Digital Signal Processing Techniques for Optical Coherence Tomography: Spectroscopic OCT and OCT Image Enhancement

by

Desmond Christopher Adler

B.Sc. Electrical Engineering
University of Alberta, Canada, 2001

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2004

© 2004 Massachusetts Institute of Technology. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of Author: ____________________________________________

Desmond C. Adler
Department of Electrical Engineering and Computer Science
May 20, 2004

Certified by: __________

James G. Fujimoto
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by: __________

Chairman, Committee on Graduate Students
Department of Electrical Engineering and Computer Science
Digital Signal Processing Techniques for Optical Coherence Tomography: Spectroscopic OCT and OCT Image Enhancement

by

Desmond Christopher Adler

Submitted to the Department of Electrical Engineering and Computer Science on May 20, 2004 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Electrical Engineering and Computer Science

Abstract

Digital signal processing (DSP) techniques were developed to improve the flexibility, functionality, and image quality of ultrahigh resolution optical coherence tomography (OCT) systems.

To reduce the dependence of OCT research systems on fixed analog electronics and to improve overall system flexibility, a digital demodulation scheme implemented entirely in software was developed. This improvement allowed rapid reconfiguration of the OCT imaging speed and source center wavelength without having to construct new analog filters and demodulators. This demodulation scheme produced a highly accurate envelope and was immune to local variations in carrier frequency.

To provide an alternative contrast modality to conventional intensity-based OCT imaging, spectroscopic OCT technology was investigated. Preliminary studies on animal models were carried out, with the ultimate goal of enabling the early detection of dysplastic lesions in epithelial tissue through spectroscopic changes not visible with conventional OCT. Various spectral analysis techniques were investigated and evaluated for their ability to provide enhanced contrast of specific tissue types. Areas of concern such as red-shifting of the spectrum with increasing imaging depth, Doppler shifts induced by the optical path length scanner, and determination of an optimal spectroscopic metric were addressed.

To improve the quality of ultrahigh resolution OCT images, wavelet processing techniques for speckle noise reduction were investigated. Spatially adaptive wavelet denoising techniques were compared to basic wavelet denoising techniques and time domain filtering. By using a set of image quality metrics, it was possible to quantify the effectiveness of the various filtering methods and determine an optimal process for removing speckle noise while maintaining feature sharpness.

Thesis Supervisor: James G. Fujimoto
Title: Professor of Electrical Engineering and Computer Science
Acknowledgements

I would like to thank my research advisor and thesis supervisor, Professor James Fujimoto, for providing the guidance, insight, and resources necessary for me to complete this work. I would like to thank Tony Ko for all of his help, suggestions, and support during my time at MIT. If it were not for his willingness to answer my endless questions and his selfless generosity with his time, I would not have been able to perform much of the work reported here. I would also like to thank Maciej Wojtkowski, Vikas Sharma, Pei-Lin Hsiung, Vivek Srinivasan, Yu Chen, Paul Herz, Andrew Kowalevicz, Aaron Aguirre, Robert Huber, Aurea Zare, and Norihiko Nishizawa for the many fruitful conversations in the lab.

I would like to gratefully acknowledge the Government of Canada, the province of Alberta, the National Science and Engineering Research Council of Canada, and the University of Alberta for providing me with the financial support and academic background necessary for a successful career as a graduate student. I hope that I will eventually be able to give back to my country some small portion of the education and opportunity that it has given to me.

Finally, I would like to thank my friends and my family from the bottom of my heart. You all stayed with me through one of the most challenging periods in my life, and have seen me through to the other side. I wouldn’t have made it without you.

For the northern lights, the hockey rinks, the mountains and the maple leaf
**TABLE OF CONTENTS**

ABSTRACT ................................................................................................................................... 3

ACKNOWLEDGEMENTS ........................................................................................................... 5

TABLE OF CONTENTS ......................................................................................................... 7

CHAPTER 1: INTRODUCTION ............................................................................................ 9
  1.1 Scope of Thesis ........................................................................................................ 9
  1.2 Background of Optical Coherence Tomography .................................................. 9
  1.3 Concept of Spectroscopic Optical Coherence Tomography ................................ 16

CHAPTER 2: APPLICATIONS OF DIGITAL SIGNAL PROCESSING IN OPTICAL COHERENCE TOMOGRAPHY ............................................................................................ 18
  2.1 Data Acquisition ....................................................................................................... 18
  2.2 Digital Demodulation .............................................................................................. 29

CHAPTER 3: SPECTROSCOPIC OPTICAL COHERENCE TOMOGRAPHY .......... 37
  3.1 Motivation for Spectroscopic OCT ........................................................................... 37
  3.2 Review of Mie Scattering Theory ........................................................................... 38
    3.2.1 General Scattering Problem ............................................................................ 39
    3.2.2 Vector Harmonic Expression of Plane Waves ............................................. 40
    3.2.3 Solution of the Scattered Field ....................................................................... 45
    3.2.4 Spatial Distribution of Scattered Light ............................................................ 48
    3.2.5 Scattering from Multiple Particles .................................................................... 52
  3.3 Visualization of Spectroscopic Information ............................................................. 53
    3.3.1 Calculation of Optical Spectra ......................................................................... 54
    3.3.2 Time / Frequency Resolution Using the Chirp Z Transform ......................... 59
    3.3.3 Pure Spectroscopic Visualization .................................................................... 60
    3.3.4 Spectroscopic / Intensity Visualization .......................................................... 63
    3.3.5 Spectroscopic / Greyscale Visualization ....................................................... 64
  3.4 Spectroscopic OCT Metrics ....................................................................................... 64
    3.4.1 Sources of Spectroscopic Noise ....................................................................... 65
    3.4.2 Center Wavelength ......................................................................................... 67
    3.4.3 Power Over Predefined Spectral Bandwidth .................................................. 68
    3.4.4 Metrics Sensitive to Spectral Modulation ....................................................... 71
    3.4.5 Gaussian Fit ..................................................................................................... 73
    3.4.6 Autocorrelation Peak Count .......................................................................... 74
    3.4.7 Autocorrelation Polynomial Fit ....................................................................... 77
    3.4.8 Autocorrelation Bandwidth ............................................................................. 78
    3.4.9 Fourier Transform of Spectra ......................................................................... 84
  3.5 Future Work with Spectroscopic OCT: Scattering Tissue ....................................... 87
CHAPTER 4: OCT IMAGE ENHANCEMENT USING WAVELET PROCESSING TECHNIQUES

4.1 Motivation for OCT Image Enhancement

4.2 Wavelet Theory
   4.2.1 Projections onto Subspaces
   4.2.2 Filter Bank Implementation
   4.2.3 Extension to Two Dimensions

4.3 Wavelet Denoising Techniques
   4.3.1 Image Quality Metrics
   4.3.2 Global Soft Thresholding
   4.3.3 SURE Soft Thresholding
   4.3.4 Spatially Adaptive Thresholding and Undecimated Wavelet Transforms
   4.3.5 Versatile Spatially Adaptive Thresholding
   4.3.6 Context-Based Spatially Adaptive Thresholding

4.4 Denoising of Ophthalmic OCT Images
   4.4.1 Algorithm Construction
   4.4.2 Time Domain Filtering Results
   4.4.3 Global Soft Thresholding and SUREShrink Results
   4.4.4 Spatially Adaptive Thresholding Results
   4.4.5 Conclusions

4.5 Future Work with Wavelet Denoising

CHAPTER 5: CONCLUSIONS

APPENDIX A: REFERENCES
CHAPTER 1: INTRODUCTION

1.1 Scope of Thesis

The main goal of this thesis project was to develop new digital signal processing (DSP) techniques to improve the flexibility, functionality, and image quality of ultrahigh resolution optical coherence tomography (OCT) systems. The field of digital signal processing is quite mature and has found many applications in other medical imaging fields such as magnetic resonance imaging, computed tomography, and positron emission tomography. In the specific area of coherent imaging, DSP technology is used in many aspects of synthetic aperture radar and medical ultrasound. To date, however, OCT systems have utilized very little of this technology. Signal filtering and demodulation are typically carried out using fixed analog electronics. Data analysis is traditionally limited to examining the intensity of backreflected light, neglecting the information contained in the spectrum of the scattered light (the exception here is Doppler flow OCT, which examines only the center wavelength of the scattered spectrum). Image processing is limited to motion artifact correction, with no attempt to improve sensitivity by noise reduction.

The specific areas investigated in this thesis project to improve OCT technology through the application of DSP include: development of a software-based digital demodulation scheme to reduce the need for fixed analog electronics, improving the flexibility of laboratory research systems and enabling rapid reconfiguration of imaging rate and source center wavelength; investigation of spectroscopic OCT, where the entire spectrum of the backscattered light is analyzed and used to provide alternative contrast modalities to traditional intensity-based OCT; and development of wavelet-based image enhancement techniques to remove speckle noise and improve signal-to-noise performance of OCT systems, specifically those utilizing low-cost, low-power superluminescent diode light sources.

1.2 Background of Optical Coherence Tomography

Optical coherence tomography (OCT) is an emerging biomedical imaging technique that provides in vivo cross-sectional images of tissue microstructure with micron-scale spatial resolution [1]. OCT operates by measuring the echo time delay of backscattered and backreflected optical waves from different layers within a biological tissue specimen. In many
ways, OCT can be thought of as an optical analogue to ultrasound imaging. However, since the time delay of reflected optical waves cannot be measured directly, an interferometric detection scheme is required. Typically, a fiber optic Michelson interferometer is used to measure the relative optical path difference between a reference arm and the various layers of the tissue sample in the sample arm. Constructive interference will occur only when the path lengths of the reference and sample arms match to within the coherence length of the light source. The source coherence length $\Delta L$ is inversely proportional to the source optical bandwidth $\Delta \lambda$. For a Gaussian optical spectrum, the coherence length is determined by

$$\Delta L = \frac{2 \ln(2) \lambda^2}{\pi \Delta \lambda}$$

(1.1)

As the coherence length decreases, the axial resolution of the OCT system improves. It is therefore desirable to use very broadband optical sources to achieve the highest possible image resolution.

By mechanically varying the path length in the reference arm (such as by translating the reference mirror with a scanning galvanometer), it is possible to obtain an interference fringe corresponding to an arbitrary path length in the tissue sample. Figure 1 shows a schematic of a typical fiber optic Michelson interferometer used for OCT imaging. To generate a single axial line (A-scan) in the image, an optical beam is incident on the specimen at a single transverse location while the reference path length is scanned to create a dataset of backscattered intensity vs. tissue depth. Each A-scan resembles a high-frequency carrier, amplitude modulated with a low-frequency “message” signal. The carrier signal contains information about the spectroscopic properties of the tissue, while the modulation contains information about the scattering structure. By translating the beam in a transverse direction, a 2D image (B-scan) can be generated. Figure 2 illustrate the concept of OCT image formation.

As the reference path is scanned through one period, the interference fringe is sampled and acquired by an A/D card. OCT systems can therefore be thought of as measuring optical interference in the time domain, where each optical wavelength $\lambda_{optical}$ present in the light source is mapped to a corresponding RF frequency $f_{RF}$ in the time domain. This mapping depends on
Figure 1. OCT measures the echo time delay of backreflected or backscattered light using low coherence interferometry. The coherence length of the light source determines the axial image resolution.

Figure 3 (right). (a) Symmetric dual-balanced interferometer. RF intensity noise from the broadband source is removed, but less power is available for imaging on the sample. (b) Asymmetric dual-balanced interferometer. Increased signal recovery is possible, with less power available for imaging on the sample. (c) Dual-balanced interferometer using circulator. 100% signal recovery with dual-balancing and 50% of source power available for imaging.

Figure 4 (below). Nonlinear galvanometer motion with sampling at evenly-spaced time intervals. Samples are not evenly spaced in z, leading to Doppler shifts in the OCT signal. Signal sees a blue shift in the middle of the scan, and red shifts at the start and end of the scan.
the velocity \( v \), at which the reference path is scanned, and is given by

\[ f_{RF} = \frac{2v}{\lambda_{optical}} \]  

(1.2)

The central RF frequency corresponding to the center wavelength of the source is often referred to as the Doppler frequency of the system, and is important for determining an appropriate sampling rate for the A/D conversion. The Doppler frequency is also the carrier frequency of the detected AM-type OCT signal as described above.

There are several basic engineering tradeoffs involved in the design and construction of an OCT system, a brief description of which is useful for understanding the work presented here. Most fundamental is perhaps the relationship between the light source parameters and imaging performance. From equation (1.1), it is evident that using a broader optical bandwidth will always produce better axial image resolution. However, the choice of center wavelength also plays an important role. Biological tissue scatters light with increasing efficiency as the center wavelength is decreased, effectively decreasing the maximum imaging depth. Also from equation (1.1), however, choosing a shorter source wavelength allows an improved axial resolution for a given source bandwidth. A second tradeoff exists between transverse resolution and depth of field. While axial resolution is determined entirely by the characteristics of the light source, transverse resolution is determined by the imaging optics used to deliver the beam onto the sample. Using high-NA focusing objectives in the sample arm leads to a narrow beam waist but a short Rayleigh range. This corresponds to a high transverse resolution, but an image that is in focus over a very short distance. Using lower NA objectives provides a better depth of field at the expense of a degraded transverse resolution.

Figure 1 shows a block diagram of a typical OCT imaging system using a single-detector symmetric Michelson interferometer. The broadband source can be a solid-state femtosecond pulse laser, a single superluminescent diode (SLD), or a set of multiplexed SLD’s. Light from the source is coupled into a single mode optical fiber and split evenly between the reference and sample arms of the interferometer. The reference mirror is translated using a scanning galvanometer, although other methods of scanning the optical path are possible [2, 3]. The photodiode \( PD_1 \) detects the interference signal between the reference and sample fields. This signal is typically bandpass filtered and demodulated on a logarithmic scale using an analog filter and logarithmic envelope detection circuit, leaving only the modulating message signal for A/D.
conversion. Since the high-frequency carrier has been removed, the sampling rate can be much lower than if the entire interference fringe were acquired. The A/D card samples the demodulated fringe and computer software processes the data for display.

Neglecting losses in the optical components, the power available for imaging on the sample is \(0.5P_0\), where \(P_0\) is the output power of the broadband source. The reference arm power is also \(0.5P_0\). The signal \(S\) generated by backscattering and backreflection in the tissue sample is split evenly by the fiber coupler, such that the total signal detected by the photodiode \(PD_1\) is \(0.5S\). The remaining signal is passed back into the broadband source. This immediately highlights a problem with the single-detector interferometer design of Figure 1; although it is simple to implement, half of the signal from the sample arm is lost in the fiber coupler. Furthermore, excess intensity noise from the broadband source will corrupt the OCT signal, degrading signal-to-noise ratio (SNR) performance. Many other interferometer designs can be used for OCT, some of which are illustrated in Figure 3.

Figure 3a shows a symmetric dual-balanced interferometer, where the intensity noise of the broadband source is removed by performing a differential measurement of the power in the two photodiodes \(PD_1\) and \(PD_2\). Light traveling in a “bar” path through a 50/50 fiber optic coupler (ie, from an upper port to a lower port or vice versa) obtains a \(-\pi/2\) phase shift relative to the light that travels in a “cross” path (ie, from an upper port to an upper port, or lower port to a lower port). Therefore the interference signal intensities at the two detectors will be perfectly out of phase, so electronic subtraction has the effect of doubling the detected signal power [4]. The intensity noise at the two detectors, however, will be in phase. The electronic subtraction will therefore cancel much of the noise induced by the broadband source [4]. The drawbacks to this approach are that only 25% of the source power is available for imaging on the sample, and that the transmission characteristics of the two couplers must be closely matched in order to provide good noise cancellation; any mismatch in the spectrum of the detected light at the two detectors will translate into reduced noise cancellation. A 50% attenuator is necessary before \(PD_1\) to equalize the intensity of the light detected by each photodiode for good noise cancellation.

Figure 3b shows an asymmetric dual-balanced interferometer, suitable for use with a high-power broadband source. By using a high split ratio coupler immediately following the broadband source, it is possible to receive a greater portion of the signal from the sample arm
while also providing noise cancellation through dual-balancing. The drawback here is that a 
smaller portion of the source power is available for imaging on the sample. Finally, figure 3c 
shows a dual-balanced interferometer implemented with an optical circulator and a fiber optic 
coupler. This is a very efficient interferometer design in terms of signal conservation. 
Neglecting system losses, 100% of the sample signal is recovered by the detectors and 50% of 
the source power is available for imaging. The main drawback here is that broadband optical 
circulators that are polarization independent are very expensive.

The field of OCT has evolved rapidly in the past several years. The earliest and most 
developed clinical application of the technology is in ophthalmology [5], since the optically 
transparent nature of the eye made it an obvious organ system for early work. These systems 
used superluminescent diode sources with relatively low optical bandwidths, and could achieve 
axial resolutions of only 10-15 μm. The development of solid-state femtosecond pulse lasers led 
to a significant improvement in image resolution for laboratory based systems, first to ~3 μm [6], 
and then to ~1 μm [7] by using double-chirped mirrors for laser cavity dispersion compensation. 
These ultrahigh resolution OCT (UHR-OCT) systems provided unprecedented image resolution 
but were not suitable for clinical use due to their large size. The development of compact, 
portable, low-threshold Ti:Sapphire femtosecond lasers operating at center wavelengths of 800 
nm [8] allowed UHR-OCT technology to be taken into ophthalmic clinics, while portable 
Cr:Forsterite lasers operating at 1300 nm [9] allowed clinical UHR-OCT gastrointestinal 
imaging.

Presently, state-of-the-art OCT systems are capable of providing axial resolutions of ~1 μm 
in laboratory systems and ~3 μm in clinical systems using Titanium:Sapphire lasers. High-speed 
imaging technology has been developed that can produce UHR-OCT images at up to a few 
frames per second [2] using traditional OCT systems, which has allowed studies of cancerous 
tissue to be conducted in vivo and in pathology laboratories. Finally, Fourier Domain OCT [10, 
11] has demonstrated the capacity for imaging rates of up to 30 frames per second with an 
improvement in sensitivity compared to traditional OCT.
1.3 Concept of Spectroscopic Optical Coherence Tomography

As described above, traditional OCT systems analyze the envelope of the demodulated interference signal to produce two-dimensional maps of backreflected / backscattered light intensity. These maps provide information on tissue microstructure by looking at the amount of light backscattered by the tissue at each point in the sample. However, there is a great deal of information encoded in the spectrum of the backscattered light that is not available with intensity-based OCT imaging. For example, different types of tissue will have different scattering and absorption spectra depending on the cell sizes, blood content, water content, and whether the area is oxygenated or deoxygenated. Although multiple biological structures may scatter the same amount of light (and therefore be indistinguishable with intensity-based OCT), the spectrum of the detected light may be different, providing an alternative contrast modality. Spectroscopic OCT may find applications in functional imaging (by detecting local variations in blood oxygenation), early detection of cancer (by detecting local variations in cell size or nuclear density), and enhanced tissue differentiation for the detection of other pathologies (by detecting changes in spectroscopic properties of tissue types not normally visible with OCT). A related technique, Doppler flow OCT, is well-developed and is capable of high-accuracy quantization of blood flow down to a few microns of flow per second [12, 13]. This method measures the Doppler shift of local spectra to produce quantitative maps of fluid flow in vivo. Earlier work on spectroscopic OCT took a similar approach, creating 2D maps of variation in center wavelength to “spectroscopically stain” standard intensity images [14]. However, this method has drawbacks that make it unsuitable for many applications, which will be fully discussed later.

The technique for spectroscopic OCT imaging is based on the short time Fourier transform (STFT) analysis. By applying sequential shifts of a local one-dimensional windowing function to each A-scan and calculating the Fourier transform of the windowed data, an approximate local spectrum can be calculated for every pixel in the image. As previously mentioned, this requires the complete interference fringe to be acquired by the A/D card. If the sample being imaged were a perfect mirror, the calculated spectra would be identical to the laser spectrum. Spectroscopic analysis produces a 4-dimensional data set consisting of x and y spatial coordinates, an RF frequency equivalent to an optical wavelength, and a spectral amplitude. To create an interpretable image, it is necessary to condense this information to three dimensions.
This forms the critical step in spectroscopic OCT analysis, as the metric chosen to describe the spectrum at each pixel strongly affects the contrast enhancement of the system as well as the types of spectroscopic properties that will be visible. The process of choosing a good spectroscopic metric, as well as preliminary results in tissue phantoms and animals, are discussed fully below.
CHAPTER 2: APPLICATIONS OF DIGITAL SIGNAL PROCESSING IN OPTICAL COHERENCE TOMOGRAPHY

2.1 Data Acquisition

In addition to common data acquisition problems such as selecting an appropriate sampling rate and managing high-speed data streaming, OCT systems also have to contend with data corruption caused by nonlinear scanning of the reference arm’s optical path. To understand the problem of nonlinear reference path scanning, it is useful to consider a mechanical galvanometer as the reference scanner, although the same problem exists for any type of scanner used in a time-domain OCT system. Consider first an ideal system, where the velocity of the galvanometer \( v_s \) is constant during the time required to acquire one A-scan. Since the A/D cards used in most OCT systems operate at a fixed sampling frequency \( f_s = 1/T_s \), the physical distance \( \Delta z \) scanned by the galvanometer over one sampling interval \( T_s \) is fixed at \( \Delta z = v_s T_s \). Now consider a non-ideal system, where the galvanometer scanning velocity is a function of time and varies significantly over the time required to acquire one A-scan. In this case, we have \( \Delta z(t) = v_s(t) T_s \), and the physical distance scanned by the galvanometer over one fixed sampling interval is no longer constant. Figure 4 illustrates the effect of nonlinear galvanometer motion in conjunction with a fixed sampling frequency. This situation is present in real OCT systems, and can be a cause of significant error in spectroscopic analysis if not corrected.

There are two implications to nonlinear reference path scanning. First, there will be artificial compression and extension of spatial features as the galvanometer accelerates and decelerates, respectively. This effect occurs in both intensity-based and spectroscopic OCT imaging, but is quite small and is usually not significant. Second, and more importantly, the nonlinear motion of the galvanometer imparts local, time varying, Doppler shifts to the spectrum of the OCT signal. This effect can be understood by considering the origin of the OCT carrier signal. As shown in equation (1.2), the RF frequency of the OCT signal is directly proportional to the galvanometer velocity. Therefore a time-varying galvanometer velocity will create a time-varying RF carrier, equivalent to inducing a frequency modulation. As the galvanometer velocity increases, the
Figure 5. Nonlinear galvanometer motion with sampling at evenly-spaced galvanometer positions. Samples are evenly spaced in z, removing the Doppler shifts in the OCT signal.

Figure 6. OCT system with calibration interferometer for simultaneous measurements of tissue backscattering and galvanometer position. CW interference fringe encodes galvanometer position as an FM-type signal. This information can be used to correct the spectrum of the OCT signal to remove Doppler shifts induced by the galvanometer.

Figure 7. Uncorrected and corrected spectra of CW calibration signal. Uncorrected spectrum shows smearing due to galvanometer motion. Corrected spectrum resembles an impulse function.

Figure 8. Uncorrected and corrected spectra of OCT signal from human oral mucosa. Corrected spectrum has more concentrated power distribution around the carrier frequency.
local spectra will become blue shifted. As the velocity decreases, the local spectra become red shifted. This is a major source of spectroscopic noise, as it leads to corruption of the local spectra and can swamp small spectroscopic signals of interest. It is therefore necessary to remove the effect of nonlinear galvanometer velocity from the OCT signal before carrying out spectroscopic analysis.

To accomplish this, several solutions are possible. One possibility is to use the Hilbert transform to obtain the phase of the carrier signal after acquisition. By linearizing the phase in a post-processing implementation, it is possible to correct the effect of galvanometer motion and remove local Doppler shifts. However, this approach is not easily converted to a real-time implementation without the use of dedicated DSP hardware. Another approach that is more suited to real-time implementation is the use of a “calibration” Michelson interferometer to trigger data acquisition at uniform reference arm intervals $\Delta z$. By linking the data acquisition trigger to the position of the reference arm as shown in figure 5, Doppler shifts caused by the galvanometer can be removed. The setup for such a system is shown in figure 6. A continuous wave (CW) narrow linewidth source, such as a helium-neon laser, is used to measure the position of the scanning reference mirror. A bulk beamsplitter cube is used to align the CW source beam coaxially with the OCT beam. A fixed mirror and second beamsplitter cube are used to form the rest of the calibration interferometer. The calibration fringe signal contains FM modulation caused only by nonlinear galvanometer motion, and is not affected by the properties of the tissue being imaged. This calibration signal is passed through a bandpass filter to remove noise outside of the RF frequency range of interest, as defined by the galvanometer mean velocity and the center wavelength of the CW calibration source.

To achieve real-time correction of galvanometer motion, it would be necessary at this point to link the analog output of the calibration detector $D2$ to the digital trigger input of the A/D card. This could be accomplished by using a Schmitt trigger circuit, which would convert the analog FM signal at $D2$ to a digital pulse train by switching logical states whenever the analog signal crosses 0V. One obstacle to overcome in this real-time implementation arises from the desire to oversample the OCT signal. Since the center wavelengths of the laser sources used in OCT typically range from 800 – 1500 nm, in order to provide a trigger pulse train at the Nyquist sampling rate we would need (at a bare minimum) a CW source with a wavelength of 400 nm. This would result in a critically sampled system, and would in reality be slightly undersampled.
due to the finite bandwidth of the OCT signal. Since most inexpensive CW sources have wavelengths around 600 nm, it would therefore be necessary to upconvert the frequency of the digital trigger signal before passing it to the A/D card. This could be done using digital electronics, or alternatively the A/D card could be programmed to acquire multiple samples at each trigger pulse.

Due to time constraints in this thesis project, a real-time implementation was not fully realized. Instead, the concept of using zero crossings of the analog calibration fringe to correct the OCT signal was applied in post-acquisition software. In the future, however, the calibration interferometer setup used in this project could be expanded to allow for a real-time implementation as described above. The post-acquisition fringe correction software is described in the following.

First, the A/D card is set to acquire two data streams, one from the OCT interferometer output and one from the calibration interferometer output. The A/D card used in this experiment (National Instruments 6052E, 16 bit resolution, \( f_s = 333 \text{ kHz} \)) is not capable of simultaneous data acquisition of multiple channels, but instead multiplexes the data together into one stream. This effectively halves the sampling rate of each channel to \( 166 \text{ kHz} \). The data from the two interferometers are separated by a de-multiplexing algorithm and stored in separate buffers in system RAM. After the complete dataset is acquired, Matlab code is used to perform correction of the OCT fringes one A-scan at a time. Since the dataset sizes are very large when dealing with spectroscopic OCT systems (the main application area for this technique), it was important to optimize the execution time of the code whenever possible. Even though the correction algorithm was implemented with a minimum of loop structures, execution time for large datasets (100 – 200 MB) was approximately 6 – 7 hours running on a Pentium-4, 2.2 GHz processor. This long execution time is primarily due to the inefficient implementation of for loop structures in Matlab, and could be improved by orders of magnitude if the algorithm were ported to C++.

The basic method of the correction algorithm is to estimate positions of the calibration fringe zero crossings, lay down evenly spaced points between pairs of zero crossings, and then interpolate the OCT fringe at these new sample points. Implicit in this method is the assumption that the galvanometer velocity is constant over one period of the calibration fringe carrier signal. This assumption is reasonable, since the calibration fringe period is typically 0.02% of the galvanometer period. It should also be noted that it is not generally possible to use the zero
crossings of the OCT signal for "self-correction." This is because the OCT signal carrier is expected to vary due to spectroscopic effects in the tissue, and self-correction would therefore remove the signal of interest that we are trying to preserve.

To estimate the locations of the zero crossings of the calibration fringes, a rapid four-step method is used. First, the sign of the sampled calibration fringe \( c[n] \) is computed as \( s[n] \). Then a two point backward difference calculation is applied to \( s[n] \), giving \( \Delta s[n] \). The points \( n_0 \) where \( \Delta s[n_0] \) is non-zero indicate samples where the sign of \( c[n] \) changes from positive to negative, or negative to positive. Therefore the points \( n_0 \) and \( n_0 + 1 \) of the sampled signal \( c[n] \) will bracket a zero crossing of the continuous time signal \( c(t) \). Once the set of points \( \{n_0\} \) have been determined, cubic interpolation is performed over the local neighborhood \([n_0 - 2, n_0 - 1, n_0, n_0 + 1, n_0 + 2, n_0 + 3]\) for all \( n_0 \) to determine the approximate zero crossing points \( \hat{n}_z[k] \), where \( 0 \leq k \leq K - 1 \) and \( K \) is the number of zero crossings in the calibration fringe. It should be noted that while \( n \) and \( n_0 \) are integer sample indices, the zero crossings \( \hat{n}_z[k] \) are not necessarily integers. The zero crossing estimation procedure is summarized as follows:

1) \( s[n] = \text{sgn}(c[n]) \)
2) \( \Delta s[n] = s[n] - s[n-1] \) (2.1)
3) \( \{n_0\} = \{n : \Delta s[n] \neq 0 \} \)
4) \( \hat{n}_z = \text{cubic interp over } [n_0 - 2, n_0 - 1, n_0, n_0 + 1, n_0 + 2, n_0 + 3] \forall \{n_0\} \)

Once the zero crossings of the calibration fringe for a given A-scan are estimated, a vector of interpolation points is created for resampling the OCT fringe. Placing \( N \) points between consecutive values of \( \hat{n}_z[k] \) has the effect of creating a new sampling function for \( c(t) \) that is evenly spaced in \( \Delta z \) instead of evenly spaced in \( \Delta t \). The new sampling function \( d[k] \) is of length \( NK \) and is mathematically given by

\[
d[k] = [\ldots \hat{n}_z[k], \hat{n}_z[k] + \Delta \hat{n}_k, \hat{n}_z[k] + 2\Delta \hat{n}_k, \ldots, \hat{n}_z[k] + (N-1)\Delta \hat{n}_k, \hat{n}_z[k+1], \ldots]
\]

\[
\Delta \hat{n}_k = \frac{\hat{n}_z[k+1] - \hat{n}_z[k]}{N}
\] (2.2)
The choice of $N$ is arbitrary, and should reflect the desired number of points to be included in the corrected A-scan. Resampling of $c(t)$ is approximated by performing a cubic interpolation of $c[n]$ over the points $d[k]$. The result of the entire process is that the local spectra of the OCT signal will be free from Doppler noise induced by the galvanometer. Figure 7 shows the uncorrected and corrected spectra of a typical calibration fringe. With a galvanometer velocity of 3.6 mm/s and a CW wavelength of 635 nm, ideally the fringe should resemble an impulse function shifted out to a carrier frequency of 11.3 kHz. The uncorrected spectrum shows a smearing effect caused by the nonlinearities in the galvanometer motion, while the corrected spectrum shows a more concentrated energy profile, indicating that the algorithm is operating properly. Figure 8 shows the uncorrected and corrected spectra of an OCT signal acquired in vivo by imaging human oral mucosa. Here, we expect the RF spectrum to be centered at a carrier frequency of ~8.8 kHz, corresponding to a source center wavelength of ~820 nm. Again, the spectrum of the corrected signal is more concentrated than the uncorrected signal. Both corrected spectra show a shift in center frequency towards DC; this is caused by the resampling procedure, which interpolates over $N$ new points between each pair of zero crossings. When $N$ is not equal to the number of points between the zero crossings in the original signal, the sampling rate of the system is effectively changed. This compresses or expands the corrected spectra, depending on whether $N$ is larger or smaller than the original number of points between zero crossings. However, this rescaling does not affect spectroscopic measurements since it is consistent for all the A-scans in a given image.

A digital bandpass filter is also included as an optional stage in the process to remove noise from the calibration signal. Since the algorithm uses zero crossings of the calibration fringe as a "ruler" to carry out the resampling operation, any spurious zero crossings will cause an accumulating error to form. The bandpass filter is an elliptical finite impulse response (FIR) filter with a user-defined passband width, passband ripple, and stopband ripple. Each A-scan in the calibration data is filtered, then the filter output is time-reversed and passed back through the filter. This technique ensures that the final filter output has no group delay, which is important to avoid introducing motion artifacts into the OCT image. The effective filtering function is the square of the prototype filter’s magnitude response. The digital filter is generally not required if the output power of the CW source is high enough, such as when a Helium-Neon laser is used as
Figure 9. Full interference fringe OCT signals for testing digital demodulation techniques. (a) Simulated OCT signal, with three distinct carrier frequencies present. (b) Section of an OCT signal acquired by imaging the cheek pouch of a Syrian golden hamster.

Figure 10. (a) Simulated OCT signal with coherent demodulation. Coherent demodulation performs poorly when the carrier frequency and phase vary with time. (b) Real OCT signal with coherent demodulation. Nulls are visible when carrier and mixing signal are out of phase.
Simulated OCT Signal and Demodulated Message Signal (Boxcar Demodulation)

![Simulated OCT Signal and Demodulated Message Signal (Boxcar Demodulation)](image)

**Figure 11.** (a) Simulated OCT signal with boxcar demodulation. DC offset errors and high-frequency errors are visible. (b) Real OCT signal with coherent demodulation. DC offset errors are visible.

OCT Signal and Demodulated Message Signal From Hamster Cheek Pouch (Boxcar Demodulation)

![OCT Signal and Demodulated Message Signal From Hamster Cheek Pouch (Boxcar Demodulation)](image)

Simulated OCT Signal and Demodulated Message Signal (Hilbert Demodulation)

![Simulated OCT Signal and Demodulated Message Signal (Hilbert Demodulation)](image)

**Figure 12.** (a) Simulated OCT signal with Hilbert demodulation. Demodulated signal tracks the envelope with high accuracy. (b) Real OCT signal with Hilbert demodulation. High accuracy demodulation is observed.
the calibration laser source. However, if a low-power laser diode is used as the calibration source, it may be necessary to use the digital filter.

As previously mentioned, the fringe correction algorithm is intended primarily for use in spectroscopic analysis of OCT data. However, as discussed in section 33.4, by appropriately designing the spectroscopic analysis technique the effects of galvanometer motion can be eliminated without the need for post-processing fringe correction or real-time corrected triggering schemes. As will be seen, this is a major advantage of the spectroscopic analysis techniques developed as part of this thesis project, since insensitivity to galvanometer noise results in a spectroscopic system that is less complex and has less computational requirements.

2.2 Digital Demodulation

As discussed in section 1.2, standard intensity-based OCT imaging acquires and analyzes only the intensity of the demodulated interferometer output. Typically, the demodulation function is combined with a logarithmic transformation and carried out in fixed analog electronics using envelope detection and a log circuit. This implementation is suitable for commercial OCT systems, where the design has been optimized and fixed. For research systems, however, it is often necessary to reconfigure system parameters that affect the demodulation circuit. Changing the center wavelength of the source, for example, may move the RF spectrum of the OCT signal outside the bandwidth of the analog demodulator. Furthermore, analog envelope detection does not generate a perfect reconstruction of the signal envelope, generating demodulation noise in the OCT signal. Finally, it is sometimes desirable to view the demodulated signal on a linear instead of a logarithmic scale. This would require switching out analog electronics in the conventional system implementation, which is time consuming and requires additional components to be kept on hand. Alternatively, the demodulation function can be carried out in real time using efficiently designed DSP software. Such an approach can provide increased demodulation accuracy, greatly improved flexibility, and decreased system cost.

As part of this thesis project, several digital demodulation schemes were investigated and evaluated for speed and accuracy. The methods examined include boxcar averaging, coherent demodulation, and Hilbert demodulation. The methods were evaluated by demodulating a
deterministic input signal \( x[n] = m[n] \cos(\omega[n]n + \phi[n]) \) designed to simulate an OCT signal, where \( m[n] \) is the message signal of interest and \( \cos(\omega[n]n + \phi[n]) \) is a carrier with time-varying frequency and phase. By calculating the mean-square error (MSE) between \( m[n] \) and the demodulated message \( \hat{m}[n] \), the methods can be quantitatively compared. MSE is defined as

\[
\text{MSE} = \frac{1}{N} \sum_{n=0}^{N-1} (m[n] - \hat{m}[n])^2
\] (2.3)

The deterministic message signal was chosen to provide the demodulation techniques with a signal having properties common to actual OCT data. Specifically, three separate carrier frequencies are present (at \( \omega_1 = \pi/8 \), \( \omega_2 = \pi/6 \), and \( \omega_3 = \pi/4 \)). The message signal consists of four Gaussian pulses of varying amplitude and width. Figure 9 shows the simulated OCT signal and a section of a real OCT signal acquired in vivo by imaging of the cheek pouch of a Syrian golden hamster. Since there is no known error-free \( m[n] \) in the case of an actual OCT signal, the MSE cannot be calculated. Rather, a qualitative comparison of the various demodulation techniques is presented for this situation.

The first technique investigated was coherent demodulation, the standard demodulation scheme used for AM transmission systems. A mixing signal at the nominal carrier frequency is used in conjunction with lowpass filtering to recover the message signal. This method is not well-suited for OCT signals, since the carrier frequency is expected to vary with axial depth due to spectroscopic properties of the tissue. Furthermore, the phase of the carrier is discontinuous at random locations in the image, leading to errors in coherent demodulation. Figure 10 shows the demodulated simulated signal and demodulated cheek pouch signal. For the simulated signal, the MSE was 0.963. Qualitatively, the real OCT signal shows that coherent demodulation performs well only in areas where the frequency and phase of the carrier match that of the mixing signal. In other areas, the demodulation is less accurate. In the extreme case of a \( \pi \) phase difference between the mixing signal and the carrier, the demodulated output is identically zero.

The second technique, boxcar averaging, is not a true demodulation method but is extremely rapid due to its low computational complexity. Furthermore, it is largely insensitive to variations
in the carrier frequency and phase. In this method, a sliding rectangular window of length \(W+1\) is applied to each A-scan and the mean of the absolute value of the modulated signal within the window is calculated, as given by

\[
\hat{m}[n] = \frac{1}{W+1} \sum_{k=-W/2}^{W/2} |x[n-k]|
\]  

(2.4)

The window length should be chosen to be approximately equal to the period of the carrier signal. The main drawbacks to this method are that it is sensitive to the choice of window length, and that it is prone to DC offset errors. It is, however, more accurate than coherent demodulation and extremely rapid. Figure 11 shows the demodulated simulated signal and demodulated cheek pouch signal. For the simulated signal, the MSE was 0.655. Qualitatively, the real OCT signal shows that the boxcar averaging technique performs better than coherent demodulation, but suffers from a DC offset error. There are also significant errors in the envelope tracking of this method, although this effect is not as severe as with coherent demodulation.

The third technique, Hilbert demodulation, had the best performance and was therefore chosen for implementation in a real-time OCT system. Hilbert demodulation is a frequency domain technique that separates the message signal from the carrier independently of the carrier phase and frequency, and is based on the decomposition of the modulated signal into two components. If the modulated signal \(x[n]\) is purely real, it can be represented as the sum of two complex signals.

\[
x[n] = x_a[n] - jx_h[n]
\]  

(2.5)

Where \(x_a[n]\) is the complex “analytic signal” and \(x_h[n]\) is the complex “Hilbert signal.” The Fourier transforms of the analytic and Hilbert signals can also be related to the Fourier transform of the original signal.

\[
X_a(\omega) = 2u(\omega)X(\omega)
\]

\[
X_h(\omega) = -j \text{sgn}(\omega)X(\omega)
\]

\[
\therefore X(\omega) = X_a(\omega) - jX_h(\omega)
\]  

(2.6)
Here, $X_a(\omega)$ is the Fourier transform of $x_a[n]$ and $X_h(\omega)$ is the Fourier transform of $x_h[n]$. The complex nature of $x_a[n]$ and $x_h[n]$ can also be seen in their Fourier transforms, as the transforms are not symmetric about the origin.

To carry out demodulation, the signal $x[n]$ is expressed in complex notation, and the analytic and Hilbert signals are calculated as

$$x[n] = m[n] \cos(\omega[n] n + \phi[n])$$

$$x[n] = m[n] \frac{1}{2} \left[ e^{j(\omega[n] n + \phi[n])} + j e^{-j(\omega[n] n + \phi[n])} \right] = x_a[n] - j x_h[n] \quad (2.7)$$

The message signal $m[n]$ can therefore be recovered by taking twice the absolute value of either the analytic signal or the Hilbert signal. The Hilbert demodulation algorithm can be carried out in four steps.

1) Take the Fourier transform of $x[n]$ to generate $X(\omega)$

2) Multiply $X(\omega)$ by $2u(\omega)$ to generate $X_a(\omega)$

3) Take the inverse Fourier transform of $X_a(\omega)$ to generate $x_a[n]$

4) Take $2|x_a[n]|$ to generate $m[n]$

As equation (2.7) shows, recovery of the message signal does not require the carrier frequency or phase to be constant over time. The Hilbert demodulation approach is therefore optimal for OCT systems. Figure 12 shows the demodulated simulated signal and demodulated cheek pouch signal. For the simulated signal, the MSE was $2.8 \times 10^{-4}$, by far the best of the three methods studied. Qualitatively, the real OCT signal shows that the Hilbert method tracks the envelope with a high degree of accuracy compared to coherent demodulation and boxcar averaging.

To ensure that the digital demodulation software is suitable for use in a real-time OCT system, it is important to implement the algorithm in a highly efficient manner. C++ code was written to carry out the demodulation, based on the use of the FFTW algorithm [15]. FFTW is a collection of C++ routines developed by researchers at MIT to provide a rapid, platform-
Figure 13. C++ algorithm for rapid digital Hilbert demodulation of OCT signals. Execution time on a Pentium-4 2.2 GHz processor running Windows 2000 is 6 ms for $2^{15}$ points. Algorithm is based on the FFTW library of functions developed by Frigo and Johnson.

Figure 14. Absorption coefficient vs. wavelength for hemoglobin, melanin, and water. These materials are the major chromophores in biological tissue, and can provide spectroscopic contrast for OCT imaging.

Figure 15. Spherical coordinate system used to describe Mie scattering. Light is incident on the particle, directed in the $xy$ plane. Nonuniform scattered light is observed in all directions.
optimized Fourier transform algorithm. FFTW is capable of adapting and optimizing itself to the hardware and software environment that it is running on. This is accomplished by carrying out a training step during the startup phase of the demodulation code. More information on the specifics of FFTW are available on the FFTW website [15].

When performing the Fourier transform of a real-valued time domain signal, FFTW makes use of the fact that the real and imaginary components of the transform will be symmetric about \( \omega = 0 \). Specifically, the real component of the transform will be even symmetric about \( \omega = 0 \) while the imaginary component of the transform will be odd symmetric about \( \omega = 0 \). For a \( N \) point real-valued time domain signal \( x[n] \), FFTW computes only the first \( N/2 \) real values and \( N/2 \) imaginary values of the transform \( X(\omega) \). This is ideal for Hilbert demodulation, since we wish to retain only the positive frequency components (i.e., the first \( N/2 \) points) of the transform of \( X(\omega) \) to generate \( X_a(\omega) \). Figure 13 illustrates a highly efficient algorithm for performing real-time Hilbert demodulation on one A-scan using FFTW and C++. The algorithm is summarized as follows.

1) Read one A-scan from the data acquisition card’s output buffer into a double precision block of memory of length \( N \), converting the data format from short to double.

2) Initialize two blocks of memory, each of length \( N \), to hold the Fourier transform of the A-scan. Fill the blocks with zeros. This step takes the place of multiplication by \( u(\omega) \) in (2.6).

3) Perform a real-to-complex FFTW operation, filling the two blocks initialized in step 2. FFTW will store the positive frequency components of the transform in the first \( N/2 \) locations in each block of memory. The first block will contain the real components of the transform, while the second block will contain the imaginary components. The two memory blocks now contain the real and imaginary components of \( X_a(\omega) \).

4) Perform a complex-to-complex inverse FFTW operation. FFTW will store the results in two new memory blocks, each of length \( N \). The first block will contain the real components of the time-domain analytic signal \( x_a[n] \), while the second block will contain the imaginary components.
5) Take the absolute value of the two memory blocks from step 4, treating each pair of values in the memory blocks as a complex number. Multiply by a scale factor of 4 to correct the absolute level and recover $m[n]$.

When carefully implemented, this algorithm is extremely fast. The FFTW algorithm execution time scales with $N \log(N)$ for complex-to-complex transforms and with $N/2 \log(N/2)$ for real-to-complex transforms. Step 5 in the process requires three multiplications, one addition, and one square root operation for every point in the original A-scan. This is roughly equivalent to $5N$ operations. Since the initialization step (step 2) needs to be performed only once for an entire data set of multiple A-scans, the total computational complexity of the method is proportional to $N \log(N) + N/2 \log(N/2) + 5N$. Not included in this estimate is the processor time required to convert the data format from short to double. For a typical full-fringe OCT data set, one A-scan can have up to $2^{15}$ points. Therefore the number of operations required for Hilbert demodulation is approximately $3.8 \times 10^5$. On a Pentium-4 2.2 GHz processor running Windows 2000, complete Hilbert demodulation of one such A-scan with $2^{15}$ points can be performed in ~6 ms. Since the time required to acquire one A-scan is typically on the order of 10-20 ms (optical path scanning at 75-150 mm/s over a scan depth of 1.5 mm), this indicates that digital Hilbert demodulation is well-suited for use in OCT imaging systems. Execution time may be further reduced by operating on short data sets, which requires modification of the operating parameters of the FFTW algorithms, or by running the software on a faster PC.

Once the C++ demodulation code was developed and tested, it was successfully integrated into data acquisition software for Optical Coherence Microscopy (OCM) by another member of the research lab. The code is being used routinely to allow rapid switching from linear to logarithmic intensity display of OCM data during image acquisition.
3.1 Motivation for Spectroscopic OCT

As discussed in section 1.3, standard optical coherence tomography systems analyze the intensity of backscattered light from biological tissue samples to generate images. For this application, the carrier signal is not required, and only the demodulated envelope signal is acquired and analyzed. Intensity-based OCT imaging is therefore limited to providing structural information about the tissue sample based on the amount of light backscattered by the various structures. However, such a technique does not allow for investigation of the spectroscopic properties of the tissue sample. Analysis of spectroscopic features can enhance the contrast of OCT systems by differentiating tissue based on something other than the intensity of the backscattered light. Contrast enhancement is observed when two tissues or materials in a sample backscatter the same amount of light, but with different distributions over the wavelength range of the OCT light source. In this case, standard intensity-based OCT would be unable to differentiate the tissues or materials, but through spectroscopic analysis, differentiation is possible.

Two fundamental sources of spectroscopic contrast in OCT imaging are wavelength-dependant absorption and wavelength-dependant scattering. The wavelength-dependant absorption profiles of melanin, oxygenated hemoglobin (HbO₂), and water are shown in figure 14. These materials are the major endogenous chromophores in biological tissue, and their presence in various concentrations can be used to provide spectroscopic contrast. For example, areas showing abnormal localized concentrations of HbO₂ may indicate angiogenesis, a marker of cancer development. Alternatively, the presence or absence of HbO₂ can also be used to detect blood vessels and to distinguish arterial from veinous blood flow.

Wavelength-dependant scattering properties provide a different source of spectroscopic contrast. In this case, the phenomena of interest is the fact that scattering particles of different size and spatial distribution produce backscattered spectra with varying degrees of modulation. Mie theory describes this process in closed form for simple systems of spherical particles under
coherent illumination. For single-particle systems, the frequency of the modulation imparted to the backscattered spectrum is proportional to the particle size. In other words, the larger the scattering sphere relative to the illuminating wavelength, the higher the frequency of the modulations imposed on the optical spectrum. The optical modulation characteristics are also related to the density of the scattering particles. Spectroscopic analysis of OCT signals can therefore be investigated for measuring (qualitatively if not quantitatively) the size and density of scattering particles in a biological tissue sample. Since the cells that compose various types of tissue typically have varying organelle sizes and densities, examining spectral modulation may lead to contrast enhancement of different cell types. Furthermore, diseased tissue such as cancer often exhibit changes in overall cell size, nuclei size, and mitochondria density. Such changes may also be indirectly visible using spectroscopic OCT imaging.

In this section of the thesis project, the development of new spectroscopic OCT data analysis techniques are described in detail. The critical aspects of information visualization, spectrum characterization (spectroscopic metric selection), software development, and preliminary imaging studies on tissue phantoms and animal models are discussed. Sources of noise unique to spectroscopic OCT are discussed, and analysis techniques that are insensitive to these noise sources are developed. Spectroscopic imaging results are presented that demonstrate the ability to detect arterial and venous blood vessels in scattering tissue. Results from imaging developmental biology specimens illustrate the ability to enhance tissue contrast based on differences in scattering particle size. Finally, preliminary data from epithelial tissue imaging showing the possibility for detection of epithelial basal membrane cells is presented. Future work is discussed, with an eye towards early cancer detection and further investigation of developmental biology specimens.

3.2 Review of Mie Scattering Theory

To understand the basis for using spectral modulation as a mode for spectroscopic contrast in OCT systems, it is useful to present a review of Mie scattering theory. Mie theory provides a justification for choosing spectral modulation as an analysis technique, and also provides insight as to the biological source of OCT signals. Measuring spectral modulation to detect the size of biological scattering particles is not new; much work has been carried out in measuring the size of epithelial cell nuclei using light scattering spectroscopy to detect cancer cells in vivo and in
vitro, as well as to characterize particle size using artificial microspheres and cellular monolayers [16-23]. Central to this work is idea that cellular organelles (most notably the nuclei) of epithelial tissue can be considered as spheroidal scatterers whose interactions with light are governed by Mie theory [24, 25]. Mie scattering theory provides a closed-form description of optical scattering from spheroidal particles as a function of particle size, refractive index, wavelength, observation angle, and optical polarization. This theory is valid when the particle size is on the order of the illuminating wavelength, or larger. When the particle size is significantly smaller than the illuminating wavelength, the closed-form Mie description becomes less accurate and Rayleigh scattering provides a more appropriate model.

The discussion begins with a brief description of the general scattering problem. The solution of Maxwell’s equations in spherical coordinates is considered. Next, a set of harmonic vector solutions to the vector wave equation in spherical coordinates is presented. These “vector harmonics” provide a complete, orthogonal basis set of functions with which we can express a plane wave (or any other wave) as an infinite series. From here expressions are shown for the expansion coefficients of the infinite series and solutions are given for the scattered field of a spherical particle. Most relevant for OCT imaging, the relationship between backscattering efficiency and illuminating wavelength is described in detail. Finally, the discussion concludes with a look at the effects of multiple scatterers. The discussion of Mie theory follows derivations in the texts by Bohren [26] and van de Hulst [27].

3.2.1 General Scattering Problem

To begin, consider a single scattering particle that is impinged upon by an incident beam. The particle will nonuniformly scatter the incident light in all directions, where the scattered intensity in a given direction depends on the properties of the particle (such as shape and refractive index) and the incident field (such as wavelength and propagation direction). Figure 15 shows the spherical coordinate system that will be used to describe scattering. The center of the scattering particle defines the origin of the coordinate system, while the direction of propagation of the incident light defines the z-axis. The unit vector \( \mathbf{e}_r \) defines the scattering direction of interest. The unit vectors \( \mathbf{e}_r \) and \( \mathbf{e}_z \) together define a scattering plane, similar to a plane of incidence in the problem of reflection at an interface. The scattering plane can be
uniquely specified by $\phi$ (the azimuthal angle) and $\theta$ (the scattering angle) except for the case of complete forward scatter or complete backscatter, when $e_r = \pm e_z$. In this case any plane containing the $z$-axis can be considered to be the scattering plane.

Assume now that the incident field is arbitrarily polarized in the $xy$-plane. The orthogonal unit vectors $e_{||}$ and $e_{\perp}$ also lie in the $xy$-plane, and therefore any incident field $E$ can be decomposed into a component along $e_{||}$ (denoted $E_{||}$) and a component along $e_{\perp}$ (denoted $E_{\perp}$). Similarly, assume that the scattered field $E_s$ is arbitrarily polarized in the plane defined by $e_{||}$ and $e_{\perp}$, such that it can be decomposed into components $E_{||}$ and $E_{\perp}$. With the problem set up, solutions to the wave equations can be expressed in the spherical coordinate system.

### 3.2.2 Vector Harmonic Expression of Plane Waves

Since Mie theory deals with scattering from spheres, a description of wave propagation in spherical coordinates is naturally required. This in turn requires a solution of the wave equation in spherical coordinates. To obtain this solution, begin with the vector wave equation (3.1) and Maxwell’s equations (3.2) in a linear, isotropic, homogenous, source-free medium.

\[
\begin{align*}
\nabla^2 E + k^2 E &= 0 \\
\nabla^2 H + k^2 H &= 0 \\
\text{where } k &= \omega \sqrt{\mu \varepsilon} \\
\nabla \cdot E &= 0 \\
\nabla \cdot H &= 0 \\
\n\nabla \times E &= j \omega \mu H \\
\n\nabla \times H &= -j \omega \varepsilon E
\end{align*}
\]  

Now, define two new vector functions $M$ and $N$ that are solutions to the vector wave equation, in analogy to $E$ and $H$.

\[
\begin{align*}
M &= \nabla \times (e_{||}) \\
N &= \frac{\nabla \times M}{k}
\end{align*}
\]
In (3.3), \( \mathbf{c} \) is an arbitrary constant vector, \( \psi \) is a scalar function, and \( k = 2\pi / \lambda \). The vector functions \( \mathbf{M} \) and \( \mathbf{N} \) will satisfy the vector wave equation (3.1) if \( \psi \) satisfies the scalar wave equation given by
\[
\nabla^2 \psi + k^2 \psi = 0 \tag{3.4}
\]
If the constant vector \( \mathbf{c} \) is chosen to be equal to the radial vector \( \mathbf{r} = r \mathbf{e}_r \), then \( \mathbf{M} \) is a solution to the vector wave equation in spherical coordinates and is given by
\[
\mathbf{M} = \nabla \times [r \psi(r, \theta, \phi)] \tag{3.5}
\]
This choice of \( \mathbf{c} \) also makes \( \mathbf{M} \) tangential to any spherical surface. The scalar wave equation in spherical coordinates is given by substituting the relationships for the spherical unit vectors into the Cartesian scalar wave equation. This gives
\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} + k^2 \psi = 0 \tag{3.6}
\]
The solution to the spherical scalar wave equation (3.6) involves Bessel and Legendre functions, both of which have spherical symmetry. After solving equation (3.6), solutions are obtained for \( \psi \) that possess both even and odd symmetries about the azimuth. The solutions to the scalar wave equation are given by
\[
\begin{align*}
\psi_{emn}(r, \theta, \phi) &= \cos(m\phi) P_n^m(\cos \theta) z_n(kr) \\
\psi_{omn}(r, \theta, \phi) &= \sin(m\phi) P_n^m(\cos \theta) z_n(kr)
\end{align*} \tag{3.7}
\]
Here, \( m \) and \( n \) are integers with \( n \geq m \), and the subscripts \( e \) and \( o \) denote even and odd symmetry respectively. The set of functions \( P_n^m \) are Legendre functions of the first kind, order \( m \), degree \( n \). The set of functions \( z_n \) can be any one of the four spherical Bessel functions.

The four spherical Bessel functions \( j_n(\rho), y_n(\rho), h_n^{(1)}(\rho), \) and \( h_n^{(2)}(\rho) \) are defined through the more familiar Bessel functions of the first and second kind, \( J_\nu \) and \( Y_\nu \). The spherical Bessel functions are given by
The last two functions in (3.8) are the spherical Bessel functions of the third kind, also called the spherical Hankel functions. In reality, any linear combination of $j_n(\rho)$ and $y_n(\rho)$ will satisfy the wave equation, but the two forms of $j_n(\rho)$ and $y_n(\rho)$ will become important later on. Like the standard Bessel functions, the spherical Bessel functions also satisfy recursion relationships whereby all higher order functions can be generated from the $0^{th}$ and $1^{st}$ order functions. These recursion relationships are given by

$$
\begin{align*}
  j_0(\rho) &= \frac{\sin \rho}{\rho} \\
  j_1(\rho) &= \frac{\sin \rho}{\rho^2} - \frac{\cos \rho}{\rho} \\
  y_0(\rho) &= -\frac{\cos \rho}{\rho} \\
  y_1(\rho) &= -\frac{\cos \rho}{\rho^2} - \frac{\sin \rho}{\rho}
\end{align*}
$$

(3.9)

$$
\begin{align*}
  z_n(\rho) &= z_{n-1}(\rho) + z_{n+1}(\rho) \\
  (2n+1) \frac{\partial}{\partial \rho} z_n(\rho) &= nz_{n-1}(\rho) - (n+1)z_{n+1}(\rho)
\end{align*}
$$

(3.10)

Looking now at $\psi_{enn}$ and $\psi_{omm}$, some interesting observations can be made. First of all, it is known that standard Bessel functions and Legendre functions form complete basis sets, similar to sinusoids. Therefore the functions $\psi_{enn}$ and $\psi_{omm}$ must themselves form a complete set, since they consist of spherical Bessel functions and Legendre functions. This means that any function satisfying the scalar wave equation in spherical coordinates can be expanded out as an infinite series of $\psi_{enn}$ and $\psi_{omm}$, in the same way that any function can be expanded as an infinite series.
of sinusoids. Finally, a series of even- and odd-symmetric solutions to the vector wave equation can be found by substituting (3.7) into (3.3). Using $\rho = kr$, this substitution gives

$$M_{\text{even}} = \frac{-m}{\sin \theta} \sin(m\phi) P_n^m(\cos \theta) z_n(\rho) e_\phi - \cos(m\phi) \frac{dP_n^m(\cos \theta)}{d\theta} z_n(\rho) e_\phi$$

(3.11)

$$M_{\text{odd}} = \frac{m}{\sin \theta} \cos(m\phi) P_n^m(\cos \theta) z_n(\rho) e_\phi - \sin(m\phi) \frac{dP_n^m(\cos \theta)}{d\theta} z_n(\rho) e_\phi$$

(3.12)

$$N_{\text{even}} = \frac{z_n(\rho)}{\rho} \cos(m\phi) n(n+1) P_n^m(\cos \theta) e_r$$

$$+ \cos(m\phi) \frac{dP_n^m(\cos \theta)}{d\theta} \frac{1}{\rho} \frac{d}{d\rho} (\rho z_n(\rho)) e_\theta$$

(3.13)

$$N_{\text{odd}} = \frac{z_n(\rho)}{\rho} \sin(m\phi) n(n+1) P_n^m(\cos \theta) e_r$$

$$+ \sin(m\phi) \frac{dP_n^m(\cos \theta)}{d\theta} \frac{1}{\rho} \frac{d}{d\rho} (\rho z_n(\rho)) e_\theta$$

(3.14)

The equations (3.11) - (3.14) are the vector harmonic solutions to the vector wave equation. They are called this since they are related to one another in increasing $m$ and $n$, somewhat like sinusoidal harmonics. It can also be shown [26] that the vector harmonics are mutually orthogonal.

Given that the vector harmonic solutions to the vector wave equation form an orthogonal, complete, basis set of functions, it is possible to expand a plane wave into an infinite series of spherical vector harmonics. This problem is basically one of solving for the expansion coefficients in the infinite series, as we will show. Solving for the expansion coefficients is the next step towards the goal of solving for the scattered field of a sphere, since many imaging systems will have incident fields that can be approximated by plane waves or superpositions of plane waves.

First, consider a plane wave that is linearly polarized along the x-axis. Although the following derivations assume x-polarization, since the particle is spherical the solutions for any other polarization state can be determined from the x-polarized solution. For example, if a
scattered field $E_s(r, \phi, \theta)$ results from an incident x-polarized field and the solution resulting from an incident y-polarized field is desired, one can simply write $E_s(r, \phi + \pi/2, \theta)$. The assumption of a plane wave is reasonable if the beam diameter is much larger than the largest scattering particle in the system. It should be noted that this assumption is not necessarily true for the interaction of OCT beams and cellular nuclei, as the OCT beam diameter can be on the same order of magnitude as the nuclei. However, the assumption is true when mitochondria and other cellular organelles are concerned. The expression for the incident wave in figure 15, propagating in the z direction, in spherical coordinates is given by

$$E_i = E_0 e^{i\omega \cos \theta} e_x$$

(3.15)

An expression for $E_i$ in terms of the vector harmonics is desired, such that

$$E_i = \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \left( B_{emn} M_{emn} + B_{cmn} M_{cmn} + A_{emn} N_{emn} + A_{cmn} N_{cmn} \right)$$

(3.16)

Here, $A$ and $B$ are the unknown expansion coefficients which depend somehow on the properties of the incident field. Since the vector harmonics are mutually orthogonal, the expansion coefficients must be of the form

$$B_{emn} = \frac{\int \int \int E_i \cdot M_{emn} \sin \theta \, d\theta \, d\phi}{\int \int \int |M_{emn}|^2 \sin \theta \, d\theta \, d\phi}$$

(3.17)

$$B_{cmn} = \frac{\int \int \int E_i \cdot M_{cmn} \sin \theta \, d\theta \, d\phi}{\int \int \int |M_{cmn}|^2 \sin \theta \, d\theta \, d\phi}$$

(3.18)

$$A_{emn} = \frac{\int \int \int E_i \cdot N_{emn} \sin \theta \, d\theta \, d\phi}{\int \int \int |N_{emn}|^2 \sin \theta \, d\theta \, d\phi}$$

(3.19)

$$A_{cmn} = \frac{\int \int \int E_i \cdot N_{cmn} \sin \theta \, d\theta \, d\phi}{\int \int \int |N_{cmn}|^2 \sin \theta \, d\theta \, d\phi}$$

(3.20)
By substituting (3.11) - (3.14) into (3.17) - (3.20), and constraining the incident field to be finite at the origin, it can be shown that

$$B_{emn} = A_{emn} = 0 \text{ for all } m, n$$
$$B_{emn} = A_{emn} = 0 \text{ for all } m \neq 1$$

(3.21)

The radial dependence of the vector harmonics must be described by the spherical Bessel function $j_n(kr)$, since it is finite at the origin. Following the notation of [26] and [27], a superscript (1) is appended to the vector harmonics to denote radial dependence on the function $j_n(kr)$. There is now a somewhat smaller expansion of the form

$$E_i = \sum_{n=1}^{\infty} \left( B_{o1n} M_{o1n}^{(1)} + A_{e1n} N_{e1n}^{(1)} \right)$$

(3.22)

The expansion coefficients $A_{e1n}$ and $B_{o1n}$ can be solved by integrating (3.19) and (3.18) respectively. This leads to the final solution of the incident electric and magnetic fields $E_i$ and $H_i$ through

$$A_{e1n} = -i(i^n) E_o (2n+1)$$
$$B_{o1n} = i^n E_o (2n+1)$$

(3.23)

$$E_i = E_o \sum_{n=1}^{\infty} \frac{i^n (2n+1)}{n(n+1)} \left( M_{o1n}^{(1)} - iN_{e1n}^{(1)} \right)$$
$$H_i = - \frac{k}{\omega \mu} E_o \sum_{n=1}^{\infty} \frac{i^n (2n+1)}{n(n+1)} \left( M_{e1n}^{(1)} + iN_{o1n}^{(1)} \right)$$

(3.24)

At this stage, the incident electric and magnetic plane waves that are impinging on the spherical particle can be expressed as an infinite summation of spherical vector harmonics. The solution to the scattered fields $E_s$ and $H_s$ can now be analytically determined.

### 3.2.3 Solution of the Scattered Field

To begin solving for the scattered field, the boundary conditions at the surface of the sphere must be considered. Specifically, there must be zero field amplitude tangential to the boundary such that
Here, $E_i$ and $H_i$ denote the fields inside the sphere. The fields inside the sphere must have the same form as (3.24) and the same radial dependence.

$$E_i = \sum_{n=1}^{\infty} E_n \left( c_n M_{eln}^{(1)} - id_n N_{eln}^{(1)} \right)$$

$$H_i = \frac{-k_i}{\omega \mu_i} \sum_{n=1}^{\infty} E_n \left( d_n M_{eln}^{(1)} + ic_n N_{eln}^{(1)} \right)$$

Here, $k_i$ is the wavenumber inside the sphere, $\mu_i$ is the magnetic permeability inside the sphere, and

$$E_n = \frac{i^n E_0 (2n+1)}{n(n+1)}$$

The solution for the scattered field can involve both $j_n(kr)$ and $y_n(kr)$, since the scattered field does not exist at the origin and both of these functions are finite for points away from the origin. Looking at the last two equations in (3.8), the radial dependence of the scattered field can be represented by either $h_n^{(1)} (kr)$ or $h_n^{(2)} (kr)$. For values of $kr \gg n^2$, the spherical Hankel functions are approximated as

$$h_n^{(1)} (kr) \approx \frac{(-i)^n e^{i kr}}{ikr}$$

$$h_n^{(2)} (kr) \approx \frac{(-i)^n e^{-i kr}}{ikr}$$

Since the first expression in (3.28) corresponds to a wave traveling away from the origin and the second expression corresponds to a wave traveling towards the origin, we select only $h_n^{(1)} (kr)$ to represent the radial dependence of the scattered field. Following the notation of [27] and [26], a superscript (3) is appended to the vector harmonics in the scattered field to denote radial dependence on the spherical Hankel function $h_n^{(1)} (kr)$. Therefore we have
\[
\begin{align*}
\mathbf{E}_z &= \sum_{n=1}^{\infty} E_n \left( ia_n N_{eln}^{(3)} - b_n M_{eln}^{(3)} \right) \\
\mathbf{H}_z &= \frac{k}{\omega \mu} \sum_{n=1}^{\infty} E_n \left( ib_n N_{eln}^{(3)} + a_n M_{eln}^{(3)} \right)
\end{align*}
\] (3.29)

In both (3.26) and (3.29), the expansion coefficients \(a_n, b_n, c_n\), and \(d_n\) relate to the scattering properties of the system, and are therefore called the “scattering coefficients”. Of principal interest are \(a_n\) and \(b_n\), since they describe the scattered field. The value of these coefficients depend on the sphere radius, incident wavelength, and refractive indices of the sphere and medium. The scattering coefficients can be solved for exactly from boundary conditions and orthogonality arguments, and by using the expressions for the vector harmonics and (3.24), (3.26), and (3.29). Assuming that the magnetic permeabilities of the sphere and medium are equal, the coefficient values are given by

\[
a_n = \frac{m \psi_n(mx) \psi'_n(x) - \psi_n(x) \psi'_n(mx)}{m \psi_n(mx) \xi'_n(x) - \xi_n(x) \psi'_n(mx)} \quad (3.30)
\]

\[
b_n = \frac{\psi_n(mx) \psi'_n(x) - m \psi_n(x) \psi'_n(mx)}{\psi_n(mx) \xi'_n(x) - m \xi_n(x) \psi'_n(mx)} \quad (3.31)
\]

Here, two new parameters \(m\) and \(x\) and two new functions \(\psi_n(\rho)\) and \(\xi_n(\rho)\) have been introduced. They are defined as

\[
m = \frac{n_p}{n_m}, \text{ the relative refractive index of the particle to the medium}
\]

\[
x = \frac{2 \pi n_m a}{\lambda}, \text{ the size parameter where } a \text{ is the sphere radius}
\]

\[
\begin{align*}
\psi_n(\rho) &= \rho j_n(\rho) \\
\xi_n(\rho) &= \rho h_n^{(1)}(\rho)
\end{align*}
\] (3.32)

The prime notation in (3.30) and (3.31) represents derivation with respect to the argument of the function. At this point, a very significant result is reached. Equations (3.29), (3.30), (3.31), and (3.11) - (3.14) allow an exact solution for the scattered field \(\mathbf{E}_s\) at any point \((r, \theta, \phi)\) as a function of the spatial coordinates, sphere radius, sphere and medium refractive indices, and incident field wavelength.
3.2.4 **Spatial Distribution of Scattered Light**

Now that a plane wave can be expressed as an infinite series of spherical waves (the vector harmonics), it is informative to look in more detail at these basis functions. The vector harmonics can be thought of as the modes excited by a plane wave incident upon a scattering sphere, in the same way that modes in an optical waveguide are excited by an incident field. In this section, the dependence of the vector harmonics on the scattering angle \( \theta \) is examined and the backscattering amplitude is calculated as a function of wavelength.

It is possible to make the form of equations (3.11) - (3.14) more compact by introducing the functions \( \pi_n(\zeta) \) and \( \tau_n(\zeta) \). These functions contain the angular dependence of the vector harmonics, and are useful for visualizing the shape of the scattered field resulting from each term in the harmonic expansion.

\[
\begin{align*}
\pi_n(\zeta) &= \frac{P_n^1(\zeta)}{\sin \theta} \\
\tau_n(\zeta) &= \frac{dP_n^1(\zeta)}{d\theta}
\end{align*}
\]

(3.33)

This gives the new expressions for the vector harmonics as

\[
\begin{align*}
M_{eln} &= -\sin(\phi)\pi_n(\cos \theta) z_n(\rho) e_\theta - \cos(\phi) \tau_n(\cos \theta) z_n(\rho) e_\phi \\
M_{oln} &= \cos(\phi)\pi_n(\cos \theta) z_n(\rho) e_\theta - \sin(\phi) \tau_n(\cos \theta) z_n(\rho) e_\phi \\
N_{eln} &= \frac{z_n(\rho)}{\rho} \cos(\phi) n(n+1) \sin(\theta) \pi_n(\cos \theta) e_r \\
&\quad + \cos(\phi) \tau_n(\cos \theta) \frac{1}{\rho} \frac{d}{d\rho} \left( \rho z_n(\rho) \right) e_\theta \\
&\quad - \sin(\phi) \pi_n(\cos \theta) \frac{1}{\rho} \frac{d}{d\rho} \left( \rho z_n(\rho) \right) e_\phi \\
N_{oln} &= \frac{z_n(\rho)}{\rho} \sin(\phi) n(n+1) \sin(\theta) \pi_n(\cos \theta) e_r \\
&\quad + \sin(\phi) \tau_n(\cos \theta) \frac{1}{\rho} \frac{d}{d\rho} \left( \rho z_n(\rho) \right) e_\theta \\
&\quad + \cos(\phi) \pi_n(\cos \theta) \frac{1}{\rho} \frac{d}{d\rho} \left( \rho z_n(\rho) \right) e_\phi
\end{align*}
\]

(3.34) - (3.37)
**Figure 16.** $\pi$ angular dependence of vector harmonics for expansion levels $n = 2$ to $n = 5$

**Figure 17.** $\pi$ angular dependence of vector harmonics for expansion levels $n = 2$ to $n = 5$

**Figure 18.** Backscattered field amplitude for spherical particles of diameter 200 nm, 800 nm, 2 $\mu$m, and 2 $\mu$m. Particle refractive index is 1.38, surrounding medium is 1.33 to simulate biological tissue. Spectral modulation increases as the particle diameter increases.

**Figure 19.** Backscattered field amplitude for spherical particles of diameter 800 nm, with randomly distributed particle spacing. Number of particles is 1, 10, 100, and 1000.
From equations (3.34) - (3.37), it is evident that the dependence of the vector harmonics on the scattering angle \( \theta \) is completely contained in the functions \( \pi_n(\zeta) \) and \( \tau_n(\zeta) \). Figure 16 and figure 17 show polar plots of \( \tau_n[\cos(\theta)] \) and \( \pi_n[\cos(\theta)] \) for the first five orders of the expansion (\( n \) ranging from 2 to 5). Blue lines indicate a positive value of the function while red lines indicate a negative value. In these plots, \( \theta = 0^\circ \) corresponds to complete forward scattering while \( \theta = 180^\circ \) corresponds to complete backscattering.

At this point, it is useful to look at the weightings of the scattering coefficients \( a_n \) and \( b_n \) of the vector harmonics in \( E_x \), as \( n \) (the expansion coefficient index, not the refractive index) increases in (3.29). For a given system, the infinite series of vector harmonics \( N_{\text{el}}^{(3)} \) and \( M_{\text{el}}^{(3)} \) can be truncated at some expansion level \( n = n_c \) and an arbitrarily small error between the estimated field and the exact field can be obtained. If the scattering coefficients are numerically calculated, the magnitudes of \( a_n \) and \( b_n \) decrease as \( n_c \) increases [26]. The contribution to \( E_x \) from the \( n^{\text{th}} \) term becomes insignificant around

\[
 n_c \approx x + 4x^{1/3} + 2 \tag{3.38}
\]

where \( x \) is the size parameter defined earlier. Therefore, as the particle size increases relative to the incident wavelength, the scattered field will be increasingly influenced by the higher-order expressions of \( \pi_n(\zeta) \) and \( \tau_n(\zeta) \). Going back to the analogy of vector harmonics as scattering modes, this can be thought of as increasing the dimensions of an optical waveguide relative to the propagating wavelength. This results in an increased number of high-order modes being supported in the system, similarly to what is observed in Mie scatterers.

Some interesting conclusions can be drawn from the plots of figure 16 and figure 17. First, all the functions are even-symmetric about a line from \( 0^\circ \) to \( 180^\circ \). This is expected, since the scattering particles are spherical and therefore have the same symmetry, and the incident light is propagating along the axis of symmetry. Also note that as \( n_c \) increases, the widths of the forward and backward scattering lobes decrease, indicating that the scattered field will be distributed in a smaller spatial area as the particle size increases. Furthermore, only half of the functions \( \pi_n(\zeta) \) and \( \tau_n(\zeta) \) have a backscattering lobe at all, whereas all the functions have a
forward scattering lobe. This suggests that as particle size increases, there will be decreased backscattering and increased forward scattering.

This result is critical for OCT imaging, and suggests that smaller scattering particles in biological tissue will produce most of the OCT signal. For example, in ophthalmic and epithelial tissue imaging, it is often observed that nucleated tissue layers show up as dark regions on the OCT tomogram, indicating a very low signal. This effect is seen with the inner nuclear layer and outer nuclear layer during ophthalmic imaging, and with the epithelial basal cell layer in epithelial tissue imaging. Other retinal layers, such as the inner plexiform layer and outer plexiform layer, which are composed of synaptic connections, show up as bright regions on OCT tomograms. This may confound efforts to detect nucleated cell layers, whether contrast is provided by intensity information or spectroscopic information, since the backscattered signal from the nuclei will be quite low.

Of central importance to spectroscopic OCT imaging is the relationship between backscattering intensity and incident wavelength for biological particle sizes. Figure 18 shows a series of backscattering spectra for particle diameters of 200 nm, 800 nm, 1 μm, and 8 μm. These spectra were calculated using a refractive index of 1.38 for the scattering particles and 1.33 for the surrounding media, in order to approximate biological tissue. From numerical calculation of the scattered field, it can be seen that as the sphere size increases, the spectral modulation also increases. This suggests that by quantifying the degree of spectral modulation in the backscattered spectrum at each pixel in an OCT image, information about the size of the scattering particles can be obtained, providing an alternative contrast modality. Such a contrast modality should provide enhanced differentiation of tissue layers where the average scattering particles size varies significantly. For example, if particular layer is composed of cells with increased mitochondria density or increased nuclear size, examining the spectral modulation may greatly enhance the contrast of these layers.

3.2.5 **Scattering from Multiple Particles**

In OCT imaging of biological tissue, the incident beam illuminates many scattering particles, not just a single sphere. This situation arises when the beam’s transverse dimension is much larger than the scattering particle size, when the particle density is such that there is likely to be more than one particle present in the beam at any given time, and when the detector collects light
over a large surface such that the contributions from multiple particles become integrated. It is useful to look at the relationship that governs this situation so as to better understand the scattering spectra detected by actual OCT systems.

First, define an active volume within which there will be multiple scattering particles. The active volume can be considered to be a cylinder defined by the incident beam’s waist \( w \) and the coherence length of the source \( l_c \). Within the active volume, under coherent illumination, contributions to the scattered field from multiple particles add coherently. Therefore the net scattered field will be roughly equal to the sum of the fields scattered by the individual spheres. If a distribution exists of \( N(a) \) particles per unit volume with radius \( a \), then the total scattered field is given approximately by

\[
E_s \approx \int N(a)E_s(a)da
\]  

(3.39)

The limits of integration are the range of sphere diameters in the distribution.

Figure 19 shows the net scattered field for various distributions of particles with uniform diameter but randomly distributed separation. As the number of particles in the distribution increases, the modulation pattern observed in the optical spectrum varies significantly. The general trend, however, is that the degree of modulation appears to increase with the number of particles. Therefore, even though the main source of backscattered signal in an OCT system may be caused by particles smaller than the wavelength, significant spectral modulation may still be observed. This is central to the following discussion of spectroscopic OCT imaging, and explains the results presented in section 33.4.

3.3 Visualization of Spectroscopic Information

Standard intensity-based OCT systems represent the information about tissue backscattering as a greyscale or false colour map image. In these systems, the intensity of the backscattered light at each pixel in the image is stored as a single value only. By acquiring and processing the full interferometric OCT signal instead of the envelope only, it is possible to calculate a local spectrum for every pixel in the image using a time-frequency analysis such as the short-time Fourier transform (STFT). By generating an optical spectrum for every A-scan in a given OCT image a 4-dimensional data set is produced, defined by transverse coordinate, axial coordinate, wavelength, and intensity. For a human reader to interpret the information contained in the
spectroscopic data set, it is necessary first to condense the data to three dimensions. This is done by choosing an appropriate spectroscopic metric (such as center wavelength) that quantifies the spectrum at each pixel with a single value (such as 800 nm). The process of choosing an appropriate metric is discussed fully in section 3.4 below. This section deals with the various methods for generating optical spectra and for visualizing the spectroscopic information after processing with a given metric.

### 3.3.1 Calculation of Optical Spectra

Previous work in spectroscopic OCT [14] has used an STFT approach to calculate local spectra at each pixel in the OCT image. The “Morlet wavelet” method described in [14] is actually not a wavelet transform in the standard sense, but is equivalent to an STFT using a Gaussian window function. Compared to other time-frequency analysis methods, such as an actual wavelet transform, the STFT provides the most meaningful transformation from time- to frequency-domain. There exists a direct mapping of RF frequency to optical wavelength when the STFT is used, via (1.2), and therefore it is easy to relate observations in the calculated spectra to physical phenomena. For example, when an STFT analysis is used, dips in the calculated spectra that match the absorption profile of known biological agents can be used to infer the presence of the agents. In other time-frequency analyses, no direct mapping from the transform coordinate to optical wavelength exists.

The STFT is calculated by taking the discrete Fourier transform (DFT) of windowed segments of each A-scan. The window is translated through the A-scan until the entire depth of the image has been analyzed. For every window location, one RF spectrum is calculated. The RF spectrum can be analyzed as-is, or converted to an optical spectrum through (1.2). Mathematically, the discrete STFT is given by

\[
X[n,k] = \sum_{m=0}^{L-1} x[n+m]w[m]e^{-j(2\pi/L)km}
\]

(3.40)

Here, \(x[n]\) is the input signal of length \(N\). The window function \(w[m]\) is of length \(L\), resulting in each STFT segment also being of length \(L\). It is possible to increment \(n\) by one point, in which case \(N\) spectra will be calculated, or by \(p\) points, in which case \(N/p\) spectra will be calculated. The window function plays an important role in an STFT analysis, since its
Figure 20. Window functions used in STFT and CZT analysis. A window with a narrow spectrum and low sidelobes is desirable to prevent corruption of the local signal spectra.

Figure 22. STFT segment from Xenopus Laevis tadpole, calculated using the CZT and FFT algorithms, each with 512 points. Frequency resolution in the area of interest using the CZT is ~30 times higher than using the FFT. The FFT data set extends from $-f_s/2$ to $+f_s/2$, while the CZT data set extends only from $f_0$ to $f_1$.

Figure 23. Complete STFT from one A-scan of 250 μm glass coverslip, calculated using the CZT algorithms with 512 points. Both surfaces of the coverslip are visible, with the reflected spectra corresponding to that of the laser source.
spectrum is convolved with the spectrum of $x[n]$ during the transform. Therefore, it is desirable to choose a window that has a narrow bandwidth and low sidelobe levels to reduce the effects of the window on the calculated spectra. Several choices of window functions are possible, with the most common being the Hanning, Hamming, Blackman, and Kaiser windows. Figure 20 shows the window functions, while figure 21 shows the spectra of the window functions. A fundamental tradeoff exists between the bandwidth of the window spectrum and the sidelobe amplitude. For the spectroscopic OCT analyses done for this thesis project, the Hamming window was used since it provides a good compromise between bandwidth and sidelobe level.

The basic element in the calculation of an STFT is the DFT. Although the DFT can be calculated very efficiently using the “fast Fourier transform” (FFT) class of algorithms, it suffers from a significant drawback when applied to spectroscopic imaging: the FFT algorithm calculates the DFT over a fixed frequency range of $0 \rightarrow 2\pi$, using $L$ points. If the discrete RF spectrum of the interferometric OCT signal does not fill the entire range $0 \rightarrow 2\pi$, the result is that unnecessary points in the transform are calculated, decreasing the efficiency of the algorithm. In other words, if the OCT system is not operating near the Nyquist sampling limit (when the full $0 \rightarrow 2\pi$ frequency range is covered), the STFT will take longer than is necessary to produce spectra with a given frequency resolution. Since it is common for OCT systems to operate at sampling rates well above the Nyquist limit to reduce aliasing noise in the A/D converter, the fixed frequency span of the FFT is inefficient.

Furthermore, the FFT frequency resolution is fixed at $2\pi/L$, where $L$ is the length of the input signal. This can be problematic for spectroscopic OCT, since it is desirable to calculate the optical spectrum with as fine a resolution as possible to improve the accuracy of the spectroscopic metrics. It is possible to increase the frequency resolution by zero-padding the input signal, but this results in a longer transform to calculate. In addition, zero-padding compounds the effect of not filling the entire range of discrete frequencies, since zero-padding by a factor of $P$ will also increase the number of unnecessary points by a factor of $P$.

To overcome the drawbacks of the FFT, a chirp z transform (CZT) was used to compute the STFT's used in spectroscopic OCT imaging. The CZT is a more general case of the DFT, where integration can be performed over an arbitrary curve in the z-plane instead of being limited to the unit circle. The DFT is given by
\[ X[k] = \sum_{n=0}^{L-1} x[n] e^{-j(2\pi/L)nk}, \quad 0 \leq k < L-1 \]  
(3.41)

while the CZT is given by

\[ X_{\text{CZT}}[k] = \sum_{n=0}^{L-1} x[n] z^n e^{-j r n}, \quad 0 \leq k < L-1 \]  
(3.42)

with \( z \in \mathbb{I} \). The choice of \( z \) controls the starting point of the transform, while the factor \( r \) controls the frequency resolution (and therefore the end point of the transform). By choosing \( z = e^{-j\omega_0} \) and \( r = (\omega_1 - \omega_0)/L \), a DFT can be calculated from \( \omega_0 \rightarrow \omega_1 \) instead of \( 0 \rightarrow 2\pi \). In this case, the CZT gives

\[ X_{\text{CZT}}[k] = \sum_{n=0}^{L-1} x[n] e^{-j(\omega_0 + (\omega_1 - \omega_0)k/L)n}, \quad 0 \leq k < L-1 \]  
(3.43)

The frequency resolution in the range of interest \( \omega_0 \rightarrow \omega_1 \) can still be increased by zero-padding, but by using the CZT there will be no extraneous points calculated that fall outside the region of interest.

The frequency region of interest can be arbitrarily adjusted to match the OCT system’s optical bandwidth and A/D sampling rate. For example, the system used for most of the spectroscopic imaging in this thesis project had a center wavelength of 800 nm, an optical bandwidth of 200 nm, and a galvanometer velocity of 3.6 mm/s. The A/D sampling frequency was \( f_s = 333 \text{ kHz} \), divided between the OCT channel and the calibration fringe channel, for an effective sampling frequency of 166 kHz per channel. Assuming that the optical wavelengths containing useful information fall between 600 nm and 1000 nm, from (1.2) the RF frequencies of interest fall between 7.2 kHz and 12 kHz. After A/D conversion, the continuous RF frequency \( f_{\text{RF}} \) [Hz] is mapped into a discrete RF frequency \( \omega_{\text{RF}} \) [rad/s] by the relationship

\[ \omega_{\text{RF}} = \frac{2\pi}{f_s} f_{\text{RF}} \]  
(3.44)

Therefore the discrete frequencies of interest fall between \( \omega_0 = 0.087\pi \) rad/s and \( \omega_1 = 0.145\pi \) rad/s. Using an FFT algorithm with \( L = 512 \) points to calculate the spectrum would result in a frequency resolution of \( \pi / 256 \) rad/s, corresponding to \( \sim 15 \) points in the range \( \omega_0 \rightarrow \omega_1 \). Using a CZT of length \( L = 512 \), all 512 points can be chosen to fall in the range \( \omega_0 \rightarrow \omega_1 \), giving a frequency resolution of \( \sim \pi \times 10^{-4} \) rad/s. Alternatively, if frequency resolution is not critical for
the spectroscopic metric being used, the CZT length can be chosen to be \( L = 15 \). This would result in the same resolution as the standard FFT, but with a factor of 30 reduction in computation time.

Figure 22 shows an STFT segment calculated from one A-scan of a Xenopus Laevis tadpole, using an FFT of length \( L = 512 \) and a CZT of length \( L = 512 \). For the CZT, \( \omega_0 = 0.087\pi \) and \( \omega_1 = 0.145\pi \). For clarity, the x-axis values are converted back to continuous RF frequencies. Note that within the frequency range containing useful information, the CZT implementation produces a more accurate spectrum than the FFT implementation. Although the additional points contained in the CZT are a form of oversampling in the frequency domain (ie, they are not required for perfect reconstruction of the signal from its transform), they produce a more accurate spectrum that is better suited for spectroscopic OCT analysis.

Figure 23 shows a complete STFT calculated from one A-scan of a glass coverslip, using the CZT method. This figure illustrates the general nature of the data collected for every A-scan in an OCT image. The coverslip was 250 \( \mu \)m thick, and produced two strong reflections at each surface. The spectra at each surface correspond roughly to the spectrum of the laser used to perform the imaging. The two spectra are not identical, possibly due to a beam-focusing effect. Since the different wavelengths in the incident beam are in focus at different axial locations in the image, the incident spectra at each axial location will vary. Furthermore, the transmission of the glass may not be entirely uniform with wavelength. In this case, the net effect is a slight blue-shift of the spectrum at the second surface of the coverslip. This illustrates the difficulty in using spectroscopic metrics based on characteristics of the spectra at absolute wavelengths. For example, calculating the center wavelengths of the two spectra in figure 23 will produce significantly different values, when in fact no “contrast” should be observed in this case between the two surfaces.

3.3.2 Time / Frequency Resolution Using the Chirp Z Transform

At this point it is useful to consider more closely the resolution of spectroscopic OCT imaging, both in the time domain and frequency domain, when a CZT is used to generate local spectra at each pixel in the image. As discussed in section 3.3.1, the frequency resolution of a standard FFT is set exclusively by the length of the signal being transformed. In an STFT analysis of a spectroscopic OCT A-scan, this corresponds to the length of the window function.
Therefore, for the FFT, there exists a simple tradeoff between time domain resolution and frequency domain resolution. As the length $L$ of the window function increases, the frequency resolution increases proportionally. However, the ability to localize spectral features in time (space) decreases proportionally.

When the CZT is used instead of the FFT for an STFT analysis, the relationship between time and frequency resolution is not as simple. A window function of length $L$ now produces a spectrum with frequency resolution $(\omega_1 - \omega_0)/L$. However, it should be emphasized that this does not provide any additional spectral information than the FFT, in the sense that an identical spectrum over the range $\omega_0 \rightarrow \omega_1$ could be generated using an FFT with zero-padding. The CZT is merely a method to generate high frequency-resolution data in a shorter time than is possible using the FFT, for a given spatial resolution $L$. From another perspective, the CZT is equivalent to oversampling the continuous Fourier transform while the FFT is equivalent to critically sampling the continuous Fourier transform. In order to obtain additional spectral information with the CZT, it is still necessary to increase the length of the window function.

It is therefore necessary to have some sense of the axial dimension over which the spectroscopic features of interest in a given image will remain relatively uniform. For example, if the spectroscopic properties of biological particles with diameter 5 $\mu$m are being investigated, then a window length of $\sim 5 \mu$m should be used for the STFT analysis. Using a shorter window will produce less accurate spectra, regardless of zero-padding (since zero-padding is equivalent to sinc interpolation between known data points). Using a longer window will mix the spectra of the particles of interest with the spectra of adjacent material, also producing less accurate spectra. In reality, it is necessary to experiment with several different window sizes to obtain the best length for a given biological tissue and spectroscopic metric.

### 3.3.3 Pure Spectroscopic Visualization

Having described the method for calculating an optical spectrum for every pixel in an OCT image, it is now important to understand how to visualize the spectroscopic data. It is convenient and intuitive to use a colour map to represent spectroscopic information, where the colour can convey specific information about the spectra at each pixel. For this thesis project, the Hue/Saturation/Luminance (HSL) colour map was used as a starting point. The HSL map allows for independent adjustment of three values (H, S, and L) for every pixel. As H varies from 0 to
Figure 24. HSL colour maps for spectroscopic OCT. First map uses hue to encode spectroscopic data, saturation to encode backscattered intensity (structural) data, and luminance is fixed at 0.7. Second map uses hue to encode spectroscopic data, but uses both saturation and luminance to encode structural data. More vivid colours and better structural contrast are obtained using the second map.

Figure 25. Spectroscopic OCT image showing center wavelength of reflected spectrum of a mirror. The chirp of the incident pulse appears as a chirp in the detected spectrum, as shown in the depth-varying center wavelength. As the reference dispersion is changed along the transverse coordinate, the chirp varies from short-long to long-short.

Figure 26. Standard in vivo intensity-based OCT image (A) and in vivo spectroscopic OCT image (B) of a Syrian golden hamster cheek pouch. Center wavelength was used for the spectroscopic metric. Blood vessels are clearly visible in the spectroscopic image, and arterial (red) and venous (green) areas can be distinguished. Spectroscopic noise is visible from red-shifting and incident pulse chirp.
1, the colour will vary from red to green to blue to violet and back to red. As S varies from 0 to 1, the colour will vary from grey to fully saturated. As L varies from 0 to 1, the colour will vary from black to white.

There are several possible ways to visualize spectroscopic OCT information. If the structural information relating to the intensity of the backscattered light is not important, then a pure spectral visualization technique is possible. In this case, the spectroscopic metric values (such as center wavelength) for each pixel can be encoded into a greyscale map or arbitrary HSL map and displayed. In reality, this technique is not a good choice, since a better understanding of the tissue being imaged can be obtained by combining the intensity (structural) and spectral information into the same image.

3.3.4 Spectroscopic / Intensity Visualization

When attempting to simultaneously display structural and spectroscopic information in a HSL colour map, it is natural to use hue to encode the spectroscopic information, and saturation and luminance to encode the intensity information. Previous work [14] used a fixed luminance value, encoding intensity information in saturation only. This results in the “darkest” areas of the image being grey, with the brightest areas being fully saturated. It is also possible to encode intensity information in saturation and luminance simultaneously, resulting in the darkest areas of the image being black and the brightest being fully saturated. Better structural contrast and more vivid colours are obtained using the second method, as illustrated in figure 24. In this case, a hue range of 0 to 0.6 was used to avoid wrapping of the map back to red. Saturation and luminance are equal, and range from 0 to 1. The second HSL colour map was used for all of the work presented in this thesis project.

Two types of spectroscopic / intensity visualization methods were investigated for this thesis project. In the first method, the intensity of the backscattered light was calculated from the STFT data by integrating the power in each local spectrum. In this case,

\[ I_1[n] = \sum_{k=0}^{L-1} X[n,k] \]  

(3.45)

This method represents a “pure” spectroscopic visualization technique, since all of the information in the displayed image comes from the calculated STFT spectra. The drawback of this method results from the time/frequency resolution tradeoff discussed in section 3.3.2. As the
window size increases, the ability to localize spectral information and structural information is degraded, resulting in an image that appears blurry in the axial direction.

To maintain the sharpness of standard OCT images, a hybrid visualization technique was investigated next. In this method, the spectral information was obtained from the STFT spectra, but the structural information was obtained directly from the demodulated A-scans. After the STFT was calculated for each A-scan, the Hilbert demodulation technique discussed in section 2.2 was applied to the A-scan. The value of the demodulated A-scan corresponding to the center of each STFT window was mapped into equal saturation and luminance values. This way, the spatial resolution of the structural data was decoupled from the spatial resolution of the spectral data. Images processed in this manner appeared to have sharp intensity (structural) features, with blurring due to the window effect limited to the pixel hue only.

3.3.5 Spectroscopic / Greyscale Visualization

In the techniques described in section 3.3.4, the entire image appears in colour. Other medical imaging techniques that make use of spectral information, such as Doppler flow ultrasound, often display the spectral data as a colour map overlaid on a greyscale map. A threshold for the spectroscopic data is selected, where “insignificant” data points are displayed in greyscale and “significant” data points in colour. The intended effect is to focus attention on the regions of the image with the most spectroscopic activity. This visualization technique was investigated for spectroscopic OCT imaging, but was found to be unsuitable for most circumstances. Due to the lack of highly focal spectroscopic properties, unlike the case of Doppler flow, the blending of greyscale and colour pixels make the images appear noisy and random.

3.4 Spectroscopic OCT Metrics

The final component in the design of the spectroscopic OCT analysis method was to evaluate different techniques for quantifying the local STFT spectra. As previously mentioned, it is necessary to condense the spectroscopic dataset to three dimensions for encoding in the colour map of 3.3.4. Although there are an unlimited number of possible metrics that can be considered, this thesis project examined two main classes of metrics: those that measure power
relative to wavelength, and those that measure the degree of spectral modulation. Within each class, several specific metrics exist.

Two metrics that fall within the first class are center wavelength and spectral power over a predefined bandwidth. Metrics of this type are useful for providing contrast based on the absorption profiles of the materials in the tissue, especially if a certain substance with a known profile is of interest. For example, with a priori knowledge of the absorption profile of blood, it is possible to quantify the power of each local spectrum over a wavelength range where blood has a characteristic absorption peak or dip. Areas with high concentrations of blood can thereby be differentiated from areas without blood. Center wavelength is a more ubiquitous spectroscopic metric, and does not map directly to known biological phenomena. It was used in earlier spectroscopic OCT work [14], but is very qualitative in nature. Both of the metrics in this class also suffer from high sensitivity to spectroscopic noise, as discussed fully below.

Metrics based on quantifying spectral modulation have not been previously investigated for spectroscopic OCT imaging, and are based on the fact that particles of different sizes impart varying degrees of spectral modulation to the backscattered light. The theory behind this effect is discussed fully in section 3.2. There are many possible ways to quantify spectral modulation, and some of them are discussed below. Techniques investigated for this thesis project were: fit to a Gaussian profile; autocorrelation function bandwidth; autocorrelation function polynomial fit; autocorrelation function peak counting; and distribution of power in the Fourier transform of the spectrum. This class of techniques is shown experimentally to be insensitive to spectroscopic noise.

3.4.1 Sources of Spectroscopic Noise

In order to design new spectroscopic metrics that are insensitive to noise, it is first necessary to understand the sources of spectroscopic noise in OCT systems. One such source, nonlinear galvanometer motion, was discussed fully in section 2.1. To briefly summarize, nonlinearity in the motion of the optical path length scanner imparts a time-varying Doppler shift to the detected spectra. This Doppler shift can be larger than variations in the spectra caused by the tissue itself, essentially swamping the system with noise. As discussed above, the Doppler shift from the galvanometer can be removed by using a reference interferometer at the cost of increased system complexity and the need for time-consuming post-processing. It would therefore be more
efficient to use a spectroscopic metric that is insensitive to Doppler shifts induced by the galvanometer.

Another source of spectroscopic OCT noise is related to the broadband light source. When a solid-state laser is used as the OCT system light source, the broad optical spectrum is generated by femtosecond pulses. These pulses are temporally chirped, such that the center wavelength of the incident spectrum varies with time and therefore with axial depth in the sample being imaged. This chirp can be removed for a single axial depth by tuning the dispersion of the reference arm to match the dispersion of the sample at that point, but there is no way to maintain a chirp-free detected signal over the entire axial range of the image since the material in the sample will induce chirp. The effect is that at isolated reflecting or backscattering surfaces in the sample being imaged, the calculated spectra also exhibit a chirp matching that of the incident pulse.

Figure 25 illustrates this effect when a mirror is placed in the sample arm. The chirp appears as a depth dependence in center wavelength, with green colours indicating a shorter center wavelength and red colours indicating a longer center wavelength. As the image was acquired, the dispersion of the reference arm was varied, changing the dispersion match between the reference and sample arms. Therefore at each transverse location, the detected signal has a slightly different chirp. The effect is that the center wavelength of the reflected spectra varies in the transverse direction, moving from short/long to long/short. In the middle of the image, the detected signal has no chirp. This effect causes spectroscopic noise by appearing as a center wavelength shift in the detected spectra at isolated reflecting or scattering surfaces, regardless of the optical properties of the tissue being imaged.

A third significant source of spectroscopic noise is related to the bulk optical scattering properties of biological tissue. The overall scattering efficiency of tissue is higher for shorter wavelengths; therefore, a general red-shifting trend is observed in the detected spectra as the axial depth is increased, since the shorter wavelengths do not penetrate as deeply into the tissue. This effect was visible in images presented in [14], and can degrade the ability to discern fine spectroscopic features of the tissue at deep axial distances.

There are other less significant sources of spectroscopic noise, such as the temporal instability of the light source. This may theoretically be a factor when using a solid state laser, but during the experiments performed for this thesis project no problems were encountered with
source spectral stability. It is important to note that the three main sources of noise described above all relate to absolute wavelength characteristics of the detected spectra; in other words, all of the major sources of noise manifest themselves as Doppler shifts. By designing spectroscopic metrics that do not rely on measurements relating to absolute wavelength, these noise sources can be removed from the system, as described below in sections 3.4.4 - 3.4.7.

### 3.4.2 Center Wavelength

Calculating the center wavelength of the detected spectra was the first spectroscopic metric reported for OCT [14]. In this metric, a center of mass calculation was performed on the local spectra at each pixel via

\[
f_C = \frac{\int f_{RF} X(f_{RF}) |df_{RF}|}{\int |X(f_{RF})| |df_{RF}|}
\]

(3.46)

The center optical wavelength \( \lambda_c \) was then calculated from (1.2). This metric is not computationally intensive, and has a straightforward interpretation. It has good performance when materials present in the tissue produce strong changes in the backscattered spectrum. For example, to detect areas of high blood concentration, center wavelength is a good metric to choose due to the characteristic absorption profiles of oxy- and deoxy-hemoglobin. These absorption profiles create significant red- and blue-shifts in the detected spectra, depending on whether the blood is arterial (predominantly oxy-hemoglobin) or veinous (predominantly deoxy-hemoglobin).

Figure 26 shows a comparison of a standard OCT image and a spectroscopic OCT image of a Syrian golden hamster cheek pouch, acquired in vivo, using center wavelength as the spectroscopic metric. In this image, green colours indicate shorter center wavelengths, while red colours indicate longer center wavelengths. The OCT system used a Ti:sapphire laser with a center wavelength of 800 nm and a bandwidth of 200 nm, resulting in a coherence length of 1.9 \( \mu \)m in air. The incident power on the sample was 2 mW. The galvanometer velocity was 3.6 mm/s, giving a detection frequency of 9 kHz and a bandwidth of 4 kHz. A calibration fringe was acquired using a reference interferometer and a second A/D channel as described in section 2.1. The OCT fringes were corrected post-acquisition using the method of section 2.1. For the STFT analysis, a CZT was used with a Hamming window of length 512 and analysis frequencies
\( \omega_0 = 0.0789\pi \), \( \omega_1 = 0.1735\pi \). After the hue values were calculated for each pixel in the image, a rolling average was performed on the hue in the transverse direction over 3 adjacent A-scans. No averaging was performed on the saturation or luminance values, which contain the structural information in the image. The colour map described in section 3.3.4 was used to visualize the data. Blood vessels are clearly visible in the spectroscopic image, including one deeply-buried vessel that is not visible in the intensity image. Spectroscopic noise from incident pulse chirp is visible at isolated reflecting surfaces. A systematic red-shift as the axial depth increases is also visible, as the image appears more green at the shallow layers and more red at the deeper layers. Galvanometer noise is not visible since the data set was corrected using the method described in section 2.1. This image demonstrates that while center wavelength is a useful spectroscopic metric for some situations, it is sensitive to all three sources of spectroscopic noise.

### 3.4.3 Power Over Predefined Spectral Bandwidth

Using power over a predefined spectral bandwidth as a metric creates contrast by calculating the spectral power within a user-defined bandwidth, and displays the result in the colour map described in section 3.3.4. The concept of this metric is analogous to histological staining, where different tissue types will preferentially absorb or reflect certain wavelengths compared to preferential absorption of dye in histology. This metric is useful when a particular substance with a known absorption or scattering spectrum is present in the tissue, and when that substance can be used to enhance the contrast of the image. For example, this metric would also be a good choice for locating blood in a tissue sample. It could also be used with exogenous contrast dyes that have characteristic absorption profiles. The metric calculates a narrow-band spectral power \( P_{BW} \) according to

\[
P_{BW} = \frac{\int |X(f_{RF})|^2 df_{RF}}{\int |X(f_{RF})|^2 df_{RF}}
\]

(3.47)

The narrow-band power calculated for each spectrum is normalized to the total power contained in that spectrum. This way, strongly reflecting areas of the tissue are not preferentially weighted in the final visualization. It should be noted that all calculations are performed in RF frequency space, not in optical wavelength space. Performing calculation in wavelength requires
Figure 27. Detected spectra from a glass coverslip and regions containing 20 µm, 5 µm, 800 nm, and 200 nm microspheres. 20 µm and 5 µm microspheres induce spectral modulation consistent with single-particle Mie scattering. 800 nm and 200 nm microspheres induce modulation consistent with multiple-particle scattering.

Figure 28. Autocorrelation of the optical spectra for a glass coverslip and regions containing 20 µm, 5 µm, 800 nm, and 200 nm microspheres. Characteristic differences are observed in the bandwidth, shape, and number of peaks in the autocorrelation functions. These properties are related to the degree of spectral modulation caused by the target.
a nonlinear resampling of the spectral data, which is time-consuming and unnecessary since the RF frequencies $f_1$ and $f_2$ bounding the bandwidth of interest can be calculated from (1.2). This metric was evaluated using the Syrian hamster cheek pouch data discussed above, and was found to provide similar contrast enhancement for detection of blood vessels to the center wavelength metric. However, the narrow-band power metric requires multiple images to be generated over several optical bandwidths to provide a complete data set. While this is useful, it is very time consuming and did not produce better results than the simpler, faster center wavelength method. Furthermore, this method is also sensitive to the three spectroscopic noise sources. In the future, if work is done involving exogenous contrast agents, this metric may be quite useful.

3.4.4 Metrics Sensitive to Spectral Modulation

Realizing that spectroscopic metrics that require an analysis of properties linked to absolute wavelength will always be affected by spectroscopic noise, it is possible to examine another class of metrics based on spectral modulation. As the discussion of Mie theory in section 3.2 describes, particles of varying size will impart varying degrees of modulation to the backscattered spectra detected by the OCT system. While other groups have shown that it is possible to extract quantitative data about the scattering particle size, these techniques require complex imaging systems such as angle-resolved interferometry [22] or polarization-sensitive interferometry [18]. Quantitative techniques based on backscattering low coherence interferometry require extensive transverse averaging of the data by using a large diameter beam on the sample [23]. Furthermore, extraction of particle size requires fitting the data to theoretical curves from Mie theory, and some a priori knowledge of the tissue's optical properties is usually required. All of these factors make it difficult to extract quantitative particle sizes from biological tissue using OCT systems, which typically have transverse beam diameters of < 20 μm, and are generally not capable of angle-resolved or polarization-sensitive measurements.

With that in mind, this new class of spectroscopic metrics was developed to provide "qualitative" information on the scattering particle size in biological tissue by quantifying some measure of the degree of modulation present in the spectra. To evaluate the suitability of using spectral modulation as a spectroscopic OCT metric, several solutions of polystyrene microspheres were imaged and their spectra calculated using the STFT/CZT method. The microspheres had a refractive index of 1.59 and were suspended in a solution with a refractive
index near 1.33. Solutions containing microspheres with diameters of 200 nm, 800 nm, 5 \( \mu \text{m} \), and 20 \( \mu \text{m} \) were imaged. The solutions had equal concentrations of microspheres by volume, so the number density of microspheres varied across the samples. Within one imaging volume (defined by the OCT beam diameter and coherence length, typically 5 \( \mu \text{m} \times 1.5 \mu \text{m} \) in tissue), the various concentrations contained 94, 1.5, \( 6 \times 10^{-3} \), and \( 9.4 \times 10^{-5} \) particles respectively. Accordingly, the 200 nm solution was expected to exhibit spectral modulation profiles consistent with multiple scattering events (section 3.2.5) while the 800 nm, 5 \( \mu \text{m} \), and 20 \( \mu \text{m} \) solutions were expected to exhibit profiles consistent with single scattering events. The imaging system was the same as that described in section 3.4.2, and the OCT data was corrected for galvanometer velocity post-acquisition as described in section 2.1.

Figure 27 shows representative spectra for a glass coverslip and the 20 \( \mu \text{m} \), 5 \( \mu \text{m} \), 800 nm, and 200 nm solutions. As expected, the 20 \( \mu \text{m} \) particles exhibit more modulation than the 5 \( \mu \text{m} \) particles, since each spectra in these cases are the results of single scattering events. Spectra for the 200 nm and 20 \( \mu \text{m} \) solutions appear similar to one another, and also appear more modulated than the single-scattering 5 \( \mu \text{m} \) microspheres. This is likely due to the fact that the 200 nm solution contained 94 particles per OCT imaging volume, while the 20 \( \mu \text{m} \) solution contained less than one particle per imaging volume. As shown in section 3.2.5, multiple scattering events tend to increase spectral modulation. It therefore appears that the scattering spectrum of \( \sim 100 \) 200 nm particles is approximately equivalent to the scattering spectrum of a single 20 5 \( \mu \text{m} \) particle, at least to the sensitivity limit of a spectroscopic OCT system. This points out an inherent problem with spectroscopic metrics based on spectral modulation: in some cases they may be unable to distinguish a single large scatterer from many smaller scatterers within a single imaging volume in certain situations. Nevertheless, the calculated spectra for the coverslip, 5 \( \mu \text{m} \) microspheres, and 200 nm microspheres all contain significantly different amounts of spectral modulation. This indicates that a metric based on spectral modulation may be able to distinguish particles of significantly different size. The question now is how best to quantify the degree of spectral modulation present at each pixel in the OCT image.
Performing a Gaussian fit is one way to quantify spectral modulation. In this metric, the individual spectra are first power normalized using

\[ X_{\text{norm}}(f_{RF}) = \frac{X(f_{RF})}{\int |X(f_{RF})|^2 df_{RF}} \]  

(3.48)

Then the center frequency is calculated using the method described in section 3.4.2 and a Gaussian function \( G(f_{RF}) \) is generated at the calculated center frequency \( f_c \), with a bandwidth equal to that of the OCT light source \( \Delta f_{RF} \).

\[ G(f_{RF}) = \exp\left[-\frac{(f_{RF} - f_c)^2}{\Delta f_{RF}^2}\right] \]  

(3.49)

This way, the metric is made insensitive to the Doppler shifts caused by spectroscopic noise sources. The residual error between the Gaussian function and the actual spectra are then calculated, and this error \( e \) is encoded into the pixel’s hue value.

\[ e = \int \left| G(f_{RF}) - X_{\text{norm}}(f_{RF}) \right|^2 df_{RF} \]  

(3.50)

Although this metric has intuitive appeal since it directly measures the degree of “non-Gaussianness”, or modulation, of the spectrum, during imaging experiments it did not prove capable of enhancing the contrast of different tissue types in biological specimens. It was capable of differentiating highly reflective surfaces (such as coverslips) from tissue, but this is of little practical use. The problem stems from the fact that spectra exhibiting any degree of modulation at all produce similar residual error values, largely regardless of the frequency or depth of modulation. Therefore areas of tissue containing scattering particles of significantly different size cannot be differentiated, only areas that have no modulation versus “some” modulation.

A more general problem with analyzing the raw spectra was uncovered during imaging of developing zebrafish embryos. These developmental biology specimens have clearly delineated boundaries between tissue regions, and so serve as a good model for evaluating the ability of spectroscopic metrics to enhance tissue contrast. When looking at the raw spectra from different locations within a homogenous tissue region, significant variance is observed in the shape of the spectra. The degree of modulation (ie, modulation depth and frequency) in homogeneous areas
was generally constant, but the exact shape of the spectra varied from pixel to pixel. This indicates that directly analyzing the modulation of the raw spectra is inadvisable, since the variance from point to point within a tissue type may exceed that across tissue types. It is therefore desirable to develop metrics that indirectly measure the degree of spectral modulation, in order to reduce the sensitivity to point-to-point variations in the spectra of homogeneous tissue types.

3.4.6 Autocorrelation Peak Count

The idea of indirectly measuring general trends in spectral modulation, as opposed to directly measuring the modulation of the raw spectra themselves, is somewhat akin to performing a statistical measurement. In other words, the calculated spectra can be thought of as outcomes of a random process, with each tissue type represented by a different random process. This model is intuitively appealing, since the spectrum observed at a given pixel will depend on the distribution of sizes, shapes, and densities of the various cellular organelles, as well as the cells themselves, in the active volume for that pixel. These random distributions will give rise to stochastic optical spectra, so a statistical analysis of the outcome (ie, the observed spectrum) at each pixel is therefore warranted. From this viewpoint, the autocorrelation function $r_{xx}(k)$ of the spectra can serve as a useful way to quantify spectral modulation. The autocorrelation function is given by

$$r_{xx}(k) = \frac{\int |X_{\text{norm}}(f_{RF})|X_{\text{norm}}(f_{RF}-k)\,df_{RF}}{\int |X_{\text{norm}}(f_{RF})|^2\,df_{RF}}$$  \hspace{1cm} (3.51)$$

Absolute values of the spectra are used to remove the phase information, which is not related to spectral modulation.

Gaussian spectra with little modulation should roll off slowly as the lag $k$ is increased, and should drop to zero relatively quickly. Spectra with high modulation should roll off rapidly as the lag is initially increased, and should then exhibit characteristic peaks when the local maxima of the modulations are aligned between $X_{\text{norm}}(f_{RF})$ and $X_{\text{norm}}(f_{RF}-k)$. Figure 28 shows the autocorrelation functions of the optical spectra obtained from a glass coverslip and from solutions containing 20 μm, 5 μm, 800 nm, and 200 nm polystyrene microspheres. The spectrum of the coverslip exhibits the lowest degree of modulation (see figure 27), while the
**Figure 29.** In vivo intensity image (A) and spectroscopic image using a polynomial fit to the spectral autocorrelation function as the metric (B). Specimen is a developing zebrafish embryo. Spectroscopic image provides contrast enhancement of the head and developing body, the nutrient sac, floating nutrient nodules, and the egg shell. Coverslip is visible at bottom of image. Red areas are highly modulated, while green areas have low modulation.

**Figure 30.** In vivo spectroscopic images using (A) center wavelength and (B) autocorrelation bandwidth as the metrics. Improved contrast between the embryo, membrane, and nutrients is observed in (B). Spectroscopic noise from galvanometer motion and incident pulse chirp is observed in (A) but not in (B).
spectrum of the 800 nm microspheres exhibit the most modulation. The 5 µm microspheres fall in between, while the 200 nm and 20 µm appear very similar. From this experiment, it is evident that the shape of the spectral autocorrelation functions change significantly depending on the degree of spectral modulation and therefore the scattering particle size. Figure 28 also illustrates that the autocorrelation functions of different sized particles can appear similar when the concentrations are not equal.

One metric designed to take advantage of this observation is the autocorrelation function peak count, which produces a count of the number of local maxima in the autocorrelation functions of the local spectra. Theoretically, spectra with higher frequency modulations should produce more local maxima in the autocorrelation functions, providing a rapid method for quantifying modulation. After testing this metric with the developing zebrafish data, however, it became clear that calculating the number of peaks was not enough to provide contrast between tissue types. This is because very small modulations in the spectrum can produce shallow peaks in \( r_{XX}(k) \), making this metric sensitive to insignificant modulations. The depth of each peak in \( r_{XX}(k) \) is also important, but to compute the local maxima and minima in \( r_{XX}(k) \), and their relative vertical separation, is computationally intensive and not easy to physically interpret.

3.4.7 Autocorrelation Polynomial Fit

To obtain a better representation of the regularity of \( r_{XX}(k) \), a polynomial curve fitting technique was investigated as a spectroscopic metric. The principle behind this metric is that as spectral modulation increases, the autocorrelation function will become more irregular in that it will have more peaks and a non-smooth envelope. To quantify this irregularity, a technique similar to the Gaussian fit of section 3.4.5 was employed, but a low-order polynomial was used in place of a Gaussian. This was done because the autocorrelation function is not expected to be well-described by a Gaussian, and a low-order polynomial can adapt to the coarse features of \( r_{XX}(k) \), leaving only the fine details as the error signal. For this metric a 5th order polynomial was generated for each autocorrelation function, providing a best fit in the least-squares error sense. The residual error \( e \) was calculated between the polynomial fit and \( r_{XX}(k) \) as in section 3.4.5, and mapped into the pixel hue.
Experimental results with this metric were quite good, as shown in figure 29. This is an example of an in vivo image of a developing zebrafish embryo, taken with the system described in section 3.4.2. In the spectroscopic image, areas in red indicate a high residual error, therefore a poor polynomial fit to $r_{xx}(k)$, an irregular $r_{xx}(k)$, and a high degree of spectral modulation. Areas in green indicate a low residual error, and therefore a low degree of spectral modulation. The developing embryo is wrapped around a nutrient-filled sac, which contains floating nutrient material and fluid. On the spectroscopic image, the embryo appears red while the nutrients appear green. The outer membrane of the embryo appears yellow. The head of the embryo, visible on the left, appears to be a mixture of yellow and red. It can be inferred that the cells in the embryo produce a high spectral modulation, while the nutrients do not. The fluid in the nutrient sac also appears to modulate the spectrum. Note that no spectroscopic noise is present; the highly reflecting glass coverslip at the bottom of the image does not appear chirped, and there is no red-shift with increasing depth. Galvanometer noise was removed during post-processing, but this metric is also insensitive to this noise source.

Although this metric produces good contrast enhancement and is also insensitive to spectroscopic noise, it suffers from two major drawbacks. First, it is extremely computationally intensive, since for each pixel a spectrum must be calculated, an autocorrelation function must be calculated, a polynomial fit must be generated, and an error calculation performed. Second, it does not have a straightforward physical interpretation since the size of the scattering particles is inferred from a polynomial fit.

3.4.8 Autocorrelation Bandwidth

To address the computational complexity issue with the polynomial fitting metric, a metric was developed based on the bandwidth of the autocorrelation function of the detected spectrum. As shown in figure 28, the bandwidth of $r_{xx}(k)$ varies significantly depending on the degree of spectral modulation present. The largest difference is seen in the region near $k = 0$, when modulated spectra will cause the autocorrelation function to fall off rapidly while non-modulated spectra will fall off slowly. Therefore, by evaluating the bandwidth of $r_{xx}(k)$, it is possible to obtain another measure of spectral modulation. The exact point at which the bandwidth is calculated is arbitrary, but should be near the peak of $r_{xx}(k)$. Experimentally, using the 90%
Figure 31. *In vivo* intensity image of a developing Xenopus Laevis (African frog) tadpole

Figure 32. *In vivo* spectroscopic image using center wavelength analysis. Galvanometer noise and red-shifting are present, and no significant contrast enhancement is obtained.

Figure 33. *In vivo* spectroscopic image using autocorrelation bandwidth analysis. No spectroscopic noise is present, and contrast between the different tissue types of the specimen is achieved.
**Figure 34.** Fourier transform of the optical spectra for a glass coverslip, 5 μm microspheres, 800 nm microspheres, and 200 nm microspheres. Data were averaged over 10 consecutive A-scans spanning 10 μm total. As spectral modulation increases, energy is distributed in higher k values.

**Figure 35.** *In vivo* intensity image (A) and spectroscopic image using the Fourier transform of the optical spectra as the metric (B). Specimen is a developing zebrafish embryo. Spectroscopic image provides contrast enhancement of the organs developing inside the head and the body. Coverslip is visible at bottom of image. Red areas are highly modulated, while blue areas have low modulation.
bandwidth was found to provide good spectroscopic contrast, although this represents an additional analysis parameter to be optimized. This calculation is much faster than the polynomial curve fitting technique, and produces images of equal quality. The autocorrelation bandwidth is also easier to physically interpret than the polynomial curve fit, which also makes it a more attractive metric.

Figure 30 illustrates the advantages of the autocorrelation bandwidth metric over a center wavelength analysis, both in terms of contrast enhancement and rejection of spectroscopic noise. The data was from the same developing zebrafish as in section 3.4.7, except that in this case the data was not corrected to remove galvanometer noise. The center wavelength analysis shows alternating red and green horizontal bars running through the image, which are the result of changes in the galvanometer velocity as the A-scans are acquired. The suspended nutrient concentrations in the nutrient sac, as well as the embryo membrane and coverslip, all show incident pulse chirp. The embryo, embryonic membrane, and nutrients are not well-differentiated. There is also a subtle red-shifting of the spectra as the axial depth increases. Conversely, the autocorrelation bandwidth analysis shows good contrast between the various sections of the specimen. Galvanometer noise and chirp noise are not present, and no red-shift is present. In this image, red areas indicate a low 90% autocorrelation bandwidth, and therefore areas of high spectral modulation. Blue areas indicate a high bandwidth and low modulation.

The autocorrelation bandwidth metric was also tested on other developmental biology specimens to further evaluate its performance. A developing Xenopus Laevis (African frog) tadpole was imaged using the spectroscopic OCT system described in section 3.4.2. The data was not corrected for galvanometer motion to compare the noise performance of the autocorrelation metric to the center wavelength metric. Figure 31 shows the intensity OCT image, figure 32 shows the center wavelength spectroscopic image, and figure 33 shows the autocorrelation bandwidth spectroscopic image. Note that in this specimen, as with the zebrafish embryo, the center wavelength analysis does not provide any significant contrast enhancement of the various tissue types within the sample. Galvanometer noise is again visible as alternating red and blue horizontal lines, and red-shifting is observed in deeper regions of the tadpole. In figure 33, no spectroscopic noise is visible. Also, good contrast enhancement is obtained between the various structures within the tadpole. The outer membrane and tail cartilage appear yellow, while the region inside the membrane appears blue. Structures inside of the developing head
appear yellow and red, while some smaller structures contain a blue (highly modulated) center. Overall, the autocorrelation function bandwidth appears to be a good metric that is insensitive to spectroscopic noise and capable of providing good contrast enhancement.

### 3.4.9 Fourier Transform of Spectra

The last spectroscopic metric investigated for this thesis project was an alternative method to quantify the degree of spectral modulation. Instead of using the autocorrelation function of the optical spectra, this metric examined the distribution of energy in the Fourier transform of the optical spectra. It is expected that highly modulated spectra will have more energy in high frequency components than spectra with low levels of modulation. Since the Fourier transforms of the absolute values of the optical spectra were analyzed, this operation is not equivalent to simply inverse-transforming the spectra to return to the time domain. If $X_{\text{norm}}(\lambda)$ is the power normalized STFT-calculated optical spectrum (as a function of optical wavelength) at a given pixel in the image, this metric examined the transform $X(k)$ given by

$$X(k) = \int X_{\text{norm}}(\lambda) e^{-j2\pi k \lambda} d\lambda$$

(3.52)

Figure 34 shows the Fourier transforms of the optical spectra for the glass coverslip, 5 μm microsphere, 800 nm microspheres, and 200 nm microspheres of Figure 27. The Fourier transforms were averaged over 10 consecutive A-scans, spanning a total of 10 μm. Data are not shown for the 20 μm solution, since motion of the larger particles due to gravity caused the particle locations to become uncorrelated over 10 A-scans. As shown in figure 34, there is a correlation between scattering particle size and the distribution of energy in the Fourier transform of the optical spectra. For the smaller particles (which exhibit more modulation), the energy in the transform is distributed in higher $k$ values. By evaluating the total power in the transform above some user-defined cutoff $k_c$, the degree of spectral modulation for each pixel can be represented by a single value and mapped into the pixel hue. This metric therefore evaluates the high-frequency power $P_c$ using

$$P_c = \int_{k-k_c}^{\infty} |X(k)| dk$$

(3.53)

When this metric was tested on the developing zebrafish and xenopus laevis tadpole data previously discussed, it produced images virtually identical to those produced by the
Figure 36. *In vivo* intensity image of Syrian golden hamster cheek pouch. Keratin layer (k), epithelium (e), and muscularis (m) are visible. Epithelial basement membrane (ebm) is somewhere at the bottom of the epithelium.

Figure 37. *In vivo* spectroscopic image of Syrian golden hamster cheek pouch using Fourier transform of the optical spectrum as the metric (A). (B) shows a 2x detail, and (C) shows a negative image of the same detail. Some subtle spectroscopic contrast is observed between the bottom of the epithelium, possibly the EBM, and the rest of the epithelium. Red areas are low spectral modulation, while green areas are high spectral modulation.
autocorrelation bandwidth metric. Figure 35 shows a comparison between a standard intensity-based OCT image and a spectroscopic image using the Fourier transform of the optical spectra as the metric. The specimen is a developing zebrafish embryo. For this analysis, the cutoff k-number was chosen to be $k_c = 0.002$. The value of $k_c$ is arbitrary, but should be chosen such that different tissue types exhibit significantly different distributions of energy above this level. The cutoff can be chosen by examining the Fourier transforms of the optical spectra at a few representative points in the image where contrast is desired, and picking a value of $k_c$ that provides good contrast between those points. Red areas in the image indicate a high amount of power above $k_c$ (high spectral modulation), while blue areas indicate a low amount of power above $k_c$ (low spectral modulation). This specimen is at a further developmental stage than the images shown in figures 29 and 30. The spectroscopic image shows contrast enhancement of the developing organs inside the head of the organism, including the brain. The body of the organism still appears highly modulated, while the membrane and nutrients appear to have low modulation.

While there is not a significant qualitative difference in the performance of this metric compared to the autocorrelation bandwidth metric, it is slightly less computationally intensive. It is not, however, any more of a direct measurement of the spectra than the autocorrelation metrics. It is merely another option for analysis, which may prove to be useful for future studies.

### 3.5 Future Work with Spectroscopic OCT: Scattering Tissue

The majority of the work on spectroscopic OCT performed for this thesis project focused on the fundamentals of developing new metrics capable of providing enhanced tissue contrast, based on the scattering particle sizes of biological tissue. Most of the testing was carried out on developmental biology specimens such as zebrafish and tadpoles, since these organisms produce relatively large amplitude OCT signals. The ultimate application of spectroscopic OCT, however, is to enhance the ability to detect diseases in humans. For this to be possible, further refinement of the techniques presented here must be made to enable spectroscopic differentiation of tissue in highly scattering samples. Some preliminary experiments were performed using in vivo images of Syrian golden hamster cheek pouch tissue, human skin, and human oral mucosa.
Figures 36 and 37 show an example of an intensity-based image and spectroscopic image of a hamster cheek pouch. The spectroscopic analysis was done using the Fourier transform of the optical spectra, with a cutoff of $k_c = 1 \times 10^8$. The window used for this experiment was a Hamming window of length 1024, since the increased length seemed to provide better results than the length 512 windows used in the rest of this project. The goal of this experiment was to try to obtain spectroscopic contrast enhancement of the epithelial basement membrane (EBM), which lies at the base of the epithelial layer. Many human cancers form in the EBM, so any way to detect disruption of this membrane would be very valuable for early diagnosis. Since EBM cells have a different size and different concentrations of organelles (higher nuclei-to-cytoplasm ratio) than other epithelial cells, their presence may be detectable using spectroscopic metrics that are sensitive to optical modulation.

The results of this experiment were not as good as those obtained in the developmental biology specimens, but do suggest that with further work it may be possible to detect early stages of cancer in scattering tissue. One difficulty in detecting the EBM lies in the fact that it has a high density of cell nuclei, which are known to be low backscattering in OCT systems. Therefore the signal returning from the EBM layer is very low, making it difficult to analyze. By taking a “negative” image, as shown in figure 37c, the brightness of the EBM region can be increased. In this image, there is a subtle spectroscopic differentiation between the bottom of the epithelium (which appears green), the middle of the epithelium (which contains red regions), and the top of the epithelium (which appears green again). This top layer may in fact be the keratinized layer, and not epithelial tissue. This bottom green layer may correspond to the EBM, but more work is necessary to correlate these results with histology stains.

More work is also necessary to fine-tune the spectroscopic metrics for the detection of EBM cells in general. Since scattering tissue does not seem to provide as much spectroscopic contrast as developmental biology specimens, it will be necessary to modify the analysis techniques presented in this thesis project to account for this. Simply tuning the parameters of the analysis (ie, the window size, colour map, cutoff $k_c$, etc) may also help in highlighting these cellular features. Further studies must be carried out on normal animals and on animal models of cancer to determine at what stage spectroscopic OCT is capable of detecting the presence of tumours. A more complete understanding of the source of the OCT signal and the spectral modulation imparted by biological tissue can be obtained by repeating the microsphere imaging experiment.
The concentrations of the solutions should be controlled to remove uncertainties due to multiple scattering particles. Finally, the spectroscopic OCT software must be rewritten in C++ to increase its efficiency and enable its widespread use in imaging studies. Overall, the results obtained to date are encouraging, and suggest that spectroscopic OCT may be useful in studies of developmental biology models and possibly for the early detection of cancer.
4.1 Motivation for OCT Image Enhancement

The final portion of this thesis project focused on developing DSP techniques to improve the quality of OCT images. In any medical imaging technique, it is important to provide the clinician with an image that is as free from noise as possible to maximize the system’s utility. This is especially true in OCT, due to the inherently low level of the detected signals compared to the noise background. The ability to reduce noise in an image is also important for the accuracy of quantitative data analysis and image segmentation. With the development of high-speed spectral domain OCT, it is becoming possible to rapidly generate full three dimensional reconstructions of biological tissue. Segmentation of these 3D reconstructions will provide a valuable tool for the diagnosis of disease, especially in ophthalmic and vascular imaging, and a significant reduction in image noise will greatly improve the accuracy of these segmentation techniques. Furthermore, the development of new broadband SLD light sources for OCT has the potential to enable widespread adoption of ultrahigh resolution OCT imaging. These light sources, however, have much lower output power than femtosecond lasers. This decrease in output power leads to a degradation in signal-to-noise ratio (SNR), which must be compensated for by image enhancement techniques.

Since OCT is a coherent imaging technique like ultrasound and synthetic aperture radar (SAR), the main source of noise is coherent speckle [28]. Speckle noise is present in an OCT system when the sample contains multiple scattering particles within the active volume of the imaging beam, as defined in section 3.2.5. The photocurrent output $i_d$ of an OCT system detector is proportional to the temporal and spatial average of the cross-correlation between the reference field $E_r$ and the backscattered sample field $E_s$.

$$i_d = K \text{Re} \langle E_r E_s^* \rangle$$ (4.1)
The detected signal can also be expressed in terms of the optical path mismatch $\Delta z$ between the reference and sample arms, the source center frequency $v_0$, and the amplitude $g(\Delta z)$ and phase of the cross-correlation $\phi(\Delta z)$ in (4.1).

$$i_d(\Delta z) = Kg(\Delta z) \cos[2\pi v_0 \Delta z + \phi(\Delta z)]$$

(4.2)

Due to the coherent nature of the OCT detection system, the backscattered waves from all of the particles within the active volume will add coherently. When the density of the scatterers becomes large, the terms $g(\Delta z)$ and $\phi(\Delta z)$ become stochastic and speckle noise appears. Speckle appears as localized regions of constructive and destructive interference in the final image, giving the image a granular appearance and making it difficult to detect the boundaries between tissue layers.

A significant amount of work has been carried out on speckle reduction for OCT imaging. Several methods based on spatial or angular compounding have been described [28-31], making use of incoherent averaging over multiple “looks” at the same location in the tissue sample. Since the speckle pattern is uncorrelated at slightly different observation positions or angles, these techniques are capable of reducing the speckle in proportion to the number of data sets acquired. The main drawback to these techniques is that multiple data sets must be acquired for every desired image, and more complicated OCT systems are required for the angle-sensitive measurements. The speckle pattern in OCT images is also uncorrelated across optical frequencies, such that incoherent averaging over multiple data sets acquired using sources with different center wavelengths will also reduce speckle noise [28, 32]. The drawback here is that multiple broadband sources are required, which dramatically increases the system cost and complexity. Time-domain filtering techniques based on the rotating kernel transform (RKT) have been investigated for OCT speckle reduction [33, 34], which produce good contrast enhancement of gross image features, but which also produce a significant blurring of edge boundaries.

Wavelet-based denoising techniques are another class of filtering methods capable of reducing coherent speckle noise. They have found widespread application in ultrasound and SAR imaging in recent years due to their excellent noise reduction properties and capability to maintain image sharpness. Application to OCT speckle reduction has been relatively limited [35, 36], with most of the work focusing on fairly simple coefficient thresholding schemes that
do not take full advantage of the properties of wavelet decompositions. The purpose of this section of the thesis project was to investigate spatially adaptive wavelet speckle reduction algorithms from the SAR and ultrasound fields, apply them to OCT images, and quantitatively evaluate their performance. A comparison is also made between the spatially adaptive wavelet techniques, more basic wavelet techniques, and time-domain filtering methods commonly used in OCT imaging.

4.2 Wavelet Theory

To understand the wavelet denoising techniques presented in section 4.3, it is first necessary to provide a brief overview of basic wavelet theory. Although wavelets are a relatively new mathematical field, there is already a large and diverse body of literature covering applications in signal processing, compression and coding, and numerical computation methods. The theory outlined here will focus on the key ideas for wavelet denoising of two-dimensional images. The starting point for this discussion is the one-dimensional wavelet transform. The discussion presented here is taken from texts by Strang and Nguyen [37], Mallat [38], and Vetterli and Kovacevic [39].

4.2.1 Projections onto Subspaces

The wavelet transform is a type of time-frequency analysis, often referred to as a multiresolution analysis or wavelet decomposition. The basic idea is that a signal \( f(t) \) of finite energy, which is part of the space of real square summable functions \( L^2(\mathbb{R}) \), can be expressed as a projection onto a set of basis functions \( \phi_{j,k}(t) \) and \( \psi_{j,k}(t) \). The functions \( \phi_{j,k}(t) \) are called scaling functions, and the functions \( \psi_{j,k}(t) \) are the wavelet functions. The basis functions are defined from the shifts and dyadic dilations of a mother scaling function \( \phi(t) \) and a mother wavelet function \( \psi(t) \), using

\[
\phi_{j,k}(t) = \sqrt{2^j} \phi\left(2^jt - k\right) \tag{4.3}
\]

\[
\psi_{j,k}(t) = \sqrt{2^j} \psi\left(2^jt - k\right) \tag{4.4}
\]
The "level" integer $j$ in the basis functions provides the "frequency" aspect of the time-frequency analysis, while the "shift" integer $k$ provides the "time" aspect. As the level $j$ increases, the scaling and wavelet functions become increasingly compressed in time, corresponding to finer (higher frequency) components of the projection.

The set of all integer shifts of the scaling function at a given level $j$, represented by $\{\phi_{j,k}(t)\}_k$, forms a subspace $V_j$ of $L^2(\mathbb{R})$ such that $V_j \subseteq L^2(\mathbb{R})$. Furthermore, $V_j$ is a subspace of the next finer level $j+1$, such that $V_j \subseteq V_{j+1}$. Therefore by considering all the levels $j \in \mathbb{Z}$, a set of nested subspaces can be built up, giving

$$\{0\} \subset \ldots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \ldots \subset L^2(\mathbb{R}) \quad (4.5)$$

In other words, a projection of $f(t)$ onto $V_j$ gives an approximation of $f(t)$ at the level $j$. This approximation can be represented by a series of approximation coefficients, given by

$$\{a_{j,k} = \langle f, \phi_{j,k} \rangle \}_{k \in \mathbb{Z}} \quad (4.6)$$

In the same way, the set of wavelet functions at a given level $j$ represented by $\{\psi_{j,k}(t)\}_k$ forms a subspace $W_j$ of $L^2(\mathbb{R})$, such that $W_j \subseteq L^2(\mathbb{R})$. It is also true that $W_j \subseteq W_{j+1}$ and

$$\{0\} \subset \ldots \subset W_{j-1} \subset W_j \subset W_{j+1} \subset \ldots \subset L^2(\mathbb{R}) \quad (4.7)$$

Projecting $f(t)$ onto $W_j$ produces the so-called detail coefficients (the name is explained below), given by

$$\{w_{j,k} = \langle f, \psi_{j,k} \rangle \}_{k \in \mathbb{Z}} \quad (4.8)$$

It is now important to consider the relationship between the wavelet function subspaces $W_j$ and the scaling function subspaces $V_j$. In order for the multiresolution analysis to be useful, the information contained in the subspaces formed by the wavelets and scaling functions should not overlap. In other words, the subspaces $V_j$ and $W_j$ should be complementary such that

$$V_j \cap W_j = \{0\} \quad (4.9)$$
$$V_j + W_j = V_{j+1} \quad (4.10)$$
$$\therefore V_j \oplus W_j = V_{j+1} \quad (4.11)$$
Looking at (4.10), it can be stated that by adding the coarse subspace $W_j$ to the coarse subspace $V_j$, a less-coarse subspace $V_{j+1}$ is formed. If $V_j$ represents an approximation of $f(t)$ at level $j$, then $W_j$ can be thought of as the details of $f(t)$ at level $j$ required to form a better approximation $V_{j+1}$. Therefore, the scaling function is associated with rough approximations of $f(t)$, and the wavelet function is associated with fine details of $f(t)$.

Using (4.11) and (4.7), it is evident that any function in $L^2(\mathbb{R})$ can be represented by a projection onto $V_0$ (the coarsest subspace of scaling functions) and an infinite series of the wavelet subspaces $W_j$. This gives

$$V_0 \oplus \bigoplus_{j=0}^{\infty} W_j = L^2(\mathbb{R})$$

(4.12)

So, the wavelet transform is a technique for projecting a function $f(t)$ onto a series of complementary subspaces $V_0$ and $W_j$, which distributes the frequency content of $f(t)$ throughout the various subspaces. One can imagine that after the frequency content of $f(t)$ is separated into a set of different signals (the approximation and detail coefficients), filtering operations can be performed and an inverse transform can be carried out to realize noise reduction. The main advantage of wavelet filtering compared to standard frequency-domain filtering (via the Fourier transform) lies in the fact that the wavelet transform separates a single input signal $f(t)$ into a group of transformed signals. This way, the transform space (often called “scale space”) contains both time and frequency information. Filtering can therefore be tailored to adapt to local frequency characteristics of the input signal, instead of operating globally on the entire signal via the Fourier transform.

### 4.2.2 Filter Bank Implementation

In reality, it is not possible to construct an infinite set of detail coefficients spanning the range $j = 0 \to \infty$ in order to perform a true wavelet transform. It is also much more common to deal with finite length, discrete-time signals $f[n]$ instead of infinite length, continuous time signals $f(t)$. The discrete wavelet transform (DWT) therefore exists to deal with the finite
Figure 38. Discrete-time analysis filter bank used to implement a 3-level discrete wavelet transform. The digital lowpass and highpass filters $H_0(z)$ and $H_1(z)$ are derived from the continuous-time scaling function and wavelet function.

Figure 39. Discrete-time synthesis filter bank used to implement a 3-level inverse discrete wavelet transform. The digital lowpass and highpass filters $F_0(z)$ and $F_1(z)$ are derived from the filters in the analysis bank.

Figure 40. (A) Block diagram of a filter bank used for a one level, two-dimensional wavelet decomposition. The lowpass and highpass filters are applied separately across the rows and columns of the input image, creating four subband images at each level of the decomposition. Iteration of the filtering is performed on the LL subband only. (B) Schematic showing a sample input image and a representation of its one level wavelet decomposition. LL subbands are lowpassed versions of the original image, HL subbands contain vertical edges, LH subbands contain horizontal edges, and HH subbands contain diagonal features.
length discrete time situation. The DWT performs a wavelet transform up to an arbitrary level by using a filter bank, which consists of a cascade of digital lowpass and highpass filters $H_0(z)$ and $H_1(z)$, and downsamplers ($\downarrow 2$). The connection between the continuous time scaling functions and the discrete time filter bank coefficients is given by the dilation equation, which allows the construction of a coarser function from the scaled shifts of a finer function.

$$\phi(t) = \sqrt{2} \sum_n h_0[n] \phi(2t - n)$$

(4.13)

A similar relation holds for the wavelet functions.

$$\psi(t) = \sqrt{2} \sum_n h_1[n] \phi(2t - n)$$

(4.14)

The filter coefficients are calculated from the inner products of the wavelet and scaling functions, using

$$h_0[n] = \langle \phi, \phi_n \rangle$$

(4.15)

$$h_1[n] = \langle \psi, \phi_n \rangle$$

(4.16)

Figure 38 shows a schematic of a one-dimensional filter bank used to perform a three level DWT. At each level $j$ of the filter bank, a series of detail coefficients $w_j[n]$ and approximation coefficients $a_j[n]$ are generated. The approximation coefficients form the input to the subsequent stage of the analysis. This “analysis” bank can be inverted to form a “synthesis” bank, used to perform the inverse discrete wavelet transform (IDWT). The analysis bank is shown in figure 39, and consists of lowpass and highpass filters $F_0(z)$ and $F_1(z)$ and upsamplers ($\uparrow 2$). While the filter coefficients in the analysis bank are calculated from the continuous time wavelet and scaling functions, the filters in the synthesis bank are calculated from the filters in the analysis bank. If the detail coefficients $w_j[n]$ and the approximation coefficients $a_j[n]$ (where $J$ is the highest level in the decomposition) are not modified before passing into the synthesis bank, then the output $y[n]$ should be equal to $x[n]$ with the possible exception of a delay. This “perfect reconstruction” condition imposes constraints on the synthesis bank filters, such that they can be determined directly from the analysis bank filters.

97
In summary, a continuous-time wavelet and its associated scaling function can be used to generate associated discrete-time digital filters, which can then be arranged in a filter bank to carry out a discrete wavelet transform. The output of the $J$ level DWT is a set of detail coefficient vectors $w_j[n]$ for $j = 0 \rightarrow J$, and a single approximation coefficient vector $a_J[n]$. At each level of the decomposition, the length of the coefficient vectors is reduced by a factor of 2 by the downsamplers. The approximation vectors resemble increasingly low-passed versions of the original signal, while the detail vectors contain increasingly coarse features of the original signal. The separation of fine features from coarse features make wavelet analysis very well-suited to noise reduction. An inverse operation can be carried out using a synthesis filter bank, converting the coefficient vectors from the scale domain back to the time domain.

4.2.3 **Extension to Two Dimensions**

The filter bank method for implementing the DWT and IDWT is easily extended to two dimensions. Two-dimensional DWT's and IDWT's are often used for image processing, and are computed by separately applying one-dimensional transforms along the rows and columns of the dataset. In the two-dimensional case, each level of the wavelet decomposition produces four outputs instead of one. These subband images at level $j$ are labeled $LL_j$, $LH_j$, $HL_j$, and $HH_j$. The $LL_j$ subband image corresponds to the approximation coefficient vector in the one-dimensional case, and has been lowpass filtered by $H_0(z)$ along its rows and columns. The $LH_j$ subband image has been lowpass filtered by $H_0(z)$ along its rows and highpass filtered by $H_1(z)$ along its columns, and contains horizontal edge details. The $HL_j$ subband image has been highpass filtered along its rows and lowpass filtered along its columns, and contains vertical edge details. Finally, the $HH_j$ subband image has been highpass filtered along its rows and columns, and contains diagonal details. Figure 40 shows a filter bank used for a one level, two-dimensional wavelet decomposition. This filter bank can also be inverted to perform a reconstruction in two dimensions.
Figure 41. Detail and approximation coefficients from a 4 level wavelet decomposition of a signal \( x[n] \) corrupted with additive white noise. Coefficients resulting from steps in the signal tend to cluster across the levels, while coefficients due to noise do not cluster.

Figure 42. Undecimated filter bank used to perform a wavelet decomposition. The prototype lowpass and highpass filters \( H_0(z) \) and \( H_1(z) \) are modified via the Noble identity at each level in the filter bank.

Figure 43. Ultrahigh resolution OCT system with broadband SLD light source. Source produces > 150 nm of optical bandwidth at a center wavelength of 890 nm. Image resolutions of ~2 \( \mu \)m in scattering tissue and ~3 \( \mu \)m in the retina are possible.

Figure 44. (A) UHR-OCT image acquired with a broadband SLD light source. Image is corrupted with speckle noise. (B) Same image after pixel thresholding to mask the effects of noise. (C) UHR-OCT image acquired with a solid state laser. SNR of image (C) is 97 dB, while the SNR of image (B) is 94 dB. Wavelet denoising is required to improve the SNR of OCT images acquired using SLD light sources.
4.3 Wavelet Denoising Techniques

To perform a filtering operation using a wavelet decomposition, the coefficients \( w_j[n] \) and \( a_j[n] \) are operated on in the scale domain. Figure 41 shows a level 4 wavelet decomposition of a sample input signal \( x[n] \), composed of small steps and corrupted with additive white noise. This input signal is the noisy “bumps” vector from the Matlab wavelet toolbox. By looking at the scale domain signals produced by the DWT, it is apparent that by operating on the detail and approximation coefficients a denoised version of the input signal can be generated. A very crude technique would be to discard all of the detail coefficients, leaving only \( a_4[n] \) for the inverse transform. This, however, would be equivalent to simply lowpass filtering \( x[n] \). A more advanced denoising technique would selectively attenuate the various detail coefficients according to some sort of statistical model, whereby the coefficients relating to noise would be strongly attenuated while the coefficients relating to real signal would be left unchanged. This way, the sharpness of the signal is not degraded by the removal of noise.

One method to determine which coefficients are “noise” and which are “signal” is through the calculation of a coefficient threshold \( T \). Coefficients with absolute values greater than \( T \) are determined to be “signal”, while coefficients with absolute value less than \( T \) are considered to be “noise.” Two main classes of thresholding exist: hard thresholding, where coefficients below the cutoff are set to zero while coefficients above the cutoff are not modified; and soft thresholding, where coefficients below the cutoff are set to zero while coefficients above the cutoff have their magnitude reduced by the value of the threshold. If the original wavelet coefficients are given by \( w_j^{(o)}[n_x,n_y] \), where \( j \) is the level of the decomposition, \( o \) is the orientation (subband) of the coefficients, and \( n_x \) and \( n_y \) are the spatial coordinates, then hard thresholding is described by

\[
\hat{w}_{j,\text{hard}}^{(o)} = \eta_{j,\text{hard}}\left(w_{j}^{(o)}[n_x,n_y]\right)
\]

(4.17)

with the thresholding function given by

\[
\eta_{j,\text{hard}}(\rho) = \begin{cases} 0, & |\rho| - T \leq 0 \\ \rho, & |\rho| - T > 0 \end{cases}
\]

(4.18)

Soft thresholding is described by
\[ w_{j,\text{soft}}^{(o)} = \eta_{T,\text{soft}} \left( w_j^{(o)} \left[ n_x, n_y \right] \right) \]  
\[ \eta_{T,\text{soft}}(\rho) = \text{sgn}(\rho) \times \max(|\rho| - T, 0) \]

In most cases, it is usually true that soft thresholding produces better results than hard thresholding, due to the sharp discontinuities in the coefficient values introduced by hard thresholding \([40, 41]\).

Generally speaking, calculation of the threshold is carried out by statistically analyzing the properties of the wavelet coefficients such that the likelihood of incorrectly characterizing a particular coefficient as "noise" or "signal" is minimized. The problem of wavelet denoising is in fact largely a problem of calculating an accurate threshold, which removes noise without blurring the image features. Many methods exist for threshold calculation, ranging from very simple to very complex. One global threshold can be applied to every coefficient in every subband, separate thresholds can be computed for each subband, or even for every individual coefficient. The threshold calculation can take into account statistical properties such as the mean and variance of the wavelet coefficients, or can also account for spatial activity in an attempt to separate edges from noise.

It should be noted that almost all wavelet denoising techniques are developed for additive noise sources, while speckle noise in OCT is known to be multiplicative in nature \([28]\). However, by performing the denoising on a logarithmically transformed version of the dataset, the speckle noise becomes additive and the denoising techniques are effective. For this thesis project, the performance of four different thresholding techniques were quantitatively evaluated: global soft thresholding, SURE soft thresholding, and two different types of spatially adaptive thresholding. The wavelet techniques were also compared to time-domain median filtering, and the simple pixel thresholding which is usually applied to OCT images to mask the effect of speckle.

### 4.3.1 Image Quality Metrics

To effectively compare multiple wavelet and time-domain denoising techniques, it is necessary to use quantitative image quality metrics instead of subjective user assessments of how well a particular technique performs. Since no single metric can provide an accurate of
assessment of overall image quality, a set of metrics is required to measure properties such as signal-to-noise ratio (SNR), contrast, sharpness of edges, and smoothness of homogeneous areas.

The first image quality metric, SNR, is commonly used to describe the performance of OCT systems. SNR measures the ratio of the global maximum signal power in an image to the noise variance in a region known to contain only noise. If \( X[n_x, n_y] \) is the two-dimensional matrix of pixel values in the final OCT image, \( X_b[n_x, n_y] \) is a background area of the image containing only noise, and \( \sigma_b \) is the pixel standard deviation in the background, then global SNR is calculated on a log scale as

\[
SNR_{\text{global}} = 10 \log \left( \frac{\max \left( X[n_x, n_y] \right)^2}{\sigma_b^2} \right) \tag{4.21}
\]

It should be noted that this definition of SNR is taken from the OCT field, and is not the same as the SNR typically used in image processing. The image processing definition of SNR is not global, but is defined over a region of interest (ROI) \( X_k[n_x, n_y] \). Often, multiple ROI’s are analyzed for a given image to fully quantify the image quality. This “local” SNR is defined as the ratio between the mean of the pixel magnitudes in an ROI to the standard deviation of the pixel magnitudes in that same ROI. On a log scale, this is given by

\[
SNR_{\text{local}} = 10 \log \left( \frac{\mu_k}{\sigma_k} \right) \tag{4.22}
\]

Local SNR provides a good measurement of the signal level relative to the noise level, but cannot measure contrast. For this reason, the contrast-to-noise ratio (CNR) is used to measure the difference between the appearance of an ROI and the appearance of a background region \( X_b[n_x, n_y] \). CNR is defined on a log scale as

\[
CNR = 10 \log \left( \frac{\mu_k - \mu_b}{\sqrt{\sigma_k^2 + \sigma_b^2}} \right) \tag{4.23}
\]

Together, CNR and SNR provide a good way to quantify how noisy an image appears. However, they say nothing about the sharpness of edges, which is another important aspect of image quality.
To quantify edge sharpness, one commonly used metric is the beta parameter $\beta$ [42-45]. This metric compares the original image $X$ to the denoised image $\hat{X}$, and calculates how much the edge sharpness has degraded due to the denoising process. $\beta$ is a global measurement, and is defined as

$$\beta = \frac{\Gamma(\Delta X - \tilde{A}X, \Delta \hat{X} - \tilde{A} \hat{X})}{\sqrt{\Gamma(\Delta X - \tilde{A}X, \Delta X - \tilde{A}X) \Gamma(\Delta \hat{X} - \tilde{A} \hat{X}, \Delta \hat{X} - \tilde{A} \hat{X})}}$$ \hspace{1cm} (4.24)

Here, $\Delta X$ and $\Delta \hat{X}$ are highpass filtered versions of the original and denoised images, obtained using a 3x3 pixel approximation of the Laplacian operator [45]. The $\Gamma$ operator is a pointwise sum-of-products operation, given by

$$\Gamma(a, b) = \sum_{n_x, n_y} a(n_x, n_y) b(n_x, n_y)$$ \hspace{1cm} (4.25)

A $\beta$ value of close to unity indicates that the edge sharpness has not significantly changed from the original image to the denoised image.

The final image quality metric used to evaluate the performance of the denoising techniques measures the smoothness of homogenous areas, which appear rough due to the effects of speckle. In comparison to edge regions, in which a sharp appearance is desired, homogenous areas of the image should appear smooth and noise-free. The “equivalent number of looks” (ENL) metric is typically used in SAR imaging to evaluate the performance of speckle reduction filters [46-48]. In SAR, multiple images (“looks”) are often acquired of the same target and the pixel values incoherently averaged to reduce speckle. This is equivalent to the angular compounding speckle reduction methods used in OCT [28-31]. ENL calculates the number of incoherent averages that would have to be performed in order to obtain the same smoothness observed in a denoised image. ENL is defined over a particular ROI, and is given by

$$ENL = \frac{\mu^2}{\sigma^2}$$ \hspace{1cm} (4.26)

Although ENL is calculated in the same way as local SNR, its application is different. To calculate ENL, the ROI being analyzed must be a homogenous area free from edges. No such restriction applies to local SNR, and in fact the ROI’s used to calculate SNR often straddle edges or boundaries in the image.
4.3.2 Global Soft Thresholding

With a robust set of image quality metrics defined, it is now possible to evaluate different wavelet techniques for noise reduction in two-dimensional images. As discussed above, the problem of wavelet noise reduction is mostly a problem of calculating an appropriate threshold $T$ for attenuating noisy wavelet coefficients. A simple, yet effective, calculation of a global threshold based on the noise variance in the original signal was presented by Donoho [49]. This method produces near-optimal results in the minimax error sense when the signal is corrupted by independent identically distributed (iid) Gaussian noise of zero mean and standard deviation $\sigma_n$. The global soft threshold $T_{g}$ is calculated only once and applied to every coefficient in the wavelet decomposition, and is given by

$$T_g = \sigma_n \sqrt{2 \log(N)}$$

(4.27)

Here, $N$ is the number of samples in the original signal.

As with most thresholding methods, it is necessary to estimate the standard deviation of the noise present in the input signal. In many cases, $\sigma_n$ is not known exactly and must be estimated from the noisy data in the input signal. For the case of image processing, the robust median estimator is usually used to estimate $\sigma_n$ [40, 50, 51]. By looking at the $HH_0$ subband, which contains the highest frequency components in the image, the noise standard deviation can be estimated using

$$\hat{\sigma}_n = \frac{\text{median}(w_0^{HH}[n_x, n_y])}{0.6745}$$

(4.28)

where $w_0^{HH}[n_x, n_y] \in HH_0$. If there is no a priori knowledge of which regions of the image are background noise and which are signal, then the entire $HH_0$ subband can be used to calculate $\hat{\sigma}_n$. In the case of OCT signals, a background area can usually be defined where only noise is present, providing a more accurate estimate of the noise standard deviation.

Another method for estimating the noise standard deviation is presented in [52], where the deviation is calculated for each individual subband instead of globally as in (4.28). In this method, the effects of the lowpass filter $h_0[n]$ and highpass filter $h_1[n]$ are taken into account, using
Here, $\hat{\sigma}$ is the estimate given by (4.28).

### 4.3.3 SURE Soft Thresholding

Although the global soft threshold described in 4.3.2 is extremely easy to calculate and provides good noise reduction performance, it tends to over-smooth images and does not adapt to the distribution of wavelet coefficients in the different subbands of the decomposition. Donoho and Johnstone [53] developed a subband-adaptive thresholding technique based on minimizing Stein’s unbiased risk estimate. The threshold $T_{SURE}$ applied to the coefficients $w_{j}^{(o)}$ in a given subband $(o)$ is chosen to minimize Stein’s unbiased risk estimate $SURE(\lambda, w_{j}^{(o)})$.

$$T_{SURE} = \arg \min \ SURE(\lambda, w_{j}^{(o)}), \ \lambda \geq 0$$

$$SURE(\lambda, w_{j}^{(o)}) = \frac{N_{j}^{(o)}}{} + \sum_{n_{x}, n_{y}} \left[ \min \left( \left\| w_{j}^{(o)} \right\|_{n_{x}, n_{y}}, \lambda \right) \right]^{2} - 2 \left[ \left( \text{# of } w_{j}^{(o)} : \left\| w_{j}^{(o)} \right\| \leq \lambda \right) \right]$$

Here, $N_{j}^{(o)}$ is the number of wavelet coefficients in subband $(o)$ at level $j$.

The more commonly-used SUREShrink algorithm is a hybrid method, whereby $T_{SURE}$ is applied only if the condition

$$\frac{1}{N_{j}^{(o)}} \sum_{n_{x}, n_{y}} \left[ \frac{\left\| w_{j}^{(o)} \right\|_{n_{x}, n_{y}}^{2}}{\sigma_{n_{x}, n_{y}}} \right]^{2} \geq \left( \frac{\log_{2} N_{j}^{(o)}}{\sqrt{N_{j}^{(o)}}} \right)^{3/2}$$

is met. Otherwise, the algorithm applies the global soft threshold discussed in 4.3.2. The SUREShrink algorithm generally provides better performance than the global soft threshold [41], since it adapts the threshold to the statistical distribution of wavelet coefficients in the individual subbands.
4.3.4 Spatially Adaptive Thresholding and Undecimated Wavelet Transforms

Although the SUREShrink algorithm is subband adaptive, it is possible to improve denoising performance by using thresholding techniques that are also spatially adaptive. This class of algorithms modifies the threshold on a per-coefficient basis by determining if each coefficient is part of a real feature in the image or noise. This improves performance by retaining low-level features corrupted by noise, where non spatially-adaptive techniques may classify the smaller coefficients incorrectly based on their amplitudes alone. The basic method used to classify coefficients as “features” or “not features” is to generate a measure of local spatial activity, by examining the neighbouring coefficients in the same subband as well as the neighbouring coefficients at the same spatial location in the next level of the decomposition. Since features such as edges tend to generate wavelet coefficients that cluster across space and across decomposition levels, the correlation of a coefficient to its neighbours is a good measure of spatial activity. Coefficients from noise, however, do not cluster across space or scales. Therefore, even when the amplitude of a signal is at the same level as the image noise, it can be distinguished from noise by examining its neighbouring wavelet coefficients.

Spatially adaptive thresholding schemes are inherently attractive for OCT images, especially images of the human retina. This is because OCT tomograms of the retina are typically composed of distinct layered structures, and so a strong spatial correlation in wavelet coefficients can be expected in much of the image. In other words, the spatial features of OCT retinal images are significantly different than the spatial features of the speckle noise which corrupts the images. Therefore, spatially adaptive thresholding should perform very well on this type of image.

For spatially adaptive thresholding schemes to provide the best performance, the coefficients from image features must be strongly correlated to neighbouring coefficients. Near sharp image features such as edges, the wavelet transform coefficients can exhibit pseudo-Gibbs phenomena [54]. The magnitude of these oscillations in the wavelet transform is related to the sharpness of the discontinuity, as well as to its absolute location and the type of wavelet being applied. The pseudo-Gibbs artifacts reduce the effectiveness of denoising techniques, but can be largely removed by computing the wavelet transforms of all possible circular shifts of the original data, denoising the expanded data set, re-shifting, and averaging the results [54]. This technique has
been successfully applied in many diverse wavelet denoising algorithms for a variety of data types [47, 52, 55-57].

This procedure, called "cycle spinning", is equivalent to performing an undecimated wavelet decomposition [51, 54, 56]. In an undecimated transform, the downsamplers in the filter bank are removed and the filters are transformed according to the Noble identities. After transformation, the prototype lowpass and highpass filters $H_0(z)$ and $H_1(z)$ are transformed into $H_0(z^{2^j})$ and $H_1(z^{2^j})$ at each level $j$ of the filter bank. This converts the filters to general bandpass shapes, except for the zeroth-level filters, which retain the lowpass and highpass characteristics. The filter coefficients should be rescaled by $1/\sqrt{2}$ to keep the energy in each subband the same. Figure 42 shows an undecimated filter bank used to perform a wavelet decomposition. The reconstruction filter bank consists of filters $F_0(z^{2^j})$ and $F_1(z^{2^j})$, with the upsamplers removed. Extension to two dimensions is straightforward.

While the undecimated wavelet transform reduces the effects of Gibbs phenomena on denoising, it increases the computational complexity of the denoising algorithm. With the removal of the downsamplers in the filter bank, the redundancy of the data in the transform grows with each stage in the decomposition and each subband image will contain the same number of coefficients as the original image has pixels. In the two-dimensional case, the zeroth stage subbands will each be redundant by a factor of 4, the first stage subbands will each be redundant by a factor of 16, the second stage by a factor of 64, and the $j$th stage by a factor of $4^j$. The undecimated transform therefore destroys the orthogonality of the standard wavelet transform, which will have implications for the denoising technique discussed in 4.3.6.

Before discussing specific spatially adaptive algorithms, it is important to make an observation about the wavelet families that should be used with these types of algorithms. Since spatial correlation between neighbouring coefficients, and correlation between coefficients at the same spatial location but in different levels of the decomposition, is important to these algorithms, only symmetric or near-symmetric wavelet filters should be used. Symmetric filters produce uniform group delays over all frequencies, which avoids two serious problems with spatially adaptive algorithms. First, sharp image features (such as edges) that contain a wide range of frequencies will not be spread into adjacent pixels due to a nonuniform group delay.
Second, the exact spatial location of the “parent” coefficient in the adjacent decomposition level can be calculated only if the group delay of the filters are constant. Otherwise, the location of the parent will depend on the frequency content of the image, reducing the effectiveness of the scale correlation calculation. For undecimated wavelet transforms, the group delays of all the filters \( H_0(z), H_1(z), F_0(z), \) and \( F_1(z) \) must be symmetric or close to symmetric, with uniform group delay in their passbands.

### 4.3.5 Versatile Spatially Adaptive Thresholding

One of the spatially adaptive wavelet denoising techniques investigated for OCT image enhancement is the algorithm proposed by Pizurica et al [52]. Many wavelet denoising techniques require \emph{a priori} assumptions about the distribution of noiseless and noisy wavelet coefficients [44, 45, 47, 56], which limits their effectiveness when applied to images containing noise from a variety of sources. The Pizurica technique attempts to overcome this limitation by developing a general spatially adaptive technique that does not depend on statistical modeling of the wavelet coefficient distributions. An undecimated transform is used, based on the cycle spinning algorithm and the quadratic spline wavelet. This wavelet is symmetric, with uniform group delay in the associated lowpass and highpass filters. A user-defined parameter can be set to control the amount of smoothing in the algorithm, allowing the loss of sharpness in the edges to be arbitrarily traded off against noise reduction.

This algorithm is based on a three-step process. After carrying out a two-dimensional undecimated wavelet transform, each wavelet coefficient \( w_j^{(o)} \) is assigned a binary classification as either “likely signal” or “likely noise”, based on the user-defined smoothing parameter and an estimation of the noise variance. Next, the local spatial activity of each coefficient is calculated by finding the average energy of the neighbouring coefficients. Finally, the probability density functions (pdf’s) of the “signal” and “noise” coefficients are analyzed to generate individual threshold values for every coefficient in the decomposition, taking the local spatial activity into account as well. After thresholding, an inverse undecimated wavelet transform is carried out. A detailed description of the algorithm follows, taken from [52].

After carrying out the wavelet decomposition and generating the coefficients \( w_j^{(o)}[n_x, n_y] \) at levels \( j = 0 \rightarrow J \) and subbands \( o \in \{LL, HL, LH, HH\} \), the first step in the algorithm is to
classify the coefficients as either “likely signal” or “likely noise.” For clarity, the wavelet coefficients at an arbitrary level and orientation (subband) will be denoted \( w_k \), where \( k \) is the spatial coordinate. An additive noise model is assumed, where \( w_k \) is the observed noisy wavelet coefficient, \( y_k \) is the unknown noise-free coefficient, and \( v_k \) is the noise contribution. This gives

\[
w_k = y_k \oplus v_k
\]  

(4.35)

The binary classification mask is modeled as a random variable \( X_k \), taking values \( x_k \) from the binary set \{0,1\}. If a particular coefficient at spatial location \( k \) is “likely signal” then \( X_k = 1 \). Otherwise, \( X_k = 0 \). The “signal” and “noise” coefficients in a particular subband are modeled as identically distributed random variables with pdf’s \( p_{W_k|X_k}(w_k|1) \) and \( p_{W_k|X_k}(w_k|0) \), respectively.

At level \( j \) and orientation \( o \), an estimate of the binary mask \( x_{k,j}^{(o)} \) is obtained by comparing each noisy coefficient to the user-defined smoothing parameter \( K \), the estimate of the noise variance in that level and orientation \( (\hat{\sigma}_j^{(o)})^2 \), and the estimated noise-free coefficient \( \hat{y}_{k,j+1}^{(o)} \) at the next coarsest level.

\[
x_{k,j}^{(o)} = \begin{cases} 
0, & \text{if } |w_{k,j}^{(o)}| |\hat{y}_{k,j+1}^{(o)}| < (K\hat{\sigma}_j^{(o)})^2 \\
1, & \text{if } |w_{k,j}^{(o)}| |\hat{y}_{k,j+1}^{(o)}| \geq (K\hat{\sigma}_j^{(o)})^2 
\end{cases}
\]  

(4.36)

Since this mask estimation technique depends on the estimated noise-free coefficients at the previous level, it is necessary to initialize the process by using

\[
\hat{y}_{k,j}^{(o)} = w_{k,j}^{(o)}
\]  

(4.37)

This implies that the wavelet coefficients in the coarsest decomposition level are noise-free, which may or may not be a valid assumption in reality. The estimate of \( (\hat{\sigma}_j^{(o)})^2 \) is obtained using (4.29) - (4.31). A possible error in the algorithm occurs here; since the authors have made use of an undecimated wavelet transform, the noise is not filtered by the same prototype lowpass and highpass filters at every stage in the filter bank. Instead, it is possible that the transformed bandpass filters with Z-transforms \( H_0(z^{j+1}) \) and \( H_1(z^{j+1}) \) should be used.
Once the binary masks have been generated for all the subbands in the current level \( j \), it is necessary to calculate the local spatial activity for each coefficient in level \( j \). Like the binary mask, the local spatial activity values are also modeled as random variables \( E_k \) with values \( \{e_k\} \). Although any indicator of spatial activity can be used, the authors in [52] use the average energy in a square window around each pixel and the energy of the parent coefficient at the same spatial location at level \( j + 1 \). This method takes advantage of the fact that coefficients from signal tend to cluster spatially and across levels, as discussed above in section 4.2.

With the binary masks and local activity values calculated, it is now possible to estimate \( \hat{y}_k \) from the noisy observation of \( w_k \). The mean squared error of \( \hat{y}_k \) is approximately minimal when

\[
\hat{y}_k = \frac{\hat{\xi}_k \hat{\eta}_k - \omega_k}{1 + \hat{\xi}_k \hat{\eta}_k}
\]

with the approximated functions \( \hat{\xi}_k \), \( \hat{\eta}_k \), and \( \hat{\rho} \) given by

\[
\hat{\xi}_k = \frac{\hat{p}_{M_k|X_k}(m_k | 1)}{\hat{p}_{M_k|X_k}(m_k | 0)}
\]

\[
\hat{\eta}_k = \frac{\hat{p}_{E_k|X_k}(e_k | 1)}{\hat{p}_{E_k|X_k}(e_k | 0)}
\]

\[
\hat{\rho} = \frac{\sum_{k=1}^{M} \hat{x}_k}{M - \sum_{k=1}^{M} \hat{x}_k}
\]

Since this algorithm assumes no \textit{a priori} knowledge of the distribution of signal and noise coefficients, no model can be used to generate the pdf’s \( p_{M_k|X_k}(m_k | x_k) \) and \( p_{E_k|X_k}(e_k | x_k) \) exactly. Their estimates \( \hat{p}_{M_k|X_k}(m_k | x_k) \) and \( \hat{p}_{E_k|X_k}(e_k | x_k) \) must therefore be calculated directly from the observed data. This is done by generating histograms of the datasets

\[
\{m_k : \hat{x}_k = 1\}
\]

\[
\{m_k : \hat{x}_k = 0\}
\]

\[
\{e_k : \hat{x}_k = 1\}
\]

\[
\{e_k : \hat{x}_k = 0\}
\]
In (4.41), \( \hat{r} \) is an estimate of the probability ratio of \( P(X_k = 1) \) to \( P(X_k = 0) \).

Typically, the histograms will be significantly asymmetric and have long tails at higher values of \( m \) and \( e \). This makes computation of the ratios (4.39) and (4.40) prone to error in the tail regions. Therefore, the authors propose a simple piecewise linear fitting to the log of (4.39) and (4.40) to correct for these errors. As discussed below in section 4.4, this simple model is not appropriate for ophthalmic OCT data and leads to degraded performance of this method. Once the functions \( \hat{\xi}_k \), \( \hat{\eta}_k \), and \( \hat{r} \) are calculated, the threshold in (4.38) can be applied to the noisy wavelet coefficients, creating a set of denoised coefficients \( \hat{y}_k \). The algorithm is completed by an inverse undecimated wavelet transform.

Although this algorithm is presented as being able to adapt to a wide variety of noise sources and conditions, its performance with ophthalmic OCT images was quite poor as discussed in section 4.4. Significant oversmoothing of the edges was observed, even with the \( K \) parameter set as low as possible. This is most likely due to the fact that the distributions of the wavelet coefficients are not well-modeled by the simple piecewise linear fit proposed by the authors.

4.3.6 **Context-Based Spatially Adaptive Thresholding**

The second spatially adaptive wavelet denoising technique investigated for OCT image enhancement is the algorithm proposed by Chang et al [40]. This technique is similar to the one described in section 4.3.5, but is based on a context coder used for image compression. To that end, the Chang technique is somewhat more complicated than the Pizurica technique, and requires some basic assumptions about the distribution of the wavelet coefficients in the image. Specifically, the algorithm assumes that the coefficients in each subband form a zero-mean generalized Gaussian distribution (GGD) with a shape parameter between 0.5 and 4.0. The additive noise is modeled as an independent identically distributed random variable. This noise model and range of shape parameters covers coefficient distributions most often found in noisy images, and seems to work well for OCT tomograms as well. The authors use an 8-tap symmlet mother wavelet, which is nearly symmetric and produces near-uniform group delay in the passbands of the associated filter bank.

As its name implies, this algorithm works by calculating the "context" of each wavelet coefficient. The context is a weighted average of neighbouring coefficients within the same
subband at level $j$, and the parent coefficient at level $j+1$. This is very similar to the idea of local spatial activity described in [52]. A threshold for each coefficient is calculated based on the ratio of the noise variance to the variance of all coefficients having a similar context.

This algorithm also makes use of the undecimated wavelet transform to improve denoising performance. However, since the orthogonality of the wavelet decomposition is lost in the undecimated case, the noise contribution $v_k$ becomes correlated within a subband. The authors therefore propose to separate the noisy coefficients $w_k$ into $2^{2(j+1)}$ uncorrelated sets at the $j$th level of the decomposition, by downsampling shifts of the coefficients.

$$\{w_j^{(o)}[2^{j+1}k_1 + k_2, 2^{j+1}k_1 + k_2]\}_{k_1, k_2 = 0}^{2^{2(j+1)} - 1}$$

(4.43)

These uncorrelated sets of coefficients are analyzed independently, and thresholds are calculated for each uncorrelated set. The process then consists of four main steps (after the undecimated wavelet transform): separation of the subbands into uncorrelated sets, calculation of the context values for each coefficient, calculation of the variance for all coefficients with a similar context, and finally calculation of the threshold values. A detailed description of the algorithm follows, taken from [40].

The starting point is again an observation of wavelet coefficients $w_k$, consisting of noise-free coefficients $y_k$ corrupted by a noise contribution $v_k$. Here it is assumed that the coefficients have already been separated into their uncorrelated sets. The noise-free coefficients are modeled as independent samples of a GGD, with variance $\sigma_y^2$. The noise coefficients are modeled as independent samples of a Gaussian distribution (GGD with a shape parameter of 2), with variance $\sigma_n^2$. Under these assumptions, the desired threshold to minimize the expected square error between the thresholded coefficients and the actual noise-free signal is approximated by [58]

$$T_B = \frac{\sigma_n^2}{\sigma_y^2}$$

(4.44)

Since $\sigma_n^2$ can be calculated using (4.28), the problem is that of finding a good estimate for the variance of the noise-free wavelet coefficients. The authors propose to calculate a value of $\sigma_y^2$ for each individual coefficient by using the concept of context modeling.
The coefficients $y_k$ are modeled as components in a discrete random field, consisting of a collection of independent zero-mean GGD's with spatially varying shape parameters and variances. To estimate the variance for the GGD random variable at a particular coefficient location, the variance of a group of coefficients with similar spatial activity ("context") is calculated. First, it is necessary to calculate the context $Z$ for each noisy coefficient $w_k$. This is done by using a weighted average of the neighbouring coefficients in a $3 \times 3$ square window, plus the parent coefficient at the same spatial location at the next-highest level in the decomposition.

$$Z[n_x, n_y] = g' u_{n_x, n_y}$$  \hspace{1cm} (4.45)

Here, $g$ is the weighting vector and $u_{n_x, n_y}$ is a $p \times 1$ vector containing the 9 neighbouring coefficients of the current coefficient. The weighting vector is calculated by minimizing the least squares error between the observed coefficients $w_k$ and the context $Z$. This gives

$$g = (U'U)^{-1} U'w$$  \hspace{1cm} (4.46)

Here, $U$ is a $N \times p$ matrix with each row being $u_{n_x, n_y}$ for all $n_x, n_y$, and $w$ is a $N \times 1$ vector containing all of the noisy wavelet coefficients $w_k$.

With the context values calculated for each coefficient in the uncorrelated sets, it is possible to calculate the variance associated with the random variable for each coefficient. For a given coefficient $w_{k_0}$ in subband $(o)$ at level $j$, having a context value $Z[n_{x_0}, n_{y_0}]$, a group of $2L + 1$ other coefficients is assembled from the same subband and level. The additional coefficients are selected such that they are the $L$ coefficients with the closest context values greater than $Z[n_{x_0}, n_{y_0}]$, and the $L$ coefficients with the closest context values less than $Z[n_{x_0}, n_{y_0}]$. This is analogous to dropping a window around $Z[n_{x_0}, n_{y_0}]$ that contains $2L + 1$ points, with $Z[n_{x_0}, n_{y_0}]$ in the middle of the window. $L$ is chosen as $L = \max(50, 0.02 \times N)$. The set of coordinates defining $\{w[n_x, n_y]\}$ that falls within this window is denoted $B_{xy}$. From this grouping, the variance $\hat{\sigma}_y^2$ is estimated as

$$\hat{\sigma}_y^2[n_{x_0}, n_{y_0}] = \max \left( \frac{1}{2L+1} \sum_{[l,m] \in B_{xy}} w[l,m]^2 - \sigma_n^2, \ 0 \right)$$  \hspace{1cm} (4.47)
The noise variance is subtracted because the \( w_k \) coefficients are noisy observations. The max statement is included to remove the possibility of calculating a negative variance when the coefficient energies are less than the noise variance. With \( \hat{\sigma}_y^2 \) estimated, the threshold can be calculated from (4.44) using \( \hat{\sigma}_y \) in place of \( \sigma_y \).

Although this thresholding scheme produces much better results than the method discussed in section 4.3.5, it does suffer from some drawbacks. First of all, the necessity of performing context calculations and variance estimates on \( 2^{2(j+1)} \) uncorrelated sets of coefficients for every subband at level \( j \) makes this method much more computationally intensive than the previous method. The additional complexity translates into a 20% increase in execution time when the code is implemented in Matlab. Furthermore, it is not able to adapt to all coefficient distributions, and is limited to GDD distributions with shape parameters between 0.5 and 4. This limitation does not seem to be serious for OCT images, but should still be pointed out.

Finally, the algorithm is fully automatic; while this may seem like a benefit, in reality it is often desired to have some sort of user-defined parameter that can trade off feature smoothing for increased noise reduction. This is especially true when low-power SLD sources are used for UHR-OCT imaging, when the noise level becomes quite high relative to the signal [59]. In section 4.4, a parameter \( K \) was added to the algorithm to enable user control of the denoising. The parameter can be thought of as a gain setting that affects only the HL subband thresholds, changing the values from \( T_B^{(m)} \) to \( KT_B^{(m)} \). Since ophthalmic OCT tomograms consist mainly of horizontal layered structures, most of the real edges in the image are contained in the LH subbands. Speckle noise, on the other hand, is approximately equal in the HL and LH subbands. Therefore by increasing the attenuation of the HL subbands only, increased noise reduction can be achieved while minimizing additional smoothing of the horizontal edges characteristic of ophthalmic OCT tomograms. As discussed below, a significant improvement in image quality was observed after optimizing the \( K \) parameter.

### 4.4 Denoising of Ophthalmic OCT Images

With four different wavelet denoising techniques and a set of image quality metrics in hand, it is now possible to evaluate their performance when applied to ophthalmic UHR-OCT images. Recently, broadband SLD light sources have been developed for UHR-OCT imaging that are
capable of providing image resolutions of ~3μm in the retina or ~2μm in scattering tissue [59]. These sources are compact, inexpensive, and easy to operate. One major limitation of these new SLD’s is that their output power is quite low compared to solid state lasers, causing a 3dB reduction in SNR when applied to ophthalmic UHR-OCT imaging. The SNR is not reduced because the incident power on the sample is lower, but because the interferometer design cannot be as efficient with a low-power source as discussed in section 1.2. UHR-OCT systems based on SLD light sources will have greater clinical utility if the SNR can be increased to a level comparable to that of traditional systems based on femtosecond pulse lasers. Therefore, low SNR images acquired with broadband SLD light sources make good test cases to evaluate the performance of the various wavelet denoising techniques discussed in section 4.3.

Figure 43 shows a schematic of the UHR-OCT system used to acquire ophthalmic images, using a broadband SLD light source (Superlum Diodes, Ltd.). The source generates > 150 nm of optical bandwidth at a center wavelength of 890 nm, with a CW output power of 4 mW. The system uses a scanning galvanometer to scan the optical path in the reference arm with a velocity of 410 mm/s. The electronic detector operates at a frequency of 1 MHz, with a bandwidth of 170 kHz. The images acquired with this system consist of 600 A-scans, acquired over a 4 second time span. The transverse resolution was 20 μm, while the axial resolution in the retina was 2.3 μm. Measured SNR using the SLD light source was 94 dB with 750 μW of incident power on the eye, compared to 97 dB typically obtained using a solid state laser with the same incident power. The scan depth was 1.5 mm in tissue, and the scan length was 6 mm.

Figure 44 shows a series of typical UHR-OCT image of a healthy human retina, acquired in vivo. Figure 44a shows a raw image acquired with the broadband SLD light source. The major intraretinal layers can be delineated, but the effect of speckle noise is severe. To mask this effect, OCT images are normally pixel thresholded by a simple algorithm which removes pixels below a certain intensity. The pixel threshold is calculated by estimating the noise variance in a background region of the image, and then removing pixels below the threshold to reduce the noise variance to a user-defined level. After pixel thresholding, the image appears as shown in figure 44b. The effect of pixel thresholding is to make the retinal features easier to identify, at the expense of losing signal information contained in the removed pixels. For comparison, a UHR-OCT image of a different healthy volunteer acquired using a solid state laser is shown in figure 44c (after pixel thresholding). The SNR of the image in figure 44c is 97 dB, 3dB higher
Figure 45. Magnitude responses and group delays for the $H_k$ filters used in an undecimated filter bank, with an 8-tap symlet mother wavelet as the generating function. Group delays within the 3 dB passband of the filters are nearly constant, with a +/- 1 pixel variation at most.

Figure 46. Magnitude responses and group delays for the $H_k$ filters used in an undecimated filter bank, with an 8-tap symlet mother wavelet as the generating function. Group delays within the 3 dB passband of the filters are nearly constant, with a +/- 1 pixel variation at most.

Figure 47. Original noisy image showing ROI's used to calculate image quality. ROI 1 is a background area, used to estimate the noise variance. ROI's 2-4 are homogenous areas, used to calculate ENL. ROI's 5-7 are edge regions, used to calculate local SNR. Global SNR and global $\beta$ parameter (sharpness) are also calculated.
than the image acquired with the solid state laser. Delineation of the intraretinal layers is more clear, especially in the choroid and weakly-scattering layers such as the external limiting membrane (ELM). For a wavelet denoising technique to be successful, it should increase the image quality of the tomograms acquired with the SLD light source to match or exceed the quality of the tomograms acquired with a solid state laser.

4.4.1 Algorithm Construction

The Matlab software package was used to evaluate the denoising techniques discussed in section 4.3. For the case of global soft thresholding and the SUREShrink algorithm using a decimated (standard) wavelet transform, the Matlab wavelet toolbox was used. For global soft thresholding, two cases were compared: first the threshold was calculated using the variance of coefficients in the entire \( HH_0 \) subband with (4.28), and then the threshold was calculated using a background region of the \( HH_0 \) subband. Very little difference in performance was observed between the two cases. For the SUREShrink algorithm, the thresholding was carried out by applying the one-dimensional algorithm along the rows and columns of the original image. For the two spatially adaptive algorithms using undecimated wavelet transforms, custom Matlab code was written to carry out the denoising. In all cases, the nearly-symmetric 8-tap symlet wavelet was used to generate the analysis and synthesis filter banks. Even-symmetric boundary extension of the original image was performed, with each edge extended by the length of the longest filter in the system.

For the spatially adaptive algorithms, the wavelet decomposition and reconstruction were implemented using digital filter banks. Since each digital filter has an associated group delay, the subband images at each decomposition level are successively shifted down and to the right with respect to the subbands in the previous level. Choosing a nearly-symmetric mother wavelet to generate the filters results in each filter having a nearly constant group delay; therefore all of the frequency components at level \( j \) are translated by almost the same amount when moving to level \( j+1 \) in the filter bank. By calculating the group delay of each filter in the analysis bank prior to generating the local spatial activity or context values for the wavelet coefficients, the coordinates of the parent coefficient at level \( j+1 \) could be determined with high precision for each coefficient at level \( j \). This is crucial for an accurate calculation of local spatial activity and
context. Figure 45 and figure 46 show the amplitude responses of the “lowpass” and “highpass” filters used in each stage of the undecimated decomposition filter bank for a 3-level analysis, along with their group delays. Note that in the 3dB passband of each filter, the group delay varies by at most +/- 1 pixel. Purely symmetric wavelets, such as biorthogonal splines, have exactly constant group delay. The cubic spline wavelet was investigated for denoising, but was found to produce slightly worse performance than the symlet wavelet.

Since filter banks were used to perform the wavelet transforms, each time-domain convolution of the filter impulse response with the rows or columns of the image results in an extension of the length of the output. The extra points at the beginning and end of the output of each stage in the filter bank do not contain accurate information, and were discarded. Specifically, if the filter at level \( j \) has length \( m(j+1) \), where \( m \) is the length of the mother wavelet, and the row or column being operated on has length \( n \), then the output will be of length \( n + m(j+1) - 1 \). The first \( m(j+1) - 1 \) and the last \( m(j+1) - 1 \) points are discarded. After the reconstruction step, the symmetrically extended boundaries are removed and the final image is analyzed with the metrics discussed in section 4.3.1.

For all of the denoising techniques, the same ROI’s were used to quantify the image quality after processing. Figure 47 shows the greyscale original OCT image used as a test case for the various denoising techniques, with the ROI’s overlaid in white. A background region was defined in an area containing no signal, and was used to estimate the noise variation. Three ROI’s were defined in homogeneous areas with signal levels ranging from low to high, and were used to calculate the ENL values. Three ROI’s were defined that straddled edge regions with signal levels ranging from low to high, and were used to calculate the local SNR values. For each denoising technique, a global SNR and global \( \beta \) value were also calculated. Table 1 shows a summary of the denoising results for all of the methods evaluated.

**4.4.2 Time Domain Filtering Results**

Two time domain filtering methods were included in this study to provide a comparison for the wavelet techniques. The first method, pixel thresholding, is the standard method to mask the appearance of noise in OCT images. All pixels with a value less than a certain cutoff value (in this case 700) are removed from the image. In the unprocessed image, the noise had a variance
Figure 48. OCT image after pixel thresholding. Noise is masked, but image appears grainy and signal appears low.

Figure 49. OCT image after 3x3 median filtering. Noise is reduced, but image appears pixelated and oversmoothed.

Figure 50. OCT image after global soft thresholding. Good noise reduction is obtained, and image is less smooth than in figure 49. Fine details appear slightly blurred.
Figure 51. OCT image after SUREShrink soft thresholding. The algorithm is too conservative, and the image appears very similar to the pixel thresholded image in figure 48.

Figure 52. OCT image after local spatial activity thresholding. The wavelet coefficients do not fit the model assumed in this algorithm, leading to large scale oversmoothing of the data.

Figure 53. OCT image after context modeling thresholding, $K = 1$, equivalent to the fully automated algorithm of Chang et all. The algorithm provides good noise reduction while maintaining edge sharpness to 99% of the original.
Figure 54. OCT image after context modeling thresholding, $K = 4$. Increasing the thresholding in the HL subband significantly improves the denoising performance without seriously compromising edge sharpness. SNR is improved by 7 dB compared to the original, with sharpness maintained to 97% of the original. K parameter allows the user to tune the denoising/sharpness tradeoff.
of 1.1x10^4, a mean of 625, and a range of 0 to 883. Figure 48 shows the resulting image after thresholding. From table 1, the local SNR, CNR, and ENL values all increase slightly except for one ROI where the contrast actually decreases. The drawback to this technique is a significant loss of edge sharpness and a very marginal increase in global SNR.

The second time domain technique was a simple 3x3 pixel median filter. Figure 49 shows the resulting image after filtering. In this figure, as with figures 50 – 54, the image is displayed after matching the background noise level to the level in figure 48 by pixel thresholding. This provides a fair comparison of image quality, without wide variances in noise to confound the reader. Viewing images comparatively in this manner results in SNR variations appearing to be increases or decreases in signal level. If the background noise levels in all of the images were not matched, the SNR variations would appear as increases or decreases in the noise level. While the local SNR, CNR, and ENL values are all dramatically better after median filtering, the image appears excessively smoothed. The low $\beta$ value also indicates loss of edge sharpness in this technique.

<table>
<thead>
<tr>
<th>ROI</th>
<th>Original Image</th>
<th>Pixel Thresholding</th>
<th>3x3 Median Filter</th>
<th>Global Soft Thresholding</th>
<th>SUREShrink</th>
<th>Local Spatial Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SNR (dB)</td>
<td>CNR (dB)</td>
<td>ENL</td>
<td>Beta</td>
<td>SNR OCT (dB)</td>
<td>SNR (dB)</td>
</tr>
<tr>
<td>1</td>
<td>9.65</td>
<td>2.26</td>
<td>85.17</td>
<td>-</td>
<td>9.72</td>
<td>2.41</td>
</tr>
<tr>
<td>2</td>
<td>9.57</td>
<td>2.01</td>
<td>82.04</td>
<td>-</td>
<td>9.85</td>
<td>2.19</td>
</tr>
<tr>
<td>3</td>
<td>8.97</td>
<td>1.41</td>
<td>62.21</td>
<td>-</td>
<td>9.96</td>
<td>1.56</td>
</tr>
<tr>
<td>4</td>
<td>6.80</td>
<td>1.12</td>
<td>22.88</td>
<td>-</td>
<td>7.71</td>
<td>1.07</td>
</tr>
<tr>
<td>5</td>
<td>7.04</td>
<td>0.67</td>
<td>25.62</td>
<td>-</td>
<td>8.99</td>
<td>0.74</td>
</tr>
<tr>
<td>6</td>
<td>8.16</td>
<td>0.71</td>
<td>42.86</td>
<td>-</td>
<td>10.44</td>
<td>0.78</td>
</tr>
<tr>
<td>Global</td>
<td>N/A</td>
<td>94.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ROI</th>
<th>Context Modeling, K = 1</th>
<th>Context Modeling, K = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SNR (dB)</td>
<td>CNR (dB)</td>
</tr>
<tr>
<td>1</td>
<td>10.46</td>
<td>2.74</td>
</tr>
<tr>
<td>2</td>
<td>10.37</td>
<td>2.43</td>
</tr>
<tr>
<td>3</td>
<td>9.76</td>
<td>1.70</td>
</tr>
<tr>
<td>4</td>
<td>7.04</td>
<td>1.23</td>
</tr>
<tr>
<td>5</td>
<td>7.46</td>
<td>0.76</td>
</tr>
<tr>
<td>6</td>
<td>8.88</td>
<td>0.85</td>
</tr>
<tr>
<td>Global</td>
<td>0.99</td>
<td>99.2</td>
</tr>
</tbody>
</table>

Table 1: Image quality metrics for various denoising techniques
4.4.3 Global Soft Thresholding and SUREShrink Results

Figure 50 shows the resulting OCT image after global soft thresholding. The local and global SNR values are increased dramatically using this method, but the image is somewhat blurred out. Thin layers appear slightly thicker, and edges appear smoothed. This is a known problem with the global soft threshold method, as discussed in section 4.3.2, since it does not adapt to local conditions in the various subbands of the wavelet decomposition. Nevertheless, the ELM layer appears brighter than in the simple pixel thresholding case, so some improvement in quality is observed.

Figure 51 shows the resulting OCT image after applying the SUREShrink thresholding algorithm. Qualitatively, this image appears almost identical to the pixel thresholded image. This is confirmed by examining table 1, as the quality metrics are virtually identical between the two cases. It seems that the heuristic SURE algorithm is too conservative for OCT data, conserving edge sharpness at the expense of noise reduction.

4.4.4 Spatially Adaptive Thresholding Results

Figure 52 shows the resulting OCT image after applying the spatially adaptive algorithm based on modeling the local spatial activity, discussed in section 4.3.5. For this image, the $K$ parameter was set as low as possible ($K = 0.15$) without causing the piecewise linear model of the pdf ratios to fail completely. Even with $K$ reduced to this low value, the image is unusable due to oversmoothing of the features. This result is disappointing, but is likely due to the fact that the pdf’s of the wavelet coefficients are not well-modeled by Gaussians, and therefore the piecewise linear model of the log of the pdf ratios is a very poor fit. This causes significant overestimation in the calculation of the thresholds, leading to an image that is far too smooth. Furthermore, the algorithm does not separate the wavelet coefficients into uncorrelated sets, which introduces correlation between noisy coefficients within a subband and further reduces the likelihood of the noise coefficients falling into a Gaussian distribution, as required by the model.

Figure 53 shows the resulting OCT image after applying the spatially adaptive algorithm based on context modeling, discussed in section 4.3.6. For this image, the $K$ parameter was set to unity. The processing was therefore exactly equivalent to the method described in [40]. The results here are quite good, with a significant improvement in all of the image quality metrics.
and only a 1% degradation in edge sharpness. This is likely due to the increased care taken by the algorithm to separate the wavelet coefficients into uncorrelated sets. The context calculation of this algorithm is also more advanced, using a weighted average of energies instead of the simple average of energies used in [52].

By varying the $K$ parameter of the context modeling algorithm, the user can increase the thresholding conditions on the HL subbands as discussed in section 4.3.6. This added features is very useful, and enables the algorithm to produce results significantly better than with $K$ set to 1. Figure 54 shows the resulting OCT image using a $K$ value of 4, which is near optimal in terms of trading off image sharpness and noise reduction for this particular image. The image quality of this tomogram is the best amongst all the methods investigated in this study. The improved ENL and SNR values compared to the $K = 1$ case make a significant difference in the qualitative appearance of the image, while the sharpness is only degraded by 3%. The global SNR improvement over the original image is 7.1dB, which is roughly equivalent to increasing the laser source power by a factor of 5.

### 4.4.5 Conclusions

Clearly, the spatially adaptive wavelet denoising technique described by Chang et al [40] is very useful for improving the image quality of ophthalmic OCT tomograms. When applied to images acquired with low-power SLD light sources, the image quality improves to a level comparable to that of images acquired with solid state light sources. By taking advantage of the horizontal banded nature of retinal tomograms, the algorithm can be modified by adding a user defined parameter $K$ to increase the thresholding on the HL subbands. This variable thresholding improves the noise reduction properties of the algorithm, without significantly affecting image sharpness since most of the edges in retinal images lie in the LH subbands. The additional parameter also allows the user to control the denoising process, instead of being locked into a fully automated algorithm. A global SNR improvement of $> 7$dB is possible when using this method once the $K$ parameter is heuristically optimized.

### 4.5 Future Work with Wavelet Denoising

While the work done on wavelet denoising for this thesis project illustrates the utility of wavelets for image enhancement of OCT tomograms, more work can be done to optimize the
process. An in-depth study comparing various symmetric and near-symmetric wavelets of varying lengths can be carried out to determine the optimal mother wavelet for OCT images. The number of decomposition levels in the wavelet transform may also affect denoising performance, and should be investigated. Furthermore, translation-invariant (TI) multiwavelets based on undecimated M-band filter banks have recently been shown to provide better denoising performance for one dimensional signals [55]. Spatially adaptive denoising schemes based on TI multiwavelets have shown additional improvements in performance [57]. These methods could be extended to two dimensions and also applied to OCT imaging. An investigation on the effects of wavelet filtering on image segmentation could be carried out, as a reduction in noise should lead to much more accurate segmentation algorithms. This could be applied to newly developed spectral OCT systems, which are capable of acquiring full three-dimensional data sets. Finally, the code used to implement the algorithms should be ported to C++. Due to Matlab’s inefficient for loop implementation, code execution time for the spatially adaptive algorithms is unnecessarily high, and needs to be reduced for the techniques to find widespread use.
CHAPTER 5: CONCLUSIONS

This thesis project has attempted to develop and apply digital signal processing techniques for data acquisition, demodulation, spectroscopic analysis, and image enhancement to the field of optical coherence tomography. Of most relevance is the work on spectroscopic imaging and image enhancement. The results presented in this thesis indicate that spectroscopic OCT imaging may have the capability to enhance the contrast of specific tissue types, potentially leading to a method for early detection of cancer. Much more work is required in this area, including a comprehensive spectroscopic study of an animal cancer model, and a detailed optimization of the spectroscopic analysis parameters specifically for the purpose of detection of cancer. The results presented using wavelets for OCT image enhancement are also very promising; OCT image quality can be improved both quantitatively ( >7dB improvement in SNR) and qualitatively, using a digital method, and with very little loss in image sharpness ( <3% degradation). Although for this project an existing wavelet filtering technique was modified to show the utility of wavelets for OCT image enhancement, new methods designed from the ground up could further improve on these initial results. It is the author’s hope that the work done for this thesis will enable others to find new applications for OCT technology, and that these new applications will some day lead to concrete improvements in human health.
APPENDIX A: REFERENCES


tomography by "path length encoded" angular compounding. *Journal of Biomedical
reduction in optical coherence tomography by frequency compounding. *Journal of
33. Rogowska, J. and Brezinski, M.E., Evaluation of the adaptive speckle suppression filter
for coronary optical coherence tomography imaging. *IEEE Transactions on Medical
34. Rogowska, J. and Brezinski, M.E., Image processing techniques for noise removal,
enhancement and segmentation of cartilage OCT images. *Physics in Medicine and
35. Choi, H., Milner, T.E., and Bovik, A.C. Speckle Noise Reduction and Segmentation on
Polarization Sensitive Optical Coherence Tomography Images. in *A New Beginning for
Human Health: Proceedings of the 25th Annual International Conference of the IEEE
Institute of Electrical and Electronics Engineers Inc.
tomography. in *Optical and Imaging Techniques for Biomonitoring III*. 1997. San Remo,
Italy: SPIE.
Cambridge Press.
Prentice-Hall, Inc.
40. Chang, S.G., Yu, B., and Vetterli, M., Spatially adaptive wavelet thresholding with
42. Sattar, F., Floreby, L., Salomonsson, G., and Lovstrom, B., Image enhancement based on
888-895.
for ultrasonic speckle suppressing. *IEEE Transactions on Medical Imaging*, 1999. 18(9):
p. 787-794.
44. Achim, A., Bezerianos, A., and Tsakalides, P., Novel Bayesian multiscale method for
speckle removal in medical ultrasound images. *IEEE Transactions on Medical Imaging*,
45. Achim, A., Tsakalides, P., and Bezerianos, A., SAR image denoising via Bayesian
wavelet shrinkage based on heavy-tailed modeling. *IEEE Transactions on Geoscience
46. Foucher, S., Benie, G.B., and Boucher, J.M., Multiscale MAP filtering of SAR images.


