ALGORITHMIC TECHNIQUES FOR MANUAL-SCANNING AND
MULTI-SENSORY REPRESENTATION OF
OPTICAL COHERENCE TOMOGRAPHY DATA

BY

ADEEL AHMAD

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Adviser:

Professor Stephen A. Boppart

ABSTRACT

Optical coherence tomography (OCT) is an emerging biomedical imaging modality that interferometrically measures the depth-resolved back-scattered light from a sample. The high resolution subsurface imaging capabilities of OCT can potentially provide valuable diagnostic feedback about tissue morphology during intra-operative applications of OCT. However, the intra-operative use of OCT is hindered by the lack of suitable sample scanning mechanisms and the difficulty in real-time interpretation of OCT data.

Screening or diagnostic procedures typically require high resolution imaging over a large field-of-view. The current scanning mechanisms, which are predominantly based on mechanically scanning the OCT beam over the specimen, have a limited scan range and lack the required flexibility for large field-of-view imaging. Moreover, OCT, being a high resolution imaging modality, requires a very high data acquisition rate and would generate huge amounts of data if used to image over a large surgical field. These extremely high data rates would make the real-time interpretation of OCT data a challenge, especially in the highly demanding operating room environment where the surgeon has to simultaneously keep track of a number of parameters. In this thesis, algorithmic techniques are proposed that may be helpful in overcoming some of these limitations.

Image acquisition over a large field-of-view with flexible scan geometry would require a manually scanned hand-held probe. This thesis describes a cross-correlation based image acquisition technique that may be used for image formation by compensating for the variable scan velocity encountered while manually scanning the probe. This thesis also describes an approach for multi-sensory representation of OCT

data by converting OCT data and images into sound. Audio rendering of OCT data when used in conjunction with the visual display may facilitate rapid interpretation of OCT data as the human auditory sense can detect more rapid transitions in the data than the visual sense and hence may be used for identifying suspected regions in real-time, which can subsequently be monitored in high resolution using the visual display.

ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge my thesis adviser, Professor Stephen Boppart, for his invaluable guidance, constant encouragement, and patience in dealing with me throughout the course of this work.

This thesis would not have been possible without the help and collaborations with a number of people. I am especially indebted to Dr. Steven G. Adie for his many insightful suggestions and for his availability whenever I needed help. The contribution from Morgan Wang, whose initial work on sonification proved to be a great foundation to build upon, is gratefully acknowledged. I would also like to thank Dr. Utkarsh Sharma and Dr. Woonggyu Jung, who have helped me immensely in my experiments throughout this work. In addition, the discussions with Prof. Scott Carney, Prof. Michelle Wang and Prof. Sever Tipei have also been very valuable. Finally, I want to thank the whole Biophotonics imaging group. Their great companionship and friendship made me look forward to coming to the lab every day.

These acknowledgments would be incomplete without mentioning my parents, who have always believed in my abilities and have been a constant support throughout my life.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ARR   A-scan Redundancy Ratio

AWGN Additive White Gaussian Noise

CCF   Cross-Correlation Function

FM   Frequency Modulation

FWHM Full Width Half Maximum

HCI   Human-Computer Interface

Jnd   Just Noticeable Difference

MA   Moving Average

MEMS Micro-Electro-Mechanical Systems

MRI   Magnetic Resonance Imaging

OCE   Optical Coherence Elastography

OCT   Optical Coherence Tomography

ODT   Optical Doppler Tomography

SD-OCT Spectral-Domain Optical Coherence Tomography

CHAPTER 1   INTRODUCTION

One of the most significant goals of surgical guidance is to provide assistance for the accurate placement and positioning of surgical tools to target an area of interest for various screening, diagnostic or therapeutic purposes. This is typically done by co-registering the pre-operative acquired images with the morphology of the tissue during a procedure [1]. Recently, imaging modalities such as fluoroscopy [2], x-ray computer tomography (CT) [3, 4], ultrasound imaging [5-7] and magnetic resonance imaging (MRI) [8, 9] have been integrated into the surgical suite to provide real-time intra-operative feedback during surgical procedures. Each of these modalities has its own advantages and limitations. While CT provides good anatomical information, the lack of sufficient contrast between soft tissues and the ionizing radiation employed during imaging limit its use in the operating room. MRI systems have been successfully employed for procedures such as needle placement and real-time monitoring of therapies such as cyro-surgery and thermal ablation [10]. However, MRI is very expensive, has relatively low resolution, and requires high precautionary measures which limit its widespread intra-operative usage. Ultrasound has been extensively studied for intra-operative imaging applications because of its simplicity and low cost. However, the contact required between the transducer and tissue and low resolution of ultrasound images make it unsuitable for many applications.

Optical imaging techniques such as visual endoscopy, fluorescence, multiphoton, confocal microscopy and optical coherence tomography (OCT) have an order of magnitude higher spatial resolution compared to other imaging techniques, have high sensitivity, are non-contact, non-ionizing and relatively inexpensive. Among these

techniques, OCT in particular is in a unique position to intra-operatively provide real-time, high-resolution, sub-surface imaging data to guide certain surgical procedures [11]. OCT has been clinically demonstrated in a diverse set of medical and surgical specialties, including ophthalmology, cardiology, oncology, dermatology and dentistry [12]. However, the limited penetration depth of OCT limits its non-invasive applications to the easily accessible parts of the human body such as the skin, oral or auditory cavities, or the eye. To access deeper regions within the body, OCT imaging is coupled with needle probes, catheters or endoscopes [13]. OCT may find its greatest application in minimally invasive surgical procedures where high resolution sub-surface images may typically be required for detection of tumor margins [14], resection of abnormal tissue, and guidance of needle biopsies, among many others [15].

The intra-operative use of OCT presents itself with numerous challenges. One of these is the requirement to image over large fields-of-view in real-time such as for screening or surgical guidance. A multi-modal approach can be employed, where a relatively low resolution imaging technique such as fluorescence imaging, MRI, PET/SPECT and ultrasound imaging is used to screen a large field-of-view and identify a suspicious region or area of interest for subsequent high resolution imaging using OCT over a relatively small field-of-view. This approach will require techniques to co-register the multi-modal image data from various modalities, which is very challenging and an active area of research [16]. However, little work has been done in combining optical imaging methods with other imaging modalities [17-19].

Another approach could be to utilize the high speed imaging capabilities of OCT for large field-of-view imaging. The introduction of spectral-domain OCT (SD-OCT) has

significantly improved the speed of data acquisition in OCT systems with scan rates of up to 25-40 kHz available in commercial OCT systems. Fast swept-source or spectrometer-based OCT systems can even extend the A-scan rates to several hundreds of kHz enabling the acquisition of high-speed real-time volumetric OCT data. Scan rates of up to 240 axial lines per second have been demonstrated using Fourier-domain-mode-locking (FDML) lasers [20, 21]. By utilizing the high scanning rate of these systems, a large field-of-view can be rapidly scanned in real-time, generating high-resolution OCT images. However, some of the potential problems associated with this brute force large field-of-view OCT imaging are the limited scan range of conventional OCT sample scanning mechanisms and the subsequent real-time interpretation of the massive amount of OCT data that would be generated. The focus of this thesis is to address some of these difficulties by proposing new techniques for OCT data acquisition and interpretation.

The organization of this chapter is as follows: Spectral-domain OCT is briefly explained in section 1.1, followed by an introduction to the current sample scanning mechanisms used for OCT imaging in section 1.2. An alternate way to represent OCT data in the form of non-speech audio signals is introduced in section 1.3. Finally, the outline of the rest of the thesis is described in section 1.4.

1.1 Spectral-domain OCT

Optical coherence tomography (OCT), first introduced in 1991, has found numerous applications within the domain of biomedical imaging, materials research, non-destructive testing and optical metrology [22, 23]. A block diagram of a typical spectral-domain OCT system is shown in Figure 1.1. Light from a broadband light source is split by a beam splitter into two arms, i.e. sample arm and a reference arm. The combination

of the light from the reference arm and back-scattered light from the internal tissue microstructures in the sample gives rise to an interference pattern. The interference pattern is spectrally decomposed by the grating and is focused onto an array of photo-detectors for detection. The Fourier transform of the spectrally decomposed interference pattern gives the depth information which constitutes a single A-scan in an OCT image. Two- or three-dimensional OCT images are formed by scanning the laser beam over the sample and appending the acquired A-scans in synchronization with the position of the beam.
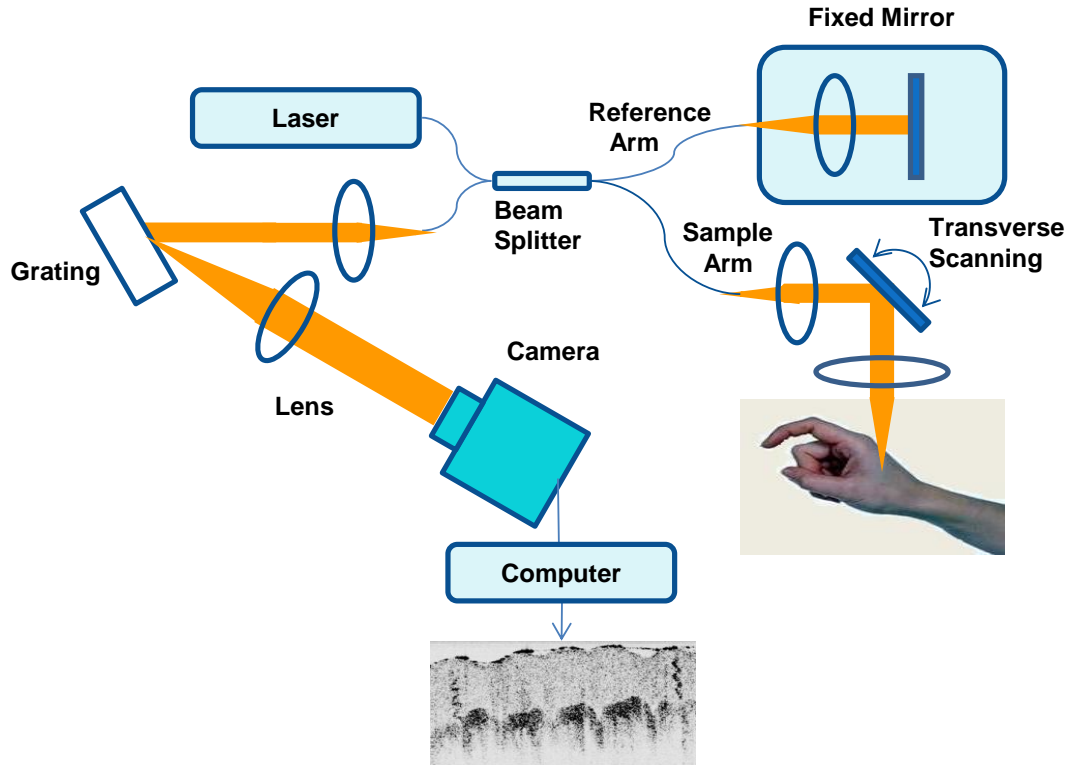

Figure 1.1: Spectral-domain optical coherence tomography (SD-OCT) system.

OCT relies on the intrinsic variations in the optical properties of tissues to generate contrast. In most of the currently used OCT systems, the differences in the spatial variation of the backscattering properties of various tissue microstructures

generate the contrast. However, any physical property that can change the amplitude, phase or polarization of the incident beam can be used to form images [24]. Various extrinsic contrast agents such as nanoparticles or microspheres have also been used to generate molecular specific contrast in OCT images, which has significantly increased the range of possible applications for OCT [21, 25].

In OCT, a broadband source having a low coherence length is used to obtain higher axial resolution and better contrast. The spectral range of the source is selected based on the penetration of light within the tissue which is dependent on both the absorption and the scattering properties of tissues. Near-infrared (700-1400 nm) light is used in the so-called 'biological window' where scattering dominates the light-tissue interaction. Light of wavelengths below 700 nm will be absorbed by melanin and hemoglobin, while wavelengths greater than 1400 nm will be strongly absorbed by the water within the tissues. In general, within this window, the longer the wavelength used, the greater the penetration depth due to the reduced scattering of light within the tissue [24].

In OCT the axial resolution and the transverse resolution are independent of each other. The axial resolution of an OCT system is determined by the full-width-half-maximum (FWHM) of the coherence length of the source. The coherence length is the width of the autocorrelation function $G(\tau)$ of the beam which is the related to the power spectrum $S(k)$ of the source by the Fourier transform relationship.

$$G(\tau) = \Im[S(k)] \tag{1.1}$$

where $\Im[\bullet]$ is the Fourier transform operator. For a source with a Gaussian spectral distribution, the coherence length l_c is given by

$$l_c = \left(\frac{2\ln(2)}{\pi} \right) \frac{\lambda_0^2}{\Delta\lambda} \qquad (1.2)$$

where λ_o is the center wavelength and $\Delta\lambda$ the bandwidth of the light source. Therefore a source with a larger bandwidth and shorter center wavelength will give a higher axial resolution in OCT. The transverse resolution in OCT depends upon the spot size of the OCT beam at the focus given by

$$2\omega_0 = \frac{4\lambda_0 f}{\pi D} \qquad (1.3)$$

where ω_o is the radius of the beam at focus, f is the focal length of the lens and D is the spot size diameter of the beam at the objective lens. Using a higher numerical aperture lens could enhance the transverse resolution but would decrease the depth of focus of the OCT beam. The typical axial and transverse resolution in OCT is around 1-10 μm and 10-20 μm respectively. The penetration depth in OCT is typically around 1-3 mm and is limited by both the multiple backscattering and the absorption of light within biological tissues. The maximum theoretical limit is determined by the following expression:

$$z_{max} = \frac{\lambda_0^2}{4n\Delta\lambda} N \qquad (1.4)$$

where N is the number of photo-detectors or number of pixels in a line scan CCD, and n is the average refractive index within the medium [12, 26].

1.2 Sample scanning mechanisms in OCT

Scanning mechanisms discussed in this thesis refer to various techniques utilized to scan a focused OCT beam over the sample surface for the purpose of OCT image formation. Interferometrically detected back-scattered light from the sample constitutes a single A-scan in OCT. A two- or three-dimensional OCT image is formed by the assembly of

sequentially acquired A-scans (a single A-scan is one column in an image) which need to be coordinated with the scanning of the beam over the sample.

Currently used OCT scanners can be classified as bench-top, hand-held or endoscopic scanners. In terms of the scanning geometry employed, they are categorized as circumferential, deflecting and translational scanning [27]. Most of these are predominantly based on mechanically scanning the beam and typically utilize galvanometers, small electric motors, piezoelectric actuators, etc. Based on the mechanisms used for beam scanning, the scanning mechanisms can be broadly classified as mechanical-scanning mechanisms and manual-scanning mechanisms.

1.2.1   Mechanical-scanning mechanisms

Mechanical-scanning mechanisms are the most widely employed scanning solutions for OCT. They have the advantage of great accuracy and the availability of a wide range of actuation mechanisms and devices. Mechanical-scanning mechanisms include techniques used in bench-top systems where the specimen is placed on a stage and image formation is done either by deflecting or translating the OCT beam over the specimen. Also included are circumferential and linear scanning techniques used in OCT based catheters or endoscopic devices.

1.2.1.1 Bench-top based mechanical-scanning mechanisms

Computer-controlled motion of galvanometer-mounted mirrors is widely used for scanning the specimen in bench-top systems. In this configuration, the tissue specimen is placed on a fixed stage and an OCT image is acquired by sequential acquisition of depth-resolved A-scans synchronized with the lateral scanning of the beam. However, the

limited angular range of the galvanometer (for a typical angular range of $\pm10^o$ the maximum scan range is approximately 2-3 cm) and the finite aperture of the objective lens (typically 1-2 cm) constrain the achievable lateral scan range. An alternative method to scan a wider area is to move the sample with uniform velocity with a motorized stage under a fixed OCT beam. This technique has the limitation of slow translation rate and inflexible scanning geometry. Although these stage-based scanning mechanisms provide excellent accuracy and are well suited for lab-based imaging of excised tissue samples, there is clearly a need for a more flexible scanning method for *in vivo* or intra-operative imaging applications.

1.2.1.2 Mechanical-scanning mechanisms in catheter-endoscopes and hand-held imaging probes

The limited penetration depth of OCT requires the incorporation of OCT beam delivery and scanning systems within various needle-based probes for imaging deep inside the tissue or OCT based catheters and endoscopes to allow intravascular imaging and imaging of hollow organs inside the human body. These devices are required to be of small size and diameter, and flexible enough to provide convenient access to tissues and organs.

A wide variety of mechanical-scanning mechanisms have been employed within these devices. In catheter-endoscopes the actuating mechanisms may be either at the distal or the proximal end of the probe. Some of the scanning mechanisms reported include micro-motors for circumferential scanning [28] and MEMS-based scanning techniques for circumferential and linear scanning [29, 30]. Motor-driven mirror mounted galvanometric techniques and piezoelectric actuators [31] have been used for beam deflection in catheters and hand-held probes. Incorporation of motor-driven linear

translation of the optical assembly inside a hand-held probe has been reported [32]. 3-D imaging by combining the MEMS motor rotational scan and linear stage transversal movement in catheters has also been demonstrated [33]. In forward-imaging devices, paired-angle-rotation scanning (PARS) has been reported which enables a linear scan pattern in front of the probe tip by using two counter-rotating-angle polished gradient-index (GRIN) lenses [34]. Although these scanning solutions provide good image quality and scan accuracy, the scan range is still limited and inflexible due to the limitations of the mechanical-scanning mechanisms. These mechanisms make the probe design complicated in addition to making the probes bulky and expensive. Mechanical-scanning mechanisms also frequently need to be customized for specific *in vivo* and intra-operative OCT imaging applications while still providing limited flexibility in choosing the scanning geometries.

1.2.2   Manual-scanning techniques

Due to the inherent limitations of mechanical-scanning mechanisms, many research groups have tried to devise alternate sample scanning techniques using hand-held probes. This is partly motivated by the availability of numerous position tracking devices and sensors for tracking surgical instruments in surgical-guidance applications. A simple hand-held, manual-scanning probe can be used to obtain OCT images of tissues and organs which might otherwise be inaccessible using standard mechanical-scanning probes. However, manually scanning a hand-held probe can cause a number of image artifacts due to variations in the scan velocity and orientation of the probe. Consequently, image formation with a manual-scanning probe requires a method to synchronize the acquired A-scans with the relative displacement between the sample and probe.

Based on the methodology to obtain positional information, the manual-scanning techniques have been divided into sensor-based and sensor-less approaches where the sensor-based technique requires an add-on sensor on the probe while the sensor-less approach utilizes the acquired data and properties of the imaging system to deduce positional information.

1.2.2.1 Sensor-based manual-scanning probes

As image-guided surgical procedures are becoming more popular, numerous companies are developing devices and sensors for position tracking of surgical instruments. A number of commercial products are available in the market. These systems attach reference markers to the probe, and sensors based on acoustic, electromagnetic or optical principles are commonly employed to spatially localize the instrument or device within the sensor's field-of-view [35]. To gain information of the position and the orientation of the probes, in general six degrees of freedom (DOF), i.e. spatial (x, y, z) co-ordinates and the rotational (raw, pith and yaw) co-ordinates, are required.

In acoustic sensing, a sound emitting source is placed on to the device to be tracked, and a receiver is used to detect the emitted sounds. The distance is determined by either measuring the propagation time of the received sound wave or the phase-difference between the sent and received signal. Acoustical position trackers are not widely used as the variations in temperature, humidity and pressure of air significantly influence the propagation speed of sound and limit the accuracy of these sensors. Moreover, these sensors require a line of sight between the source and the detectors.

Electromagnetic tracking can be based on either alternating current (AC) or direct current (DC) fields. An electromagnetic receiver is mounted on a probe or the instrument

to be tracked and placed within the AC or DC fields; the electric currents induced due to the variation of the position of the probe are then used to estimate the spatial location of the probe. Ferromagnetism and eddy currents, however, affect the accuracy of these devices. Moreover, magnetic-field-based sensing systems are highly susceptible to electromagnetic interference which limits their utility in the operating room environment [36].

Optical tracking systems can in general be divided into two main classes: passive tracking and active tracking. Active tracking systems are based on mounting optical emitters (typically IR LEDs) to the probe and tracking them using a camera or any other suitable sensor. These systems have the drawback that the optical emitters will require a power source of their own which will either have to be mounted on to the probe itself or supplied through wires. This makes the probe bulky and its use inconvenient. Passive tracking systems have also been reported, where instead of using light emitters, retro-reflecting spheres or specific geometric patterns are attached to the probe and are subsequently recognized by optical sensing systems. Computation algorithms are then used to estimate the position and pose of specific markers [37]. At least three sensors are required to get the full position and orientation of the probe. Optical tracking systems are currently the most widely used systems. Optical tracking systems, like acoustic sensors, require a line of sight for operation; however, in general they are more accurate with studies showing that these systems have an accuracy and precision of around 0.5 mm [38].

Other devices based on accelerometers and gyro-meters have also been used for tracking. Using a combination of technologies such as magnetic and optical sensing has

also been reported. However, all sensor-based devices have to be carefully calibrated, typically have sub-millimeter spatial resolution, and the operating distances need to be within the range of the mounted sensor and the base unit.

A distinction must be made between tracking surgical instruments inside the body and outside it. Most of the above mentioned techniques require that the instrument be placed outside the human body. In many cases it is very useful to track the position of instruments such as catheters, needles or endoscopes within the human body. Most often an imaging technique such as ultrasound, fluoroscopy or MRI is used for this purpose. As optical or acoustic sensing requires a line of sight, the only viable sensor-based tracking technology that can be used inside the human body is based on electromagnetic sensing. Many techniques have been reported which are based on placing small electromagnetic sensors at the tip of needles or catheters for real-time positional information inside the human body [39]. Typical accuracies reported for these systems are around sub-millimeter level [40]. Some companies have reported sensor sizes of less than 1 mm in diameter that can fit into needles as small as 16 gauge [41, 42]. The only reported work for image formation with a sensor-based manually scanned probe in OCT is based on passive optical positional tracking with an accuracy of 6 μm along two axes and 19 μm along the third axis [43].

1.2.2.2 Sensor-less freehand scanning

Most of the sensor-based technologies described above are utilized for tracking surgical instruments in image-guided procedures and not necessarily for image formation. Some of these have been used for image formation in 3D ultrasound. However, OCT, being a high resolution imaging modality where the typical transverse resolution is in the range

of 10-20 μm, would require much more accuracy than what is currently commercially available with these sensor based techniques. Hence, there is a need for sensor-less based tracking. Position tracking without the use of an external position sensor can offer significant advantages. It would not only make the probe design simple, but give a more flexible scanning geometry and remove many of the constraints imposed by the position sensor. A sensor-less approach will require the positional information to be deduced by the acquired data or images. In this section we describe some motion estimation methods used in other imaging modalities.

Sensor-less hand-held scanning has been studied extensively in ultrasound imaging. Some of the most common techniques used are based on speckle decorrelation. Speckle is a common phenomenon in coherent imaging systems, which arises from the random interference from the back-scattered light from objects that are within the resolution volume of the imaging system. Speckle depends upon the distribution of the scatterers within the sample, the system parameters and the beam characteristics. Speckle decorrelation has been used in ultrasound for velocity estimation, elastography and 3D image formation with varying degree of success [44, 45]. Theoretically, these techniques only work for a fully developed speckle pattern. Most of the studies have come to the conclusion that speckle information alone would not be sufficient for accurate motion estimation as the speckle size and distribution are influenced by a number of parameters.

Previous work in OCT has shown that OCT data can be used to obtain information about the displacements and velocity of the scatterers within the sample. For example, in optical coherence elastography (OCE) the decorrelation in the speckle pattern has been utilized to measure the stiffness of tissues [46] while in optical Doppler

13

tomography (ODT) the velocity of moving structures is obtained by utilizing the change in phase or frequency of the back-scattered signals. Techniques based on the cross-correlation have been used for example in motion artifact correction [47] and phase stabilization [48] in OCT. However, no prior work has been done in image assembly of OCT images using the cross-correlation techniques.

1.3 Multi-sensory data representation

In most real-world situations, we rely on our multi-sensory input capabilities to obtain information about our immediate surroundings. The sets of information obtained from our different senses complement each other. It therefore makes sense to use the same multi-sensory capability to interpret the increasingly vast amounts of scientific data being generated with the availability of inexpensive hardware resources and computational power. A multi-sensory input may be valuable in those scenarios where our visual sense is busy with other tasks or where we require rapid data interpretation.

An important component of an image-guided system is the way data is represented and conveyed to the user in the surgical field. Stereoscopic, virtual-reality and augmented-reality techniques have been used in surgical guidance applications [1]. In OCT, the traditional way is to represent the data in image form using a visual display. The tissue structure, morphology, and beam attenuation are encoded in the intensities of the back-scattered light which constitutes a single A-scan in OCT. These A-scans are assembled together to form a B-mode image and then displayed on the screen.

Often it is desirable to image over large fields-of-view in real-time such as for screening or surgical guidance. Given the high resolution capabilities of OCT and the desire to image over a large field-of-view, high data acquisition rates are required, which

would make the real-time interpretation of OCT data a challenge. In these scenarios a multi-sensory representation of OCT data may facilitate rapid interpretation.

1.3.1   Sonification

The process of converting data into non-speech audio signals or waveforms for the purpose of conveying information about the data is known as sonification. Sonification of scientific data can be a valuable extension to the traditional visual display of data. Some of the inherent advantages of audio representation of data include:

- The complementary nature of sound data

In many scenarios the information we obtain from our audio sense complements our visual sensory information.

- Superior temporal resolution of the human auditory system

The human auditory system has far superior temporal resolution compared to the human visual system. Auditory representation can make the perception of patterns and sudden changes in the data more easily noticeable.

- Ability to monitor parallel streams

The human auditory system has the ability to monitor many parallel streams together. These properties can be used to represent multi-modal and multi-dimensional data.

- Faster processing of data

Audio rendering of data is in general computationally much simpler than rendering 3D visualizations of the same data.

- Localization of sound

The ability of the human auditory system to localize sound has been used for warning alarms.

Auditory information has been utilized in a number of different instruments or devices, such as in Geiger counters, electrophysiological recordings, warning alarms, and representation of multi-dimensional and multi-modal data [49, 50]. Audio rendering of visual scenes and audio cues have been used as a navigational aid by the visually disabled. Sounds have been extensively used in human-computer interfaces (HCI) and as an extension to visualization of complex data.

1.3.2   Sonification of biomedical data

Over the years, the complexity of biomedical data has increased tremendously. This calls for increasingly complex analysis and novel representation methods. Sophisticated methods to provide intra-operative information to the surgical team in real-time have been developed such as three-dimensional visualization, and virtual- and augmented-reality systems.

However, many of these systems have surprisingly made little use of the capabilities of the human auditory perception. Physicians have relied on auditory information for decades to diagnose illness or listen to body sounds as is evident from the ubiquitous use of the stethoscope in medical practice. Audio representation is well suited for scenarios in which data is being acquired continuously and the user is only interested in abnormal or aberrant values. In biomedical applications, sonification has been used for providing audio feedback for manual positioning of surgical instruments [51], surgical navigational systems, analysis of EEG signals from the brain [52], physiological monitoring, heart rate variability [53] and interpretation of image data [54] and texture [55]. Audio output has also been utilized in Doppler ultrasound [56] and Doppler OCT [57] .

1.3.3 Sonification of OCT data

Auditory representation of OCT data may be more beneficial than the conventional visual display in many situations. This is especially true in the highly demanding operating room environment where the surgeon has to simultaneously keep track of a number of parameters. The addition of an audio channel can free the visual sense for other tasks. It is known that human auditory perception is very sensitive to slight changes in the temporal characteristics of sound and can detect even small changes in the frequency of a signal [58]. These properties can be exploited to provide a faster method of tissue classification and identification of morphological landmarks in time-sensitive image-guided surgical procedures such as screening, tumor resection or needle biopsy, and may complement the visual representation of OCT data. Sonification may also find applications where non-image data is collected such as procedures which use forward sensing devices where only axial scan data is acquired [59, 60].

1.4 Thesis outline

The main contribution of this thesis is the development of algorithmic techniques which would potentially make the adaptation of OCT imaging for intra-operative applications more feasible. An image acquisition technique for OCT is proposed that can enable manual-scanning of the sample at an extended scan range with flexible scan geometries. Techniques for sonification of OCT data are also described which when combined with the visual display of OCT data, may enable the utilization of our multi-sensory capabilities for rapid interpretation of OCT data.

The organization of the remaining chapters of the thesis is as follows. Chapter 1 is followed by the description of the cross-correlation based image acquisition technique in

Chapter 2. Chapter 2 also includes image assembly results from different tissue samples and biological tissues by using a sensor-less manual-scanning mechanism. Audio representation of OCT images and data is described in Chapter 3, which is followed by conclusions and future work recommendation in the final chapter.

CHAPTER 2   CROSS-CORRELATION BASED IMAGE ASSEMBLY

This chapter describes algorithms for image and data reconstruction from sequentially acquired A-scans while manually scanning an OCT beam over the sample. Some of the work described in this chapter has been reported in a recent publication [61]. The algorithm is based on estimating the movement of the probe (or sample) by utilizing the cross-correlation of the consecutively acquired A-scans. The proposed method can not only provide a simpler and less expensive scanning solution with an extended field-of-view and greater penetration depth inside the tissue, but may also allow greater flexibility and freedom of movement while acquiring OCT images. The details of the cross-correlation based algorithm for image reconstruction are explained, followed by simulation results for the longitudinal A-scans assembly and experimental results for the lateral manual-scanning case.

2.1   Cross-correlation based motion estimation approach

Figure 2.1 shows the displacement of the OCT beam along the transverse (lateral) direction and longitudinal (axial) direction. The sample volume of the object probed by the beam will constitute a single A-scan. OCT signal *s(r)* can be modeled as the convolution of the point spread function of the system $g(\bullet)$ with a certain scatterer distribution function $\eta(\bullet)$

$$s(r) = \int_{-\infty}^{\infty} \eta(r')g(r-r')dr' \qquad (2.1)$$

where *r* is a position vector denoting spatial co-ordinates *(x,y,z)*.
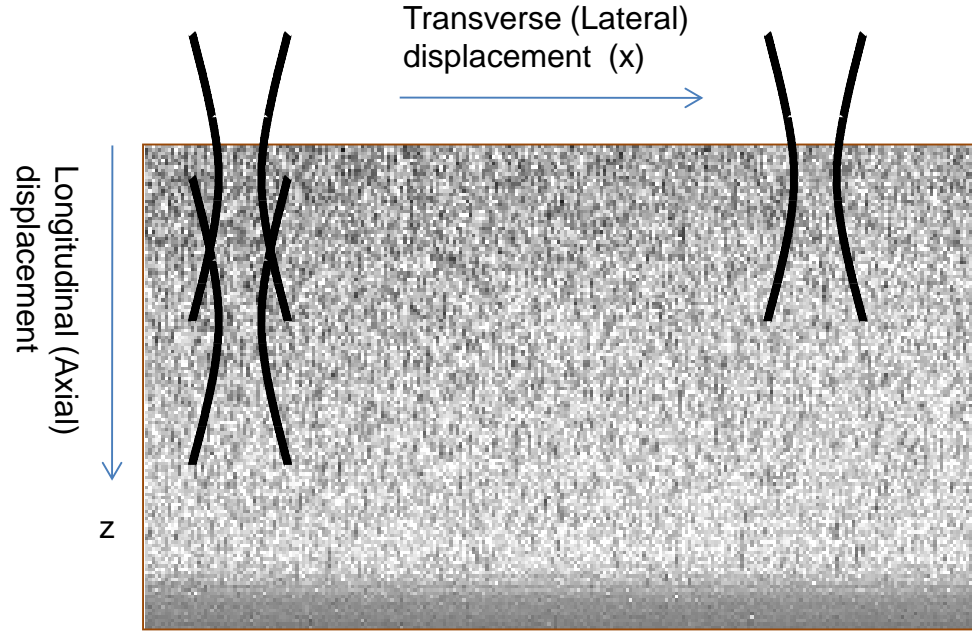
Figure 2.1: OCT beam displacements along the longitudinal and lateral directions.

The cross-correlation $p(r_1,r_2)$ between the two signals collected at positions r_1 and r_2 can be expressed as

$$\rho(r_1,r_2) = \ <s(r_1)s^*(r_2)> \ = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} <\eta(r')\eta^*(r''-\Delta r)> g(r_1-r')g^*(r_2-r'')dr'dr'' \qquad (2.2)$$

where $< \bullet >$ is the ensemble average and $\Delta r = (r_2 - r_1)$ is the displacement of the probe between the two measurements. This indicates that the cross-correlation function will depend not only on the point spread function but also on the beam displacement Δr, and the object properties. Thus, estimation of the displacement Δr will be challenging unless some simplifying assumptions are made or something is known *a priori* about the object properties.

One simplification often made is that the object is assumed to have a homogeneous distribution of scatterers and uncorrelated microscopic structures [62]. In this case we can express $<\eta(r')\eta^*(r''-\Delta r)>$ as a delta function. If c is the average scattering strength, then Equation (2.2) can now be written as

20

$$\rho(r_1, r_2) = c^2 \int_{-\infty}^{\infty} g(r_1 - r')g(r_1 + \Delta r - r')dr' \tag{2.3}$$

which shows that for a homogenous distribution of scatterers the cross-correlation function is equivalent to the autocorrelation of the point spread function. This assumption may be true for certain tissue phantoms but would rarely be true for biological tissues, hence causing a bias in the displacement estimation.

However, if we sufficiently oversample, the cross-correlation based approach may be used for image-assembly. An OCT image is a sequential assembly of uniformly spaced A-scans. If the motion of the probe is constrained below a certain threshold determined by the A-scan acquisition rate, then consecutive A-scans within one resolution volume will have high cross-correlation due to the regions of overlap, depending on the amount of oversampling of the sample. Non-uniform movement of the probe will cause non-uniform sampling of the sample which, in turn, causes variability in the cross-correlation between adjacent A-scans. Slower scan velocities will result in sequential A-scans with higher correlation, while faster scan velocities will result in reduced correlation between successive A-scans.

The cross-correlation based approach is used for image reconstruction by utilizing the correlation information. For an extended A-scan image assembly along the longitudinal direction, the lag of the cross-correlation function has been used to find the axial displacement of the probe, while for lateral manual-scanning the cross-correlation coefficient (which corresponds to a lag of zero) has been used to estimate the lateral displacement. The details of the algorithm, the constraints imposed on image acquisition and the limitations of this technique are described in the subsequent sections.

2.2 Longitudinal A-scan assembly algorithm

Forward-imaging devices have some advantages compared to the side-imaging needles as they can collect data before the device is introduced into the tissue, which can be helpful in many clinical procedures [63]. Long A-scan assembly along the longitudinal direction using these devices will enable the acquisition of OCT data over several millimeters inside the tissue much beyond the penetration depth of OCT beam which is typically around 1-3 mm.

The proposed method shown in Figure 2.2 can be used to assemble a long A-scan along the longitudinal direction. A series of A-scans are acquired at different depths using an OCT forward-imaging needle. A-scans are then co-registered and synchronized with the location of the needle without utilizing any external position sensors. The velocity of
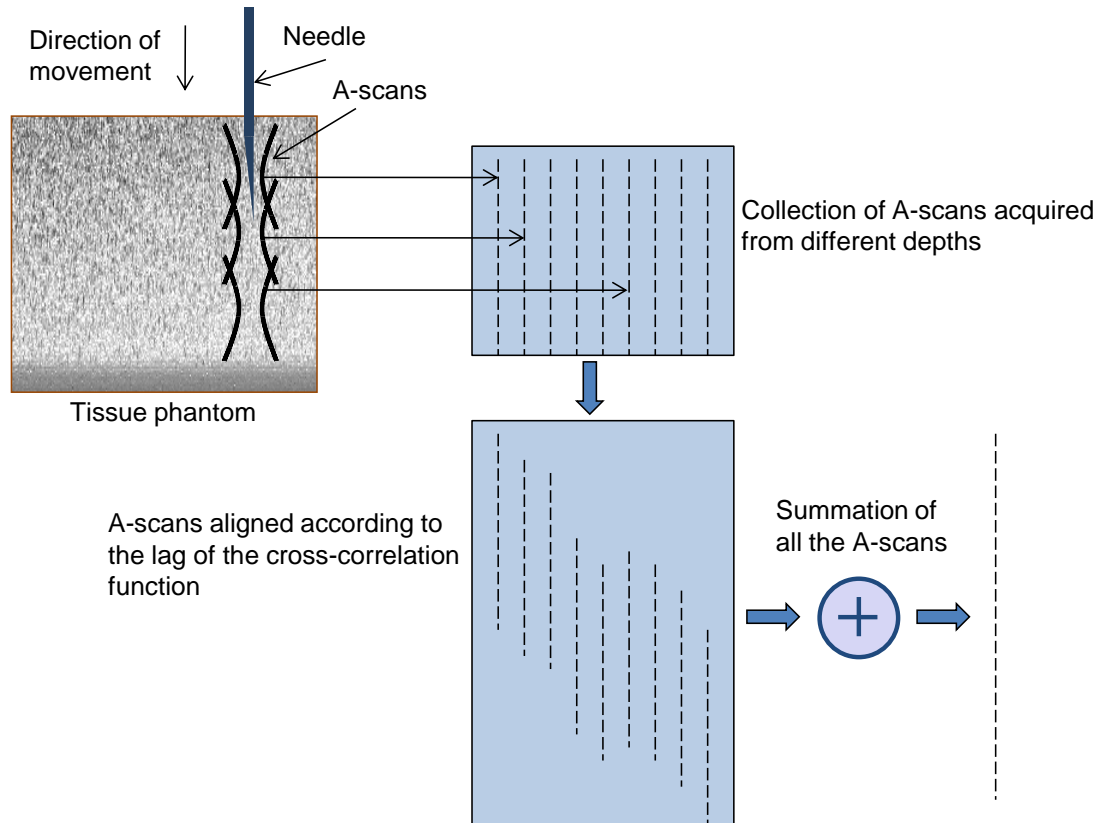

Figure 2.2: Methodology for longitudinal A-scans assembly.

the needle is constrained below a certain threshold (depending upon the A-scan acquisition rate) such that consecutive A-scans have considerable regions of overlap and hence high degrees of correlation between them. The maximum allowable velocity v can be determined by

$$\zeta = \frac{f_s \Delta z}{v}$$ (2.4)

for a given A-scan acquisition rate f_s, axial resolution Δz, and sampling factor ζ. A value of $\zeta < 2$ indicates undersampling and $\zeta > 2$ oversampling while sampling at the Nyquist rate occurs when $\zeta = 2$.

If sufficient oversampling of the object takes place, then the information contained in the consecutive correlated A-scans may be used to deduce the amount of movement of the needle and hence, reconstruct a long A-scan over several millimeters to centimeters of depth. In general the amount of correlation will depend upon the beam parameters and the object properties. The A-scans are aligned together depending on the estimated lag determined by the cross-correlation function and are then finally summed together. However, prior to adding the A-scans together to get an extended A-scan, the attenuation of the beam and the distortion due to the beam shape must be compensated for each A-scan.

2.2.1   Simulations for longitudinal A-scan assembly

Synthetic OCT data was obtained by convolving the point spread function of the system with a certain random point scatterer distribution function and propagating the beam back to the detector [64]. To simulate the attenuation of the beam as it penetrates inside the tissue, each A-scan was multiplied by an exponentially decaying function. Finally some additive white Gaussian noise (AWGN) was added to each A-scan. To simulate the

forward scanning of the probe, the beam was displaced longitudinally by a random number which was generated subject to the condition that the A-scans are oversampled. Following this procedure, a series of A-scans were acquired which had overlapping regions along the axial (longitudinal) direction due to oversampling.

The cross-correlation values at a lag z_o can be written as

$$E\,[\Im\{S(k,z_1)\}.\Im\{S^*(k,z_1+z_o)\}] \tag{2.5}$$

where $S(k,z)$ is the complex signal obtained at longitudinal position z, $E\,[\bullet]$ is the expected value, $\Im\,[\bullet]$ is the Fourier transform and k is the wave number. The displacement of the probe $\hat{z}_o$ is found by evaluating the cross-correlation at each lag and finding the maxima of the cross-correlation function by the expression

$$\hat{z}_o = \arg\max_{z_o} E\,[\Im\{S(k,z_1)\}.\Im\{S^*(k,z_1+z_o)\}] \tag{2.6}$$

This problem is similar to the time delay estimation techniques studied extensively in signal processing and communication theory [65].

In an ideal case, the cross-correlation function (CCF) will have one well defined peak corresponding to the displacement of the beam. However, there may be several local peaks due to the structure of the object and at times the maxima of the CCF may not necessarily correspond to the actual displacement of the beam. Based on the lag values obtained, all the A-scans are aligned with each other and then summed together to get the extended longitudinal A-scan.

Figure 2.3 shows sample A-scans and the cross-correlation function (CCF). The A-scan in Figure 2.3(b) has been displaced and the corresponding CCF has a well defined sharp peak where the lag value at this peak will be an estimate of the displacement of the two A-scans.
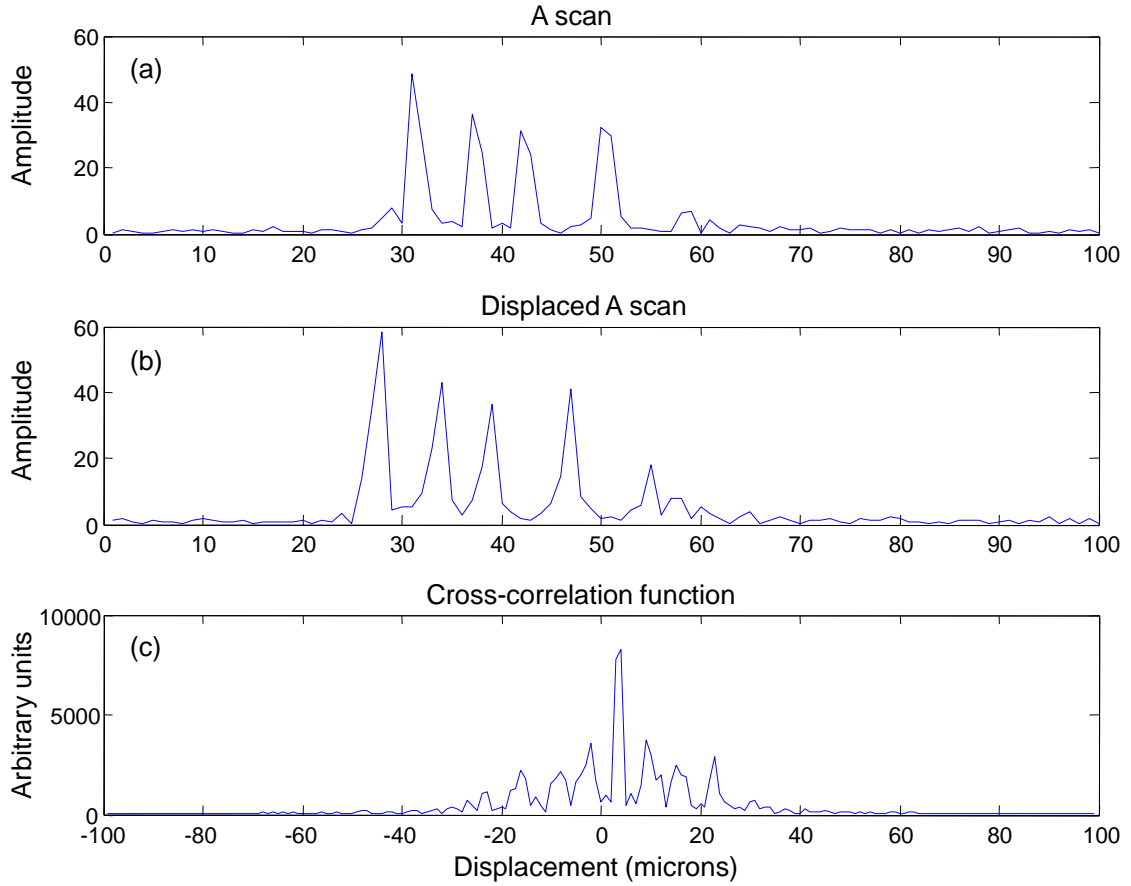
Figure 2.3: Cross-correlation function between two A-scans. (a) Simulated A-scan data. (b) A-scan data collected after shifting the beam position along the depth. (c) Cross-correlation function where the peak corresponds to the displacement between the two A-scans.

However, there may be cases when the actual displacement of the beam would not correspond to the maxima of the CCF. To investigate this effect, an object with scatterers with periodic spacing of distance 9 μm was simulated. Figure 2.4 shows two A-scans that are displaced from each other at a distance of 1 μm. However, the CCF has a maxima at a lag value of -8 μm because the scatterer at position 10 is now at the focus and has a larger back-scattered signal in the displaced A-scan.
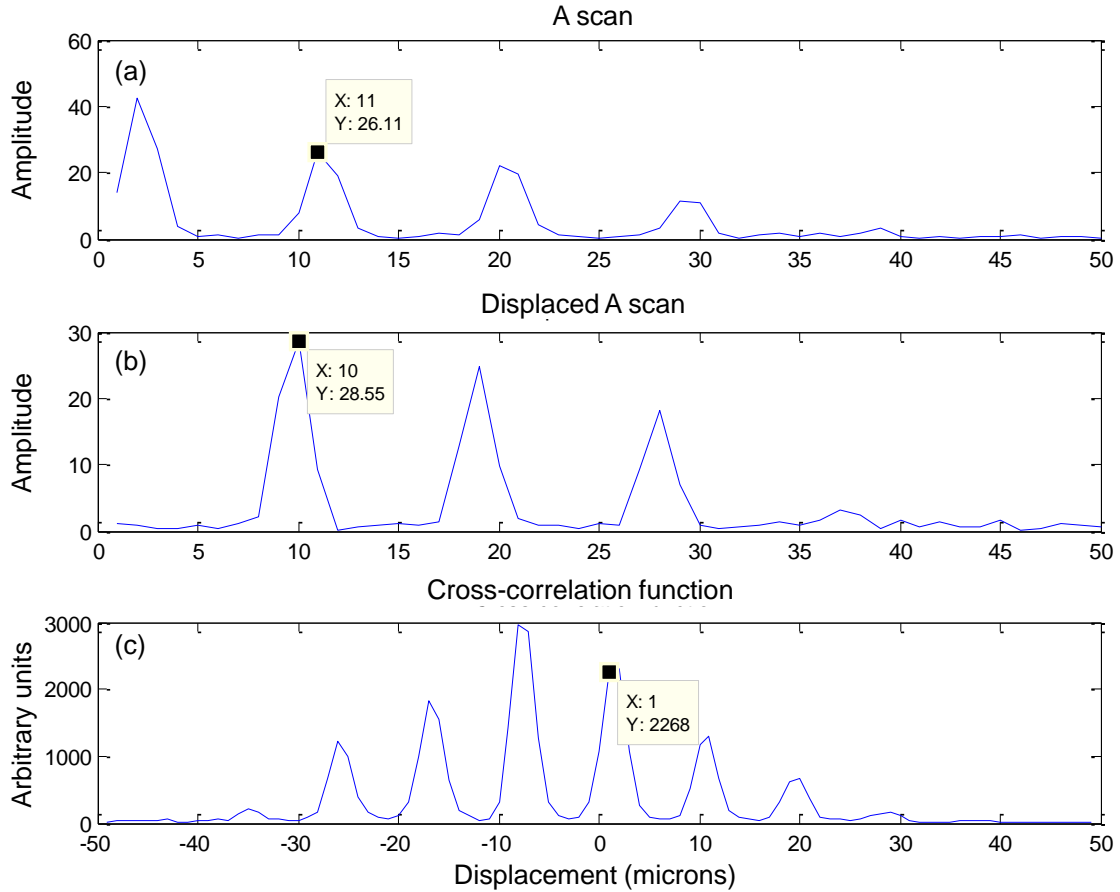
Figure 2.4: Cross-correlation function and the estimation of the lag.

One simple way to remedy this situation is to search for the maxima of CCF in a local region. The criteria of the selection of this 'lag search' region may require some *a priori* knowledge of the object structure. If the mean spacing of the scatterers or the amount of oversampling in the acquired dataset is known, the search of the lag can be limited to within these limits. Assuming than an object has a periodic scatterer distribution of *d* μm, the successful application of this technique would require the displacement of the probe to be constrained less than *d* μm and the search for the peak must be limited to within lag values corresponding to *d* μm. Practically this requirement will put a constraint on the maximum displacement of the probe allowed between the collections of two consecutive A-scans. However, with high speed OCT systems

available today this may well be within the limits of normal human hand movement. For example, at 25 kHz scan rate and 4 µm axial resolution, Equation 2.4 requires the maximum velocity of hand movement to be less than 1 cm/s to achieve a sampling factor of 10.

Figure 2.5(a) shows the simulated lags and the estimated lags for this dataset. It can clearly be seen that the estimated lags have significant errors which are primarily due to the periodicity of the scattering structure. However, by utilizing *a priori* knowledge of the object structure and limiting the search to only 10 positive lag values (the movement of the probe in simulations was constrained up to 9 µm), significantly more accurate lag estimations were obtained, as shown in Figure 2.5(b).
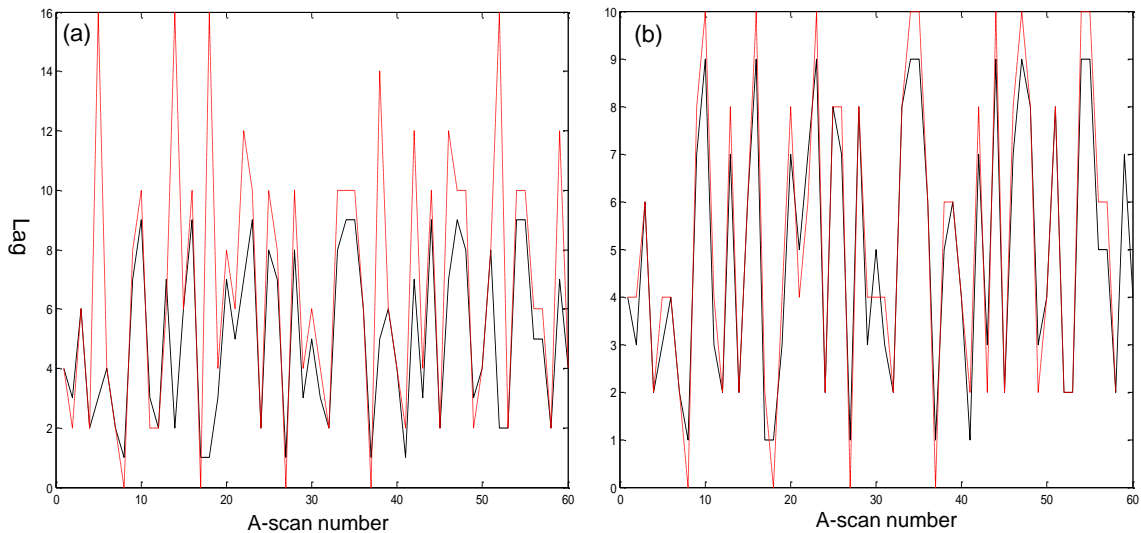

Figure 2.5: Simulated and estimated lag values based on the cross-correlation function peak. (a) Without limiting the search of the peak in the CCF. (b) Constraining the search of the peak to 10 lag values. Solid black lines are the actual lag values while the dashed red line is the estimation of the lag.

Figure 2.6 shows results with a random distribution of scatterers. The scatterers were randomly distributed between 5 and 20 µm as shown in Figure 2.6(a). The beam was randomly moved with a displacement between 1 and 9 µm. The collected A-scans

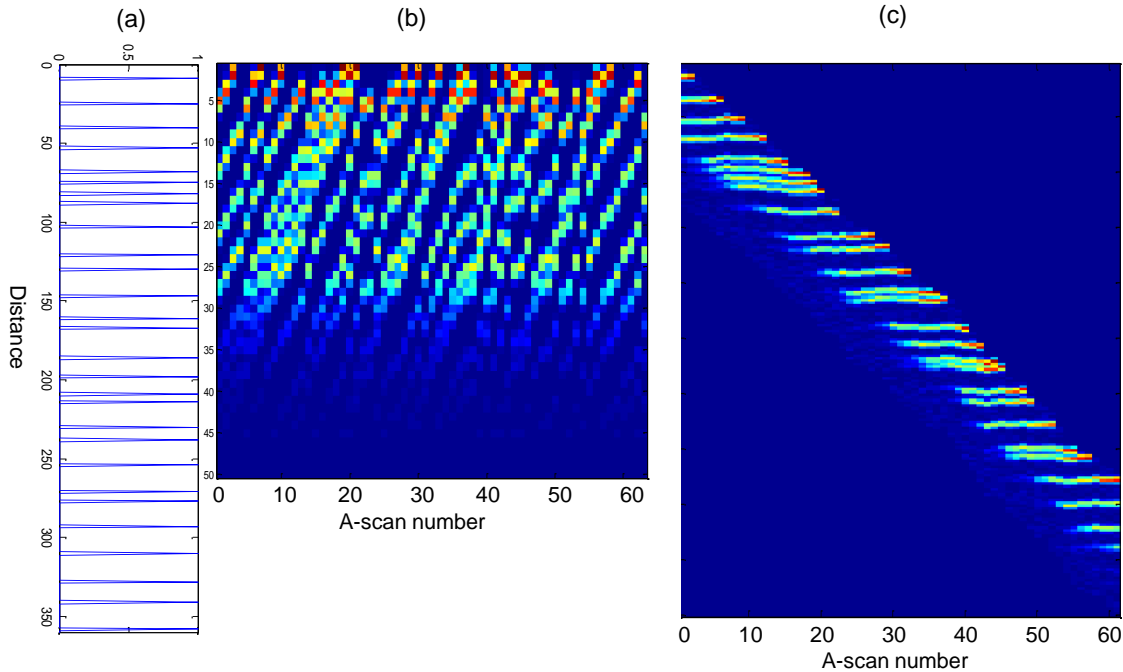are shown in Figure 2.6(b) while Figure 2.6(c) shows the A-scans after aligning them with each other.

Figure 2.6: (a) Random distribution of the scatterers for a long A-scan. (b) The A-scans appended together without positional information of the beam. (c) The alignment of the A-scans based on the peak of the cross-correlation function.

The actual and the estimated lags are shown in Figure 2.7(d). In this case the peaks were found by constraining the search for the maximum value of the CCF to positive lags values (because the probe was simulated to move in only one direction). Reconstructions based on the lag values are shown in Figure 2.7(c) where it is noticeable that the peaks and the peak spacing correspond to the original scatterer distribution reasonably well.
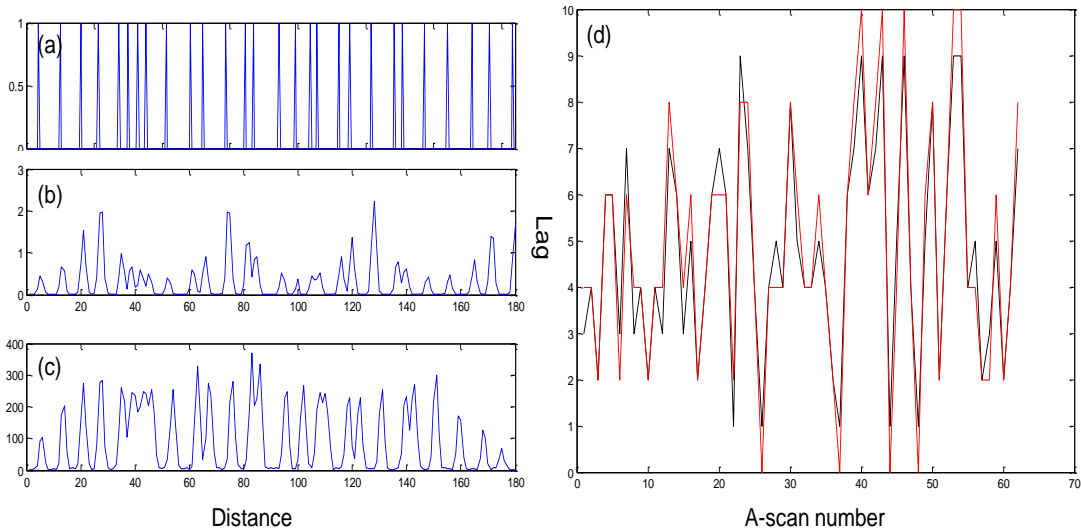
Figure 2.7: (a) Original distribution of scatterers. (b) Scatterers after simulating with the OCT forward model. (c) The assembled A-scan based on the cross-correlation function. (d) The actual (solid black line) and the estimated (red dashed lines) lag values.

2.2.2   Limitations

The algorithm presented for longitudinal A-scans assembly has not been verified experimentally and therefore is rather limited in scope. This technique will work best in probes which have a low numerical aperture and a large Rayleigh range so that consecutive A-scans have a large overlapping area. However, this method will fail in those scenarios where the cross-correlation function does not have a well defined peak. This will occur for example in objects that have very similar features along the depth. If sufficient oversampling of the sample takes place and the velocity of the probe is constrained, then the performance can be improved by constraining the search of the maxima of the CCF to small lag values. Other methods for finding the peak of the CCF such as the Maximum Likelihood or generalized cross-correlation functions may also be tried. However, these results and conclusions have been drawn by simulations only and they need to be verified by acquiring actual measurements from a forward-imaging needle.

2.3 Lateral manual-scanning image assembly algorithm

In this section, a cross-correlation based image assembly algorithm for lateral manual-scanning is described. The image assembly is done based on the degree of correlation between laterally acquired A-scans which can be measured by the Pearson cross-correlation coefficient given by

$$\rho(i, j) = \frac{< (I_i - \mu_i)(I_j - \mu_j) >}{\sigma_i \sigma_j}$$

(2.7)

where $<\bullet>$ is the expected value, I_i and I_j are the intensities of the sequential A-scans, and μ_i, μ_j and σ_i, σ_j are the means and standard deviations, respectively, of the corresponding i^{th} and j^{th} A-scans. Identical A-scans would correspond to perfect correlation (ρ=1) whereas highly uncorrelated A-scans exhibit zero or no correlation (ρ=0). The cross-correlation between adjacent scans will depend not only upon the sample structure, but also on the sampling factor, speckle pattern [66], and the signal-to-noise ratio of the images.

The block diagram of the algorithm used for image assembly is shown in Figure 2.8(a). The first A-scan is selected as the reference and the cross-correlation coefficients with the subsequently acquired A-scans are computed. By selecting an appropriate threshold based on the sample properties, all A-scans that fall within the resolution volume can be discarded (as they contain the same information) and only uniformly spaced A-scans are used for image assembly. When the cross-correlation coefficient falls below the selected threshold, the displacement is deemed to satisfy the desired sampling criteria, and that particular A-scan is appended to the assembled image. This assembled A-scan is now selected as the new reference and the steps are repeated until the algorithm iterates through all the acquired A-scans. The goal of the algorithm presented here is to

discard all oversampled regions of a manually scanned image, essentially reconstructing the OCT image by assembling A-scans which are equally sampled in distance rather than equally sampled in time.
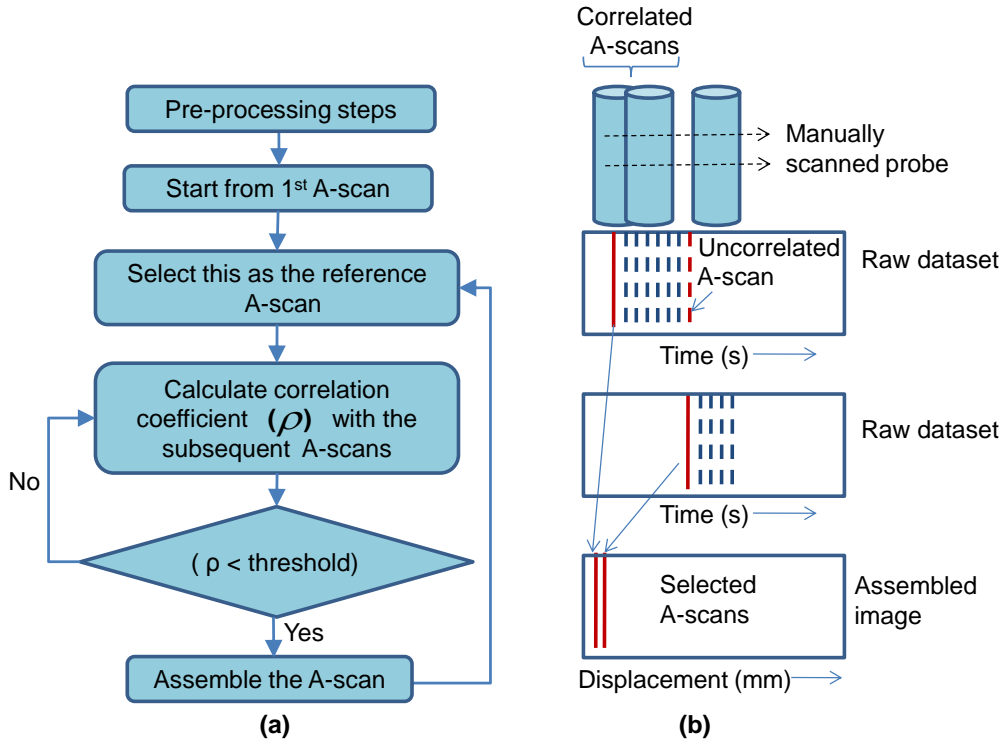

Figure 2.8: (a) Flow chart representation of the algorithm. (b) The cross-correlation between A-scans decreases with the lateral displacement of the beam. The raw dataset contains A-scans uniformly placed in time, but due to non-uniform manual-scanning, the successive A-scans have non-uniform displacement. The assembled image consists of A-scans selected by the algorithm which are uniformly spaced in distance.

2.3.1   Pre-processing steps

Several pre-processing steps were performed prior to computing the cross-correlation coefficients. Noise contributions were minimized by truncating each A-scan so that only the portion containing sample information was selected. To make the algorithm more robust to variations in the sample structure and to increase the dependency on the speckle pattern from the sample, the output from a two-dimensional moving-average (MA) filter was subtracted from the raw image. The size of the MA filter should be of the order of several resolution elements (both in axial and lateral direction). Whereas a lower value of

31

MA filter size will result in loss of useful speckle information, a higher value will make the decorrelation curves less sensitive to the slowly varying sample structure and attenuated signal in the axial direction. The size of the filter along the axial dimension was independent of the lateral scan velocity and was chosen to be around 5-6 times the axial resolution elements. While it is critical to choose the optimum filter size in the axial direction to ensure robustness of the algorithm, the requirement of the choice of filter size in the lateral direction was relatively relaxed and depended on the average scan velocity.

2.3.2   Sampling criteria

An inherent assumption of this technique is the requirement that the acquired data be sufficiently oversampled. Oversampling in this context means sampling more than twice within the transverse resolution of the OCT system, which depends upon both the transverse resolution of the OCT system and the lateral step size. Due to the high A-scan rates available with current systems, and to fully reconstruct the features of the sample, OCT images are usually oversampled. The sampling factor for a manually scanned system along the lateral direction is defined similarly to Equation (2.4), i.e. $\zeta = \frac{f_s \Delta x}{v}$, where Δx is the transverse resolution of the OCT system which is equal to the diameter of the beam ($1/e^2$ intensity) at the focus in the sample arm, f_s is the A-scan acquisition rate (Hz) and v is the velocity of the moving sample or probe. Similar to the longitudinal scanning case, a value of $\zeta < 2$ indicates undersampling and $\zeta > 2$ oversampling, and sampling at the Nyquist rate occurs when $\zeta = 2$. This equation can be used to calculate the maximum velocity with which the probe or sample can move relative to each other by solving for v for a sampling factor of $\zeta = 2$.

2.3.3   Threshold selection

Selection of an appropriate threshold value is essential for the proper working of this technique. A decorrelation curve plotted for a sample depicts the decrease in the correlation coefficient value as a function of lateral displacement between two A-scans. Based on the decorrelation curve of a sample, a threshold can be determined, corresponding to the desired sampling factor for the assembled image.

2.4   Image assembly results for lateral sensor-less manual-scanning

In this section, experimental results showing the decorrelation curves and images assembled from laterally manually scanned tissue phantoms and biological tissues are shown. In order to test the validity of the hypothesis that consecutive A-scans will be decorrelated outside the OCT resolution volume, decorrelation curves were plotted for different tissue samples. The ability of the technique to differentiate between different scan velocities was then investigated by moving the sample with uniform scan velocities and observing the decorrelation curves. Subsequently, experiments were carried out for image assembly using both tissue phantoms and biological tissues. The human tissue used in this study was acquired and handled under a protocol approved by the Institutional Review Boards at the University of Illinois at Urbana-Champaign and Carle Foundation Hospital (Urbana, IL).

2.4.1   Experimental setup

A spectral-domain OCT system was used to perform the experiments. A Ti-sapphire laser with 800 nm center wavelength and 90 nm bandwidth was used, providing an axial resolution of 5 μm. The power in the sample arm was 10 mW and the samples were

imaged with a 40 mm lens producing a transverse resolution of approximately 16 μm. The experiments for manual-scanning were conducted at a line scan rate of 1 kHz and an exposure time of 200 μs for the line scan camera. The sensitivity of the system at 1 kHz was measured to be 96 dB. The relatively low scan rate was chosen to allow sufficient time for manually translating the sample under the fixed OCT beam.

The experiments were conducted by mounting the sample onto a manually movable stage and the position of the beam was kept fixed. The computer-controlled translational stage axes were aligned with the axes of a manually movable spring loaded translational stage in order to obtain OCT images of the same cross-sectional planes within a sample while employing two different scanning mechanisms.

2.4.2   Results

2.4.2.1 Decorrelation curves

The variation in the correlation coefficient between two A-scans as a function of lateral displacement can be shown by the decorrelation curves. Figure 2.9 shows average decorrelation curves for several tissue phantom samples and biological tissues. The cross-correlation coefficients were obtained by an ensemble average of 400 A-scans at each lateral displacement. A-scans that are within a resolution volume are expected to be highly correlated while those outside of the resolution volume of the reference A-scan are expected to have little correlation. The decorrelation length, which is measured as the decrease in the cross-correlation coefficient to *1/e* of its maximum value, would be approximately equal to the lateral resolution of the system, which governs the lateral speckle size in OCT images of scattering tissues [66]. The results, however, show that there exists some degree of variability in the coefficients at each lateral position and this

variability increases with an increase in the lateral displacement. This variability may be due to a number of factors which include noise, speckle and image features. The decorrelation length may be higher in samples containing prominent structural features as is evident in the case of adipose tissue which contains highly regular structural features typical of adipose cells. Despite this variability, in general, the cross-correlation coefficient values tend to decrease with increasing lateral separation.
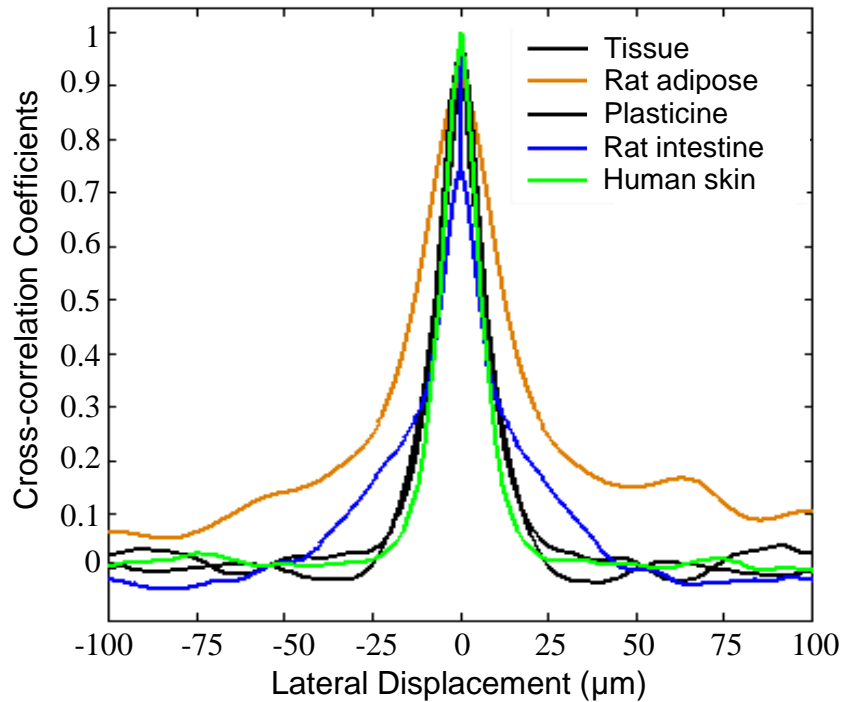

Figure 2.9: Decorrelation curves obtained from galvanometer-scanned images of several tissue phantom samples and biological tissues (negative distance corresponds to the cross-correlation between the current A-scan and previously acquired A-scans).
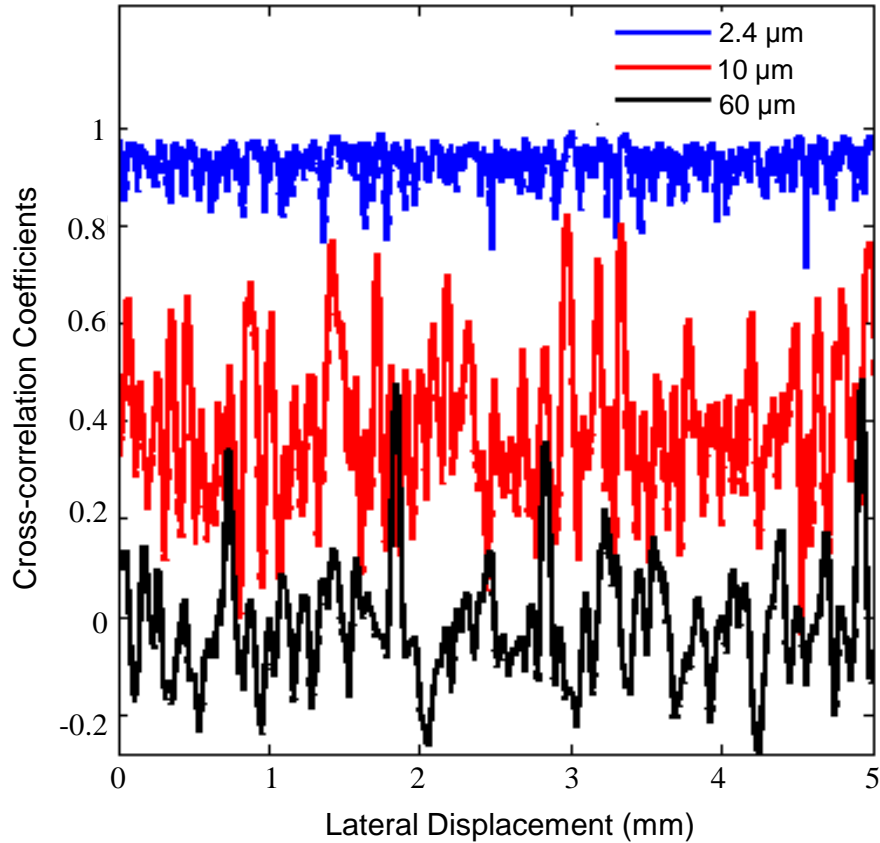
Figure 2.10: Variations in the cross-correlation coefficient values for tissue phantom.

The preprocessing steps (MA filter size and the A-scan truncation range) may also cause variations in the decorrelation lengths between different samples as the correlation coefficients may be influenced by the contributions of noise and the varying beam diameter within the truncated A-scans. All the decorrelation curves converge to a low correlation value for a lateral displacement well beyond the transverse resolution of the system. The decorrelation curves obtained in Figure 2.9 were obtained by averaging over several A-scans. Figure 2.10 shows the variations in the cross-correlation coefficient values (with no averaging) at three different displacements. These curves suggest that the data has to be significantly oversampled and the cross-correlation coefficient values need

to be averaged over several A-scans for a more accurate estimation of the cross-correlation coefficient value.

2.4.2.2 Velocity estimation during lateral scanning

For proof-of-principle, a silicone-based tissue phantom was created with 2-5 μm sized titanium dioxide (TiO_2) scattering particles. A standard galvanometer-scanned OCT image of the phantom was acquired. The means and the standard deviations of the cross-correlation coefficients for over 2000 A-scans at different lateral displacements were computed and are shown in Figure 2.11(a). The OCT beam was then held fixed while the sample was moved along the lateral direction with 5 different velocities using a computer-controlled movable stage. A threshold of 0.8 corresponding to a sampling factor of 4 was selected from the decorrelation curve. The images were then downsampled using the algorithm for the chosen sampling factor. The higher the velocity, the lower the sampling factor and the fewer the A-scans selected per resolution element. The A-scan redundancy ratio (ARR) for the downsampled assembled image was then calculated where ARR is defined as the number of A-scans compared for each selected A-scan. The mean and the standard deviation of the A-scan redundancy ratio (ARR) are shown in the blue curve in Figure 2.11(b).
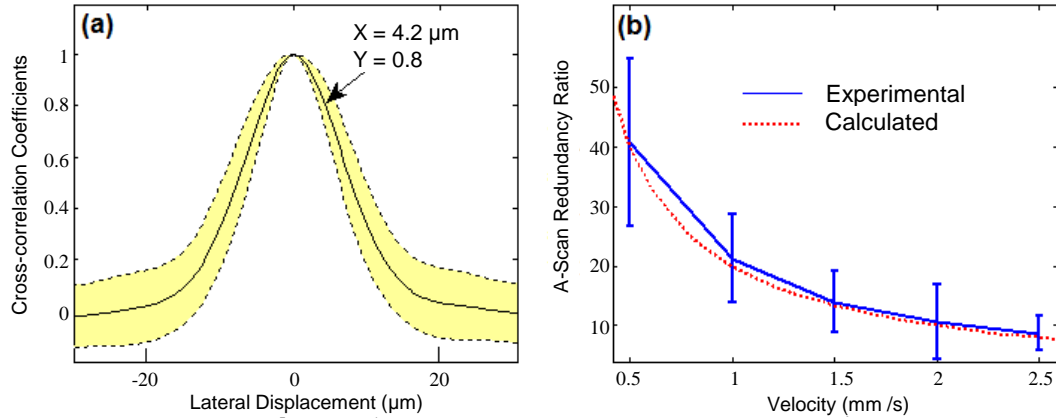
Figure 2.11: Results with a silicone-based tissue phantom with titanium dioxide (TiO$_2$) scattering particles. (a) Decorrelation curve as a function of lateral distance. The solid curve is the mean and the dotted curves are the standard deviations of the correlation coefficients. (b) A-scan redundancy ratio (ARR) as computed by the algorithm for various sample scan velocities. The error bars show one standard deviation above and below the mean.

Equation (2.4) was used to calculate the actual sampling factor for different scan velocities for the given A-scan rate of 5 kHz and transverse resolution of 16 μm. The calculated sampling factor was then divided by the desired sampling factor (equal to 4 in this case) to calculate the relative sampling ratio between the raw and assembled image and is plotted as the red dotted curve in Figure 2.11(b). The results show that the experimentally obtained results are in good agreement with the numerically predicted values. The algorithm is able to compensate for variations in scan velocity by adjusting the periodicity of A-scan selection from the raw image data set. The curves suggest that this algorithm can provide better results for highly oversampled raw images, which would occur with slower scanning velocities or with advanced OCT systems with exceptionally fast A-scan acquisition rates.

2.4.2.3 Image assembly for tissue phantoms

A set of experiments were conducted to perform image assembly by moving the sample with a non-uniform scanned velocity. All images have been log-normalized and displayed in the inverted gray scale.

A tissue phantom with titanium dioxide (TiO_2) scattering particles (size < 5 μm) was prepared for imaging. Figure 2.12(a) shows the tissue phantom uniformly sampled in time and distance by moving the sample with uniform velocity by a computer-controlled stage. Figure 2.12(b) consists of 5000 A-scans acquired over duration of 5 s by non-uniform scan velocity of the sample. The OCT beam was held fixed and the phantom was translated along the lateral direction with different velocities by programming a movable stage. Approximately 1900 A-scans were acquired while the sample moved at a velocity of 2.5 mm/s and 0.5 mm/s, respectively, and roughly 1200 A-scans were acquired during the stop interval in between.
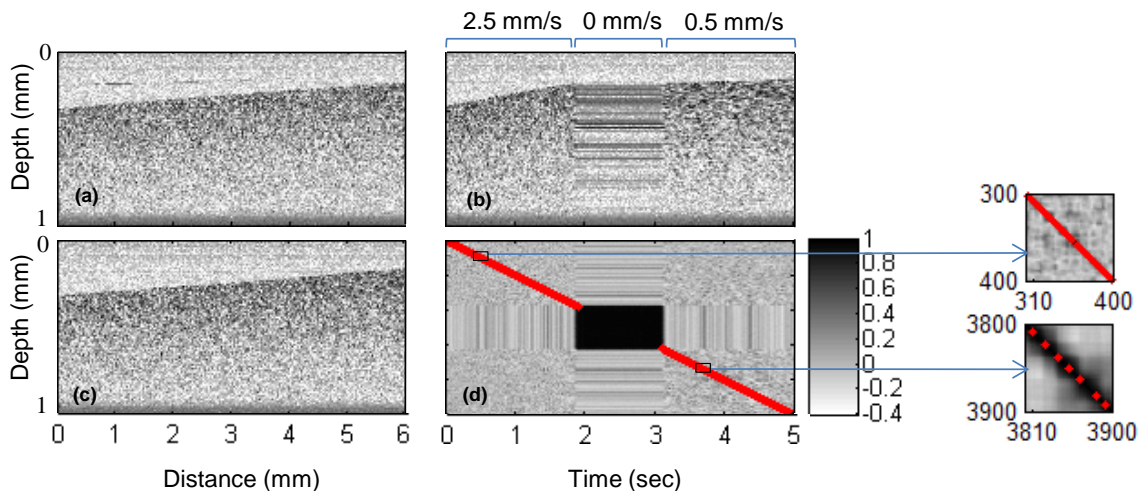

Figure 2.12: Image assembly for a silicone-based tissue phantom with titanium dioxide (TiO_2) scattering particles. (a) Motorized stage scanned image (uniformly sampled in distance and time). (b) Non-uniformly scanned image (sampled non-uniformly in distance but uniformly in time). (c) Assembled image using A-scan selection algorithm (compensated for non-uniform sampling in distance). (d) Cross-correlation matrix with red points showing the A-scans selected for image assembly by the algorithm.

To aid visualization of the A-scan selection process from the algorithm, the correlation matrix is displayed in the form of a 2-D image (Figure 2.12(d)). Each row shows the variation of the cross-correlation coefficients as a function of the adjacent A-scans. A solid diagonal line would correspond to the fact that the A-scans are perfectly correlated with themselves. The red points show the A-scans selected by the algorithm for assembling the image. The spacing of these red points will vary depending on the degree of sampling of the A-scans. In a relatively homogeneous sample, as shown in Figure 2.12, the spacing of the red points varied proportionally with the degree of oversampling. The zoomed-in areas show the different ARR corresponding to different sample scan velocities. It should be noted that calculating the complete cross-correlation matrix is not necessary for image assembly. Rather it is merely shown here to aid in visualizing the variations of cross-correlation coefficients with lateral displacement, where dark regions correspond to little or no movement and lighter regions correspond to rapid movements. Figure 2.12(c) shows the result after correcting for non-uniform sampling in distance. A threshold value of 0.7 was used for image assembly corresponding to a sampling factor of 2. The algorithm selected approximately 580 and 135 A-scans from the regions corresponding to the velocities 2.5 mm/s and 0.5 mm/s, respectively, making the assembled image uniformly sampled with a sampling factor of 1.95-2.30.

Figure 2.13 shows the result with a more aggressively scanned tissue phantom with multiple stops while scanning. In this case, the phantom was mounted on a spring-loaded movable stage (not controlled by the computer) under a fixed OCT beam. Figure 2.13(a) shows the uniformly scanned image while Figure 2.13(b) shows the manually

scanned image. Figure 2.13(c) shows the final image assembly clearly showing that the algorithm has been able to remove many of the distortions while manually scanning the sample.
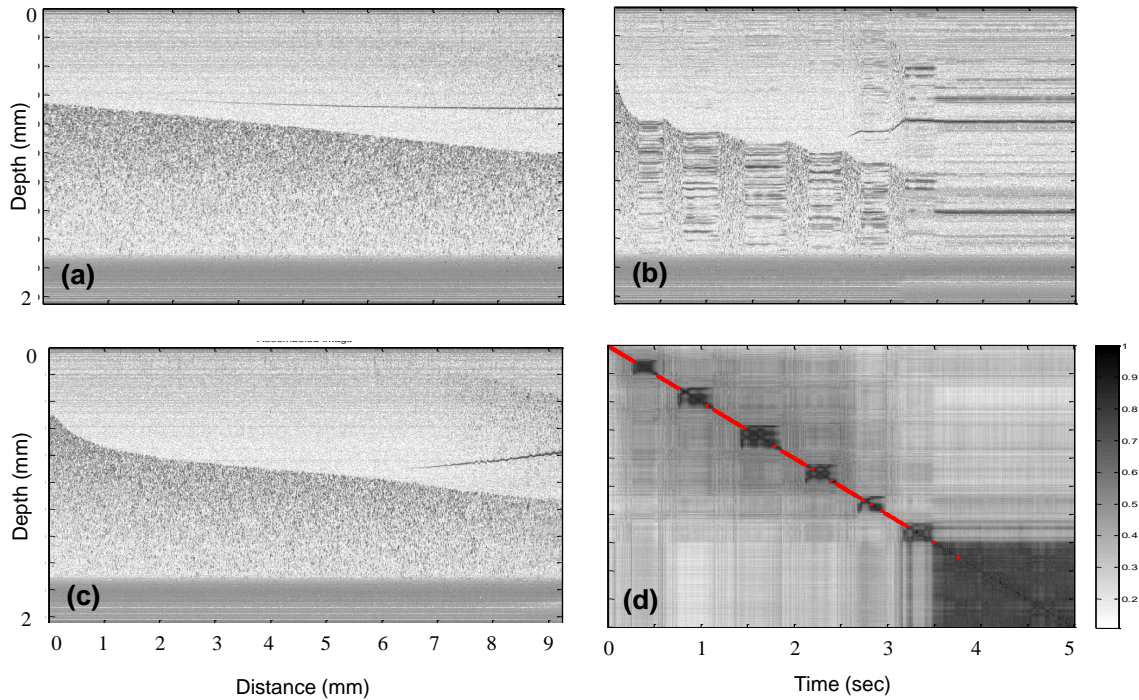

Figure 2.13: Image assembly for a silicone-based tissue phantom with titanium dioxide (TiO_2) scattering particles. (a) Motorized stage scanned image. (b) Non-uniformly scanned image. (c) Assembled image using A-scan selection algorithm. (d) Cross-correlation matrix with red points showing the A-scans selected for image assembly by the algorithm.

Figure 2.14 shows the result from manually scanning a plasticine sample over a distance of 1 cm. Plasticine is a highly scattering medium with limited penetration depth. The surface features were placed in the sample by manually scratching and carving the surface with a razor blade. The algorithm is able to assemble most of the surface image features using a threshold value of 0.3. The average sampling factor in Figure 2.14(b) is estimated to be 4. The assembled image in Figure 2.14(c) consists of ~ 600 A-scans. The algorithm fails near the 4 mm mark in Figure 2.14(a) due to the highly similar surface features which are misinterpreted as stops by the algorithm.
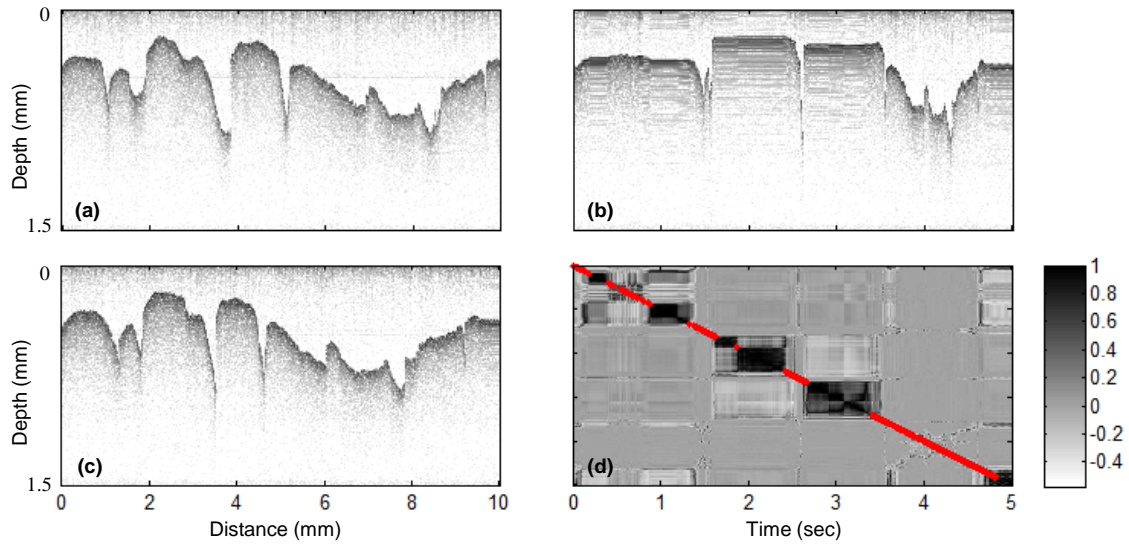
Figure 2.14: Image assembly for a plasticine sample over a sample length of 1 cm. (a) Uniformly scanned image using a motorized stage. (b) Non-uniform hand-scanned image. (c) Assembled image. (d) Cross-correlation matrix with red points showing the A-scans selected for image assembly.

Another example of image assembly is shown in Figure 2.15 which shows the result of hand scanning of an agarose based sample with titanium dioxide scattering particles forming internal structures. The structures seen within the sample are probably due to the aggregation of the particles. The sample was scanned by hand over a distance of 10 mm. In the uncorrected image Figure 2.15(b), the sample structure is not apparent while Figure 2.15(c) shows reasonably good resemblance with the uniformly scanned image of Figure 2.15(a). The difference at the start of the image may be due to the misalignment of the scanning plane while manually scanning the sample.
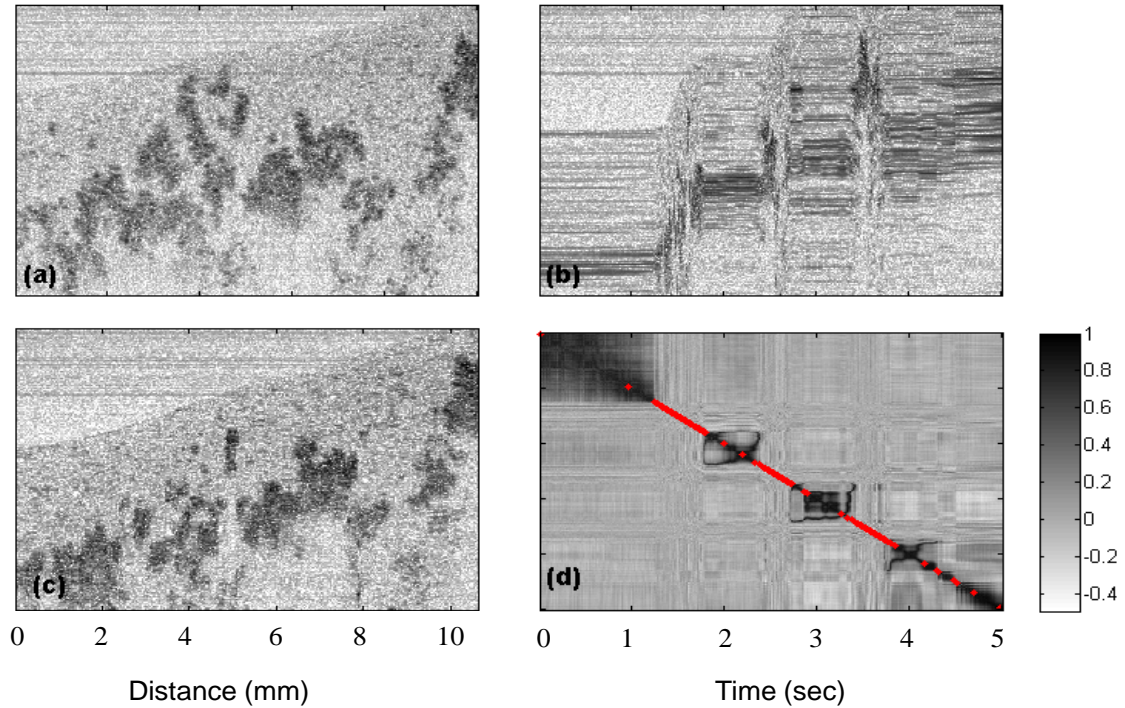
Figure 2.15: Image assembly for an agarose-based tissue phantom with titanium dioxide (TiO$_2$) scattering particles. (a) Uniformly scanned image using a motorized stage. (b) Non-uniform hand-scanned image. (c) Assembled image. (d) Cross-correlation matrix with red points showing the A-scans selected for image assembly.

2.4.2.4 Image assembly for biological tissues

Image assembly with biological tissues proved to be much more challenging compared with tissue phantoms. In Figure 2.16, human adipose tissue was manually scanned over a distance of 1.5 mm. A threshold value of 0.7 was used due to the relatively slow rate of decorrelation of adipose tissue. A line scan rate of 5 kHz was used for A-scan acquisitions. The assembled image in Figure 2.16(c) has good correlation with the galvanometer-scanned image in Figure 2.16(a). Limitations in the algorithm are apparent at the lateral displacement of 0.3-0.5 mm in Figure 2.16(c) which are believed to be due to hand vibrations while trying to hold the spring-loaded stage still. These vibrations cause rapid decorrelation of the A-scans which the algorithm misinterprets as valid probe displacements.

Figure 2.16: Image assembly for human adipose tissue. (a) Galvanometer-scanned image. (b) Non-uniformly hand-scanned image. (c) Assembled image. (d) Cross-correlation matrix with red points showing the A-scans selected for image assembly.

Another example with rat adipose tissue is shown in Figure 2.17. The assembled image shows that, although the algorithm has does a decent job in assembling the image from the manually scanned data, it is unable to reconstruct the fine structures within the sample.
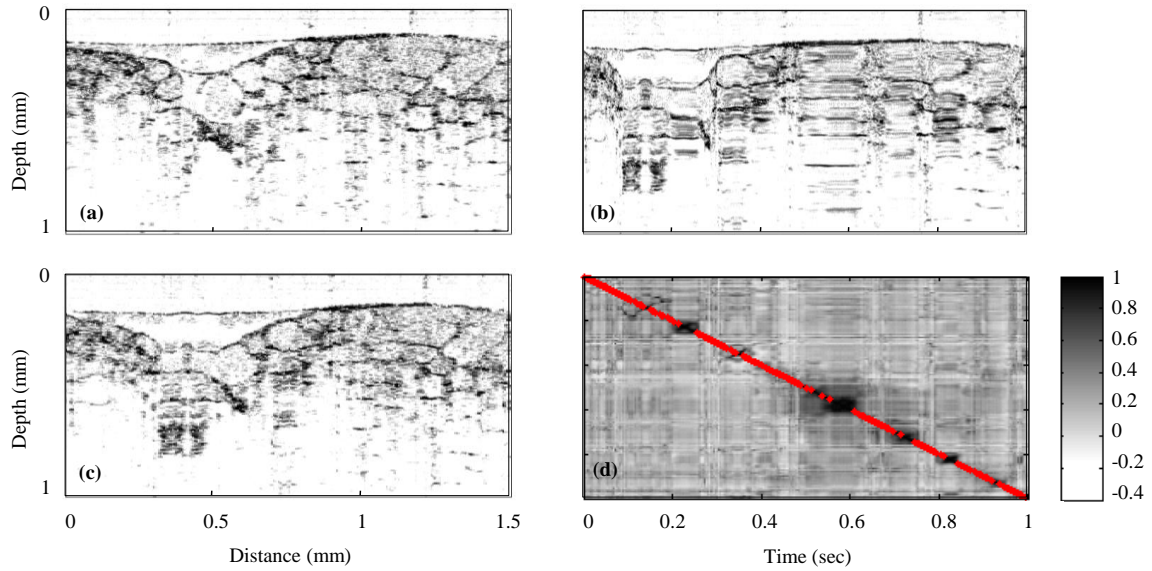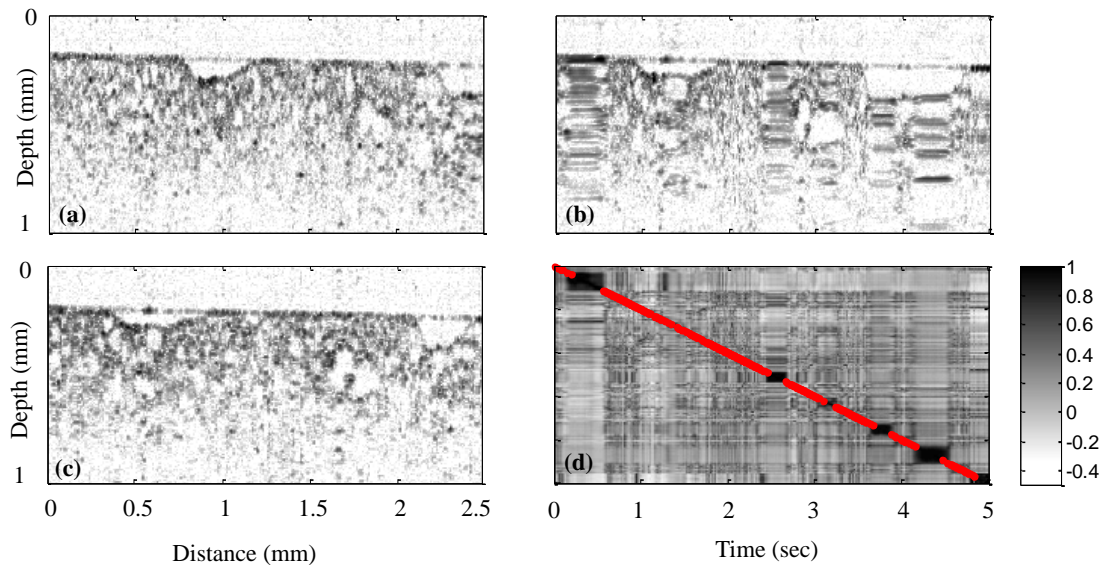

Figure 2.17: Image assembly for Rat adipose tissue. (a) Galvanometer-scanned image. (b) Non-uniformly hand-scanned image. (c) Assembled image. (d) Cross-correlation matrix with red points showing the A-scans selected for image assembly.

44

Figure 2.18 shows the results with manual-scanning of human breast tissue containing both tumor and adipose tissues. The sample was moved manually under a fixed OCT beam along a distance of 1 cm. The sample was stopped for various lengths of time during scanning as shown in Figure 2.18(b). The assembled image in Figure 2.18(c) shows that the algorithm is able to remove the artifacts associated with abrupt stops during manual-scanning. The differences between the motorized-stage scanned image and the hand-scanned image are likely due to the misalignment of the scanning plane while manually moving the sample by hand without using the translational stage.
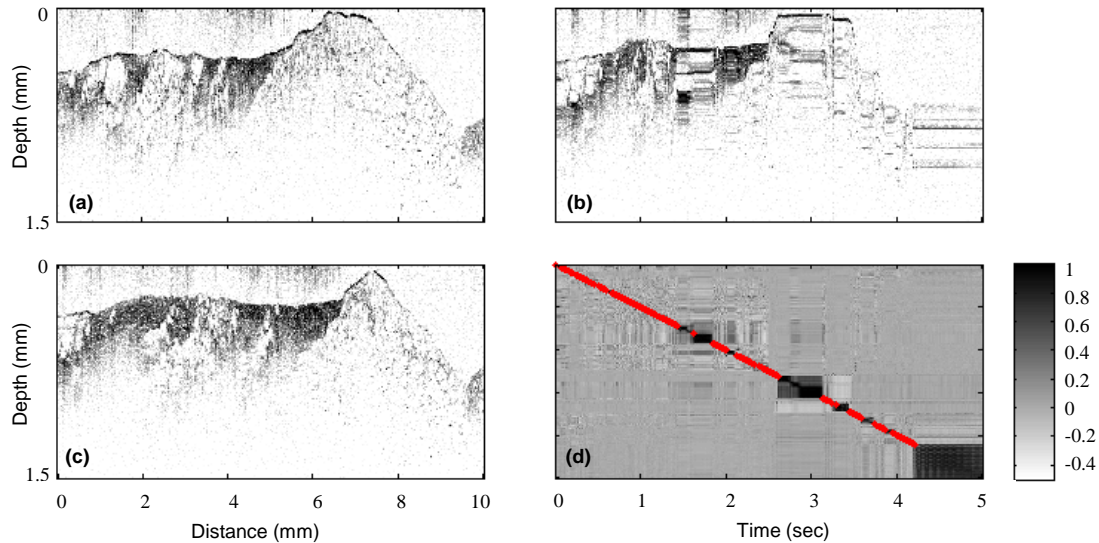
Figure 2.18: Image assembly for a human breast tissue over a sample length of 1 cm. (a) Uniformly scanned image using a motorized stage. (b) Non-uniform hand-scanned image. (c) Assembled image. (d) Cross-correlation matrix with red points showing the A-scans selected for image assembly.

2.4.3   Limitations

Although successful image assembly has been demonstrated using phantom samples and biological tissues, the present technique has a number of limitations that need to be overcome when performing real-time *in vivo* imaging.

This technique relies on the selection of an appropriate threshold for accurate image reconstruction. The current selection method for the optimal threshold was based

on the decorrelation curves of a particular sample similar to those shown in Figure 2.9. However, as the results show, there exist inter-sample and intra-sample variations in the cross-correlation coefficients at each lateral displacement. These variations can be attributed to inhomogeneity of structural features in tissues, variability in speckle patterns, and noise in the system. Due to these variations, selection of a single threshold throughout the image assembly procedure may produce errors in the final reconstructed image. One possible solution is to adaptively vary the threshold based on the tissue properties. Another possible solution to this problem could be to select a very low threshold value based on the fact that, at distances beyond the transverse resolution of the system, the cross-correlation coefficient values become very small. However, this would make the final reconstructed image very pixilated as very few A-scans will be selected for image formation. As the success of this technique depends significantly on how well the decorrelation curves represent the actual lateral displacement, further studies need to be done to investigate the dependency of various parameters on the decorrelation curves.

The algorithmic techniques described here operate under the assumptions that scanning is unidirectional along one dimension with no relative change in the orientation between the probe and sample. However, in practice, hand jitter and probe rotation while manually moving the probe will invalidate many of these assumptions, resulting in image artifacts. More sophisticated algorithms can be designed to compensate for these effects. Cross-correlating a block of A-scans rather than single A-scans can potentially be used to detect change in scan direction and orientation of the probe [35]. By selecting depth-dependent regions along each A-scan from which to do cross-correlations between adjacent A-scan regions, it may be possible to track angular out-of-plane displacements

as well. The intended application for this approach is to assemble large images over scan ranges that exceed the capabilities of current galvanometers and computer-controlled scanning techniques. In cases for a hand-held probe, needle-probe, or catheter, the precise in-plane orientation of the acquired data may not be as critical as capturing adjacent A-scans over large lateral distances.

This technique may impose certain limitations on the speed of image acquisition. Our experimental results show that a sampling factor as high as ~ 50 may be necessary for good results. Hence the velocity of the probe has to be constrained so that the sample is sufficiently oversampled. However, this limitation could be easily countered due to the availability of high-speed OCT systems [4, 34]. While commercial OCT systems have scan rates in the range of 25-40 kHz, fast swept-source or spectrometer-based OCT systems can extend the A-scan rates to several hundreds of kHz. For instance, a typical Fourier-domain OCT system with a 25 kHz A-scan rate may allow a probe with 16 μm lateral resolution to be moved with a maximum velocity of 8 mm/s while still allowing a sampling factor of 50. Hence fast OCT systems would allow reasonable freedom to allow free-hand manual-scanning without compromising the effectiveness of the algorithm. Moving faster than the threshold will result in undersampling. A possible extension of this algorithm could be to detect undersampling by comparing the cross-correlation coefficients of adjacent A-scans against a lower threshold value. Subsequently, interpolation algorithms can be used to estimate the missing A-scans [36]. Another potential extension could be to use the phase information to detect lateral displacement. Phase information has been found useful in a number of other techniques such as Doppler

OCT and elastography. In this analysis, the rapid variability of phase between adjacent A-scans made it an unreliable metric for quantifying the amount of lateral displacement.

CHAPTER 3   SONIFICATION OF OCT DATA

This chapter describes the sonification of OCT data and images. A portion of the initial work reported in this chapter was done in collaboration with Morgan Wang. The OCT images used for sonification were acquired by Steven Adie and Adam Zysk. All of these individuals were members from the Biophotonics Imaging Laboratory at the Beckman Institute. Some of the results shown in this chapter have been adapted from recently published work [67].

A multi-sensory representation of OCT data may be an effective method to enhance the real-time interpretation of OCT data and images. Audio representation (known as sonification) may also be useful in those scenarios in which non-image data is acquired. This chapter discusses various techniques of sonification of data, the psycho-acoustic principles behind any sonification system design, and the results of sonification of OCT data in A-scan mode sonification and image-mode sonification. The basic aim of the sonification presented here is to distinguish between normal and tumor tissues based on the rendered audio signals.

3.1   Methods for sonification of data

The audio representation of data is known as sonification. Various methods of translating data into auditory signals have been reported in the literature [68]. Although, there is no agreed upon classification of sonification techniques, most of the sonification methods can be divided into the following main categories.

3.1.1   Audifications

Audification is an audible playback of the data samples where each data value directly represents a sound waveform sample. Audification is effective if the physical mechanisms that have generated the data samples are similar to the physics of natural sound transmission through air. The frequency range of data should be within the human audible range, i.e. 20-20,000 Hz, and there should be enough data points to play sound for a sufficient duration. However, for most real-world data these conditions are rarely met. Hence, typically the data has to be re-sampled, the frequencies shifted to within the human audible range, and appropriate scaling of the data has to be done prior to audible playback of the sound [69]. Very few successful audifications of data have been demonstrated. Audification of seismic data is an exception, primarily because the physics of seismic vibrations transmitted through earth is very similar to sounds transmitted through air.

3.1.2   Earcons

Earcons have been defined as non-verbal audio messages consisting of short, rhythmic sequences of pitches that can have variable intensity, timbre and amplitude to represent different categories of data. Simple earcons can be combined in a tree-like hierarchy to produce more complex sounds. For example earcons representing *delete, create, file, program* could be combined to produce more complex earcons representing for example *create file* or *delete program*. Earcons have found extensive applications in human-computer interfaces (HCI). However, earcons require data to be categorized in distinct units which is not possible in many applications [70].

3.1.3   Auditory icons

Auditory icons are based on associating a unique sound to a specific message (or signal) [71]. In contrast to earcons, auditory icons have a direct analogy between the sound and the event. For example, auditory icons (unique sounds) can be assigned to normal and tumor tissues in biomedical data. The limitation of auditory icons is that they require *a priori* classification of data. However, some authors have used parameterized auditory icons to represent the variability in the data.

3.1.4   Parameter-mapped sonification

Parameter-mapped sonification is the most widely used technique for sonification. It is based on calculating features (parameters) from the data and mapping them onto sound attributes such as pitch, loudness, tempo, and duration. This would typically require scaling the range of the data parameters by using various linear/nonlinear mapping functions to fit within the human sensation of these sound attributes. This method does not require prior categorization of data and hence has been widely applied for process monitoring. However, interfaces based on parameter-mapped sonification are more difficult to design and interpretation may require extensive user training.

3.1.5   Model-based sonification

Another approach is model-based sonification where the data parameters control a parameterized sound model to generate non-speech audio [68].

3.2   Psycho-acoustic principles

The psycho-acoustic properties of human hearing have been widely studied and are well understood. Psychological perception of sound can be described in terms of loudness,

pitch, timbre and time. These sound attributes in turn depend on the physical properties of intensity, frequency, waveform and duration of the sound waves. The relationship between the physical properties and the psycho-acoustic perception of sound must be considered for sonification design. Moreover, the interactions among the sound attributes must also be taken into account, especially when more than one of these variables is to be manipulated simultaneously in the same sonification.

The psycho-acoustic concepts of *critical band* and *just noticeable difference (jnd)* have the strongest implication for a good sonification system design. Critical band is the frequency dependent bandwidth at which sound energies interact with each other. It can be approximated by the relationship $\Delta f = 25 + 75(1 + 1.4f^2)^{0.69}$ where f is the center frequency in kHz [72]. The phenomenon of critical band is believed to be due to the excitation of the same region of the basilar membrane by different tones (that lie within this band). Just noticeable difference (jnd) characterizes the ability to differentiate between two nearly equal stimuli. In general, people are much better at making relative judgments of the sound attributes than at making an absolute judgment.

3.2.1   Loudness

Loudness of a sound is the magnitude of the auditory sensation produced by the sound [73]. The human perception of loudness is primarily dependent upon the sound pressure (intensity) values; however, the frequency, duration and bandwidth of sound also influence the perceived loudness.

According to the Steven's power law, loudness (*L*) is related to the sound intensity *I* by the relationship $L = kI^{0.3}$ where k is a constant which depends upon the

units used and individual perception. The loudness level of sound is generally measured in units of *phons* which are curves of constant loudness. The human sensitivity to loudness as a function of frequency is shown by the Equal-Loudness contours (in units of phons) in Figure 3.1 [73]. Some general trends can be seen from these curves. These curves show that the sensitivity of the human ear to loudness is greatest between 3000 and 4000 Hz and it decreases at lower frequencies. The curves suggest that the variation in perceived loudness may be minimized by using a frequency range of approximately 800-2000 Hz. Loudness is also dependent on duration: with increasing duration of a sound the perception of loudness stabilizes after about 100 ms. If frequencies of two different sounds lie within the critical band, then the perceived loudness will only slightly change; however, if they lie outside the critical band the perceived loudness may greatly increase.

The *jnd* in intensity between two sounds is about 1 dB (a 12% change in intensity). However, in real life scenarios, a change of 3 dB (50% change in intensity) is easily detectable by humans [72]. Using loudness as a sound attribute for sonification is challenging as the human auditory system has a tendency to adjust itself to the loudness level of sound; loudness is affected by the distance from the source, and the loudness of a sound may be masked by other sounds in the environment [72, 74].
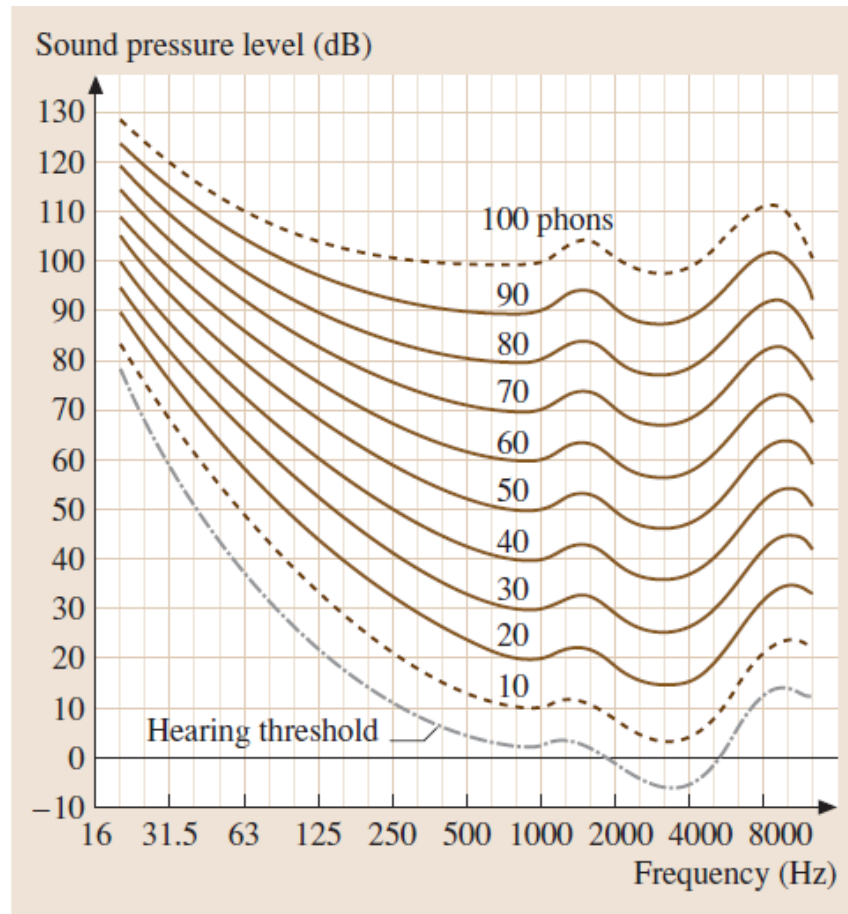
Figure 3.1: Equal-loudness contours (from B. Moore, "Psychoacoustics," in *Springer Handbook of Acoustics*, T. D. Rossing, ed., pp. 459-501, 2007) [74].

3.2.2   Pitch

The sensation of pitch can be defined as the perceived fundamental frequency of the sound. Perception of pitch is primarily dependent upon the frequency of the sound wave, but is also a function of the intensity level and duration of sound. In harmonic sounds, pitch depends on the fundamental frequency while for inharmonic sounds it is a function of the amplitude-weighted mean of the spectral components. The audible range for most humans is from 20-20,000 Hz. On average, sound should have duration of at least 13 ms to be ascribed as a definite pitch. A short duration sound will be heard as a click rather than a pure tone.

Pitch discrimination between two tones depends on their intensity and frequency difference as shown in Figure 3.2. These curves show that the human ear is more sensitive to frequency changes at the mid-frequency region between 1-4 kHz. The *jnd* for pitch is typically about 1/30th of the critical bandwidth at a particular frequency. Although the human ear has sensitivity up to around 20 kHz, sensitivity of the human ear drops significantly at higher frequencies. Thus, it is reasonable to use frequencies in the middle of the audible range, i.e. 100-5000 Hz, so that the sound is audible in most circumstances.
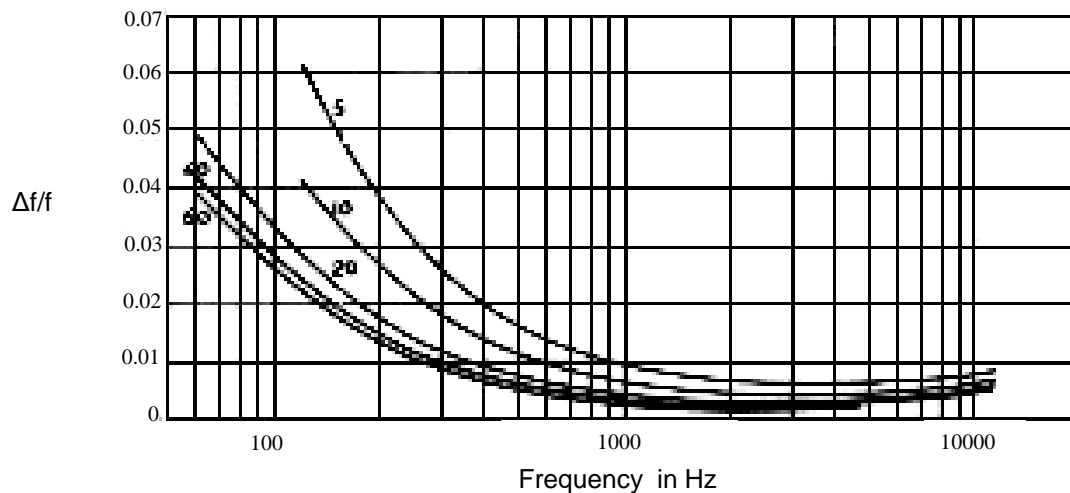
Figure 3.2: Sensitivity of pitch as a function of frequency and intensity of sound. Numbers on the curves indicate levels above the threshold of hearing. (from H. F. Olson, *Music, Physics and Engineering*, Dover Publications,1967) [73].

3.2.3   Timbre

Those characteristics of sound which enable the human auditory system to distinguish between sounds of similar pitch and loudness are, by definition, timbre. Timbre perception depends upon the harmonic content, temporal evolution and the vibrato and tremolo properties of the sound waves. Timbre may be useful to represent multiple data streams simultaneously.

3.3   Sonification of OCT data

In this section, different sonification approaches of OCT data are described. A well designed sonification system must be fast, the sound must make intuitive sense, and the sounds must effectively extract important diagnostic features from the OCT data.

A simple method of sonification is the use of earcons or auditory icons. For earcons, suppose two parameters are calculated from data and the range of values in these parameters are further categorized as low, medium and high values. Each of the parameters can be assigned a unique sound and some property of this sound can be changed based on the low, medium or high categories. Earcons can potentially be used for OCT data but the main difficulties are the requirement of prior categorization of data.

A more simple method is the use of auditory icons, where each A-scan can be classified as either normal or tumor. Based on this classification, a unique predetermined sound can be assigned to each A-scan. This requires the use of classification algorithms to distinguish between tissue types. Real-time image interpretation is thus performed by the classification algorithm, which is subject to its own sensitivity and specificity limitations [75].

One method could be the direct audification of spatial-domain intensity data which has the potential advantage of high speed due to low processing requirements. However, this audification may be noisy due to the presence of speckle in OCT. Audification has in general been successfully applied in those situations where the physical model of the data generating mechanisms is very similar to the physics of sound transmission through air. OCT datasets in general will not meet this requirement.

However, removing the noise from the OCT data and re-sampling the data to fit within the audible range may be an interesting approach to try for some OCT datasets.

Another method was primarily adapted from previous work [76]. In this method the horizontal dimension represents time and the vertical direction represents pitch for a 2-D OCT image. The number of sinusoids is determined by the number of vertical pixels in each column. A frequency range of 20-10,000 Hz was logarithmically mapped to the number of pixels (in this case 1024). The intensity of each pixel is the amplitude of the corresponding sine wave. The final sound is the summation of all sinusoids from a particular column in an image and was played for duration of 10 ms keeping in mind the time required to perceive a pure tone by the human auditory system. Although this method worked reasonably well, it required excessive computation due to the large number of sinusoids required for sound synthesis.

3.4   Parameter-mapped sonification

Parameter-mapped sonification, which is based on the mapping of features extracted from data into sound attributes, was chosen for sonification of OCT data. In this method the sensitivity of the human auditory system is utilized for tissue classification. Hence, it does not require prior classification of the data. The method of parameter-mapped sonification is illustrated in Figure 3.3. Characteristics of A-scans or image parameters that can be used to classify tissue types are chosen, and the extracted parameters are then mapped to a set of sound attributes for sonification. The sound attributes depend on the sound synthesis algorithm employed. The mapping from the data parameters to the sound attributes is performed while considering the psycho-acoustic response of the human auditory system.
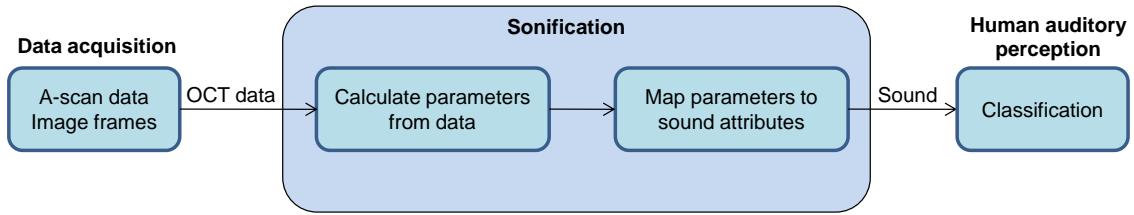
57

Figure 3.3: Parameter-mapped sonification for OCT.

3.4.1 Parameter extraction from OCT data

A range of different parameters have been reported for tissue classification in OCT. These parameters depend on the nature of the data and may typically be the slope, intensity variations, spatial frequencies, periodicity of A-scans, and textural features of OCT images [59, 75, 77, 78]. The extracted parameters must have good discriminating power and must be computationally simple in order to meet the real-time requirements of a sonification system. In this sonification to distinguish between adipose and tumor tissue, the slope and the energy in the Fourier space regions of the A-scans have been used as data parameters.

OCT images of human breast adipose and tumor tissue, and the corresponding A-scans, are shown in Figure 3.4. Adipose tissues mostly consist of lipid-filled fat cells called adipocytes which are typically of 20-160 microns in size. These cells give adipose tissue its regular structure. Tumor is characterized by the loss of normal tissue architecture, alteration in the nuclear size and shape, and an increase in the nucleus to cytoplasm ratio which is defined as the diameter of the nucleus divided by the diameter of the cytoplasm. Optically, the nucleus is a highly scattering organelle. Hence tumor tissue in an OCT image is very dense and highly scattering and the resulting OCT signal as it propagates through the tissue suffers from more attenuation compared to adipose tissue.

The different attenuation causes a difference in the slopes of the (logarithmically mapped) A-scans with tumor having a higher (steeper, more attenuating) slope compared to adipose tissue. The slopes were calculated by selecting values in the A-scans that were above a certain threshold (~ 30 % of the maximum value) and then fitting a linear function to these values.

Previous work has shown that the unique Fourier signatures from tissue types can be used for classification [75]. The Fourier transforms of A-scans are shown in Figure 3.4 (e). Adipose, due to its regular structure, has more energy concentrated in the low frequency regions while tumor has more high frequency content. The Fourier transform was normalized to a unit area and was divided into non-overlapping regions. The squares of the areas under the curve corresponding to the low frequencies (labeled as region I), middle frequencies (II) and high frequencies (III) were selected as the three spectral parameters in our sonification.
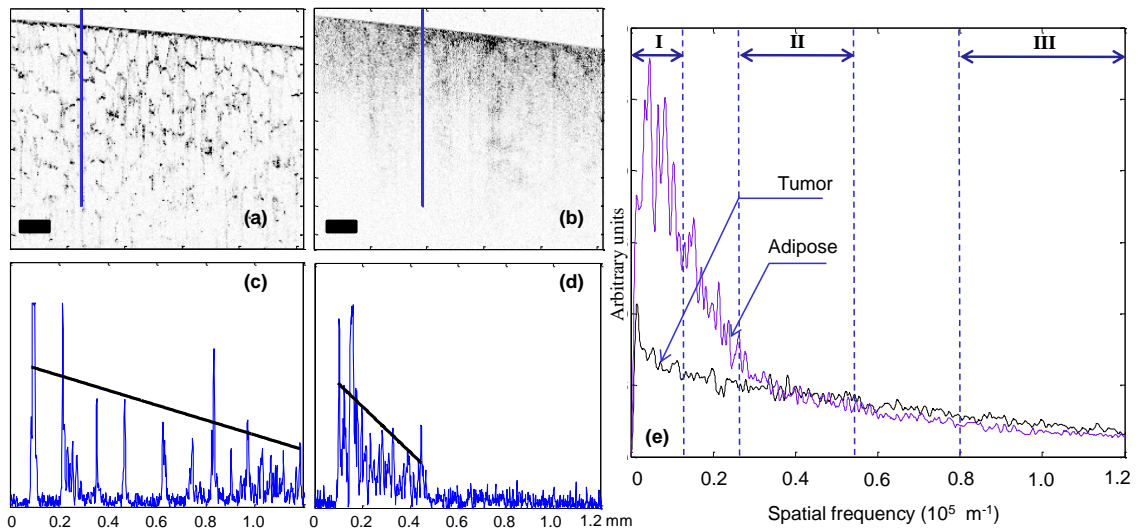
Figure 3.4: OCT image and data analysis from human breast. (a) Adipose tissue and (b) tumor tissue. (c) A-scan corresponding to the highlighted line from the adipose tissue. (d) A-scan corresponding to the highlighted line from the tumor tissue. (e) Normalized Fourier transforms (average of 150 A-scans). Roman numerals indicate the regions corresponding to the three spectral parameters. Scale bars represent 200 μm. The superimposed lines in (c) and (d) represent the 1st order fit to the A-scan for calculation of slope.

3.4.2   Sound synthesis

A number of sound synthesis methods such as additive synthesis, subtractive synthesis, frequency modulation (FM) synthesis, and granular synthesis can be used to generate sound [79]. For any given application there is no preferred technique, with each having its own merits and demerits.

FM synthesis was used, which has the advantage of generating a rich variety of sounds with the control of only a few parameters. The FM signal can be mathematically described as

$$A.\cos(2\pi f_c t + M \sin 2\pi f_m t) \tag{3.1}$$

where f_c is the carrier frequency, f_m is the modulating frequency, A is the amplitude and M is the modulating index. In this technique, the carrier wave frequency f_c is modulated by the modulating wave frequency f_m. The FM modulated signal consists of a complex tone with frequency components separated from one another by the modulating frequency as shown in Figure 3.5. However, if there are reflected side frequencies (due to sidebands falling into the negative frequency domain of the spectrum) then the ratio f_c/f_m would determine the position of the components in the spectrum [80]. The amplitude of the components can be determined by the Bessel function which would be a function of the modulating index M, where higher values of M imply that more spectral energy will be dispersed among the frequency components, as can be seen from Figure 3.5(b).
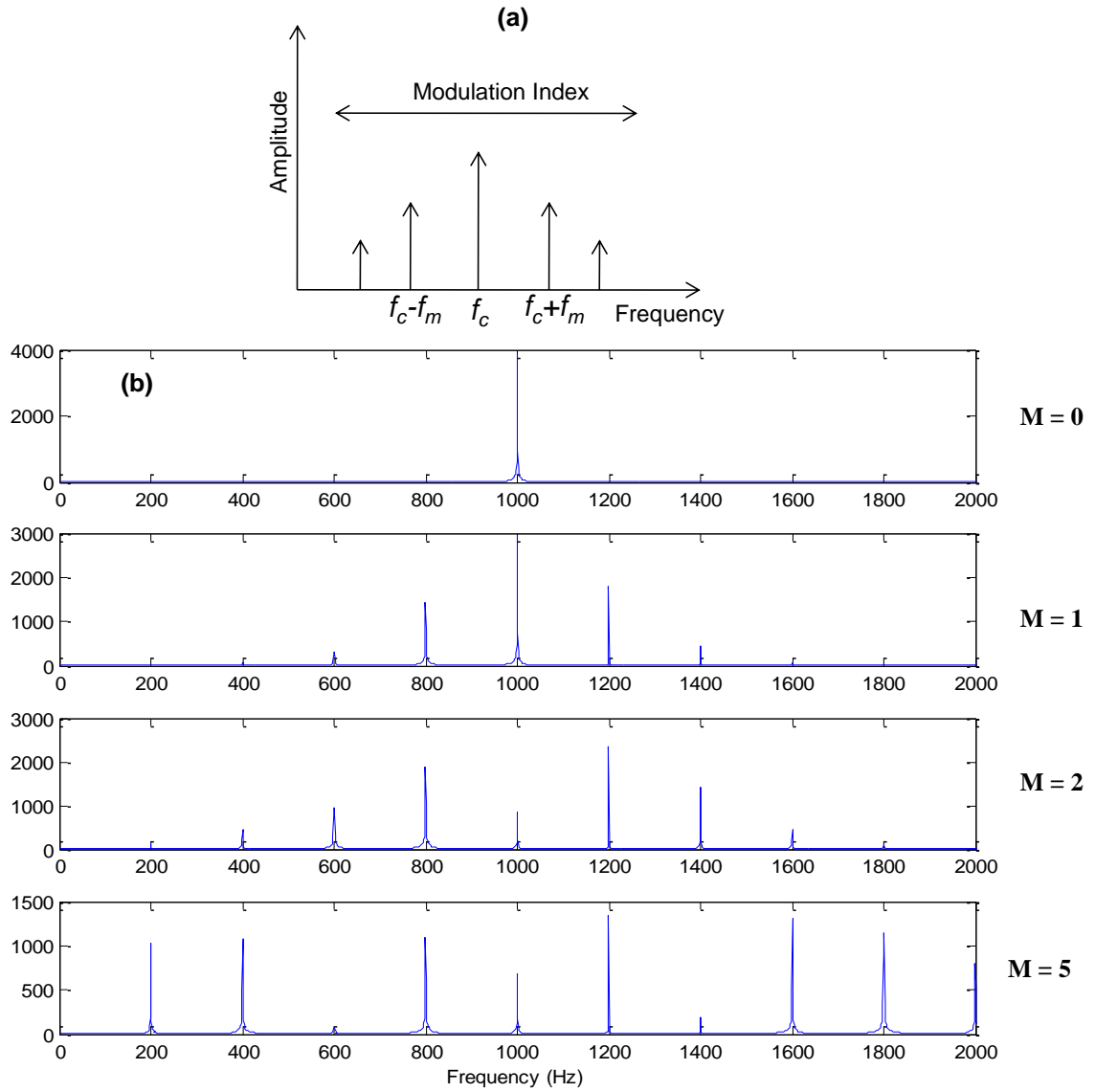
Figure 3.5: (a) Spectral components in FM synthesis. (b) FM synthesis. Amplitude=1, carrier frequency f_c=1,000 Hz, modulation frequency f_m=200 Hz.

3.4.3   Parameter mapping to frequency-modulation synthesis model

The parameters extracted from the OCT data can be mapped to any or all of the attributes. The slope of the A-scans and the spectral features corresponding to the low (I), middle (II) and high frequency (III) regions of the Fourier spectrum of the data as shown in Figure 3.4(e) were selected as the significant parameters. These parameters were

mapped after appropriate scaling into the carrier frequency f_c, modulation index M, amplitude A and modulating frequency f_m, respectively, where $f_m = $ [(Energy in region III) x f_c)].

Interpretation of the mapping is shown in Figure 3.6. The slopes of the A-scans are mapped to the pitch, the high frequency content determines the separation of the spectral components relative to the carrier frequency, while the low frequency content will determine spectral energy within these spectral components. The final synthesized sound is strongly influenced by the choice of carrier frequency. In these datasets, slope was the variable with the greatest discriminating power and hence was mapped into the carrier frequency. These mappings made the sonifications of tumor and adipose tissues have non-overlapping audio spectra and made the perceived sound of tumor of higher pitch. This makes intuitive sense as the Fourier spectra of tumor has more high frequency content compared to that of adipose tissue.
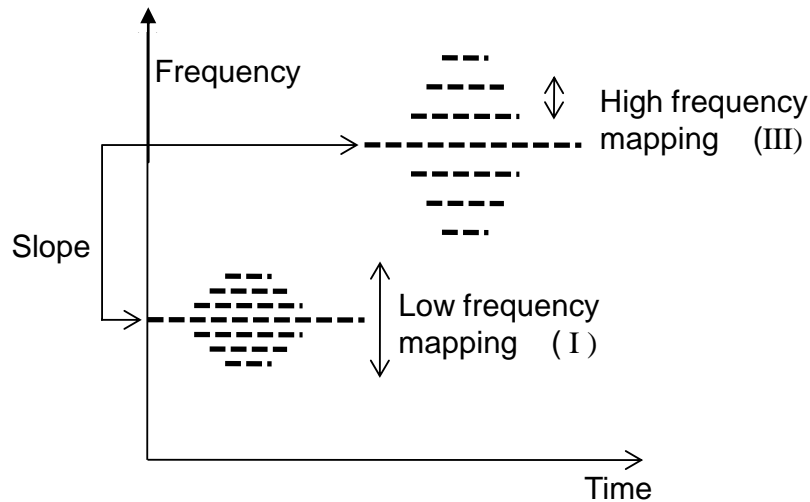

Figure 3.6: Mapping of parameters for sonification via FM synthesis.

Sonification of the data was performed on a block by block basis, with the final output containing a clicking sound at the end boundary between blocks. These artifacts were minimized by multiplying the sonified block by an envelope having linearly rising and decaying slopes at the edge of the blocks as shown in Figure 3.7.
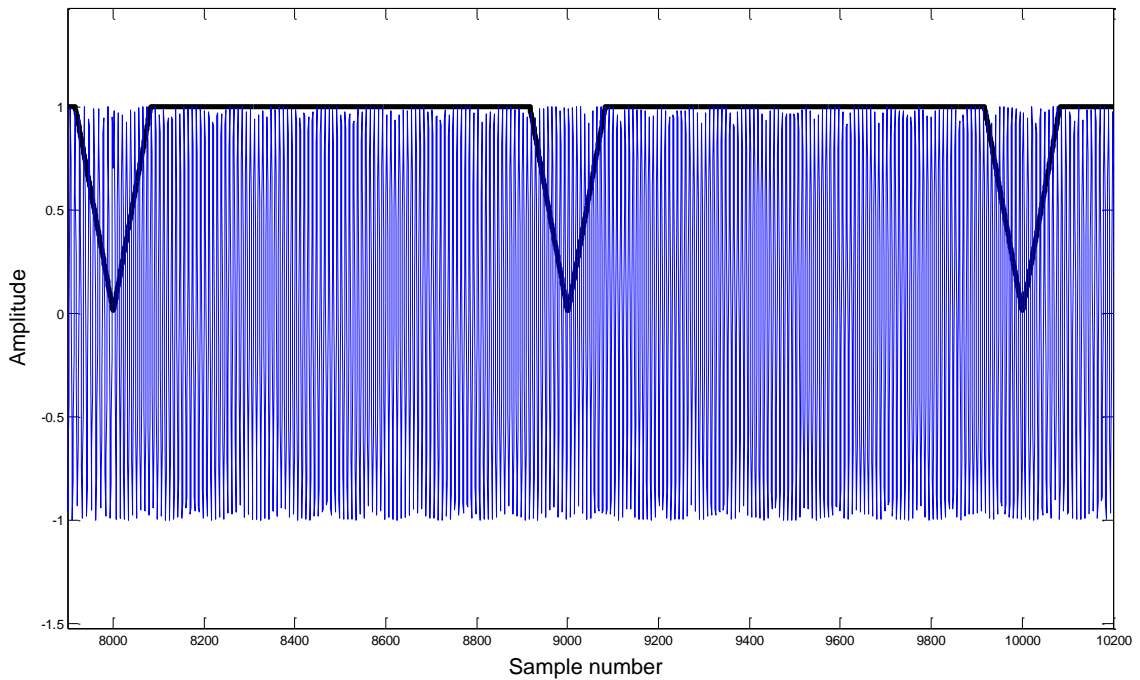
Figure 3.7: Multiplication of each block by an envelope to remove clicking sounds generated between blocks.

3.4.4   Sound rendering modes

The sonification of OCT data has been organized into two modes, i.e. A-scan and image-mode sonification. In the A-scan sonification mode each individual A-scan (or a group of A-scans for faster playback) is sonified. Although this mode has high resolution, it has the limitation of being non-real-time as the typical A-scan acquisition rate (~ 0.1 ms for an A-scan rate of 10 kHz) will be much higher than the playback time (~100 ms) of the sound. A playback time of 100 ms was chosen based on the tone perception of the human ears.

Image-mode sonification may be used for real-time sonification of the data. In the image-mode each frame is played for the duration of the playback time of the sound, and hence is much faster than the A-scan mode. In this mode each frame is divided into a certain number of blocks and for each block the average value of the parameters are calculated and mapped into sound, as shown in Figure 3.8. The final synthesized sound consists of summation of the waveforms from each individual block. The sonification (parameter calculation + sound synthesis) of each block is independent of all the other blocks. Hence, these calculations can be done in parallel for each block, which will significantly decrease the computational time for each frame. However, this mode will have a lower resolution than the A-scan mode (where the resolution depends on the number of divisions of each frame).
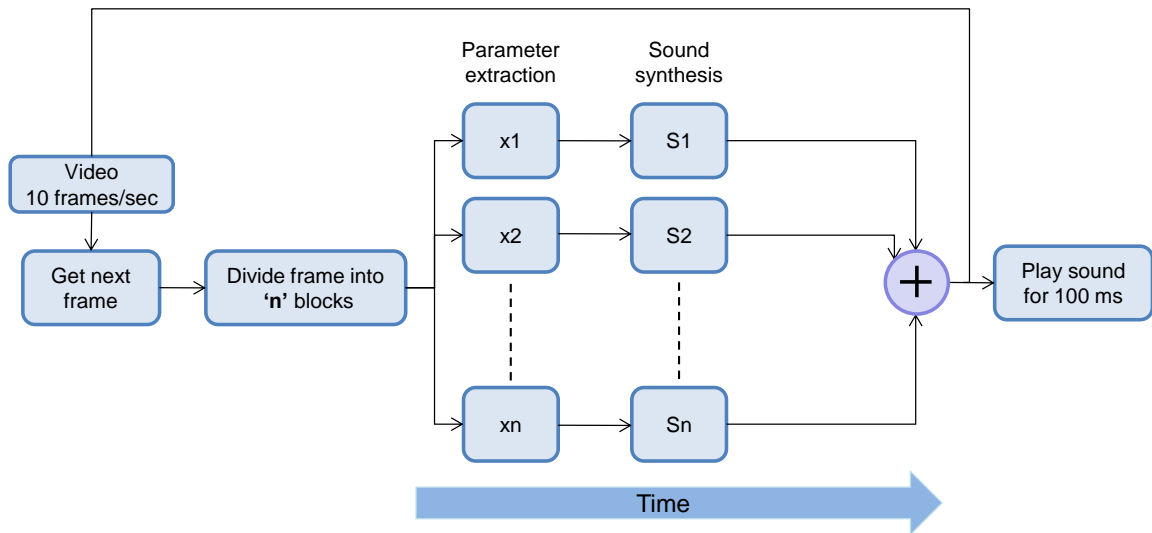

Figure 3.8: Block diagram for the image mode.

3.5   Results

Sound was synthesized using Matlab and played at a sample rate of 10 kSamples/s. The results obtained by sonification in the A-scan mode and the image mode are shown below.

3.5.1   A-scan mode

In the A-scan mode the acquired A-scans are grouped together in the form of bins, each 10 A-scans wide. The data parameters are calculated for each A-scan and averaged together for each bin. Each bin is played for duration of 100 ms. The mappings of the A-scans parameters obtained from adipose and tumor tissues are shown in Table 3.1, which shows that adipose will have a sound of lower pitch with the spectral components more closely spaced to each other and energy more widely dispersed among them. In contrast, tumor will have a sound of a higher pitch with relatively large spacing between the spectral components, and with most of the energy concentrated within the carrier frequency (due to low modulation index M).

Table 3.1. A-scan parameter mapping to FM synthesis

FM synthesis parameters	A-scan parameters	Adipose	Tumor
Carrier frequency (f_c)	Slope	Low	High
Modulation index (M)	Low frequency content *(I)*	High	Low
Amplitude (A)	Middle frequency content *(II)*	Moderate	Moderate
Modulation frequency (f_m)	High frequency content *(III)*	Low	High

Note: Roman numerals refer to frequency bands shown in Figure 3.4.

Figure 3.9 shows the sonification of a 2D OCT image containing a tumor margin. The extracted parameters from the image are displayed in image form in Figure 3.9(a). The dataset was acquired using a spectral-domain OCT system with 800 nm center

wavelength and 70 nm bandwidth, providing an axial resolution of 4 μm. The spectrogram (computed using the short time Fourier transform) shown in Figure 3.9(c) displays the frequency components of the sound at each time instant and is helpful in visualizing the sonification results. Results demonstrate that tumor and adipose have distinct sounds.
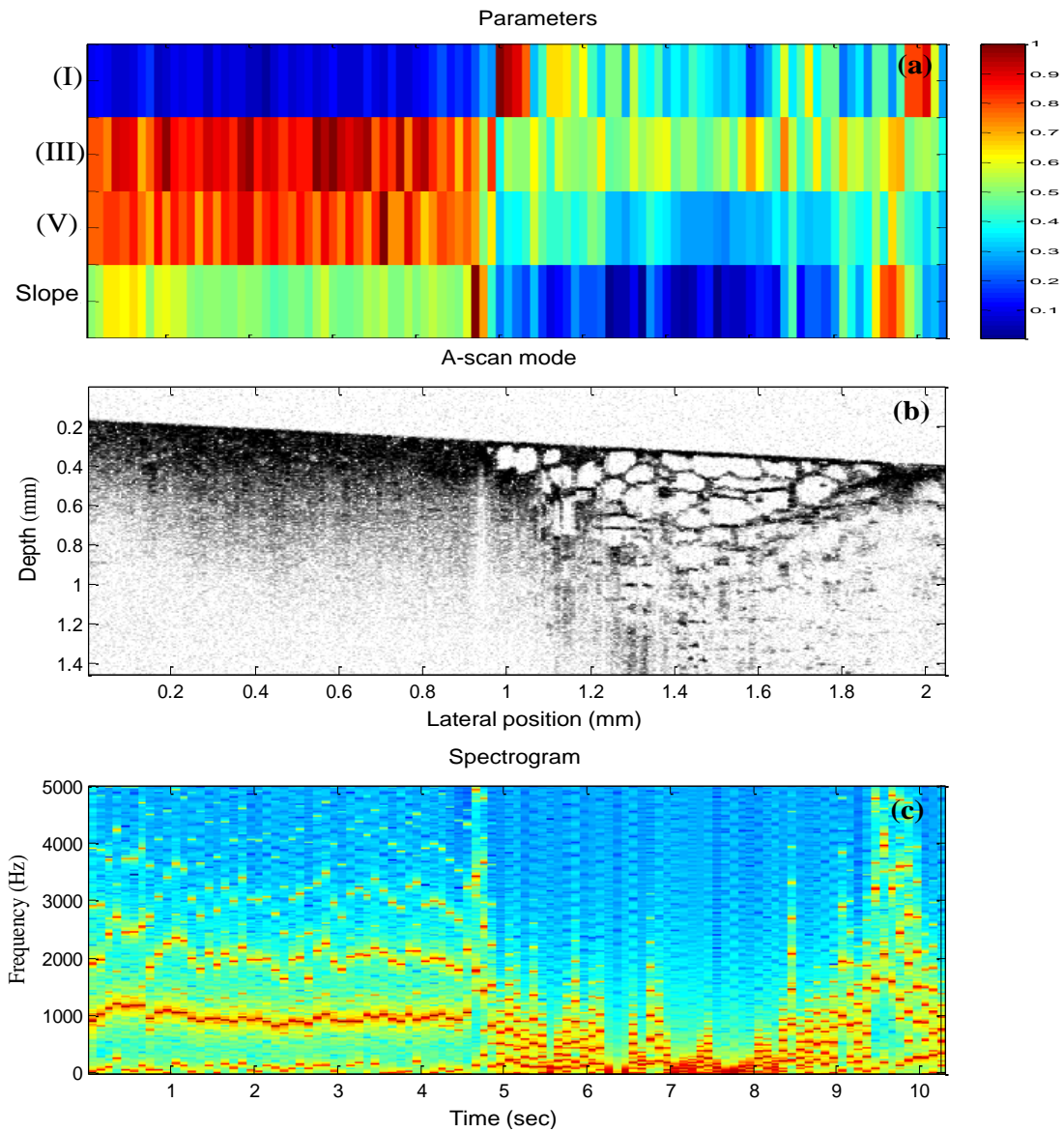

Figure 3.9: Sonification for A-scan mode. (a) Parameters extracted from image (for display purposes they have been normalized between zero and one). (b) Tissue containing tumor (left side of the image) and adipose (right side of the image). (c) Spectrogram of the output sound.

3.5.2 Image-mode

The results from image-mode are shown in Figure 3.10. The sonification is applied to a three-dimensional dataset of dimensions 1.7 mm x 3 mm x 5 mm containing both adipose and tumor tissues. The dataset was acquired intra-operatively using a 1310 nm spectral-domain OCT system with 11 μm axial and 20 μm transverse resolution. Each frame was divided into 10 blocks and sonification was done based on the scheme shown in Figure 3.8. A part of the dataset (after 30 s) is played backwards to highlight the distinction in sonification of adipose and tumor tissues.

The spectrogram in Figure 3.10(b) demonstrates that the sound of tumor has higher frequency content than that of adipose tissue. A frame consisting mainly of adipose tissue had a somewhat 'bubbly' sound at low frequencies while tumor had a more random and intense sound at high frequencies. The change of sound was especially evident at the boundary frames between tissue types. In order to test the human perceptive capability of distinguishing between adipose and tumor tissue, a frame consisting of tumor tissue was inserted (at position 37) in between 190 frames consisting of adipose tissues. If only image data is displayed, then the rapid transition of adipose-tumor-adipose may be missed if the user does not pay attention to the visual display at that particular instant of time. However, the addition of another information channel in the form of audio in conjunction with the visual display made this abrupt transition more easily noticeable during data acquisition.
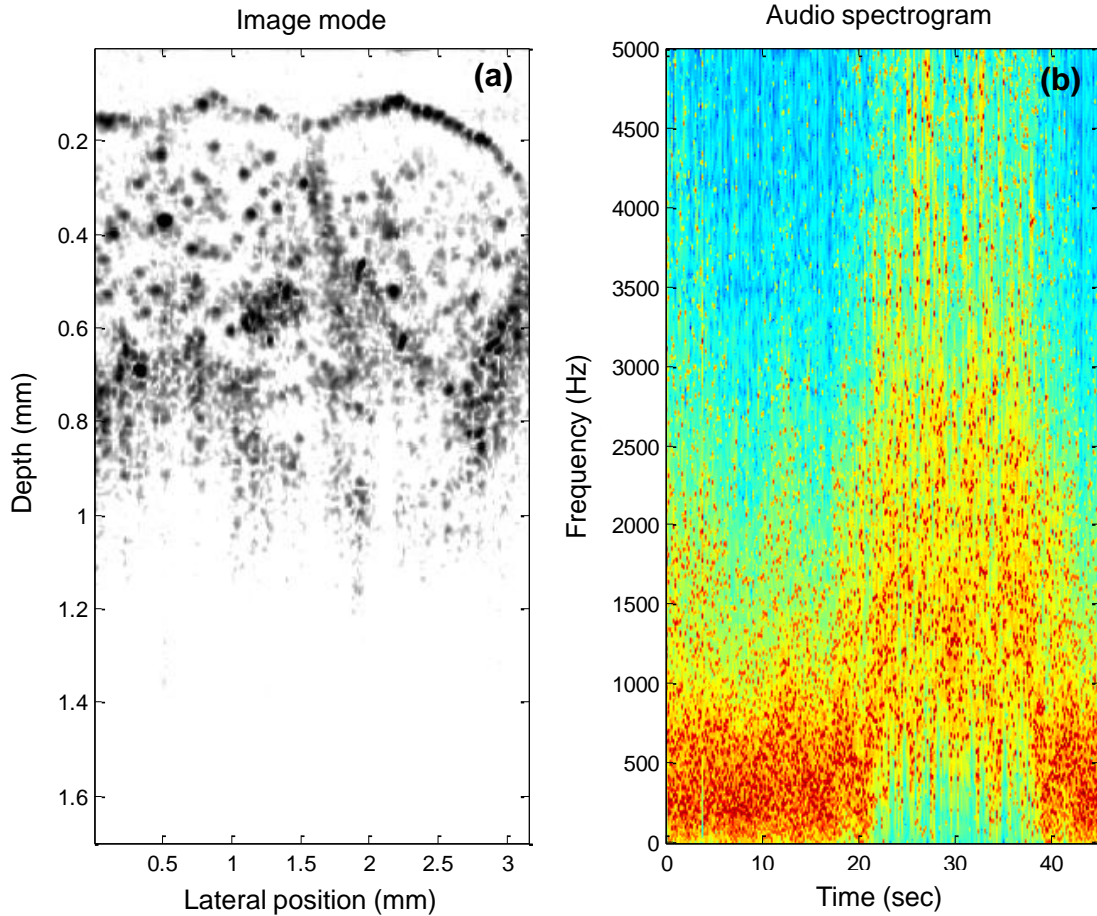
Figure 3.10: Sonification using the image mode. (a) A single frame from a three-dimensional volumetric dataset, which consists of 450 frames played at 10 frames per second. (b) Audio spectrogram of the output sound where each frame in the three-dimensional volume now corresponds to a playback time of 100 ms, and the audio spectrum from each frame is represented by a single column in this spectrogram.

3.6 Limitations

The aim of this sonification was to use the human auditory perception as a classifier to distinguish between normal and abnormal tissue types based on the sonification of OCT data. Sonification of data, however, has some fundamental drawbacks and limitations. Audio perception like all other human senses is very subjective. Different individuals will perceive the same sound very differently. Moreover, there could be potential interference from other sound sources such as speech. Another limiting factor is that sound attributes are not completely independent of each other. For example, loudness has frequency

dependence while pitch perception also depends on the intensity levels, which may cause misinterpretation of mapped data features. To compensate for these effects, the sound attributes must be carefully chosen.

One of the main challenges in sonification is finding the most efficient mapping of data parameters into sound attributes. However, currently there is no optimized approach and the sonification technique will depend to a great extent on the type and form of the data, individual perception and preference of sound, and the computational requirements. With this in mind, the current sonification scheme may not be optimum for every OCT dataset. Datasets and tissue types will have vastly different distinguishing parameters and a sonification system would need experimentation with different mappings, synthesis techniques and tuning of the parameters to customize it to the unique properties of the datasets employed.

CHAPTER 4   CONCLUSIONS AND FUTURE WORK

4.1   Summary and conclusions

This thesis describes a novel way to acquire and represent OCT data and images that may especially be useful in clinical and surgical applications of OCT. Conventional mechanical-scanning mechanisms due to their limited scan range and fixed scan geometry have limited use in a surgical setting while the sensor-based hand-held probes are somewhat inconvenient to use and have limited positional accuracy. In this thesis, a new cross-correlation based approach has been described which utilizes the structural and speckle information from the acquired A-scans to estimate the relative position of the probe. This algorithm may be used in manually scanned hand-held probes and needles for image formation. Several examples of image assembly using tissue phantoms and biological tissues were shown. However, in its current form, the technique has only been implemented for lateral manual-scanning. It is not fully automatic and requires user input to select an appropriate threshold and an area of interest to truncate the A-scans.

In addition, a new way to represent OCT data in the form of audio signals has been described. This representation may complement the more traditional visual display and enable the user to utilize multi-sensory perception capabilities for interpretation of OCT data under real-time imaging conditions, such as during surgical procedures. An estimate of the tumor location may be gauged using audio feedback, and subsequent analysis of the image data from the suspected region using tissue classification algorithms has the potential to more precisely determine locations of tumor within tissues. Sonification is to be used as an extension rather than a replacement for the traditional

visual display. Such a multi-sensory approach has the potential to improve the real-time interpretation of data during high-speed OCT imaging.

4.2   Future work

In this section some of the potential future work and possible applications of the techniques developed in this thesis are described.

4.2.1   Sensor-less manual-scanning

4.2.1.1 A-scan assembly along the longitudinal direction

Simulation results presented for A-scan assembly along the longitudinal direction need to be verified by taking actual experimental measurements from a forward-imaging device. The simulation results show that longitudinal A-scan assembly may not be trivial and may require prior assumptions about the object structure and properties. However, implementing more sophisticated correlation functions and exploring the extensive literature of time delay estimation techniques used in signal processing and communications may provide possible solutions.

4.2.1.2 Combining the cross-correlation based approach with sensor-based tracking

The cross-correlation based image acquisition technique can be combined with sensor-based positional tracking devices for more accurate image formation. These sensors can give the orientation and direction information during a scan, while the cross-correlation based approach can complement this information by deducing position at a much higher resolution than is possible with these devices.

4.2.1.3 Mosaicing

Mosaicing can be used to image over a large field-of-view by using the images acquired from hand-held scanners. Numerous image Mosaicing or registration methods have been reported based on different techniques such as cross-correlations, mutual information etc. [16]. Adapting these techniques for OCT imaging can potentially provide a more robust and practical way of large field-of-view imaging. However, the intensive computation required and the lack of structural features in many OCT images will make the adaptation of these techniques for OCT image formation very challenging.

4.2.2   Sonification

4.2.2.1 Improvements in sonification system design

Numerous investigations can be performed to improve the current method of sonification of OCT data. For example, more complex sound attributes such as vibrato and tremolo of the tones can be used. Additional sound dimensionality such as stereo, where different parameters could be mapped to the left and right ear, can be added to the current sonification. Experimentation with different mappings, and different variations in the scaling and polarity on the audio rendering can further improve performance. Moreover, depending on the datasets employed, additional data parameters based on the histograms, A-scan peaks, standard deviation (for A-scan data) or textural parameters (for image-data) can be incorporated for parameter-mapped sonification.

4.2.2.2 Human performance evaluation

Performance of human subjects also needs to be evaluated in distinguishing between tissue types based on audio feedback, visual image feedback and a combined multi-sensory feedback. For this evaluation, the present work needs to be expanded by incorporating more tissue data, including normal stroma tissue that is more optically similar to tumor tissue.

4.2.2.3 Real-time implementation

Sonification will be especially useful if done in conjunction with the acquisition of A-scans in real-time. For real-time performance, the calculation of the parameters and the subsequent mapping into sound attributes must be done faster than the data acquisition rate. A parallel implementation of the scheme given in Figure 3.8 can be used for real-time performance utilizing either commercially available sound synthesizers or using parallel programming techniques [81]. For real-time sonification, the data will either have to be downsampled or averaged. This may not be a problem as auditory feedback is intended to be a fast and efficient screening method for identification of important data features, which alert the user to suspicious areas of tissue. For more detailed recognition and visualization, the user may look at the high resolution image on the screen. The speed of real-time sonification can be increased by decreasing the playback time of sound (100 ms was used in the sonification examples in this thesis); however, it will decrease the resolution of the sonification, and clicks rather than sound tones will be heard.

4.2.2.4 Sonification of non-image data

Many applications involve the use of forward sensing devices such as for the collection of low-coherence interferometry (LCI) data or refractive index measurements. Sonification of these types of data during image acquisition may be a more conducive and intuitive method of interpretation in real-time before subsequent analysis.

REFERENCES

[1]   T. M. Peters, "Image-guidance for surgical procedures," *Phys Med Biol,* vol. 51, no. 14, pp. R505-R540, 2006.

[2]   S. A. M. Baert, E. B. van de Kraats, T. van Walsum, M. A. Viergever, and W. J. Niessen, "Three-dimensional guide-wire reconstruction from biplane image sequences for integrated display in 3-D vasculature," *IEEE Trans Med Imaging,* vol. 22, no. 10, pp. 1252-1258, 2003.

[3]   J. H. Siewerdsen et al., "Volume CT with a flat-panel detector on a mobile, isocentric C-arm: Pre-clinical investigation in guidance of minimally invasive surgery," *Med Phys,* vol. 32, no. 1, pp. 241-254, 2005.

[4]   K. Masamune et al., "System for robotically assisted percutaneous procedures with computed tomography guidance," *Comput Aided Surg,* vol. 6, no. 6, pp. 370 - 383, 2001.

[5]   A. Fenster, K. J. M. Surry, G. R. Mills, and D. B. Downey, "3D ultrasound guided breast biopsy system," *Ultrasonics,* vol. 42, no. 1-9, pp. 769-774, 2004.

[6]   H. H. Holm, and B. Skjoldbye, "Interventional ultrasound," *Ultras Med Biol,* vol. 22, no. 7, pp. 773-789, 1996.

[7]   J. Machi, and B. Sigel, "Operative ultrasound in general surgery," *Am J Surg,* vol. 172, no. 1, pp. 15-20, 1996.

[8]   G. R. Sutherland, T. Kaibara, and D. F. Louw, "Intraoperative MR at 1.5 Tesla--experience and future directions," *Acta Neurochir Suppl.,* vol. 85, pp. 21-8, 2003.

[9]   B. Albayrak, A. F. Samdani, and P. M. Black, "Intra-operative magnetic resonance imaging in neurosurgery," *Acta Neurochirurgica,* vol. 146, no. 6, pp. 543-557, 2004.

[10]   Y. C. Chung, J. L. Duerk, and J. S. Lewin, "Generation and observation of radio frequency thermal lesion ablation for interventional magnetic resonance imaging," *Invest Radiol,* vol. 32, no. 8, pp. 466-74, Aug, 1997.

[11]   Z. Yaqoob, J. Wu, E. J. McDowell, X. Heng, and C. Yang, "Methods and application areas of endoscopic optical coherence tomography," *J Biomed Opt,* vol. 11, no. 6, pp. 063001-19, 2006.

[12]   W. Drexler, and J. G. Fujimoto, *Optical Coherence Tomography: Technology and Applications*, New York: Springer, 2008.

[13]   J. M. Zara, and C. A. Lingley-Papadopoulos, "Endoscopic OCT approaches toward cancer diagnosis," *IEEE J. Sel. Top. Quantum Electron.,* vol. 14, no. 1, pp. 70-81, 2008.

[14]   F. T. Nguyen et al., "Intraoperative evaluation of breast tumor margins with optical coherence tomography," *Cancer Res,* vol. 69, no. 22, pp. 8790-8796, November 15, 2009.

[15]   A. M. Zysk, F. T. Nguyen, A. L. Oldenburg, D. L. Marks, and S. A. Boppart, "Optical coherence tomography: a review of clinical development from bench to bedside," *J. Biomed. Opt,* vol. 12, no. 5, pp. 051403-21, 2007.

[16]   W. R. Crum, T. Hartkens, and D. L. G. Hill, "Non-rigid image registration: theory and practice," *Br J Radiol,* vol. 77, no. suppl_2, pp. S140-153, December 1, 2004.

[17]   J. Culver, W. Akers, and S. Achilefu, "Multimodality Molecular Imaging with Combined Optical and SPECT/PET Modalities," *J Nucl Med,* vol. 49, no. 2, pp. 169-172, February 1, 2008.

[18]   R. L. Barbour et al., "MRI-Guided Optical Tomography: Prospects and Computation for a New Imaging Method," *IEEE Comput. Sci. Eng.,* vol. 2, no. 4, pp. 63-77, 1995.

[19]   Y. Shuai et al., "Co-registered optical coherence tomography and fluorescence molecular imaging for simultaneous morphological and molecular imaging," *Phys Med Biol,* vol. 55, no. 1, pp. 191, 2010.

[20]   R. Huber, M. Wojtkowski, and J. G. Fujimoto, "Fourier domain mode locking (FDML): A new laser operating regime and applications for optical coherence tomography," *Opt Express,* vol. 14, no. 8, pp. 3225-3237, 2006.

[21]   D. C. Adler, S.-W. Huang, R. Huber, and J. G. Fujimoto, "Photothermal detection of gold nanoparticles using phase-sensitive optical coherence tomography," *Opt. Express,* vol. 16, no. 7, pp. 4376-4393, 2008.

[22]   D. Huang et al., "Optical coherence tomography," *Science,* vol. 254, no. n5035, pp. 1178-1181, 1991.

[23] D. Stifter, "Beyond biomedicine: a review of alternative applications and developments for optical coherence tomography," *Applied Physics B: Lasers and Optics,* vol. 88, no. 3, pp. 337-357, 2007.

[24] J. M. Schmitt, "Optical coherence tomography (OCT): a review," *IEEE J. Sel. Top. Quantum Electron.,* vol. 5, no. 4, pp. 1205-1215, 1999.

[25] A. L. Oldenburg, F. J.-J. Toublan, K. S. Suslick, A. Wei, and S. A. Boppart, "Magnetomotive contrast for *in vivo* optical coherence tomography," *Opt. Express,* vol. 13, no. 17, pp. 6597-6614, 2005.

[26] P. H. Tomlins, and R. K. Wang, "Theory, developments and applications of optical coherence tomography," *Journal of Physics D: Applied Physics*, no. 15, pp. 2519, 2005.

[27] F. I. Feldchtein, V. M. Gelikonov, and G. V. Gelikonov, "Design of OCT scanners," in *Handbook of Optical Coherence Tomography*, B. E. Bouma and G. J. Tearney, eds., pp. 125-142: Marcel Dekker, Inc, 2002.

[28] P. R. Herz et al., "Micromotor endoscope catheter for *in vivo*, ultrahigh-resolutionoptical coherence tomography," *Opt. Lett.,* vol. 29, no. 19, pp. 2261-2263, 2004.

[29] J. M. Zara, S. Yazdanfar, K. D. Rao, J. A. Izatt, and S. W. Smith, "Electrostatic micromachine scanning mirror for optical coherence tomography," *Opt. Lett.,* vol. 28, no. 8, pp. 628-630, 2003.

[30] P. H. Tran, D. S. Mukai, M. Brenner, and Z. Chen, "*In vivo* endoscopic optical coherence tomography by use of a rotational microelectromechanical system probe," *Opt. Lett.,* vol. 29, no. 11, pp. 1236-1238, 2004.

[31] K. H. Gilchrist, R. P. McNabb, J. A. Izatt, and S. Grego, "Piezoelectric scanning mirrors for endoscopic optical coherence tomography," *Journal of Micromechanics and Microengineering,* vol. 19, no. 9, pp. 095012, 2009.

[32] B. Colston et al., "Dental OCT," *Opt. Express,* vol. 3, no. 6, pp. 230-238, 1998.

[33] J. Su, J. Zhang, L. Yu, and Z. Chen, "*In vivo* three-dimensional microelectromechanical endoscopic swept sourceoptical coherence tomography," *Opt. Express,* vol. 15, no. 16, pp. 10390-10396, 2007.

[34] J. Wu et al., "Paired-angle-rotation scanning optical coherence tomography forward-imaging probe," *Opt. Lett.,* vol. 31, no. 9, pp. 1265-1267, 2006.

[35] R. L. Galloway, "The process and development of image guided procedures," *Annu. Rev. Biomed. Eng,* vol. 3, no. 1, pp. 83-108, 2001.

[36] L. Mercier, T. Langø, F. Lindseth, and D. L. Collins, "A review of calibration techniques for freehand 3-D ultrasound systems," *Ultras Med Biol,* vol. 31, no. 4, pp. 449-471, 2005.

[37] Claron Technology Inc, Micron tracker, website URL: http://www.clarontech.com/measurement.php

[38] A. D. Wiles, D. G. Thompson, and D. D. Frantz, "Accuracy assessment and interpretation for optical tracking systems," in *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*, San Diego, CA, USA, 2004, pp. 421-432.

[39] B. J. Wood et al., "Navigation with electromagnetic tracking for interventional radiology procedures: a feasibility study," *J Vasc Interv Radiol,* vol. 16, no. 4, pp. 493-505, Apr, 2005.

[40] D. D. Frantz, A. D. Wiles, S. E. Leis, and S. R. Kirsch, "Accuracy assessment protocols for electromagnetic tracking systems," *Phys Med Biol,* vol. 48, no. 14, pp. 2241-2251, 2003.

[41] Ascension Technology Inc, trackSTAR, website URL: http://www.ascension-tech.com/medical/trakSTAR.php.

[42] CIVCO Medical Solutions Inc, eTRAX Needle Guidance System, website URL: http://www.civco-etrax.com/etrax/.

[43] J. Ren, J. Wu, E. J. McDowell, and C. Yang, "Manual-scanning optical coherence tomography probe based on position tracking," *Opt. Lett.,* vol. 34, no. 21, pp. 3400-3402, 2009.

[44] P. Hassenpflug, R. W. Prager, G. M. Treece, and A. H. Gee, "Speckle classification for sensorless freehand 3-D ultrasound," *Ultras Med Biol,* vol. 31, no. 11, pp. 1499-1508, 2005.

[45] T. A. Tuthill, J. F. Krucker, J. B. Fowlkes, and P. L. Carson, "Automated three-dimensional US frame positioning computed from elevational speckle decorrelation," *Radiology,* vol. 209, no. 2, pp. 575-582, 1998.

[46] S. J. Kirkpatrick, R. K. Wang, and D. D. Duncan, "OCT-based elastography for large and small deformations," *Opt. Express,* vol. 14, no. 24, pp. 11585-11597, 2006.

[47] R. J. Zawadzki et al., "Correction of motion artifacts and scanning beam distortions in 3D ophthalmic optical coherence tomography imaging," in *Ophthalmic Technologies XVII*, San Jose, CA, USA, 2007, pp. 642607-11.

[48] P. H. Tomlins, and R. K. Wang, "Digital phase stabilization to improve detection sensitivity for optical coherence tomography," *Measurement Science and Technology,* vol. 18, no. 11, pp. 3365-3372, 2007.

[49] E. S. Yeung, "Pattern recognition by audio representation of multivariate analytical data," *Anal Chem,* vol. 52, pp. 1120-1123, 1980.

[50] S. Barrass, and G. Kramer, "Using sonification," *Multimedia Systems,* vol. 7, no. 1, pp. 23-31, 1999.

[51] E. Jovanov et al., "Tactical audio and acoustic rendering in biomedical applications," *Information Technology in Biomedicine, IEEE Transactions on,* vol. 3, no. 2, pp. 109-118, 1999.

[52] G. Baier, T. Hermann, and U. Stephani, "Event-based sonification of EEG rhythms in real time," *Clinical Neurophysiology,* vol. 118, no. 6, pp. 1377-1386, 2007.

[53] M. Ballora, B. Pennycook, P. C. Ivanov, A. Goldberger, and L. Glass, "Detection of obstructive sleep apnea through auditory display of heart rate variability," in *Computers in Cardiology*, 2000, pp. 739-740.

[54] A. D. N. Edwards, G. Hines, and A. Hunt, "Segmentation of Biological Cell Images for Sonification," in *Congress on Image and Signal Processing, CISP*, 2008, pp. 128-132.

[55] A. C. G. Martins, R. M. Rangayyan, and R. A. Ruschioni, "Audification and sonification of texture in images," *J. Electron. Imaging,* vol. 10, no. 3, pp. 690-705, 2001.

[56] H. F. Routh, "Doppler ultrasound," *IEEE Eng Med Biol Mag,* vol. 15, no. 6, pp. 31-40, 1996.

[57] V. X. D. Yang et al., "High speed, wide velocity dynamic range Doppler optical coherence tomography (Part III): *in vivo* endoscopic imaging of blood flow in the rat and human gastrointestinal tracts," *Opt. Express,* vol. 11, no. 19, pp. 2416-2424, 2003.

[58] E. Jovanov, D. Starcevic, V. Radivojevic, A. Samardzic, and V. Simeunovic, "Perceptualization of biomedical data. An experimental environment for visualization and sonification of brain electrical activity," *IEEE Eng Med Biol Mag,* vol. 18, no. 1, pp. 50-55, 1999.

[59] B. D. Goldberg et al., "Automated algorithm for differentiation of human breast tissue using low coherence interferometry for fine needle aspiration biopsy guidance," *J Biomed Opt,* vol. 13, no. 1, pp. 014014-8, 2008.

[60] A. M. Zysk, D. L. Marks, D. Y. Liu, and S. A. Boppart, "Needle-based reflection refractometry of scattering samples using coherence-gated detection," *Opt. Express,* vol. 15, no. 8, pp. 4787-4794, 2007.

[61] A. Ahmad, S. G. Adie, E. J. Chaney, U. Sharma, and S. A. Boppart, "Cross-correlation-based image acquisition technique for manually-scanned optical coherence tomography," *Opt. Express,* vol. 17, no. 10, pp. 8125-8136, 2009.

[62] L. Pai-Chi, C. Chong-Jing, and Y. Chih-Kuang, "On velocity estimation using speckle decorrelation, " *IEEE Trans Ultrason Ferroelectr Freq Control,* vol. 48, no. 4, pp. 1084-1091, 2001.

[63] S. A. Boppart et al., "Forward-imaging instruments for optical coherence tomography," *Opt. Lett.,* vol. 22, no. 21, pp. 1618-1620, 1997.

[64] T. S. Ralston, D. L. Marks, P. S. Carney, and S. A. Boppart, "Inverse scattering for optical coherence tomography," *J Opt Soc Am A Opt Image Sci Vis,* vol. 23, no. 5, pp. 1027-37, May, 2006.

[65] C. Knapp, and G. Carter, "The generalized correlation method for estimation of time delay," *Acoustics, Speech and Signal Processing, IEEE Transactions on,* vol. 24, no. 4, pp. 320-327, 1976.

[66] J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *J Biomed Opt,* vol. 4, no. 1, pp. 95-105, 1999.

[67] A. Ahmad, S. G. Adie, M. Wang, and S. A. Boppart, "Sonification of optical coherence tomography data and images," *Opt. Express,* vol. 18, no. 10, pp. 9934-9944, 2010.

[68] T. Hermann, and H. Ritter, "Sound and meaning in auditory data display," *Proceedings of the IEEE,* vol. 92, no. 4, pp. 730-741, 2004.

[69] G. Kramer. Ed., *Auditory Display-Sonification, Audification, and Auditory Interfaces*, Reading, MA: Addison-Wesley, 1994.

[70]   S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "A detailed investigation into the effectiveness of earcons,"in *Auditory Display*, G. Kramer, ed., pp. 471-498: Reading, MA: Addison Wesley, 1994.

[71]   W. W. Gaver, "Synthesizing auditory icons," in *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*, Amsterdam, The Netherlands, 1993, pp. 228-235.

[72]   E. Zwicker, and H. Fastl, *Psychoacoustics - Facts and Models*, 2nd ed., Berlin: Springer, 1999.

[73]   H. F. Olson, *Music, Physics and Engineering*, Dover Publications, 1967.

[74]   B. Moore, "Psychoacoustics," in *Springer Handbook of Acoustics*, T. D. Rossing, ed., pp. 459-501: Springer, 2007.

[75]   A. M. Zysk, and S. A. Boppart, "Computational methods for analysis of human breast tumor tissue in optical coherence tomography images," *J. Biomed Opt,* vol. 11, no. 5, pp. 054015, 2006.

[76]   P. B. L. Meijer, "An experimental system for auditory image representations," *IEEE Trans Biomed Eng,* vol. 39, no. 2, pp. 112-121, 1992.

[77]   X. Qi, J. M. V. Sivak, G. Isenberg, J. E. Willis, and A. M. Rollins, "Computer-aided diagnosis of dysplasia in Barrett's esophagus using endoscopic optical coherence tomography," *J. Biomed Opt,* vol. 11, no. 4, pp. 044010-10, 2006.

[78]   K. W. Gossage, T. S. Tkaczyk, J. J. Rodriguez, and J. K. Barton, "Texture analysis of optical coherence tomography images: feasibility for tissue classification," *J. Biomed Opt,* vol. 8, no. 3, pp. 570-575, 2003.

[79]   C. Scaletti, "Sound synthesis algorithms for auditory data representations," in *Auditory Display*, G. Kramer, ed., pp. 471-498: Reading, MA: Addison Wesley, 1994.

[80]   J. M. Chowning, "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation," *Journal of the Audio Engineering Society,* vol. 21, no. 7, pp. 526-534, 1973.

[81]   H. G. Kaper, E. Wiebel, and S. Tipei, "Data sonification and sound visualization," *Computing in Science & Engineering,* vol. 1, no. 4, pp. 48-58, 1999.