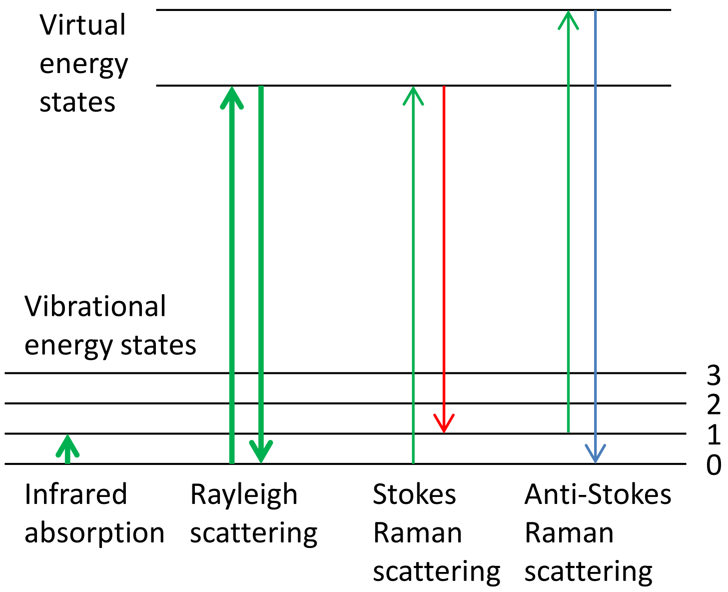


Figure 2-3. Process of Raman (inelastic) scattering. In normal absorption the molecule is simply excited to a higher energy state, absorbing the light. In Rayleigh (elastic) scattering, the molecule is excited to a virtual energy state, but then returns to the same state. In Raman scattering, the molecule returns to a different energy state. Raman scattering occurs much less frequently than Rayleigh scattering.

Raman spectroscopy, operates on a different physical process from IR absorption spectroscopy (change in polarizability instead of change in dipole moment) and so can probe different molecular vibrational modes [36]. The advantages of Raman spectroscopy include the fact that the spectrum is produced by monochromatic light (such as laser light), negating the need for tuneable sources, and the excitation can be done using sources that emit in the visible. However, the Raman effect is very weak, and the need for high power sources can lead to heating and even sample damage [35]. Molecular vibrations that are Raman and IR active will have peak positions in a matching position, the shape of the spectrum will often be very different [36].

2.1.3 Baseline correction

There are two main sources of baseline problem that could arise in the systems used in this thesis. The first is additional absorption from substrates. There will be reflection at each surface of the substrate, as well as absorption by the substrate. If the refractive index of the substrate and the absorption of the substrate are spectrally flat, this will add a constant absorption at all wavelengths, and could be relatively straightforward to ratio out. This constant offset can also be completely removed simply by taking a background with the substrate only.

The second source is from scattering. The two main sources of scattering (assuming an optically flat sample) are Rayleigh and Mie scattering (see section 2.3.1.2) [37]. Rayleigh scattering is strongly wavelength dependent (with λ^4 dependence), while Mie scattering is more wavelength independent [37]. The actual wavelength dependence can be quite complex in biological specimens such as tissue, as there will be a wide range of scattering centre sizes. In these more complex cases, it may be easier to correct for the baseline through practical methods (such as ratioing against a non-absorbing reference sample) rather than fitting a polynomial function [1]. Another alternative is a processing step such extended multiplicative scatter correction, which corrects a spectrum by using a similar reference spectrum to correct for scattering effects and other baseline effects [38].

2.2 Spectroscopy tools

When taking an absorption spectrum, a source capable of emitting across a range of wavelengths is required. The source can be scanned across its range to build up a spectrum, or the FTIR method can be used.

2.2.1 Spectroscopy configurations

When acquiring an absorption spectrum, there are a number of different setups that we can implement depending on the sample to be probed. For example, if the sample is small and thin, it will be easy to place it in an optical pathway and do a transmission measurement. If the sample is large and thick (thick enough that to attempt a transmission measurement would attenuate the light completely) a reflection measurement is most suitable. There are variations on reflection based on what information is actually being collected and the setup of the measurement. Transflection can be considered as a third configuration, effectively a double-pass transmission setup, in which the light travels through the sample, reflects off a metal-coated substrate and travels back through the sample. However, transflection has inherent issues that result in the data from such a setup being untrustworthy due to the electric-field standing wave at the reflective interface giving rise to a spectrum that is nonlinear across wavelengths [39-41].

2.2.1.1 Transmission

The transmission (or transmissivity) of a sample is the ratio between the intensity of light that enters the sample and the light that is detected on the other side, unabsorbed and not reflected. It is described by equation 2-2 below

$$T = \frac{I}{I_o} \tag{2-2}$$

where I is the intensity of light exiting the sample and I_0 is the intensity of the light entering the sample. T can be expressed as either a decimal number between 0 and 1, or more commonly as a percentage. The transmission configuration is shown in Figure 2-4, along with other loss mechanisms apart for absorption that may be present.

The resulting transmission is often expressed as absorbance, as it can the data more straightforward to manipulate, and it is more common to use it in the fields of chemistry and biology. The absorbance is given by equation 2-3 below

$$A = -\log_{10}\left(\frac{I}{I_o}\right) \tag{2-3}$$

The absorbance should not be confused with the absorptance (sometimes referred to as absorptivity), which is the ratio of radiation absorbed to the incident radiation.

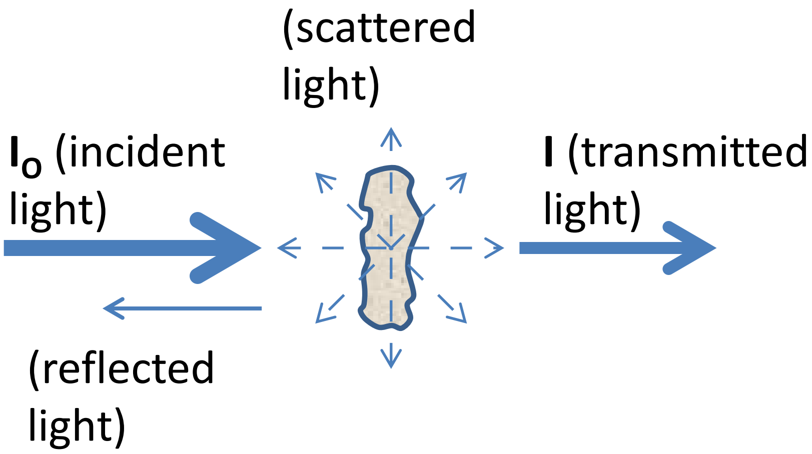


Figure 2-4. Paths that light will take as it travels through an absorbing sample. Some proportion of the light will be reflected, from the first surface of the sample and possibly from layers within the sample. Some of the light that enters the sample may be scattered, in all directions, but not necessarily equally in all directions. Light will be absorbed within the sample, resulting in the light that exits the sample being less intense than that which entered.

These equations lead to the Beer-Lambert Law, which states that the absorbance is linearly proportional to the concentration on absorbers. The fraction of light absorbed by each layer of equal thickness is the same. There are certain select situations where the Beer-Lambert law does not hold, and the absorbance is nonlinear in relation to the concentration. These situations include the light having a linewidth wider than the absorption, non-parallel rays in the sample (such that the light experiences different path lengths), extremely intense light that saturates the absorption mechanism (similar to the Moss-Burstein effect, see section 3.3.3.2.1) and fluorescence/phosphorescence of the sample.

Although the transmission and absorbance are not taking scattering and reflection into account, it is possible to account for them if the final result sought is not a value for absorbance/transmission. As with equation 3-3 (see section 3.1.3), we can subtract the absorbance at one reference wavelength from the wavelength of interest. If the scattering and reflection are equal at the two wavelengths, after subtraction only the difference in absorbance of the sample will be left. The scattering and reflection are unlikely to be constant across wavelengths, but between closely spaced wavelengths, the difference may be small enough to be negligible. This is expressed in equation 2-4, where δ is the additional value in the absorbance due to the scattering or reflection, denoted with superscripts s and r respectively.

$$\Delta A = [A_1 + \delta_1^s + \delta_1^r] - [A_2 + \delta_2^s + \delta_2^r] \approx A_1 - A_2 \qquad \text{2-4}$$

2.2.1.2 Reflection

The reflection (or reflectivity) of a sample is the fraction of light reflected at an interface of the sample. The transmission, T, the reflection, R and the absorptance, $\mathcal{A}$ are related such that

$$T + R + \mathcal{A} = 1 \qquad \text{2-5}$$

Reflection can be split into two types – specular or diffuse reflection (as shown in Figure 2-5), though most real objects have a mixture of both.

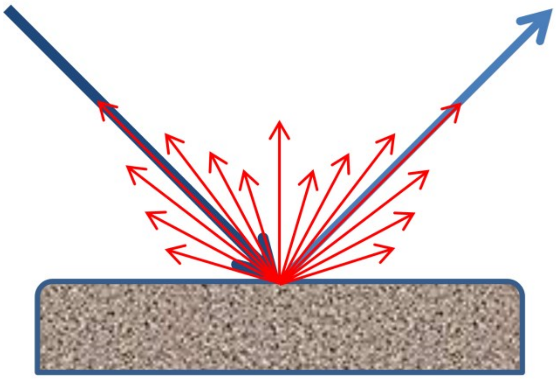


Figure 2-5. Two types of reflection – specular (light blue arrow) and diffuse reflection (red arrows). Specular reflection occurs at the topmost surface that the light is incident on, and is reflected at an angle equal to that which the incident light creates with the surface normal. Diffuse reflection originates from below the topmost surface and occurs in many directions. With most real (non-Lambertian) surfaces, there may still be a 'preferred' direction, not necessarily the same as shown in the figure. Most surfaces will exhibit a combination of diffuse and specular reflection.

2.2.1.2.1 Specular reflection

Specular reflection is the reflection you see from polished metal, or a mirror. All the light is reflected at the same angle it was incident on the surface at. The fraction of light reflected in dielectrics is determined by the Fresnel equations, which (for unpolarised light at normal incidence) is given as

$$R = \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 \qquad \text{2-6}$$

where n_2 is the refractive index of the medium at which reflection takes place and n_1 is the index of the incident medium. Almost all materials can be polished to remove irregularities of similar size with the wavelength of light in order to give specular reflection.

2.2.1.2.2 Diffuse reflection

Diffuse reflection is the reflection you see from paper in the visible or any other matte surface. Light incident from a single direction is diffusely reflected when the light is reflected at many angles. A surface that appears to reflect light with the same brightness at all angles is known as a Lambertian surface. Such a surface follows Lambert's cosine rule, which says that the power per solid angle observed (from an ideal diffusely reflecting surface) is proportional to the cosine of the angle between the line of observation and the surface normal [42]. The surface appears to have the same brightness since, as the power is reduced as the observer's angle moves off-normal, the solid angle from the observed angle is reduced proportionally.

Diffuse reflection is not largely due to surface roughness – a flat polished surface will have a component of specular reflection, but this does not exclude the existence of diffuse reflection. This is thanks to most of the diffusely reflected light occurring due to scatterers beneath the surface [43]. The resulting diffuse reflection is a sum of the random walks of partially transmitted and reflected rays at each particulate interface – the diffuse reflection is seen when these rays eventually exit the material in random directions.

A non-Lambertian surface (the more common case) will have an angular distribution of reflections that is difficult to predict, though some propose that there will be three 'lobes' from diffuse reflection of normal materials – a forescatter lobe, in the direction of the specular reflection, a backscatter lobe in the direction of the incident light and a normal lobe spread around the surface normal [44].

2.2.2 Spectroscopy methods

There are two basic elements required when doing a spectroscopic measurement – a light source and a detector. In this section, the basics of various types of light source that may be used in the MIR are discussed, following a quick introduction to parameters and metrics of interest.

The radiance is a measure of the amount of radiation that is emitted from a surface and falls within a given solid angle in a specified direction. It has units $W.sr^{-1}.m^{-2}$, whilst the spectral radiance, the radiance in one span of wavelength, as a function of wavelength has units $W.sr^{-1}.m^{-3}$.

Radiance, L, is defined as

$$L = \frac{d^2\Phi}{dAd\Omega\cos\theta} \approx \frac{\Phi}{\Omega A\cos\theta} \qquad 2\text{-}7$$

where Φ is the total radiant flux, or power, emitted, θ is the angle between the surface normal and the specified direction, A is the area of the surface and Ω is the solid angle created by the measurement.

For some laser and laser-like sources that give a collimated output, the radiance can be very high, as the solid angle is very small. When calculating the radiance at a detector, A is the area of the detector and Ω is the solid angle subtended by the detector.

Radiance is sometimes known as 'brightness', especially in the field of laser physics.

2.2.2.1 FTIR spectroscopy

Fourier transform infrared spectroscopy (FTIR) is a technique that can be used to collect an absorption spectrum without having to scan through a single wavelength at a time. Thanks to an FTIR's ability to collect a spectrum over a wide spectral region in a short amount of time, dispersive spectrometers which have to collect a spectrum point by point are rarely used to collect absorption spectra in the infrared.

The FTIR technique relies on, and takes its name from, the Fourier transform's ability to convert data between reciprocal spaces (such as between time and frequency). An FTIR spectrometer is constructed like a Michelson interferometer (an instrument for measuring temporal coherence), with the mirror on one arm being moved by a motor, and interference occurring at the detector [1]. The whole spectrum is not taken quite simultaneously – for each position of the mirror, a different combination of wavelengths will be experiencing constructive/destructive interference [1]. Software then converts this data (an 'interferogram') into a spectrum. The interferogram is modulated by absorption from the sample when it is present [1]. A figure demonstrating the method of operation is shown in Figure 2-6.

The advantages of FTIR over dispersive spectrometers include the speed at which a wide spectrum can be acquired. A higher SNR can be acquired in the same amount of time as a dispersive spectrometer takes – this is known as Fellgett's advantage, which states that the SNR will be improved in multiplexed measurements over direct measurements when detector noise is the dominant noise source [1]. This is the main reason Fourier transform spectroscopy has not replaced dispersive spectroscopy in the UV and visible wavelengths, where the noise is dominated by photon shot noise (see section 2.5.1.1.1). Another key advantage is throughput – a monochromator requires the slit to be narrowed to improve spectral resolution, eventually attenuating the signal. Using FTIR however, spectral resolution can be improved without reducing apertures [1].

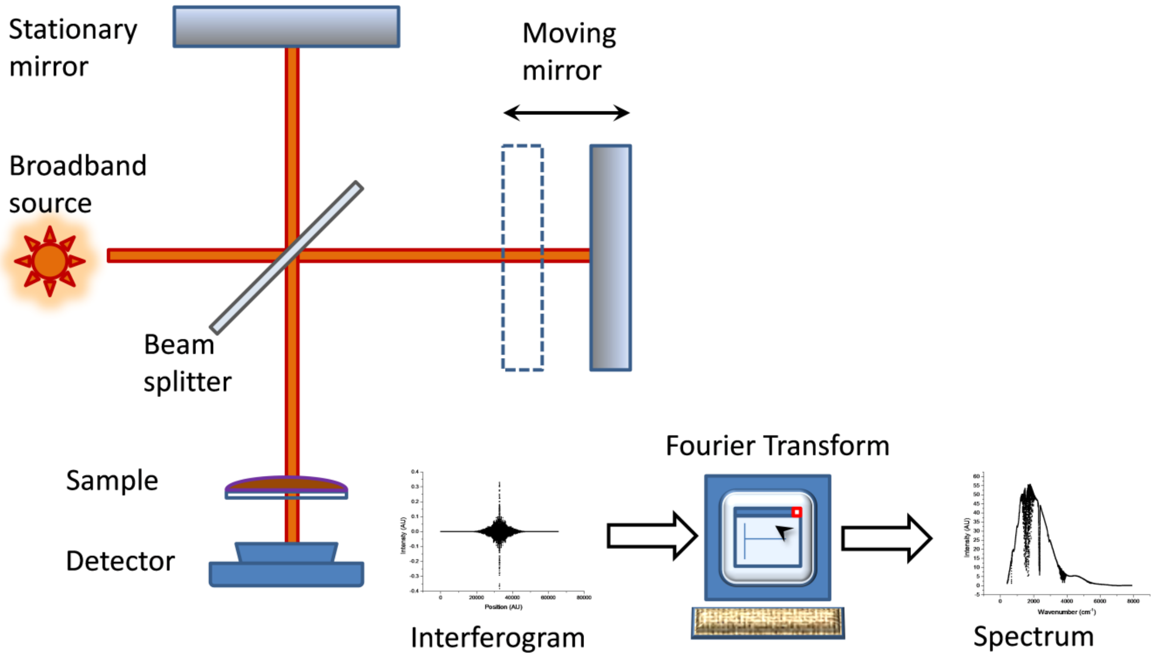


Figure 2-6. Operating mechanism of an FTIR spectrometer. The data collection process acts much like a Michelson interferometer, with two arms, one of which has a moving mirror. A broadband thermal source is used and interference occurs at the detector after the light has passed through the sample. For each position of the moving mirror, a different combination of wavelengths experience constructive/destructive interference at the detector. The output from the detector is intensity as a function of time, called an interferogram. Software then performs a Fourier transform on the data to plot intensity against frequency.

2.2.2.2 Thermal source & bandpass filters

All real-world objects emit radiation due to the thermal motion of charged particles within [45]. The interatomic collisions change the kinetic energy of atoms and molecules, resulting in charge acceleration and dipole oscillation, producing electromagnetic radiation [45]. The radiation spectrum corresponds to the spectrum of energies and accelerations present [45].

A black body is an ideal absorber that can absorb all radiation, at all wavelengths and angles of incidence [45]. At thermal equilibrium, a back body is also an ideal emitter – it has emissivity $\varepsilon = 1$ and emits radiation isotropically [45]. Real materials will have an emissivity of $\varepsilon < 1$, radiating at a fraction of the amount of black bodies.

The spectrum of wavelengths emitted from a thermal source at a particular temperature is dictated by Planck's law, resulting in an emission spectrum that changes with the temperature of the emitter [45]. The emission spectra from black bodies at various temperatures can be seen in Figure 2-7. Each increase in temperature increases the emission at every wavelength, and shifts the peak emission wavelength to shorter wavelengths.

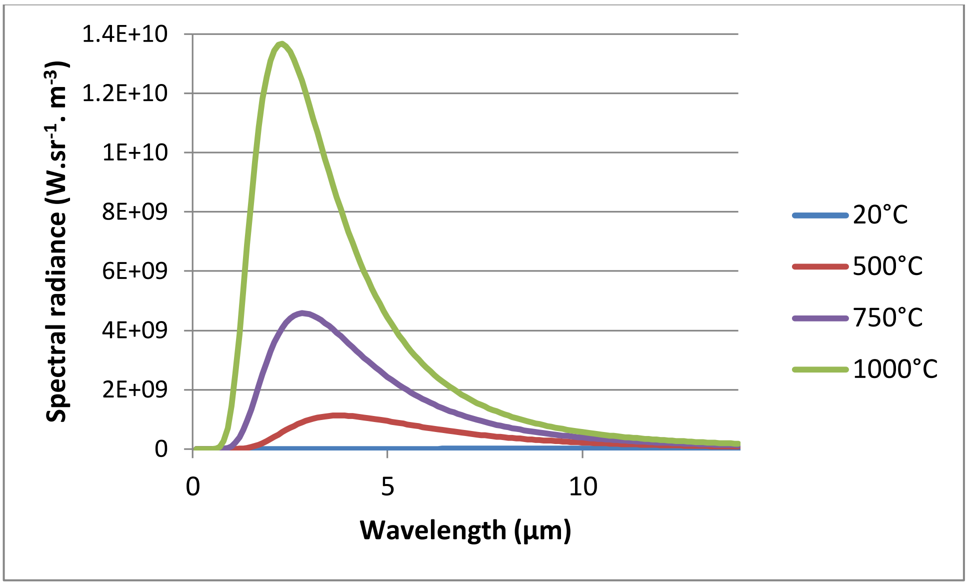


Figure 2-7. Black body emission following Planck's law. Each increase in temperature increases the emission at every wavelength, and shifts the peak emission wavelength to shorter wavelengths. The peak emission wavelengths of 20°C, 500°C, 750°C and 1000°C are 9.88µm, 3.75µm, 2.83µm and 2.28µm respectively. The emission at 3µm is over a magnitude greater at 1000°C than at 500°C, but less than three times greater at 8µm.

With a broadband source, it is possible to build a spectrum over time by using a dispersive spectrometer to select only a very narrow band of wavelengths at a time. The disadvantages of this method compared to FTIR have been touched on in section 2.2.2.1. However, if a whole spectrum is not required, and only a few select wavelengths are required, the FTIR method collects a lot of unnecessary data. In this case, the SNR possible in a given acquisition time may be better using a dispersive spectrometer.

Instead of a dispersive element selecting the wavelength, it is possible to use narrow bandpass filters to select the wavelength. This has the advantage that the spatial intensity distribution of the signal has no chance of being modified for the use of imaging, and intensity is not lost in higher order diffraction modes, in the case of a grating element. However, a grating element allows the selected wavelength to be changed by changing the angle, being able to select any arbitrary wavelength in its operating range – to change the wavelength with bandpass filters, a new filter will have to be used, adding cost. As long as the wavelengths to be investigated are predetermined and known, the cost will be fixed. The wavelengths required for the purposes of he work in this thesis are determined by previous work that has identified the spectral peaks of most interest [27, 46].

2.2.2.3 Tuneable source

A tuneable laser-based source can output near-monochromatic light, but can vary that output across a range of wavelengths. The laser-based sources may not have the same overall output power, but will have a higher optical power spectral density as all the power is concentrated at one wavelength. Thus they should give a higher SNR than thermal sources that are scanned through their wavelengths. In context of spectral imaging, the spatial coherence of these laser-based sources can add complications (see section 2.3.1.1), but their collimated output enables diffraction-limited imaging while using all the available power. The output from such sources is often pulsed. If the pulse is sufficiently short, events on the scale of that pulse may be captured, leading to a very high temporal resolution.

There are white-light laser sources available in the visible and near-IR regions of the spectrum (supercontinuum generation due to nonlinear processes like four-wave mixing in photonic crystal fibre), but laser-based sources in the MIR have monochromatic outputs. Therefore any spectroscopy work has to scan the output wavelength across a range to build up a spectrum.

Laser-based sources have the advantage that they are diffraction-limited sources and have high spectral radiance, which can lead to higher SNR data than thermal sources. One key disadvantage they have when it comes to imaging applications however is the spatial coherence, which can have a detrimental effect on image quality. Steps must be taken to reduce the coherence if laser sources are to be used for imaging (see section 2.3.1.1).

In previous decades, a system with a nonlinear crystal was required to have a tuneable laser source in the MIR. In the past few years, tuneable quantum cascade lasers (QCLs) have become commercially available [47].

Quantum cascade lasers (QCLs) are semiconductor lasers that utilise intersubband transitions such that a single electron can generate multiple photons [48]. They can generate light in the MIR and Far-IR, and have been demonstrated operating in the range from 2.6µm to 250µm [49, 50]. A QCL has a periodic series of nanometre-scale semiconductor layers of varying composition, forming a quantum well superlattice that creates a staircase-like band structure when under suitable electrical potential [48]. The wavelengths emitted are largely determined by the quantum well layer thicknesses as opposed to the material the layers are composed of [48]. After an electron makes an intersubband transition, it can tunnel into the next period of the structure to emit again, thus cascading down the staircase [48].

QCLs are capable of emitting over a range of wavelengths, and are most commonly used as tuneable devices when operating as a laser. To fix the wavelength that the QCL lases at, a very short cavity can be made by cleaving the ends of the device [48]. Including a distributed Bragg reflector can further suppress unwanted modes [51]. In an external cavity configuration, the cleaved facets of the device will have an anti-reflection coating, leaving the device to act as a classical gain medium [52]. A possible way of tuning the output can be to include a frequency-selective element such as a diffraction grating [52].

As QCLs typically produce less intense pulses than an OPO system, they are more suitable for applications in which unwanted nonlinearities may manifest (see section 3.3.3.2) [53, 54]. However, due to the coherent nature of their output, when imaging, the coherence of the light will still have to be dealt with.

2.2.2.4 Relative strengths and weaknesses

Thermal sources are convenient for imaging applications, as there are no coherence considerations to worry about. Additionally, as they operate in a continuous-wave (CW) mode, nonlinearities are much less of a concern. Being a broadband source, they are ideal for use in an FTIR spectrometer, with only a synchrotron source being the other viable option in the MIR. They also tend to be relatively cheap and small-scale. The disadvantage they have when compared to other sources is their spectral radiance. Other sources tend to have higher power per wavelength, and can produce collimated outputs, appropriate for diffraction-limited imaging. When using a non-FTIR setup, much of the radiant power is discarded as the output wavelength is selected. The linewidth of the output in this case would also typically be wider than is afforded by laser sources.

Laser-based sources are portable, with high spectral radiance and diffraction-limited outputs, though considerably more expensive than thermal sources. They will only output one discrete wavelength at a time, so to acquire a spectrum, the wavelength must be step scanned. Using it with an FTIR setup is not possible. The two main drawbacks with laser-based sources are the coherent light that cause problems for imaging applications and the high intensity pulses produced that may cause nonlinearities. For details on how these problems are resolved, see sections 3.3.3.1 and 3.3.3.2 respectively.

Not discussed in detail is a synchrotron source, due to the fact that they require large dedicated facilities to operate – it would be impossible to create a device that may one day see widespread use. In the IR, the synchrotron light is broadband, so can easily be incorporated into FTIR spectrometers as a much more intense and diffraction limited (i.e. a higher spectral radiance) source over normal thermal sources, giving larger SNRs.

2.2.2.4.1 Spectral Resolution

The spectral resolution determines the smallest spectral feature that could potentially be seen in a spectrum.

In an FTIR spectrometer, this is generally determined by speed at which the mirror travels – the slower it moves, in smaller steps, the more data points will be gathered as it sweeps across the range, so when the interferogram is Fourier transformed, there will be a smaller spectral distance between data points [55]. A common practice is to 'zero-pad' or 'zero-fill' the interferogram before transforming to artificially increase the number of data points [1]. This is akin to interpolating between the 'real' data points, so if there were a spectral feature that was thin enough to lay in between the 'real' data points, it would not be revealed by zero-padding.

The best spectral resolution attainable by an FITR spectrometer is also determined by the maximum path length difference between the two arms [55]. If using flat mirrors, increasing this distance can lead to alignment problems due to walk off, but these can be corrected for by using corner-cube mirrors (which always reflect light in the incident direction) [1].

For a thermal source with filters, the spectral resolution is determined by how narrow the pass band is. The spectral resolution is typically much larger than an FITR or laser-based system.

For laser-based systems, the spectral resolution is determined by the linewidth of the source and when building a spectrum, the smallest change in output wavelength a tuneable source can give.

When looking at biological samples, the absorption peaks tend to be much broader than atomic-like transitions that are usually observed in solid state structures such as quantum wells, so having a good spectral resolution is less of a critical issue. However, when creating a continuous spectrum, the spectral resolution of the instrument should still be less than half of the narrowest feature that wants to be measured, due to the Nyquist limit. For a spectrometer that uses a discrete wavelength source, the spectral resolution of a continuous spectrum will be a convolution of the spectral step size and the source linewidth [56].

2.2.3 IR detectors

To cover the 6-10µm spectral region, there are only a few choices of detectors open to us. Various detectors, with their detectivity and effective operating wavelengths are shown in Figure 2-8. There are only 3 types of detector available for our desired range – the bolometric 'SiB' (liquid helium cooled silicon bolometer), the pyrolelectric 'DLaTGS' (deuterated L-alanine doped triglycine sulphate) and the MCT (liquid nitrogen cooled mercury cadmium telluride) photodetector. Figure 2-8 does not quite tell the whole story of options – for example, the camera used in the *digistain* work in this thesis is a silicon bolometric camera, but operates uncooled. This will decrease its detectivity compared to the cooled detector plotted in Figure 2-8. Most detectors will work better when they are cooled due to a reduction in noise, but cooling is only optional with some detectors, whilst essential in others.

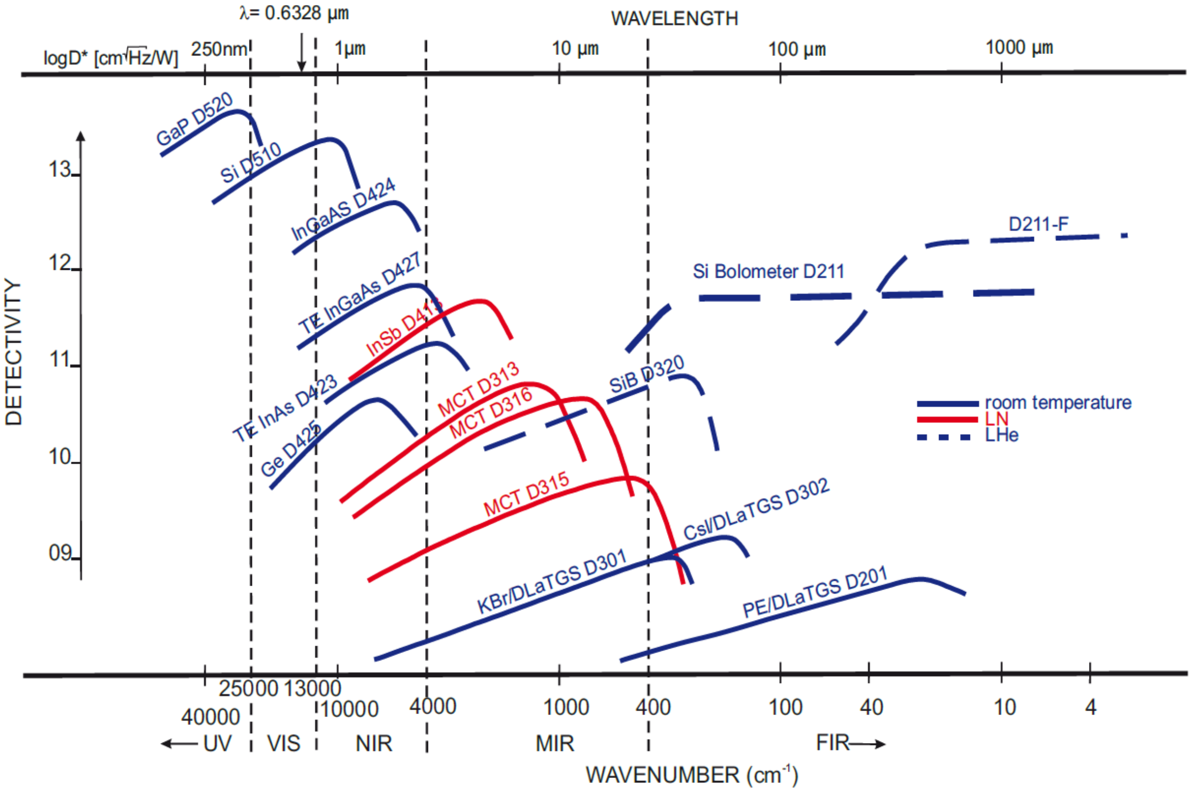


Figure 2-8. Detectivity of detectors across the UV to the FIR. To cover the 6-10µm range, there are 3 types of detector that are suitable, some of which need a cooling mechanism to reduce noise. Each of these 3 detectors has a very different operating principle – the bolometric SiB, the pyroelectric DLaTGS and the photodetector MCT. The detectivity of detectors in the MIR is much lower than those in the NIR and visible regions of the spectrum. Image taken from reference [57].

Bolometric and photodetectors are currently the most used detectors for commercially-produced cameras.

There are a number of parameters that are available for consideration when it comes to choosing a detector. Depending on the application, a certain parameter may be of greater importance than others. For most, the detectivity will be of most importance, as it takes into account the efficiency and intrinsic noise of the detector. However, in the case of trying to obtain an oscilloscope trace of a short laser pulse, response time is most important.

There are other parameters that must be considered, such as the operating spectral range, though parameters like detectivity are often plotted as a function of wavelength. In the case of a camera, the pixel pitch could also be important for concerns such as the spatial resolution (see section 2.3.1.3.3).

To the work in this thesis, the dark current is of greatest practical importance, as it influences the settings used during image capture, and contribute to the need to take 'background' images (see section 3.1.1). When a reverse bias is applied to a photodetector, there is a small 'dark current' that is set up. The dark current is present even when there is no incident light, and is due to the random generation of electron-hole pairs, which are then carried away as charge carriers before they have time to recombine. Dark current can be partially accounted for by taking a suitable background, though there will still be a time-varying noise signal due to the shot noise (see section 2.5.1.1.1).

2.2.3.1 Thermal detectors

Thermal detectors generally have a flat spectral response, which can make them useful for calibrating optical systems, but tend to be slower and noisier [58]. Although a thermal detector will respond to radiation at all wavelengths, significantly increasing the background level, a suitable filter can be used to restrict detection to a narrower band of wavelengths. Thermal detectors fall into three types – pyroelectric, thermoelectric and bolometric detectors [59].

Pyroelectric are usually made of triglycine sulphate (TGS) or deuterated TGS (DTGS), and generate a temporary voltage when they undergo a temperature change [58]. This makes them most suitable for detecting rapid changes in temperature [60].

Thermoelectric detectors also generate a voltage, but the voltage is permanent as long as light is incident. One type of thermoelectric detector, a thermocouple, one part of the detector is kept at a constant temperature and the other is exposed to a different temperature. If the thermocouple is made of two different conductors with different properties, a net voltage is created [59].

Bolometric detectors work through the use of a material with a temperature-dependent electrical resistance. As the material is heated, the resistance decreases. In the case of imaging, a camera is usually made up of microbolometers, with each microbolometer being a pixel [59].

A typical design for a microbolometer is as follows, and is represented in Figure 2-9. A material that absorbs IR light sits on the top surface of the pixel, exposed to incoming light. Below is a gap to thermally isolate the readout circuitry from the absorbing layer. Below that is a reflector to direct any unabsorbed light back into the absorber. It is the resistance change of the IR absorber which is then measured and processed to give an output.

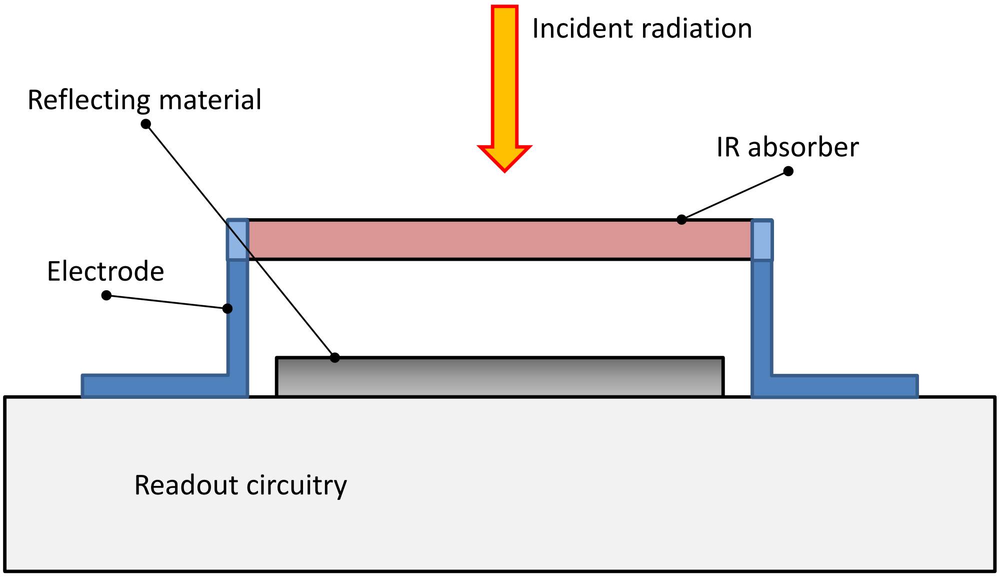


Figure 2-9. Typical microbolometer design, comprising a thermally isolated IR absorbing element that exhibits large changes in resistance with temperature change that sits above readout circuitry, connected to electrodes. Below, a mirror reflects IR light that was not absorbed as it passed through the element back up to increase the fraction of light absorbed. The resistance change in the element is measured to give an output.

Bolometric detectors do not require cooling, reducing their overall cost. However, a non-cooled detector will be less sensitive than a cooled one. Additionally, the speed with which the detector can respond to incoming light is limited by the thermal heat capacity and the thermal conductance of the IR absorber, meaning they often cannot operate as quickly as photodetectors [58]. The material used as the IR absorber must show large changes in resistance with temperature. The two most commonly used materials in microbolometers are amorphous silicon (as it can easily be integrated into the fabrication process) and vanadium oxide, which has a high thermal resistance coefficient [58].

2.2.3.2 Photodetectors

Photodetectors are detectors that respond only to an incoming photon (i.e. they only respond to electromagnetic radiation) and only operate within a band of wavelengths. A photodetector is made of semiconductor, as a p-n junction or PIN structure – when a photon with sufficient energy is absorbed, it excites a charge carrier across the band gap.

Although photodetectors can be operated at room temperature, they work best when cooled to reduce the noise contribution from thermally excited carriers, and in practice are always operated like this for use in the MIR [61]. Most commonly, the detectors are cooled with liquid nitrogen to 77K, though they can also be cooled by other means such as through the use of a Stirling engine [61].

Two common semiconductor materials used for photodetectors in the MIR are indium antimonide and mercury cadmium telluride [57]. Indium antimonide (InSb) is a material used for photodiodes from 1~6.7μm [57]. Mercury Cadmium Telluride (HgCdTe), also known as MCT, can be used across the whole MIR part of the spectrum, working from 1-25μm [57].

Unlike thermal detectors, photodetectors have a range of wavelengths in which they can operate. For this reason, the detector material has to be chosen to match the signal. A photodetector will respond weakly or not at all to photons with an energy below the bandgap energy, which may make them less sensitive to background thermal noise than thermal detectors without a filter [62].

2.3 Imaging in the MIR

Conducting experiments in the MIR can be quite different from working in the visible. The rules of optics remain the same, but the sources, components and detectors that can be used are usually different, even in the physics behind their operation. Sources, detectors and optics tend to be underdeveloped in the MIR compared to the visible, and so the cost of components is much greater. Due to the high absorption of water (see section 2.4.2.4), wet samples are very hard to look at.

As everything is glowing in the MIR, special care has to be taken to properly account for background radiation. As people move about in the vicinity of a detector, the radiation they emit can shift spatially, changing the background, adding random fluctuations to the measurement.

Optical materials commonly used in the visible part of the spectrum (BK7, SiO_2) for the manufacturing of optical components like lenses and windows usually do not transmit in the mid infrared.

Materials that are commonly used that can transmit in a wide range of the mid infrared include germanium (Ge), zinc selenide (ZnSe), barium fluoride (BaF_2), calcium fluoride (CaF_2) and AMTIR-3 (GeSbSe) [63].

Of greatest interest for imaging is germanium, due to its relatively low dispersion [63]. This lends itself to working at multiple wavelengths, as the focus stays relatively static across a wide spectral range. It transmits well between 2-16µm. When using germanium, visible light cannot be used for alignment effectively. It also has a very high refractive index for optical materials ($n \approx 4$), meaning a lens can be made with much flatter surfaces than other materials would need (which reduces spherical aberration and Petzval field curvature) [64]. This high n helps to reduce distortion that can be present when imaging. Uncoated germanium has high Fresnel losses due to its high refractive index (~36%), but with a suitable antireflection coating germanium can have transmission of 90% through a few millimetres. One problem with germanium is that its refractive index can change a lot with temperature – it has a $\frac{dn}{dT}$ = 0.000396, roughly 25 times the magnitude of the thermal response of BaF_2 [63]. This will shift the focus of a lens if it has to operate in a range of temperatures. Additionally, the transmission of Ge in the infrared can drop to nearly 0% at temperatures of 100°C and above [63].

The dispersion of the lens across the MIR is of particular importance when combining images at different wavelengths – each pixel in the image must have the same part of the sample imaged onto it if the quantitative numbers are to be trusted. When looking at the amide and phosphodiester peaks, we are looking at wavelengths of about 6 and 8µm (see section 2.4.2). The imaging system we use must be able to image at those two wavelengths without the need for refocusing to maintain pixel registration. For how different materials compare for maintaining focus across wavelengths, see Figure 2-10.

Any problems with dispersion and transmission range could be negated if reflective optics were used. A mirror reflects at the same angle regardless of wavelength, and aluminium or gold coatings have high reflectivity from the visible through to the longer wavelengths of the MIR [63]. However, the implementation of reflective optics have practical problems such as reduced optical throughput and increased aberrations [65].

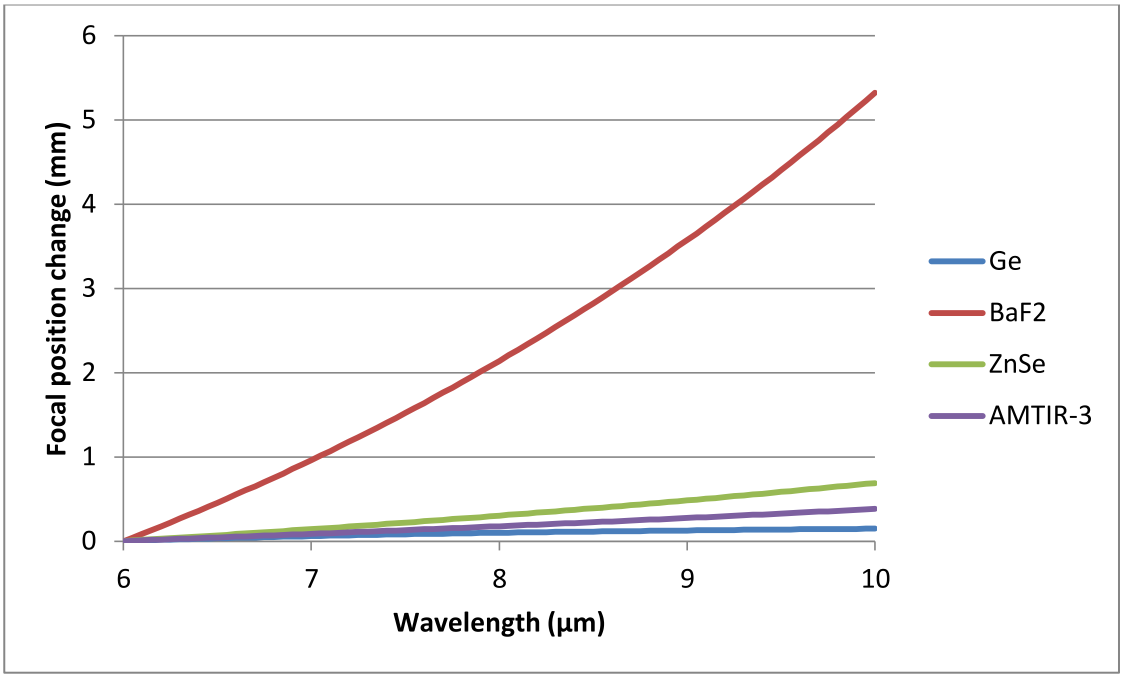


Figure 2-10. Shift in focal length for MIR materials between 6-10µm, for a plano-convex lens designed to have a 50mm focal length at 6µm. The shift is greatest for BaF_2 *and smallest for Ge. Ge has the smallest change in focal length due to it having both the lowest dispersion and the highest refractive index.*

When imaging tissue, the organelles and cellular substructures dictate what type of scattering will be observed (see section 2.3.1.2). For example, the cell nucleus size is on the order of the wavelengths being used, and so Mie scattering is the main scattering route caused by cell organelles.

In the MIR, everything warm emits radiation, which can interfere considerably with imaging measurements unless proper background procedures are taken. The presence of background radiation is usually not a problem (just requiring a subtraction), but if the background is changing with time, in a time period faster than the time taken for image acquisition, then a normal data acquisition will encounter problems. In this case, the most straightforward way to account for the variation in background is to break the entire acquisition into smaller time periods to minimise the effect (see section 3.3.2.3).

2.3.1 Optical theory

When designing an imaging system, the optics used play a crucial role in how well the system performs. Choosing the right optics allows us to get the best light throughput, resolution and image quality for the magnification that we want. This section covers the basics of lens design, properties of light such as coherence and a discussion of the diffraction limit and how it applies to array detectors.

2.3.1.1 Coherence and how to remove it

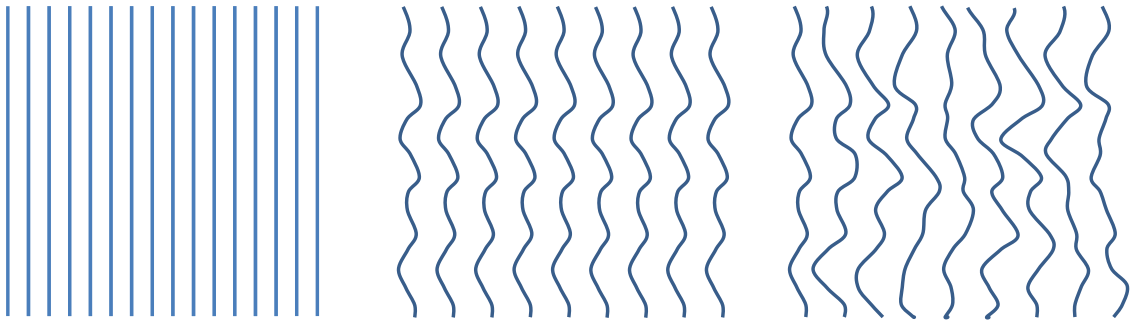
Spatial coherence describes the uniformity of wavefronts emitted from a source, transverse to the direction of propagation (see Figure 2-11). A filament bulb for instance emits at different points along its length, but there is no fixed phase relationship between any two points, meaning the wavefront emitted from the filament will change with time. Two effects of spatial coherence can be seen from the output of lasers (which tend to have both high spatial and temporal coherence) in a speckle pattern and diffraction fringes from a knife edge.

Temporal coherence is normally dependent on how monochromatic a source is – it is a measure of how a wave interferes with a copy of itself after the copy has been delayed by some time (see Figure 2-11).

A Michelson interferometer or Mach-Zender interferometer can be used to measure the temporal coherence. When imaging, a monochromatic source can lead to problems when there are closely-spaced flat optical surfaces, leading to the fabry-perot fringes overlaid on the image.

Spatial coherence is often a problem when imaging, both in a transmission and a reflection setup. Using a coherent source such as a laser for imaging in transmission produces a ‘ringing’ effect from the edges and coherent scattering (see Figure 2-12). When imaging in reflection, this coherent scattering is commonly referred to as a speckle pattern. Polychromatic light often obscures the speckle, but the effect is purely from the spatial coherence of the source, and so there are certain circumstances in which polychromatic light can still exhibit a speckle pattern [66].

A few methods have been employed to remove this coherence. The most straightforward is the use of a diffuser. A diffuser is simply a window with a rough surface profile that can change the shape of the wavefronts passing through it. Ground glass is an example of a visible-spectrum diffuser – it still transmits light, but the incoming light has been scattered as it passes through. To produce a time-averaged uniform illumination and better speckle reduction, the diffuser is moved, either through rotation [67] or lateral movement [68].



Plane wave,
infinite coherence length

Varying wavefront,
infinite coherence length

Varying wavefront,
finite coherence length

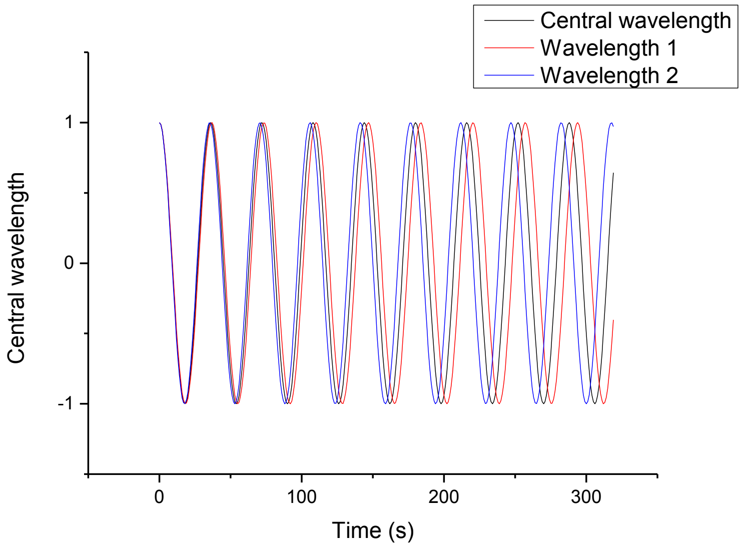


Figure 2-11. Behaviour of spatially and temporally coherent light (top), and only temporally coherent light (bottom). The top-left image shows a plane wave with infinite coherence length – the light is spatially and temporally coherent. The top-middle image shows a non-plane wave that also has infinite coherence length – the light is temporally coherent, but spatially incoherent. In the top-right image, a non-plane wave with a finite coherence wave is shown – the light is spatially and temporally incoherent. The bottom image shows the effect of a temporally incoherent light source. A wave that starts out in phase is oscillating out of phase after a few periods due to the wave consisting of different frequencies. Only a near-monochromatic source will have high temporal coherence.

A light pipe (simply a pipe with reflective inner walls) can also be used to randomise the light that has been scattered by the diffuser even further [69]. The setup used in reference [69] used a static diffuser, but rapidly moved the incident beam across the surface of the diffuser.

Other methods to remove the spatial coherence include rapidly varying the incident angle of the incoming light [70] and using diffractive optical elements [71].

Figure 2-12. Transmission image of tissue using a QCL source at 1400cm^{-1}*, demonstrating coherent scattering effects. The centre vertical line has no tissue present and should be present. It can be seen that there is a textured, grainy appearance in this area that is also overlaid across the whole image. This pattern is due to coherent scattering from the tissue due to the spatial coherence of the source. This pattern will not only decrease image quality, but the pixel values cannot be trusted for quantitative information also.*

2.3.1.1.1 Speckle

Although a speckle pattern is produced due to the spatial coherence of the light, temporally incoherent (i.e. non-monochromatic) light may remove the speckle pattern. This is because the speckle pattern is dependent on wavelength, and so patterns from individual wavelength components may be overlaid, presenting a uniform intensity distribution with low contrast. An example speckle pattern is shown in Figure 2-13.

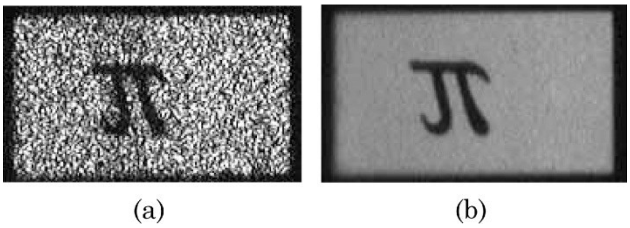


*Figure 2-13. Example of laser speckle pattern. Shown left (a), the illumination appears nonuniform on a relatively short length scale, with some illumination appearing in the dark "**π**" symbol. The image on the right (b) is when the coherence has been destroyed, with a much more even illumination and no bright points in the **π** symbol. Image from reference [69].*

Speckle can be divided into two different types of pattern – objective and subjective patterns. An objective pattern is observed when the speckle pattern is incident on another diffuse surface before being imaged. In this case, the pattern is fixed, and looks the same regardless of viewing distance, angle or aperture of imaging optics. There may actually be a secondary speckle pattern from this second surface, but typically it is the primary pattern that is observed.

The subjective pattern is the pattern we see when we look at a laser beam on a wall – we are viewing the pattern directly. In this case, the size of the speckles is dependent on the cone of acceptance of the optical imaging system (assuming a fixed wavelength). By decreasing the aperture in our imaging system, the speckle size will increase. In a 2D imaging system, the size of the speckle is on the order of the imaging system's spatial resolution. This is because in the imaging system, only two points separated by the spatial resolution length have unrelated intensities at the image.

2.3.1.2 Scattering

There are considered to be three types of scattering commonly seen in optics – Rayleigh, Mie and geometric. The scattering case depends on the size of the scattering object. There is also Raman (inelastic) scattering, which is touched upon in section 0.

Rayleigh scattering is the situation in which the wavelength of light is much larger than the scattering particles (ratio of circumference of particle to wavelength <0.1) [37]. Rayleigh scattering is caused by the electric polarisability of the particles [37]. The oscillating electric field exerts a force on the charges within a particle, causing them to oscillate at the same frequency [37]. The scattered light we see is the radiation from the particle that has become a small radiating dipole.

The intensity of the scattered light has a proportionality to the wavelength of light that goes $I \propto \frac{1}{\lambda^4}$, and the direction of the scattered light can be described by $I \propto (1 + \cos^2 \theta)$, where $\theta = 0$ is the incident direction of light [72]. Scattering at right angles is half the forward intensity, and the forward intensity is equal to the backward intensity.

For particle sizes on the order of a wavelength, Mie scattering is observed [37]. The intensity is greatest in the forward direction – the larger the particle, the more anisotropic (in the forward direction) the scattering is [37].

Geometric scattering is not the same process as in Rayleigh and Mie scattering – it is simply refraction. This type of "scattering" occurs when the objects are more than 10 times greater than the wavelength of light.

2.3.1.3 Spatial resolution

The spatial resolution of an imaging system defines the smallest object that can be imaged. With a fixed detector, simply increasing the magnification will only go so far until we reach the diffraction limit [63]. As the point at which two points fail to be resolved can be subjective, there are various definitions for what constitutes two resolved points. The most commonly used is the Rayleigh criterion, which is usually used to quote the resolution of a system [63].

The importance of having high spatial resolution is dependent on the application. If the application requires imaging of cellular structures, the spatial resolution must be as high as possible, at the diffraction limit, for the MIR [63]. For an application like the *digistain* system, the spatial resolution is not of as great importance. The *digistain* system needs to image cells at a more macro level – it is the phosphodiester-to-amide ratio (PA ratio) of the cell that needs to be determined. If the resolution was high enough that cellular structures such as the nucleus could be seen, a pixel-binning algorithm would have to be applied to reduce the spatial resolution.

2.3.1.3.1 The diffraction limit

Light in the far field obeys the diffraction limit, which states that there is a minimum spot size that light can possibly be focused to, and is related to the wavelength of light and the smallest aperture of the system [73]. A typical value for the smallest spot size that light can be focused to, is half the wavelength of light [73]. A larger aperture (or smaller wavelength) will allow for larger k-vectors to be collected by the lens, allowing for more information on the source to reach the image plane [74]. This will allow for a better reconstruction of the original point source [74]. Standards for determining the resolution of a microscope are ultimately expressions of the diffraction limit.

2.3.1.3.2 The Rayleigh criterion

The Rayleigh criterion for resolution states that two points are considered resolved when the maximum of one point spread function (PSF) overlaps with the minimum of another PSF [75]. Cases of resolved and unresolved points according to the Rayleigh criterion are shown in Figure 2-14.

With the approximation of small angles, the spatial resolution can be written as

$$\Delta l = 1.22\frac{f\lambda}{D} \qquad \text{2-8}$$

where Δl is the distance from one point to another (the spatial resolution) and f is the focal length of the lens being used, λ is the wavelength of light and D is the diameter of the lens' aperture [75]. This is for the best case scenario, when the lens is imaging an object at infinity to a distance f away from the lens. Often we image with only a single lens, in which case we replace *f* with the distance between the lens and the image [73].

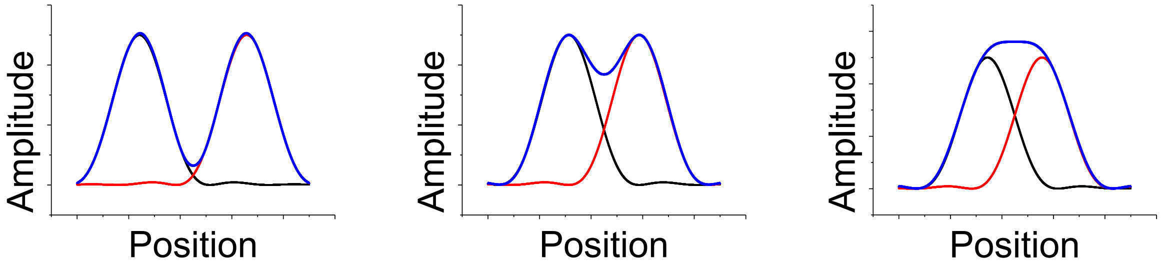


Figure 2-14. Representation of three cases (resolved, limit of resolved and unresolved from left to right) as defined by the Rayleigh criterion. The red and black curves are the intensities of the two points being imaged, the blue curves are the sum of those curves to show what will be seen by an observer. As the points get closer together, the dip between them gets smaller making it harder to see that there are two separate points. Eventually, the points get so close that any dip between them disappears completely.

2.3.1.3.3 Pixel resolution

When imaging onto a screen and observing the screen, the resolution is determined by the Rayleigh criterion as normal. When we image onto a detector, the fixed size of its pixels starts to play a role in the resolution.

If we consider a 1:1 magnification, with an f-number of 1, at 8µm, the spatial resolution will be 9.76µm. Thus two points 9.76µm apart would be imaged onto our detector, also 9.76µm apart. However, if the size of the pixels on our detector is 20µm, these two points would be imaged on the same pixel or adjacent pixels. The detector would not be able to distinguish between these two points. The spatial resolution in this case would be close to the pixel pitch, 20µm, with each point being imaged onto a pixel either side of a central pixel.

At higher magnifications, this will not be an issue, as the magnification will be large enough to image each point onto pixels that are separated by another pixel. However, as the magnification increases to even higher values, those two points will start to be imaged onto pixels that are separated by more than one other pixel. At this point, we start oversampling. Although there may not be a loss in spatial resolution, we will not be using the resolution of the detector to its full potential. There is a band of magnifications that we can use to avoid losing spatial resolution unnecessarily, and also avoid oversampling. The diffraction limit also plays a role [75]. Using the parameters of the *digistain* system (see section 3.2.1) in a 2D model, the useful band of magnifications range from roughly 3x-9x.

2.3.1.4 Optical performance

The limiting factors of the spatial resolution of an imaging system have already been covered in section 2.3.1.3. These factors are namely the aperture (usually created by the lens) and the pixel size of the detector. The theoretical spatial resolution can be calculated or predicted in a simple model, but the empirical spatial resolution of the system may differ due to unaccounted factors. The standard way to measure the spatial resolution is with the 1951 USAF resolution test chart, a pattern of lines of decreasing size. The resolution of the system will be the smallest lines that the system is able to resolve. There is scope for subjectivity here, but if we use a digital camera we can look at the pixel values and still follow the Rayleigh criterion [75].

The field of view, the size of area actually being imaged, is also an important consideration. There is a trade-off between the spatial resolution of a system and the field of view, assuming detector size and pixel pitch remains constant. Ideally, one could match the field of view to the size of the sample area under study, but if magnification is increased to improve spatial resolution, the field of view is decreased. If the field of view is smaller than the area under investigation, multiple images will have to be stitched together.

The f-number of a system is largely defined by the lens used, and is a measure of how much light can be collected by the system. The f-number is defined as

$$f/\# = \frac{f}{D} \tag{2-9}$$

where f is the focal length of the lens and D is the diameter of the aperture (usually the diameter of the lens) [76]. When imaging an object that is not at infinity, f is replaced with the distance from the image and the lens [76]. The lower the f-number, the more light a lens can collect.

2.4 Biological tissue and biological molecular vibrations

This section will present some background information on cell and tissue structure that will be useful in understanding later discussions regarding tissue biopsy samples. If the samples looked at are not dehydrated, then water content of the cells plays an important role due to the high absorption of water in the MIR.

2.4.1 Cell structure

An animal cell is an extremely complex system, however for the purposes of understanding this thesis, only the very basics need to be considered. In the simplest model, the cell can be separated into the cell membrane, the cytoplasm and the cell nucleus. An animal cell has a typical cell diameter of tens of microns, whilst the cell nucleus is roughly 6μm in size [77], roughly on the order of the diffraction limit in the MIR.

The cell nucleus is an example of an organelle, an internal structure that plays some role in the functioning of the cell. The nucleus contains the code for cellular function and replication, the DNA, and can regulate the gene expression to control cell activity. A double membrane separates the nucleus from the rest of the cell, with nuclear pores that allow the movement of large molecules like proteins and RNA. In the bulk of the nucleus (the nucleoplasm) is contained the DNA. Usually the DNA is stored in a DNA-protein complex called chromatin, and during cell division it takes the form of chromosomes [78].

The cell membrane is responsible for controlling the movement of substances in and out of the cell, and can control the size of the cell. The cell membrane is made of a lipid bilayer, the main constituent being phospholipid, a molecule that has a hydrophilic head group and a hydrophobic lipid tail. Proteins are abundant in the cell membrane, accounting for around 50% of membrane volume [79], and play a role in identification, enzymatic activity and cell-cell or cytoskeleton contact in addition to transporting molecules.

The cytoplasm's main constituent is water, but this does not need to be considered when the cell is dehydrated, as with tissue biopsies. The endomembrane system in the cytoplasm is mainly concerned with the manufacture of proteins, lipids and steroids and the transport of these within the cell or for excretion.

There is also a type of vesicle, termed 'granules', which function as secretory vesicles. An example is a melanosome, which synthesises, stores and transports melanin.

2.4.2 Molecular vibrations present in biological tissue, 6-10μm

Absorption peaks in the mid-infrared stem from molecular vibrations, such as those that can be found in biological tissue. Our light sources tend to have a common operating range of ~1000-1666cm^{-1} (~6-10μm), and there are plenty of molecular vibrations that fall within this window. A typical MIR spectrum of biological tissue in this spectral region is shown in Figure 2-15, with the main peaks of interest to this work pointed out. There are however further vibrational modes (amide A, stretching modes from lipid molecules) in the longer-wave MIR, overtones (harmonics) in the NIR and lower energy pure rotational modes in the far-IR (see section 2.1.2.1).

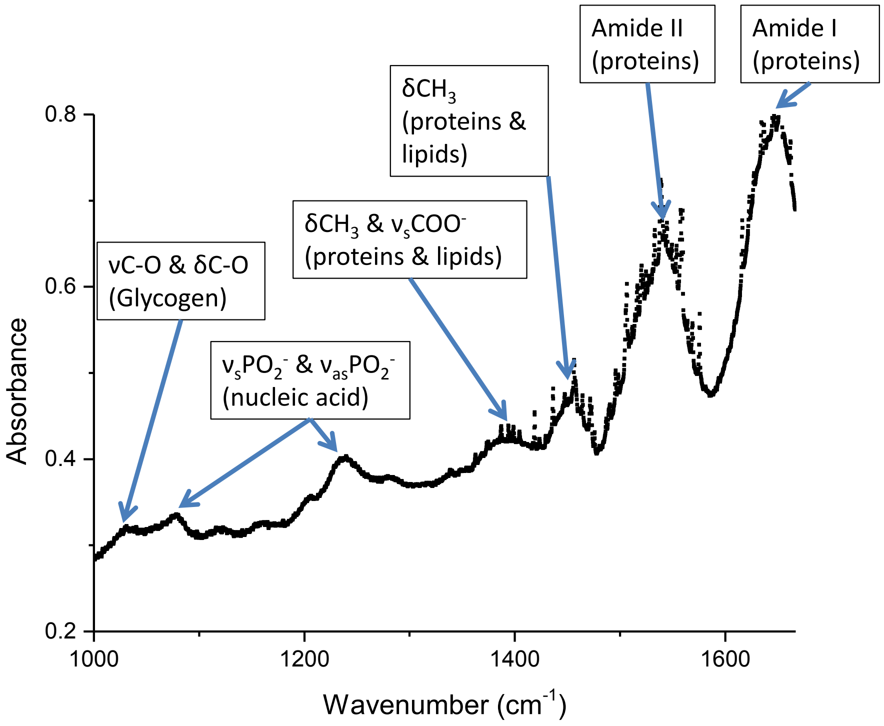


Figure 2-15. Typical spectrum of tissue between 6-10µm, with noteworthy peaks annotated. The spectrum was taken using oesophageal tissue using an FTIR spectrometer. The largest peaks are the two amide peaks that correspond to proteins in tissue. The amide I peak is at 1653cm^{-1}*, the amide II peak at 1545cm*$^{-1}$*. The other two peaks of importance to this thesis are the phosphodiester (*vPO_2^-*) peaks that correspond to nucleic acid i.e. DNA and RNA. The symmetric stretch mode peak is at 1080cm*$^{-1}$*, the larger antisymmetric stretch mode peak is at 1238cm*$^{-1}$*. Information on peak attributions from reference [46].*

2.4.2.1 Nucleic Acid (phosphodiester) absorption peaks

The cell nucleus contains the information for cell replication, DNA. DNA has a double helix structure, and contains 4 base pairs that encode the cell information. The single bases are connected on one strand of the double helix by a sugar-phosphate backbone, comprised of a pentose sugar joined by phosphate groups that form the phosphodiester bonds. The phosphodiester vibrations in the MIR that can be attributed solely to nucleic acid (DNA and RNA) are at 1080cm^{-1} and 1237cm^{-1}, corresponding to symmetric and antisymmetric stretching modes respectively [46]. There is also a phosphodiester mode at 995cm^{-1} that is solely attributed to RNA, and a phosphodiester mode at 964cm^{-1} that be attributed to nucleic acid or phosphorylated proteins (when phosphate is added to a protein – this is often a switch to control enzymes).

Although the cell nucleus will have a high concentration of nucleic acid compared with the rest of the cell, RNA is also a nucleic acid with the same absorption peaks as DNA. RNA comes in a number of forms and plays various roles, including using the information in DNA to build proteins in the cell. Due to the function of RNA, there will be some RNA present in the nucleus, and the mass ratio of RNA:DNA present will be close to 1:1 [80]. One study showed that the nucleic acid found in the nucleus (by mass) is of the same order as that found outside the nucleus in animal cells [80]. A similar study shows that roughly half of the mass of total nucleic acid is found within the nucleus in Chinese hamster ovary cells [81].

The phosphodiester bond is also present elsewhere, such as in phospholipid bilayers present in the cell (the cell membrane, liposomes and micelles). The most common phospholipid is phosphatidylcholine, which has the same phosphodiester bonds present in its infrared spectrum [82].

Despite sources outside of the nucleus that can contribute to the phosphodiester absorption peaks, the phosphodiester bond can still be used as an indication of the concentration of cell nuclei [27]. In an early *digistain* study, the absorbance at the phosphodiester peak was shown to have a near-linear correlation with the mean nuclear count, the nuclei being counted manually using higher-resolution visible images. This is an important point, as it validates the hypothesis of using the phosphodiester absorbance as a marker of nucleus size, a factor that is then used in creating the *digistain* images (see section 3.1.3).

2.4.2.2 Protein (amide) absorption peaks

Proteins are an extremely wide and varied group of molecules with a correspondingly varied group of vibrational modes that can be attributed to them. A functional group that is common to many proteins are the amide groups. Those that fall in our spectral range are the amide I peak, with the strongest contribution being from the stretching of the carbonyl (C=O) dipole, at 1653cm^{-1} and the amide II peak, attributed to N-H bending coupled to C-N stretching, at 1545cm^{-1} [46, 83]. There is also a broad amide III peak in the range that can be found from 1230-1300cm^{-1} that could overlap with a phosphodiester peak of nucleic acids [84]. The amide I and II peaks are each made up of sub-peaks that are associated with protein substructures such as α-helical segments, β-sheets sections, turns and unordered regions [85]. As a result, the amide peaks can change shape and slightly shift position depending on the main protein substructure present [85]. For example, haemoglobin is predominantly α-helical, with an amide I sub-peak centred on 1656cm^{-1}, whilst ribonuclease has high β-structure with a sub-peak centred on 1636cm^{-1} [85].

The amide peaks can give an indication of the protein concentration in tissue. Proteins are contained throughout the cell, including the cell nucleus – the presence of amide functional groups indicates the presence of biological tissue.

2.4.2.3 Lipid absorption peaks

Lipids have biological roles primarily as energy stores, for signalling and a as component of the cell membrane (the phospholipid bilayer). The lipids only have one unique peak in the 1000-1666cm^{-1} region, a bending mode of the methylene (CH_2) functional group [46]. There are two other peaks, attributed to bending modes in methyl, but these vibrational modes can also be found in proteins [46]. Lipids may indicate the presence of biological tissue, but the strength of the absorption is much weaker than the amide peaks of protein [46]. For this reason, the amide absorption peak is used in the *digistain* process instead of the lipid peak to give a better SNR and contrast in molecular concentration.

2.4.2.4 Water absorption peaks

Water absorbs heavily in the MIR, and can be detrimental to experiments. Absorption spectra can often show artefacts from the absorption in water vapour in air if the ‘background’ and ‘signal’ measurements are taken at different times, due to very slight variations in atmospheric composition. This can be accounted for by purging the system with nitrogen, which does not absorb in the MIR. Water also absorbs across the 6-10µm region (see Figure 2-16) – the transmission in this region through 100µm of liquid water will be 0.5% at most and 10^{-10}% at lowest [86].

Note that the majority of the peaks we see in the background spectra from an FTIR (see Figure 2-17) are from water *vapour*. H_2O has a small moment of inertia on rotation, giving rise to many modes. However, in liquid phase, rotations tend to be restricted, so the spectrum from vapour and liquid differ slightly. Water vapour, along with CO_2 absorption can play an important role in a spectrum if not properly accounted for (see section 2.1.2.1.1).

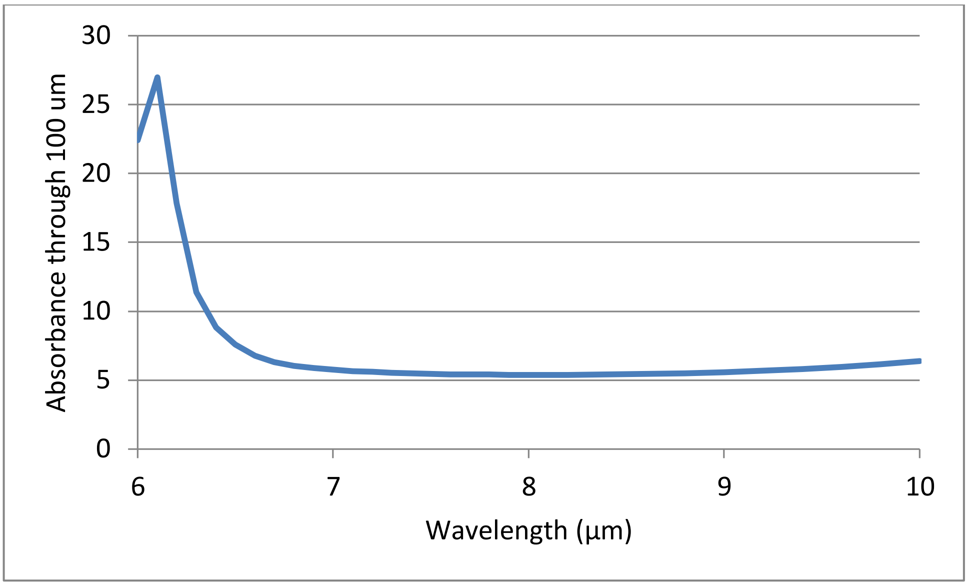


Figure 2-16. Magnitude of absorbance through 100µm of liquid water between 6-10µm. The absorbance is high throughout the entire region, but is particularly large at shorter wavelengths in this range. The maximum transmission through 100µm of liquid water in this range is 0.5%.

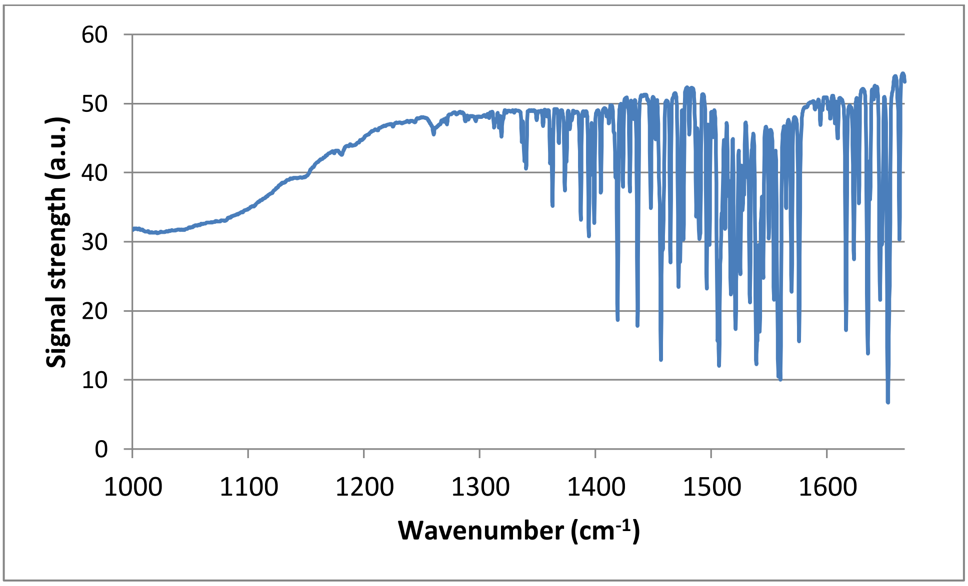


Figure 2-17. Spectrum of water vapour in atmosphere between 6-10µm. The sharp peaks from approximately 1300cm^{-1} upwards are from water vapour. The effect of these peaks can be removed from spectra by proper background subtraction. The absorption effect from water vapour can be most apparent when using sources with a very narrow linewidth – the signal power from the source will drop at wavelengths lying on one of these sharp peaks.

2.5 *Digistain* and QCL systems

The relative advantages and disadvantages of the sources, spectroscopy methods and detectors have been covered in sections 2.2.2.4 and 2.2.3. If having to acquire a detailed spectrum, the thermal source with bandpass filters and tuneable QCL source could not compete with an FTIR technique for speed. However, if only having to investigate a limited number of pre-determined wavelengths, systems that probe discrete wavelengths can have a speed advantage. Additionally, the QCL source has higher spectral radiance than a thermal source.

In particular, when it comes to imaging, the *digistain* and QCL systems have a distinct advantage. In the past, building a spectral image in the MIR involved using a scanning system, whereby a single channel detector was used, and adjacent points on a sample were measured to build up an image. This took considerable time (hours) to create an image with a large pixel resolution [26]. Raman spectroscopy still relies on scanning techniques to create an image, but FTIR systems can now be used in conjunction with a 2D detector array to acquire spectral images. However, commercial systems are sold with detectors with a pixel resolution no larger than 128x128, possibly due to issues of temporal resolution and data throughput required. The QCL and *digistain* systems have more than four times and six times as many pixels respectively in their detectors. This leads to more data being able to be collected at once, in turn reducing the data acquisition times.

The single-minded application for the *digistain* can be considered both a strength and a weakness. By focusing only on cancer diagnosis, only four wavelengths (see section 3.1.3) need to be investigated at a time which allows for acquisition times to be shorter than an FTIR system could offer. It has previously been shown that the four wavelengths used are sufficient to determine the nuclear concentration of an area of a sample [27]. If additional wavelengths are required, the speed advantage over an FTIR system will be decreased or removed. If different wavelengths need to be investigated, more bandpass filters will be required for the *digistain* system, adding to the total cost of the system. This is not an issue for the QCL system though, which can tune to any arbitrary wavelength within its tuning range.

The *digistain* system has been built as a prototype device, with cost and footprint kept in mind. The goal is that one day a *digistain* device may sit in a clinical lab and compliment or maybe even replace the current grading procedure. The QCL system on the other hand is not currently being designed with commercialisation in mind. Investigations with the QCL system are designed to assess its suitability as a quantitative tool in a transmission geometry, and the wide tuning range adds freedom to look at other samples and molecular vibrational modes beyond only cancer tissue biopsies. Additionally, the higher spectral radiance of the QCL radiation opens the possibility of doing reflection measurements. With regards to biological tissue, this could be in the context of doing remote *in vivo* measurements on living samples. More generally, reflection imaging could allow samples that are otherwise unsuitable for a transmission setup to be studied, such as relatively thick samples that could attenuate transmitted light completely.

2.5.1 Signal to noise ratio (SNR)

In order to compare the *digistain* and QCL systems with other spectroscopic devices, it is useful to have a common figure of merit. The signal to noise (SNR) is a measure used in many fields and applications that compares the size of the signal being measured to the background noise as a ratio. The SNR can be evaluated by an operator visually estimating the size of the signal, and the noise (the 'wiggle' in what should be a uniform value) on the signal. An alternative way to calculate the SNR is given in equation 2-10

$$SNR = \frac{\mu}{\sigma} \tag{2-10}$$

where μ is the mean signal value and σ is the standard deviation in the signal.

When averaging over multiple readings to get a final result, the SNR increases with the square root of the number of averages N:

$$SNR \propto \sqrt{N} \tag{2-11}$$

Thus to increase the SNR by a factor of two, the averages needed (and the time required) must increase by a factor of four.

For a spectral imaging system, the SNR could refer to the SNR of the image (affecting the 'graininess' of an image), or to the SNR of the spectrum (affecting the 'wiggle' on a spectrum of a material that should be spectrally flat). The SNR can also be used to describe the repeatability of a result, though it is more common to state this as only a multiple of the standard deviation, as an uncertainty in the measurement.

To increase the SNR, the signal can be increased and maximised. In an optical system, this would often consist of maximising throughput, using low f-number configurations to increase the amount of light collected. Maximising the signal can usually only go so far, and the SNR is limited by sources of noise in the system.

2.5.1.1 Sources of noise

Noise in an infrared imaging system can come from a variety of sources, but can generally be divided into three categories – electronic noise at the output of the detector that is measuring the radiation, environmental noise due to the thermal fluctuations of a system and its surroundings and source noise, variations in the output from the illumination source [63, 87, 88].

In some cases, phase sensitive detection (i.e. using a lock in amplifier) can move the measurements to a different part of the noise spectrum with less noise, but this may not always be a feasible option for the application.

2.5.1.1.1 Electronic noise

The electronic noise is characteristic to all electronic circuits, and can arise from noise within the detector element or from the readout circuitry.

Johnson noise is due to the random thermal motion (Brownian motion) of charge carriers inside any electrical conductor [87]. The Johnson noise can be reduced by cooling the detector system. The root mean square (RMS) voltage due to this noise is proportional to the square root of the temperature. Johnson noise is present even when there is no applied voltage.

Generation-recombination noise is often only present when current flows and is dependent on the type and quality of the material [87]. In semiconductors, it is due to the fluctuation of charge carrier pairs and recombination due to thermal excitation.

Shot noise is due to the discrete nature of electric charge and photons [87]. Electrical shot noise is attributed to the fluctuations of electric current when charge carriers cross a gap (such as in capacitors). Although the arrivals of the electrons may be constant on a macro level, electrons are particles, leading to the charge arriving in packets with a Poisson distribution. Shot noise in a photodetector is due to the discrete nature of photons.

Flicker noise, also known as 1/f noise (being inversely proportional to the frequency), can be due to a variety of mechanisms. Most commonly it is due to fluctuations in carrier lifetime in semiconductor devices, or due to resistance fluctuations at contacts [87]. Flicker noise is still present at higher frequencies, but is typically overshadowed by other noise effects.

2.5.1.1.2 Environmental noise

In the visible spectrum, environmental noise can be a problem, but turning off the lights or enclosing the experiment in a light-tight box is usually a suitable solution. In the MIR, everything gives off radiation as described by Planck's law (see section 2.2.2.2). These can be corrected for with proper background subtraction, but thermal fluctuations will still be present. This could be from electrical components, the presence of people, or the ambient temperature as the room heats during the day and cools at night.

Most thermal fluctuation is slow compared to the sources of electronic noise – if we can sample faster than the time it takes the thermal background to change, we can reduce the effect of thermal fluctuations.

2.5.1.1.3 Source noise

The output from a thermal source can have some variation in power output due to variations in the driving power source and hence thermal fluctuations. Pulsed laser operation can suffer from a larger source of noise in pulse-to-pulse variation, when not every pulse has the same energy. For a system that only integrates over a period of time, allowing for only a single pulse to register, this can have a large effect on the detected power (especially if ratioed against a background later on, using a different pulse, as in the QCL imaging system). The only way to account for this would be to use a beamsplitter to send the pulse to a secondary detector to measure its intensity, using this value to normalise the value on the primary detector.

A pulsed source introduces another potential problem – if the detector refresh rate is not synchronised with the repetition rate of the source, a variable number of pulses can be detected in each frame. This effect will become less apparent the longer the integration time of the detector is relative to the pulse period, and can be minimised by the averaging over multiple acquisitions.

2.5.2 Current cancer detection and diagnosis procedures

In regards to the detection of breast cancer, the most effective method is examination, checking for lumps. This can either be performed through self-examination or examination by medical staff. Another technique is mammography, a method that uses low-energy X-rays to image the breast, looking for changes in tissue density and microcalcifications, though the technique may not bring any advantage over physical examination in some developed countries [89].

After detection, a biopsy may be taken for diagnosis to be looked at by a histopathologist after H&E staining. Although H&E staining is carried out on all biopsies for diagnosis under an optical microscope, there are other techniques for the detection and diagnosis of cancers. The H&E process and grading system is described in detail in sections 1.1.2.1 and 4.2.2.1.

It is the diagnosis that will then decide what treatment will follow. With breast cancer, a number of factors will play a role in the treatment decision [9]. The tumour grade, following the H&E slide inspection, is used, as well as the tumour size [9]. Whether the cancer cells are 'ER (oestrogen receptor) positive' also plays a role, mainly in deciding on whether to administer hormone treatment [9]. Lastly, if the tumour has already been surgically removed, a number of underarm lymph nodes will also be removed and inspected for cancer cells [9]. This indicates the risk of the cancer metastasising (spreading), and is also considered when choosing a treatment. Depending on all these factors, if the cancer has progressed considerably, an aggressive treatment such as chemotherapy may be suggested [9].

The drug administered for hormone therapy is tamoxifen [90]. Some breast cancer cells require oestrogen to grow – oestrogen binds to and activates receptors in these cells [91]. Tamoxifen is processed by the body into a competitive antigen, binding to the receptor but not activating it, preventing activation by oestrogen [91, 92]. This blocks the growth of the cancerous cells. Tamoxifen should only be used to treat breast cancer if it will work, as it has negative side effects such as bone loss in some premenopausal women, and can increase the risk of some types of uterine cancer [93, 94]. Raloxifene, a drug that has a similar effect to tamoxifen, may have reduced side effects [95].

There are a number of other tests available mainly to indicate the suitability of hormone therapy as a treatment option, including genomic tests such as 'Oncotype DX', the 'Mammostrat' test and the 'MammaPrint' test, and non-genomic tests such as IHC (ImmunoHistoChemistry) [96].

2.5.2.1 Previous work of IR vibrational spectroscopy for cancer diagnosis

Due to their ability to probe the chemical composition of a sample and detect subtle changes, infrared spectroscopy techniques offer a number of potential advantages as clinical diagnostic tools such as *in vivo* tissue diagnosis; objective, non-destructive discrimination between diseased states across a range of pathologies; and detection of diseased states prior to visible structural changes [97].

Raman spectroscopy has been shown to be able to distinguish between pathological subtypes in the oesophagus [6], stomach [98], and in colorectal cancer [99], among others [97]. Similarly, FTIR spectroscopy has been shown to be able to distinguish healthy and diseased tissue at various stages of pathological development in a range of cancer types [97, 100-106].

Much work has also been done on multichannel imaging IR vibrational spectroscopy. In addition to the work performed in the group, these works include the generation of Raman images that are able to distinguish different tissue types in a human polyp colon section (also discussed in section 5) [28]; identification of normal and malignant colonic tissue after extended multiplicative scatter correction [107]; and tissue classification and identification of cellular abnormality in axillary lymph nodes and detection of metastatic breast cancer cells [105]. Aside from Raman and FTIR imaging, work has also been done using QCL sources, imaging tissue at discrete wavelengths. These works include the imaging of mouse intestinal tissue [108]; imaging of breast tissue at the amide I band at 1655^{-1} [109]; and imaging of brain cancer core slices [110]. Reflection imaging with a QCL has been attempted, although not of tissue samples; diethyl phthalate was imaged on a diffuse gold surface, and the contrast was confirmed to be mainly due to the absorption at $1072cm^{-1}$ [111].

3 System design and optimisation

There are two systems to discuss in this thesis – the QCL system and the thermal source-based *digistain* system. As they are both spectral imaging systems, they share many similarities in the principles and considerations for operation.

This chapter looks first at the data being collected and the basic routine taken to get this data. It also looks at the specific components used in both the QCL and *digistain* systems, as well as some specific implementations, such as with the diffuser to destroy the spatial coherence of the laser source. Following the description of the hardware of the systems, some specific issues that were overcome during the development of each system are discussed, and the approaches taken to maximise SNR. The systems have some shared problems, such as the changing thermal background, and some more particular issues, such as a nonlinear detector response (seen only when using the QCL source).

When looking at tissue samples, the data is presented as a ratio between the phosphodiester and amide absorbance. Some of the methods to display the results in the best way to the user are discussed, as well as how to process the data to avoid nonsensical results.

The hardware used for data acquisition and the subsequent processing of the data is all handled by a program made using LabVIEW software.

3.1 Data collection routine

The way in which the data is collected and processed is similar for both the *digistain* and laser setups. The absorbance needs to be calculated at various wavelengths, which is calculated directly from the transmission (see section 2.2.1.1). Due to the necessity of background correction, a transmission image will be made up of four separate images. These four images will themselves be an average of many images.

With a 'reflection' image, the difficulties of using another sample surface with perfect diffuse reflectivity to get quantitative data means one of the steps for background correction is not required, and only two (rather than four) images are used to make a reflection image. The image will not have quantitative data, but will still show contrast.

The phosphodiester-to-amide (PA) ratio is mentioned regularly in this thesis, and the equation to get the PA ratio is described in this section with a small discussion of intrinsic problems that may arise and how to get around these problems.

3.1.1 Transmission image creation

When doing spectral imaging in transmission mode, a value for the transmission across the image is required. This is calculated based on equation 2-2 in section 2.2.1.1. To record only the intensity of the source light, a background has to be taken to eliminate IR radiation from the surrounding environment or artefacts in the readout from the camera. The transmission at a single wavelength is calculated from four separate images as described in equation 3-1.

$$T = \frac{S - E_S}{B - E_B} \tag{3-1}$$

where S is the sample image, when the sample is being imaged with the source 'on', and E_S is the environmental background when the sample is being imaged but the light source is 'off' or being blocked. B is the background, when the sample is not in the optical path but the source is on, and E_B is the environmental background when the sample is not in the optical path and the source is off or being blocked.

As described in section 3.3.2.3, to avoid drift in the value of the four images, the four images themselves can be made up of many images that were not taken consecutively.

The background images, B and E_B are taken when the sample is removed. Tissue samples necessarily have to be on substrates (BaF_2 or CaF_2) that have non-zero absorption that is almost spectrally flat across the spectral range being investigated. To measure the transmission of the sample alone, the background images, B and E_B should be taken when the sample is not in the field of view, but a part of the substrate without sample on it, or an identical substrate, is being imaged. As explained in section 3.3.2.1.1, taking the background images with a substrate in place also has the advantage of maintaining the same intensity distribution at the camera as the sample images.

As there is a risk that the denominator of equation 3-1 is equal to zero, the transmission image undergoes a routine to replace any pixels equal to $\pm\infty$ with the median of its nearest neighbours, similar to the routine described in section 3.4.2.

3.1.2 Reflection image creation

In diffuse reflection imaging, a similar approach to transmission imaging as in section 3.1.1 can be taken, following equation 3-1. To acquire reliable quantitative information, the background images, B and E_B should be taken with the sample removed, but a target should be inserted to reflect the light with no absorption taking place. In the visible, usually a calibration target is used, such as Spectralon, a brand name for a fluoropolymer that has very high diffuse reflection across the UV, visible and NIR. In the NIR and MIR, there is also 'Infragold', a reflection coating that has high diffuse reflection. For a quantitative number though, the calibration target used to take the background images should reflect light in the same way as the sample in question, but with zero absorption taking place, a difficult condition to satisfy when using samples with an irregular surface. The more immediate goal is to obtain qualitative information. For this reason, the reflection image is made up of only two images, as shown in equation 3-2

$$R = S - E_S \tag{3-2}$$

where S is the sample image, when the sample is being imaged and the source is 'on', and E_S is the environmental background when the sample is being imaged but the light source is 'off' or being blocked. If using light sources that can be turned on and off (such as the QCL unit used in this thesis), then no moving parts are required to make a reflection image.

3.1.3 Phosphodiester-amide ratio image creation

Once transmission data has been collected, in the case of the *digistain*, the images at the four different wavelengths are combined in a way to display the amount of phosphodiester bonds (corresponding to nucleic acid) present relative to the amount of amide bonds (corresponding to proteins). The basic way to do this is expressed in equation 3-3:

$$PA_{ratio} = \frac{A_{P\,peak} - A_{P\,base}}{A_{A\,peak} - A_{A\,base}} \tag{3-3}$$

where A_X is the absorbance at spectral feature X. The subscripts refer to the phosphodiester (P) and amide (A) features, at the 'peak' (tip of the peak) and 'base' (bottom of the peak) of those features. By subtracting the base from the peak, only peak height is regarded, largely removing the effect of a baseline in the spectrum, as shown in Figure 3-1.

The transmission images at each wavelength are converted into absorbance to circumvent the effect of differing thicknesses and nonuniform optical path length of the illumination light within the sample or between samples. Assuming no saturation of the absorption mechanism, if the thickness of a sample is increased by a factor of x, the absorbance at all four wavelengths will increase by a factor of x, hence PA_{ratio} will remain unchanged.

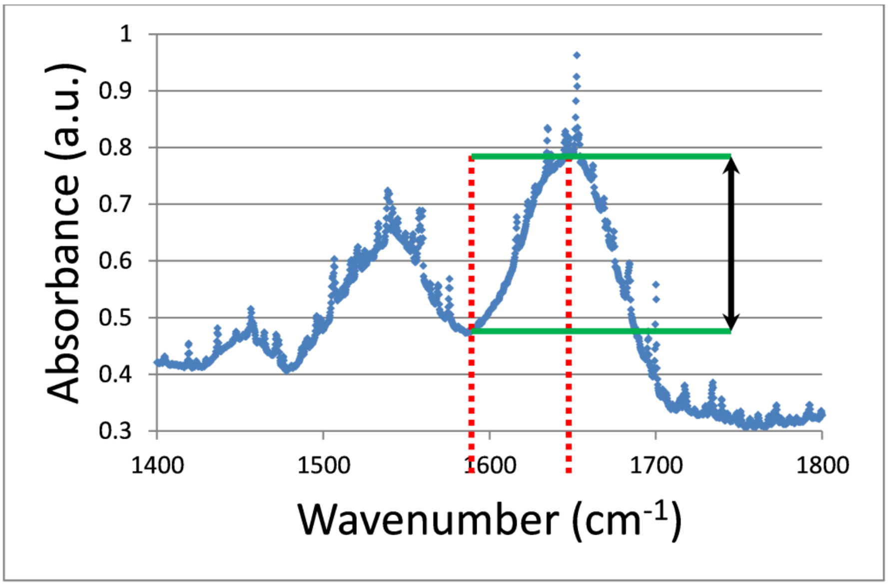


Figure 3-1. Process of taking the absorbance peak height to remove baseline effects. Two wavelengths are used, one corresponding to the highest point of the absorption peak, the other serving as a baseline reference. By subtracting the two absorbances, only the magnitude of the peak height remains.

When imaging tissue the absorbance value at the 'peak' wavelength is generally greater than at the 'base' wavelength. However, this may not always be the case across the entire image; some areas may have higher absorbance at the base wavelength than the peak. This is possibly due to a low presence of the molecules that contain the spectral peak being studied, and another molecule's spectral peak overlapping with the base wavelength. This brings significant problems when taking a ratio – when some parts of the numerator or denominator images are above zero and some parts are below, the areas with a higher concentration of phosphodiester bonds will not necessarily always appear as such. What is needed is to bring all parts of the image above zero. The obvious way to do this would be to add constants to the numerator and denominator. As long as the constants are always kept the same, different samples can be compared reliably. However, if just adding a constant, the ratio image would no longer be independent of sample thickness.

Instead, a method that retains the benefits of using absorbance in the ratio to remain thickness-independent is employed. The 'peak' absorbances are increased slightly to make sure they are always larger than the 'base' absorbances, as detailed in equation 3-4:

$$PA_{ratio} = \frac{(X.A_{P\,peak}) - A_{P\,base}}{(X.A_{A\,peak}) - A_{A\,base}} \qquad 3\text{-}4$$

The ‘peak’ absorbances are multiplied by a factor X. As long as X is kept constant across all samples, the values of the ratio will be independent of sample thickness and can be compared on the same scale. The larger X is, the closer the ratio gets to being a ratio of the peaks without the baseline correction, so X should be kept as low as possible. As Figure 3-2 shows, the value of X need only be slightly above 1 for the ratio to enter a positive regime for typical absorbance values obtained by the *digistain* system when looking at a tissue biopsy.

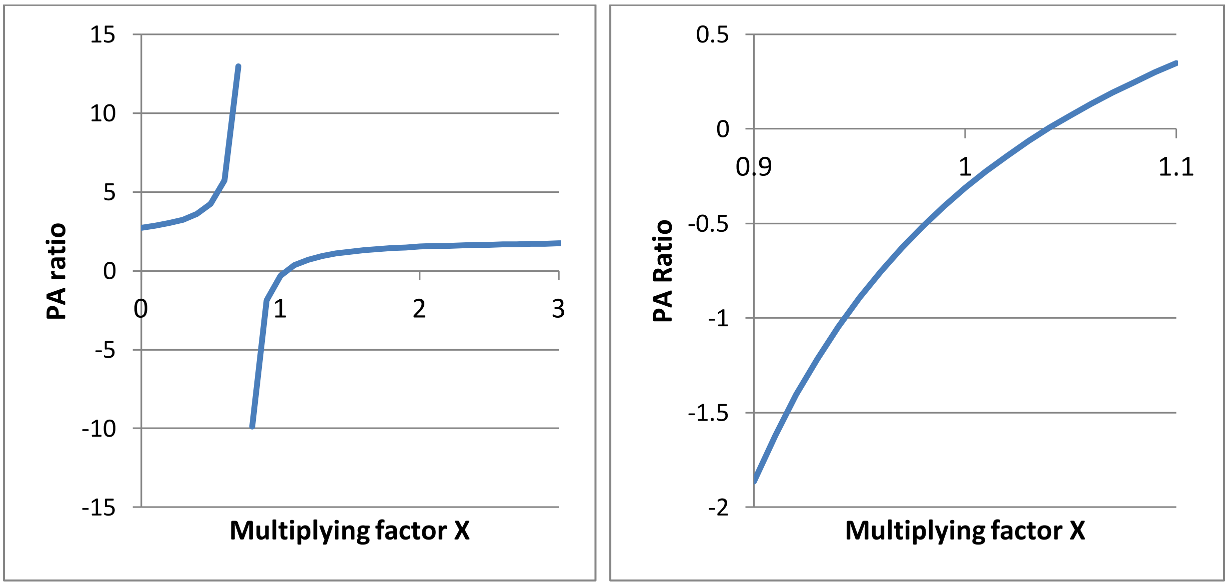


Figure 3-2. Effect of multiplying factor X on a typical PA ratio value. Looking at the left chart, when X is slightly above 0, the PA ratio is positive due to the numerator and denominator both being negative. As one of the numerator or denominator becomes positive, the ratio switches to being negative. As X increases to greater than 1, the ratio becomes positive once more. Focusing on the region around 1 when the line crosses the x-axis (right chart), it can be seen that X only needs to be slightly larger than 1 to get a positive PA ratio. Although in this case X=1.05 will be enough, to make sure that we can use the same value for X across samples and in different areas of the image, it is safer to use a larger value still for X such as X=1.2.

Since the practice of using a blank substrate during the acquisition of the background images, B and E_B, has been implemented, the problem of negative numerators/denominators has become less of a problem such that the equation needs no modification, remaining as equation 3-3. This may be because although BaF_2 and CaF_2 are considered spectrally flat between certain wavelengths, they may not have been flat enough, with the differences being large enough to be picked up by the spectral imaging system. Using a blank substrate during the background procedure removes any spectral modulation caused by the substrate.

3.2 Hardware & setup

The setups for the two systems are quite similar, with common techniques used and common layouts. The hardware used in the laser setup and the *digistain* setup could in theory be used almost interchangeably, though the instrumentation in the laser setup should have higher spectral radiance and sensitivity, as well as more flexibility (having a continuous tuning range). The *digistain* has been implemented in a transmission setup only, whilst the laser setup can be changed between transmission and reflection modes.

3.2.1 *Digistain* system

The *digistain* system uses a thermal element as the light source that emits across a wide range of wavelengths simultaneously. The wavelengths are then selected by a set of narrow bandpass filters. The *digistain* is designed to work at 4 specific wavelengths only, at a fixed magnification and only in transmission mode for work with tissue biopsies. As the *digistain* system is designed with possible future commercialisation in mind, cost is taken into consideration leading to decisions such as using a thermal source instead of laser sources, and an uncooled bolometric detector instead of a potentially more sensitive cooled one. The setup also has a ‘visible’ arm to work with H&E stained samples for comparison with the *digistain* images. A photograph of the *digistain* setup is shown in Figure 3-4 on page 66, with the key components marked.

3.2.1.1 Thermal source

The thermal source used is an “IR-35” from Scitech Instruments [112]. This is a coiled filament that when operated in a steady state, is held at 950°C (a peak wavelength of ~2.4µm) [112]. The element has an emissivity of ~0.75 and is positioned at the focus of a parabolic reflector to give a brighter, more directional output [112]. The manufacturer estimates that the presence of the parabolic reflector increases the on-axis output by a factor of 11 [112].

3.2.1.2 Bandpass filters

Four bandpass filters are used to select the wavelengths from the broadband thermal source. The filters are custom-made, and consist of both single filters and combinations of filters that result in only a narrow band being passed. The peak transmission of the filters ranges from roughly 50 to 90%. The centre wavelengths of the filters are designed to match as closely as possible with the centre wavelengths of the absorbance peaks and a point at the base of each peak. Coincidentally, the filters roughly match the width of the vibrational mode peaks. As a consequence, there is some non-symmetrical overlap between the 'peak' and 'base' filters. The transmission of the filters is displayed in Figure 3-3, along with a typical absorption spectrum of tissue.

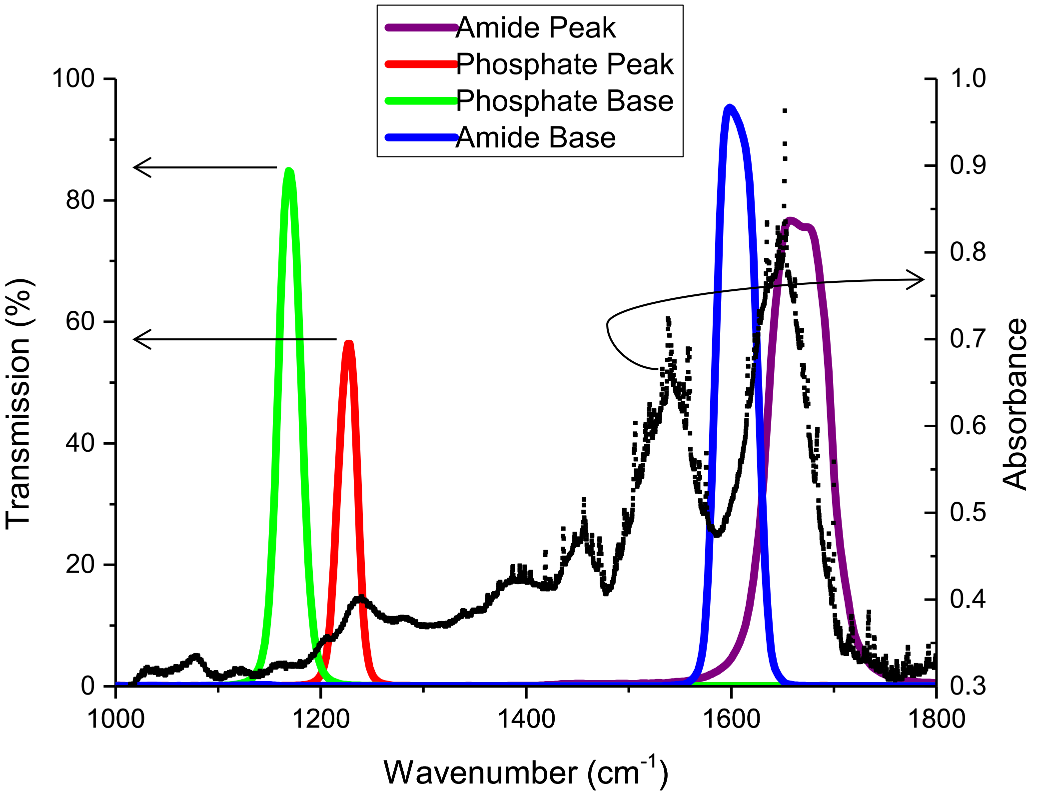


Figure 3-3. Transmission of digistain bandpass filters (coloured traces, measured on left axis) overlaid on a typical tissue absorption spectrum (black data points, measured on right axis). The filters are centred on the amide I and the antisymmetric phosphodiester peaks. One filter is for the peak absorption, while the other is targeted at the 'base' of the peak for proper background correction. Although the transmission of the 'peak' filters is lower than the 'base' filters, due to the shape of a black body emission curve, the 'peak' wavelengths may still see higher throughput. The amide I wavelengths will have greater intensity than the phosphodiester wavelengths after passing through the bandpass filters.

Ideally, the filters could be narrower to avoid overlapping with their 'peak'/'base' counterpart, with higher peak transmissions to increase the SNR and a better centring on the absorption peaks of the tissue spectrum.

3.2.1.3 FLIR bolometric camera

The camera is an uncooled silicon microbolometer array, with 384x288 pixels, each pixel being a 35µm square. The camera can detect light at wavelengths ranging from ~2-14µm. The camera outputs pixel values at 14-bit resolution and has a fixed refresh rate of 50Hz and a 60µs integration time.

3.2.1.4 *Digistain* transmission setup

The *digistain* transmission setup has all components aligned in a line along a single axis. The setup is shown in Figure 3-4.

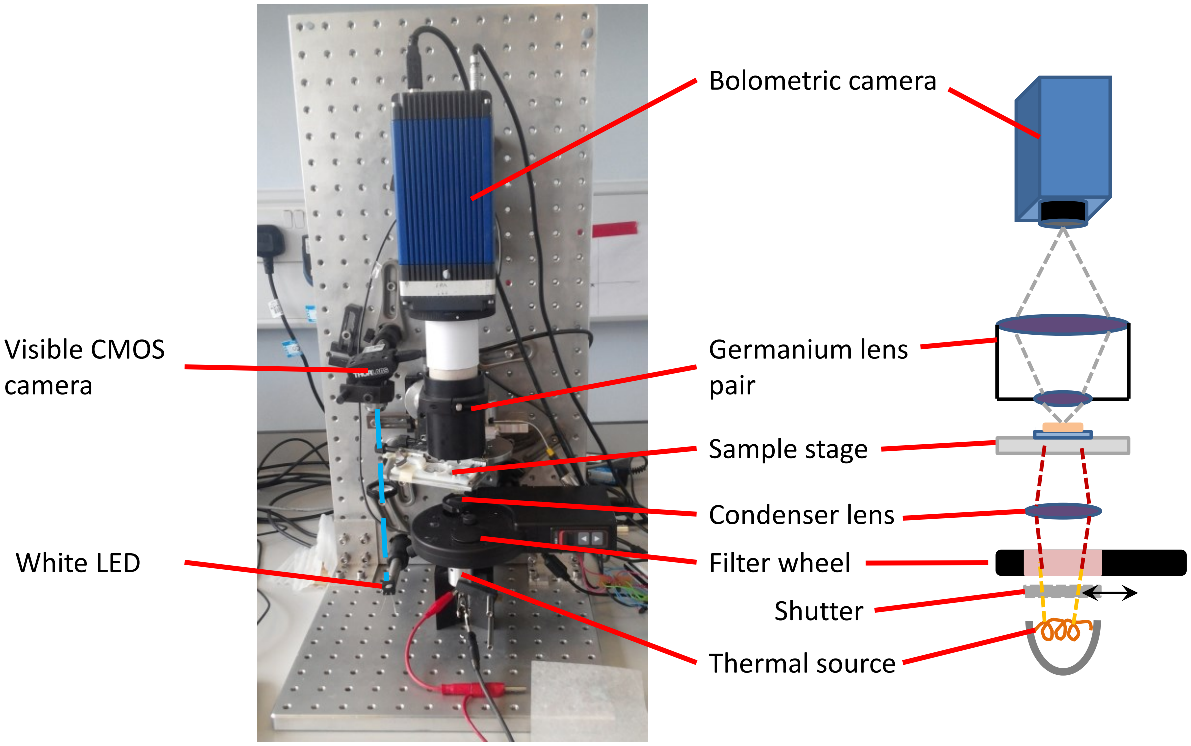


Figure 3-4. Digistain transmission setup photo, with IR line diagram to the right. At the base of the unit is the thermal source that emits broadband radiation. Above that is the shutter that effectively turns the source 'on' and 'off', followed by the switchable bandpass filters. A germanium condenser lens captures some of the diverging light from the source and concentrates the collimated light into the field of view. The sample (on a 1 inch diameter disc) is held within a rectangular adaptor, the adaptor itself held on a modified microscope stage. This is connected to a DC motor that is able to move the sample in and out of the field of view. Two custom-made lenses within a lens tube image the sample onto the IR camera. The 'visible line' (see section 3.2.1.5) is to the left of the main IR line. A CMOS camera is at the top, illuminated by a white LED from below. The sample can be held within the same modified microscope stage when being imaged by the visible line.

The setup is assembled on a small breadboard, with the components arranged vertically, giving it a footprint of 30cmx30cm. The thermal source at the bottom is driven by an external power supply. Above is the motorised flipper mount that acts as a shutter to turn the thermal source on and off. The blocking element is a gold mirror on a glass substrate. The mirrored surface faces the source to reflect the vast majority of the MIR light, whilst optical glass, a strong absorber in the MIR, should efficiently block any remaining light. There is some concern that the shutter element itself could heat up and start emitting light in the MIR. The risk is minimised thanks to the glass substrate; glass is a very poor conductor of heat relative to other solid materials (such as metals).

Between the source and the condenser lens is the filter wheel. The filter wheel can hold 6 filters and switch between them, controlled by a computer. The time to switch between adjacent positions is quoted by the manufacturer as 0.9 seconds. The filter wheel could be placed at any point along the optical path, and there are arguments for various positions. Placing the filters immediately after the source would reduce the risk of the shutter heating up, but the filters themselves could heat up and contribute additional radiation. Placing them just in front of the camera will remove the majority of the background thermal radiation, but then there is a risk of overheating the sample. The current position will cut the danger of sample overheating, whilst cutting out the majority of MIR emission from the shutter that could occur.

A single germanium lens with an 8-12μm antireflection coating (although the transmission of the lens at 6μm is still 70%) concentrates the collimated and diverging light from the source onto the sample. The positions of the source and condenser lens have been optimised to maximise throughput.

At the image plane, a microscope stage part with two directions of movement holds a microscope slide-sized object (76mmx30mm). It is attached to a DC motor to move the sample in and out of the field of view, controlled by a computer. To view a standard H&E slide, the slide can simply be placed on the stage. The MIR slides used are usually circular discs – to fix them on the microscope stage, a 3D-printed adapter is used. The rectangular adapter is gripped as a standard microscope slide would be. The adapter holds two BaF_2 discs, one for the sample and another for background purposes (see section 3.3.2.1.1).

The objective lens used in the *digistain* setup was designed by Eddie Judd and manufactured by Davin Optronics Ltd. It was designed to work at a fixed magnification of 3x, and maintaining pixel registration primarily at wavelengths of 6 and 8μm, where there are amide and phosphodiester peaks.

The resulting objective lens is made of two germanium lenses, a meniscus lens at the front and a plano-convex lens behind. One of the front lens surfaces is aspheric to improve performance. Splitting the focusing between two lenses should reduce distortion, and allows for a slightly lower f-number (better light collection) than would be viable with a single lens. The lenses have been anti-reflection coated for the 6-10µm region to give high throughput. The lenses provide a spatial resolution of 25µm at the image plane as verified experimentally, with a magnification of 3x.

Between the imaging lenses and the camera is a cardboard tube to eliminate stay radiation (from 'bright' thermal sources such as people in the room and warming electrical components). Enclosing the entire setup in a MIR-opaque box would further assist in reducing background intensity.

3.2.1.5 Visible line of *digistain* system

As well as the IR system, there is a visible line in the setup that can image in the visible, for H&E stained samples on microscope slides. The sample is illuminated by a white LED, and is imaged onto a 1280x1024 pixel CMOS colour camera with 3.6µm square pixels. The imaging lens is a standard microscope objective, a non-infinity corrected achromatic lens with 10x magnification for the operating tube length. The visible line can be seen to the left in Figure 3-4.

The visible line is intended to work with standard H&E slides for the sake of comparison with the IR images. The visible image is taken when a normal microscope slide is inserted into the microscope stage instead of the rectangular adapter. The H&E stained tissue can be imaged when the sample is in what is the 'out' position for the normal IR line. There is no complex procedure to take the visible image as in the IR, and the process takes no more than a second. Once the visible image is taken, it can be matched (through automatic rotation and cropping) and displayed side by side with its corresponding *digistain* image for ease of comparison.

3.2.2 Laser-based system

The laser-based system uses a cooled semiconductor photodetector as the camera, and a continuously tuneable quantum cascade laser (QCL) as the light source. This wide range of output wavelengths gives relative flexibility, as any spectral feature of interest within the operating range of the laser can be probed. Using a laser source, however, necessitates the use of a diffuser to deal with the spatially coherent light that leads to unwanted imaging artefacts. Due to the higher sensitivity of the detector and spectral radiance of the source, the laser system is able to operate in a reflection mode as well as a transmission mode. The transmission mode is most suitable for looking at thin samples, such as biopsied tissue samples, whereas the reflection mode is more suitable for thicker samples and has potential in-vivo imaging from, for example, skin. Setups of the transmission and reflection modes are shown in Figure 3-6 and Figure 3-7 respectively, on pages 71 and 72.

3.2.2.1 CEDIP camera

The camera is a 320x240 pixel array detector, each pixel being a 30µm square piece of HgCdTe. It operates at 77K, being cooled by a Stirling engine.

The integration time (IT) and the refresh rate can be independently controlled, although there is a maximum repetition rate that the camera can be operated at that is dependent on the IT. The minimum rate the camera can operate at is 25Hz, with a maximum of 259Hz when reading the full array. Due to the way the camera settings can be used, a longer IT operating at a maximum refresh rate will increase the acquisition time within one second. Therefore (for an experiment in which time resolution is of low importance) the IT and refresh rate should be chosen to maximise the total IT over one second, by operating at the maximum IT that does not saturate the 14-bit pixel values.

3.2.2.2 Quantum Cascade Laser (QCL)

The QCL is produced by Block Engineering and is tuneable in the range 1000-1667cm^{-1} (wavelengths ~6-10µm). The entire range is provided by 3 separate chips, with each chip providing as range of ~200-250cm^{-1}. The laser operates pulsed, with the duty cycle quoted as 5%. When triggering internally, the repetition rate can range from 63 up to 200kHz; whilst when using external triggering, the repetition rate can go up to ~1.3MHz. The pulse length can be varied from 32 to 480ns.

There is a slight but noticeable shift in beam position and intensity distribution as the laser tunes across the entire spectral region, shifting between the chips. The pulse shape is not always identical for a fixed pulse length, although for these spectral imaging experiments, it is the stability of the pulse-to-pulse power and the time-averaged power which is more important. A typical value for the pulse-to-pulse peak power variation is 4%. In a typical setup with the CEDIP camera, 100 pulses might be captured in a single ‘frame’ of the camera, meaning the frame-to-frame variation would be 0.4%. Additionally, as a single acquisition is made up of hundreds or thousands of frames, the pulse-to-pulse variation of the laser becomes relatively insignificant when compared to other noise sources, such as the presence of a diffuser.

A power spectrum of the QCL output is shown in the upper image in Figure 3-5. In the lower image in Figure 3-5 is shown the power spectrum recorded by the CEDIP detector that takes into account factors such as the detector sensitivity and atmospheric absorption.

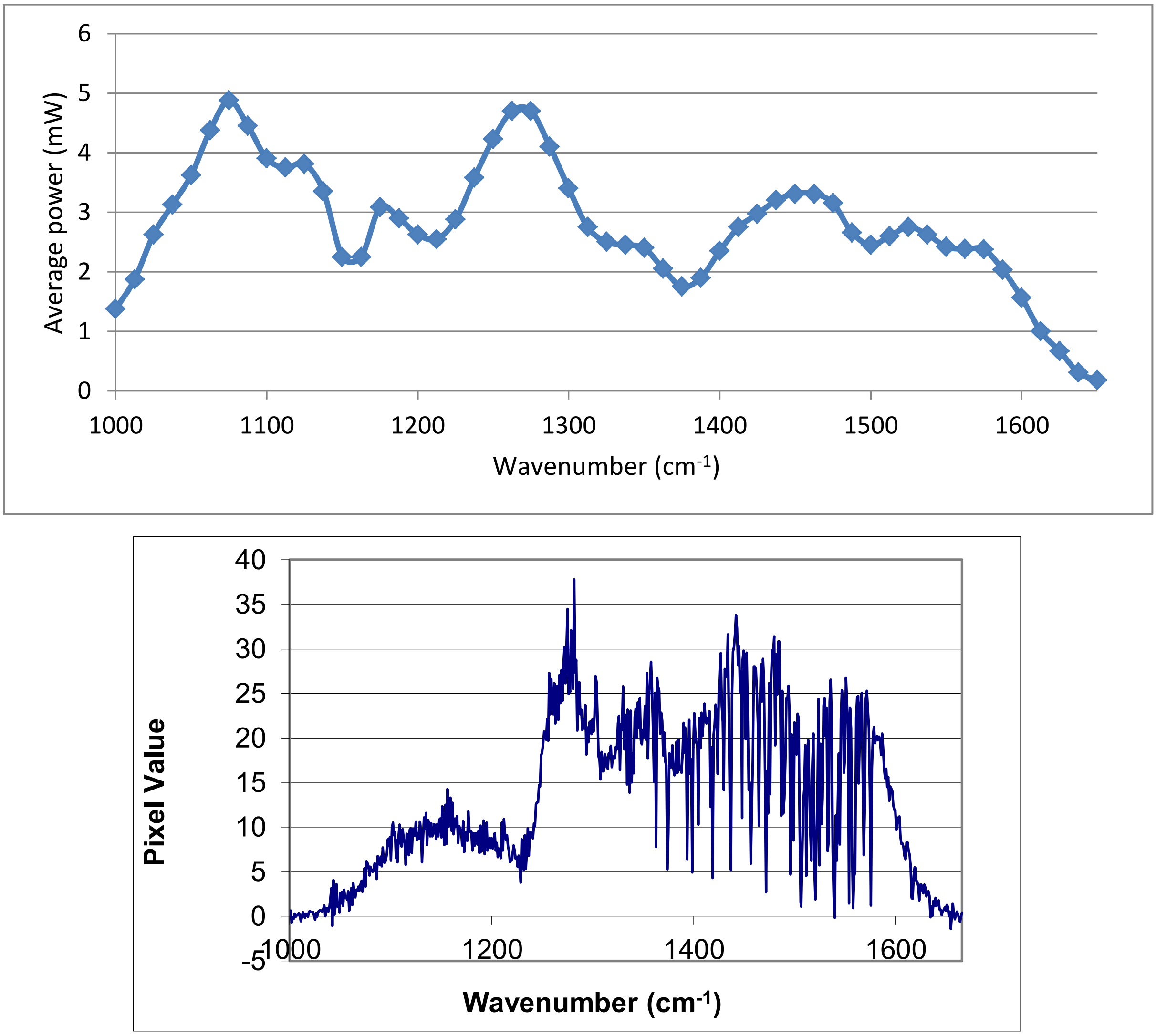


Figure 3-5. The power spectrum of the Block Engineering QCL, as specified in the supplied documentation (upper graph). The power spectrum as seen by the CEDIP camera will differ, as detector response, atmospheric absorption and optics used are taken into account. It was found in practice that the range 1000-1100cm^{-1} was unusable for most setups due to the detected power being too weak. Wavenumbers above 1600cm^{-1} can have a similar problem. The lower graph shows the power spectrum recorded by the CEDIP camera after diffuse reflection from aluminium. Note the spikes of low power in the range 1400-1600cm^{-1}, attributed to absorption by water vapour (see section 2.4.2.4).

3.2.2.3 QCL transmission setup

The QCL transmission setup is outlined in Figure 3-6. The diffuser serves to destroy the coherence of the beam. To collect any scattered light, and match the illumination to the field of view, the diffuser is placed a distance $4f$ from the image plane with a lens at the $2f$ point. This gives a one-to-one magnification of the beam and diffuser. If the magnification of the image changes, the condenser lens can be adjusted to expand or reduce the distribution of the illumination light to fit the field of view. It was found that this setup performed better at removing coherence than a setup in which the diffuser is out of focus at the image plane. Germanium is the preferred material for the lenses, due to its low dispersion that will maintain pixel registration across wavelengths. Other studies using a MIR QCL have utilised reflective optics to achieve this [4, 113]. The sample at the image plane is held within a motorised flipping mount that can be controlled by a computer using LabVIEW.

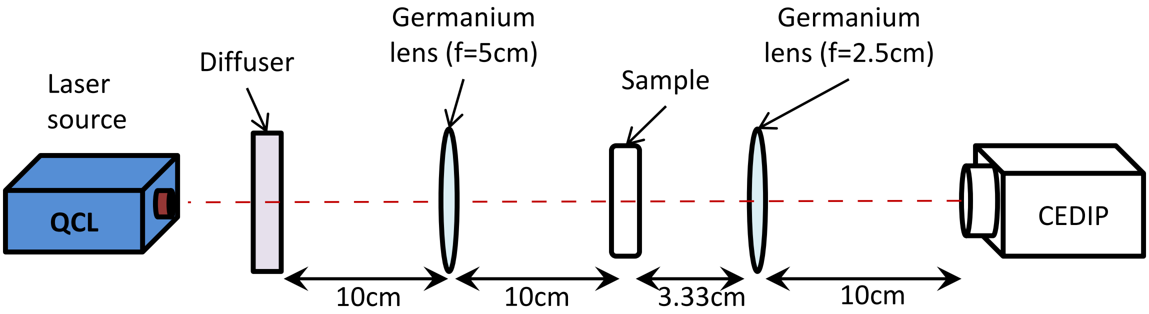


Figure 3-6. Laser transmission setup with approximate dimensions used. The collimated output from the QCL first passes through a moving diffuser to scramble the wavefronts and remove the coherence of the source. A germanium lens with an 8-12μm wavelength antireflection coating then focuses the beam and images the diffuser onto the sample in the image plane in a 4f configuration. The non-uniform illumination, arising from the Gaussian output profile of the QCL and modification by optical elements, is always accounted for during the background correction. A second germanium lens with identical antireflection coating images the sample onto the CEDIP camera.

The condenser and objective lenses have a standard MIR antireflection coating from wavelengths ~8-12μm to account for the high Fresnel losses from high-index germanium. The transmission falls off towards 6μm, but still remains above 70%.

This setup actually ends up throwing away almost 50% of the light intensity (mainly due to the diffuser), but in this case it is of little concern, as the intensity of the illumination light will have to be attenuated to reduce the effect of nonlinearities in the detected intensity (see section 3.3.3.2).

3.2.2.4 QCL reflection setup

The reflection setup differs slightly from the transmission setup in section 3.2.2.3, and is shown in Figure 3-7. The main change is that the laser beam is incident from a different direction. The sample holder at the image plane may change – if looking at a sample mounted on a BaF_2 disc, the same motorised flipper mount may be used (although there is usually no need to move the sample in and out). If looking at someone's arm, there should be a fixed structure at the image plane that the person can brace their arm against to keep it steady. The incoming light should not be incident on the normal of the sample's surface to avoid any specular components being captured by the imaging lens, which will be normal to the sample surface.

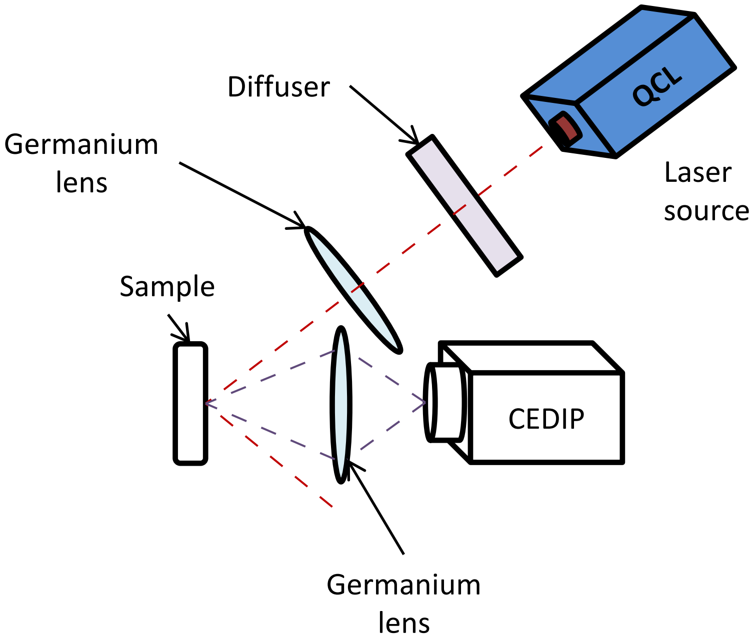


Figure 3-7. Diffuse reflection imaging laser setup. The laser beam passes through the diffuser to destroy the spatial coherence of the light. The diffuser is imaged onto the sample surface at a magnification which fills the field of view. The main specular is not picked up by the camera, whilst the diffuse reflection is collected by the objective lens.

3.2.2.5 Diffuser to remove coherence

The coherence effects seen when imaging come from the spatial coherence of the source (see section 2.3.1.1). It is the spatial coherence that has to be changed. A single window with a varying surface profile on one surface is enough to modify the wavefront of light as it passes through.

When a diffuser is used to remove coherence, it is usually rotated, primarily to make sure the intensity distribution is uniform after being time-averaged. This approach has been utilised in other studies using an MIR QCL [113]. The problem with this implementation in practice is that the movement is not sufficiently randomised, such that the rotation can be seen overlaid on the image, as seen in Figure 3-8.

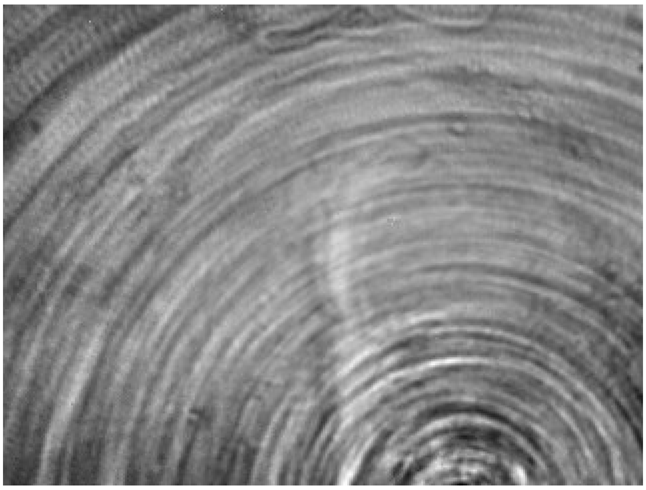


Figure 3-8. Transmission image of tissue biopsy in which a pattern from a rotating diffuser can clearly be seen overlaid. The centre of rotation is off-image, to make sure that there is no part of the beam which does not experience a time-varying signal. The rotation overlay is thought to be present due to a lack of randomness, and a slight change in the position of the diffuser.

The design from Ref. [68] (Kubota, 2010), was used, partly to remove the rotational overlay seen, partly because the paper concludes that their method of diffuser movement improves speckle reduction over a rotating diffuser. That paper utilised what is essentially a vibrating tuning fork, with the diffuser resting on one arm of fork. Instead of a tuning fork, a loudspeaker was used here (see Figure 3-9). To remove the periodic motion, the diffuser was not firmly attached to the loudspeaker cone (see right image of Figure 3-9), resulting in the motion of the diffuser being much more random. When it was firmly fixed onto the speaker, there was still evidence of an overlain pattern, but a pattern is not detectable when the diffuser is detached.

Critical illumination (in which the diffuser is imaged at the sample plane) was chosen as it appeared to give the best illumination and the highest SNR.

The diffuser was made by sanding the surface of a BaF_2 disk to roughen the surface. P60 was determined to be the best sandpaper grade by balancing two factors – SNR, and whether coherence was destroyed. The finer the grit, the higher the SNR, probably because the coarser grits scatter the light to wider angles, meaning not all the light is being collected. On the other hand, the finer grits do not do a good enough job of removing the coherence. A balance should be found in which the diffuser is made with as fine a grit as possible, whilst still removing coherence sufficiently.

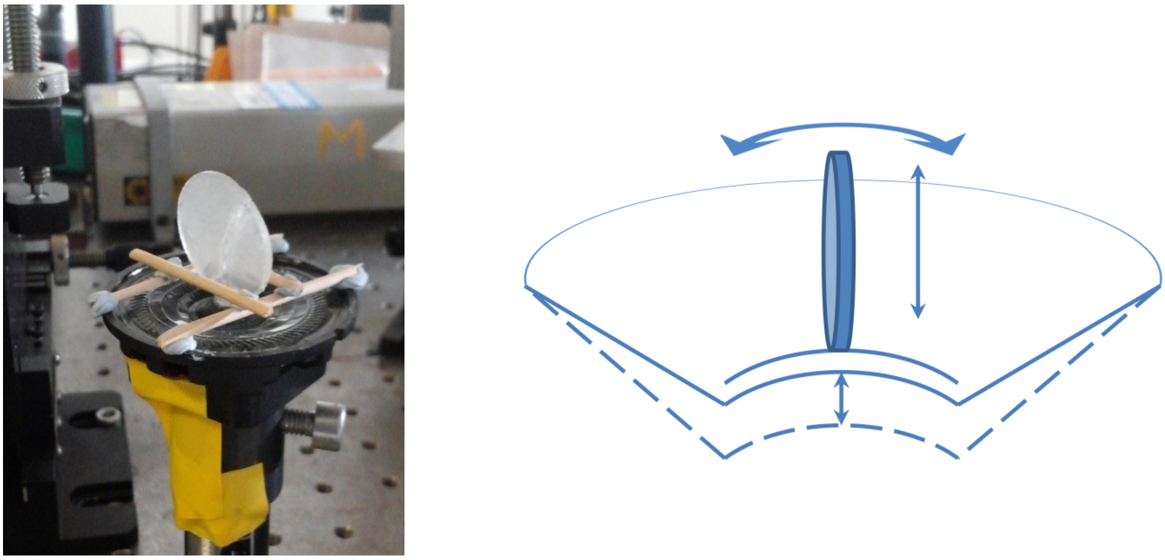


Figure 3-9. Photo of the diffuser implementation on a loudspeaker (left) and diagram showing diffuser and loudspeaker cone movement (right). When the loudspeaker is on, being driven a fixed frequency (usually 50-100Hz), the diffuser bounces about on the loudspeaker diaphragm. The diffuser is also allowed to pivot slightly. This motion is able to randomise the part of the diffuser the light passes through, giving a uniform time-averaged illumination and removing spatial coherence.

Commercially available diffusers designed for the visible and NIR part of the spectrum cover a wide range of surface roughness – from 10µm surface features (~5 times the operating wavelength, on the boundary of the Mie regime) to 115µm surface features (~330 times the operating wavelength). The most common diffuser for the visible spectrum appears to be one with a roughness of ~124 times the operating wavelength, far beyond the Rayleigh or Mie scattering regimes (see section 2.3.1.2). If we apply this roughness to the MIR, it would correspond to surface features of roughly 990µm, requiring a sandpaper grade of P20. However, P60 grit sandpaper (~35 times the wavelength size) was found to give the best performance, putting the diffuser still in the range of geometric scattering.

There are low-cost commercially available diffusers that operate in the IR – reflective diffusers, essentially metal-coated glass diffusers. However, using them in practice showed that the homemade BaF_2 diffuser performed better than reflective diffusers.

Diffuser designs beyond two roughened surfaces were also investigated, specifically, a diffuser that had inclusions within it of a different refractive index (a porous material would have the same effect). One effort was made, with the help of Simon Bastians at the Department of Chemistry at Imperial College, to make such a diffuser, using KBr powder and a compression vice. Although the resulting disc was opaque to the eye (so scattering of visible light was certainly happening), it had no scattering effect on MIR light at all. Combining two powders of differing refractive index may still work. NaCl was also used to try to make a porous material using a different manufacturing method [114], but could not be used, possibly due to highly-absorbing water being locked within the diffuser.

Ways to remove the coherence without a diffuser have been attempted based on the ideas mentioned in section 2.3.1.1, such as a spinning wedge to vary angle or a light pipe, but none have proven sufficiently effective in the MIR experiments to remove coherence. However, the diffuser settled on(the roughened BaF_2 disk on the loudspeaker) actually accounts for the bulk of the image noise seen – longer averages are required to ensure identical illuminations when taking the background and sample images.

3.3 Development choices and considerations

When developing the systems, a key goal is maximising the SNR (increasing the reliability and repeatability of the results) without having to increase acquisition times.

As described in section 2.5.1, the SNR can refer to two separate measures – the SNR across an image (primarily reflecting the differences between individual pixel responses and readout noise), and the SNR of a spectrum (reflecting how accurately a spectrum may be measured).

The method implemented to get the SNR on an image is to take a '100% T' image in transmission mode. With this method, an image is taken in accordance with the routine described in section 2.2.1.1, with no sample being inserted, and the SNR calculated as in equation 2-10. The mean value of all the pixel values will always be ~1 (i.e. ~100% T). This will give an SNR that is related to the random differences between the pixels, and the random distribution of pixel values around 1, but this value alone may not tell the whole story.

If there were a case where 90% of the pixels had a value of 1 and 10% random pixels had a value of 0.9, it would give the same SNR as an image that has those 10% of pixels all in one corner, maybe because of obstruction or another problem with the illumination. These cases are represented in Figure 3-10. The second case, in which there is an obstruction, would not give an accurate value for the SNR of the image, as the difference between pixels is not due to the random noise. To account for this potential effect, a second measure for the image SNR is used, called 'macro SNR'. To calculate the macro SNR, pixel oversampling is used (a post-processing equivalent to pixel binning). The pixels in a 16x16 square, say, are combined into a single 'super pixel'. The SNR in this low-resolution image of super pixels is taken. If the noise on the image is randomly distributed, the macro SNR should increase as the normal SNR does. Looking at how the macro SNR behaves can indicate if there is an issue with the illumination that otherwise may have been too subtle to notice.

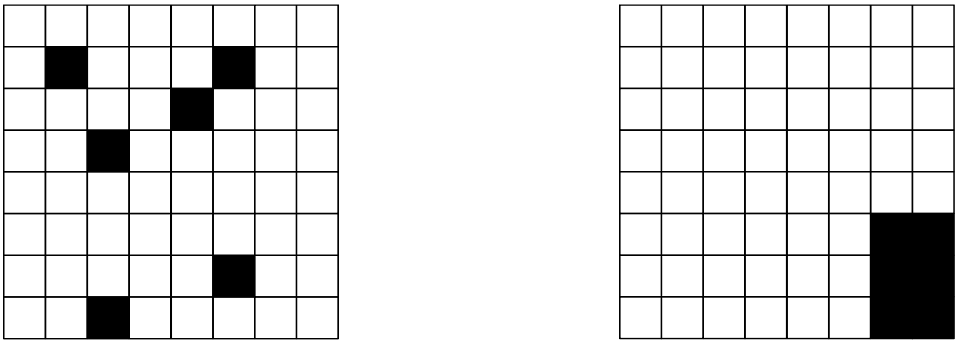


Figure 3-10. Demonstration of a non-random distribution of pixel values affecting the measurement of SNR across the image. The white pixels represent pixels with a value of 1, black pixels represent pixels with a value of 0.9. In the figure on the left, the 0.9 pixels are randomly distributed, as is expected. In the figure on the right, the 0.9 pixels are concentrated in one corner, probably due to some obstruction in that corner of the image. The right image will have its measured SNR artificially decreased.

The spectral SNR can be evaluated in a similar manner to the SNR of an image: a 100% T spectrum (a spectrum when no absorbing sample is inserted into the beam path) can be taken, acquiring data across a range of wavelengths, and equation 2-10 used to calculate the SNR (how 'noisy' the spectrum is). Note that this value for SNR may not be reliable if systematic errors are present. Systematic errors may arise from a nonlinear response of the detector to incident radiation from a QCL (see section 3.3.3.2).

The SNR (both spectral SNR, and SNR on an image) will usually be dependent on wavelength, as the sources and detectors will have a wavelength-dependent output and response.

3.3.1 Shared camera considerations

Although the cameras operate on different detecting principles (bolometric vs. photodetector), there are a few common operating issues that are shared. These include the practice of using a non-uniformity correction (NUC) to account for differing pixel responses and the influence of the thermal background.

3.3.1.1 Non-Uniformity Correction (NUC)

The IR cameras used are made up of tens of thousands of detecting elements (pixels). Each element's response to light will be different from the others, meaning that even when exposed to a uniformly-illuminated scene, there will be a lot of apparent noise and variation in intensity across the image. To account for this, it is common for a non-uniformity correction (NUC) to be used. The 'non-uniformity' in this case refers to the pixel response, not the illumination. The NUC brings each pixel's response to radiation in line with an average response of the camera, as shown in figure Figure 3-11.

Figure 3-11. Effect of NUC to provide a uniform response to radiation. The black dotted lines show a few individual pixels' responses to incident light. The pixels have different 'resting values' when there is low radiation, and a different response gradient as the radiation intensity increases. The thick red line is the average of all the individual pixel responses, and will be the response of all the pixels once the NUC has been applied.

The NUC process works by giving the camera two different exposures that will give two pixel values, and then for every pixel fitting a straight line equation to the response. The two values can be obtained by exposing the camera to two different levels of intensity, usually two sources of different temperature. The way to generate an NUC is as follows.

The detector response, Y, of an individual pixel in row i column j is a function of the incident radiation X:

$$Y_{ij} = m_{ij}X + b_{ij} \tag{3-5}$$

$$\bar{Y} = \bar{m}X + \bar{b} \tag{3-6}$$

where m is the gradient, b is the y-intercept. The bar (¯) accent denotes the average of the whole pixel array. Coefficients α and β are assigned to bring each pixel's response in line with the average response:

$$Y_{ij}^{NUC} = \alpha_{ij} Y_{ij} + \beta_{ij} \quad 3\text{-}7$$

$$Y_{ij}^{NUC} = \bar{Y} \quad 3\text{-}8$$

Y_{ij}^{NUC} is the pixel reading after the NUC has been applied, equal to the average of the pixel array. After exposing the camera to two temperature sources (T_1 and T_2), it can be shown that α can be expressed as

$$\alpha_{ij} = \frac{\bar{Y}(T_1) - \bar{Y}(T_2)}{Y_{ij}(T_1) - Y_{ij}(T_2)} \quad 3\text{-}9$$

After calculating α, β too can be easily calculated.

$$\beta_{ij} = \bar{Y}(T_n) - \alpha_{ij} Y_{ij}(T_n) \quad 3\text{-}10$$

where T_n is one of the two temperatures.

The NUC could be wavelength independent in theory, but will certainly not be if the intensity distribution shifts from one wavelength to the other. Additionally, if each pixel has a different response to different wavelengths also, it would not be wavelength independent.

The NUC can be done with two uniform sources, or the normal illumination source used in the system can be used (on and off). The fact that the source used often does not give a uniform intensity distribution is not an issue as long the NUC is not used in a scenario where the source has changed or moved. Additionally, as the NUC process assumes a linear response from the pixels to radiation, a NUC will not have the desired effect if the detector's pixels have a nonlinear response to radiation, as can sometimes be seen when using the QCL source.

An NUC does not have any effect on a 'Transmission' or 'Absorbance' image as defined in section 3.1.1, as the NUC gets completely ratioed out. When making a 'Reflection' image, the presence of the α coefficient is still in effect, so in this case a correctly performed NUC is important. The NUC is also useful for alignment of the system when intensities and contrast are low, as it can show features that were previously undetectable.

3.3.1.2 Increasing pixel value of IR camera

During normal operation of the IR cameras, it has been observed that the pixel value usually increases over time. This happens during the daytime regardless of whether there is a MIR light source purposefully incident on the detector or not.

A key point in the operation of these systems is taking the Environmental Backgrounds (E_S, E_B) to get only the signal that is transmitted/reflected. This is because the resting pixel value of the cameras is never zero. The largest contribution to this is most likely due to the fact that everything in a room is roughly at room temperature, and so will radiate at least some infrared radiation, in accordance with Planck's law.

When in operation, there are two reasons that the pixel value would change and increase over time. The first is the natural heating of a room when the sun rises and sets during a day, added to the fact that the presence of people in a room also causes it to warm up. The effect of the room heating up on the resting pixel value is demonstrated in Figure 3-12. As people leave in the evening to go home, the resting pixel value drops sharply and continues to drop, increasing once more in the morning.

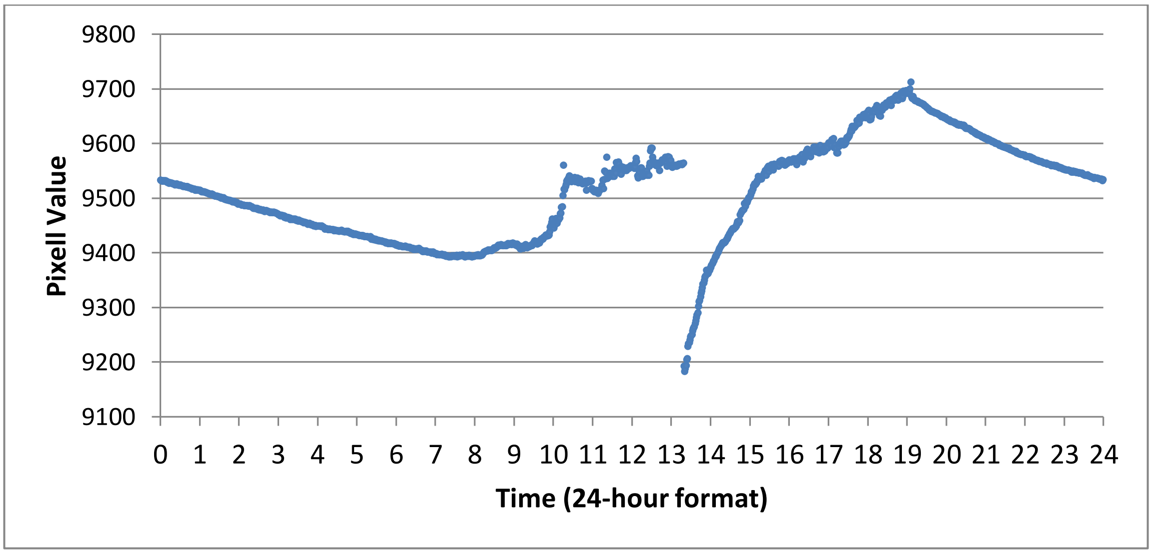


Figure 3-12. Average resting pixel of the CEDIP IR camera throughout the day (early November) with no intentional IR source incident on the detector. The acquisition started at just after 1pm, after which the pixel value started increasing immediately. At approximately 7pm, there were no people present in the room, and it can be seen that the pixel value starts decreasing until the morning. This is likely due to two factors. Firstly, the body heat of people will no longer be contributing to the pixel value, and after the sun sets, the room cools along with the environment outside. When the sun rises once more and people enter the room in the morning, the pixel value increases once more. By 1pm the pixel value is greater than it was the day before. This may be because of the CEDIP enclosure itself heating up as it runs, contributing some additional thermal radiation.

The second effect has been observed only in the CEDIP and occurs when a MIR light source is directed at the camera. In this case, the pixel value rises and falls at a sharper rate than with no incident radiation in a timescale of the order of minutes, as shown in Figure 3-13. This is thought to be due to the IR source heating up elements in the camera itself, and temperatures changing as the cooling system in the camera works to maintain the sensor at its set temperature of 77K.

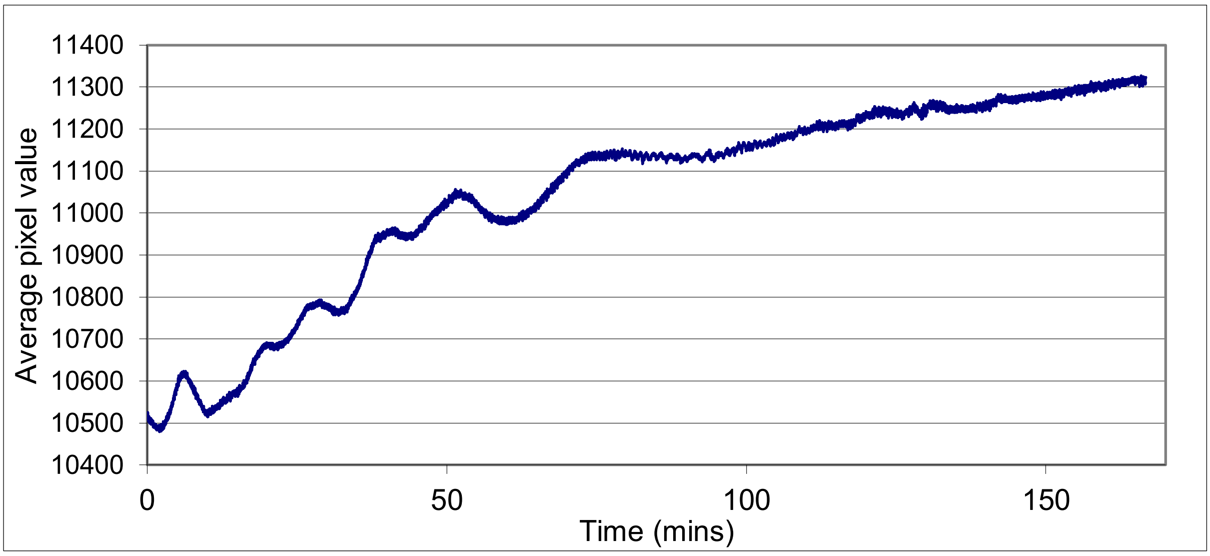


Figure 3-13. Change in average pixel value as a QCL source (time-averaged power of 4mW, 272ns pulses) is consistently incident on the CEDIP IR camera. The pixel value increases, as is expected, but at a more rapid rate than if there were no MIR source incident. As the internal cooling system works to cool the detector (to 77K) however, the pixel value decreases for periods of time. After some time the rate of increasing pixel value stabilises, probably due to the cooling mechanism reaching a stable point.

It is mainly due to this always-changing pixel value that the interleaving approach is taken to reduce noise (see section 3.3.2.3). The interleaving reduces the time between the environmental background being taken and the sample image being taken, bringing them closer to a scenario where they have the same resting pixel value.

The overall pixel value of the camera changes over time, but the distribution over the camera array also changes due to the presence of warm or hot objects. If a person is in a position where radiation they emit can be incident on the camera array and they then move, the distribution on the camera array will change on a relatively short time scale. Any motorised optical components can also cause a problem as the motors that drive them heat up with time, acting as significantly bright MIR sources that does not necessarily have a steady output. To remove these problems, it is important to employ some sort of shielding to block the radiation from these sources reaching the camera.

Narrow bandpass filters can also be used to block most of this radiation, as described in section 3.3.3.3, though this will only be relevant when working at certain wavelengths with the laser-based system.

3.3.2 *Digistain*-specific considerations

The *digistain* setup has some unique aspects that can be optimised to maximise the SNR of the system. The illumination is a large thermal source, quite different from the collimated laser source that is provided by a QCL. The way the light is collected and concentrated on the sample affects how the background correction can be done. The mechanical solution for moving the sample in and out also introduces a potential issue that can affect the results. The interleaving routine previously touched upon is described in detail in this section.

3.3.2.1 Optimising sample illumination

The thermal source is given some directionality by a parabolic reflector, as described in section 3.2.1.1. However, condenser optics are still required to collect all the light, as there is some diverging light that is not affected by the parabolic reflector, and the light that is collimated has a beam diameter larger than the field of view at the typical 3x magnification.

This can be done using a single lens. Note that although this setup appears to deviate from the Beer-Lambert Law, as outlined in section 2.2.1.1 (as the light will not be travelling parallel through the sample), the error is removed during the ratioing procedure between the phosphodiester and amide wavelengths, as outlined in section 3.1.3.

A purely empirical approach is taken towards the positioning of this condenser lens. When the distribution of the light at the image plane is no longer an issue, the amount of light reaching the detector needs to be maximised. The intensity distribution at the image plane is not uniform, with the intensity in the centre of the image roughly twice that in the very corners. This is turn will mean that the SNR is not uniform across the image.

3.3.2.1.1 Effect of substrate thickness on illumination

Another consideration with regards to the illumination is the thickness of the substrate. There is an issue with the fact that the spectrum of the substrate will be mixed in with the response of the sample. This is not a huge problem with BaF_2 substrates in the 6-10µm region, where BaF_2 is quite spectrally flat with a transmission of ~95% for a few millimetres of BaF_2. As long as the same thickness of substrate is used, different samples can still be compared. If the light is passing with parallel rays through the sample, the absorption from the substrate (added to every pixel equally) is the only thing to be considered. However, if the light is not travelling parallel through the sample, another problem can arise.

Adding a window into the optical path shifts the spatial focal point of a lens. When the sample is removed (along with its substrate), the light will be shifted slightly along the optical axis, changing the distribution of light at the camera. As a result, the background image in the denominator of equation 3-1 is not seeing the same distribution of light as the sample image, modifying the results. This can be solved by taking the background image with an identical substrate in place instead of just empty space. Care must be taken to make sure the two substrates, with and without the sample, are orientated at the same angle. A substrate thickness of 2mm was used (see section 4.3.1.1). Ideally the substrate should be as thin as possible whilst still be structurally sound to minimise any effect caused by misalignment between the sample and background substrates.

The focus of the image will not change in this setup, as there are no optical components changed between the sample and the detector.

3.3.2.2 Restricting sample movement

A key aspect to making the *digistain* results reliable is pixel registration – between wavelengths, the same area of the sample is being imaged to the same pixel every time, allowing us to combine images from different wavelengths. This is the main reason why a germanium objective lens is used, maintaining focus and magnification thanks to germanium's relatively low dispersion.

There is a practical consideration during the data acquisition however, in that if the sample were to shift during the acquisition, pixel registration would no longer be maintained. The sample is moved in and out of the field of view numerous times during a data acquisition, which can cause the sample position to shift.

This is due to two factors – firstly, the centrifugal force and the jarring action as the movement comes to a halt can move the x- and y-position of the microscope slide holder. Removing the lubrication in the gears of the x-y mechanism was tried to make it more resistant to movement, but this had an insufficient effect.

The second aspect of the unwanted movement is the movement of the microscope slide adapter. Primarily due to the jarring action mentioned before, the slide adapter can move very slightly as it is held within the modified microscope stage during the acquisition.

The solution to this movement problem is to use sticky tape to secure all sources of possible movement. This adds a small amount of time for the user to secure a new sample before an acquisition. This issue could be removed in a future iteration of the *digistain* setup.

3.3.2.3 Interleaving images to increase SNR

As was discussed in section 3.3.1.2, the pixel value of the camera changes over time. The longer the time between taking each of the four images that make up the final image (for a transmission image, see section 3.1.1), the more unreliable the final values will be. The longest delay in acquiring the data takes place when the sample is moved in and out of the field of view, and when the light source is turned on or off. These individually take 1.5 seconds at most each time, and are accepted as delays that cannot be removed. The other aspect that causes a delay between the four images being taken is the fact that the 4 separate images that make up a transmission image are each an average of many images. Since $SNR \propto \sqrt{N}$, N being the number of images used to make the average, we want to use a high number of averages to make up our image. However, the more averages we take, the bigger the time delay becomes. To counter this problem for large averaging numbers, instead of taking, say S (sample image) all at once, we take it in chunks. This is called an "interleaving" approach. The process is demonstrated in Figure 3-14.

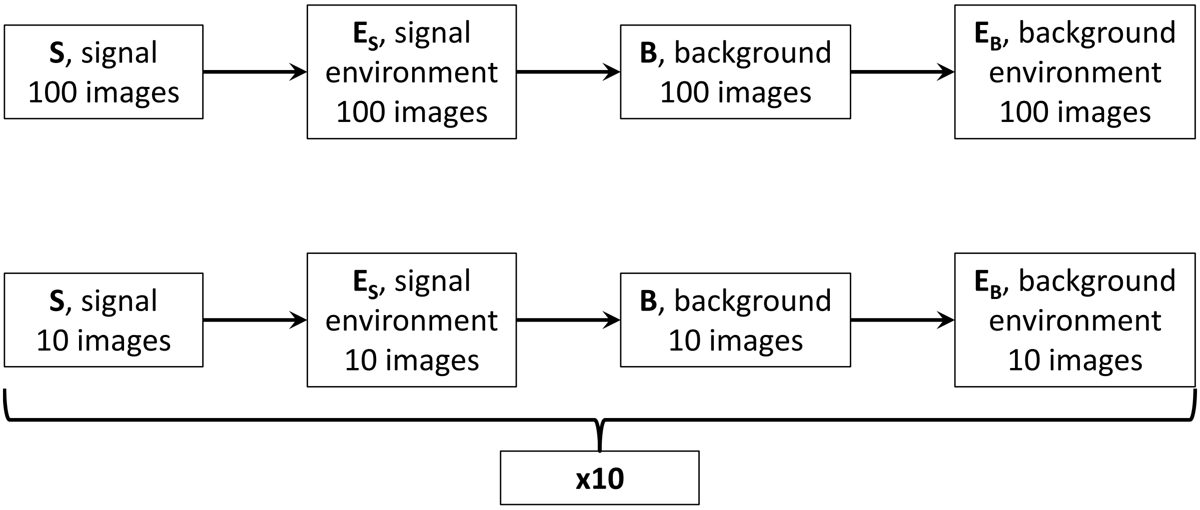


Figure 3-14. Image interleaving routine algorithm. The top process shows what happens during a 'normal' data acquisition to create a transmission image, with S, E_S, B and E_B being the four images that are used to create a transmission image, as described in section 3.1.1. The four images are taken one after the other in succession. The bottom process is the interleaving approach – a lower number of averages for each of the four images is taken, to minimise the time taken between each image, but the cycle is repeated to make sure the total number of averages is the same.

The effect of not using an interleaving approach can be seen in Figure 3-15. The measured SNR across the image does not increase with the square root of the number of averages, N, as is expected.

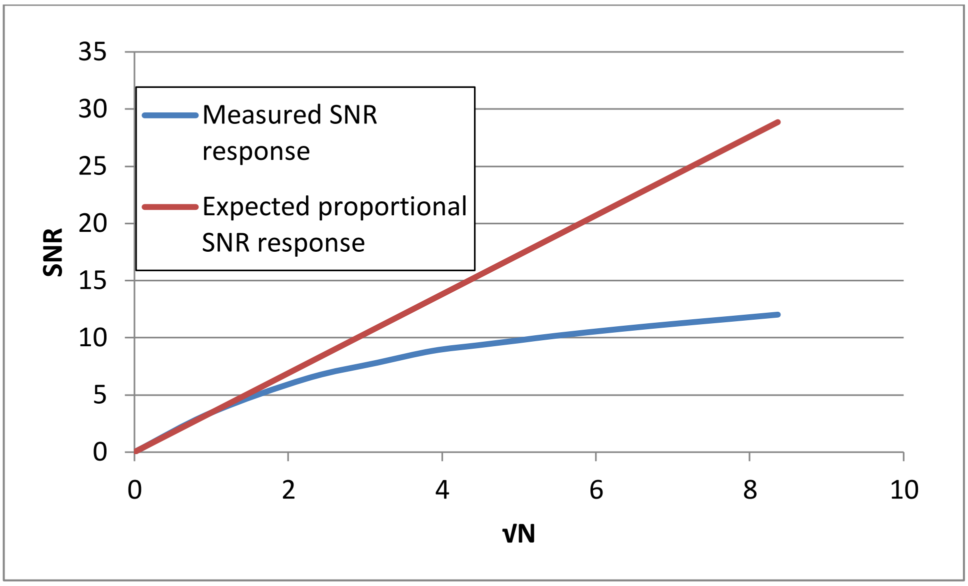


Figure 3-15. Graph showing the difference in SNR between the expected SNR and the measured SNR when no interleaving process is used. According to theory, the SNR should increase proportionally with the square root of the number of averages, N. The straight line shown is a continuation of the line for N=1. The measured behaviour of the SNR starts to plateau with increasing N, getting further from the theoretical prediction. The lower the number of averages, the higher the SNR per image will be. If we want to scale the SNR more effectively, we can use the interleaving approach. When interleaving, the 'chunks' of images do scale with the square root of N. The SNR can be maximised with respect to N by setting N=1. However, if we wish to maximise SNR against total acquisition time, a larger N should be used.

Using the interleaving technique, we can increase the SNR, since the interleaved chunks do increase in the expected manner. i.e. if the SNR for 10 averages taken normally is x, the SNR for 10 averages repeated 4 times will be $2x$.

From Figure 3-15, it may seem that it is best to choose N=1 and interleave to scale the SNR to greater values. However, as each chunk of the interleaving process adds additional time to the acquisition, to move samples and shutters, to maximise SNR with respect to time, a value N≠1 should be used. The best values for the number of averages and number of interleaves to optimise the SNR for a set acquisition time are determined by the *digistain* program by referring to a saved SNR curve and accounting for the time each step of the acquisition takes.

3.3.3 Laser-based system considerations

As with the *digistain* setup, there are some considerations particular to the laser-based system. These stem from the light source used (a pulsed coherent QCL source), and the detector (a photodetector comprised of semiconductor elements). The solutions to get around the problems the source and detector introduce have an adverse effect on the SNR.

3.3.3.1 Removing coherence

The outputs from the laser-based sources are collimated, so that focusing optics may not be necessary to concentrate the illumination into the field of view. Doing a 100% T SNR test (the SNR across the image) like this, an SNR of 20 per image can be obtained – an SNR in the thousands could be achieved from just a few minutes of data acquisition. The SNR can vary significantly however, it is thought due to interference fringes created from parallel surfaces in optics within the camera, which move over time due to small (<1nm/<0.2cm^{-1}) fluctuations in the output wavelength. The fringes are due to the coherence of the source. However, the more pressing issue with the coherence of the source is the effect it has on image quality. Due to coherent scattering and 'ringing' from the edges of the imaged object (see Figure 3-16), the data at each pixel on the camera ceases to be quantitatively reliable. The spatial coherence of the source needs to be destroyed to do absorption imaging.

The diffuser implementation (see section 3.2.2.5) is shown to destroy the spatial coherence sufficiently to remove evidence of coherent scattering, but it is the diffuser that now limits the SNR that can be obtained (see Figure 3-17). As the intensity distribution at the camera is now random during each frame of the camera, a large number of images are required to get a uniform illumination, and the SNR increases with $\sqrt{N}$, the square root of the number of images taken. The noise introduced by the diffuser is usually greater than the noise from the change in pixel value over time, meaning an interleaving approach outlined in section 3.3.2.3, is not required for averages over at least 5000 images. The SNR across the image obtained with the diffuser is now typically slightly less than 1 per image – the introduction of the diffuser has reduced the SNR by about 20x from what it was previously. However, in practice the difference is unlikely to be so large, as straightforward illumination with no diffuser would usually be required to be attenuated to remove nonlinearities (see section 3.3.3.2).

The diffuser brings in noise, but it also reduces the amount of light reaching the detector, through absorption and reflection and scattering losses at the diffuser surface. It is estimated that the inclusion of the diffuser cuts the signal throughput to 58% of what is possible without the diffuser.

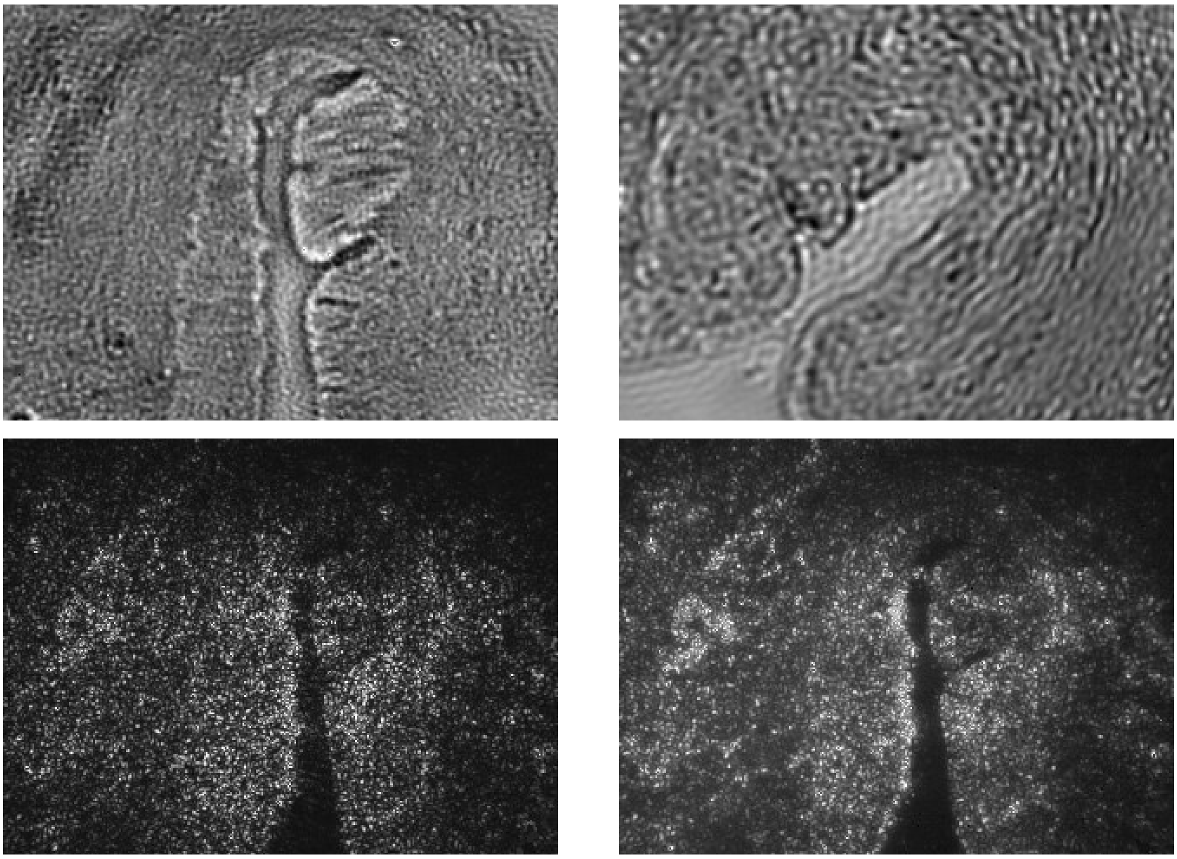


Figure 3-16. Effect of coherent scattering on image quality. The top left image shows a region of an oesophagus biopsy at ~3x magnification. There is no tissue present in the vertical column in the centre of the image, yet there appears to be a texture in that region – this is due to the coherent scattering of the light. The 'graininess' across the image can be removed with the introduction of a diffuser, although it should be noted that the coherence, although not noticeable, may still be there. The image in the top right is an image of the same sample, but at higher (~10x) magnification, with a diffuser present. The spatial coherence was thought to be removed with the at 3x magnification. However, the coherent 'ringing' effect can be seen near the sample edges. Bottom left is shown a diffuse reflection image with no diffuser, and bottom right the same image taken when a diffuser is in place to shown the improvement in coherent speckle reduction.

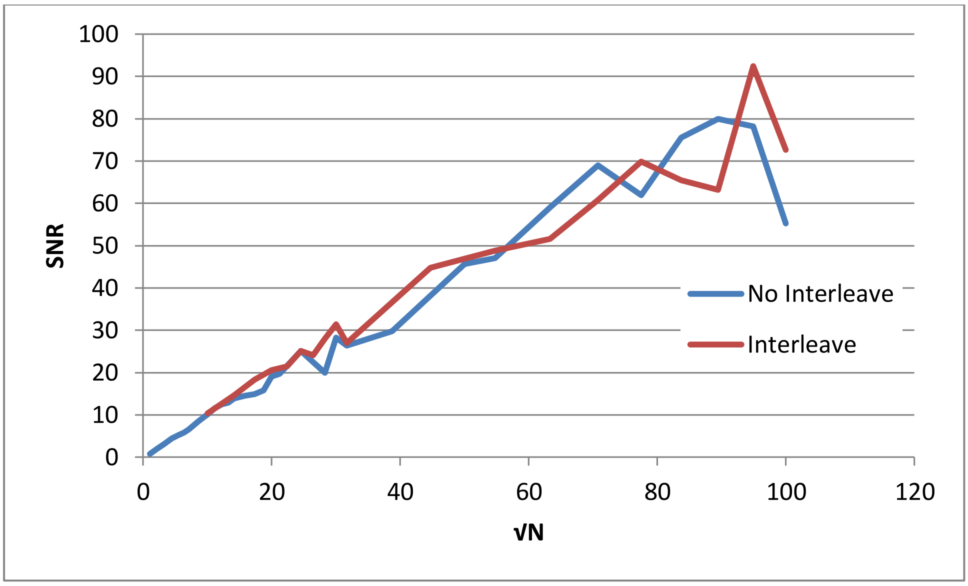


Figure 3-17. SNR scaling with the square root on the number of averages, N, with a diffuser in place. The SNR scales approximately linearly with $\sqrt{N}$*, indicating that the problem that led to the need for interleaving, a changing thermal background and distribution on the imaging array, is no longer the main source of noise. The noise from the diffuser comes from the effect that for each position of the diffuser, there are some parts of the imaging array that do not see any radiation. Many averages are required to get a smooth intensity distribution. This type of noise now scales linearly with* $\sqrt{N}$*.*

The diffuser is also required when doing reflection imaging to reduce the appearance of a speckle pattern. The speckle pattern creates a large range of pixel values, which can hide finer details present. When using the diffuser in a reflection mode, there still appears to be some areas which are speckled, but it is thought that this is due to the surface roughness of the sample being imaged. This is supported by the observation that when the sample is moved to another position the speckle-like pattern moves with it, as shown in Figure 3-18.

Figure 3-18 shows a situation contrary to what we would expect to see if this was actually coherent speckle, which should change as the source of speckle moves relative to the observer [115]. If not from interference, the speckle-like effect must be originating from features on or in the sample. The transmission images do not show up any such features within the sample, meaning the bright spots must be originating from the sample surface.

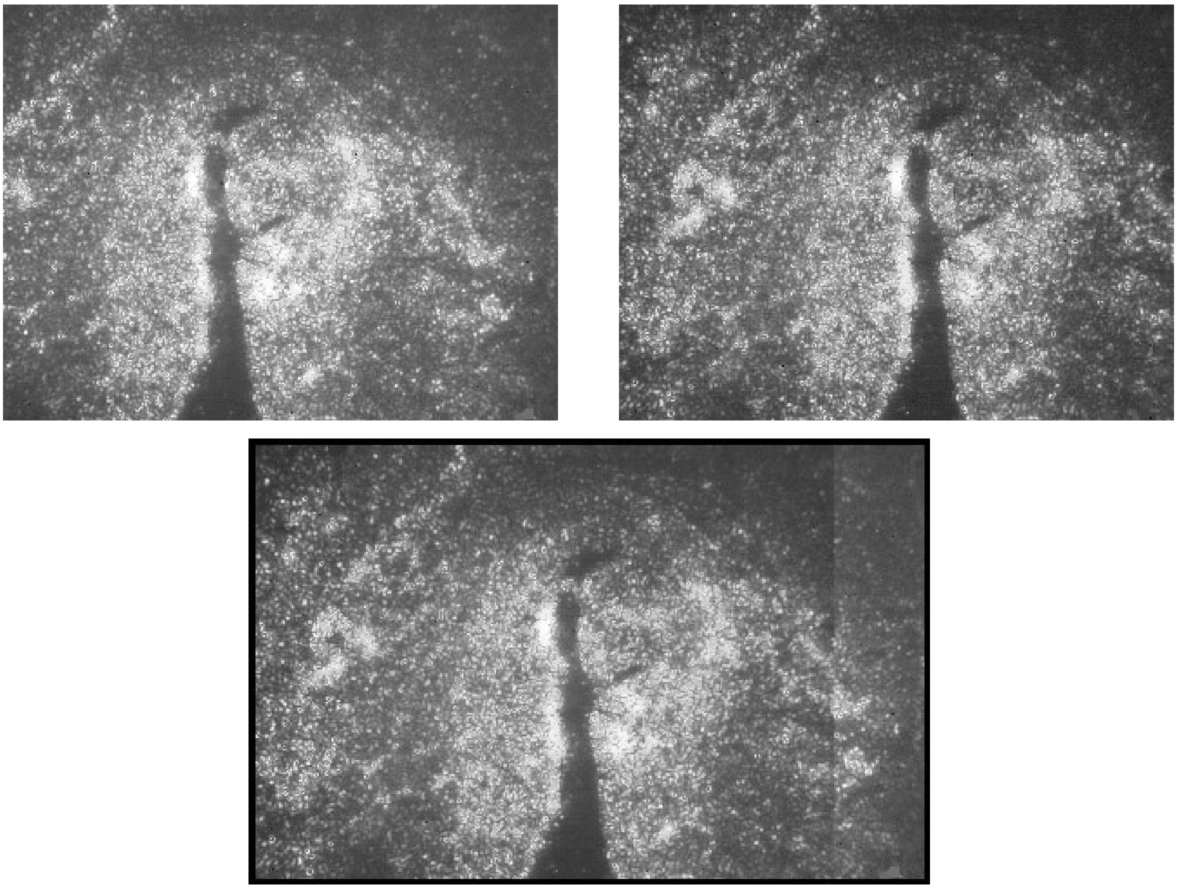


Figure 3-18. Reflection image showing the movement of the speckle-like pattern with the sample position (top left, top right), and the combination of those images. The bright spots are coincident on the same locations on the sample. The images were taken using the reflection imaging setup as shown in Figure 3-7. Judging by eye, the speckle-like pattern appears to move with the sample. If the pattern was from coherent speckle, the pattern should shift, such that when the images are combined, the speckle effect would be reduced. The lower image shows the images combined, but with seemingly no improvement in the speckle-like pattern within the overlapping region. This indicates that the pattern does not originate from coherent scattering, instead being a result of the surface roughness.

3.3.3.2 Nonlinear response of CEDIP camera to pulsed light from QCL

To acquire quantitatively reliable information, the response of the camera to the incident light must be linear. If it is in any way nonlinear, the transmission and absorption data will not tell us the relative concentrations of absorbing molecules present. It is at higher powers, supplied by the QCL, that the nonlinear response of the CEDIP camera is observed. A typical curve with fixed, measured attenuating filters is shown in Figure 3-19 to demonstrate the nonlinear response with increasing power. The camera appears to be saturating at higher intensities, but it is not due to intensity alone, as it is not the pixel value of the detector that is being saturated. The nonlinear response occurs because we are using a pulsed source. Although the time-averaged power might be relatively low, the instantaneous intensity during a pulse will be much greater.

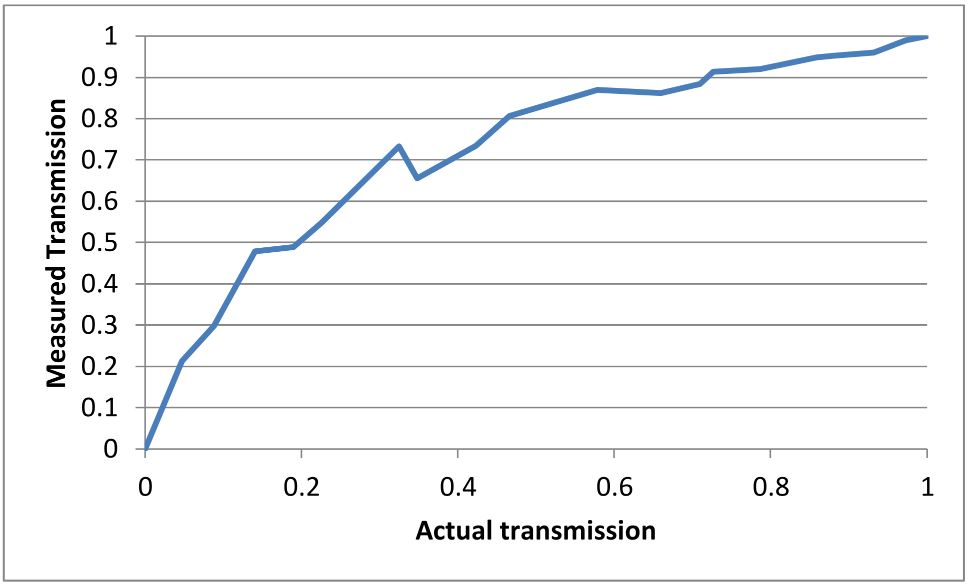


Figure 3-19. Typical nonlinear response curve of the CEDIP camera to the pulsed QCL source, an average response of the imaging array with 256ns laser pulses at 1250cm^{-1} wavenumber incident. The radiation was increased using attenuating filters combined to give different transmissions. The known transmission of the filters is plotted on the x-axis, the transmission as measured by the CEDIP camera is plotted on the y-axis. The curve shown is not smooth, but the general shape will be accurate.

The degree of nonlinearity will vary across wavelengths due to the output intensity variation and the detector response. This means that the linearity will need to be checked for every wavelength in use, and will also have to be rechecked if the optical setup changes.

3.3.3.2.1 Moss-Burstein effect

It is thought that the nonlinear response is due to the Moss-Burstein effect, whereby photons within a pulse can pass undetected if the recombination time of the promoted electrons in the semiconductor pixel element is of the same order as the pulse duration [62]. If the leading part of the pulse excites all the available electrons from the valence band, the trailing part of the pulse can pass without exciting any electrons if it can pass before electrons decay back to the valence band (see Figure 3-20). In this way, part of a pulse can pass undetected. Due to curved energy bands, the effect should be more noticeable at longer wavelengths, as there will be less available empty states in the conduction band [62].

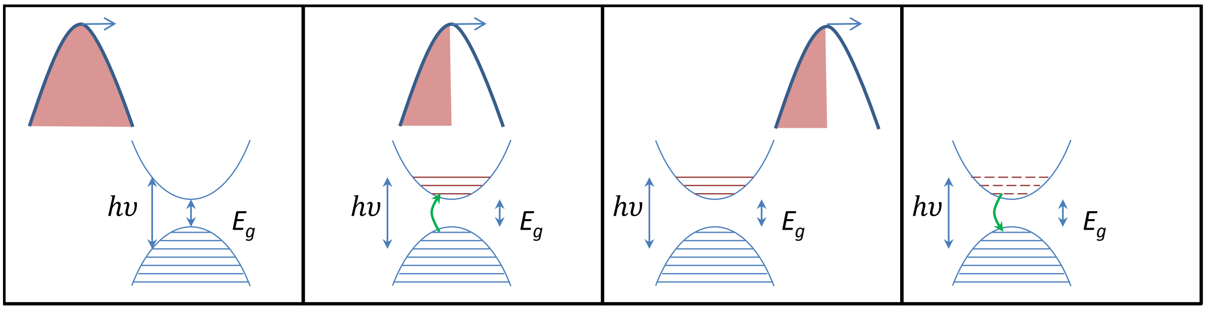


Figure 3-20. The Moss-Burstein effect, that can lead to nonlinearities in the detector, as the entire QCL pulse is not being detected. In the leftmost panel, the red pulse of light approaches the detector, and the band structure diagram shows electrons in the valence band only. The energy of the photons, $\boldsymbol{h\nu}$*, is also shown. In the second panel, as photons are absorbed, promoting electrons to the conduction band, the available energy states in the conduction band are all filled when only the first half of the pulse has been detected. If the energy states in the conduction band ore not made available before the pulse passes, the second half of the pulse passes undetected, as in the third panel. The fourth panel shows the energy states in the conduction band becoming available, either by decay back to the valence band or the promoted carriers being extracted, long after the pulse has passed. The effect is stronger at longer wavelengths which have lower energy photons.*

3.3.3.2.2 Nonlinear calibration

If we allow the response to remain nonlinear, we can create a calibration curve to convert the signal to data with a linear response. As pixels will likely receive a different amount of light, a calibration curve for each pixel is required. An equation that can describe the saturation mechanism is not known, but looking at a typical nonlinear curve, the response appears to be roughly logarithmic (see Figure 3-19). It was found that the best equation for a fit was an exponential of the form

$$y = ae^{bx} + c \qquad \text{3-11}$$

where x is the measured transmission and y is the actual transmission. Applying the conditions $y(0) = 0$ and $y(1) = 1$, the fitting equation becomes

$$y = \frac{e^{bx} - 1}{e^{b} - 1} \qquad \text{3-12}$$

This equation can also fit a linear response (with solution $b \to 0$).

This fit can deal with slight nonlinearity, but does not deal well with strong nonlinearities, when near-saturation occurs before reaching 100% (see Figure 3-21).

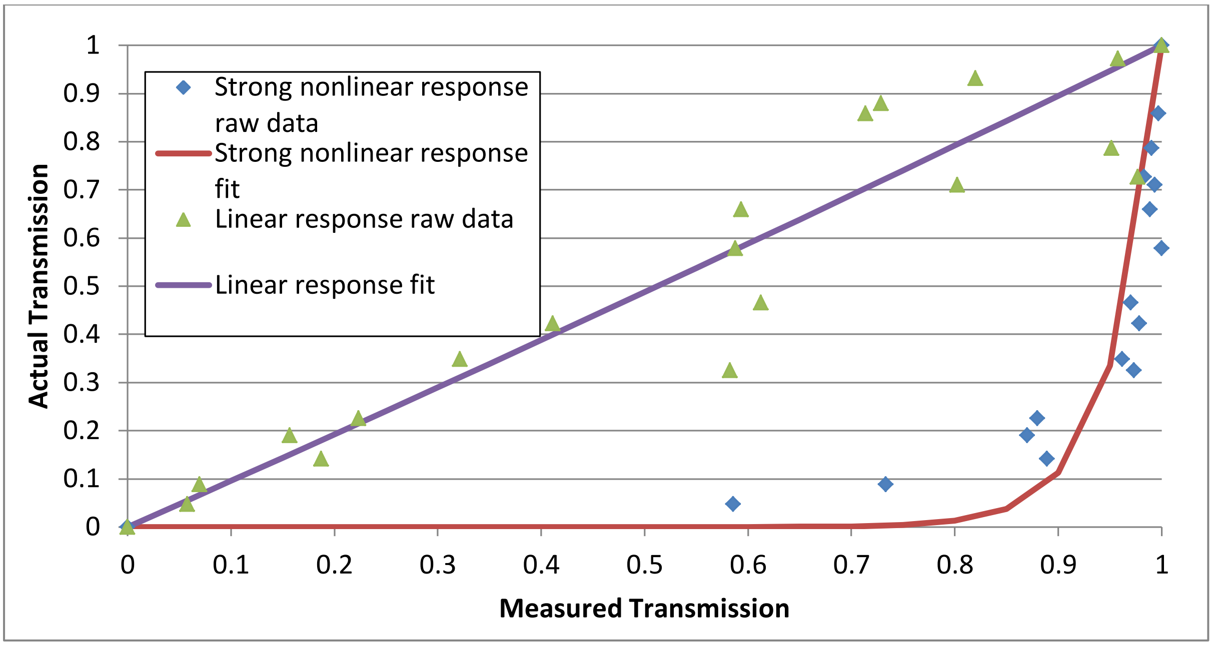


Figure 3-21. Demonstration of how the fitting equation 3-12 deals with the two extremes of nonlinearity. The green and purple points show the data and fit respectively for what appears to be a linear response (a pixel on the edge of the imaging array, under the same conditions as described in Figure 3-19). The blue and red points show the data and fit respectively for what appears to be a strongly nonlinear response (a pixel in the centre of the imaging array). The linear fit appears to do well – the fit is almost linear. For the strongly nonlinear data set, however, it is clear that fit is missing the mark considerably. The function we have chosen to fit with is simply not the correct form to be able to represent the data set.

As the fit is applied to each pixel individually, and not as an average of the whole, the fit must be able to deal with the most nonlinearly-responding pixels, usually in the centre of the image due to the Gaussian distribution of the illuminating light. Figure 3-22 shows the fit of the nonlinear response of the average of the array, and the average of the fits for each pixel. The most nonlinear pixel responses cause a poorer fit to be achieved. As a result, the light will sometimes have to be attenuated to a less nonlinear regime before nonlinear fitting takes place.

Even if the nonlinear curve can be fitted well, there is the problem of how errors propagate in a nonlinear function. At higher transmissions, where the response is ‘steeper’, a small error in the measured transmission can translate into a much bigger error in the actual transmission acquired after translation with the fitting equation (see Figure 3-23). For a strongly nonlinear response, error in the transmission value from noise originating from the detector could be amplified after being acted on by a nonlinear fitting function.

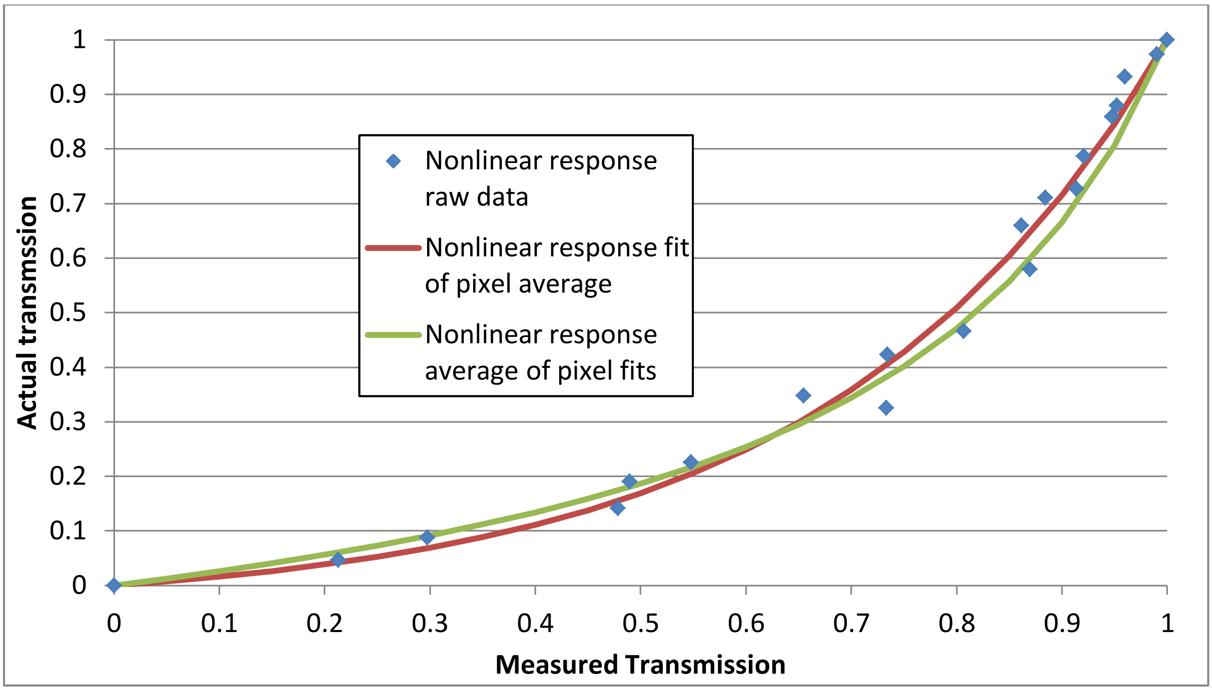


Figure 3-22. A fit (red) on the average of the pixel array (blue), and the average of the fits on all individual pixels (green). If the fitting function was able to deal with all levels of nonlinearity, the two fits should be identical. The average of the fits (green) deviates from the fit of the averages (red) due to the fits' inability to cope with the strongly nonlinear pixels that are exposed to more light.

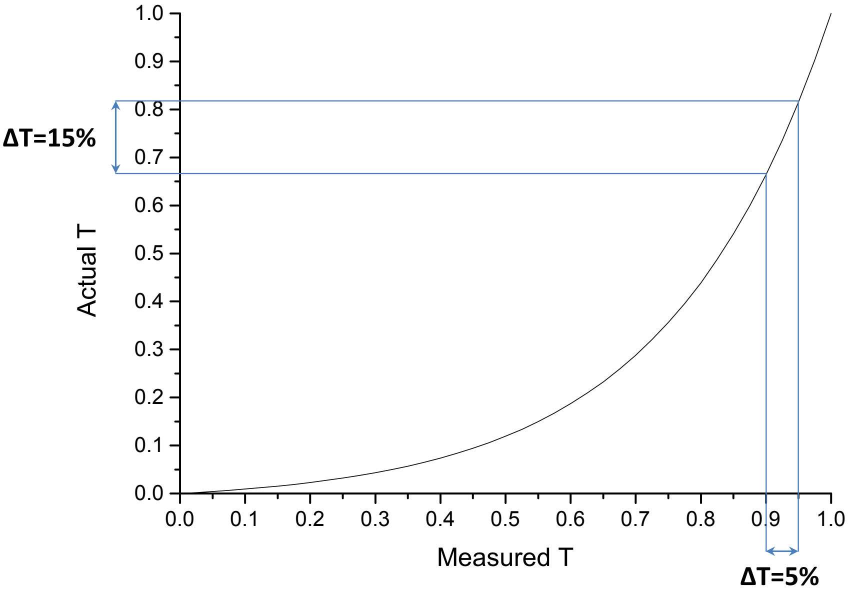


Figure 3-23. Propagation of errors for a nonlinear response correction. At high measured transmissions, a small variation in the measured transmission, possibly due to noise in the detector or readout electronics, can translate into a large variation in the determined transmission after the fitted function has been applied. Conversely, at low measured transmissions, a variation in the measured transmission will translate into a smaller variation in the determined transmission.

The bigger problem still, with fitting a function, is a practical one – the data given to the fitting routines are not always monotonic. It is suspected that the reason for this is that the curves are acquired by inserting various combinations of neutral density filters (with transmission measured in an FTIR) into the beam path. As the filters are roughly 1mm or 2mm thick, the beam can shift noticeably if multiple filters are combined, with the filters being at a slight angle relative to each other. Although a rough fit can be obtained, the error in this fit will be larger the more scattered the data to be fitted is.

Due to the difficulties of obtaining a reliable fit, and the time-consuming nature, it is much more desirable to work in a linear regime from the outset.

3.3.3.2.3 Pulse length control

To reach a linear regime, the beam could simply be attenuated enough such that all pixels behave linearly. This would undoubtedly work, but this may involve throwing away up to 95% of all the light. Taking into account that the pixels not in the centre of the image may be behaving linearly already, we would be throwing away a lot of the light unnecessarily for the sake of the central pixels that receive the most light.

The nonlinearity comes from the fact that we are using a pulsed source with intense pulses. If we could move towards a more CW mode of operation whilst maintaining the same time-averaged power, the nonlinearity problem may be able to be sidestepped altogether.

The pulse length and the repetition rate can be controlled, within the confines of the 5% duty cycle. It was found that when operating at the shortest pulse length, 32ns, the peak power dropped noticeably as viewed on an oscilloscope from a single channel detector (see Figure 3-24). When the 32ns pulse was used with the CEDIP camera, the nonlinear response may no longer be observed. Although the peak power of the pulse has decreased, the whole pulse is being detected by the camera, meaning that through control of the repetition rate, the original pixel value can be returned to or surpassed (since there is no saturation occurring – the actual power incident on the detector is almost certainly receiving less optical power, it is only the pixel value from the camera that has increased). Note that with 32ns pulsewidths, pulses can still return a nonlinear response under certain conditions, such as when the light is focused into a small area, so the nonlinearity should still be checked.

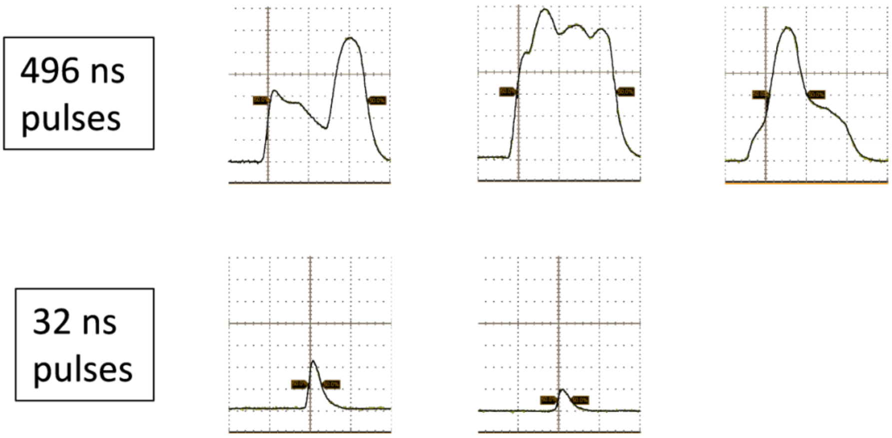


Figure 3-24. Difference in pulse intensity between the maximum pulsewidth (496ns) and the narrowest pulsewidth (32ns). The pulses are shown on the same voltage and time scale. The variation in pulse shape that can be seen can be quite dramatic at 496ns pulse length. The shape may be influenced by nonlinearity though – the detector used was a photodetector, so could have been suffering from the Moss-Burstein effect also. The 32ns pulses are considerably smaller in power than the 496ns pulses, and are likely relatively even smaller than displayed, due to the possible nonlinear response of the detector.

3.3.3.3 Use of *digistain* filters in the QCL setup

The intensities being detected in reflection mode are much lower than in transmission mode, and so the effect of the thermal background is more prominent. The noise from thermal fluctuations can be reduced by using the bandpass filters previously being used in the *digistain* setup (see section 3.2.1.2). If the wavelengths being used match the bandpass filter's central wavelength, most of the radiation from every other thermal source can be excluded. The filters do not have 100% transmission at their central wavelength, so the signal will be slightly reduced, but the noise will be reduced by an even greater amount. Another benefit is that it reduces the resting pixel value, meaning we can use longer integration times. As pointed out in section 3.2.2.1, this allows us to increase the time we acquire data for without increasing acquisition time, improving the SNR further.

3.4 Software design

After the data collection routine and, in some cases, the creation of a ratio image, there are a limited number of post-acquisition processes that the data goes through before being presented to the user. All these processes are primarily to improve the way in which the data is displayed to the user in the form of an image, and most techniques have no risk of changing the underlying pixel values. Doing this is beneficial to the clinicians or other users who may be using a future iteration of the *digistain* system, allowing them to easily pick out the areas of interest to focus on.

The most intrusive process is a noise filter that acts on a small percentage of pixels that are the highest or lowest within the image. In the past, more complex and aggressive routines were used to reduce noise. However, as the system improved and the SNR of the data increased, the need for these aggressive routines was diminished.

The image is usually displayed with a standard 'rainbow' colour gradient that goes as a gradient from blue to green to red, from low-valued to high-valued pixels. Usually, the colour gradient is mapped from the lowest to the highest pixel values present in the image. This means that a single very low-valued pixel could cause the rest of the image to appear uniformly red, even if there would normally be variation seen within the image. It is possible though to control how the colour gradient is mapped, negating the effect of outlier pixels. In cases when the range of pixel values mapped is fixed, the noise filter will not be necessary at all.

3.4.1 Image stitching

If the area to be imaged is larger than the field of view, the images must be stitched together to create a single larger image. If the sample translation stage were indexed (and ideally motorised and computer-controlled), this would be a simple task of putting bordering images together. That is not the case, however. Automatic stitching in the software may be possible, but this was neither practical nor time-permissive within LabVIEW.

Instead, the solution that was implemented was a manual stitching program – the images are loaded onto a blank canvas and the user can drag and drop the images to line up with one another. There needs to be some overlap between images to give the user a common reference point, but this need be only a few lines of pixels – in the samples looked at, there is enough distinct features to identify common regions. Once the positioning of the images is determined, the images are combined in a weighted average (proportional to the distance from the centre of the image). This weighted average is implemented to account for the lower SNR and possible defocusing at the edges of the image.

3.4.2 Noise filters

There are a number of software noise filters that could be used to tidy up the final image [116]. The one implemented is relatively simple to avoid modifying the data without due cause, known as a bad pixel replacement (BPR) [26]. These pixels may be 'bad' due to dead or 'on' pixels in the camera, or unusually high as an artefact of dividing by near-zero.

This filter aims to modify those pixels that lie outside of a certain number of standard deviations of the image as a whole [26]. The number of standard deviations is set by the user, with 4 or 5 standard deviations found to remove the outliers that lower the contrast of the image without modifying too many pixels as to obscure any details. Pixels outside the cut-off point are assigned the median value of their nearest neighbours, using the pixels within a square patch centred on the pixel in question. This process is then repeated a number of times, to remove any small clusters of bad pixels.

If the image had a normal (Gaussian) distribution, the BPR would modify <0.0001% of pixels (with the number of standard deviations set to 5). However, when imaging tissue samples, the distribution of pixel values should not be Gaussian. As a result of this, and the fact that more than just one image is undergoing the BPR procedure multiple times, a typical value for the number of pixels modified by the BPR is actually about 4%.

3.4.3 Removing empty space of sample

In areas where there is no tissue present, the absorption will be near to zero. When it comes to making a ratio image, the modulus of these pixels can become very large. To mask these pixels and prevent them interfering with the BPR procedure, empty space in the image is set to a certain value based on the amide peak image. In the case of tissue, anywhere there is tissue will have absorption at the amide wavelengths. The areas that have no tissue present have to be ignored when doing the BPR, as they will increase the standard deviation of the whole image. The effect this has on the displayed PA ratio image is shown in Figure 3-25.

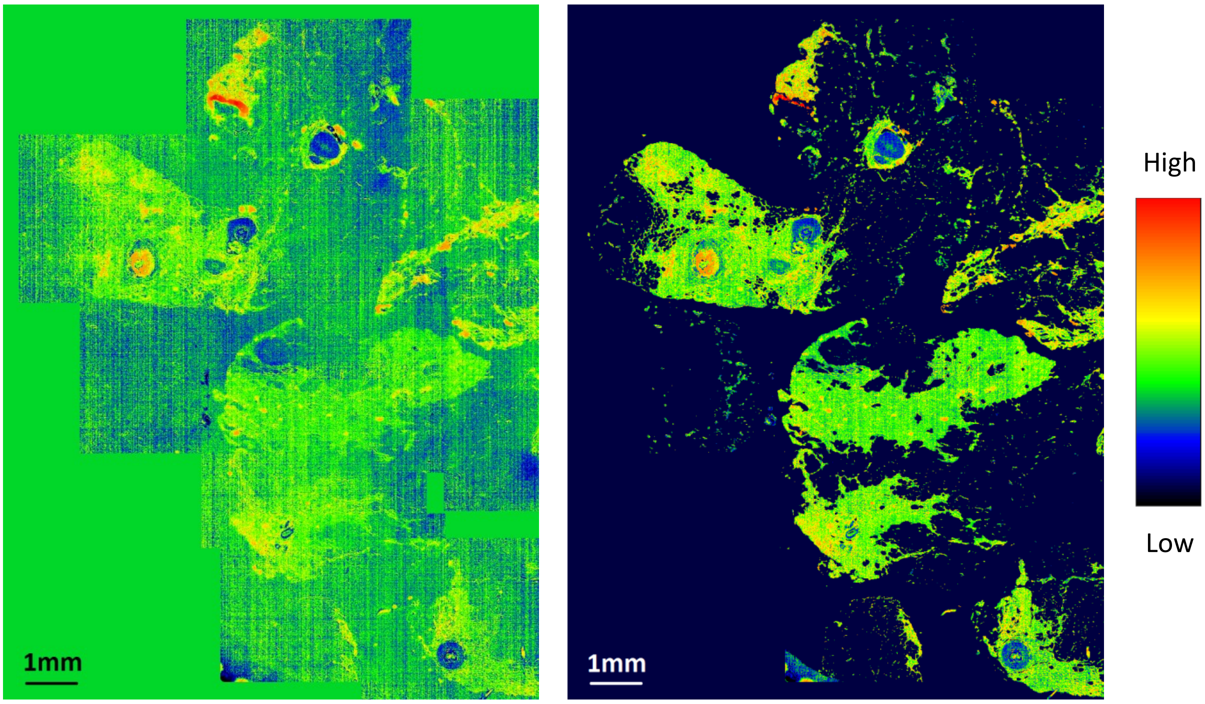


Figure 3-25. Effect of filtering out areas without tissue present on the PA ratio image. The image on the left has not had the areas without tissue present masked. Although the areas without tissue are not high/low enough to modify the contrast by much, it is difficult to tell where the tissue boundaries are. The large areas of block green on the edge of the image are a result of the image stitching process. The image on the right has had areas of no tissue present masked, using information from the amide Peak absorbance image. The boundaries of the tissue are much more visible.

4 Information on samples

The samples that can be looked at by the spectral imaging systems can be any sample that has spectral features that fall within the range we can probe and detect at (wavelength of ~6-10μm), but the samples of greatest interest and convenience have predominantly been cancer tissue biopsies. Two types of cancers have been looked at – oesophageal cancer (from the throat), and breast cancer. Both types have different subtypes, usually classified by the exact origin of the cancer and the cell types involved. There is terminology that medical professionals use when discussing these cancers, which will be explained in this section.

The particular biopsy samples looked at will be described in sufficient detail such that the MIR spectral imaging results in section 5 and their significance can be understood. The grading procedure for breast cancer will also be discussed in more detail to highlight the factors that contribute to the cancer diagnosis procedure to provide context for the results from the *digistain* (the phosphodiester-to-amide ratio) experiments.

Although there are other samples looked at (e.g. plastic), any required background on the absorption spectrum will be given when appropriate in section 5.

4.1 Glossary of terms

This section contains a glossary all terms that may be necessary to understand a description of the biopsies investigated [117].

Epithelium lines the cavities and surfaces of structures throughout the body, and also forms many glands. Cells in epithelium are densely packed like bricks in a wall. All epithelial cells rest on a basement membrane which acts as a selectively permeable membrane. Epithelium that is a layer one cell thick is known as **simple epithelium**, whilst if it is two or more cells thick, it is called **stratified epithelium. Squamous epithelium** has cells that appear 'flattened'.

Stroma is often called fibrous tissue. Stromal cells are connective tissue cells of any organ.

Malignancy is the tendency of a medical condition, especially tumours, to become progressively worse and to potentially result in death. Malignancy in cancers is characterised by anaplasia, invasiveness and metastasis.

A **tumour** is an abnormal mass of tissue as a result of abnormal growth or division of cells. If the tumour is not benign, it is also called **neoplasm.**

A **benign tumour** is a mass of cells that lacks the ability to invade neighbouring tissue or metastasize. Generally they have a slower growth rate than malignant tumours and the tumour cells are usually more differentiated (i.e. cells have their normal features).

Cellular differentiation is the process by which less specialised cells become a more specialised cell type. It is the differentiation (or anaplasia) which plays a large role in the grading of tumours, based on the resemblance of the tumour to the tissue of origin. Poorly differentiated neoplasms have a high ratio of nucleus to cytoplasm with numerous mitoses (cell divisions).

Necrosis is a form of cell injury that results in the premature death of cells in living tissue by autolysis (the destruction of a cell through the action of its own enzymes). Cells that die due to necrosis do not follow the naturally occurring, programmed, cellular death (apoptosis). It results in a uncontrolled release of products of cell death into the intracellular space, initiating an inflammatory response which prevents phagocytosis (elimination of dead cells).

Anaplasia is a reverse in the differentiation in cells, also known as dedifferentiation. It is a loss of structural and functional differentiation. Anaplastic cells display marked **pleomorphism** (variability in size, shape and staining of cells and/or nuclei). The nuclei are characteristically large. The nuclear-cytoplasmic ratio may approach 1:1 instead of the normal 1:4 or 1:6. Anaplastic cells usually fail to develop in recognisable patterns of orientation to one another. They may grow in sheets, with total loss of communal structures, such as gland formation or stratified squamous architecture.

Dysplasia is the earliest form of pre-cancerous lesion which pathologists can recognise in a biopsy. The risk of low grade dysplasia transforming into high grade is low. Dysplasia consists of an expansion of immature cells, with a corresponding decrease in the number and location of mature cells. This is linked to the differentiation – immature cells may not have acquired their specialised structures. The term is typically used when the cellular abnormality is restricted to the originating tissue (often *in-situ* cases). As it is precancerous, appearing before carcinoma develops, it is not included in the usual grading systems.

Carcinoma is tumour tissue derived from epithelial cells that have transformed, exhibiting abnormal malignant properties.

A **carcinoma in situ** is when the carcinoma is highly localised. An example is when malignant cells may be confined within ducts. The diagnosis of in situ neoplasia is important as such lesions may progress to become invasive. Treatment at a preinvasive stage is often completely curative.

An **adenocarcinoma** results from neoplasia of epithelial tissue that has a glandular origin, glandular characteristics, or both.

Metastasis is the spread of a cancer to a non-adjacent part. The 'new' tumour is known as a metastatic (or secondary) tumour. The cells are similar to the original tumour, meaning that if breast cancer metastasises to the lungs, the tumour is metastatic breast cancer, not lung cancer.

Adenosis is when the lobules are enlarged and sometimes touch one another, if they do touch this makes it simpler for the examiner to feel it or find it in general. This is not cancerous, but can be a precursor.

4.2 Healthy and cancerous biological tissue

The basic common cell structure has already been discussed in section 2.4. In this section, different tissue structures found in the body, relevant to the samples being investigated, are discussed. There are four basic tissue types in animal tissue – connective, epithelial, muscle and nervous tissue. Connective and epithelial tissue is of greatest interest for this thesis.

Connective tissue is often described simply as 'stroma' when in an organ, and provides a supportive framework for other cells. Stromal cells have been found to have an effect on cancer growth and progression, due to their control of growth factors [118].

Epithelial tissue is the tissue that lines surfaces of many structures in the body. This is not just limited to skin, but also includes the gut and throat lining, and of interest to breast cancer, the lining of milk ducts.

There are also particular structures unique to the oesophagus and breast that are discussed in sections 4.2.1 and 4.2.2, along with tissues types of interest to this thesis.

4.2.1 Particular to oesophagus

The oesophagus is the part of the throat dealing with food, pushing food to the stomach with waves of muscle contractions [119]. In the samples looked at in this thesis, only the innermost layer of the oesophagus (i.e. the surface that would be in contact with food) is looked at. This is the epithelium, which in the oesophagus is comprised of stratified squamous epithelial cells [119]. The surface of the oesophagus is often arranged with folds in the structure [119]. Below the epithelial tissue, near the surface of the oesophagus, is typically a layer of connective tissue, and below that, a thin layer of muscle [119]. A typical view of the epithelium and connective tissue that can be seen in the oesophagus is shown in Figure 4-1.

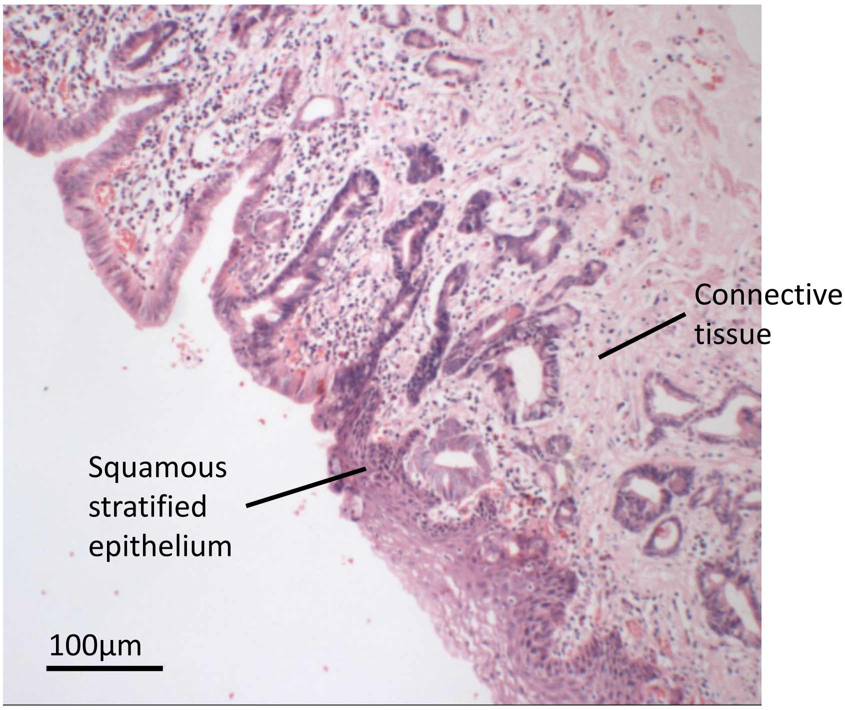


Figure 4-1. Typical view of H&E stained oesophagus biopsy, with the squamous stratified epithelium and underlying connective tissue marked. The epithelium appears a dark purple, whist the connective tissue appears a light pink. Below the connective tissue is usually a layer of muscle that contracts to aid in moving food down the oesophagus.

The most common tumour growth in the oesophagus is squamous cell carcinoma, though adenocarcinomas may be found in the lower end on the oesophagus [117]. Obstructive symptoms typically do not occur until lesions (abnormalities) are well advanced, with extensive invasion of the locally surrounding tissue [117].

4.2.2 Particular to breast

The breast has a number of more specialised structures present for the purposes of producing and transporting breastmilk. Of particular interest are lobules, small lobe-like structures found in the mammary gland that are involved in the production of milk, and ducts, the vessels that transport milk from the mammary glands towards the nipple. These structures have an inner lining of epithelial cells. In non-lactating women, the duct can be blocked by a plug of fibrous protein to prevent bacteria from entering the duct. The other tissue in the breast, the tissue that surrounds these structures, includes fat and fibrous tissue.

There are various acronyms used when discussing breast cancer, reflecting the origin of the cancer and how localised the cancer is. Note that the term ‘carcinoma’ is used when the tumour is derived from epithelial cells, as is the normal case for breast cancer.

Ductal carcinoma in situ (DCIS) is a tumour that has grown from the lining of a milk duct and is still wholly contained within the milk duct. DCIS rarely produces symptoms of a lump that can be felt, so it is usually found through screening mammography. Invasive ductal carcinoma (IDC) is a tumour that broken the confines of the duct and is now growing into surrounding tissue. This happens after the DCIS tumour grows, distending the duct and eventually breaching the basement membrane. Due to the difficulty of detecting DCIS via physical inspection, tumours are often invasive when first diagnosed, but will have previously undergone an in situ stage. The vast majority (~80%) of breast cancer is of ductal origin [117].

4.2.2.1 Three-tier grading system of breast cancer

Breast cancer uses a three-tier grading system that takes three different factors into consideration [120]. In grade 1 (low grade), the cells are well-differentiated. In grade 3 (High grade), the cells are poorly differentiated. Grade 2 (intermediate grade) is a case between grade 1 and 3. Grading is distinct from staging, which is based on how the cancer has spread/is localised.

Each of the three factors that go into the grading is given a score between 1 and 3 [120]. These scores are added up, with the sum determining the grade of the cancer. 3-5 points is grade 1, 6-7 points is grade 2 and 8-9 points is grade 3. The three factors used in grading are as follows [120]:

1. “Tubule formation”. The overall appearance of the tumour has to be considered. It looks at the tissue structures present. Normal tissue forms normal duct structures (i.e. tubules). Cancerous tissue stops communicating with each other and the structures become less orderly. 1 point: tubular formation in more than 75% of tumour. 2 points: tubular formation in 10-75%. 3 points: tubular formation in less than 10%.
2. Nuclear pleomorphism. 1 point: nuclei with ‘minimal’ variation. 2 points: moderate variation. 3 points: marked variation.
3. Mitotic count. How many dividing cells are seen in high magnification. Higher mitotic activity leads to more points, dependent on the magnification and microscope being used.

4.3 Samples under investigation in this thesis

The tissue samples looked at consist of different biopsy types. Core biopsies return a relatively smaller area to look at, but often are a less invasive procedure for the patient, and are usually taken at an early stage for the purposes of diagnosis and prognosis [121]. Excision biopsies are much larger, and usually take place after diagnosis, as a measure to completely remove a tumour. Typically more than just the tumour will be removed, with some surrounding tissue also removed to ensure that the tumour has been completely extracted.

"Core" tissue samples have not been surgically removed with a scalpel, but are instead extracted with a large-bore hypodermic needle (see Figure 4-2) [121]. The hollow needle extracts a relatively thin core from the tissue (~1-2mm in diameter). As surrounding tissue that is not abnormal will likely be extracted, it is possible to see if the cancer is at an invasive or non-invasive stage. This type of biopsy is most commonly used with biopsies looking for breast cancer [121].

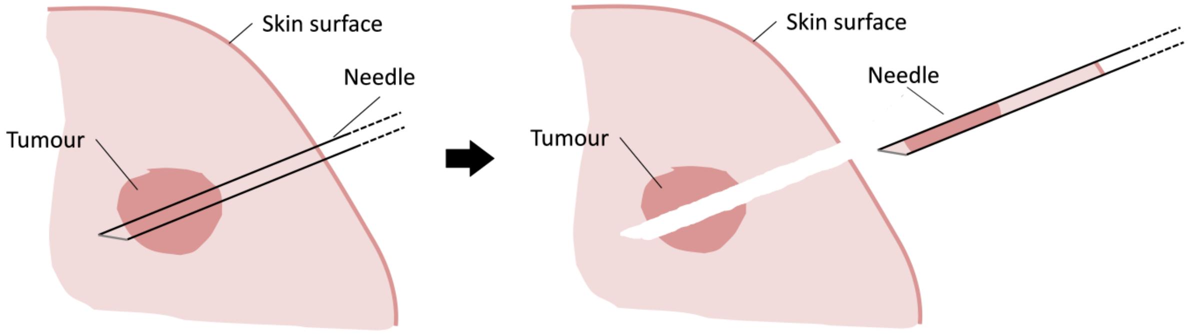


Figure 4-2. Core biopsy process, whereby a large-bore hypodermic needle is inserted into the area that contains tumour, and extracts a 'core' of tissue. Healthy tissue will also be extracted. This type of biopsy usually takes places to extract a sample of tissue to be graded by a histopathologist. This process contrasts to an excision biopsy, in which a scalpel is used to remove the entire cancer, a much more invasive process that will leave more scarring on the patient.

Although the IR images taken are compared with H&E stained visible samples to identify regions of interest, the two are not the exact same samples. The two will in fact be adjacent slices of the same tissue, with one slice going through the normal staining procedure, and the other being prepared for viewing in the IR, stopping short of the staining procedure. The slices are considered to be close enough spatially that the cell types and structures will be similar.

4.3.1 Sample preparation

The samples under investigation underwent a process almost identical to the process for H&E staining, with the difference being that the dyes are not applied near the end of the process. The process is described as follows. The sample preparation process does not have an effect on the cellular structures and molecules that are probed in the MIR.

Following a biopsy, the extracted tissue is treated with formaldehyde solution. This serves the main purpose of preservation, to prevent decay and stabilise the tissue structurally and at a cellular level in preparation for further processes. The formaldehyde does not have an adverse effect on the spectral signal that is under investigation in the MIR [55], should any still remain.

The next stage is to embed the tissue in paraffin wax in preparation for slicing the tissue into thin sections. This happens after alcohol immersions to soften the tissue for wax impregnation and also serves to remove water from the tissue.

After embedding in paraffin blocks, the blocks can be stored for many years without degradation of the tissue. When required, a microtome is used to cut a thin slice of tissue (~5-10μm) off the block. The slice falls onto a liquid, but remains on the surface due to surface tension, where it is scooped from below to adhere it to a microscope slide. The slide is then washed in xylene to remove the paraffin.

If doing H&E staining, the dyes would be applied at this stage and fixing of the coverslip with DPX transparent adhesive would occur. When investigating in the IR, the dyes are not required.

4.3.1.1 Substrates used for samples

The substrates for the samples to view in the MIR are 1-inch diameter BaF_2 or CaF_2 circular discs. Initially, there was a 2mm thick substrate, along with a 1mm coverslip. The coverslip was attached with DPX glue. When preparing visible samples, the DPX, which is transparent in the visible, is allowed to mix with the tissue sample. However, DPX is not transparent in the MIR, so the glue must only be placed on the edges of the substrate.

A common problem with the addition of the coverslip is the creation of Fabry-Perot interference fringes, due to the small air gap between the substrate and the coverslip. This can be removed by later repositioning the coverslip. Due to this problem, it was decided to not use a coverslip at all. The tissue adheres to the substrate well, and there has been no evidence of tissue detaching itself from the substrate.

BaF_2 can transmit at longer wavelengths than CaF_2 (up to 11µm rather than up to 9µm), but it is also typically more than twice the cost of CaF_2 and is more hygroscopic (can dissolve in water), so cleaning and reusing slides is less feasible [122]. Slides can usually be washed and reused with water and a mild detergent, although this has not been verified with the dry tissue samples being worked with.

Using CaF_2 instead of BaF_2 is thought to reduce transmission at the 'Phosphodiester base' wavelength by an estimated 5% using the *digistain* filters due to the cutoff wavelength being slightly shorter for CaF_2. CaF_2 is actually marginally more transmissive than BaF_2 at shorter wavelengths however, so the overall reduction in SNR from using CaF_2 substrates will be minimal.

4.3.2 5933 oesophagus sample

One set of samples that have been looked at extensively in previous work are a selection of oesophagus biopsies provided by Cancer Research UK and the University of Oxford under a Local Research Ethics Committee permission Number 07-Q1604-17. One sample in particular that is of interest is sample number 5933, as it has the practical advantage of having healthy and cancerous tissue of the same tissue type in the same field of view at 3x magnification using the IR cameras from both systems.

The area is question is shown in Figure 4-3. There is a fold in the epithelium (the wall of the oesophagus) which delineates two regions, with one area being healthy, normal epithelium and another being classified as high grade dysplasia, a pre-cancerous stage of tissue. Pre-cancerous tissue does not yet contain cancerous cells, only signifying an increased likelihood that cancer will occur. The classification of the sample was carried out by Bill Otto and Sir Nick Wright of Cancer Research UK.

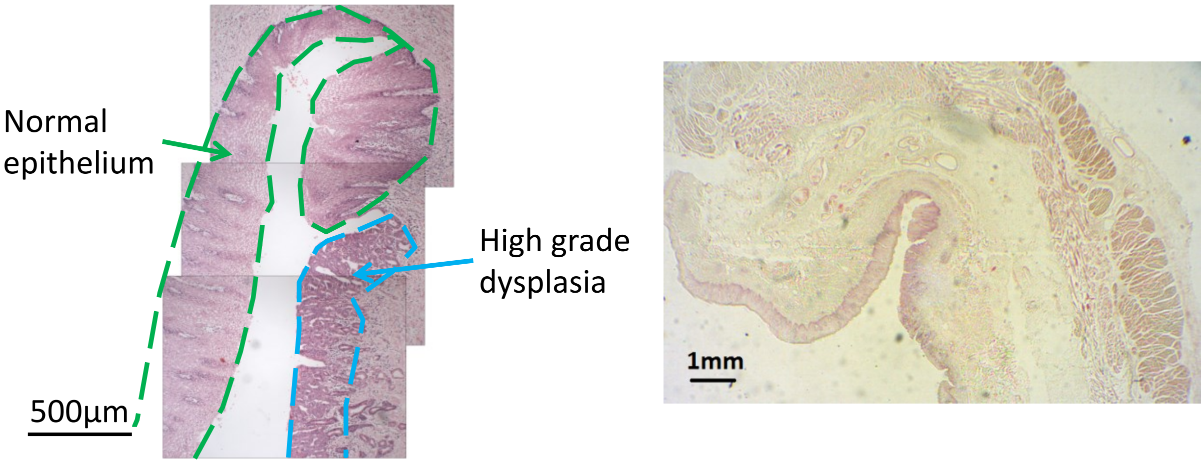


Figure 4-3. H&E stained image of the 5933 oesophagus biopsy. The tissue stained a darker pink is the epithelial tissue, the surrounding tissue stained a paler pink is stroma. The length (long axis) of the structure shown in the left image is approximately 2-3mm. The areas of healthy epithelium and high grade dysplasia are marked. Dysplasia is considered the earliest stage that a histopathologist can identify a change in the cells, and is a stage before being classified as a carcinoma.

4.3.3 CS1x breast core samples

An initial batch of 26 samples of breast core biopsies was acquired from Charing Cross Hospital under a Local Research Ethics Committee permission Number R12074. These core biopsies are varied in the tissue types and structures present and the grade of tumour present. These include DCIS, IDC, blood, fibrous tissue, adenosis, epithelium, lymphocytes, and regions of necrosis. An example of the core biopsy samples are shown in Figure 4-4 as H&E stained images, with regions of interest marked when possible.

These samples have been described in detail and graded by Sami Shousha and Charles Coombes at Charing Cross Hospital.

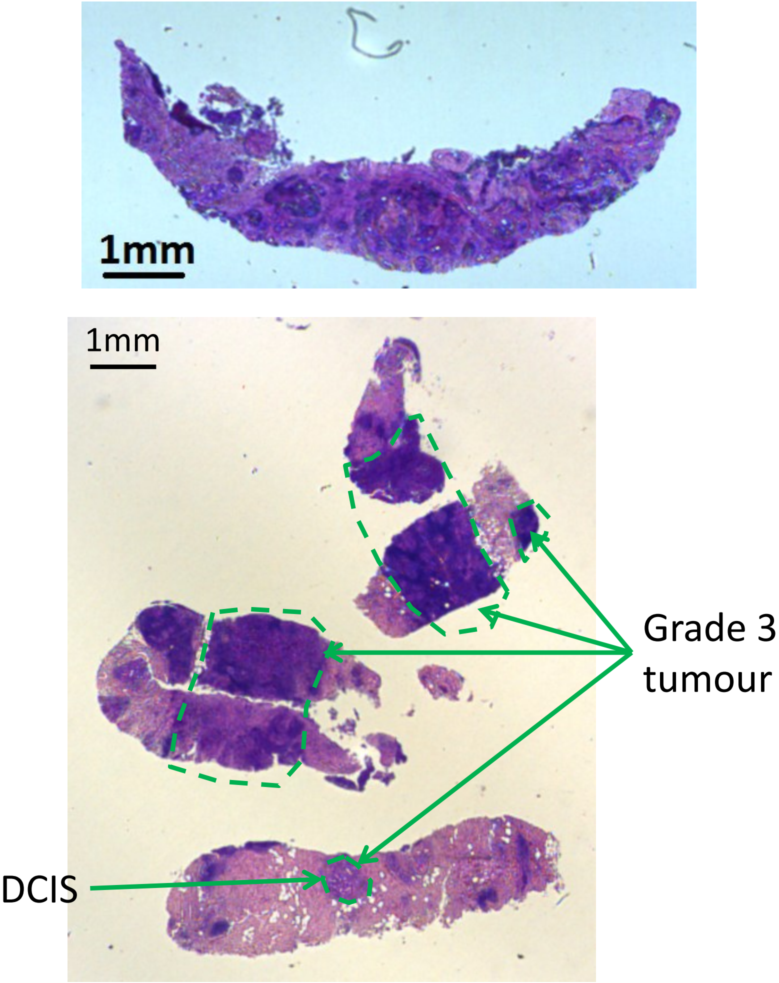


Figure 4-4. H&E stained images of breast core biopsies, samples named CS12-436 (top) and CS12-2626 A1 (bottom). The width of the images shown is approximately 8mm. The regions of cancerous tissue are marked in the bottom image, whilst in the top image, the tumour is mixed with other tissue in the right hand side. CS12-436 contains tumour classified as grade 1 and CS12-2626 A1 contains tumour classified as grade 3. The majority of the rest of the tissue is fibrous (connective) tissue.

5 Results and discussion

The results obtained from the *digistain* and laser based systems are presented in this section (mainly looking at the cancer biopsy samples described in section 4.3), reflecting the current stage of development of each system.

Note that uncertainty values are quoted as and calculated using the standard deviation, not the standard error of the mean. The uncertainty is quoted to one standard deviation. In the context of tissue imaging and creating phosphodiester-to-amide (PA) ratios, each pixel will contain information on the PA ratio of a cell or a group of cells. There should be natural variation in the PA ratio of similar cells. To convey this variation, the standard deviation is used. The standard error in contrast is typically used in situations where there is one true value, to convey how precise the data is.

As a reference to put the transmission images of tissue biopsies in context, the image shown in Figure 5-1 shows a Raman image of a human polyp colon section taken with a point-by-point scanning system [28]. The false colour image shows different tissue types as identified by the Raman spectrum at each pixel. The colours were assigned using self-organising maps in combination with principal component analysis [28]. The pixel resolution of the image is slightly lower than can be achieved with a single *digistain* image, although the spatial resolution is considerably better (each pixel imaging a point 2µm in size, as oppose to the *digistain* imaging a point 12µm in size). The image took approximately 45 hours to acquire [28].

Scanning systems will always be slower than those with multi-channel detectors, such as those with an FPA (focal plane array). Commercial FTIR systems with FPA detectors have been available for a number of years now. As a reference, the best FPA FTIR systems take approximately 7 minutes to acquire an image with similar pixel resolution to the *digistain* using suitable settings for “real-world conditions”, although the SNR that can be achieved is not stated [123]. At approximately the same pixel resolution, the *digistain* system currently takes 3 minutes 20 seconds to image at the four wavelengths of interest to create a PA ratio image, whilst the imaging system that uses the QCL source takes approximately 11 minutes to image at the four wavelengths of interest when operating in a transmission mode.

Both the *digistain* and QCL imaging systems are magnitudes faster than the Raman system with which Figure 5-1 was acquired, and the *digistain* system is also faster than the best commercial FPA FTIR systems. The QCL imaging system is slower than the best FPA FTIR systems, but without knowing the SNR of the FPA FTIR system, a direct comparison cannot be made. The *digistain* system in particular compares well with other infrared spectral imaging systems, also considering the fact that it is built with relatively cheaper ‘low-tech’ components. The QCL imaging system is slightly slower with acquisition time, but has the benefit of being able to operate in a diffuse reflection mode, imaging on live human skin, a first for this region of the spectrum.

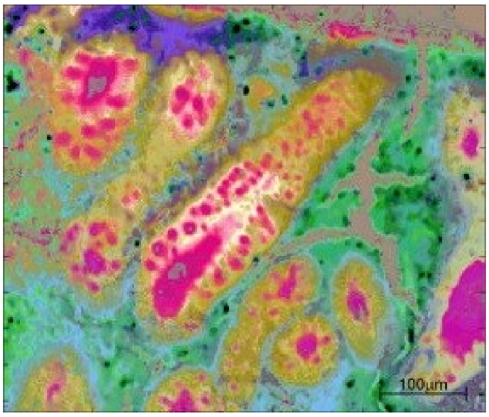


Figure 5-1. False-colour Raman image distinguishing different tissue types in a human polyp colon section. The image is generated from point-by-point scanned Raman spectra. The pixel resolution is 267x267, the area shown in less than 600x600µm, and it took ~45 hours to acquire the data. Magenta shows mucin pools, green shows connective tissue, yellow shows cytoplasmic protein and blue shows basement membrane. Image taken from [28].

5.1 *Digistain* images of tissue samples

The *digistain* system is being used only to create a PA ratio image, after taking images at four different wavelengths. Two different sample sets were looked at – the 5933 oesophagus sample and the CS1x breast core samples – both in a transmission geometry.

There is a certain level of judgement when it comes to making a *digistain* PA ratio image, since there is currently no fixed goal for the SNR we need to achieve – the higher the SNR is, the better. The image quality (SNR) will improve with acquisition time, but we need to set a limit on how long we spend taking the images. The current length of time is fairly arbitrary, chosen mainly due to ‘historical’ reasons – at one stage in the *digistain*’s development, a 20 minute acquisition time was deemed sufficiently long to show contrast between tissue types in samples. As the system was improved, the acquisition time was modified to maintain the same SNR that was previously achieved with 20 minutes acquisition. The current setup requires an acquisition time of 3 minutes and 20 seconds to give an average SNR of 215 over the four acquisition wavelengths. This value for the SNR is the SNR across the image. Since the *digistain* only operates at four wavelengths, a reliable reading for the SNR on a spectrum cannot be obtained, though values for the repeatability of a measurement at a single wavelength can be calculated. The repeatability errors are small enough relative to the intrinsic variation in tissue absorbance to be negligible (the errors are less than 1%).

5.1.1 *Digistain* imaging of the 5933 oesophagus sample

The 5933 oesophagus sample described in section 4.3.2 was imaged using the *digistain* setup and routine described in section 3, being imaged at 4 different wavelength bands and combined to produce a PA ratio image. The amide and phosphodiester absorbance peak heights are shown in Figure 5-2. For reference for which areas correspond to which tissue types, refer to Figure 4-3 on page 106. From these pictures, it can be seen that the epithelium has noticeably different amide content from the surrounding tissue, although from the 'amide peak height' image in Figure 5-2 there does not appear to be a difference between the high grade dysplasia (HGD) epithelium and the normal tissue. The phosphodiester content is shown to be more uniform across the imaged tissue (there is less contrast between the epithelium and the surrounding tissue), although certain regions of the normal epithelium have lower phosphodiester content than the HGD epithelium.

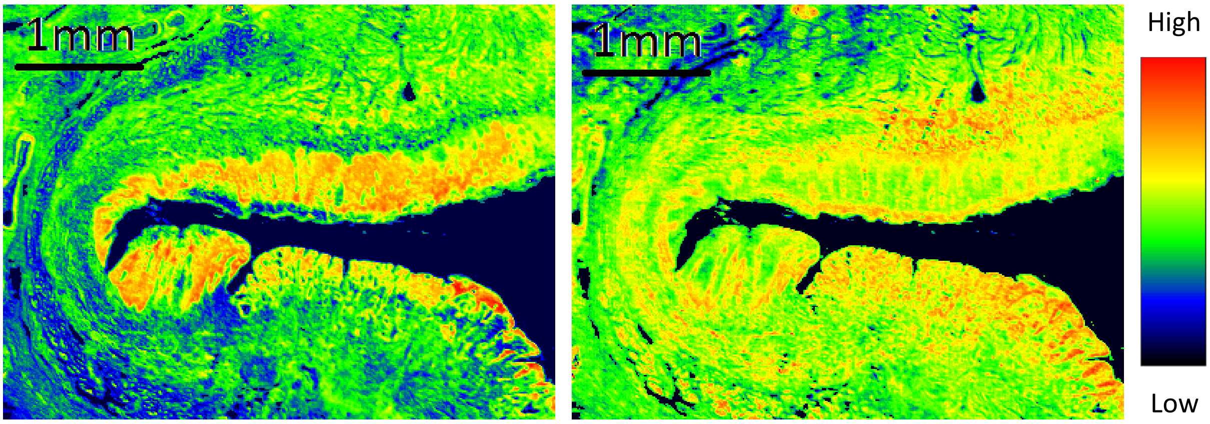


Figure 5-2. Amide peak height (left) and phosphodiester peak height (right) absorbance images of oesophagus 5933 sample. From the amide peak image, it can be seen that the epithelium has a higher amide concentration than the surrounding tissue, though the concentration appears to be fairly uniform within the epithelium. There is less contrast across the sample in the phosphodiester peak image, but there are small differences in absorbance noticeable between areas of epithelium – the area of high grade dysplasia has higher phosphodiester content than that of the healthy epithelium. The images were captured during an earlier iteration of the digistain system, and took roughly half an hour each to acquire, with an estimated average SNR of 395 over the four acquisition wavelengths.

The PA ratio image is shown in Figure 5-3. Note that this image was produced by setting the variable X in equation 3-4 to 1.2, increasing the importance of the 'peak' wavelength to remove problems of having to divide by a value close to zero (see section 3.1.3). The image in Figure 5-3 shows that the region of HGD generally has a higher PA ratio, and this is substantiated by the numbers. 5 areas of 5x5 pixels in both the normal tissue (white boxes) and the HGD (black boxes) were probed to attain a typical value for the PA ratio in each region. The normal tissue was found to have a mean PA ratio of 0.61 ± 0.05, whilst the HGD was found to have a mean PA ratio of 0.95 ± 0.11. The colour scale has been adjusted to highlight the differences between the two regions of epithelium.

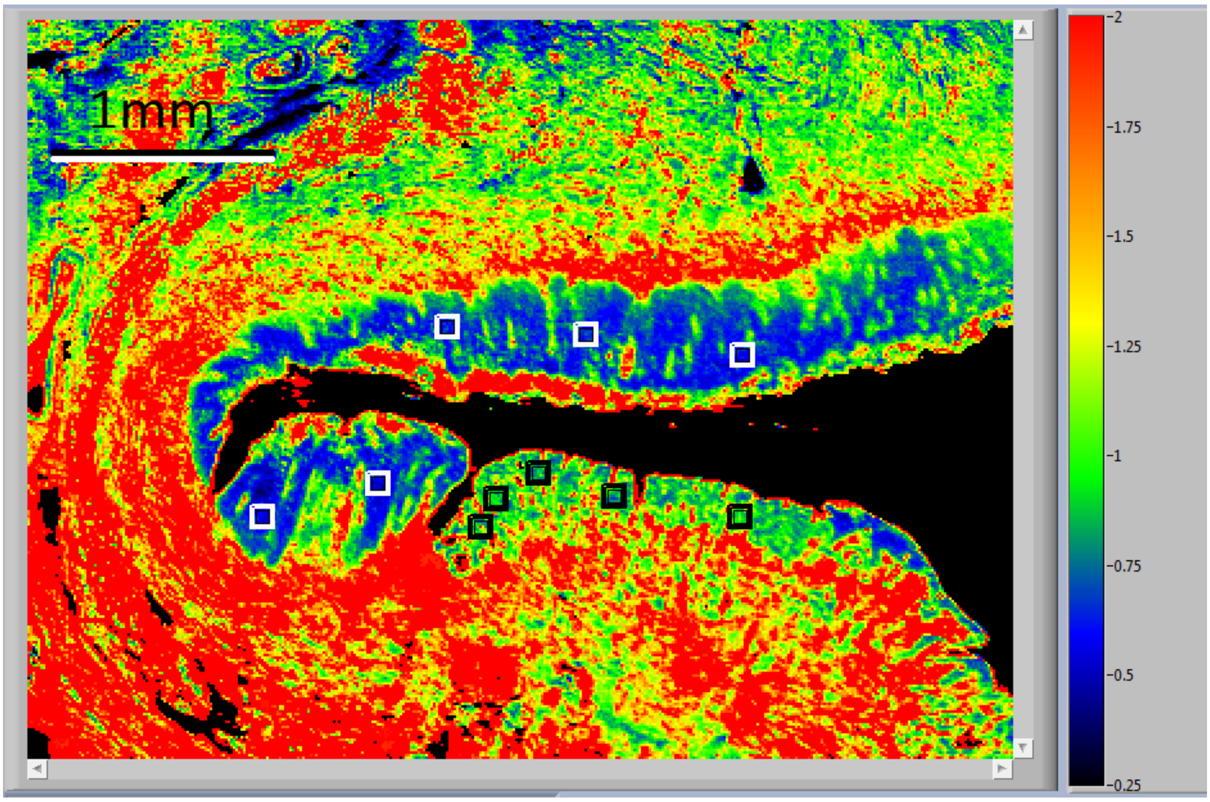


Figure 5-3. Phosphodiester-to-amide (PA) ratio image, made from a ratio of the two images shown in Figure 5-2. The colour scale on the right hand side shows the maximum and minimum pixel values in the image, but the colour ramp has been restricted to a small region of the scale to better highlight differences within the epithelial tissue. 5 boxes of 5x5 pixels were drawn in the healthy epithelium and that classed as HGD (white and black boxes respectively). The average ratio value in the boxes for the healthy tissue was 0.61 ± 0.05, whilst the average of the HGD was 0.95 ± 0.11. This shows two distinct values for the PA ratio in the two areas, with low probability of an overlap in values. The streak of red in the lower part of the upper epithelium is muscle.

The tissue immediately surrounding the epithelium obviously has a much higher PA ratio than either region of epithelium. This is because different tissue types naturally have higher PA ratios than others, as they have a different chemical composition. For example, as is shown in section 5.1.2, fibrous tissue always has a higher PA ratio than epithelial tissue. The type of visual representation as in Figure 5-3 is of most use when comparing regions of similar tissue type. For this reason, if a clinician is to use the information from the *digistain* image, they must also have a normal H&E stained image from which they can identify different tissue types.

More work could be done to characterise the range of PA ratio values that different tissue types may have so that the visible H&E stained image is not necessary to identify tissue types. However, as can be seen in Figure 5-3, there are areas that are not epithelium that have the same or similar colour. Given the number of tissue types often present, there is a high probability that some will have overlapping PA ratios. However, it may be possible, for example, to incorporate the morphology of areas with similar PA ratios to identify the tissue type.

High grade dysplasia is a pre-cancerous stage, a step before being classified as ‘grade 1’. The fact that there is a visible difference in the two epithelial regions in the *digistain* image that is backed up by the underlying PA ratio pixel values is compelling evidence that this method can be used to distinguish between healthy and abnormal tissue.

5.1.2 CS1x breast core samples

Through discussions with the medical staff at Charing Cross Hospital, it became apparent that the problem that histopathologists face is not in identifying where the regions of cancerous tissue is – this is relatively straightforward to do by eye – but is in the accurate grading of biopsies. We acquired a total of 26 breast core biopsies which contained areas in the biopsies with tissue that had been graded as grade 1, 2 or 3. As mentioned in section 4.3.3, these core samples contain various tissue types in every sample. Any tumour is often mixed in among fibrous tissue.

An example of the PA ratio image of CS1x samples in shown in Figure 5-4. Figure 5-4 (top) is an example of a core biopsy in which the tumour is quite mixed with other tissue such as fibrous tissue. Figure 5-4 (bottom) has regions of tumour which are better defined. For reference for which areas correspond to which tissue types, refer to Figure 4-4 on page 107.

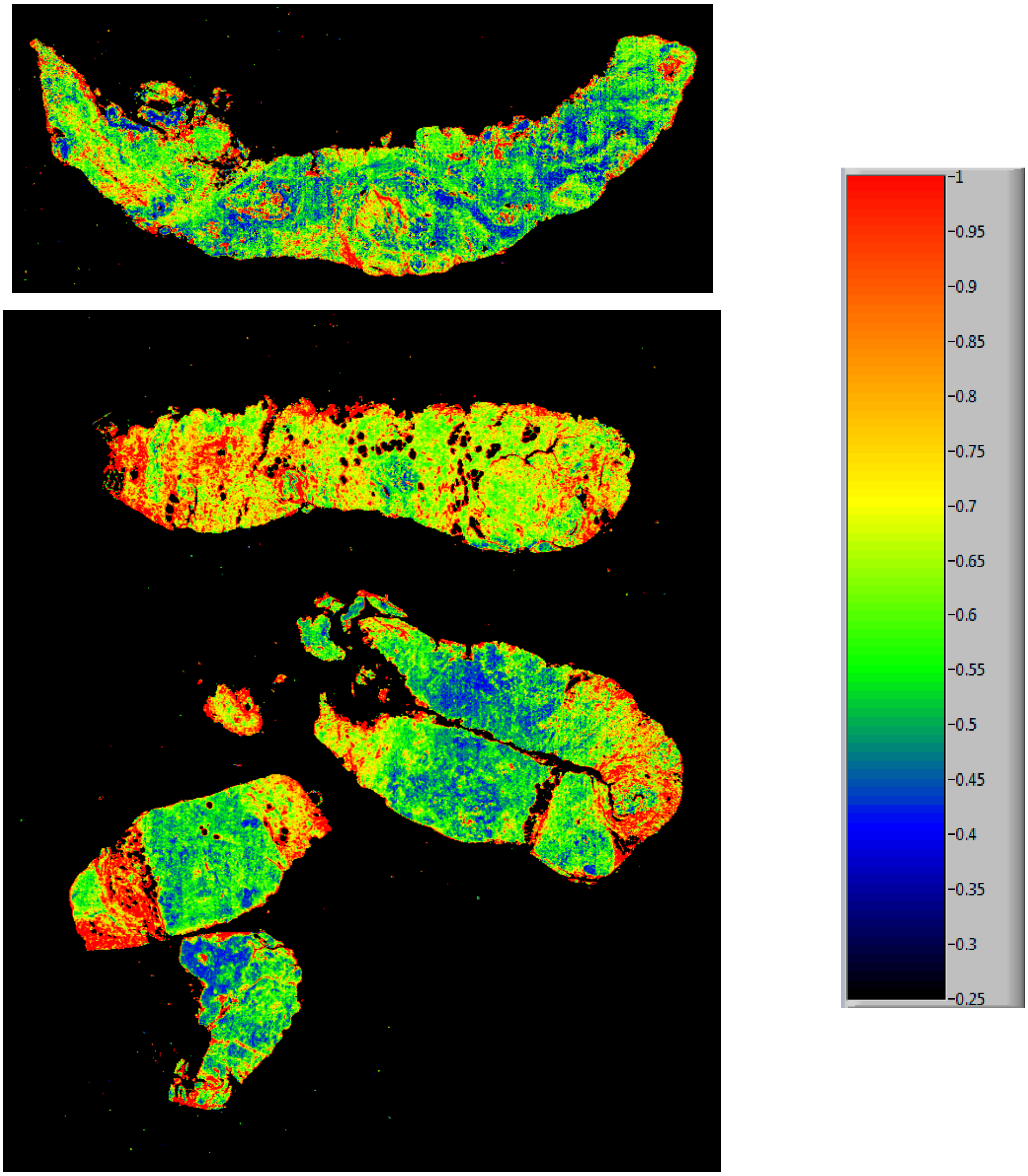


Figure 5-4. PA ratio images of samples CS12-436 (top) and CS12-2626 A1 (bottom) acquired with the digistain system. The width of the imaged area of CS12-436 is 8.3mm and has cancerous tissue classified as grade 2. The width of the imaged area of CS12-2626 A1 is 8.5mm and has cancerous tissue classified as grade 3. In these images, the areas of tumour appear blue or green. Although we expect areas of cancerous tissue to have a higher PA ratio than healthy tissue, there are no areas of healthy tissue in the images to compare with. The areas that appear yellow/red, with a high PA ratio, are mostly fibrous tissue that has an intrinsically higher PA ratio. In CS12-436, the tumour and fibrous tissue is mixed at a relatively microscopic level, meaning if an average of a large area of pixels were taken, the result would definitely be skewed by the presence of fibrous tissue. CS12-2626 A1 in contrast has much larger areas of tumour in well-defined patches. It took a total of 10 mins to acquire the CS12-436 image, and a total of 30 mins for the CS12-2626 A1 image.

From these images, it can clearly be seen that different tissue types appear as a different colour corresponding to a different PA ratio. The samples provided only contained cancerous epithelium, with no areas of normal healthy epithelium, so a direct comparison within the image cannot be made in this case. Tumour tissue appears primarily as a blue colour, with some areas also appearing green. Fibrous tissue (stroma) ranges from yellow-green through to red. Lymphocytes appear a yellow colour; glands appear as a green colour. On other samples with additional tissue types, blood, areas of necrosis and calcification appeared as red. There appears there can be overlap between tissue types – an area of green could be tumour, or could be fibrous tissue. This can be explained by looking at the colour scale used – green can range from a PA ratio of 0.5 up to 0.7. By choosing a different colour scale, differences may be better highlighted. For this reason, focus should be given to the underlying values of the PA ratio, not the colour that an area appears.

Across 19 of the samples, areas of 4x4 pixels were taken of cancerous tissue, with a total of 198 areas investigated to get an average value for the graded regions. To try to minimise the problem of fibrous tissue being mixed in with the regions, only the darkest blue (i.e. lowest PA ratio) regions were chosen.

Note that with the CS1x samples, the factor X in equation 3-4 was set to X=1. X was able to be set to 1 after the approach of taking a background with a blank substrate disc in place was adopted.

The results of the PA ratio sampling are displayed in Figure 5-5. The points at each grade are an average of the areas, with the standard deviation as error bars. It can be seen that the average PA ratio for grade 3 areas is higher than that for grade 1 areas, although the size of the error bars is large enough that they may easily overlap.

From the size of the standard deviation error bars in Figure 5-5, it can be seen that there is a very large degree of variation within each category. It is thought that this is due to the difficulty in ensuring that only pure tumour tissue is being investigated. Although it was attempted to look at regions without fibrous tissue (which naturally has a high PA ratio) by selecting the regions with the lowest PA ratios, there is no guarantee that this was successful when looking at the breast core biopsies that always have some mixing of tissue types. The problem may be even more complex if mixing between tissue types on a single cell level is present.

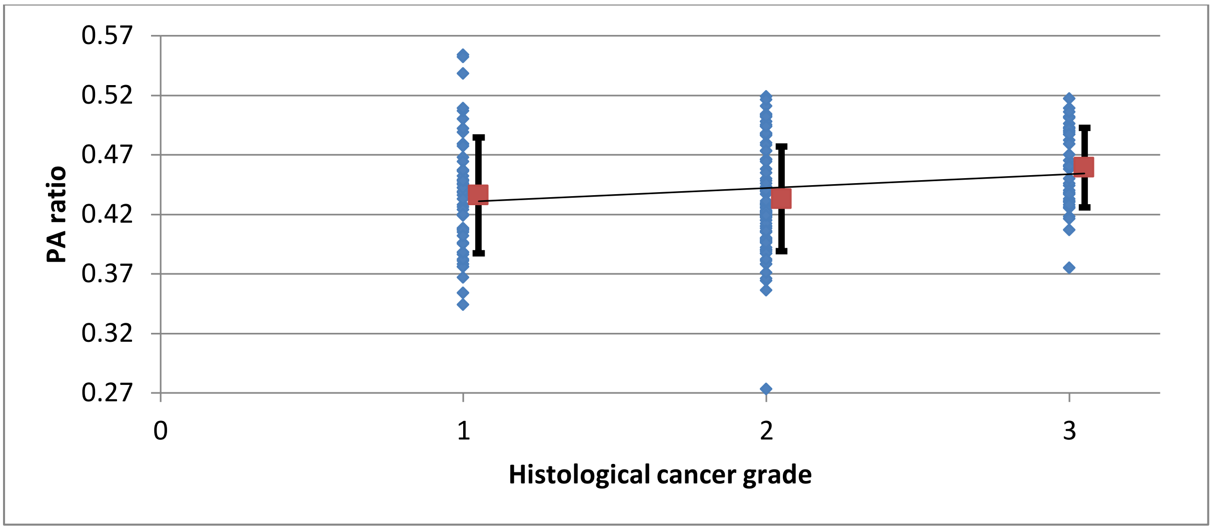


Figure 5-5. PA ratios of 198 graded areas of 4x4 pixels from 19 CS1x samples. The spread of pixel values for all grades is large enough that any potential differences between grades will be hidden. The average PA ratio value for each grade is also plotted with the standard deviation as error bars. The average PA ratio value for grades 1, 2 and 3 are 0.436, 0.433 and 0.459 respectively. Although the grade 3 PA ratio is greater than the grade 1 value as we would expect, due to the relatively large errors, no confident correlation between the grading and the PA ratio can be drawn. The reason for these large errors could be due to an inability to pick out regions of pure tumour, including other tissue types in the measurement. The errors bars on each of the 198 values are not shown, and are small enough (relative to the spread of values) to not affect the overall error.

5.1.2.1 Next steps in analysis

As demonstrated in section 5.1.2 , the results with core biopsies show promise, but are partially dependent on the user's input in drawing the correct region around the tumour to exclude other tissue types that will affect the average PA ratio value. There are further steps in analysis that could be taken to help make the results more repeatable between different users, with reduced user input. Implementing these steps would first require more samples with which to acquire the correct data.

5.1.2.1.1 Extracting a reliable value for the PA ratio

What is really needed, to verify whether there is any correlation between the PA ratio obtained with the *digistain* and the grading of the tissue, are samples that have large areas of pure tumour. Only when we are positive that no other tissue is interfering can we draw conclusions. During the initial stages of cancer diagnosis, the majority of biopsies will be core biopsies, as it is the least invasive and scarring procedure. To be able to work with core biopsies would be beneficial to allow the *digistain* technique to be used without modifying current working practices.

To have large areas of tumour, it will be necessary to acquire new samples, of excised tumour. Excised biopsies are a more invasive type of biopsy, carried out when a tumour has to be surgically removed. The tumour lump will provide a higher chance of being able to probe areas of pure tumour due to its larger size. This should still require the help of histopathologists in identifying regions within the excised biopsies that consist of pure tumour. Once these regions have been identified, accurate PA ratio values for grades 1, 2 and 3 tumour can be obtained.

When the PA ratio for each grade has been determined using the pure tumour, it may be possible to use this information as a calibration to then work with core biopsy samples. Using the histogram of the PA ratio values, it may be possible to acquire a 'spectrum' of the PA ratio values present. It may be possible to separate all the individual peaks that correspond to different tissue types. An example histogram of one of the samples is shown in Figure 5-6, in which distinct peaks for tumour tissue and fibrous tissue can be seen. Figure 5-6 also demonstrates the difficulty that may be faced when trying to separate peaks – the peaks may be so close together that accurately being able to extract the individual peaks will be dependent on the computational technique. In a situation where the peaks certainly cannot be identified by eye, there is no way to verify the accuracy or reliability of the method, and changing some parameters will likely yield different results for the same data.

One method for separating the peaks in the histogram would be an iterative least-squares curve fitting, whereby we try to fit a number of functions of the correct form. It is not yet known the exact form of these curves – due to the subtraction and division of the absorbances at each wavelength, there is no function that we can determine analytically, so we must make a guess at what type of function may fit the curve. From initial observations, an asymmetric heavy-tailed distribution, such as a function of the form

$$e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} \tag{5-1}$$

may do well. One problem with this method is determining how many peaks to fit, though with appropriate bounds on the fitting parameters, it may be possible. This step will only be possible when samples with areas of pure tumour are obtained to ascertain the actual shapes of the PA ratio histogram. Furthermore, there is no guarantee that the PA ratio curves will have the same form – fibrous tissue may have a slightly differently shaped histogram. To determine if this is true, some samples with areas of pure fibrous tissue should also be investigated.

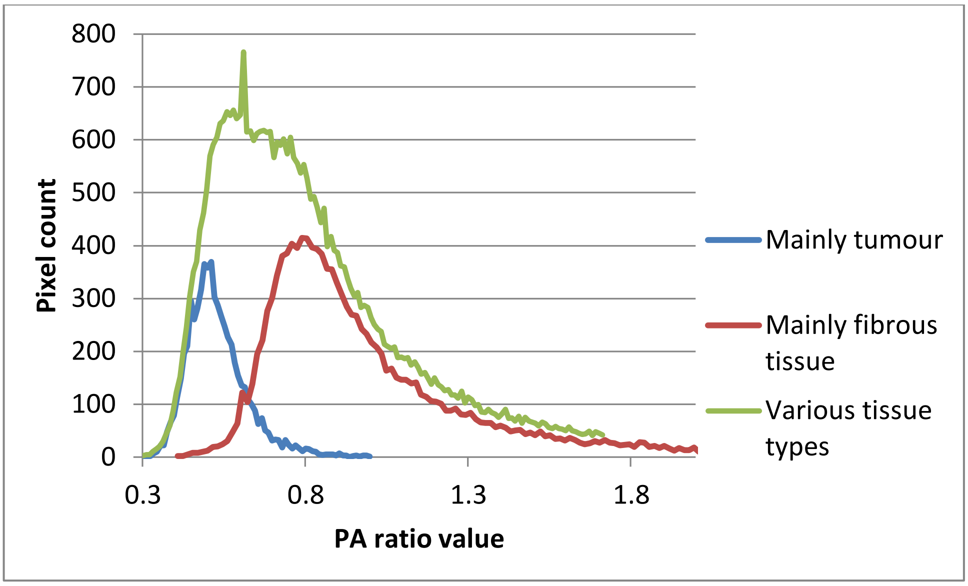


Figure 5-6. Typical histogram of a CS1x PA ratio image. The blue histogram is of an area identified as cancerous, so should contain mainly tumour (though some other tissue types will likely be present). The red histogram is of mainly fibrous tissue. The green histogram is of an entire image that will have all the tissue types included. It can be seen that the tumour and fibrous tissue form distinct peaks, but there is at least one more peak from another tissue type needed to make the histogram of the whole image. The histograms can be thought of as peaks in a spectrum – a peak may have subpeaks from different contributions that add to the overall shape of the peak. We may be able to use techniques usually used to separate spectral peaks to separate the contributions from different tissue types. However, without the knowledge of what the individual peak shapes are, and how many peaks there are, the close proximity of the peaks will add complications to doing this accurately.

It can be seen in Figure 5-6 that by simply setting a single window for tumour tissue, we may be excluding areas of the tumour, or including other tissue types. One possible way to work with only a single window though may be to also use spatial information. Pixels at the upper boundary of the window that are surrounded by values that are near or above the upper range of the window are likely to be an area of a different tissue type, whilst pixels surrounded by values firmly within the window are more likely to be part of the statistical variation in possible values for tumour. This should increase the chance that only cancerous tissue is used to measure morphology, but is still not ideal, as it could exclude very small areas of cancerous tissue embedded in other tissue types.

There are other methods that may be suitable for resolving overlapping peaks within the broad histogram. For example, in reference [85] the amide I peak is deconvolved so that the individual component bands can be resolved.

Showing a correlation between PA ratio and grade of tumour will be a step in the right direction, but the grading system is intended to be a reflection of patient survivability – someone diagnosed with grade 1 cancer will have a higher probability of surviving longer than someone with grade 3 cancer. If we have the patient histories that correspond to all the samples, it may be possible to correlate the PA ratio with patient survival. Although tumours are binned into grades, they are not discrete, so the PA ratio should also be continuous. If the continuous gradient of the PA ratio can be correlated with the survival rate of patients, this information should be of more use to those who have to decide the treatment for cancer patients.

5.1.2.1.2 Identifying tumour via morphology

The PA ratio that we get from the *digistain* images may only be a part of the picture. When a histopathologist grades a tumour, they look at factors such as the morphology of the tissue. The images acquired with the *digistain* can be used to provide quantised information on the morphology of the samples. Two such metrics for the morphology are the anisotropy and the typical "length scale".

There are cases in which the cancerous growth is noticeable by the anisotropy – one example is invasive lobular carcinoma (ILC), in which the cancerous growth is typified by cells growing in single file, invading other tissue. The *digistain* does not have the resolution at the moment to detect this kind of structural detail, so it is suspected that currently, the anisotropy will not be as useful as other measures of morphology. Less obvious forms of anisotropy may be present however, whose patterns only become apparent when looking at statistically large data sets.

The "length scale" is a quantification of how 'mottled' or 'speckled' the growth of the tumour appears. At the moment, there is no hypothesis for how the grade of tumour would correlate with the length scale, so this approach would consist of measuring the length scale of many samples and seeing if any statistical correlation arises between the length scale and the tumour grade, either alone or in conjunction with other metrics such as the PA ratio.

These approaches to quantise the morphology of the tumour should only be tackled when we can use the PA ratio information to separate tumour from other tissue, as we are mainly concerned with the morphology of the cancerous tissue.

5.2 QCL-based imaging

The laser-based system is more flexible than the *digistain*, being able to tune to arbitrary wavelengths in its tuning range.

When looking at the results, we are looking for two things – structure, contained in the 2D image, and spectral response, which corresponds to the chemical composition. One difficult aspect of chemical imaging that is easy to overlook is the choosing of the sample to look at – ideally, it should be a sample that has variations in chemical composition across an area being imaged, the area typically a square ranging from a few millimetres to a couple of centimetres. Additionally, the sample should have areas with the same underlying structure in regions where the chemical composition varies, so that we can attribute any changes in contrast in the image solely to the chemical composition.

The transmission results are quantitative (if the nonlinear response of the detector has been removed, in the case of the QCL system), whilst the reflection results can only provide qualitative information. The reflection mode setup looked at both ‘cold’ samples and *in vivo* diffuse reflection imaging of human skin, marking the first time this has been possible in the MIR region of the spectrum.

5.2.1 Results from QCL transmission setup

Acquiring transmission data using the QCL takes considerably more effort than the *digistain* setup due to the problems of coherent imaging and nonlinearities. However, the laser-based system is more flexible and has potential for higher SNR data collection.

As one of the main samples we look at, it is important that the spectrum of tissue that we can acquire using the QCL matches what we would get from an FTIR. A sparse spectrum acquired using the QCL system is shown in Figure 5-7, overlaid on a bulk tissue spectrum acquired using the FTIR. The QCL points were calculated from the average of the entire array from the CEDIP camera. The fact that the QCL points do not line up directly with the FTIR spectrum is not important, as it is unlikely that the exact same part of tissue is being probed by both systems. Figure 5-7 gives no indication of the quantitative reliability of the QCL system, but, qualitatively, it shows that the QCL system is able to pick out the spectral peaks present.

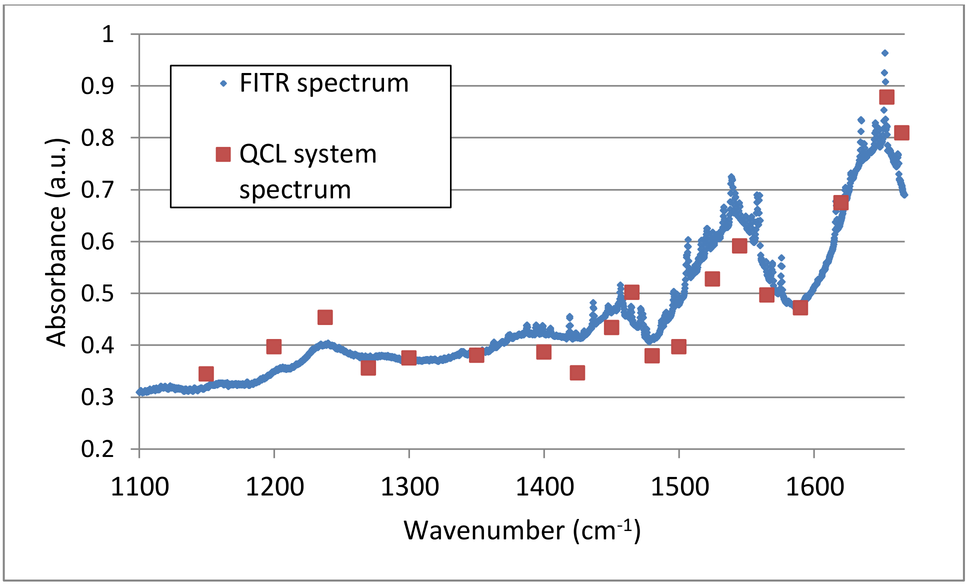


Figure 5-7. Spectrum of bulk tissue from FTIR (blue) and sparse spectrum of a different area of tissue taken with the laser-based system (red). The two spectra cannot be compared directly, being from different areas which will have intrinsically different absorbance across the spectral range. The QCL spectrum does roughly follow the FTIR spectrum though, showing that qualitatively, the QCL system is able to give a spectrum that can match an FTIR spectrum, following the same pattern of peaks and troughs.

5.2.1.1 5933 oesophagus sample

The sample described in section 4.3.2, the 5933 oesophagus biopsy, was investigated at the QCL wavelengths of interest for biological tissue, the amide and phosphodiester wavelengths. The 4 wavelengths investigated were similar to the wavelengths at which the *digistain* images (see section 3.2.1.2). The wavenumbers used with the QCL system are 1654cm^{-1}, 1590cm^{-1}, 1238cm^{-1} and 1176cm^{-1} for the amide I peak, amide I base, phosphodiester peak and phosphodiester base respectively. However, the linewidths are much narrower than can be achieved with the *digistain*, with greater accuracy to the peak and base locations. Due to the flexibility of the QCL source, we also have the choice to look at the other amide and phosphodiester wavelengths, the amide II peak at 1545cm^{-1} and the symmetric phosphodiester mode at 1080cm^{-1}. However, it was found that the amide I peak and the antisymmetric phosphodiester mode gave results with more contrast, probably due to the fact that the absorption peaks are larger in magnitude.

The images at the 4 wavelengths are combined in the same way as the *digistain* images, described in section 3.1.3, with the absorbance images resulting in a PA ratio image. One such PA ratio of the 5933 sample is shown in Figure 5-8. The H&E stained visible image is also shown for reference. For reference for which areas correspond to which tissue types, refer to Figure 4-3 on page 106.

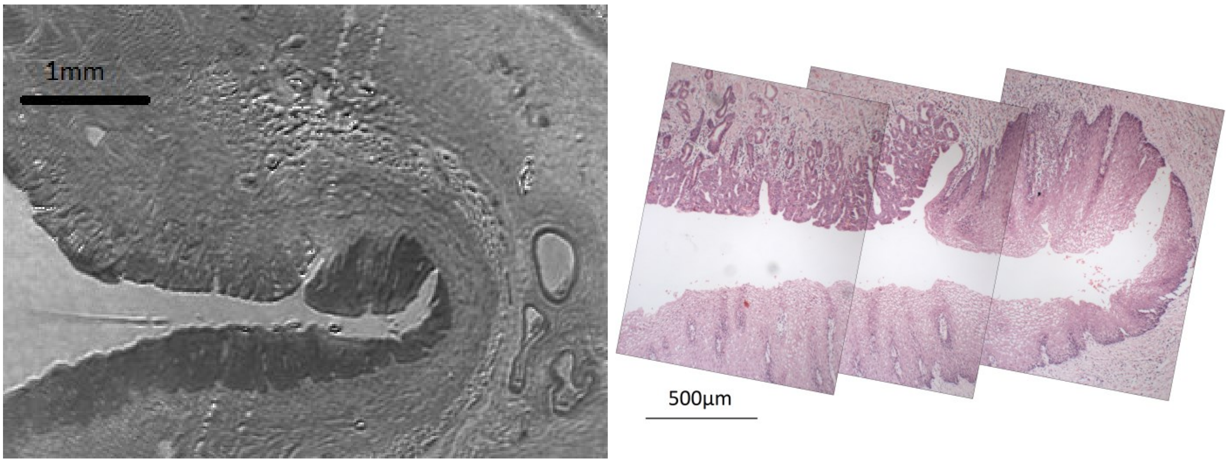


Figure 5-8. Grayscale PA ratio image acquired using the QCL system (left) with 2x magnification and the corresponding H&E stained image (right). The epithelium appears to have a lower PA ratio than the surrounding tissue, as was also seen with the digistain results. The upper-left layer of epithelium (in the H&E image), identified as high grade dysplasia, appears to have a higher PA ratio than the normal healthy epithelium.

The region described as high grade dysplasia (HGD) appears brighter than the region of normal epithelium. Looking at the values for the PA ratio in these two regions supports the idea that the HGD has a higher PA ratio. This was confirmed by looking at the average value of a collection of areas of 5x5 pixels. These areas are shown in Figure 5-9. The green areas have an average value of 0.67 ± 0.04. The red areas have an average value of 0.82 ± 0.04. The results were an average of 10,000 transmission images at each wavelength collected as 1,000 images, interleaved 10 times. This took an estimated 11 mins to collect all the data required to create the PA ratio image.

Note that these PA ratio values cannot be compared with those values obtained using the *digistain*, due to the differing nature of the sources. The QCL has a very narrow linewidth in comparison to the *digistain* bandpass filters, leading to the absorbance at each of the four wavelengths being different. The bandpass filters, being relatively wide compared to the QCL linewidth, will be measuring absorption from wavelengths not included within the QCL measurements. Due to the shape of the bandpass filters, these wavelengths will be weighted differently in contributing to the overall absorption measured when using each filter. Qualitatively, both QCL and *digistain* systems place the HGD regions as having a higher PA ratio than the healthy epithelium – a PA ratio of 0.82 ± 0.04 in HGD with 0.67 ± 0.04 in normal epithelium, for the QCL system; and a PA ratio of 0.95 ± 0.11 in HGD with 0.61 ± 0.05 in normal epithelium, for the *digistain* system.

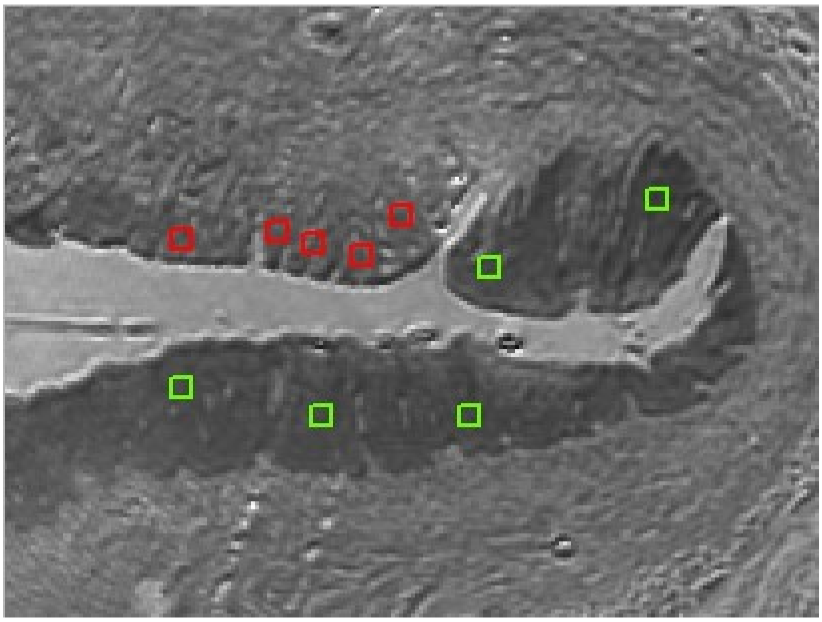


Figure 5-9. A region of the QCL PA ratio image as in Figure 5-8, with areas investigated marked. The red boxes investigated 5x5 pixel areas of high grade dysplasia (HGD) epithelium; the green boxes the healthy epithelium. The red areas have an average PA ratio value of 0.82 ± 0.04, the green areas an average value of 0.67 ± 0.04. The HGD region has a higher average PA ratio value than the healthy epithelium, with the uncertainty indicating a low probability of overlap between the two values. The PA ratio image took approximately 11 minutes to capture, with an estimated average SNR of 85 over the wavelengths used.

The error in the PA ratio value also cannot be directly compared between QCL and *digstain* systems, as the areas measured in each image are not identical. The only value that can be directly compared is the SNR. The in the most recent design iterations, the SNR of the *digistain* within a given acquisition time is over 4 times greater that the QCL system can achieve. As previously discussed, it is thought this difference is largely due to the noise that the diffuser introduces in the QCL system.

5.2.1.2 Comparison of spectrum acquired with QCL with FTIR spectrum

Plastic has some distinct absorption peaks in the infrared, so this was used to see how well the spectrum from a QCL matched the shape of an FTIR spectrum for a homogenous sample. The two spectra are shown in Figure 5-10.

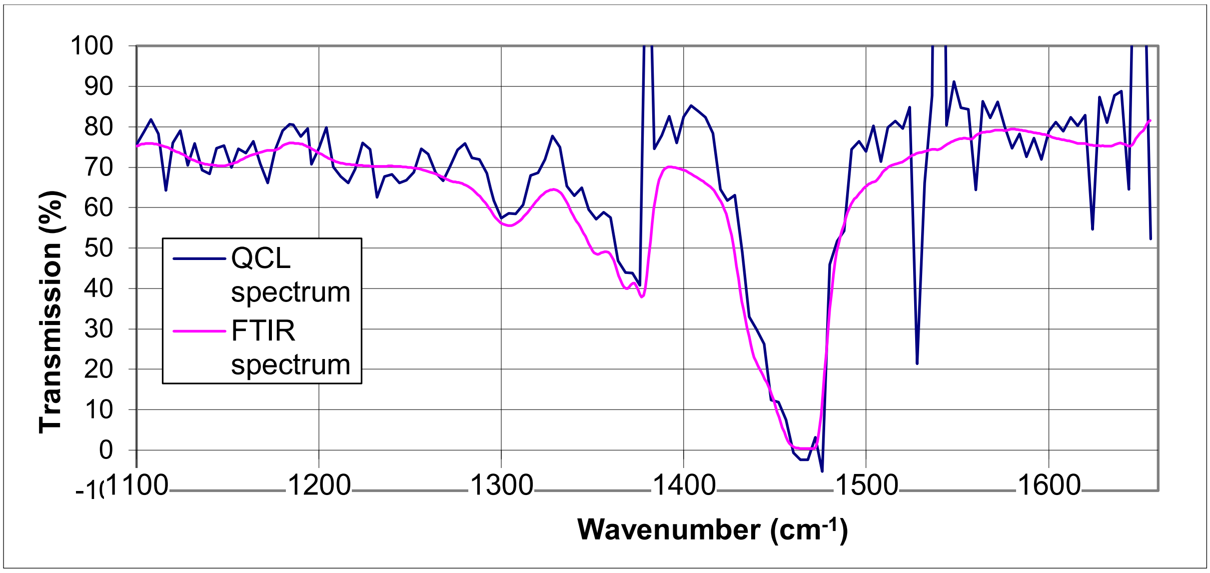


Figure 5-10. Transmission spectrum of a thin layer of plastic (polyethylene) from an FTIR (pink) and the QCL system (blue). The QCL spectrum has a resolution of 4cm^{-1}, the FTIR spectrum a resolution of 1cm^{-1}. The QCL spectrum does roughly follow the FTIR spectrum, with two notable differences. The first difference is the point-to-point variation, the noise on the spectrum. The FTIR spectrum is very smooth along the spectrum, reflecting its low noise, whilst the QCL spectrum varies considerably from point to point. The second difference is the approximate transmission value. Ignoring the noise on the spectrum, the value for transmission is generally greater on the QCL spectrum. The most likely reason for this is the nonlinear response, which presumably must not have been completely removed at all wavelengths when this spectrum was taken.

The QCL spectra shown in this section are an average of 1,000 images, with 32ns pulses from the QCL. The acquisition time per wavenumber is roughly 20 seconds. The QCL spectrum was taken with a resolution of 4cm^{-1}. The two match quite closely, with the exception of some sudden peaks, likely due to a failed image acquisition or a wavenumber that results in a very weak signal (either due to low output or high atmospheric absorption). Note however, that there are some bands of wavenumbers at which the transmission is consistently higher in the QCL spectrum than the FITR spectrum (such as 1300-1420cm^{-1}). It is thought that this is a symptom of the detector nonlinearity problem – when nonlinearity is present, the transmission will be measured as higher than it actually is. The nonlinearity is wavenumber-specific, so this spectrum is able to give an indication of which wavenumbers are more susceptible to nonlinearities.

5.2.1.3 SNR of spectrum acquired with QCL

To estimate the spectral SNR, we take a 100% T spectrum as detailed in section 3.3. This is done by taking transmission images at a range of wavelengths when there is no sample being imaged. The 100% T spectrum was taken at 4cm^{-1} resolution. The SNR across the entire spectrum is calculated using equation 2-10. The mean transmission was calculated as being 100.4%, with the SNR equal to 14.7. This is roughly two orders of magnitude smaller than the SNR that a commercial FTIR system can achieve in the same acquisition time [124]. This spectrum was taken under the same conditions as Figure 5-10, so if a nonlinear response is still present, the SNR of the spectrum acquired with the QCL will not be reliable.

5.2.2 Reflection setup and the effect of surface roughness

The relatively high power of the QCL source makes it feasible to do diffuse reflection imaging, in which only a small fraction of the light may be reflected. In practice, it is difficult to get quantitative results from diffuse reflection imaging – the image would have to be ratioed against an identical surface that displays no absorption. Having a reference sample to image that has exactly the same morphology and surface roughness is a difficult condition to satisfy, so only contrast within an image of a sample that is known to have regions of different chemical makeup is currently sought.

The nonuniformity of the illumination could be accounted for if needed by using a Fourier transform to remove the Gaussian-like profile of the intensity distribution. By removing very low-frequency components from the image, the contrast should improve for the features we are interested in.

One of the main problems encountered was being able to attribute any contrast to differences in the chemical composition and not differences in surface features – a rough metallic surface will direct more light towards the camera than a smooth surface with a setup to image diffuse reflection. However, this would not be due to diffuse reflection, as this reflected light from the top surface has not penetrated the surface. It is in fact the specular reflection, but from a non-flat surface that varies on a relatively short length scale.

Another important aspect of the surface being imaged is its ‘macro’ flatness. For example, an arm can be considered to be cylindrical in shape. When imaging an arm, the curvature of the surface can become apparent. As shown in Figure 5-11, there can appear to be contrast where there should be none when the surface is not flat.

Due to this effect of the curvature, the sample being imaged should be kept as flat as possible. This could be done by pressing a soft sample (such as an arm) against a flat optical window. This would present a flat surface to be imaged, but may also change the way in which the light is scattered. When it is not possible to present a flat surface, the expected possible boundary between regions of contrast is placed in such a way that, even if there is contrast resultant from the curvature, any contrast from differences in chemical composition would still be visible and attributable.

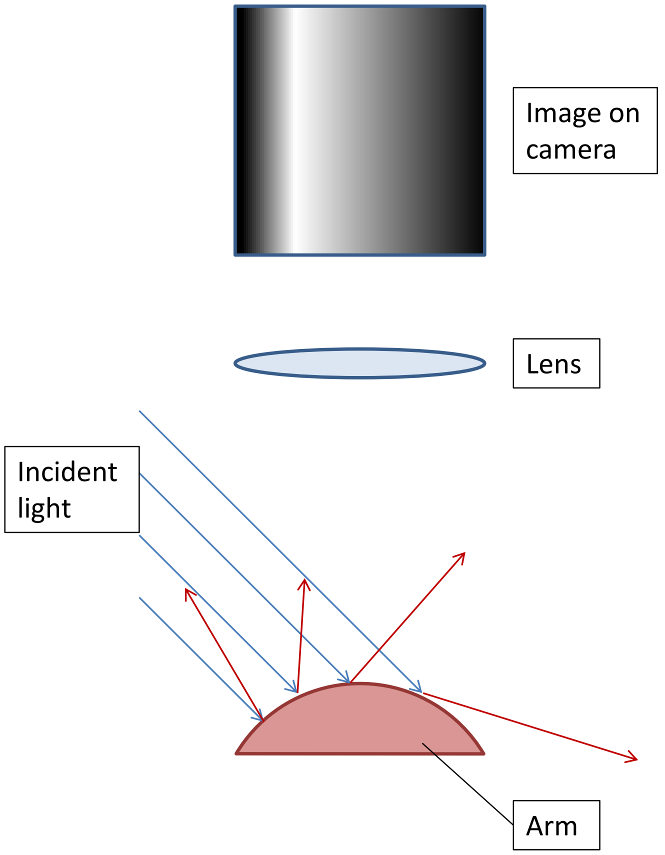


Figure 5-11. Diagram highlighting that specular reflection from a curved surface can interfere with diffuse reflection measurement. If the surface being imaged were flat, the specular reflection component would be directed beyond the imaging lens, ensuring only diffuse reflection is looked at. As is shown in the diagram, with a curved surface, certain parts of the arm specularly reflect such that they are picked up by the camera. A possible intensity distribution on the camera is shown. This will be overlaid on the diffuse reflection, which itself will likely be modified compared to that from a flat surface. The specular components should be of higher intensity than the diffuse components, masking any contrast that may have been present.

The images in this section are displayed with a grayscale colour palette instead of the false colour 'rainbow' palette used elsewhere in this thesis. This is because the reflection images typically display lower contrast than the transmission images, and the human eye is better able to distinguish small differences in brightness than subtle differences in colour.

Early attempts at reflection imaging showed signs of an effect that would later prove to be a considerable problem – contrast could often be evident when there were no differences in chemical composition of the object being imaged. For example, the markings of a metal ruler would appear brighter than the smooth surface of the surrounding metal. No paint was present, ruling out differences in absorbance. The only difference was surface roughness – the ruler markings were rougher than the surrounding metal.

Silicone rubber compound, a type of glue/sealant was found to be non-transmissive at MIR wavelengths, as was 'Nivea moisturising sun spray' (sun block). Contrast between areas with and without the substances, when applied on substrates of aluminium and paper, was observed.

One of the aims of reflection imaging here is to image live, warm tissue. To take a step towards that, silicone and sun block was also imaged, with an arm as the substrate. This marks the first step towards *in vivo* diffuse reflection imaging in the MIR, with contrast arising from differences at the topmost surface of the skin. The results are shown in Figure 5-12. In both images, the substance is in the dark corner of the image, and it can be seen that the regions with substance applied to the skin do not diffusely reflect as much as pure skin. The images appear very 'mottled' – this is due to the uneven surface provided by skin.

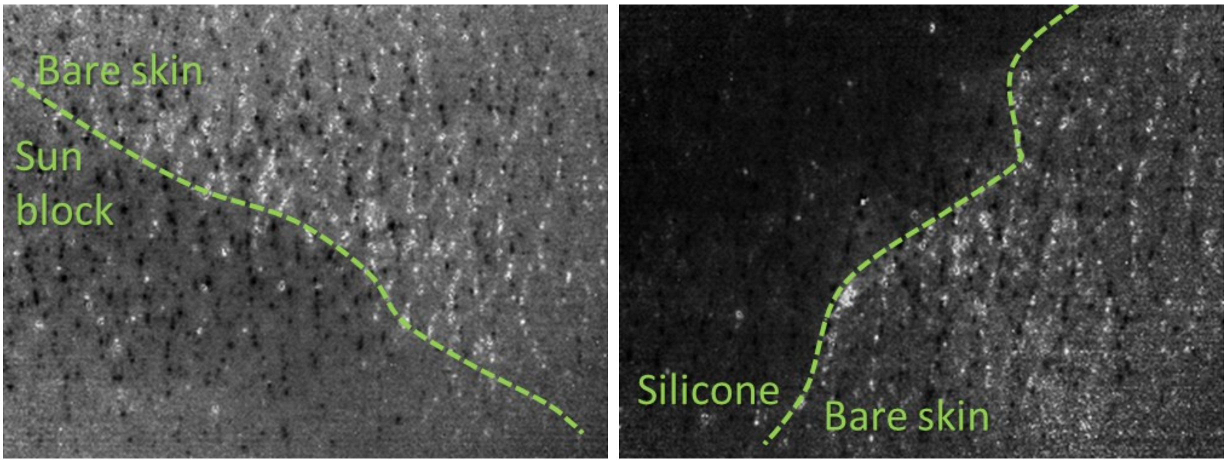


Figure 5-12. Reflection images of silicone rubber compound on a person's arm (left, silicone is in top right corner) and sun block on a person's arm (right, sun block is in bottom left corner). The images are an average of 100 images, taken at 1250cm^{-1}*. As previously highlighted, the contrast may not be solely due to differences in absorption. These two images have been ratioed against an area of arm with no substance applied, but the problem of nonidentical surfaces between each image of the ratio is evident. The texture of the arm comes through in the bright spots scattered across the images. The dark spots are where the bright spots were in the image with no substances. If exactly the same areas were imaged, the bright and dark spots should not be visible. The area being imaged is approximately 1cm square.*

Even with images that have contrast, we cannot confidently assign the contrast to absorption of the infrared wavelengths. The silicone and sun block appear shiny in the visible, with a large specular reflection component. This may also be the case in the MIR, meaning the contrast seen could be mainly due to the specular reflection from the shinier substance. This effect was also apparent when looking at Parma ham, in which the fat is 'shinier' that the meat, and so is less able to diffusely reflect light, appearing dark in reflection images. Although fat and meaty sections should have different absorption spectra in the MIR, the contrast seen was again attributed solely to the surface roughness of the sample.

5.2.2.1 5933 oesophagus sample

To ascertain that contrast comes from differences in absorption due to chemical differences in the samples, a familiar sample that has been confirmed to be suitable for use in a transmission setup was used – the thin 5933 oesophagus samples (see section 4.3.2). This sample typically has between 20% and 70% transmission at the wavelengths we are interested, with much of the rest of the light presumably being absorbed. Only a fraction of the light should have a chance of being reflected. Furthermore, the sample is only approximately 5µm in thickness, allowing for relatively little space for a random walk before the light can exit the sample again. However, it is a familiar sample that is known to display differences in contrast between regions, so serves well as a sample for proof of principle.

Figure 5-13 shows the 5933 sample viewed in reflection mode at 1234cm^{-1} (the phosphodiester peak location). The structure is visible, and we can see contrast between different areas of the sample. Note that the boundary of contrast in the reflection image is not at the edge of the epithelium – the *digistain* image (bottom image in Figure 5-13) shows the boundary of the contrast, which is clearly beyond the epithelium.

Although there is contrast in the reflection image of Figure 5-13, when we refer back to the images we saw with the transmission setup, the image shows a situation contrary to what would be expected if absorption were taking place to produce contrast. In the upper image of Figure 5-13, the epithelium and the tissue immediately surrounding it is brighter, so one might assume that it is absorbing less. When we look at the transmission setup images (lower image of Figure 5-13) however, it is apparent that the epithelium generally absorbs on par with the tissue surrounding it, and is the case across all MIR wavelengths that we can investigate. In general, the epithelium and the tissue immediately surrounding it absorbs more than the rest of the tissue that is in the field of view of the transmission image. As for the reason for this apparent reversal of what we expect to see, the contrast is again attributed to differences in surface roughness. The image in Figure 5-13 was taken with a diffuser in place to remove the spatial coherence, but the image still has small bright spots. It is thought that each of the bright points is in fact a specular reflection, and the epithelium appears brighter because it is rougher, the effect described in sections 3.3.3.1 and 5.2.2. If that is the case, the specular reflection from the rough surface is overwhelming any real diffuse reflection that may be present and producing a different contrast image. The picture does not allow us to distinguish contrast between the regions within the epithelium that have been classified as normal tissue and high grade dysplasia.

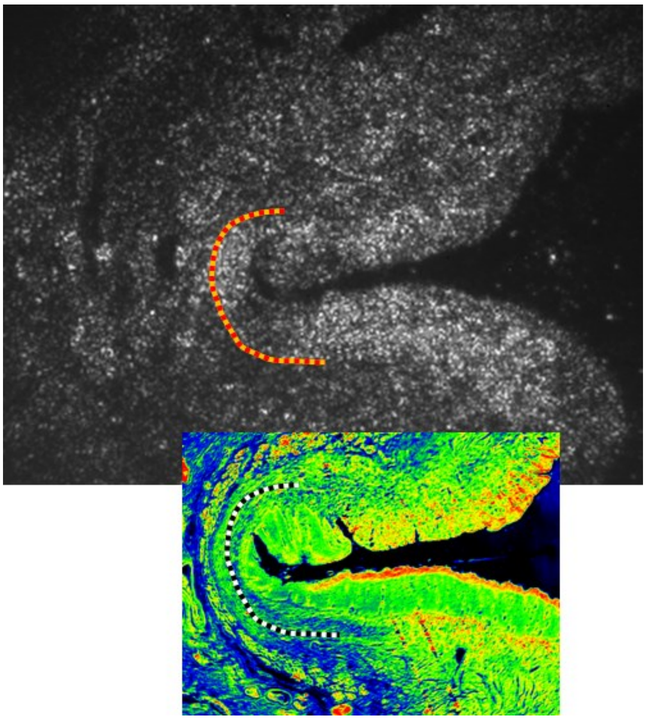


Figure 5-13. Oesophagus biopsy 5933 viewed at ~1234cm^{-1} (phosphodiester peak) with the QCL in reflection mode (top) and a corresponding absorbance image from the digistain (bottom, taken in a transmission setup). The images have been scaled and marked to show the matching regions and the contrast boundary, denoted by the dashed lines. The reflection image shows greatest intensity in the epithelium and the tissue immediately surrounding the epithelium, with no discernible contrast between the areas of normal and high grade dysplasia tissue. This would suggest at least that the epithelium is absorbing less than its surrounding tissue. However, the digistain absorbance image shows that on average the epithelium and the tissue immediately surrounding it appear to have higher absorbance than the rest of the tissue being imaged. This suggests that the contrast is not due to absorption within the sample, but instead only differences in surface roughness.

5.2.2.2 *In vivo* imaging of mole on skin

A long term vision of the reflection studies is to apply it to the imaging of skin cancer. When a patient has a melanoma, a type of skin cancer, the tissue immediately surrounding it may be at a cancerous stage also without exhibiting any visual differences from normal skin. To ensure the entire cancer is removed, a surgeon will often remove more than is necessary. If it were possible to image how large the cancer actually is, the surgeon could remove only the areas they wanted to and no more, reducing cosmetic damage to the patient and possibly reducing the risk of infection.

As melanomas often develop from moles, it was attempted to image a mole on skin, something that has never been done in the MIR before. If imaging a melanoma, we could look at the amide phosphodiester ratio again, or only the phosphodiester, but in a normal (benign) mole, the only difference from normal skin should be the higher amount of melanin present. There are different types of melanin, with eumelanin and pheomelanin contributing to the visual distinction of moles from normal skin.

Based on the literature on moles and related conditions, there should be “an appreciable amount” of melanin present in the stratum corneum (the uppermost layer of skin) at the location of moles relative to skin without an increased number of melanocytes (the cells that produce melanin and distribute it to surrounding cells) [125, 126]. This is an important assumption, as it leads us to believe that differences in melanin concentration should be detectable in the top layer of the skin in and around a mole. Melanocytes are usually concentrated at the base of the epidermis, suggesting melanin should be distributed throughout the entire epidermis.

Eumelanin and pheomelanin has absorption peaks at about 1380cm^{-1}, 1580cm^{-1} and 1490cm^{-1}, although the peaks are so broad that imaging at any wavenumber between ~1200-1600cm^{-1} should show some contrast with normal skin [127, 128]. There should be relatively low absorption from molecular vibrations present in tissue at 1380cm^{-1} compared with the other peaks, making this spectral location a good candidate wavenumber for the investigation of melanin. Additionally, absorption due to water is lowest (of the three peak locations) at 1380cm^{-1}, which is an important consideration when dealing with live tissue instead of the dehydrated cancer biopsies investigated in the transmission setup. For these reasons, 1380cm^{-1} was the wavenumber chosen at which to image skin and moles.

Attempts were made to acquire an image that showed contrast that could be attributed to the presence of melanin in a mole, but no contrast was seen at 1380cm^{-1}. There are a few possible reasons that no contrast could be seen. Since we are looking at a warm object, there will be a considerable bright thermal background that could be interfering. As mentioned in section 3.3.3.3, we can use the *digistain* filters to block out the thermal background to improve image quality. There are no filters available for the 1380cm^{-1} peak, but the Amide 'base' filter used in the *digistain* system (centred on 1598cm^{-1}) will work for the 1580cm^{-1} peak. With the *digistain* filter, there was still no contrast to be seen when imaging a patch of skin with a mole. It is interesting to note that the resultant reflection image was not noticeably improved in quality from before the filter was being used, suggesting that the background correction procedure was working well.

Another reason has been touched on already in sections 3.3.3.1 and 5.2.2 – there may be contrast in the diffuse reflection there, but the diffuse signal is being overwhelmed by the points of specular reflection that arise from a surface that is uneven on a short length scale. This issue may be removed in the future, but a larger problem here may be to do with how far through and how deep into the skin the diffuse signal that eventually reaches the camera actually penetrates.

Even from travelling through only the 10-20μm thick stratum corneum with relatively low water content [129], and assuming no scattering, only 80% of the light entering the skin could possibly be detected exiting the skin due to water absorption. There have been studies that have reached the conclusion that using MIR QCLs with higher powers in a reflection setup, reflected light can be detected exiting the skin after reaching the dermis layer in porcine and human skin [130, 131], but in these studies the maximum detected power was still less than 0.1% of the input power.

Due to specular reflection components reaching the detector, scattering of the light and absorption from water and tissue, the resultant SNR of the diffusely reflected signal will be too low to see contrast from the presence of a higher concentration of melanin. The SNR can be increased by increasing the acquisition time, although without even a weak signal to work with and the fact that increasing the acquisition time increases the time that a person would have to sit in one place keep their arm still starts to make this approach impractical. The current setup coupled with a higher-power laser however may yet yield the images being sought. The current time-averaged power produced by the QCL is 5mW – the peak power can be up to 120mW. The maximum permissible exposure (MPE, at the cornea of the human eye) in the MIR wavelengths for a 100ns pulse with the same spot size as the QCL is a pulse power of 19.6MW [132]. Operating at the same repetition rate as the QCL, the time-averaged power would need to be ~350W. This is at least 4 orders of magnitude greater than what the QCL can currently supply.

6 Conclusions

This section presents an overview of the work in this thesis, highlighting the achievements of the *digistain* system and the laser-based system. Following that are suggestions for future work, focusing on the next steps that could be taken to improve the systems, and reiterating the immediate barriers that still have to be overcome in their development.

6.1 Overview of thesis

In this thesis different avenues of spectral imaging in the MIR have been investigated, with distinct systems developed. The *digistain* system has been shown to be able to distinguish between different tissue types using only four wavelengths. It has also been shown to distinguish between healthy tissue and tissue in a very early (pre-) cancerous stage for tissue of the same type. The system compares favourably to other systems that may be available – it is magnitudes quicker than a Raman scanning system, and also faster than a commercial FTIR system with an FPA detector. The *digistain* system can be faster than other systems, as it has a very narrow focus for the applications it can be used for – it is designed solely for the imaging of tissue biopsies, and cancer diagnosis, and so only requires four spectral positions to be probed. In contrast, the other two systems mentioned will acquire full spectra, which will give them greater flexibility. The narrow scope of the *digistain* system is one of its strengths, allowing for good quality images and sufficiently low-noise data using relatively simple components.

The laser-based system using a tuneable QCL, operating in a transmission mode, has also been shown to be able to distinguish between healthy and pre-cancerous tissue of the same tissue type, using the same process as the *digistain*. This was only possible after removing the spatial coherence of the QCL and being able to remove (or at least largely reduce) the nonlinearity of the detector response to the QCL radiation. The process of removing the spatial coherence, using a moving diffuser, has a negative effect on the SNR of the resulting data. Mainly because of this reduction in the SNR, the QCL system does not perform at a better SNR than the *digistain* system despite using hardware that should have increased sensitivity and power output. The laser system is however fully flexible, as the wavelengths that can be probed are continuously tuneable.

Spectra of known materials have been taken with the QCL system. The spectra show good agreement in shape and position of peaks with spectra taken using a commercial FTIR system. However, there can be slight differences in quantitative data between the two spectra at some wavelengths, indicating that the nonlinear response of the detector to the laser light has not been fully eradicated. Once the nonlinearity has been confirmed to be completely removed, the system can be used as a reliable, automated spectral imaging system that can be used to record quantitative data.

Working in a diffuse reflection imaging mode, the QCL system has been demonstrated to be able to show contrast between surfaces that have substances of differing chemical composition applied. These surfaces include metal, paper and live human skin, a first for the MIR region of the spectrum. However, it is not certain that the contrast seen is due to the differences in chemical makeup causing absorption, or simply variances in surface roughness that cause differences in the way the light scatters and is reflected. A tissue sample with a known absorption distribution was investigated, which confirmed the chemical absorption was not the main mechanism behind the contrast seen.

Human skin was probed in the reflection mode, as a sample which largely has uniform surface roughness in certain areas. Any contrast that may be observed would be due to absorption due to the differing molecules that lie below the surface of the skin. No contrast was observed, which was attributed to the high water absorption in live tissue attenuating the signal. Live tissue is not the ideal test sample to test a fresh system in a field with no known publications on similar setups that look for chemical differences in tissue in the MIR.

6.2 Suggestions for further work

The *digistain* system may be improved through incremental improvements in the SNR of the system. This may come about from upgrading the components used (e.g. a brighter thermal source, using bandpass filters with higher peak transmissions) or tweaking the setup. This will reduce the time required to acquire data.

The more immediate goal is to be able to show a correlation between the phosphodiester-to-amide (PA) ratio and histological grade and, later, patient survival rate. This process is already currently underway, with a large number of tissue sections being studied to gain statistically significant results. The issue still remains of separating out different tissue types when they are mixed on short length scales. This may be solved by methods such as the one described in section 5.1.2.1, by which histogram ‘peaks’ can be sorted out after some calibration of the peak positions and shapes has occurred. If this does not prove possible, other bandpass filters may be used to utilise other spectral features, such as those mentioned in section 2.4.2 to better label tissue types. It is not known yet what the differences between tissue types at other wavelengths may be. There are also two other images that may yet be used to help distinguish tissue types – the baseline-corrected phosphodiester and amide peak images, the numerator and denominator of equation 3-3. Using one or both of these images may lead to the tissue types being better separated, although the thickness-dependence of the images could lead to images being dealt with on a case-by-case basis.

The *digistain* technique is proving capable when investigating cancer biopsies, but it need not be limited to only cancer, possibly also being successful at investigating other samples and biopsies, perhaps by choosing different bandpass filters.

The QCL system still has large room for improvement, thanks to the spatial coherence of the source and the nonlinear response of the detector. Using a QCL that can provide a CW output should improve the nonlinearity problem such that it is no longer an issue that needs to be considered. As already stated, the largest source of noise is the diffuser that is used to remove the spatial coherence. Although it was the only implementation found to remove coherence, another way not yet tested may provide improvements. The diffuser is currently imaged in the sample plane, requiring a long averaging period to build up a repeatable uniform illumination. If the diffuser can be placed such that its image is not in the sample plane, whilst still removing coherence, the SNR of the system should increase considerably. Alternatively, a system that does not use a diffuser at all may be possible.

These improvements apply to the transmission setup. In the reflection setup, the biggest obstacle to overcome is how to deal with differences in surface roughness. It is this that limits the type of samples that can be looked at with diffuse reflection imaging, as the sample should have uniform surface roughness, and ideally be relatively smooth and flat. Live biological samples in particular seem particularly ill-suited for investigation because of this, along with the high water content causing attenuation. Aside from choosing different samples to investigate, it may be worth looking at specular components of reflection. This approach would move away from determining solely the chemical absorption, as other factors (such as the refractive index) will affect the contrast, but in many cases will supply a greater signal to work with.

Bibliography

1. Smith, B.C., Fundamentals of Fourier Transform Infrared Spectroscopy, Second Edition. 2011: Taylor & Francis.

2. Phillips, M.C. and N. Ho, Infrared hyperspectral imaging using a broadly tunable external cavity quantum cascade laser and microbolometer focal plane array. Opt. Express, 2008. 16(3): p. 1836-1845.

3. Fellgett, P.B., On the Ultimate Sensitivity and Practical Performance of Radiation Detectors. Journal of the Optical Society of America, 1949. 39(11): p. 970-976.

4. Yeh, K., et al., Fast Infrared Chemical Imaging with a Quantum Cascade Laser. Analytical Chemistry, 2014. 87(1): p. 485-493.

5. Jensen, P., et al., Observer variability in the assessment of type and dysplasia of colorectal adenomas, analyzed using kappa statistics. Diseases of the Colon & Rectum, 1995. 38(2): p. 195-198.

6. Kendall, C., et al., Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus. The Journal of Pathology, 2003. 200(5): p. 602-609.

7. Melville, D.M., et al., Observer study of the grading of dysplasia in ulcerative colitis: Comparison with clinical outcome. Human Pathology, 1989. 20(10): p. 1008-1014.

8. Agarwal, V. and P. O’Neill, Adjuvant! Online as a Decision-making Tool in Early Breast Cancer — a UK National Survey. Clinical Oncology, 2011. 23(2): p. 159-160.

9. Oratz, R., et al., Impact of a Commercial Reference Laboratory Test Recurrence Score on Decision Making in Early-Stage Breast Cancer. Journal of Oncology Practice, 2007. 3(4): p. 182-186.

10. Why is early diagnosis important? ; Available from: http://www.cancerresearchuk.org/cancer-info/spotcancerearly/key-signs-and-symptoms/whyisearlydiagnosisimportant/.

11. Early detection of cancer. Available from: http://www.who.int/cancer/detection/en/.

12. Knowles, M. and P. Selby, Introduction to the Cellular and Molecular Biology of Cancer. 2005: OUP Oxford.

13. Bohunicky, B. and S.A. Mousa, Biosensors: the new wave in cancer diagnosis. Nanotechnology, Science and Applications, 2011. 4: p. 1-10.

14. Guck, J., et al., Optical Deformability as an Inherent Cell Marker for Testing Malignant Transformation and Metastatic Competence. Biophysical Journal, 2005. 88(5): p. 3689-3698.

15. Le, N.D.B., M. Yazdani, and V.M. Rotello, Array-based sensing using nanoparticles: an alternative approach for cancer diagnostics. Nanomedicine, 2014. 9(10): p. 1487-1498.

16. Iorio, M.V. and C.M. Croce, MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. Vol. 4. 2012. 143-159.

17. Huang, X. and M.A. El-Sayed, Gold nanoparticles: Optical properties and implementations in cancer diagnosis and photothermal therapy. Journal of Advanced Research, 2010. 1(1): p. 13-28.

18. Pansare, V.J., et al., Review of Long-Wavelength Optical and NIR Imaging Materials: Contrast Agents, Fluorophores, and Multifunctional Nano Carriers. Chemistry of Materials, 2012. 24(5): p. 812-827.

19. Marcott, C., G.M. Story, and R.K. Dukor, Infrared Spectral Imaging of H&E-Stained Breast Tissue Biopsies. Microscopy and Microanalysis, 2004. 10(SupplementS02): p. 182-183.

20. Zhang, L., et al., Tryptophan as the fingerprint for distinguishing aggressiveness among breast cancer cell lines using native fluorescence spectroscopy. Journal of Biomedical Optics, 2014. 19(3): p. 037005-037005.

21. Veta, M., et al., Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. Mod Pathol, 2012. 25(12): p. 1559-1565.

22. Veta, M., et al., Automatic Nuclei Segmentation in H&E Stained Breast Cancer Histopathology Images. PLoS ONE, 2013. 8(7): p. e70221.

23. Cancer Fact sheet N°297 Updated February 2014. Available from: http://www.who.int/mediacentre/factsheets/fs297/en/index.html.

24. All cancers combined Key Facts. Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/Allcancerscombined/.

25. Amrania, H., Ultrafast Mid-Infrared Spectroscopic Imaging with Biomedical Applications. 2009, Imperial College London.

26. McCrow, A., Laser based mid-infrared spectroscopic imaging: exploring a novel method for application in cancer diagnosis. 2011, Imperial College London.

27. Amrania, H., et al., Digistain: a digital staining instrument for histopathology. Opt. Express, 2012. 20(7): p. 7290-7299.

28. Lloyd, G.R., et al., Histological imaging of a human colon polyp sample using Raman spectroscopy and self organising maps. Vibrational Spectroscopy, 2012. 60(0): p. 43-49.

29. Turgeon, M.L., Clinical hematology: theory and procedures. Vol. 936. 2005: Lippincott Williams & Wilkins.

30. Salzer, R. and H.W. Siesler, Infrared and Raman Spectroscopic Imaging. 2009: Wiley.

31. Cameron, D.G. and D.J. Moffatt, Deconvolution, derivation, and smoothing of spectra using Fourier transforms. Journal of Testing and Evaluation, 1984. 12(2): p. 78-85.

32. Shaner, N.C., P.A. Steinbach, and R.Y. Tsien, A guide to choosing fluorescent proteins. Nat Meth, 2005. 2(12): p. 905-909.

33. Swift, S. and L. Trinkle-Mulcahy. Basic principles of FRAP, FLIM and FRET.

34. Chang, C.-W., et al., Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. Soil Science Society of America Journal, 2001. 65(2): p. 480-490.

35. Rehman, I., Z. Movasaghi, and S. Rehman, Vibrational Spectroscopy for Tissue Analysis. 2012: Taylor & Francis.

36. Lewis, I.R. and H. Edwards, Handbook of Raman spectroscopy: from the research laboratory to the process line. 2001: CRC Press.

37. Björn, L.O., Photobiology: the science of life and light. 2007: Springer.

38. Martens, H. and E. Stark, Extended multiplicative signal correction and spectral interference subtraction: New preprocessing methods for near infrared spectroscopy. Journal of Pharmaceutical and Biomedical Analysis, 1991. 9(8): p. 625-635.

39. Bassan, P., et al., The inherent problem of transflection-mode infrared spectroscopic microscopy and the ramifications for biomedical single point and imaging applications. Analyst, 2013. 138(1): p. 144-157.

40. Brooke, H., et al., A Study of Electric Field Standing Waves on Reflection Microspectroscopy of Polystyrene Particles. Applied Spectroscopy, 2009. 63(11): p. 1293-1302.

41. Filik, J., et al., Electric field standing wave artefacts in FTIR micro-spectroscopy of biological materials. Analyst, 2012. 137(4): p. 853-861.

42. Warren, J.S., Modern optical engineering: the design of optical systems. McGraw-Hill, Inc., Figure, 1990: p. 228, 256.

43. Hanrahan, P. and W. Krueger, Reflection from layered surfaces due to subsurface scattering, in Proceedings of the 20th annual conference on Computer graphics and interactive techniques. 1993, ACM: Anaheim, CA. p. 165-174.

44. Tagare, H.D., A theory of photometric stereo for a class of diffuse non-Lambertian surfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991. 13(2): p. 133-152.

45. Siegel, R., Thermal radiation heat transfer. Vol. 1. 2001: CRC press.

46. Maziak, D.E., et al., Fourier-transform infrared spectroscopic study of characteristic molecular structure in cancer cells of esophagus: An exploratory study. Cancer detection and prevention, 2007. 31(3): p. 244-253.

47. Arnaud, C.H., Quantum Cascade Lasers Push Infrared Imaging Closer To The Clinic, in Chemical & Engineering News. 2015, ACS. p. 25-26.

48. Pereira, M.F. and O. Shulika, Terahertz and Mid Infrared Radiation. Terahertz and Mid Infrared Radiation: Generation, Detection and Applications, 2011. 1.

49. Cathabard, O., et al., Quantum cascade lasers emitting near 2.6µm. Applied Physics Letters, 2010. 96(14): p. 141110.

50. Walther, C., et al., Quantum cascade lasers operating from 1.2to1.6THz. Applied Physics Letters, 2007. 91(13): p. 131122.

51. Faist, J., Quantum Cascade Lasers. 2013: OUP Oxford.

52. Welzel, S., New Enhanced Sensitivity Infrared Laser Spectroscopy Techniques Applied to Reactive Plasmas and Trace Gas Detection. 2009: Logos Verlag Berlin.

53. Vodopyanov, K.L., Parametric generation of tunable infrared radiation in ZnGeP2 and GaSe pumped at 3 µm. Journal of the Optical Society of America B, 1993. 10(9): p. 1723-1729.

54. Vodopyanov, K.L., et al., ZnGeP2 optical parametric oscillator with 3.8—12.4-µm tunability. Optics Letters, 2000. 25(11): p. 841-843.

55. Bhargava, R. and I.W. Levin, Spectrochemical analysis using infrared multichannel detectors. 2008: Wiley. com.

56. Rao, K.N., Spectroscopy of the earth's Atmosphere and interstellar Medium. 2012: Elsevier Science.

57. DigiTect detectors for Bruker Optics VERTEX and TENSOR series FT-IR-specrtometers. Available from: http://www.bruker.com/products/infrared-near-infrared-and-raman-spectroscopy/opus-software/downloads.html.

58. Rogalski, A., Infrared Detectors, Second Edition. 2010: Taylor & Francis.

59. Kruse, P.W., Uncooled Thermal Imaging: Arrays, Systems, and Applications. 2001: Society of Photo Optical.

60. Fraden, J., Handbook of Modern Sensors: Physics, Designs, and Applications. 2010: Springer.

61. Williams, T., Thermal Imaging Cameras: Characteristics and Performance. 2009: CRC Press.

62. Razeghi, M., Fundamentals of Solid State Engineering. 2009: Springer.

63. Mann, A., Infrared Optics and Zoom Lenses. 2009: Society of Photo Optical.

64. Riedl, M.J., Optical design fundamentals for infrared systems. Vol. 48. 2001: SPIE press.

65. Braun, K., et al., Novel Parabolic Mirror Microscope Illuminated with Cylindrical Vector Beams for Confocal and Tip Enhanced Super Resolution Imaging, in Fringe 2013, W. Osten, Editor. 2014, Springer Berlin Heidelberg. p. 791-796.

66. McKechnie, T.S., Image-plane speckle in partially coherent illumination. Optical and Quantum Electronics, 1976. 8(1): p. 61-67.

67. Lowenthal, S. and D. Joyeux, Speckle Removal by a Slowly Moving Diffuser Associated with a Motionless Diffuser. Journal of the Optical Society of America, 1971. 61(7): p. 847-851.

68. Kubota, S. and J.W. Goodman, Very efficient speckle contrast reduction realized by moving diffuser device. Applied Optics, 2010. 49(23): p. 4385-4391.

69. Akram, M.N., et al., Laser speckle reduction due to spatial and angular diversity introduced by fast scanning micromirror. Appl. Opt., 2010. 49(17): p. 3297-3304.

70. Mattheyses, A.L., K. Shaw, and D. Axelrod, Effective elimination of laser interference fringing in fluorescence microscopy by spinning azimuthal incidence angle. Microscopy Research and Technique, 2006. 69(8): p. 642-647.

71. Ouyang, G., et al., Speckle reduction using a motionless diffractive optical element. Opt. Lett., 2010. 35(17): p. 2852-2854.

72. Seinfeld, J.H. and S.N. Pandis, Atmospheric chemistry and physics: from air pollution to climate change. 2006: Wiley.

73. Allewell, N.M., L.O. Narhi, and I. Rayment, Molecular Biophysics for the Life Sciences. 2013: Springer New York.

74. Lewerenz, H.J., Photons in Natural and Life Sciences: An Interdisciplinary Approach. 2012: Springer.

75. Michalet, X. and S. Weiss, Using photon statistics to boost microscopy resolution. Proceedings of the National Academy of Sciences of the United States of America, 2006. 103(13): p. 4797-4798.

76. Wayne, R.O., Light and Video Microscopy. 2013: Elsevier Science.

77. Alberts, B., Molecular biology of the cell. 2000: Garland Science.

78. Busch, H., Chromatin. 2013: Elsevier Science.

79. Alberts, B., et al., Membrane Proteins. 2002.

80. Cheung, M.C., et al., Intracellular protein and nucleic acid measured in eight cell types using deep-ultraviolet mass mapping. Cytometry Part A, 2013. 83A(6): p. 540-551.

81. Cheung, M.C., et al., Deep ultraviolet mapping of intracellular protein and nucleic acid in femtograms per pixel. Cytometry Part A, 2011. 79A(11): p. 920-932.

82. Hübner, W. and H.H. Mantsch, Orientation of specifically 13C=O labeled phosphatidylcholine multilayers from polarized attenuated total reflection FT-IR spectroscopy. Biophysical Journal, 1991. 59(6): p. 1261-1272.

83. Hayashi, T. and S. Mukamel, Vibrational-exciton couplings for the amide I, II, III, and A modes of peptides. The Journal of Physical Chemistry B, 2007. 111(37): p. 11032-11046.

84. Kong, J. and S. Yu, Fourier transform infrared spectroscopic analysis of protein secondary structures. Acta Biochimica et Biophysica Sinica, 2007. 39(8): p. 549-559.

85. Susi, H. and D.M. Byler, [13] Resolution-enhanced fourier transform infrared spectroscopy of enzymes, in Methods in Enzymology, S.N.T. C. H. W. Hirs, Editor. 1986, Academic Press. p. 290-311.

86. Hale, G.M. and M.R. Querry, Optical Constants of Water in the 200-nm to 200-μm Wavelength Region. Applied Optics, 1973. 12(3): p. 555-563.

87. Driggers, R.G., M.H. Friedman, and J. Nichols, Introduction to Infrared and Electro-optical Systems. 2012: Artech House.

88. Siegman, A.E., Lasers. 1986: University Science Books.

89. Miller, A.B., et al., Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ, 2014. 348.

90. Kaplan, C., et al., Willingness to use tamoxifen to prevent breast cancer among diverse women. Breast Cancer Research and Treatment, 2012. 133(1): p. 357-366.

91. Wang, D.-Y., et al., Identification of Estrogen-Responsive Genes by Complementary Deoxyribonucleic Acid Microarray and Characterization of a Novel Early Estrogen-Induced Gene: EEIG1. Molecular Endocrinology, 2004. 18(2): p. 402-411.

92. Desta, Z., et al., Comprehensive Evaluation of Tamoxifen Sequential Biotransformation by the Human Cytochrome P450 System in Vitro: Prominent Roles for CYP3A and CYP2D6. Journal of Pharmacology and Experimental Therapeutics, 2004. 310(3): p. 1062-1075.

93. Grilli, S., Tamoxifen (TAM): the dispute goes on. Annal-Istituto Superiore di Sanita, 2006. 42(2): p. 170.

94. Vehmanen, L., et al., Tamoxifen Treatment After Adjuvant Chemotherapy Has Opposite Effects on Bone Mineral Density in Premenopausal Patients Depending on Menstrual Status. Journal of Clinical Oncology, 2006. 24(4): p. 675-680.

95. Vogel, V.G., et al., Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes: The nsabp study of tamoxifen and raloxifene (star) p-2 trial. JAMA, 2006. 295(23): p. 2727-2741.

96. Breast Cancer Tests: Screening, Diagnosis, and Monitoring Available from: http://www.breastcancer.org/symptoms/testing/types.

97. Old, O.J., et al., Vibrational spectroscopy for cancer diagnostics. Analytical Methods, 2014. 6(12): p. 3901-3917.

98. Teh, S.K., et al., Diagnostic potential of near-infrared Raman spectroscopy in the stomach: differentiating dysplasia from normal tissue. Br J Cancer, 2008. 98(2): p. 457-465.

99. Molckovsky, A., et al., Diagnostic potential of near-infrared Raman spectroscopy in the colon: Differentiating adenomatous from hyperplastic polyps. Gastrointestinal Endoscopy, 2003. 57(3): p. 396-402.

100. Quaroni, L. and A.G. Casson, Characterization of Barrett esophagus and esophageal adenocarcinoma by Fourier-transform infrared microscopy. Analyst, 2009. 134(6): p. 1240-1246.

101. Wang, T.D., et al., Detection of endogenous biomolecules in Barrett's esophagus by Fourier transform infrared spectroscopy. Proceedings of the National Academy of Sciences, 2007. 104(40): p. 15864-15869.

102. Li, Q.-B., et al., Diagnosis of gastric inflammation and malignancy in endoscopic biopsies based on Fourier transform infrared spectroscopy. Clinical Chemistry, 2005. 51(2): p. 346-350.

103. Khanmohammadi, M., et al., Application of Linear Discriminant Analysis and Attenuated Total Reflectance Fourier Transform Infrared Microspectroscopy for Diagnosis of Colon Cancer. Pathology & Oncology Research, 2011. 17(2): p. 435-441.

104. Xie, Y., et al., Diagnosis of colon cancer with Fourier transform infrared spectroscopy on the malignant colon tissue samples. Chinese medical journal, 2011. 124(16): p. 2517-2521.

105. Bird, B., et al., Detection of breast micro-metastases in axillary lymph nodes by infrared micro-spectral imaging. Analyst, 2009. 134(6): p. 1067-1076.

106. Isabelle, M., et al., Lymph node pathology using optical spectroscopy in cancer diagnostics. Spectroscopy, 2008. 22: p. 97-104.

107. Nallala, J., et al., Infrared spectral imaging as a novel approach for histopathological recognition in colon cancer diagnosis. Journal of Biomedical Optics, 2012. 17(11): p. 116013-116013.

108. Kröger, N., et al., Quantum cascade laser–based hyperspectral imaging of biological tissue. Journal of Biomedical Optics, 2014. 19(11): p. 111607-111607.

109. Bassan, P., et al., Large scale infrared imaging of tissue micro arrays (TMAs) using a tunable Quantum Cascade Laser (QCL) based microscope. Analyst, 2014. 139(16): p. 3856-3859.

110. Clemens, G., et al., Quantum cascade laser-based mid-infrared spectrochemical imaging of tissues and biofluids.

111. Goyal, A., et al., Active hyperspectral imaging using a quantum cascade laser (QCL) array and digital-pixel focal plane array (DFPA) camera. Optics Express, 2014. 22(12): p. 14392-14401.

112. New Infrared Source IR-30 Series. Available from: http://www.scitec.uk.com/irsources/ir-30.

113. Kole, M.R., et al., Discrete frequency infrared microspectroscopy and imaging with a tunable quantum cascade laser. Analytical chemistry, 2012. 84(23): p. 10366-10372.

114. Bjuggren, M., L. Krummenacher, and L. Mattsson, Infrared diffuser for application in transmission and reflection at 10.6 µm. Applied Optics, 1997. 36(4): p. 940-948.

115. Giglio, M., S. Musazzi, and U. Perini, Distance measurement from a moving object based on speckle velocity detection. Applied Optics, 1981. 20(5): p. 721-722.

116. Russ, J.C., The Image Processing Handbook, Sixth Edition. 2011: CRC Press.

117. Wheater, P., et al., Basic Histopathology—A Colour Atlas and Text. 1985: Churchill Livingstone.

118. Wiseman, B.S. and Z. Werb, Stromal Effects on Mammary Gland Development and Breast Cancer. Science, 2002. 296(5570): p. 1046-1049.

119. Kuo, B. and D. Urma, Esophagus-anatomy and development. GI Motility online, 2006.

120. Frkovic-Grazio, S. and M. Bracko, Long term prognostic value of Nottingham histological grade and its components in early (pT1N0M0) breast carcinoma. Journal of Clinical Pathology, 2002. 55(2): p. 88-92.

121. Judd, S.J., Breast Cancer Sourcebook: Basic Consumer Health Information about Breast Cancer. 2004: Omnigraphics, Incorporated.

122. Tobin, M.J., et al., FTIR spectroscopy of single live cells in aqueous media by synchrotron IR microscopy using microfabricated sample holders. Vibrational Spectroscopy, 2010. 53(1): p. 34-38.

123. FTIR Microscopes & Imaging Systems — Cary 620 FTIR Microscopes. Available from: http://www.chem.agilent.com/en-US/products-services/Instruments-Systems/Molecular-Spectroscopy/Cary-620-FTIR-Microscopes/Pages/default.aspx.

124. Highest available Signal-To-Noise performance, delivering superior sensitivity and analytical performance. 2011; Available from: http://www.chem.agilent.com/Library/technicaloverviews/Public/si-1353.pdf.

125. Cohen, L.M., et al., Hypermelanotic nevus: clinical, histopathologic, and ultrastructural features in 316 cases. The American journal of dermatopathology, 1997. 19(1): p. 23-30.

126. Stegmaier, O.C. and S.W.M. Becker Jr, Incidence of Melanocytic Nevi in Young Adults1. The Journal of Investigative Dermatology, 1960. 34(2): p. 125-129.

127. Galván, I., et al., Vibrational characterization of pheomelanin and trichochrome F by Raman spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2013. 110(0): p. 55-59.

128. Barden, H.E., et al., Morphological and geochemical evidence of eumelanin preservation in the feathers of the Early Cretaceous bird, Gansus yumenensis. PloS one, 2011. 6(10): p. e25494.

129. Kottmann, J., et al., Glucose sensing in human epidermis using mid-infrared photoacoustic detection. Biomedical Optics Express, 2012. 3(4): p. 667-680.

130. Liakat, S., et al., Mid-infrared (λ = 8.4–9.9 μm) light scattering from porcine tissue. Applied Physics Letters, 2012. 101(9): p. -.

131. Michel, A.P.M., et al., In vivo measurement of mid-infrared light scattering from human skin. Biomedical Optics Express, 2013. 4(4): p. 520-530.

132. Commission, I.E., IEC 60825-1,«. Safety of Laser Products—Part, 2001. 1.

Appendices

Appendix: summary of permission for third party copyright works

Page Number	Type of work: text, figure, map, etc.	Source work	Copyright holder & year	Work out of copyright	Permission to re-use	Permission requested	permission refused	Orphan work
Page 35	figure	Guide for Infrared Spectroscopy, reference [57]	© 2011 Bruker Optics			✓		
Page 43	figure	Applied Optics (2010), vol 49, pages 3297-3304	© 2010 Optical Society of America		✓			
Page 109	figure	Vibrational Spectroscopy (2012), vol 60, pages 43-49	© 2013 Elsevier B.V		✓			

Permission to use figures on page 109 can be found at the following URL:

https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=a40b420d-c654-4baa-a2e5-c76c12828b8e

Permission to use the figure on page 43 is displayed on the following pages.

Permission to use the figure on page 35 has been requested on 27th June 2014 and again on 11th July 2014 and has not been granted or refused. Proof of request is displayed on the following pages.

E-theses letter template: request to reproduce an extract from a third party's published work

27 Jun. 14

Dear OSA,

I am completing [or have completed] my PhD thesis at Imperial College London entitled ‘Absorption spectral imaging in the mid-infrared and its application in cancer diagnosis’.

I seek your permission to reprint, in my thesis an extract from Applied Optics, Vol. 49, Issue 17, pp. 3297-3304 (2010). The extract to be reproduced is: Figure 6 on page 3301.

I would like to include the extract in the printed examination copy of my thesis and also the electronic version which will be added to Spiral, Imperial's online repository http://spiral.imperial.ac.uk/ and made available to the public under a Creative Commons Attribution-NonCommercial-NoDerivs licence.

If you are happy to grant me all the permissions requested, please return a signed copy of this letter. If you wish to grant only some of the permissions requested, please list these and then sign.

Yours sincerely,

Laurence Drummond

Reply:

Dear Mr. Drummond,

Thank you for contacting The Optical Society.

OSA considers your requested use of its copyrighted material to be Fair Use under United States Copyright Law. It is requested that a complete citation of the original material be included in any publication.

Let me know if you have any questions.

Kind Regards,

Susannah Lehman

Email and attachment sent to email address info@brukeroptics.de :

E-theses letter template: request to reproduce an extract from a third party's published work

27 Jun. 14

Dear Bruker Optics,

I am completing my PhD thesis at Imperial College London entitled ‘Absorption spectral imaging in the mid-infrared and its application in cancer diagnosis’.

I seek your permission to reprint, in my thesis an extract from: Guide for Infrared Spectroscopy, 2011, available for download from http://www.bruker.com/products/infrared-near-infrared-and-raman-spectroscopy/opus-software/downloads.html . The extract to be reproduced is: Figure labelled ‘Detectors’ on page 7.

I would like to include the extract in the printed examination copy of my thesis and also the electronic version which will be added to Spiral, Imperial's online repository http://spiral.imperial.ac.uk/ and made available to the public under a Creative Commons Attribution-NonCommercial-NoDerivs licence.

If you are happy to grant me all the permissions requested, please return a signed copy of this letter. If you wish to grant only some of the permissions requested, please list these and then sign.

Yours sincerely,

Laurence Drummond

Acknowledgements

I would like to pay tribute to a number of people who helped me in their own way throughout the years.

First and foremost, I would like to thank Prof Chris Phillips, my supervisor for four years. It is difficult to put into words my gratitude towards him for all that he has done, and so I will just leave it at thank you.

A special thank you must go towards Dr Hemmel Amrania, whose insight into physics research I always appreciated, and was always willing to lend a hand.

I am grateful to my brothers in arms, Frank Murphy and Alex Bak, for maintaining a state of surrealism and general euphoria day in, day out in H901 as we worked to push back the frontiers of science.

Additional thanks go towards colleagues past and present – among them, Megumi Yoshida, Dr Ned Yoxall and Dr Andy McCrow.

Those at Imperial who played a similarly crucial role during my time there include the ladies in H902, Carolyn, Bhavna and Juraci, who kept us well watered and fed, and Martin Morris, who over the course of four years spent considerable time helping me to get various computers to work.

Further thanks go to Dr Matthew Baker and Prof Daniel Rueckert for their input in the editing of my thesis and agreeing to be my examiners.

I also wish to thank my parents, Andy and Annette, for seemingly having the confidence that I could go the distance.

Lastly, I wish to thank my wife Goeun for always being there for me, and on a more practical note, without whose computer I could not have finished writing my thesis.