SOFTWARE FRAMEWORK FOR MIDDLE EAR
OPTICAL COHERENCE TOMOGRAPHY

by

Joshua Dennis Farrell

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
July, 2023

Dalhousie University is located in Mi'kma'ki, the
ancestral and unceded territory of the Mi'kmaq.
We are all Treaty people.

To my brother, Adam, for inspiring me to think outside the box and to always question the norms. You are sorely missed.

Table of Contents

List of Tables

List of Figures

xiv

Abstract

Originally developed for ophthalmic imaging, optical coherence tomography (OCT), a non-invasive depth-resolved optical imaging modality, has recently been successfully applied to imaging of the middle ear through the intact eardrum [1]. Middle ear OCT (ME-OCT) offers real-time structural and functional images of the middle ear health at point of care safely and without exposing the patient to ionizing radiation. As the use of ME-OCT technology evolves from feasibility studies to clinical application there is a need for system-level software engineering to integrate it into clinical workflow to enable clinicians to image and visualize diagnostic information independently and in a manner that can keep up with clinical workflow.

This thesis introduces the design, implementation, verification, and validation of a flexible software framework meeting the performance and usability requirements for a ME-OCT imaging system suitable for real-world, clinical deployment. The framework provides a custom-built rendering engine capable of volumetric rendering of ME-OCT datasets at real-time interactive rates. Additionally, the rendering engine introduces two novel rendering techniques, digital tympanotomy for the real-time removal of the ear drum to visualize the underlying middle ear structures, and Doppler animation for intuitive visualization of the middle ear's acoustic response to applied stimuli. This rendering engine is integrated into a purpose-built software architecture providing a turn-key imaging experience for integration into the clinical workflow without requiring close engineering support. The architecture was extended to incorporate two novel diagnostic imaging modalities: geometrically accurate, live, continuous volumetric ME-OCT imaging, and real-time ME-OCT angiography (ME-OCTA) for the direct, non-invasive, visualization of depth-resolved middle ear vasculature and dynamics in-vivo. Residual geometrical error was assessed by co-registration corrected ME-OCT and micro-CT datasets from both a 3D printed imaging phantom and cadaveric temporal bone. While in-vivo continuous volumetric imaging was demonstrated on a healthy adult volunteer during a dynamic pressurization maneuver. Real-time, phase-sensitive, 2D and 3D ME-OCTA in-vivo imaging was demonstrated for the first time and applied to visualization of middle ear vasculature and the stapedius reflex.

Beyond ME-OCT, this research could be adapted for robotic intra-surgical applications, including surgical planning and real-time navigation, while also enabling post-surgery disease progression monitoring.

List of Abbreviations Used

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
4D	Four Dimensional
ABR	Auditory Brainstem Response
ALU	Arithmetic Logic Unit
AOM	Acute Otitis Media
API	Application Programming Interface
BOE	Biomedical Optics Express
CHL	Conductive Hearing Loss
CI	Cochlear Implant
CLAHE	Contrast-Limited Adaptive Histogram Equalization
CLV	Constant Linear Velocity
COM	Chronic Otitis Media
CPU	Central Processing Unit
CSOM	Chronic Suppurative Otitis Media
CT	Computed Tomography
CUDA	Compute Unified Device Architecture
CVL-SC	Constant Linear Velocity Spiral Scanning
DAC	Digital-to-Analog Converter
DAQ	Data Acquisition Card
DC-SC	Discretized Spiral Scanning
DICOM	Digital Imaging and Communication in Medicine
DRAM	Dynamic Random-Access Memory
ENT	Ear, Nose, and Throat
ET	Eustachian Tube
ETD	Eustachian Tube Dysfunction
FD-OCT	Fourier Domain Optical Coherency Tomography
FFT	Fast Fourier Transform
FOV	Field Of View
FPS	Frames Per Second
GPU	Graphics Processing Unit

GRIN	Gradient Index
GUI	Graphical User Interface
HMVC	Hierarchical Model View Controller
HRCT	High-Resolution Computed Tomography
IMJ	Incudomallear Joint
ISJ	Incudostapedial Joint
JBO	Journal of Biomedical Optics
LUT	Lookup Table
MDL	Motorized Delay Line
MEMS	Micro Electromechanical System
ME-OCT	Middle Ear Optical Coherence Tomography
ME-OCTA	Middle Ear Optical Coherence Tomography Angiography
Micro-CT	Micro-Computed Tomography
MRI	Magnetic Resonance Imaging
MVC	Model View Controller
OAG	Optical Angiography
OCT	Optical Coherence Tomography
OCTA	Optical Coherence Tomography Angiography
OCT-DV	Optical Coherence Tomography Doppler Vibrometry
OM	Otitis Media
OMAG	Optical Microangiography
OME	Otitis Media with Effusion
OpenCV	Open-Source Computer Vision Library
OpenGL	Open Graphics Library
PACS	Picture Archiving and Communication System
PCA	Principal Component Analysis
PET	Patulous Eustachian Tube
PORP	Partial Ossicular Replacement Prosthesis
PSF	Point-Spread-Function
RGB	Red, Green, Blue
RGBA	Red, Green, Blue, Alpha
RMM	RAPIDS Memory Manager
RMSE	Root Mean Square Error

ROI	Region Of Interest
RPCA	Robust Principal Component Analysis
SD-OCT	Spectral Domain Optical Coherence Tomography
SIMD	Single Instruction, Multiple Data
SIMT	Single Instruction, Multiple Thread
SLSQP	Sequential Least Squares Programming
SM	Streaming Multiprocessors
SNHL	Sensorial Neural Hearing Loss
SNR	Signal-to-Noise Ratio
SP	Streaming Processors
SSADA	Split-Spectrum Amplitude-Decorrelation Angiography
SS-OCT	Swept-Source Optical Coherence Tomography
STFT	Short-Time Fourier Transform
TBB	Thread Building Blocks
TD-OCT	Time Domain Optical Coherence Tomography
TM	Tympanic Membrane
TOF	Time-Of-Flight
TORP	Total Ossicular Replacement Prosthesis
UI	User Interface
UML	Unified Modelling Language
UV	Ultraviolet
VCSEL	Vertical Cavity Surface-Emitting Laser
VT-DBR	Vernier-Tuned Distributed Bragg-Reflector

Acknowledgements

Now that my long journey of completing my PhD in biomedical engineering, here at Dalhousie University, is ending I would like to acknowledge and thank all of those that have offered unwavering support, encouragement, and guidance over the course of this academic adventure.

To my loving and beautiful wife, Rose Panikulam, whose constant love, and support has been instrumental in the completion of my PhD journey. Thank you for helping shoulder the burden of all the trials and tribulations over these many years, for sticking with me through the ups and downs, and for always believing in me. This degree is just as much your degree at the end of the day.

To my supervisor, Dr. Robert Adamson, for the opportunity of a lifetime and his unwavering commitment and guidance towards helping me finishing my degree. Thank you for the unparalleled mentorship, growth opportunities, and lessons you have offered me throughout this experience. I am very fortunate to have you as my supervisor and would not have it any other way.

To my committee members, Dr. David P. Morris, and Dr. Jeremy Brown, in generously donating their time, resources, and energy over the years such that I may grow as a researcher and pushing me to be the best I could be. Thank you both for the invaluable insights, and rationales you have given me within the fields of Medicine and Biomedical Engineering, and for your endless patience as you listen to me ramble on about software engineering for hours on end. It was a privilege and an honour to have collaborated with such knowledgeable, amazing, mentors.

To all the friends and colleagues whom I have been fortunate enough to work alongside of throughout this degree who have shared both their time and skills in the completion of my thesis work.

I thank my older brother Matthew Farrell for his invaluable contributions in the completion of the clinical user interface for the software architecture. Your attention to detail, design instincts, and ability to discuss and breakdown software problems has and always will be an inspiration. I could not have done this without you, my dude.

Chapter 1

Introduction

This thesis details the design and development of a real-time, volumetric, software architecture for optical coherence tomography (OCT) of the human middle ear for clinical otology at the point-of-care and its extension to include advanced OCT imaging modalities. The user-friendly, real-time software developed using the architecture lowers the barrier for clinical adoption of middle ear OCT (ME-OCT) to provide clinical otologists with novel tools and insights into middle ear diagnostics and pathologies. This introductory chapter presents a review of middle ear OCT (ME-OCT) imaging, and challenges faced in its application to clinical otology. Finally, this chapter presents the format of this thesis work and provides a summary of the subsequent chapters.

1.1 Middle Ear Optical Coherence Tomography

Over the past three decades, OCT has become a standard-of-care diagnostic technology in ophthalmology, providing real-time structural and diagnostic information on pathologies of the retina [2]. Beyond ophthalmology, OCT has been extensively investigated for a wide variety of other clinical imaging applications including: larynx [3], skin [4], esophagus [5], lung [6], and nose [7]. The feasibility of ME-OCT was first demonstrated in 2001 by Pitris et al. [8]. However, clinical development was slowed due to limitations in the TD-OCT technology of the period in terms of speed, sensitivity and imaging depth [9]. It was only with the advent of swept-source OCT technologies combined with highly phase stable swept laser sources such as the Vernier-tuned distributed Bragg-reflector (VT-DBR) akinetic swept laser [10] in the 2010s that real-time in-vivo ME-OCT become clinically viable [1].

In addition to structural information, SS-OCT can provide functional information on the response of the middle ear to sound through a method known as OCT Doppler vibrometry (OCT-DV). Applying a sound stimulus into the ear canal introduces a Doppler shift in the detected inference signal. This Doppler shift manifests as an optical phase shift

between A-lines acquired sequentially in time with the phase shift modulated at the acoustic stimulus frequency and with an amplitude proportional to the structure's acoustic mobility. The motion of middle ear structures can be extracted from these A-lines by cross-correlating the acoustic phase information with a complex phasor at the known stimulus frequency or by Fourier transformation of the A-line phase. By performing this acoustic phase analysis at every pixel along a line, the depth-resolved magnitude and phase of the acoustic response of the middle ear to the applied sound stimulus can be measured [1].

Several groups have demonstrated the potential of ME-OCT as a new diagnostic tool for clinical otology [11,12]. Applications investigated include classification and characterization of otitis media (OM) [13,14], postoperative evaluation of tympanoplasty success [15], TM thickness measurement [16], delineation of cholesteatoma [17], discrimination of otosclerotic ears and healthy ears [18], and modelling the vibrational modal response of the TM and ossicles [19,20]. This thesis focuses on the efforts of our lab over the last decade to translate ME-OCT to clinical use [1,9,18] with the aim of improving on current standard-of-care diagnostic and imaging technologies while fulfilling the need for a non-invasive, point-of-care, diagnostic imaging system for assessing middle ear disease and conductive hearing loss.

1.1.1 Optical Coherence Tomography Imaging System for Clinical Otology

Obtaining clear OCT images of the middle ear is challenging as the middle ear lies at the end of a 3 cm long ear canal that is often narrow and curved. Because of this anatomical limitation, ME-OCT lends itself well to an entocentric imaging geometry where a wide angular FOV is imaged from a small aperture so as to capture the entire ~10 mm diameter of the TM in a single acquisition [1,19,21].

Image data in this work was collected using a custom OCT handpiece specifically designed to navigate the anatomical restrictions of the ear canal to image the middle ear through the tympanic membrane (Fig. 1c–e). The OCT imaging optics contained in this handpiece consisted of a two-axis, 3.6 mm diameter micro electromechanical system (MEMS) mirror (Mirrorcle A8L1.1, USA) [22] used to scan an OCT beam over $\pm 6.2°$ and a 4-lens pupil relay that imaged the mirror axis onto the entrance pupil of a graded index (GRIN) rod lens that served as the objective. The GRIN lens imaged the mirror axis onto its exit pupil so that the exit pupil acted as a point of convergence for OCT scan lines. The

entocentric design of the optics enabled scanning over $30°$ in both lateral dimensions. Following the manufacturer's recommendations [23], the $3\,dB$ bandwidth of the signal used to drive the MEMS mirror was limited to 200 Hz on each axis. The control voltages for the MEMS mirrors were generated using a 10-bit digital-to-analog converter (DAC) and amplified by a high voltage amplifier to deliver a drive signal to the mirrors of $155\,V_p$. This imaging handpiece was coupled to a custom SS-OCT engine (Fig. 1a and Fig. 1b) consisting of an akinetic swept source laser (Insight Photonics SLE-1001) with a center wavelength of $1550\,nm$, a sweep range of $35.4\,nm$, a nominal sweep rate of $200\,kHz$, and a $3dB$ axial and lateral resolution in air of $40\,\mu m$ and $40\,\mu m$ respectively. While the coherence length of the swept laser permitted scanning over as much as 200 mm in depth[10], we limited the scan range to 10.9 mm [9] which was adequate to image the full depth range of the human middle ear from the most lateral aspects of the tympanic membrane to the most medial visible parts of the cochlear promontory. A motorized delay line (MDL) was inserted into the reference arm to provide a variable delay that could move the image plane axially to accommodate for differences in patient ear canal length.

The work presented in this thesis consists of a series of innovations in the software used to control this system, to acquire and process data and to present the resulting images and diagnostic information to clinicians.
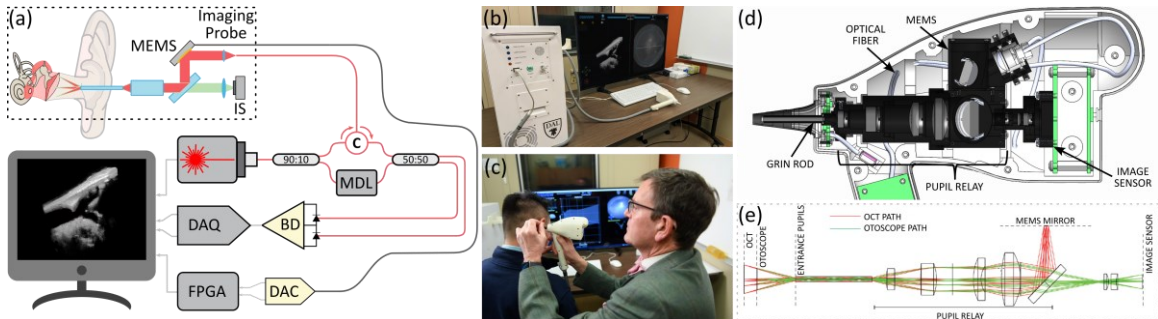
Fig. 1:   Middle ear optical coherence tomography (ME-OCT) imaging system. (a) ME-OCT system schematic. (b) ME-OCT clinical system. (c) ME-OCT imaging at the point-of-care using a handheld imaging probe with an otoscopic form factor. (d) Computer-aided design (CAD) cross-sectional view of the handpiece. (e) Optical model of the handpiece's entocentric imaging optics. IS, image sensor; MDL, motorized delay line; DAQ, data acquisition card; FPGA, field programmable gate array; DAC, digital to analog converter; BD, balance detector. C, three-way fiberoptic circulator.

1.2   Thesis Summary

This thesis is comprised of seven chapters including introduction, background, and conclusion chapters. A brief overview of each chapter in the main body of the thesis is

provided here to give context to the work. More thorough motivations for each research study are provided in their respective chapters. Overall, this thesis represents the significant contributions of the author in translating ME-OCT from the research lab to the clinic. The presented improvements to software lowered the barrier for ME-OCT to be used in the clinic, at the point-of-care, and provided improved diagnostic capabilities through geometrically accurate 4D imaging and real-time OCT Angiography (OCTA) of the middle ear giving clinicians new diagnostic tools to explore a wide range of pathologies.

Chapter 2 presents a detailed technical background and literature review on middle ear anatomy, physiology, pre-surgical diagnosis, pathology, OCT, graphics processing unit (GPU) architecture, and ME-OCT signal and image processing. This chapter explores current techniques and research in each field, identifies future research avenues, and lays the groundwork for the inquiries and discussions in subsequent chapters.

Chapter 3 details the design and implementing of a real-time 3D renderer for ME OCT image data. The renderer uses GPU-accelerated single-pass ray-casting to deliver faithful renders of ME-OCT volumetric data (Fig. 26), suitable for various in-vivo applications and real-time visualization and manipulation. The rendering engine enabled novel visualization techniques such as digital tympanotomy and volumetric Doppler animation. Notably, as shown in Fig. 27 and Supplementary File B, digital tympanotomy showed potential in producing images akin to those from conventional otomicroscopy during surgical exploration. Additionally, the renderer's ability to animate and depict the acoustic response of the middle ear to pure-tone sound stimuli was successfully demonstrated using a ME-OCT Doppler vibrometry dataset from a cadaveric temporal bone (Fig. 28 and Supplementary File C). Offering a more intuitive visualization of pathological middle ear function at the point-of-care. The development and application of a reliable, real-time, volumetric rendering engine for ME-OCT data visualization and manipulation was a key component of several subsequent advancements detailed in this thesis, including real-time 2D and 3D ME-OCT angiography (ME-OCTA) and geometrically accurate 4D ME-OCT imaging.

Chapter 4 outlines the creation of a real-time software framework, integrating the rendering engine discussed in Chapter 3, to enhance our ME-OCT imaging system's utility at the point-of-care. This GPU-accelerated, multithreaded processing framework,

combined with a clinically focused user interface (UI), allows for 2D B-mode processing at rates well in excess of 20 FPS while maintaining a responsive user interface and running multiple concurrent imaging pipelines without sacrificing performance. As shown in Fig. 46 and Supplementary File E, the architecture not only aligns with our design requirements for real-time clinical imaging, Section 4.2.1, but also serves as a basis for advanced imaging modalities like geometrically accurate 4D and real-time ME-OCTA imaging covered in Chapter 5 and Chapter 6 respectively. This thesis work has supported various clinical studies in our lab [24,25] exploring the applications of ME-OCT to post-surgical follow-up and intraoperative cochlear implants imaging.

Chapter 5 introduces a novel discretized spiral scanning (DC-SC) protocol and a real-time geometric correction model to facilitate geometrically accurate 4D ME-OCT imaging in-vivo. By integrating DC-SC with real-time volumetric scan conversion and lateral distortion correction, the system could achieve resolution-limited geometric accuracy over 92% of a $30° \times 30° \times 10.9$ mm FOV as shown in Fig. 52. Geometrical accuracy was verified and validated through the co-registration of corrected OCT and micro-CT datasets with a 3D printed imaging phantom (Fig. 53) and cadaveric temporal bone (Fig. 54) respectively. The imaging system incorporating geometric correction demonstrated a geometric accuracy of 52 μm to 75 μm (Table 4) for measuring anatomical distances, surpassing conventional clinical CT and closely matching resolutions commonly found in surgical stereomicroscopes. While ME-OCT cannot replace clinical CT because of its limited penetration depth in tissue, it can complement it, offering images of the middle ear areas visible through the tympanic membrane with excellent geometric fidelity and soft tissue contrast, potentially enhancing CT images or reducing the number of CT images needed to track longitudinal changes in middle ear pathology over time. Finally, 4D ME-OCT imaging was used to visualize the real-time structural changes occurring in the middle ear of a healthy adult volunteer during a Valsalva maneuver. Volume rates of 0.22 vol/s over the full $30° \times 30° \times 10.9\ mm$ field of view FOV and 3.56 vol/s over a limited $7.5° \times 7.5° \times 10.9\ mm$ FOV were achieved as shown in Supplementary File H. The ability to obtain accurate, real-time 4D images of the middle ear anatomy provides a unique ability to image dynamic middle ear processes and may be important for future applications such as surgical guidance and robotic navigation.

Chapter 6 covers the implementation of phase-variance OCTA algorithms to image middle ear vasculature. OCTA techniques originally developed for ophthalmic imaging were adapted to address unique challenged created by the anatomy and physiology of the middle ear. For example, one notable challenge arises from the differential bulk motion present between the tympanic membrane and underlying middle ear structures, arising from factors like changes in ear canal pressure or pressure on the canal walls caused by the handpiece during imaging. By applying a novel depth-dependent localized bulk motion correction, we obtained 3D angiograms (Fig. 56) of the middle ear showing recognizable vascular networks for the first time. Middle ear OCTA has potential in numerous clinical applications such as assessing graft vascularization, predicting ossicular necrosis risks, and directly visualizing the stapedius reflex as demonstrated for the first time in Fig. 59 and Supplementary File I. As the adoption for middle ear OCT grows, middle ear OCTA techniques could offer valuable diagnostic insights to otolaryngologists, aiding in the treatment and diagnosis of diverse middle ear diseases.

Chapter 2

Background

2.1 Anatomy and Physiology of the Middle Ear

The middle ear, Fig. 2, is a well vascularized, mucous lined, air-filled cavity containing a complex three-dimensional dynamic structure responsible for the mechanical transmission of sound from the external auditory canal (or meatus) to the cochlea of the inner ear. Housed within the temporal bone on both sides of the skull [26], the middle ear consists of the tympanic membrane (TM), the ossicular chain, and supporting middle ear muscles and ligaments. The TM is a thin, cone-shaped, membrane $\sim 100\ \mu m - 380 \mu m$ thick [27] that separates the external auditory canal from the middle ear space which couples sound to the ossicular chain. The ossicular chain is comprised of three of the smallest bones in the human body: the malleus, incus, and stapes. The malleus attaches to the TM and the stapes attaches to the oval window of the cochlea with the incus connecting the two. Average dimensions of the malleus, incus and stapes are $8.0\ mm \times 2.7\ mm$, $6.8\ mm \times 5.3\ mm$, and $3.5\ mm \times 2.4\ mm$ respectively [28]. The area difference between the TM and oval window serves to increase the pressure seen at the input to the cochlea relative to the sound pressure in the ear canal. Further pressure gain is obtained through lever-like action of the ossicles generated by the differences in the length of the long processes of the malleus and incus [29]. Two middle ear muscles, the stapedius connected to the stapes superstructure and the tensor tympani connected to the malleus, serve to modulate acoustic pressure through the ossicular chain to protect against overstimulation of the inner ear, reduce autophony, and mask out background noise [30]. Called the stapedius reflex, both the stapedius and tensor tympani contract in response to loud noises, stiffening the ossicular chain to increase middle ear impedance and reduce sound transmission to the inner ear [30].

Conductive hearing loss (CHL) occurs when the normal amplification of sound through the middle ear is disrupted either through a fixation or discontinuity of the ossicular chain or perforation of the TM. CHL is distinct from sensorial neural hearing loss (SNHL)

7

which is associated with disruption of the inner ear processes that convert sound vibration into neuroelectric signals [31]. With a prevalence of 3.8% [32], and a 5-year incidence of 2.3% [33] CHL constitutes a substantial health and quality of life burden on hundreds of millions of adults worldwide [34]. Fortunately, CHL is often treatable with surgical intervention, however, accurate pre-surgical diagnostics are essential for both surgeons in selecting appropriate treatment options and for patients to be able to provide informed consent to those treatments.
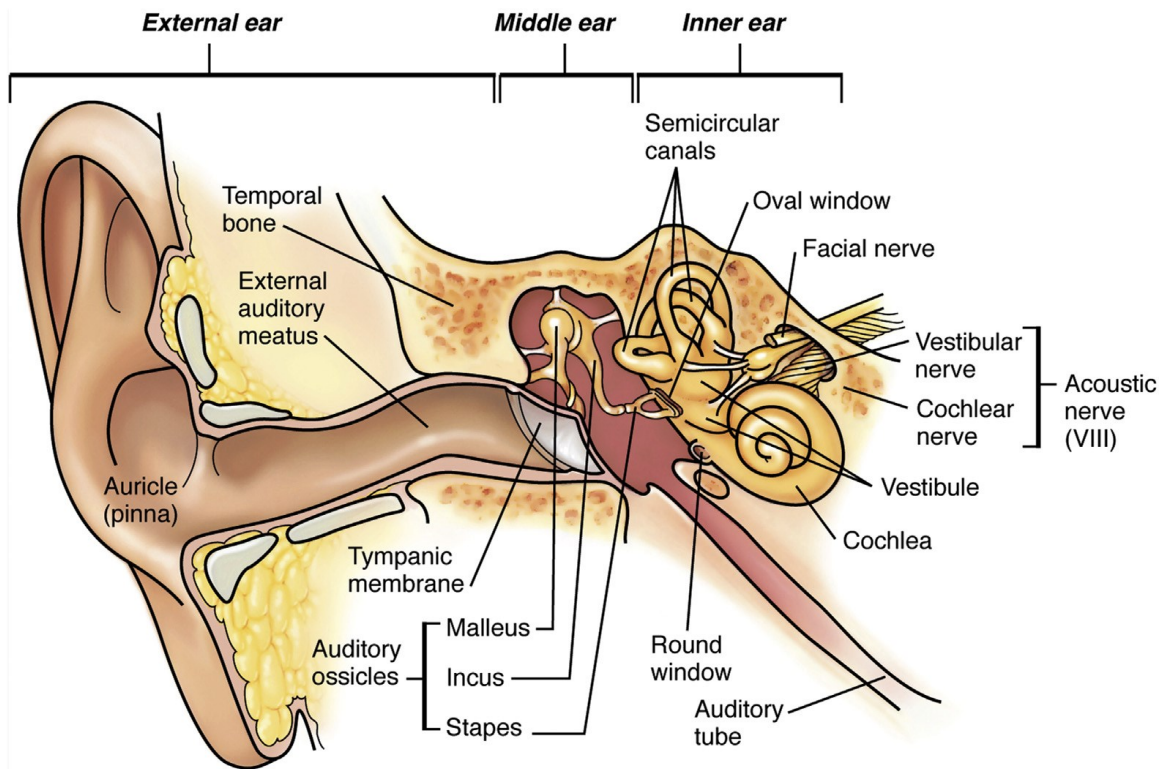
Fig. 2:   *Anatomical overview of the ear. From [29] and reproduced with permission.*

2.2   Current Standard-of-Care Pre-Surgical Diagnostics

Traditional pre-surgical diagnosis of ear disease is based on three pillars: (i) patient history, (ii) otomicroscopy, and (iii) audiometry [35]. Radiological imaging is also employed to provide additional diagnostic information on suspected pathologies. The following section will provide a review of the most common diagnostic tools used in otology.

2.2.1   Otomicroscopy

Otomicroscopy is a routine, non-invasive, examination of the ear using optical imaging instruments such as otoscopes, surgical microscopes, and endoscopes. Otomicroscopy can typically detect diseases affecting the external auditory canal and TM (Fig. 3a). However,

assessment of the underlying middle ear is difficult due to the opacity of the TM. Only when the TM is perforated, or when it is cut and lifted away during surgery, can the underlying structures of the middle ear be clearly visualized (Fig. 3b).
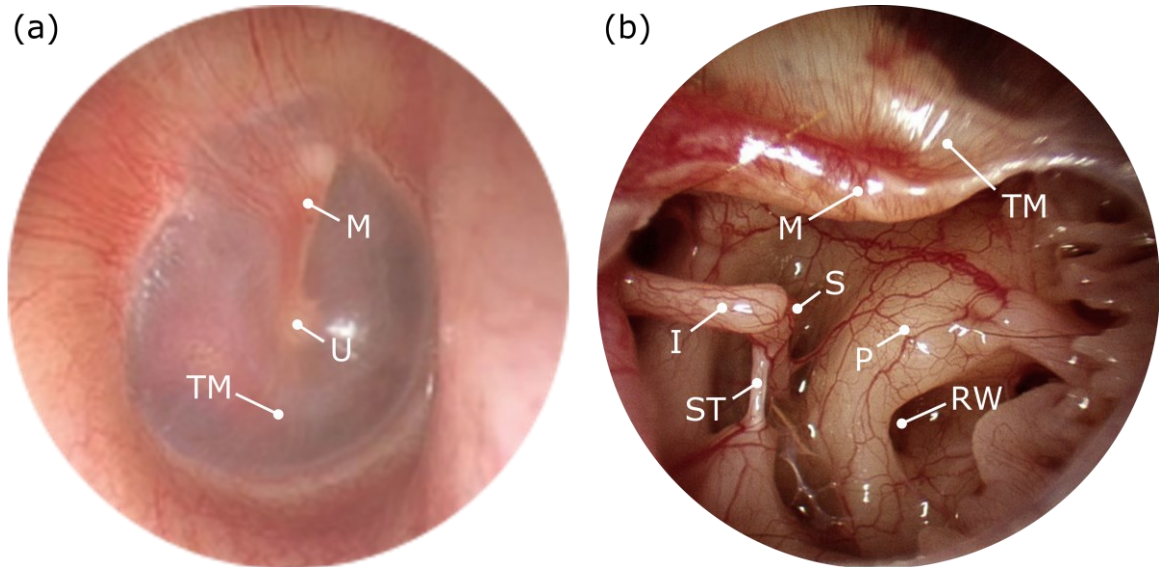
Fig. 3:   Otomicroscopic views of the middle ear. (a) Otoscopic view of the intact TM. (b) Endoscopic view of the middle ear as seen during surgery. Derived from [36] and reproduced with permission. M, malleus; U, umbo; TM, tympanic membrane; S, stapes; I, incus; ST, stapedius tendon; P, promontory; RW, round window.

Otoscope

An otoscope is a small handheld instrument consisting of a light source and magnifying lens used to visualize the external ear canal and TM. A disposable plastic speculum, attached to the distal end of the otoscope, is used to help navigate the instrument through the ear canal to obtain a clearer view of the underlying ear structures. Otoscopes are routinely used during ear examinations and provide invaluable information for treatment planning.

Otoscopes are an important diagnostic tool for many middle ear diseases such as cholesteatoma, otosclerosis, tympanosclerosis, perforation, and fibrosis. For cholesteatoma, otoscopy is crucial for evaluating the presence of cholesteatoma where a retraction pocket or foul-smelling discharge is a hallmark of its presence [37]. In otosclerosis, the presence of a reddish hue over the promontory is a clear indicator of the pathology under otoscopy [38]. Additionally, otoscopes function as a key step for the diagnosis of ear infections such as otitis media with effusion (OME) [39]. However, while otoscopes have a high sensitivity (89.7%) in diagnosing OME, they have low specificity

(71.4%) [40]. This is due to several limitations associated with otoscopes including limited field of view (FOV), poor illumination, and obstructions of the ear canal preventing clear lines-of-sight. Besides pneumatic otoscopy, it provides no functional imaging information and is primarily used for visually inspecting middle ear structure.

Surgical Microscope

Binocular surgical microscopes are the gold-standard visualization instrument for ear examination. Surgical stereomicroscopes are typically used in conjunction with a larger ear speculum, separate from the stereomicroscope, inserted into a patient's ear while they are supine, and their head tilted. This provides increased stability over the otoscope along with improved illumination, FOV and, unlike the otoscope, a sense of depth perception. As the microscope is mounted to an articulating arm, the clinicians' hands are free, allowing more free-range over the examination procedure. As compared to otoscopes, stereomicroscopes can also incorporate better optics that provide more magnification. Modern binocular surgical microscopes can have magnifications up to 40X and resolutions of $\sim 35 \ \mu m$ [41]. As a result, microscopes have a higher accuracy in diagnosing middle ear pathology as compared to otoscopes [42]. For example, microscopes have been shown to have an improved sensitivity (94.4%) and specificity (93.8%) in diagnosing OME [43].

Endoscope

Over the past few years, the use of rigid endoscopes has become increasingly common in both the clinic and in the operating room. As compared to stereomicroscopes, endoscopes offer a wide-angle FOV, improved illumination and a greater depth of field for imaging small middle ear recesses that allows for the direct visualization of many middle ear pathologies including retraction pockets or perforations [44]. The use of endoscopes in the clinic has generated a measurable improvement in patient-centered care and surgical outcomes when used in conjunction with conventional surgical stereomicroscopes [45].

2.2.2 Audiometry

Audiometry is comprised of two distinct classes: subjective and objective. Subjective audiometry is a qualitative measure of a patient's hearing performance by testing their reaction to various acoustic stimuli. Although widely used, subjective audiology inherently has a low specificity for most causes of conductive hearing loss [46]. Subjective

audiometry tools include tuning fork tests, and pure-tone audiometry. Other subjective audiometry tools such as cochlear recruitment tests and speech audiometry focusing on sensor and neurological function are outside the scope of this work and will not be discussed [46,47]. Objective audiometry is a quantitative measure of a patient's hearing performance by direct observation of physiological reaction to applied acoustic stimulus. Objective audiometry tools include tympanometry, stapedius reflex tests, auditory brainstem response (ABR), otoacoustic emissions, and electrocochleography. Although in this work we will only be focusing on the first two tools mentioned. Since these physiological reactions are the combined result of the entire hearing system the accuracy of objective audiometry is inherently limited [46]. Objective audiometry is best used in conjunction with subjective audiometry to provide extra information for differential diagnostics and is not a reliable predictor of pathology on its own [46,48].

Tuning Fork Tests

Tuning fork tests are the simplest subjective measure that can be done with relatively simple tools and serve to help orientate surgeons on possible pathologies. While these tests cannot specifically quantify the amount of hearing loss they can help localize and distinguish the type of hearing loss. Two common tuning fork tests are: the Weber test and the Rinne test. The Weber is only applicable to unilateral hearing loss and works by placing a vibrating tuning fork on the skull of the patient. If the patient has SNHL the sound is perceived in the healthy ear. However, if the sound is perceived in the unhealthy ear, then CHL is suspected. This phenomenon is a result of the tuning fork signal being unimpeded by the disrupted conduction pathway of the middle ear (i.e., more signal reaches the inner ear) [46]. The Rinne test provides a measure of sound perception between bone and air conduction. Here the tuning fork is placed on the patient's mastoid (behind their ear) and is moved in front of the patient's ear when they can no longer perceive the tuning fork via bone conduction. If the tuning fork can be heard from the front of the ear (via air conduction) then SNHL is suspected and if it cannot then CHL is suspected [46]. However, both the Weber and Rinne test are only valid in cases where a mixture of SNHL and CHL is not present [46].

11

Pure-Tone Audiometry

Air-bone pure-tone audiometry is considered the gold standard audiometric test and is used to assess the level and type of hearing loss (SNHL and/or CHL). Pure-tone audiometry has a sensitivity of 91.4% and a specificity of 93.5% in a detecting SNHL [49]. Pure-tone audiometry plots the patient reported hearing thresholds for both air and bone conduction for frequencies between $125\ Hz - 8\ kHz$ and $250\ Hz - 4\ kHz$ in octaves (doubling of each subsequent frequency step) respectively. For bone conduction, frequencies below $250\ Hz$ cannot be reliably measured as its tactile sensation interferes with its acoustic perception, and at frequencies above $4\ kHz$ the bone oscillator can produce audible sound interference [46]. This plot of measured thresholds is called an *audiogram* where each threshold is measured in $dBHL$ (hearing level) with the reference level ($0\ dBHL$) calibrated per patient population. Hearing loss is observed if any of the plotted thresholds exceed $20\ dBHL$ [47]. The difference between the air and bone thresholds is called the *air-bone* gap. The presence of an air-bone gap estimates the amount of CHL present and defines the maximum improvement possible through surgical intervention [47]. Audiogram shape can be an indicator of the type of disease affecting hearing. For example, an elevated air-bone gap at low frequencies could suggest Meniere's disease while an elevated air-bone gap at high frequencies could indicate noise induced hearing damage [47]. However, audiograms lack the accuracy to distinguish the underlying pathologies causing hearing loss and only provide clinically suggestive findings [50,51].

Tympanometry

Tympanometry is used to quantify compliance of the middle ear by measuring acoustic admittance which is the ratio of the acoustic volume velocity in the external auditory canal to the applied acoustic pressure. Admittance is the inverse of impedance and is indicative of the ease with which sound is transmitted through the middle ear. A high admittance corresponds to higher mobility of the TM and ossicular chain, whereas a low admittance corresponds to a lower mobility.

In tympanometry, the external auditory canal is pressurized and stimulated with an acoustic tone of $226\ Hz$ from a calibrated sound source with a known output admittance. A microphone in the external auditory canal measures the pressure generated by the sound source and the volume velocity from which the acoustic admittance is calculated. This

measurement is repeated as the pressure is varied over a range of a few hundred decapascals ($daPa$), or equivalently mmH_2O where $1\ mmH_2O$ equates to ${\sim}0.981\ daPa$, centered around nominal atmospheric pressure. The resultant plot of admittance vs pressure is called a *tympanogram*. The pressure range used in tympanometry depends on the suspected underlying pathology where a range of $\pm 200\ mmH_2O$ is used in healthy ears whereas an extended negative pressure range is used in ears with a retracted TM [48].

Three parameters of a tympanogram used to draw clinical conclusions of underlying pathology are its shape, location of peak admittance, and the value of this peak admittance. Each different shape of a tympanogram is given a type with each type being associated with a suspected pathology. A type *A* tympanogram corresponds to a healthy ear where the peak admittance is located around nominal atmospheric pressure of $0\ daPa$. A flattening of this shape is a type *B* tympanogram and often indicates fluid in the middle ear. A negative pressure peak is a type *C* tympanogram and indicates negative middle ear pressure possibly due to Eustachian tube dysfunction [47]. Numerous different tympanogram types, besides just the ones previously listed, have been associated with different underlying causes of CHL [52]. Unfortunately, since tympanogram measurements are dominated by the admittance of the TM, which varies greatly between individuals, rather than the ossicular chain, tympanometry has low specificity for ossicular disorders [46,53]. Thus, tympanometry is only considered diagnostically suggestive for middle ear pathologies apart from otitis media [54].

Stapedius Reflex Test

Another use for tympanometry is the detection of the stapedius reflex. The stapedius reflex increases the impedance of the middle ear in response to a loud sound stimulus. The triggering pathway for the stapedius reflex involves an interaction between the middle and inner ear, and the facial and hearing nerves of the brain. Since the nerves involved in this pathway innervates both ears, the stapedius reflex causes bilateral impedance changes even when sounds are presented unilaterally. Thus, the stapedius reflex can be triggered either ipsilaterally or contralaterally [46].

To measure the ipsilateral stapedius reflex, an acoustic tone at frequencies of $500\ Hz$, $1\ KHz$, $2\ kHz$, or $4\ kHz$ is played through the tympanometry probe using a reflex triggering amplitude between $70\ dBHL$ and $110\ dBHL$ for $1\ second$. For measurement of

the contralateral stapedius reflex a headphone is used to simulate the ear opposite to the tympanometry probe. Both measurements are done using a static ear canal pressure where middle ear admittance is maximized [46]. The reflex threshold is found by gradually increasing the stimulus amplitude from $70\ dBHL$ until the reflex is triggered as determined by an observed increase in middle ear impedance.

The level at which the stapedius reflex is triggered depends on loudness of the applied stimulus relative to the patient's hearing threshold. For cases of CHL the reflex threshold increases by the amount of hearing loss present. For SNHL the reflex threshold is constant up to a reported hearing loss of $50\ dB$ and increases linearly with further hearing loss [46]. An absent reflex may hint at possible neurological disfunction in the absence of any detectable hearing loss.

2.2.3 Radiological Imaging

Current standard-of-care non-invasive imaging technologies used in clinical otology include high-resolution computed tomography (HRCT) and magnetic resonance imaging (MRI). For preoperative imaging these modalities are reserved for cases where their use will have an impact on diagnosis, surgical treatment, or mitigating risk from possible complications during surgery [35]. Use of HRCT and MRI is recommended for postoperative follow-up of cochlear implant (CI) surgeries or surgeries involving the temporal bone [35]. However, the use of these imaging technologies is ultimately limited by their lack of availability at the point-of-care. Typically, patients need to be referred to a radiologist for imaging where it can take several weeks to get results.
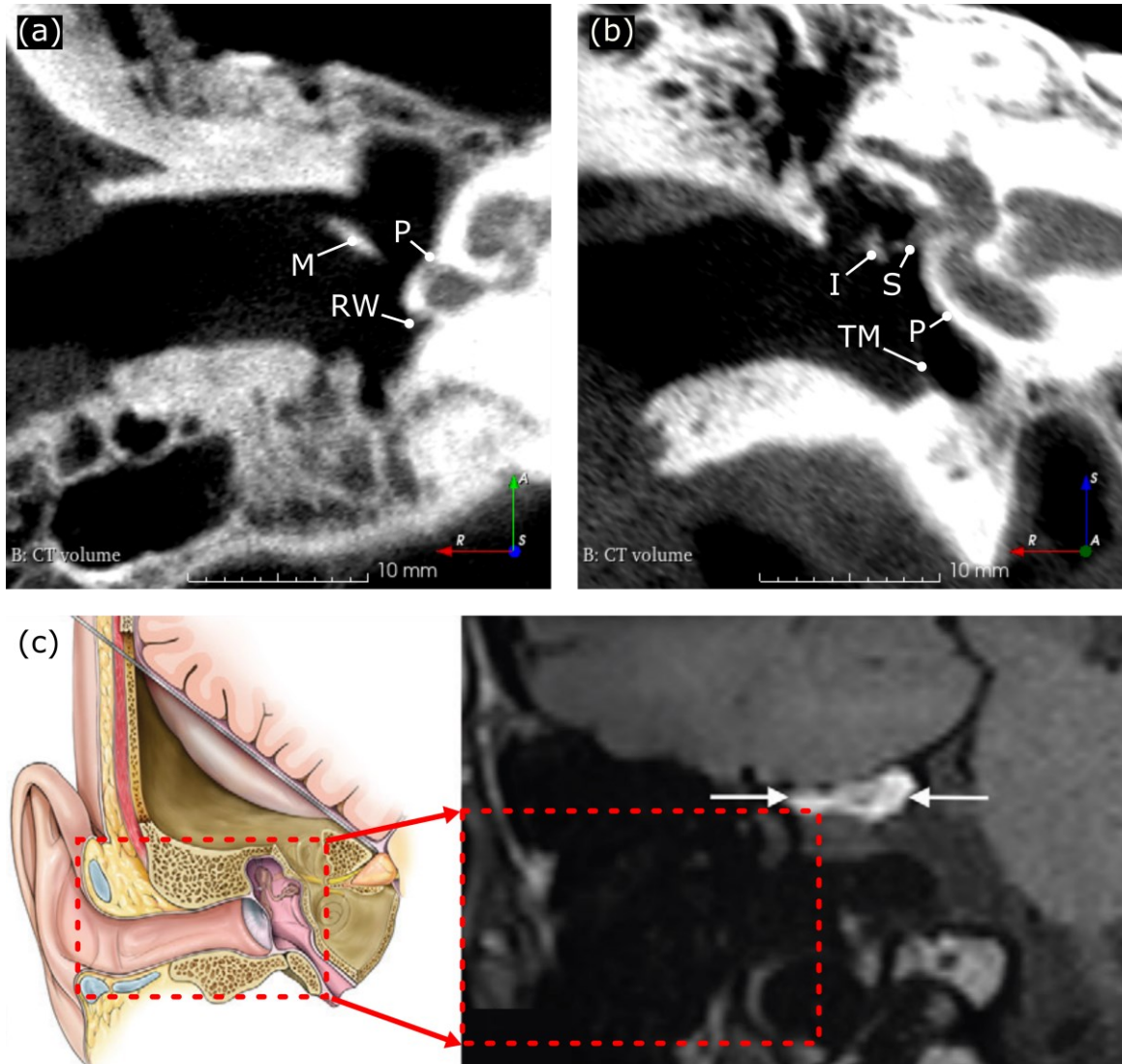
Fig. 4: *Radiological images of the middle ear. (a) and (b) axial and coronal computed tomography (CT) images of the human middle ear. (c) Magnetic resonance (MR) coronal image of the human middle ear outlined by the dotted red box. White arrows highlight a graft of fat tissues laid along the roof of the inner auditory canal. Derived from Ref. [55]. M, malleus; P, promontory; RW, round window; I, incus; S, stapes; TM, tympanic membrane.*

High-Resolution Computed Tomography

HRCT is capable of capturing the entire cross-sectional area of the ear anatomy and is routinely used to detect pathology affecting the temporal bone [56]. HRCT is reserved only for situations where the diagnostic benefit is clear as the radiation dose it delivers to patients increases risk of developing complications [57]. For example, high resolution CT significantly increases the risk of developing cataracts later in life [12]. Indications for use of HRCT include the presence of cholesteatoma, recurrent inflammation/infection and/or otosclerosis. In preoperative imaging, the extra anatomical information provided by HRCT can improve presurgical planning, thereby improving safety and efficacy of the subsequent

surgery [59]. HRCT is highly accurate in detecting small boney erosions in the lateral semicircular canal where it has a sensitivity and specificity of 100% and 97.73% respectively [56]. However, due to poor soft tissue contrast, HRCT can sometimes fail to adequately distinguish the extent of cholesteatoma and only has a sensitivity and specificity of 87% and 66% respectively in that application [56]. Furthermore, with a spatial resolution of $\sim 0.25\ mm$ and slice resolution of $1 - 2\ mm$ HRCT is too low-resolution to image and reliably detect pathology involving the ossicles, Fig. 4a and Fig. 4b. For example, it only has a sensitivity and specificity of 80% and 46.1% respectively for the detection of ossicular erosion [56].

Magnetic Resonance Imaging

Like HRCT, MRI can provide cross-sectional images of the middle ear. MRI has a similar resolution to HRCT but does not deliver ionizing radiation. Given MRI's improved intracranial soft tissue contrast over HRCT, MRI is the modality of choice to detect soft-tissue complications such as abscesses or tumors [60]. However, because MRI images experience strong susceptibility artefacts due to the air-tissue interfaces in the middle ear HRCT is preferred over MRI for imaging for most non-invasive middle ear imaging applications. An example of such susceptibility artefacts can be seen in Fig. 4c [59] highlighted by the dotted red box. Low-field MRI could potentially overcome these susceptibility artefacts, however, further development is needed to demonstrate clinical utility [61].

2.3 Pathology of the Middle Ear

In the context of pathology, the middle ear is approached as one part of a more complex, interrelated system, called the middle ear cleft, comprising (alongside the middle ear) the eustachian (auditory) tube and mastoid air cells [62]. The middle ear cleft is often described as a *miniature lung* that breathes through the eustachian tube due to its similarities with the lungs. Both the lungs and the middle ear are mucosa-lined structures that require air-filled cavities to function properly [63]. This 'miniature lung' is responsible for keeping the middle ear space aerated, maintaining normative pressure levels, and clearing of any outside contaminants and/or mucosal secretions. Disruption of these functions is the root cause of most common middle ear pathologies [59,62–64].

Eustachian Tube Dysfunction

The Eustachian tube (ET) is a small tubular structure connecting the nasal system to the middle ear cavity (also called the auditory tube as in Fig. 1) and is normally closed at rest [26]. The main function of the ET is middle ear pressure regulation, ventilation, mucosal secretion drainage, and protection of the middle ear from pathogens and sounds [9]. Eustachian tube dysfunction (ETD) is classified as the inability of the ET to perform these main functions. Prevalence of ETD has been previously reported to be between 0.9% and 4.6% [26, 27]. ETD has been detected in up to 70% of middle ear surgeries which implies a strong a functional correlation between ET function and other middle ear pathologies [2]. There are two types of ETD: obstructive ETD, commonly caused by mucosal blockages resulting from inflammation or infections, and patulous ETD, believed to be caused by an abnormally open ET during rest. The exact cause for either type of ETD is not currently well understood [27].

Patients with an obstructed ET report symptoms such as impaired hearing, fullness of the ear, tinnitus, and/or difficulties with balance. Obstruction of the ET creates negative pressure within the middle ear cavity that, if persistent, leads to complications such as otitis media (serrous or chronic), atelectasis (retraction of parts or all the TM) and cholesteatoma [2]. Diagnosis of an obstructed ET is generally based on clinical examination and a patient's medical history to identify underlying causes. Failure to repressurize the middle ear through pressure equalization methods combined with patient reported symptoms are common diagnostic signs [26]. Otoscopy and audiometry are commonly used to inform diagnosis where the presence of an intact TM, a tympanogram with either a flat shape or peak at a negative pressure around $-100\ mmH_2O$, and an air-bone gap of $15\ dBHL$ or greater is one definition for clinical diagnosis used by the UK Institute of Hearing Research [28]. However, these diagnostic signs are not considered clinically sufficient or reliable as many patients with obstructed ET can present with normal middle ear pressure or be asymptomatic. Moreover, an intact TM does not rule out obstruction of the ET [26].

Patients with patulous Eustachian tube (PET) commonly present symptoms of autophony, abnormally loud hearing of one's own voice and breathing. This is believed to be caused by an increase in sound conduction between the nasal system and middle ear cavity resulting from the ET being open at rest [29]. Traditionally, PET is diagnosed by

17

otoscopically observing movement on the TM corresponding to a patient's breathing. Difficulties arise however, when the amplitude of low frequency stimulus is too low to be visually observed by the clinician. This is further complicated by the possibility of other conditions mimicking the symptoms of PET such as superior canal dehiscence [30].

Cholesteatomas

Cholesteatomas are non-cancerous epidermal tissue growths that form within the middle ear cavity. As they slowly, yet inexorably grow in size, cholesteatomas lead to a multitude of complications such as hearing loss and bone resorption [59]. One sequela of cholesteatoma-induced bone resorption is the thinning of the bone of the inner ear creating a labyrinthine fistula (opening) which often leads to vestibular (balance) impairment [65]. Early signs of a cholesteatoma include the presence of ear discharge, CHL, and tinnitus that eventually leads to neurological effects such as facial paralysis and altered taste [66]. Cholesteatomas can either be congenital or acquired [67,68]. In acquired cholesteatomas the two most widely accepted theories around their primary development are the retraction pocket theory and the proliferation theory. Both theories suggest disruption of the TM's self-cleaning mechanism as the root cause of cholesteatoma formation. In retraction pocket theory this disruption is caused by negative middle ear pressure, due to obstructive ETD, creating deep retraction pockets of the TM trapping debris. In proliferation theory, disruption is caused by papillary ingrowth into the attic space of the middle ear cavity spurred by inflammation of the TM. Another theory, consistent with observed secondary cholesteatoma development, is the immigration theory. This theory postulates invasion of epithelium tissue into the middle ear through TM perforation sites [59]. Histologically, cholesteatomas are comprised of three layers: matrix, perimatrix and cystic content. The matrix layer consists of epithelial tissue like that of regular skin. The perimatrix is a supporting collection of connective tissue, new vessels, and inflammatory cells externally adjacent to the matrix. Finally the cystic content primarily consists of keratinized cell debris located within the matrix layer [59,68,69].

Congenital cholesteatomas have a prevalence between $2-5\%$ and are three times more likely to occur in males [59]. Acquired cholesteatomas were found to have a prevalence of ~0.4% amongst 3056 members of a northern Israel population [70]. However, there have been few studies conducted to examine acquired cholesteatomas

prevalence in different populations, making reported rates highly variable. For example a study of patients in Iowa reported a prevalence rate of only 0.01% [59].

Chronic Suppurative Otitis Media

Otitis media (OM) is an infection of the middle ear cleft consisting of a continuum of stages ranging from acute to chronic. Infection is caused by bacterial or viral infiltration into the middle ear cleft through either a puncture in the TM and/or nasopharyngeal secretion reflux from the ET due to ETD [71]. Chronic suppurative otitis media (CSOM) has been traditionally defined as a chronic discharge of the middle ear alongside a perforated TM leading to permanent alteration of middle ear structures. Patients with CSOM typically present with hearing loss and foul smelling discharge [59]. CSOM cause many histopathological changes to the middle ear with the formation of granulation tissue being the most common. Granulation occurs when vascularizing connective tissues herniate out from ruptured openings in the mucosa's basement membrane and epithelial cell lining because of inflammatory processes and bacterial toxins [72]. Alongside granulation tissue, bacterial biofilms have been shown to be present in CSOM. This suggests an important link between their formation and the progression of acute otitis media (AOM) to CSOM, however, their pathological connection to CSOM is not currently well understood [73–75]. In $3 - 5\%$ of cases CSOM may lead to adhesive otitis media where the TM is partially, or totally, bound to the promontory or ossicles by fibrous adhesions which persists even during forced reaeration of the middle ear space [76]. If inadequately treated CSOM can cause serious complications ranging from destruction of the inner ear leading to SNHL, vestibular dysfunction, erosion of the ossicles from fibrous adhesions leading to CHL, tympanosclerosis, cholesteatoma formation, meningitis, and mortality [59,76–79].

Prevalence of CSOM varies depending on the population with the lowest rates occurring in developed countries ($> 1\%$) to the highest rates being amongst Australian aboriginals ($4 - 46\%$) [59,80,81]. Globally CSOM has an average prevalence rate of ~3.07% with poorer populations, particularly children within developing countries affected more than wealthier industrialized nations [82]. CSOM is the leading cause of hearing loss in these developing countries [83].

Glomus Tumour

Glomus tumours, also known as paragangliomas, are typically benign tumours which may occur at any location in the body arising from cells of the paraganglia. They are commonly labeled according to their location with names such as the glomus tympanicum, glomus jugulare, carotid body, or glomus vagale and occur most often in the temporal bone or neck [84]. Of these, glomus tympanicum is the most common primary tumour of the middle ear and the second most common tumour of the temporal bone. Glomus tympanicum tumours typically form along the tympanic or auricular nerves of the middle ear [84]. Fortunately, glomus tympanicum tumours are relatively rare having an incidence rate of 1 per 1.3 million [85]. However, they are known to be hereditary with as many as 50% of paragangliomas being associated with a previous family history [86]. Glomus tympanicum are most prevalent in the 5th to 6th decades of life and can present as a conductive hearing loss and pulsatile tinnitus due to its highly vascular nature [87]. Larger tumours lead to more serious complications such as vertigo, facial palsy or SNHL.

Under otomicroscopic examination, glomus tympanicum tumours can often be seen as bright reddish, pulsatile, masses behind the TM [87]. For further evaluation, HRCT is the preferred imaging where they present as well-defined round soft tissue masses originating from the cochlear promontory and extending into the middle ear space [88]. Due to their highly vascularized nature, these tumours also have high contrast under MRI. Contrast-enhanced angiographic imaging is typically performed alongside conventional HRCT and MRI to help determine tumour perfusion and feeding vasculature which is useful for pre-surgical planning and treatment [88].

Middle Ear Congenital Malformations

Congenital malformations of the middle ear are structural and functional abnormalities that mainly occur during fetal development. Congenital malformation of the middle ear is often found in combination with external ear malformations and has an incidence of 1 in 15 thousand births. Congenital malformations primarily affect the right ear and are typically unilateral [89,90]. Isolated middle ear malformations, without external ear deformities, have been reported to occur in less than 10% of children presenting with CHL [89].

Both the ossicles and tympanic cavity can be affected by middle ear malformations that change the size, configuration, and number of these structures with varying degrees of

severity. Ossicular malformations may affect only parts of the ossicular chain or may affect the entire chain. Malformations involving the incus and ISJ separation are the most common [91]. Malformations of the promontory are also common. These affect the facial nerve trajectory and the oval and round windows [91]. HRCT, alongside typical audiometric diagnostics, is crucial for a clear diagnosis of congenital middle ear malformation and the subsequent surgical reconstruction effort, although many clinicians are very cautious about conducting HRCT imaging in children due to the radiation exposure risk.

Otosclerosis

Otosclerosis is a localized disease primarily affecting the boney labyrinth of the inner ear and the stapes footplate. While being a pathology of the inner ear, otosclerosis mainly presents clinically as a CHL due to the fixation of the ossicular chain at the stapes footplate and/or obstruction of the round window membrane due to the formation of osteosclerotic plaque [76]. The cause of otosclerosis is still largely unknown despite the numerous studies investigating its origin [92]. Otosclerosis is most prevalent in the Caucasian population with a rate of 3/1000. Were ~60% of patients report a family history of the disease which suggests a genetic origin [92]. A persistent measles virus infection of the boney labyrinth has also been suggested as a possible cause of osteosclerosis [76]. Patients with otosclerosis present with hearing loss, tinnitus, and dizziness. Typically hearing loss manifests as a progressive CHL without a history of head trauma or recurrent pathology. CHL associated with otosclerosis is usually noticed when a patient's normal hearing threshold is reduced by $15 - 25\ dB$. In some cases (~10%) patients can present with either a combination of SNHL and CHL or pure SNHL in the case of pure retrofenestral otosclerosis [76].
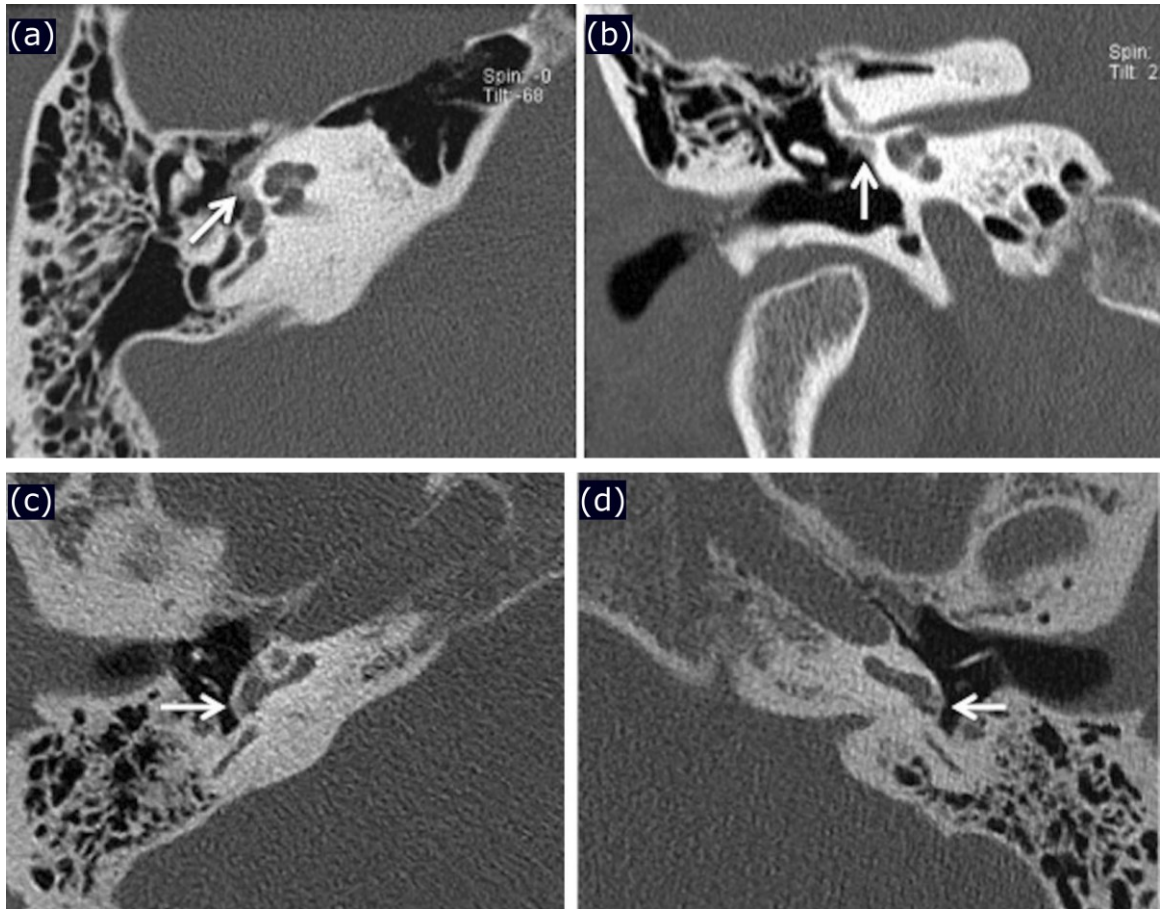
Fig. 5:   HRCT images depicting the different stages of otosclerosis. Axial (a) and (b) coronal images of the right-side middle ear of an adult patient with CHL and fenestral otospongiosis. White arrows point to the hypodense demineralized of the promontory. Right (c) and left (d) images of the middle ear in a patient with bilateral otosclerosis. White arrows point to the deposit of otosclerotic plaque causing narrowing of the round windows. From [93] and used according to the Creative Commons Attribution License [94].

Histologically, otosclerosis has two phases: active and stabilized. In the active stage, referred to as otospongiosis, the normal bone of the inner ear labyrinth is replaced by demineralized, highly vascularized, spongy tissue as shown in Fig. 5a and Fig. 5b. In the stabilization, or otosclerotic, phase this spongy tissue undergoes mineralization resulting in dense bone deposits (or plaques) as shown in Fig. 5c and Fig. 5d [76]. Osteosclerosis can be further characterized based on its location: fenestral and/or retrofenestral. Fenestral otosclerosis involves the stapes, oval window and round window and is the most common form of otosclerosis [76]. Retrofenestral otosclerosis affects the cochlea and is commonly found alongside fenestral otosclerosis [76].

The primary treatment option for otosclerosis is stapedotomy where a small hole is drilled into the oval window and piston attached to the long process of the incus is inserted

through it [65]. Stapedotomy has a success rate of ~95% in closing an air-bone gap within 10 $dBHL$ and is one of the most successful hearing restoration surgeries [76]. Despite its high success rate, stapedotomy requires specialized tools and skills and so is only undertaken by a subset of practicing otolaryngologists.

TM perforations

The most common TM perforations encountered clinically are those resulting from direct mechanical trauma due to cotton swabs, scuba diving, barotrauma caused by slapping and/or explosions, and head injury to only name a few [65]. Spontaneous perforation can also arise from complications of AOM with the risk of perforation increasing with recurring AOM [95]. Symptoms of TM perforations include a sudden cessation of pain, brief episode of tinnitus and vertigo, and hearing loss due to disruption of the TM's ability to transmit sound vibrations. Perforations can lead to further complications such as cholesteatoma or middle ear infection if the middle ear cavity remains exposed through the external auditory canal. Fortunately, ~80% of TM perforations heal spontaneously without clinical intervention [65]. However, if the perforation fails to heal given sufficient time or is too large to heal spontaneously then surgical reconstruction of the TM, through tympanoplasty, to close the perforation can be performed [65]. Although, tympanoplasty success can be impacted by ETD where Podoshin et al. reported more than a twofold increase in graft rejection in the presence of ETD [96].

Ossicular Trauma

Besides perforation of the TM, physical trauma may also lead to CHL associated with disruption of the ossicular chain. In such cases, temporal bone fractures are often also present leading to further complications such as SNHL, vestibular disfunction or facial palsy [97,98]. Ossicular trauma is typically classified into incudostapedial joint (ISJ) or incudomallear joint (IMJ) separation, dislocations of the incus, malleus, or stapes superstructures, and ossicular fracture [97,99]. Separation of the ISJ is the most common form of ossicular trauma because of its fragile connection to the stapes superstructure and lack of direct muscular attachments [99].

Diagnosis of ossicular trauma causing CHL is often delayed and is only suspected when CHL persists after the patient has recovered from the main traumatic injury. HRCT

is the modality of choice for pre-surgical diagnosis of ossicular trauma. However, it requires radiologists to have extensive knowledge of ossicular anatomy to detect abnormalities and can be complicated depending on the type of injury [100]. Thus, the heterogeneous nature of ossicular trauma requires a tailored surgical approach [97].

Ossiculoplasty

A common surgical treatment option for CHL arising from pathology of the middle ear, is replacing portions of the ossicular chain with a partial or total ossicular replacement prosthesis (PORP/TORP) in an attempt to restore the disrupted auditory pathway through the middle ear. PORPs are used in cases where there is an intact stapes superstructure while TORPs are used when the stapes is missing. The prostheses act to create a direct acoustic coupling between the tympanic membrane and the cochlea.

Currently about 15% of primary ossiculoplasty surgeries [101] and up to 50% of revision ossiculoplasty surgeries [102] fail to result in a good hearing outcome. Improved pre-surgical diagnostics can potentially increase surgical success by giving surgeons a better picture of the underlying pathology preoperatively and can allow patients to select more conservative treatment options such as hearing aids in cases where surgery is unlikely to provide benefit.

2.4 Optical Coherence Tomography

Optical coherence tomography (OCT) is a non-invasive optical imaging technique that constructs depth resolved images of tissue by measuring the time-of-flight (TOF) of reflected light originating from scatterers at different depths within a sample. TOF information is measured indirectly through low-coherence interferometry as originally demonstrated by Fujimoto et al. in 1991 using a time domain OCT (TD-OCT) setup [103]. As depicted in Fig. 6, TD-OCT works by splitting the light from a low-coherence light source between two optical paths, or arms, via a coupler. One arm directs the light towards a reference mirror and the other towards the tissue sample to be imaged. The backscattered light from both arms is then recombined in the coupler and the resulting interference pattern, or interferogram, is recorded by a photodetector. Signals from both arms only coherently interfere when the optical path length difference between them is within the coherence length of the light source. Therefore, by changing the path length by moving the

24

reference mirror creates a depth resolved axial image line (A-line) of the tissue sample scattering amplitude. By scanning the sample beam laterally across the image, these A-lines can be spliced together to form 2D brightness mode (B-mode) or 3D volumetric images. A-line rates of TD-OCT systems are limited by the rate at which the reference mirror can be moved over its range and typically cannot exceed 1 kHz [104].
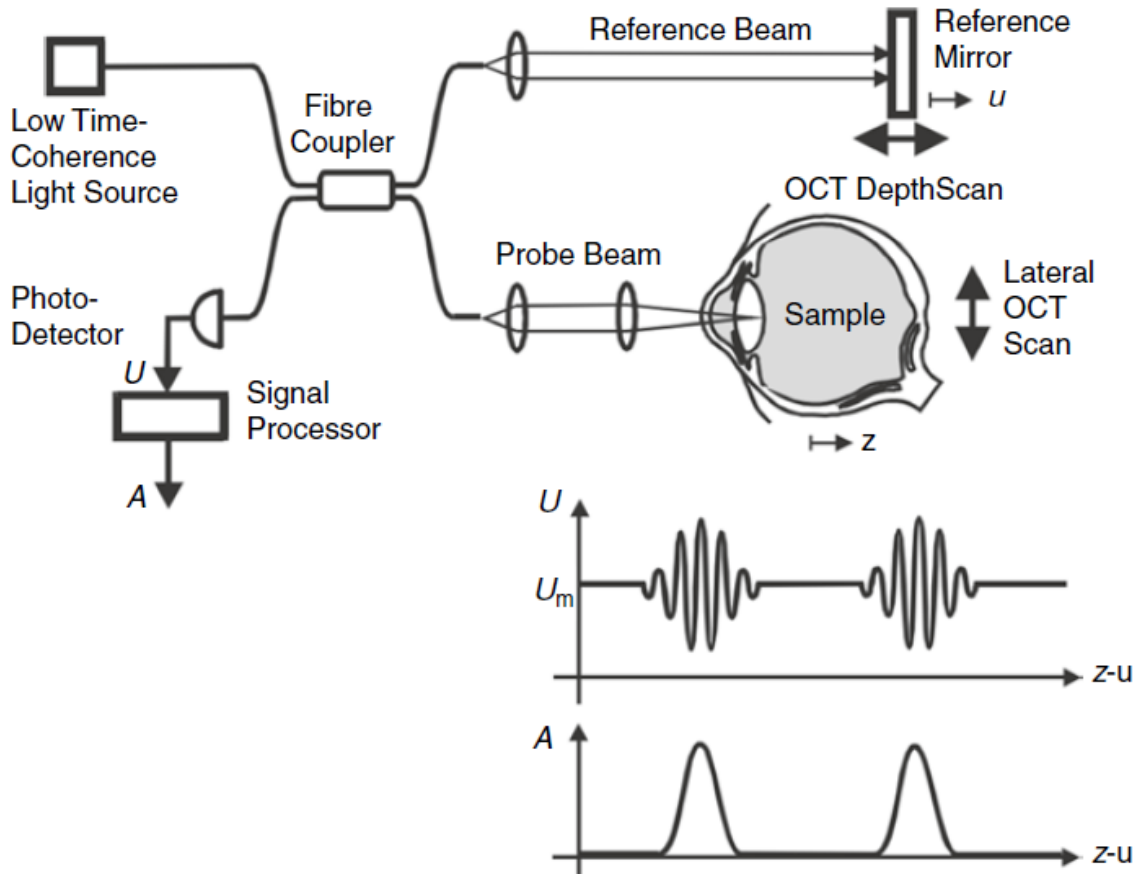
Fig. 6: *Time-Domain Optical Coherence Tomography (TD-OCT) system. From [104] and used according to the Creative Commons Attribution License [94].*

2.4.1 Fourier Domain Optical Coherence Tomography

Fourier domain optical coherence tomography (FD-OCT) was introduced after TD-OCT as a more efficient approach. In FD-OCT, the spectrum of the interferogram across all depths is measured directly without the need to modulate the reference mirror. An A-line can then be constructed by simply taking the inverse Fourier transform of the measured spectrum. Two different methods of FD-OCT are in widespread use, as depicted in Fig. 7, spectral domain OCT (SD-OCT) and swept-source OCT (SS-OCT). SD-OCT, Fig. 7a, replaces the photodetector of the time domain system with a diffraction grating and a high-

speed line scan camera. The diffraction grating spatially separates the different wavelengths of the light source into different wavelength bins corresponding to the elements on the line scan camera. SS-OCT, Fig. 7b, replaces the broadband light source of the time domain system with a swept-source laser. The sample is imaged by a laser beam with an instantaneously narrow spectral line width which is swept across a broad range of wavelengths. This results in each bin of the spectral interferogram being sequentially recorded as the laser sweeps through wavelengths using a similar photodetector setup as the time domain system. A-line rates of SD-OCT and SS-OCT systems are limited by the line scan camera's readout speed and laser sweep rate respectively but are orders of magnitude faster than typical TD-OCT systems and can range from tens of kHz to MHz [104].
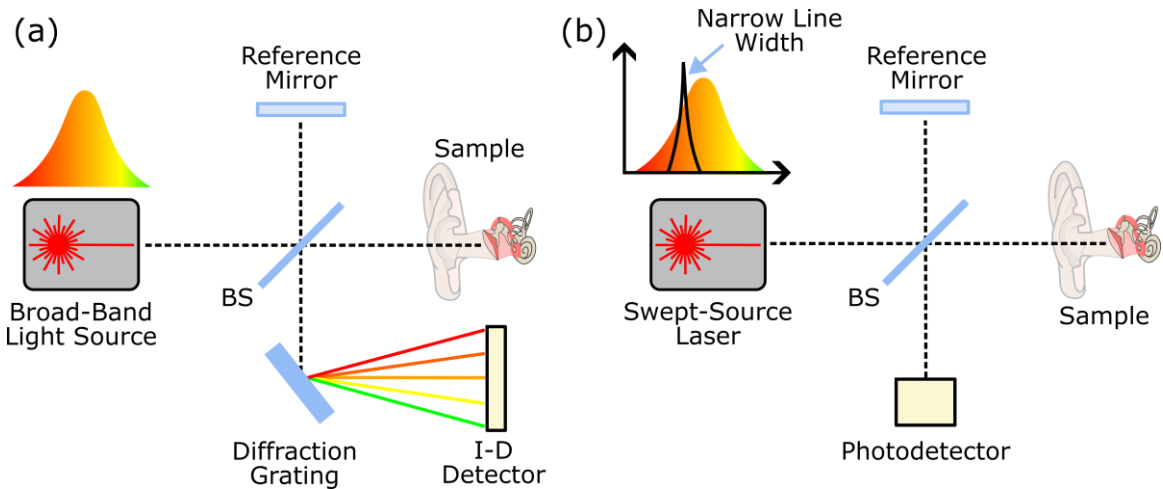
Fig. 7:   Configuration of the two different types of Fourier Domain Optical Coherence Tomography (FD-OCT) setups.
(a) Spectral Domain OCT (SD-OCT). (b) Swept-Source OCT (SS-OCT). Derived from [104] and used according to the
Creative Commons Attribution License [94].

2.5   Optical Aberrations in ME-OCT Imaging Systems

ME-OCT imaging systems are susceptible to aberrations introduced through imperfect imaging optics or a non-linear relationship between scanning mirror drive signals and scanning mirror angle. Aberrations are higher-order deviations from an ideal first-order optical system that cause image distortion, reducing geometrically accuracy, and defocusing that deteriorates image quality. Aberrations that cause image distortion include geometrical and lateral distortion whereas those causing defocusing include spherical aberration, coma, astigmatism, and chromatic aberration.

Geometrical Distortion

To construct OCT images, our ME-OCT system uses a surface mounted MEMS mirror to scan the swept-source laser beam laterally across the image volume. Due to the entocentric geometry of the optics, images are acquired within a spherical coordinate space $[r, \theta, \phi]$. This imaging geometry produces a "fan-beam distortion" [105] when OCT images are displayed without correcting for the fact that data was collected along non-parallel scan lines. This geometrical distortion is illustrated in Fig. 8 where a planar surface $D(r, \theta, \phi)$ imaged within the spherical coordinate space of the scanning optics leads to an apparently curved image surface $D(x, y, z)$ when rendered in a Cartesian coordinate space. Generally, geometrical distortion causes planar surfaces to appear to bend away from the center of the scan in proportion to the beam angles (θ, ϕ) [106].
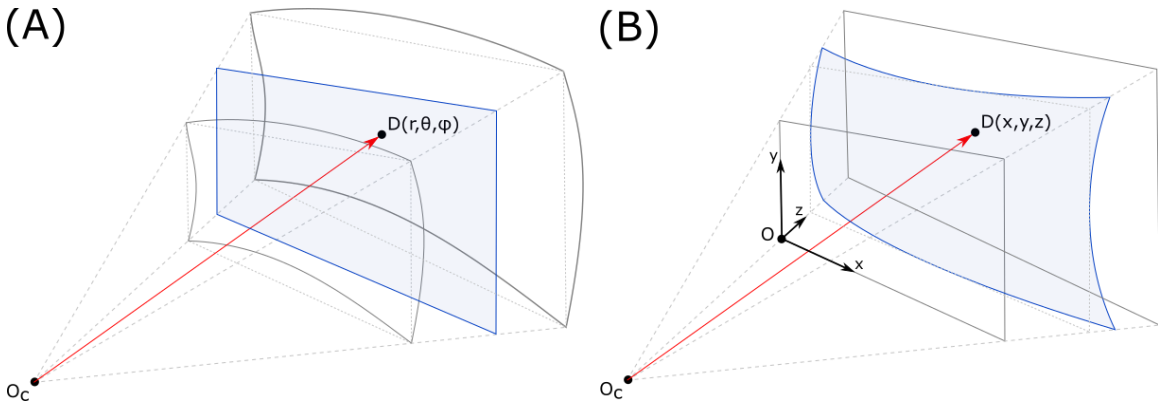
Fig. 8:   Geometrical distortion of the ME-OCT imaging system where a planar surface $D(r, \theta, \phi)$ captured within the spherical cooridante space of the scannign optics (a) leads to a curved surface $D(x, y, z)$ when rendered in a Cartesian coordinate space (b).

Lateral Distortion

Lateral distortion occurs when the transverse magnification, and focal length, across the imaging optics is a function of lateral distance from the optical axis. Lateral distortion manifests as a deformation of the image, independently from the defocussing effects of other aberrations [107]. As shown in Fig. 9, the two main types of lateral distortion are pincushion distortion and barrel distortion. With pincushion distortion (Fig. 9b) lateral magnification increases with the radial distance from the center causing image points to pinch inwards along the horizontal and vertical axes. In contrast, barrel distortion, Fig. 9a, occurs when magnification decreases with radial distance from the center causing image points to bulge outwards along the horizontal and vertical axes.
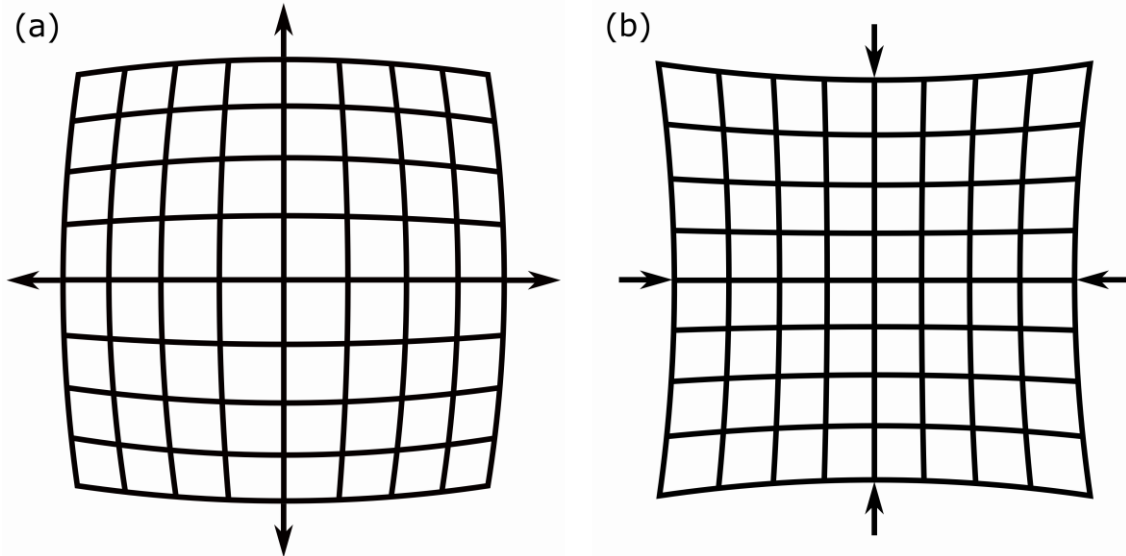
Fig. 9:   Lateral distortion resulting from (a) barrel distortion and (b) pincushion distortion.

MEMS mirrors also exhibit inter-axis crosstalk in which the sensitivity of mirror tilt angle to applied voltage in the x axis depends on the tilt angle in the y-axis and vice versa [34–36] leading to lateral distortion of the displayed OCT image. Using a similar MEMS scanner to the one in our system, Kim et al. determined angular error during dual axis scanning due to crosstalk to be as high as 6.13% depending on MEMS scanner alignment [170]. Furthermore, the dependence of MEMS mirror angle on drive voltage is non-linear with mirror angle and tends to saturate at the ends of the drive voltage range.

Spherical Aberrations

Spherical aberration arises from the tendency of spherical lenses to exhibit a dependency of focal length on the distance from the center of the lens as shown in Fig. 10 [107]. The effective variation in focal length across the lens surface causes images to be blurred uniformly across the lateral FOV.
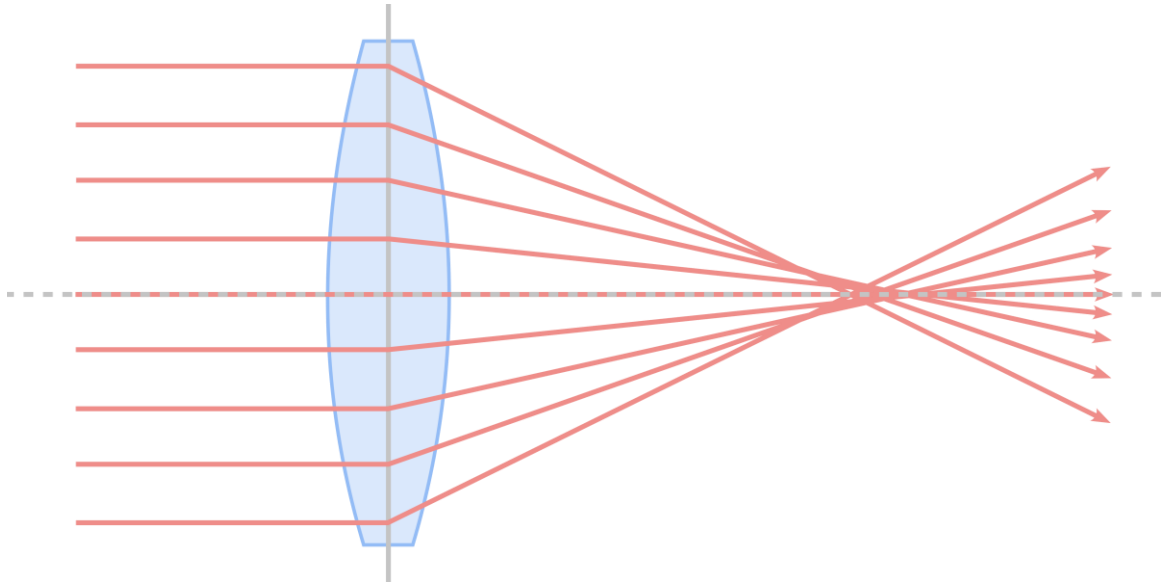
Fig. 10: Spherical aberration is a variation in focal length with distance from the lens center that occurs in spherical lenses and causes blurring of the image across the lateral field of view.

Coma

Coma arises when an optical system fails to magnify off-axis image points uniformly across the image plane resulting in a blurred shape that resembles a comet with a head and a tail [108]. Off-axis light passing through the optical system is refracted with different powers along its surface depending on the light's angle of incidence. As shown in Fig. 11, this results in a comet-like elongation of image points as the light converges at points off-center from the optical system's central focal point in proportion to radial distance from the image center. Coma is most prominent near the edges of the lateral FOV.
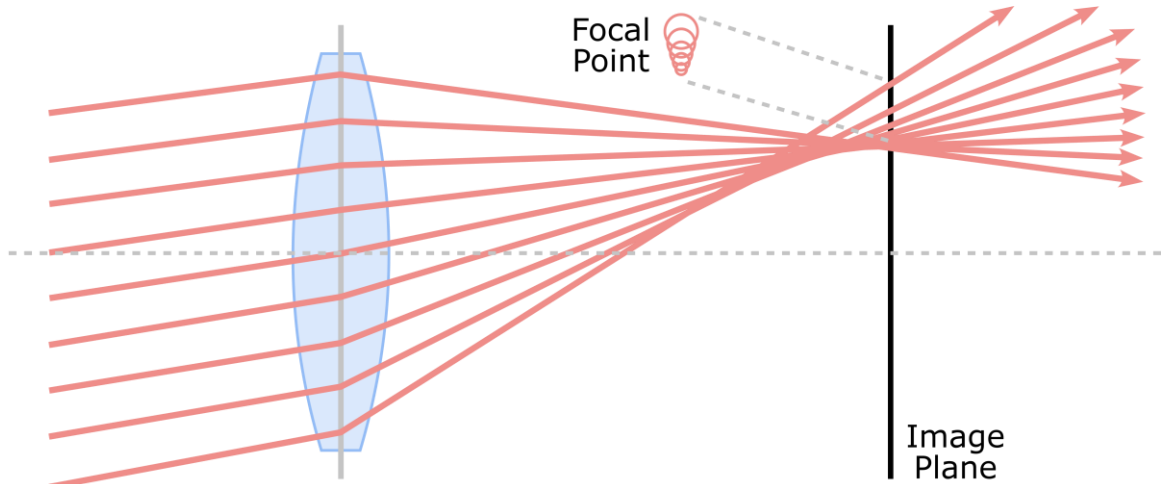
Fig. 11:   Coma results in a comet-like elongation of focal points as the light converges at points off-center from the optical systems central focal point in proportion to radial distance from the optical axis.

Astigmatism

As shown in Fig. 12, astigmatism manifests when off-axis light passes through an optical system with different focal lengths along the tangential (vertical) and sagittal (horizontal) planes. Depending on the focal plane, this results in tangential or sagittal astigmatism. In tangential astigmatism, image points are elongated radially relative to the center of the image. Whereas in sagittal astigmatism image points are elongated along concentric rings radiating outwards from the center of the image. In both types of astigmatism, the elongation of the image points is proportional to the radial distance from the image center. Thus, astigmatism is most noticeable near the edges of the lateral FOV.
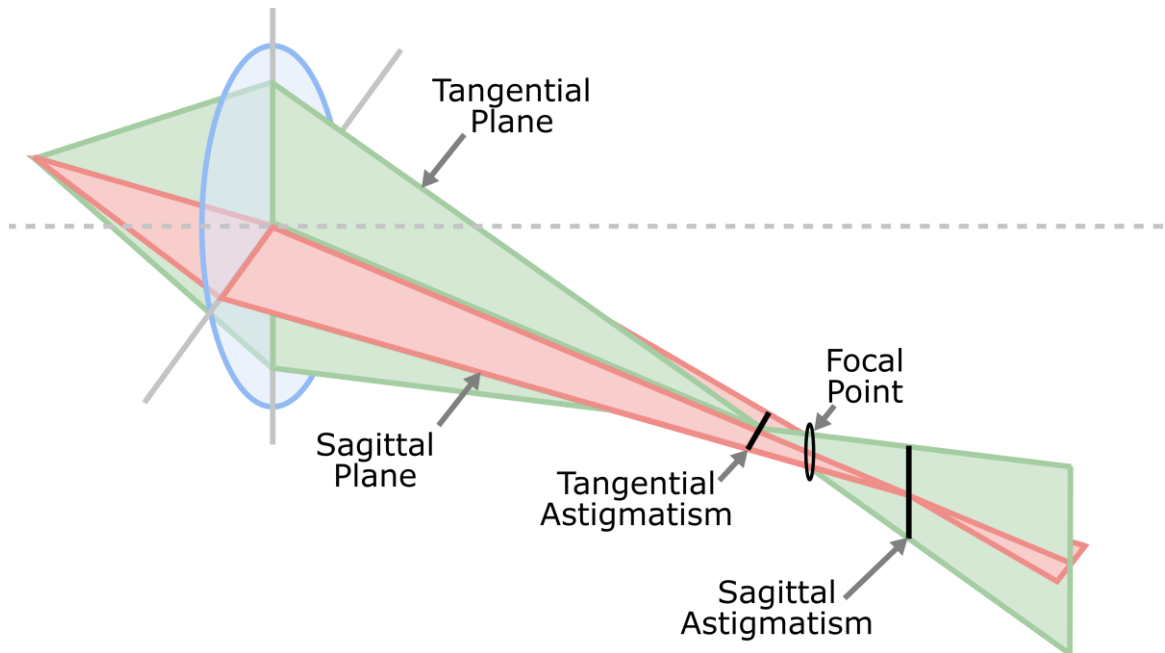
Fig. 12: Astigmatism due to different focal lengths along the tangential and sagittal planes of an optical system.

Chromatic Aberration

Chromatic aberrations occur due to the variation of the refractive index of a lens on wavelength causing light of different wavelengths or "colours" to be focused to different points. Chromatic aberration is comprised of longitudinal (axial) chromatic aberration, Fig. 13a, which causes coloured fringes to appear around image points uniformly in an image and lateral chromatic aberration, Fig. 13b, which causes dispersion and blurring of off-axis image points.
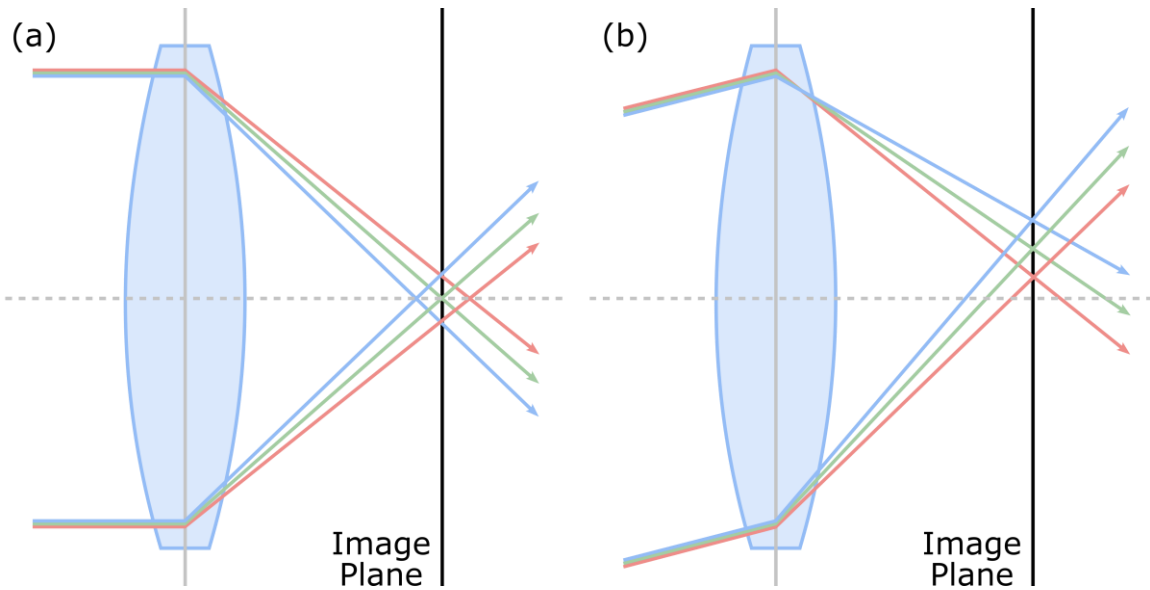
Fig. 13: Longitudinal (a) and lateral (b) chromatic aberrations which occur due to the variation of the refractive index of an optical system on wavelength causing light of different wavelengths to refract along different paths.

2.5.1 Handpiece Design and Aberration Trade-off

In the current ME-OCT imaging system, Section 1.1.1, the optical design of the handpiece utilized a GRIN rod lens to position the imaging origin close to the TM and achieve a large numerical aperture. However, GRIN rods have been shown to suffer more from chromatic aberrations, coma, and astigmatism [109–112] than conventional optics. At the laser's operating wavelength of 1550 nm, dispersion is small enough that chromatic dispersion is negligible, but astigmatism and coma result in noticeable aberrations in images taken with the system.

When originally designing the handpiece, it was uncertain whether an endoscopic or an otoscopic formfactor would be best suited for in-vivo patient imaging at the point-of-care. A GRIN rod would enable the handpiece to fulfill both formfactors and allow it to be modified from its default otoscopic form factor to an endoscopic one relatively easily. However, because otoscopic imaging results were satisfactory, the conversion was never attempted. In the latest iteration of the system (which was completed after the experimental work described in this thesis), the handpiece uses conventional (non-GRIN) lenses and has an otoscopic form factor. This design choice significantly reduces the coma and astigmatism of the imaging optics compared to the handpiece used in the studies described in this thesis.

32

2.6 Optical Coherence Tomography Angiography

OCT angiography (OCTA) is a functional extension of OCT that is widely used in ophthalmology for imaging retinal vasculature [113,114], particularly in the diagnosis of diabetic retinopathy and glaucoma [115–117]. OCTA works by taking a sequential time series of images and using interframe differences in phase and/or amplitude to obtain contrast on vasculature. As shown in Fig. 14, static structures in OCT scans do not change between successive acquisitions, whereas the motion of blood cells causes the amplitude and phase in vascularized regions to vary rapidly. This produces contrast between these vascular and avascular regions when the change in phase or amplitude is tracked over time. To extract the angiographic signal, numerous algorithms have been developed over the years, but they can be generally classified into four groups: Doppler OCT, phase variance, amplitude variance and complex methods [118]. Doppler OCT provides a quantitative measurement of blood flow by looking at phase shifts between adjacent scans and operates using the same principles as OCT-DV. The other three methods are considered *true* OCTA and produce a qualitative map of vasculature called an *angiogram* [118]. Enface angiograms are typically constructed through intensity projection of a volumetric stack of many 2D cross-sectional angiograms. Regardless of the method, the main differences between these algorithms comes down to how they acquire adjacent scans, how motion differences are calculated, and how they deal with bulk motion [119].
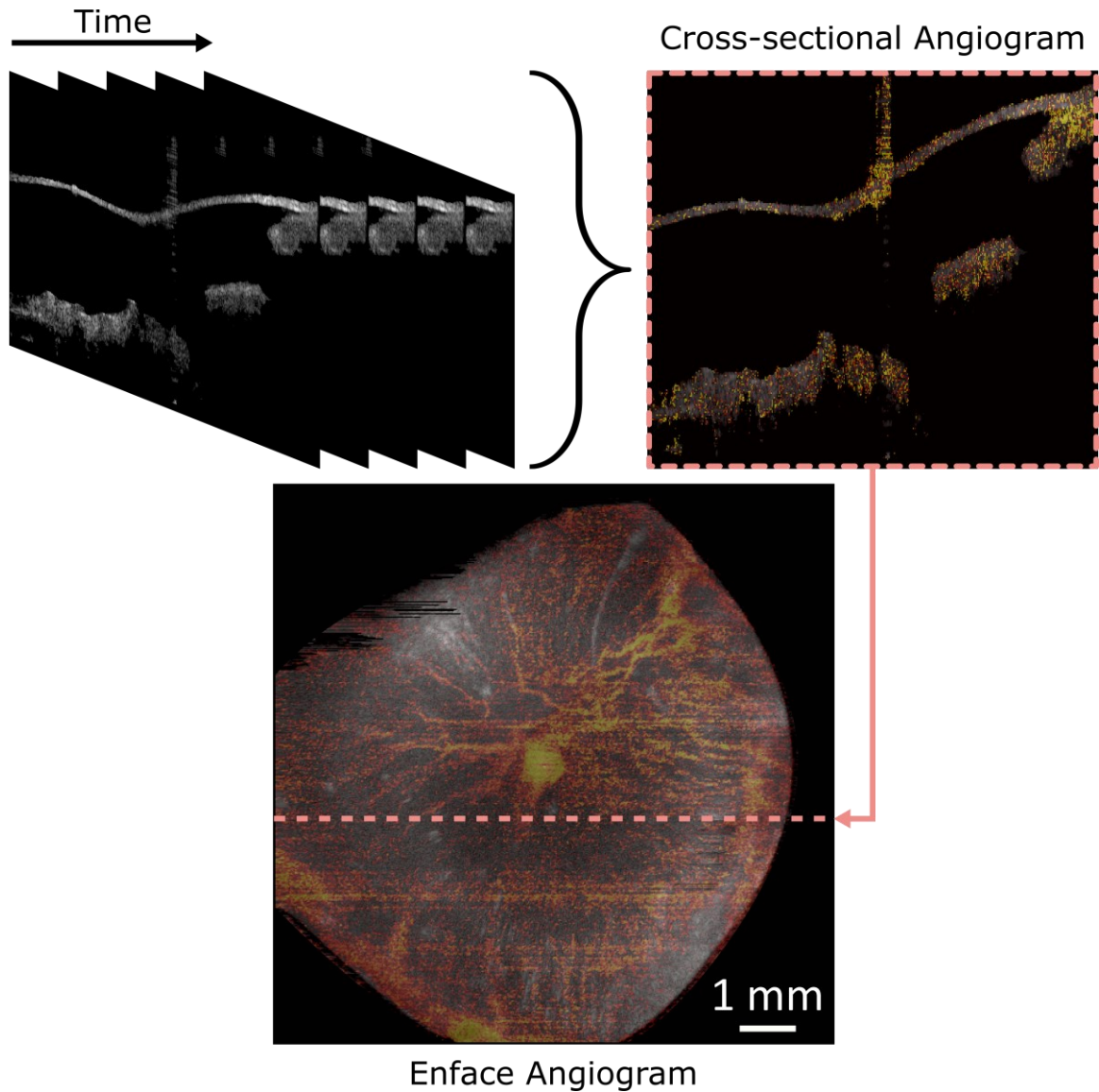
Fig. 14: *A sequential time series of OCT scans are taken at the same location and the differences between sequential images are used to produce a vasculature contrast image called an angiogram. An enface angiograms can be constructed through intensity projection of a stack of many cross-sectional angiograms.*

2.6.1 Bulk Motion

Bulk motion is defined as the common-mode differential phase noise of all pixels within each A-line. It arises due to the relative motion between the OCT system and the sample either through physiological processes within the sample such as heartbeat or breathing or through the OCT imaging probe when being manipulated by the operator. Without correcting for this relative motion, the resultant bulk motion phase noise dominates the interframe angiographic signal resulting in poor vasculature contrast. Common approaches for bulk motion correction involve edge detection [120,121], amplitude weighted phase

averaging [122,123] and the histogram method [124,125]. In all these approaches, the interframe bulk motion is estimated for each A-line and subtracted from each pixel in the A-line. The resultant bulk motion corrected A-line can then be used as input to the myriad of angiography algorithms to extract the angiography signal for display.

Edge Detection

Edge detection methods for bulk motion correction relies on the presence of bright, avascular, reflectors on the sample surface. Such structure on the surface can be segmented to obtain the bulk motion value for each A-line separately from any angiographic signal. One such edge detection described by Fujimoto et al. [120] outlines a two-step process where the surface is first smoothed using a small 3×3 filter window followed by segmentation of the blurred surface within a band of ± 5 pixels. The central value of this segmented surface is then chosen as the bulk motion term for the entire A-line.

Amplitude Weighted Phase Averaging

In samples without a bright, avascular, reflective surface, amplitude weighted phase averaging can be used instead. Phase averaging assumes that angiography signal from the pixels containing vasculature within an A-line is small compared to the value of bulk motion [122]. Thus, the average of the pixels phase values will be the value for bulk motion. An amplitude weighted phase average is used to limit the phase contribution from areas of low signal that would otherwise distort the result.

Histogram Method

For samples where the angiography signal from vasculature is not negligible, in comparison to bulk motion, the previously described amplitude weighted phase averaging method breaks down. In these cases, the phase average is skewed by the angiography signal offsetting the calculated bulk motion value. This results in correction artefacts manifesting as false angiography signals both above and below the vasculature in the final OCTA image [125].

The histogram approach seeks to fix this correction artefact by constructing a histogram of axial phase differences for each individual A-line and looking for the mode of the phase differences. Here it is assumed that the number of pixels containing static tissue within each A-line is large compared to those pixels containing vasculature. Thus the

mode of the histogram describes the phase change seen by the static tissue component and represents a good estimate of the bulk motion [125].

2.6.2 Doppler OCT

The use of Doppler OCT for the non-invasive imaging of in-vivo blood flow was first demonstrated by Chen et al. using a TD-OCT system [126]. In Doppler OCT the velocity of the moving blood cells causes a time dependent beat frequency f_s to develop in the measured OCT interferogram for each depth sample. This results in a shift of the local sample's frequency spectrum by f_s relative to the nominal center frequency of the OCT system. The value of f_s for each depth sample can be extracted by applying a series of windowed short-time Fourier transforms (STFT) across the interferogram [127]. Assuming the angle θ between the OCT sample beam, and the direction of blood flow is known, the mean blood flow velocity V_s can be calculated for each sample according to:

$$V_s = \frac{f_s c}{2 n_t v_0 \cos \theta}$$ 1

Where v_0 is the light source's center frequency, n_t the local tissue refractive index, and c the speed of light [127]. The sensitivity of V_s is a function of acquisition time and θ, where higher sensitivities can be achieved for longer acquisition times or smaller θ. Here the main limitations of Doppler OCT become evident as Chen et al. reported that it took ~3 minutes to acquire Doppler OCT images spanning a region of $1\,mm \times 1\,mm$ with a $10\,\mu m$ spatial resolution to achieve a velocity sensitivity of $\sim 100\,\mu m/s$ [126]. Furthermore, Doppler OCT cannot measure blood flow in vasculature that is perpendicular to the OCT sample beam, as Eq. (1) is undefined for $\theta = 90°$.

2.6.3 Phase Variance Angiography

Chen et al. where also the first to demonstrate phase variance OCT using a FD-OCT system [128]. They showed it was possible to construct a qualitative map of blood flow turbulence through calculation of the Doppler variance σ^2:

$$\sigma^2 = \frac{1}{T^2} \left\{ 1 - \frac{\left| \frac{1}{N} \sum_{j=1}^{N} [f_j(z) \cdot f_{j+1}^*(z)] \right|}{\frac{1}{N} \sum_{j=1}^{N} [f_j(z) \cdot f_j^*(z)]} \right\}$$ 2

Where T is the time difference between adjacent A-lines, N is the number of Adjacent A-lines used, $f_j(z)$ is the complex OCT signal after taking the Fourier transform of the measured interferogram, $f_j^*(z)$ is its complex conjugate [129] and j indexes the line number. Using this approach, we do not need to explicitly calculate the Doppler frequency shift as in traditional Doppler-OCT. However, the Doppler variance method is sensitive to any phase noise introduced either through the lack of phase stability in the FD-OCT used and/or in-vivo motion [128].

The method of optical angiography (OAG) was later introduced by Wang et al. as a way to efficiently filter out phase noise artefacts by introducing a frequency shift in the recorded blood flow measurements [130]. This frequency shift was introduced in hardware via modulation of the reference arm length of their SD-OCT system by means of a mirror attached to a piezoelectric stage. This caused a positive shift in the Doppler frequency proportional to the modulation frequency of the reference arm as sequential A-lines were acquired. By applying the Hilbert transform across the sequential A-lines at each depth, before applying the Fast Fourier transform (FFT), the angiographic signal of interest is isolated to the negative portion of the resultant FFT spectrum [130]. Here the frequency modulation effectively acts as a high-pass filter where any measured velocity equal or greater than the velocity of the modulation results in the measured velocity being aliased into the negative frequency spectrum. Using OAG, Wang et al were able to acquire a 3D angiographic volume ($2.2\ mm \times 2.2\ mm \times 1.7\ mm$) in $\sim 50s$ with a velocity sensitivity of $0.4\ mm/s$.

2.6.4   Amplitude Variance Angiography

Another solution to the phase noise sensitivity of the phase variance method was proposed by Liu et al. [131]. Here they modified Eq. (2) by simply taking the absolute value of the A-line variance before averaging according to:

$$\sigma^2 = \frac{1}{T^2}\left\{1 - \frac{\frac{1}{N}\sum_{j=1}^{N}\left|f_j(z)\cdot f_{j+1}^*(z)\right|}{\frac{1}{N}\sum_{j=1}^{N}\left[f_j(z)\cdot f_j^*(z)\right]}\right\} \qquad 3$$

While simple, this amplitude variance method is effective at eliminating artefacts caused by phase unstable FD-OCT systems, although can still be susceptible to motion

artefacts [131]. However, throwing away the phase information can cause difficulties in situations where blood flow contrast is only provided by small Doppler phase shifts instead of moving scatters [132].

Huang et al. would expand on amplitude variance by introducing split-spectrum amplitude-decorrelation angiography (SSADA) [133]. SSADA works by splitting the spectrum of each A-line into four or more distinct frequency bins through convolution with a gaussian bandpass filter bank. The speckle decorrelation within each of these bins is then calculated and averaged together to produce the decorrelation image $\bar{D}(x,z)$ according to:

$$\bar{D}(x,z) = 1 - \frac{1}{N-1}\frac{1}{M}\sum_{n=1}^{N-1}\sum_{m=1}^{M}\frac{A_{n,m}(x,z)A_{n+1,m}(x,z)}{\left[\frac{1}{2}A_{n,m}(x,z)^2 + \frac{1}{2}A_{n+1,m}(x,z)^2\right]} \qquad 4$$

Where x and z are the lateral and depth indices of the samples within each B-mode image, n is the B-mode slice index, m is the split-spectrum index, $A_{n,m}(x,z)$ is the sample amplitude, N is the number of sequential B-mode images used, and M the number of split-spectra [133].

Splitting the A-line spectrum decreases the axial resolution to better capture bulk motion in the axial direction and reduces phase noise. Huang et al. applied this method to an SS-OCT system where the axial resolution was $5\,\mu m$, as compared to the lateral resolution of $18\,\mu m$. This made phase noise induced by axial pulsatile motion of vasculature disproportionally affect their measurements. Here they split the spectrum as to decrease the axial resolution such that it matched their systems lateral resolution. The angiographic signal was extracted from eight sequentially acquired B-mode images where noisy frames, affected by lateral bulk motion between frames, was greatly reduced. Using SSADA, Huang et al. where able to capture detailed angiogram volumes ($3.0\,mm \times 3.0\,mm \times 2.9\,mm$) in only $3.2\,seconds$ [133].

2.6.5   Complex Angiography

Spurred by the availability of faster, phase-stable, FD-OCT systems, Wang et al. developed the method of optical microangiography (OMAG) as an extension and simplification of their earlier OAG method [134]. In OMAG, the full complex OCT signal is used which provides the benefits of both phase and amplitude methods. That being the sensitivity of

phase variance and the robustness against bulk motion noise provided by amplitude variance. Complex OMAG was also shown to provide superior performance over both phase and amplitude methods [134]. Furthermore, as modern FD-OCT systems are inherently more phase stable and offer faster scanning rates than older systems, there is often no longer a need to introduce a frequency modulation in the reference arm (via piezo electronics) as was done in OAG.

OMAG works by taking the running differential complex average between successive samples of either sequentially acquired A-lines or B-mode images:

$$OMAG_C = \frac{1}{N-1} \sum_{i=1}^{N-1} (|C_{i+1} - C_i|) \qquad\qquad 5$$

Where C_i is the complex element in the ith sequential acquired A-line or B-mode image, and N is the number of acquired A-lines or B-mode images. Each scanning configuration has its own advantages and disadvantages where the best configuration ultimately depends on the context in which OMAG will be used. For example, if sequential A-lines are used then bulk motion phase noise will not be an issue provided the OCT systems has a sufficiently fast A-line rate. However, as the time interval between adjacent A-lines is relatively small (on the order of microseconds) the blood flow velocity sensitivity is lower and so contrast on smaller vasculature is poor [134]. If sequential B-mode images are used then the time between acquisitions at the same location is increased (typically to millisecond timescales) making the method more sensitive to slow-moving blood in smaller vessels [134]. However, this longer time interval also introduces more bulk motion phase noise within the angiogram that must be corrected for, before the application of OMAG, through bulk motion correction algorithms as discussed in Section 2.6.1.

2.7 GPU Architecture, Processing and Rendering

In 2003 the performance of serial algorithms started to plateau due to limitations in processor speeds imposed by energy and power dissipation constraints [135]. This limitation spurred a paradigm shift in software development towards parallel processing and concurrency. Modern architectural design of both the CPU (central processing unit) and the GPU (graphics processing unit) has been greatly influenced by this fundamental shift towards concurrency in modern software. However, both the CPU and GPU have

taken vastly different approaches in providing resources for concurrent execution. The architectural design of the CPU has focused on speeding up sequential algorithms by allowing multiple sets of serial instructions, or threads, to be run in parallel. This approach has made use of sophisticated control logic and large on-chip memory caches to reduce memory and instruction latency, however, this limits the CPU to only being able to manage a handful of arithmetic logic units (ALU) running parallel threads. In contrast, the GPU has focused on increasing the throughput of sequential algorithms by running thousands of lightweight threads in parallel. GPUs are comprised of many processing cores called streaming multiprocessors (SM), with relatively simple control logic and limited cache size. Each lightweight GPU thread is run on the many streaming processors (SP) contained within each SM [136]. Compared to CPU threads, GPU threads are much higher latency and generally operate at lower clock rates.

Due to their complementary properties, the GPU and CPU are often used together in heterogeneous systems to obtain optimal application performance gains [137]. For example, CPUs excel at task parallelism due to their low latency. Task parallelism arises when there are multiple sets of instructions, or tasks, that can be performed independently of each other on multiple CPU cores simultaneously. GPUs excel at data parallelism due to their high throughput. Data parallelism arises when there are many data elements that can be operated on independently across multiple SMs simultaneously. For medical signal and image processing applications, the CPU is used to accelerate core control flow whereas GPUs are used to accelerate computations on large data arrays. The high latency of GPU threads makes it unsuited for control flow acceleration. Likewise, the limited number of cores in CPUs makes them scale poorly for large data arrays [136].

2.7.1 Compute Unified Device Architecture Programming Model

The compute unified device architecture (CUDA) is a programming model originally introduced by the Nvidia corporation in 2007 [138]. CUDA provides software level abstractions that allow developers to implement general-purpose parallel algorithms, or kernels, on the GPU's physical architecture. One of those abstractions is the two-level thread hierarchy which divides groups of threads into blocks which are then organized into a grid. The dimensionality of the grid and blocks are specified by the developer when executing a kernel (i.e., a computer program written for the GPU) from the host CPU.

GPUs follow the single instruction, multiple data (SIMD) model where a single instruction is used across all threads to manipulate multiple streams of data simultaneously [139]. However, due to the physical organization of the GPU into multiple SMs, a GPU is not a true SIMD device, as not all SIMD capable SMs are run simultaneously. CUDA emulates such a device by multithreading each SM such that, from the perspective of the developer, it can be seen and treated as a true SIMD device. Nvidia calls this multithreaded SIMD programming model single instruction, multiple thread (SIMT) [140].

2.7.2 OpenGL Rendering Pipeline

The Open Graphics Library (OpenGL) is a graphical rendering application programming interface (API) first introduced by Khronos in 1992 [141]. OpenGL provides a purpose built, GPU accelerated, graphical rendering pipeline which takes as input 3D coordinates, called vertices, and transforms them into 2D pixels on the computer screen called fragments. As shown in Fig. 15, OpenGL's rendering pipeline can be divided into four main sections where each section takes as input the output from previous sections.
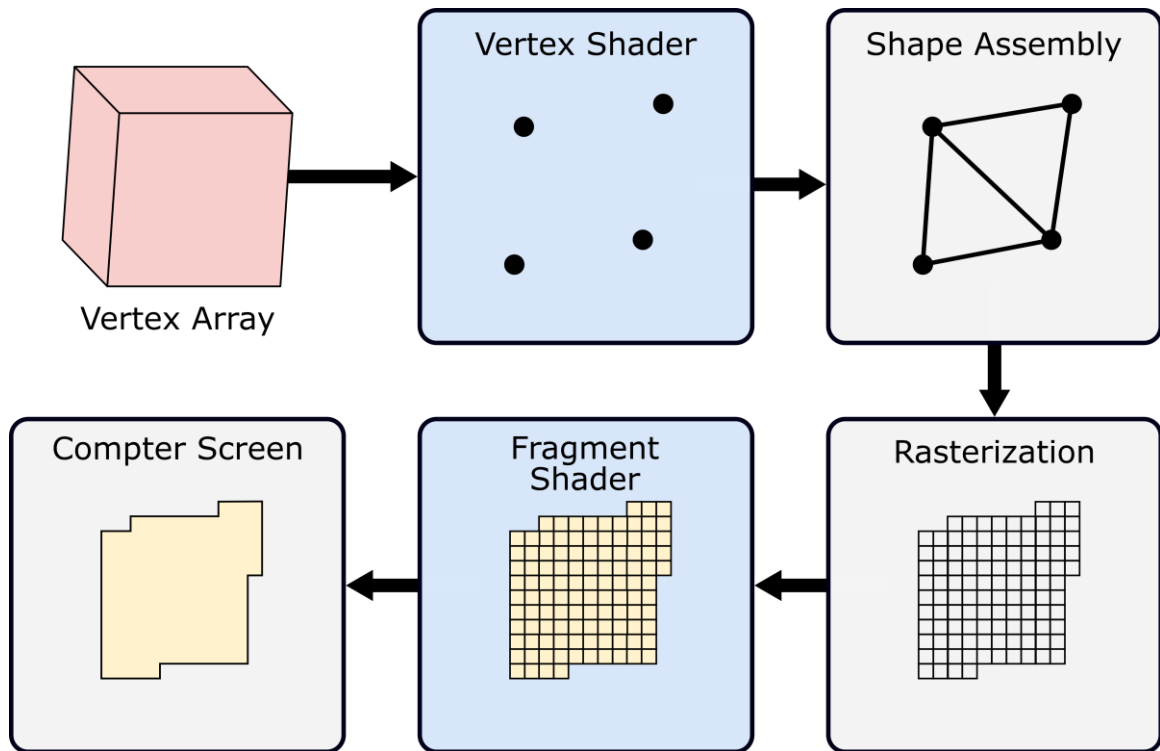
Fig. 15: OpenGL rendering pipeline. The job of the rendering pipeline is to take an array of 3D coordinates called vertices and transform them into 2D pixels or fragments to be displayed to the computer screen. The pipeline consists of four parts: The vertex shader, shape assembly, rasterization, and the fragment shader.

The first section of this pipeline is called the vertex shader, where it takes the given user vertices and applies any user specified transformations to prepare them for the next step of the pipeline. Besides 3D coordinates, each vertex can also be specified to contain several other attributes such as colour values or texture coordinates. Following this, the shape assembly section of the pipeline organizes the vertex data into sets of geometric shapes such as triangles. Each geometric shape is then mapped onto a set of 2D pixels (fragments) in a process called rasterization. Only the area bounded by the geometrical shapes as seen from the computer screen are mapped to pixels. The fragment shader is then responsible for colouring in each 2D pixel based on specified details of the final scene including lights, shadows, and vertex colour before displaying it to the computer screen.

The rendering pipeline is semi-automated in that developers only have programmatic control over the vertex and fragment shader. Furthermore, developers are constrained to work within the specified inputs and outputs of each shader and must abide by the general rules laid out by the overall pipeline. Thus, in contrast to CUDA, general purpose computations are not generally supported by OpenGL rendering pipelines. Later versions of OpenGL introduced compute shaders to provide some support for general

purpose computations, although it lacks many of the features and support otherwise found in CUDA [142].

2.7.3   GPU Memory

Optimizing GPU memory utilization is the most important area for achieving high performance in parallel algorithms [143]. Performance is maximized by optimizing for GPU memory bandwidth and latency. Bandwidth describes how much data can be transferred (typically in bytes per second) and latency describes the amount of time it takes to move the data (typically measured in clock cycles). The GPU memory space (graphics card) is physically separate from the host memory space (CPU) with the GPU memory itself being organised into four memory types: global, texture, shared, and registers (local thread memory) [140]. As data moves through this memory space, from host memory, through to local thread memory, bandwidth and latency improve significantly [144]. Therefore, an important principle in implementing an efficient parallel algorithm is to keep as much data local to the thread as possible and to limit its movement across this memory space.

Host to Device Memory Transfer

The first area of optimization is to minimize data transfer between the host and device. Bandwidth between host and device memory is limited by the physical PCIe slot connecting the two and is orders of magnitudes slower than the peak theoretical bandwidth of the GPU's global memory. For optimal performance, all data is transferred to the device once at the beginning of a computation and is left there for the duration, even if it means running kernels that do not provide noticeable performance gains as compared to their serial counterparts [143]. Further, for optimal performance, data transfers should be batched into a single large transfer instead of a collection of smaller ones to reduce memory overhead and improve transfer times during kernel initialization. All GPU memory should be allocated only once and should be reused during the duration of kernel execution as allocating and deallocating GPU memory is a slow process [143].

Global Memory

The most plentiful type of memory in the GPU is global memory which resides in the dynamic random-access memory (DRAM) outside of the SMs. Global memory has the

highest latency and lowest bandwidth of all the memory types due the physical characteristics of DRAM. Briefly, DRAM encodes bits of information using the charge of thousands of extremely small capacitors strung together into data lines, thus, its response time is limited by the rise time of the capacitors [136]. Data lines do not support random read-write access. If one wishes to access a particular data element one must read from or write to the entire data line. The length of a data line is typically 32-bytes, 64-bytes, or 128-bytes [137]. Therefore, to efficiently utilize global memory, data must be stored contiguously such that global access to these data lines are coalesced. Coalesced data access is the single most important part of an efficient algorithm [143]. Non-coalesced access would require the retrieval of multiple data lines even if the data required is less than the size of each data line. For example, if each thread in a kernel required 16 bytes of data but that data was spread out randomly across global memory, the GPU would, at worst, retrieve 16 separate data lines for each thread. If each data line was 128-bytes, this would equate to 2048 bytes ($2 \ kB$) of data. This inefficient organization of data within memory wastes memory bandwidth and dramatically increases kernel execution times. In contrast, if those 16 bytes of data were stored contiguously, then only one data line would have to be retrieved across multiple coalesced threads (1 data line per 8 threads).

Texture Memory

Texture memory is a special read-only memory space residing within global memory that is cached locally on the GPU. After texture memory data is initially retrieved, subsequent reads from texture memory do not incur further bandwidth costs [143]. However, for optimal performance, texture memory should still be accessed with coalesced memory reads. In addition to data caching, texture memory provides hardware accelerated support for filtering operations like nearest-neighbour and linear interpolation. Reads from texture memory can index data arrays using use floating point coordinates instead of discrete integer coordinates. When floating point coordinates are used data values are automatically interpolated on hardware, according to specified user preference, from nearby data elements. This can offer significant improvements to processing bandwidth as compared approaches that rely on software-based interpolation.

Shared Memory

Depending on the problem (for example matrix multiplication) coalesced memory access might not be possible. In such cases, shared memory can be used to avoid significant slowdown. Shared memory is located on each SM and, as the name suggests, is shared across all threads in a SM. Most importantly, shared memory does not suffer from coalescing problems and is significantly faster than global memory [144].

Local Thread Memory (Registers)

The fastest memory type available to threads are 4-byte registers that are used to hold frequently accessed variables private to each thread and persist for the lifetime of the kernel. While fast, registers are a scarce resource as each SM only has a limited number available that must be divided up and shared amongst threads. On older Fermi architectures, each thread was limited to 64 registers, while newer Ampere architecture allow up to 256 registers per thread [137,145]. While this may seem like a lot, it can quickly be used up when using larger data types and/or the previously mentioned tile method is used (due to the increased indexing overhead required). If a thread goes beyond this limit, any extra register usage will spill over to the local memory in DRAM. Contrary to its name, local memory is simply global memory specifically reserved for register spilling and as such it has the exact same properties as normal global memory. Therefore, one must avoid register spilling to maintain kernel performance [137].

2.8 Real-time Processing of Middle Ear SS-OCT Data on the GPU

Fig. 16 shows the chain of processing steps involved in transforming SS-OCT interferograms into A-lines that can in-turn be processed into displayable structural B-mode images and Doppler vibrometry lines. All these steps were implemented as separate GPU kernels using Nvidia's Compute Unified Device Architecture (CUDA).
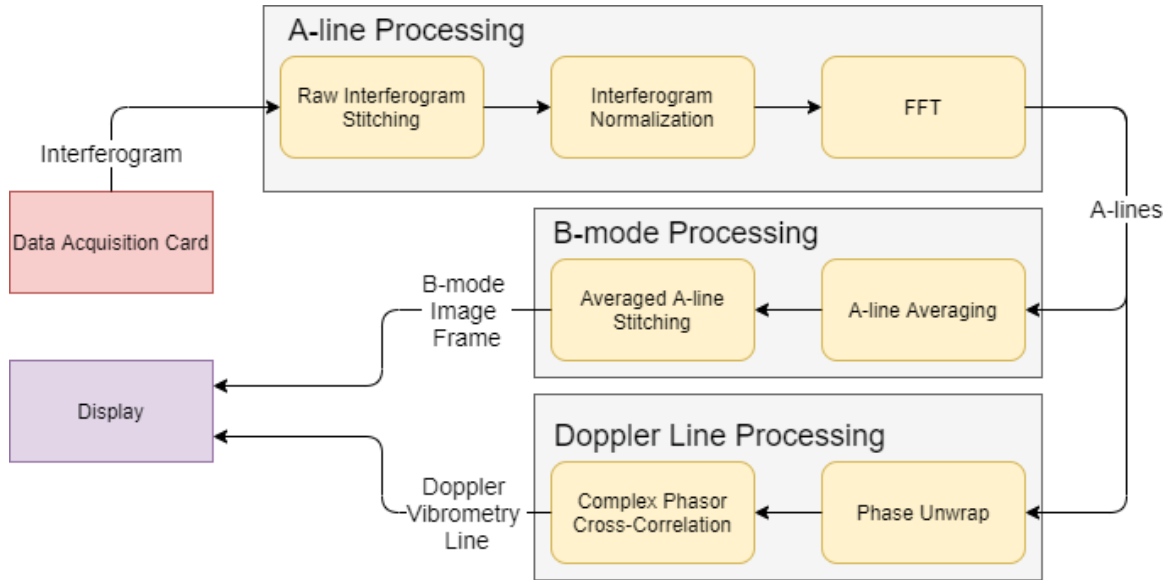
Fig. 16: Processing steps involved in transforming SS-OCT interferograms into A-lines, structural B-mode images, and Doppler vibrometry lines using the digital locked-in method. SS-OCT, swept-source optical coherence tomography.

Starting with A-line processing, depicted in Fig. 17, the raw interferograms acquired from the data acquisition card (DAQ) first need to be stitched together to remove non-linear segments of the laser sweep. The VT-DBR swept-source laser used in our system experiences non-linear mode-hops along its otherwise linear-in-frequency sweep. Fortunately, the laser's internal calibration process indicates where along the sweep these mode-hops occur which allows them to be cropped out. The valid parts are then stitched together to form a valid interferogram. Following interferogram stitching we apply a normalization vector across the interferogram. Normalization is an element-wise multiplication that corrects for systemic amplitude ripples and phase dispersion introduced through frequency dependent index of refraction within the optical system components and tissue. The spectrogram is then multiplied by a Hamming window to control dynamic range and point-spread-function (PSF) width. Following normalization, a 1D Fast Fourier Transform (FFT) is applied along the interferogram to convert its frequency encoded depth information into an A-line containing the spatial and phase information of the sample.
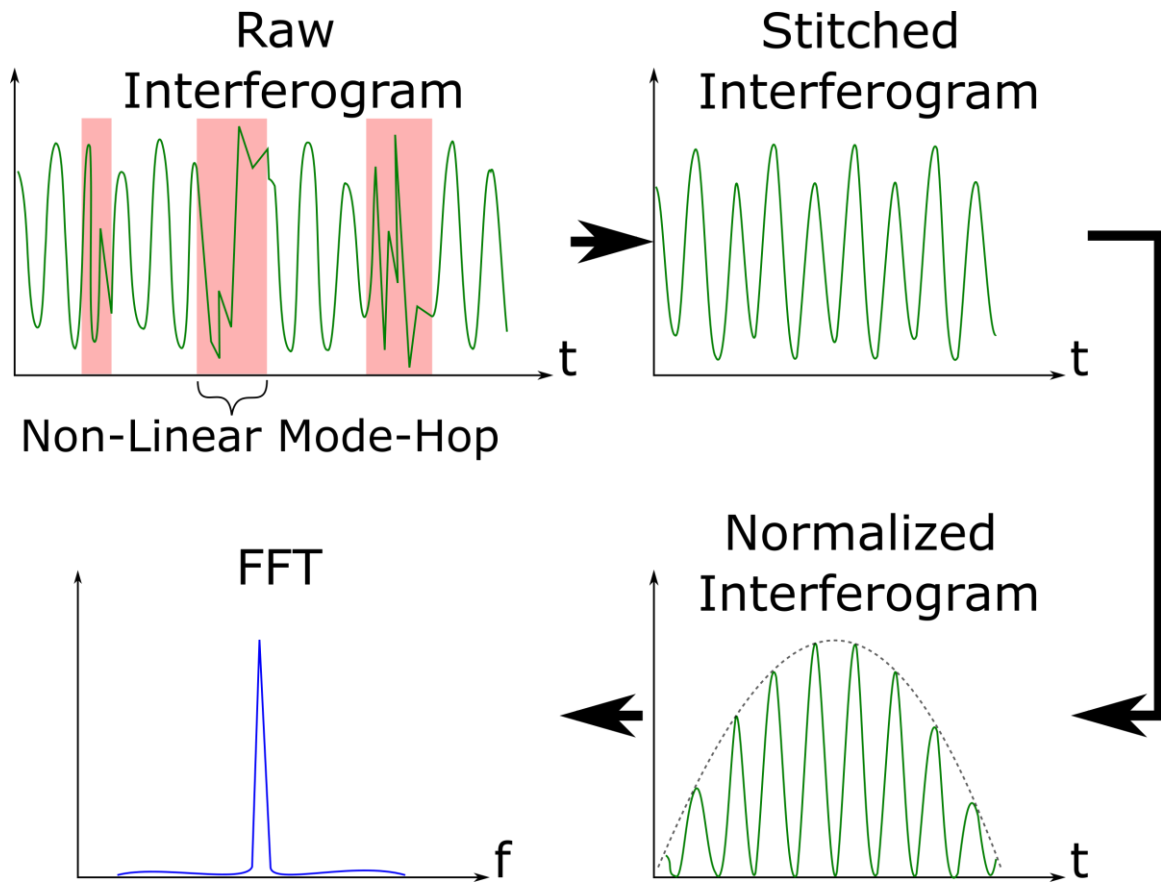
Fig. 17: Processing of SS-OCT interferograms into A-lines. From left to right the interferograms are stitched to remove non-linear laser mode-hops, normalized, and the FFT taken.

From A-lines we can either construct 2D, cross-sectional, structural B-mode images or extract Doppler vibrometry lines. To construct B-mode images we configure the DAQ to acquire groups of N sequential interferograms per image line and start scanning the laser laterally across the sample. These N interferograms are then run through the A-line processing steps for the removal of invalid data, stitching of remaining data, normalization, windowing and FFT to produce a group of N complex valued A-lines. In the first step of B-mode processing, the group of N A-lines are combined through complex averaging to produce a single image line. These image lines are then stitched together laterally to form the final B-mode frame depicting the cross-sectional structural information of the imaged sample. We average together sequential A-lines instead of sequential B-mode frames due to limitations in how fast the scanning mirror can transverse across the field of view.
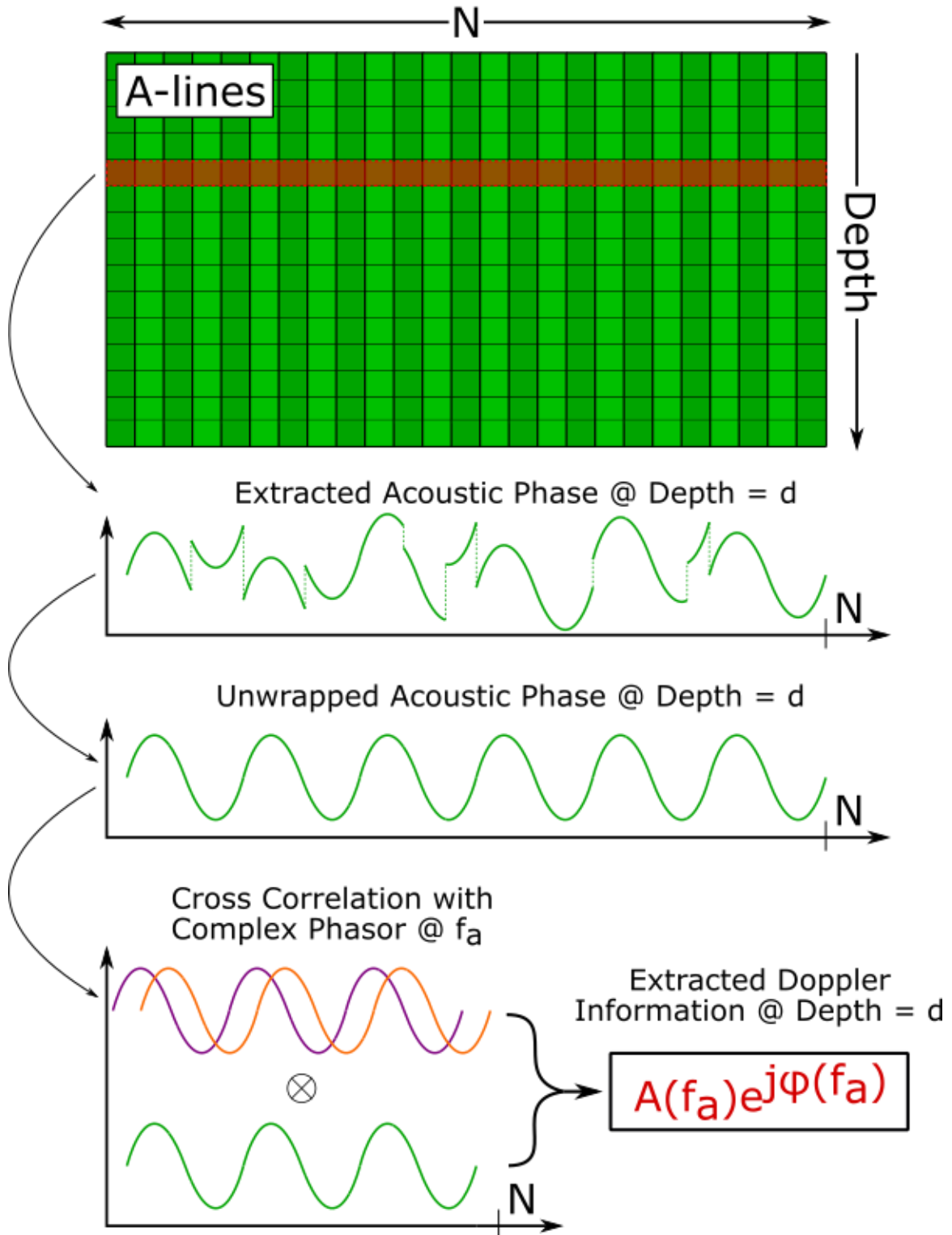
Fig. 18: Extraction of Doppler vibrometry information from A-lines. From top to bottom: phase information at $Depth = d$ is extracted across N adjacent A-lines, extracted phase is unwrapped, unwrapped phase is cross-correlated with a complex phase at the stimulus frequency of interest f_a resulting in the recovered Doppler information at d. This process is done at each depth to construct a full Doppler line.

At the time the work for Chapter 3 was completed, we utilized a digital lock-in method [1] to extract Doppler vibrometry information from the processed A-lines. Illustrated in Fig. 18, digital lock-in Doppler vibrometry consists of two processing steps: phase unwrapping and complex phasor cross-correlation. In the first step, a non-iterative phase unwrapping algorithm is applied across a group of N unaveraged A-lines at each pixel/depth. The number of A-lines N is chosen such that an integer number of acoustic periods of the applied stimulus tone is captured within the A-line group. Following unwrapping, the phase at each pixel is multiplied by a complex phasor at the stimulus frequency of interest and the product summed across the N lines. The result being a single depth-resolved Doppler vibrometry line containing the magnitude and phase of the acoustic response of the middle ear to the applied sound stimulus. Doppler vibrometry imaging continues until the running average of extracted Doppler lines reaches the desired noise floor (typically $100\ pm$ RMS noise floor is reached in 2 seconds for in-vivo imaging).

In the current ME-OCT imaging system, as described in Section 1.1.1, we use a FFT-based Doppler vibrometry method as originally described by Applegate et al. [19]. Here the Doppler vibrometry information is extracted by first calculating the depth-resolved phase difference between N adjacent A-lines followed by taking the FFT across this phase difference for each depth in the resultant vibrometry line. As in the digital lock-in method, the number N is chosen such that it contains an integer number of acoustic periods of the applied sound stimulus. The resultant Doppler vibrometry lines are then combined via a running average until the desired noise floor is reached. An advantages of the FFT method, as compared to the digital lock-in method, is that it requires no phase-unwrapping which simplifies the implementation, and that it provides the complete depth-resolved frequency spectrum of the middle ear's acoustic response to the applied sound stimulus instead of the magnitude and phase at a single frequency.

Chapter 3

Real-time Rendering Techniques for Volumetric Middle Ear Optical Coherence Tomography

3.1 Abstract

In this chapter we introduce the design and development of a real-time volumetric rendering engine for ME-OCT. We then apply this rendering engine to the development of novel methods for the visualization of ME-OCT data, namely digital tympanotomy and animation of ME-OCT Doppler vibrometric datasets. The rendering engine uses a single-pass ray-caster technique implemented using OpenGL to produce direct visualizations of ME-OCT datasets at real-time rates. Conversion of ME-OCT datasets to bitmaps suitable for rendering was implemented using CUDA as were the digital tympanotomy and Doppler animation visualization features. Validation of the Doppler animation was performed on a cadaveric middle ear.

3.2 Introduction

Due to the complex 3D structure and functional dynamics of the middle ear, 3D structural and functional visualization of the middle ear at real-time interactive rates has significant diagnostic value. While several groups have demonstrated real-time 2D ME-OCT [146,147], they did not have the means to render 3D datasets in real-time [19,148]. For ophthalmic OCT, Carrasco-Zevallos et al. [149] demonstrated real-time 3D visualization in-vivo utilizing a GPU accelerated ray-casting approach [150]. In this study we developed the first real-time 3D visualization engine for ME-OCT using a similar ray-casting approach to render 3D structural and functional ME-OCT data in real-time, the results of which were published in [1].

There exists several excellent open-source medical rendering toolkits such as 3D Slicer [151,152], ITK [153], and VTK [154] that are specifically tailored for rapid extensibility with the main purpose of offline (non-real-time) clinical imaging research [152,155,156]. Although these are well developed, mature software libraries, at the time

50

this work was completed they had minimal native support for GPU acceleration via OpenGL or CUDA, making them inadequate for real-time processing [157,158]. These toolkits tend to be used for offline post-processed 3D imaging tasks such as viewing and manipulating CT or MRI data.

Volumetric ray-casting has several benefits over other rendering techniques. First, it preserves the geometric relationships of structures within the dataset. This means that geometrically faithful measurements can be made directly from the 3D render. Second, rendering time is independent of the dataset contents, allowing easier separation of processing and rendering code. Finally, ray-casting renders via line-of-sight, directly emulating how ME-OCT data is acquired. The ray-casting method lends itself to efficient implementation on modern GPU hardware.

To provide context for the design choices made in the ray-casting engine development, in the sections below we briefly describe the visualizations that we generate with the engine. Engine design details follow subsequently in Section 3.3.1.

Digital Tympanotomy

At the wavelength of light used in our SS-OCT system $1550\ nm$, we can image middle ear structures through the intact TM with a two-way signal loss of approximately $12\ dB$ [1]. With a typical dynamic range of $50\ dB$, this leaves roughly $38\ dB$ of dynamic range for imaging structures lying distal to the TM including the ossicular chain and supporting soft tissue structures. A useful feature of the ray-casting approach applicable to middle ear imaging is the ability to remove voxel data along the ray-casting directions, allowing proximal structures (i.e., the TM) to be removed from the render. In analogy to exploratory tympanotomy, a surgical procedure in which the tympanic membrane is cut and lifted away to visualize the middle ear space, we call this technique digital tympanotomy. Digital tympanotomy (Fig. 19a) provides clinicians with a non-invasive view of the ear like what they would see in the operating room after the TM has been cut and lifted as shown in Fig. 19b.

51
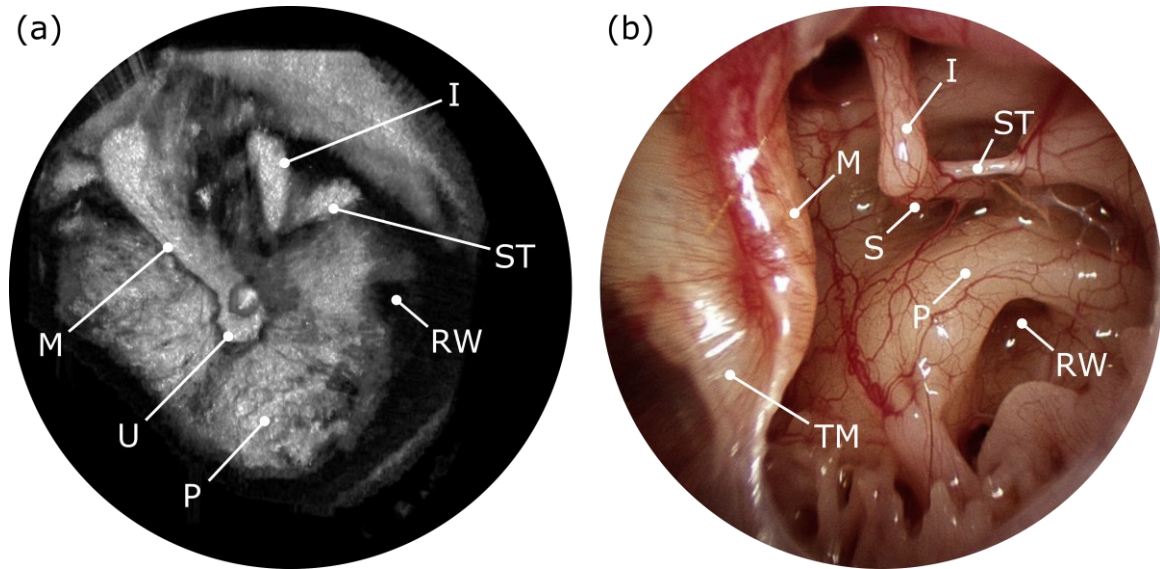
Fig. 19: Digital tympanotomy vs exploratory tympanotomy. (a) Enface view of the underlying middle ear structures obtained through digital tympanotomy. (b) Endoscopic view of the middle ear structures obtained during surgery. Derived from [36] and reproduced with permission. M, malleus; U, umbo; TM, tympanic membrane; S, stapes; I, incus; ST, stapedius tendon; P, promontory; RW, round window.

Doppler Animation

Doppler vibrometry datasets are raster-scanned datasets in which multiple A-lines are collected at each spatial location across the sample being imaged along a uniform grid. Recall from Section 1.1, that if sound is played during data collection and the structures being imaged move in response to that sound, then the phase of the OCT A-lines will be modulated by the sound. The magnitude and phase of this modulation can be extracted for each voxel, as outlined in Section 2.8, and used to create an animation showing exaggerated motion of the middle ear structures in response to this sound. This provides an intuitive picture of the ear's response to sound, including any pathological response such as fixed points that should normally vibrate or discontinuities exhibiting excess vibration and decoupling from structures they should be coupled to. Because the ME-OCT system only measures a projection of the motion along the direction of the A-line, only the axial velocity component of motion can be accessed and animated in this way rather than the full 3D velocity vector. As a result, the exaggerated animation contains some quantitative inaccuracy, but it is nevertheless useful in conveying a qualitative picture of the middle ear response.

The rendering engine supports displaying mobility amplitude as a colour overlay (Fig. 20) or displaying both amplitude and phase through an animation. Either of these

52

operations can be performed with minimal computational load within the ray-casting approach.
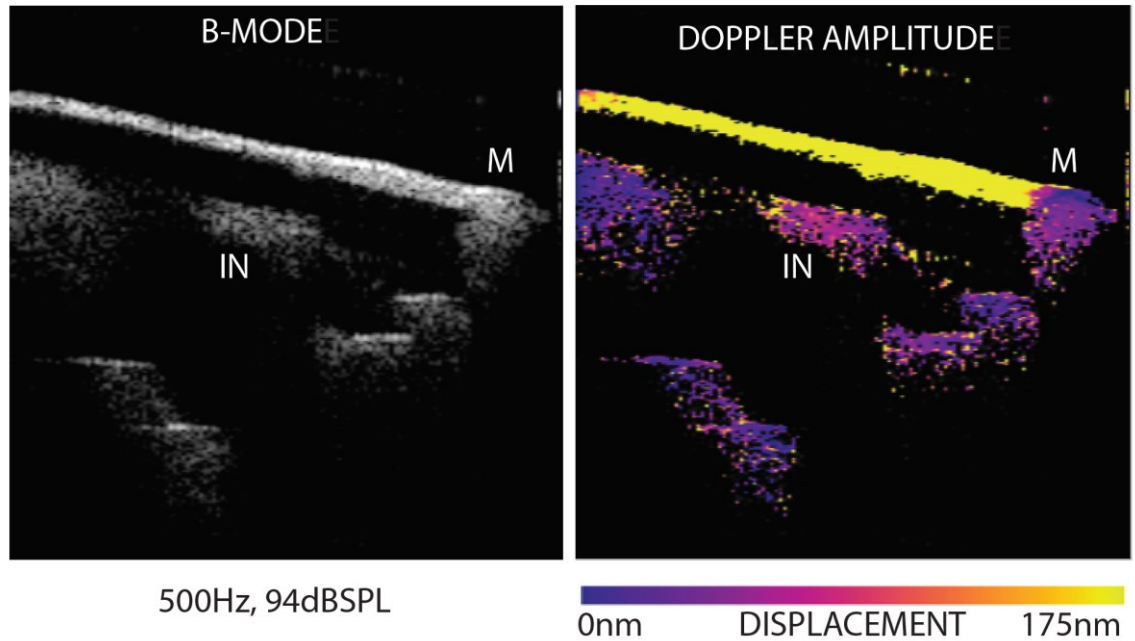
Fig. 20: False colour overlay onto a B-mode (left) representing Doppler vibrational amplitude (right) *[159]*.

3.3 Methods

3.3.1 Volumetric Rendering Utilizing GPU Accelerated Ray-Casting

Before rendering volumetric ME-OCT data, a suitable image volume needs to be constructed. This is done by raster scanning across the ME volume of interest and stitching together the resulting cross-sectional 2D B-mode image stack into a 3D dataset. As shown in Fig. 21, to display a 3D dataset, a view of the data is projected onto the 2D viewing surface of the computer screen, and the user manipulates the dataset to view it from different angles.
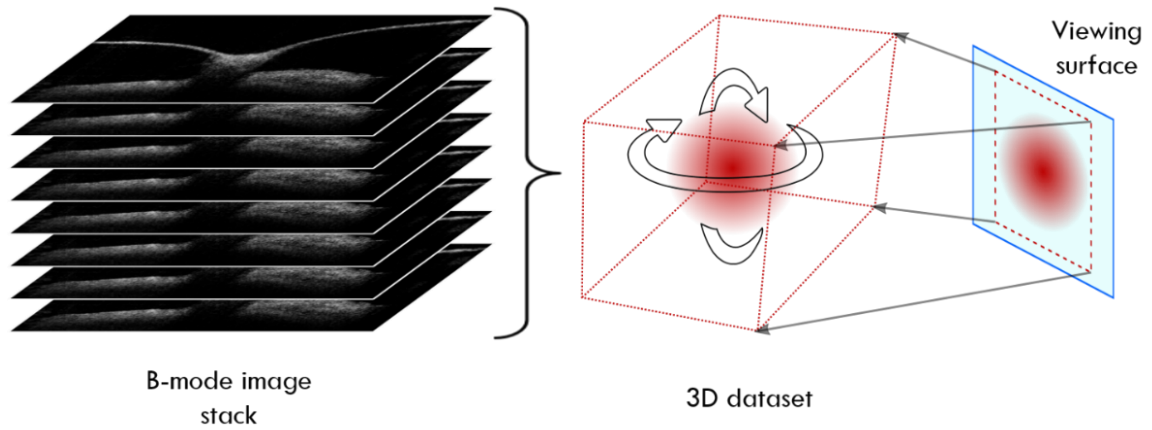
Fig. 21: Volumetric dataset construction from a stack of B-mode images and its manipulation and projection onto a 2D viewing surface such as the computer screen.
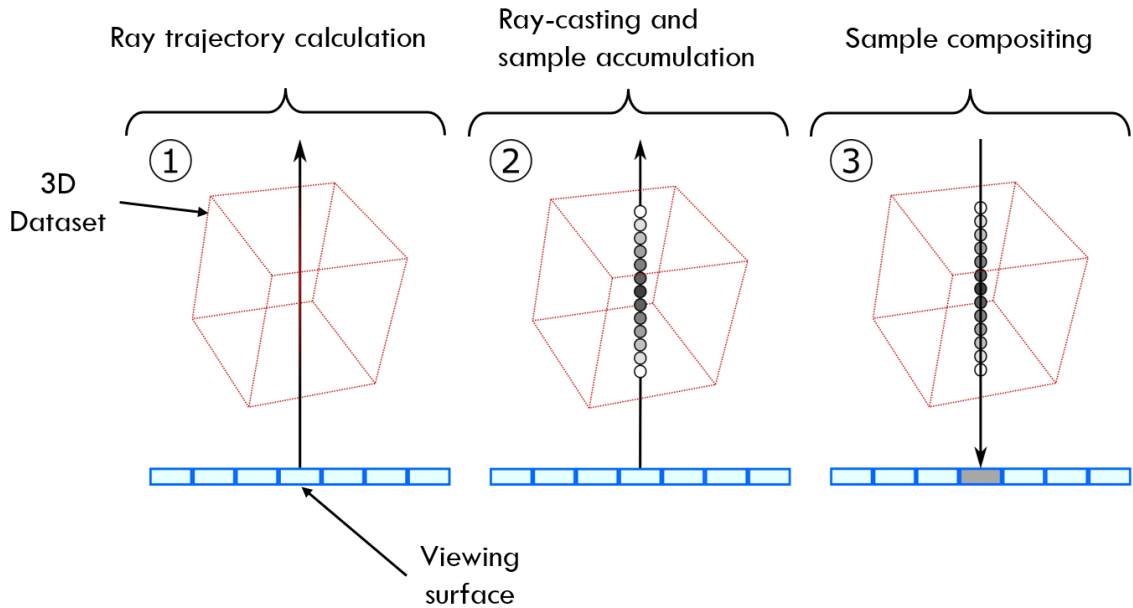
Fig. 22: Single-pass volumetric ray-caster rendering steps including: (1) Ray trajectory calculation, (2) ray-casting and sample accumulation, and (3) sample composition.

A single-pass ray-casting technique [160] was implemented using OpenGL to render 3D ME-OCT datasets in real-time. Ray-casting is an optical model describing the emission and absorption of light rays through a volume [161]. Single-pass refers to the number of rendering iterations required in an image frame to fully display the volume. Initially, a simple geometrical shape, a bounding box, is rendered at the location in the 3D scene where the volumetric data is to be displayed as shown in Fig. 21. Transformations applied to this bounding box directly affect the orientation of the displayed volumetric dataset. From here the first step of ray-casting, Fig. 22, involves calculating the trajectory

of 3D vectors called 'rays' through the dataset originating from each pixel in the viewing surface. Each ray is described according to the parametric ray equation:

$$ray(t, x, y, z) = o + t\, d(x, y, z) \qquad\qquad 6$$

Where o is the ray origin, $d(x, y,\ z)$ is the normalized ray direction vector, t is the parameterized location of the ray along the ray direction, (x, y) are the screen's pixel coordinates and (z) is the depth. Computation of each ray's direction vector is done by subtracting the bounding box's (x, y, z) coordinates from the ray's origin (the 2D viewing surface). Next, the minimum and maximum extents each ray needs to traverse are determined. This is done using the slab method [162] to determine intersection points between the ray and the bounding-box t_{min}, t_{max}. The second step, Fig. 22, involves casting the rays through the volume and accumulating sample data at evenly spaced intervals along the min and max extents of the ray. Accumulation consists of an optical integral representing accumulated emission and absorption of light along the ray. The optical integral depends on the optical model used with alpha-blending being the most common as it gives a realistic appearance to solid surfaces [163]. Alpha-blending computes the ray's accumulated colour intensity I through a weighted average of accumulated sample points according to [163]:

$$I = \sum_{i=0}^{n} \alpha_i C_i \prod_{j=0}^{i-1} (1 - \alpha_j) \qquad\qquad 7$$

Where C_i represents a sample point's colour as a red, green, and blue (RGB) value and α_i represents a sample point's opacity as an alpha value giving each sample point a 4-channel RGB-alpha (RGBA) colour spectrum. Each RGBA sample is weighted according to the ray's accumulated opacity α_i at the sample's current location i along the ray and by the product of the transmission coefficients $(1 - \alpha_j)$ at all points j proximal to i. The number of sample points n along the ray trajectory is chosen as to balance rendering performance with quality where a large value of n increases rendering quality at the expense of performance and vise-versa. In this work we used $n = 256$ sample points. For B-mode ME-OCT images, all three RGB colour channels are set to the same value resulting in a greyscale image. However, for other functional imaging modes such as Doppler or

angiography these three colour channels can be used to produce false-coloured overlays on top of the B-mode data.

Within this ray-casting framework, illumination of the volume by an external light source can be modeled by multiplying each sample's intensity by a term representing the illumination effects for each colour in the Blinn-Phong shading model [164]:

$$I = \sum_{i=0}^{n} \alpha_i C_i (C_a + C_d + C_s) \prod_{j=0}^{i-1} (1 - \alpha_j) \qquad 8$$

The Blinn-Phong shading model (Fig. 23d) is comprised of a sum of three contributions to illumination: ambient illumination C_a (Fig. 23a) which represents global illumination, diffuse illumination C_d (Fig. 23b) representing uniform directional lighting, and specular illumination C_s (Fig. 23c) representing specular reflections from a point source. For this work, the light source was placed behind the viewing screen to simulate the way illumination is delivered in surgical microscopes used in otomicroscopy. For ME-OCT volumes the Blinn-Phong shading model can enhance otherwise difficult to spot features in the volume. For example, Fig. 23f and Fig. 23e shows a ME-OCT volume acquired in-vivo from a healthy adult volunteer rendered with and without Blinn-Phong shading respectively. As can be seen, the Blinn-Phong shading in Fig. 23f highlights more detail on the surface of the TM as compared to Fig. 23e and provides an enhanced sense of depth.

The final rendering step, Fig. 22, involves taking the accumulated value of the computed optical integral I and compositing it back onto the 2D viewing surface once all the sample points n in the ray are traversed. Alternatively, the optical integral calculation, Eq. (7), can be optimized by terminating and compositing the ray early after the accumulated alpha channel α_j is saturated (i.e., fully opaque). As light can no longer be absorbed by more distal points along the ray, the rest of the ray extent does not affect the intensity or color of the pixel on the screen and traversal can be terminated.
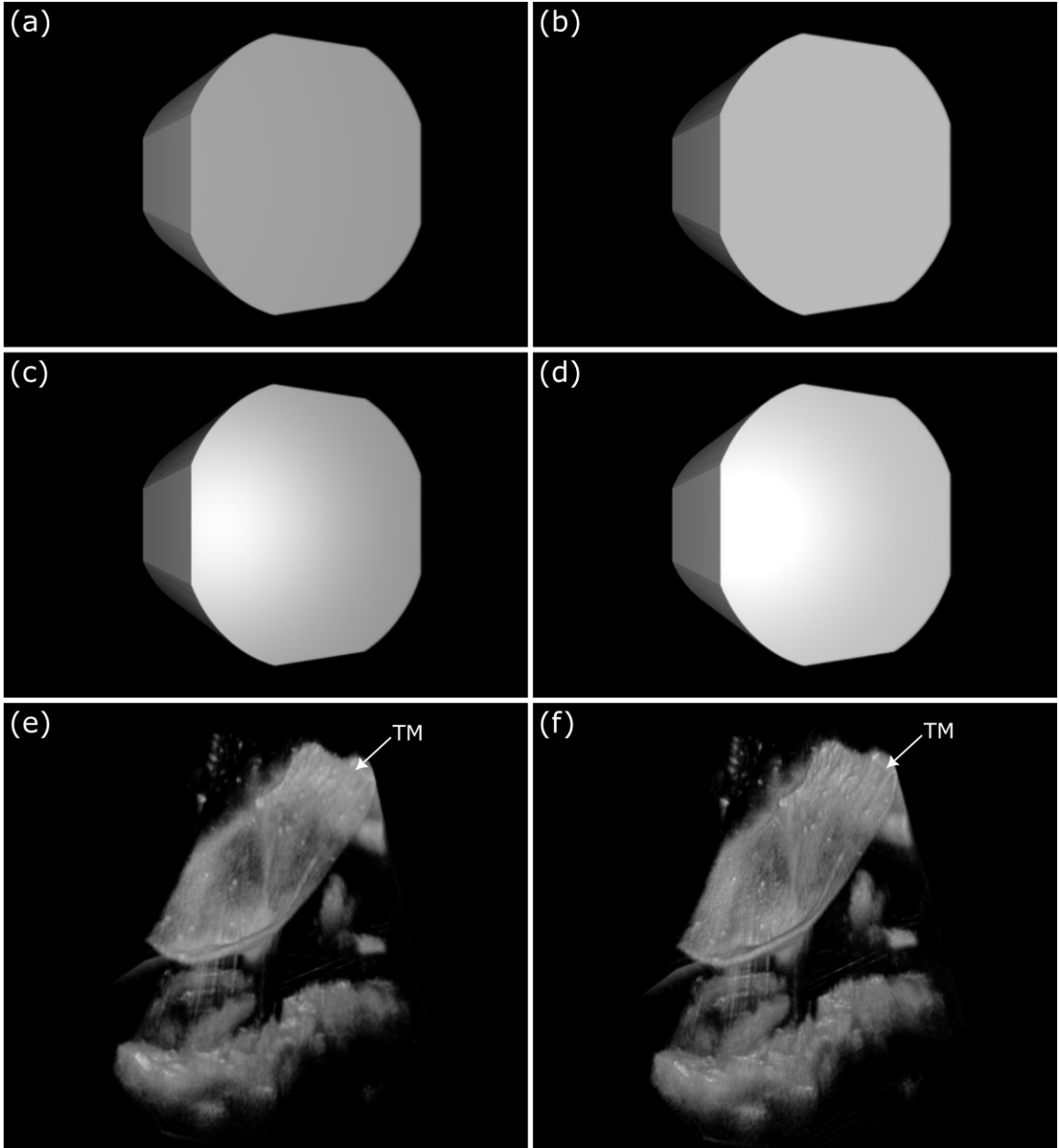
Fig. 23: Blinn-Phong shading model is composed of three light contributions including (a) ambient, (b) diffuse, and (c) specular. (d) Combined shading from all three light contributions. (e) ME-OCT volume collected from a healthy adult volunteer rendered without Blinn-Phong shading. (f) The same ME-OCT volume as (e) rendered with Blinn-Phong shading. TM, tympanic membrane.

The single-pass ray-caster expects as input a volumetric bitmap of RGBA colour values. To convert the ME-OCT data volume to this volumetric bitmap format, a CUDA-based data preprocessor was integrated into the rendering engine. The preprocessed volumetric bitmap was handed over to the ray-caster using OpenGL-CUDA interop functionality which simply transfers ownership of the data from CUDA to OpenGL. This avoids any memory transfer overhead and prevents any latency between bitmap generation

57

and display. Using CUDA for bitmap generation also enables other, more sophisticated preprocessing steps such as those required for digital tympanometry and Doppler Animation to be performed prior to rendering. These will be explored in the next two sections.

3.3.2 Segmentation and Removal of the Tympanic Membrane in ME-OCT Volumes

For digital tympanotomy, identification and segmentation of the TM is performed within the CUDA preprocessor. For each image line in the volume a ray is cast starting from the top of the image line and ending at the bottom of the image line. Each cast ray utilizes a pre-configured threshold to detect when there is a sharp discontinuity in voxel intensity (i.e., an air-tissue interface) to identify the most proximal intersection of the ray and the TM. From this most proximal TM voxel, a user-specified number of voxels are set to be fully transparent thus removing them from the image and allowing the user to visualize more distal structures when looking at the middle ear enface.
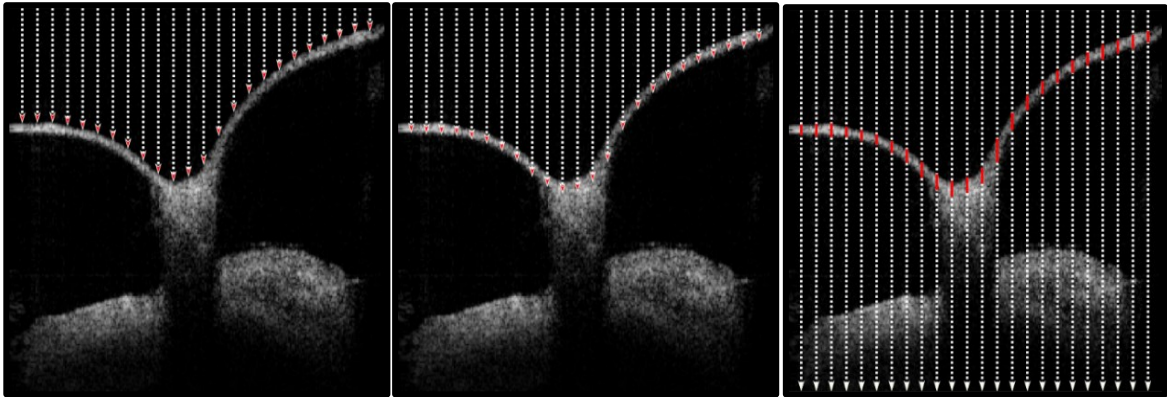
Fig. 24: Identification and segmentation of the TM by each cast ray as it traverses from the top of the ME-OCT dataset to the bottom.

The digital tympanotomy segmentation process is illustrated in Fig. 24. Voxels identified as part of the TM are removed by setting their alpha values to zero such that they no longer contribute to the optical integral, Eq. (7), during rendering. Due to the parallelizability of this approach, the entire TM can be identified and segmented in real-time while simultaneously being rendered through volumetric ray-casting. As the ray-caster only works on the RGBA bitmap volume output from the CUDA preprocessor, it is agnostic to what preprocessing steps have been performed. As a result, removal of the TM does not impact rendering performance.

3.3.3 Recreation of the Middle Ears Acoustic Response using Doppler Vibrometry Datasets

The generation of exaggerated animated motion of the middle ear's acoustic response was performed in the CUDA-based preprocessor. The first step in animating the dataset is to calculate new positions along the z-dimension for each sample in the dataset, n in Fig. 25, based on an exaggerated range of periodic vibrational amplitude. The new position of each voxel is calculated according to:

$$n_i = n_{(i-1)} + A_D M \cos(2\pi i S - \phi_D) \qquad\qquad 9$$

where A_D is the measured Doppler vibration amplitude, M is the motion exaggeration factor, i is the running frame index, S is the animation speed factor, and ϕ_D is the measured Doppler vibrational phase offset. Doppler vibrational phase is used to offset the animation starting position for each sample in the dataset according to the measured phase of its response. The running frame index multiplied by the speed factor determines the rate at which phase accumulates and the pixels move. The motion exaggeration factor determines how much the displacement is exaggerated in the animated visualization as compared to its measured value. Without the exaggeration factor, the motion of middle ear structures would be imperceptibly small.
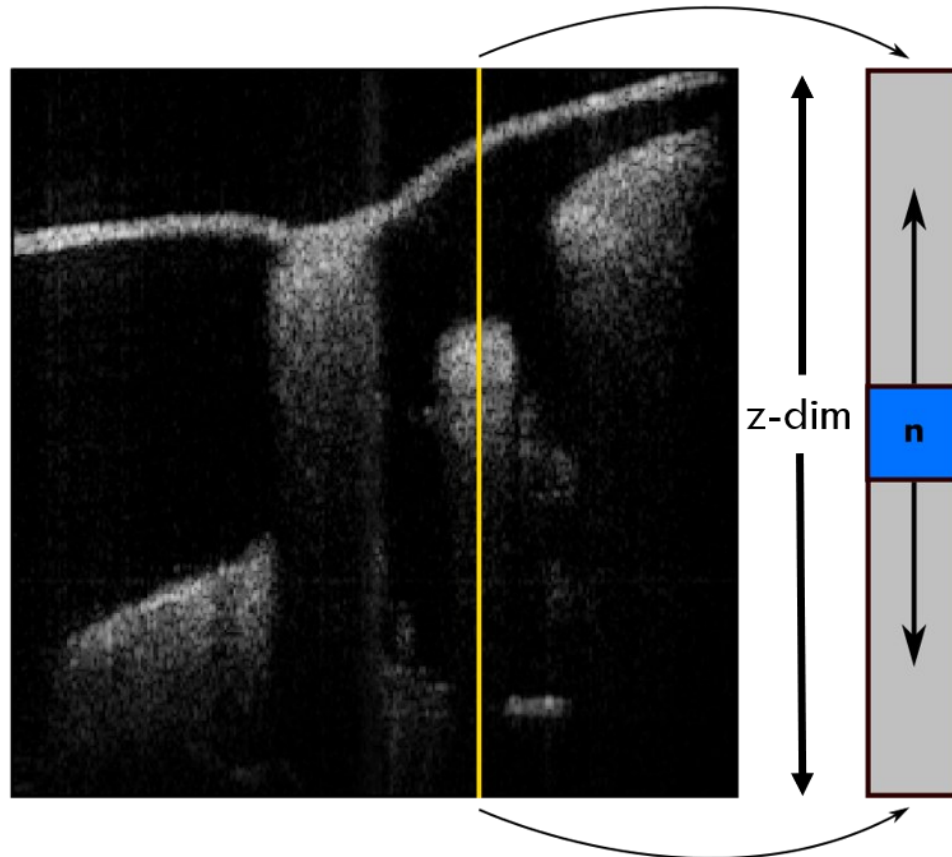
Fig. 25: Doppler animation where a new position is calculated for every sample position n along the depth dimension z within each image line based on the Doppler vibrometry dataset.

Applying motion to the data samples creates the possibility of collisions between a data sample that has been displaced from its original position due to motion and a data sample that is stationary or between two data samples that move out of phase towards each other. To resolve these collisions, we introduced a reconciliation procedure where the data sample with the highest intensity that occupies a given voxel 'wins' the voxel and determines its intensity. This simple approach was found to provide good visual fidelity without visually obvious artefacts associated with collisions. As the animation calculation is applied along independent image lines it can be applied to both cross-sectional and volumetric Doppler vibrometry ME-OCT datasets.

3.4   Results

Fig. 26 shows a collection of real-time ray-cast renders of volumetric ME-OCT datasets collected in patients. As can be seen, the rendering engine produces clear, qualitatively accurate, representations of acquired ME-OCT volumetric datasets for various commonly

encountered situations including healthy volunteers (Fig. 26a), TM perforations (Fig. 26b), and visualization of synthetic materials located within the middle ear such as tympanostomy tubes with granulation tissue (Fig. 26c). An accompanying video of Fig. 26a can be found in Supplementary File A showing the rendering engine facilitating real-time rendering and manipulation of the ME-OCT volumetric dataset. The rendering engine was capped at 60 frames per second (FPS) to provide smooth, lag-free, manipulation of the datasets and prevent over utilization of GPU resources for other processing applications.
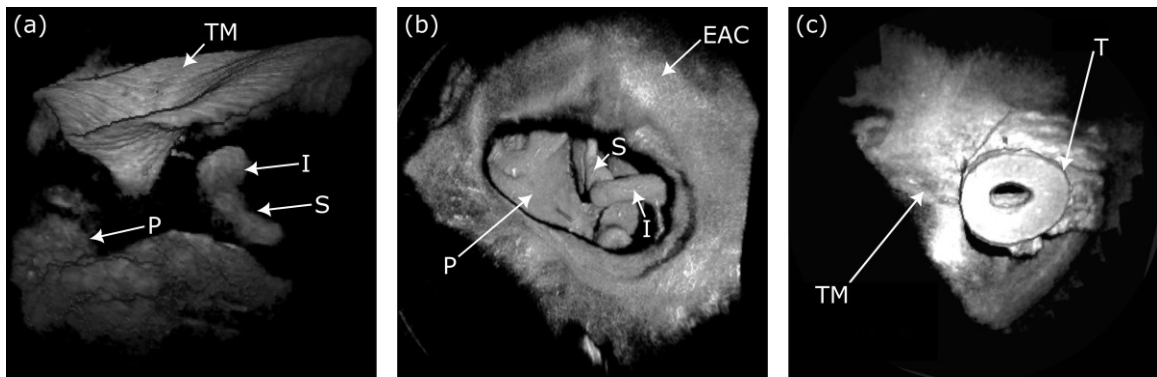
Fig. 26: In-vivo volumetric ME-OCT renders of (a) a healthy adult volunteer, (b) a perforated TM exposing underlying middle ear structures, and (c) a TM with an inserted tympanostomy tube. From [1,159]. A video of (a) can be found in Supplementary File A. TM, tympanic membrane; I, incus; S, stapes; P, promontory; EAC, external auditory canal; T, tympanostomy tube. Adapted with permission from [1] © Optical Publishing Group.

Digital Tympanometry

Fig. 27 depicts a series of snapshots of an in-vivo ME-OCT volume render using digital tympanometry to remove the TM and observe the underlying middle ear structures. The number of pixels removed is manually controlled by the user as to account for TMs of varying thicknesses. Identification and segmentation of the TM works best in situations where a clearly delineated TM surface is present. The video corresponding to Fig. 27 can be found in Supplementary File B.
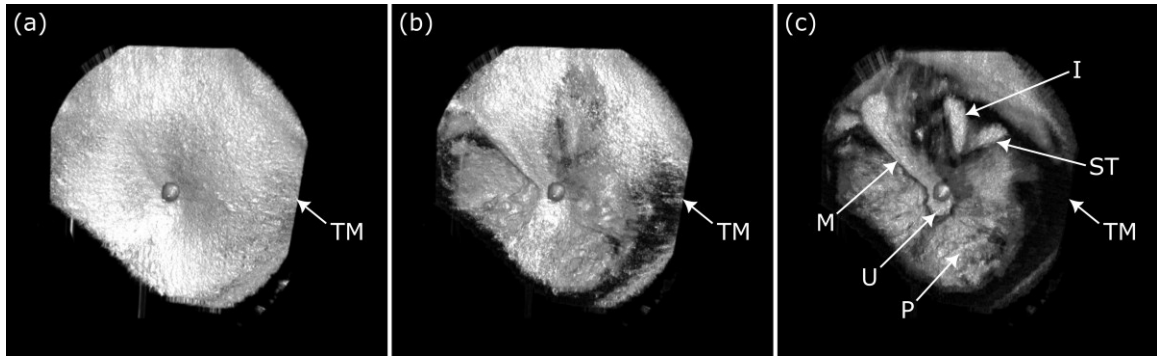
Fig. 27: ME-OCT volume render undergoing digital tympanotomy from (a) the initial enface view, to (b) TM partially removed to (c) TM fully removed and the underlying middle ear structures visualized. The associated digital tympanotomy video can be found in Supplementary File B. TM, tympanic membrane; M, malleus; U, umbo; P, promontory; ST, stapedius tendon; I, incus.

Doppler Animation

Fig. 28 depicts the cross-sectional (Fig. 28a − Fig. 28c) and volumetric animation (Fig. 28d − Fig. 28f) of a ME-OCT Doppler vibrometry dataset of an ex-vivo middle ear cadaver model with the external ear canal partially drilled away to aid in alignment during imaging. As outlined in [165], the dataset was configured to contain $128 \times 128 \times 330$ voxels covering a FOV of $10\ mm \times 10\ mm$ and took $\sim 70\ minutes$ to acquire. A pure tone was presented through a tube speaker (Etymotic Research, ER3A), located within $3\ mm$ of the TM, at a stimulus frequency of $500\ Hz$ with a nominal acoustic pressure of $100\ dB\ SPL.$. Sound pressure was regulated through a calibrated tube microphone (Etymotic Research, ER7C Series B). A false colour overlay corresponding to a dynamic displacement range between $0 - 1000\ nm$ was applied to the cross-sectional Doppler animation. The motion scale factor M was selected such that the reconstructed acoustic response of the middle ear stayed within the top boundary of the renderer. Any animation moving data outside of this boundary resulted in clipping of the dataset. The animation speed factor S was chosen to allow easy perception of the simulated acoustic response. Although, any value could have been chosen to suit personal preference so long as it resulted in animation that would be smoothly perceived within the 60 FPS cap of the renderer. Videos of both the cross-sectional and volumetric Doppler animations can be found in Supplementary File C and Supplementary File D respectively.
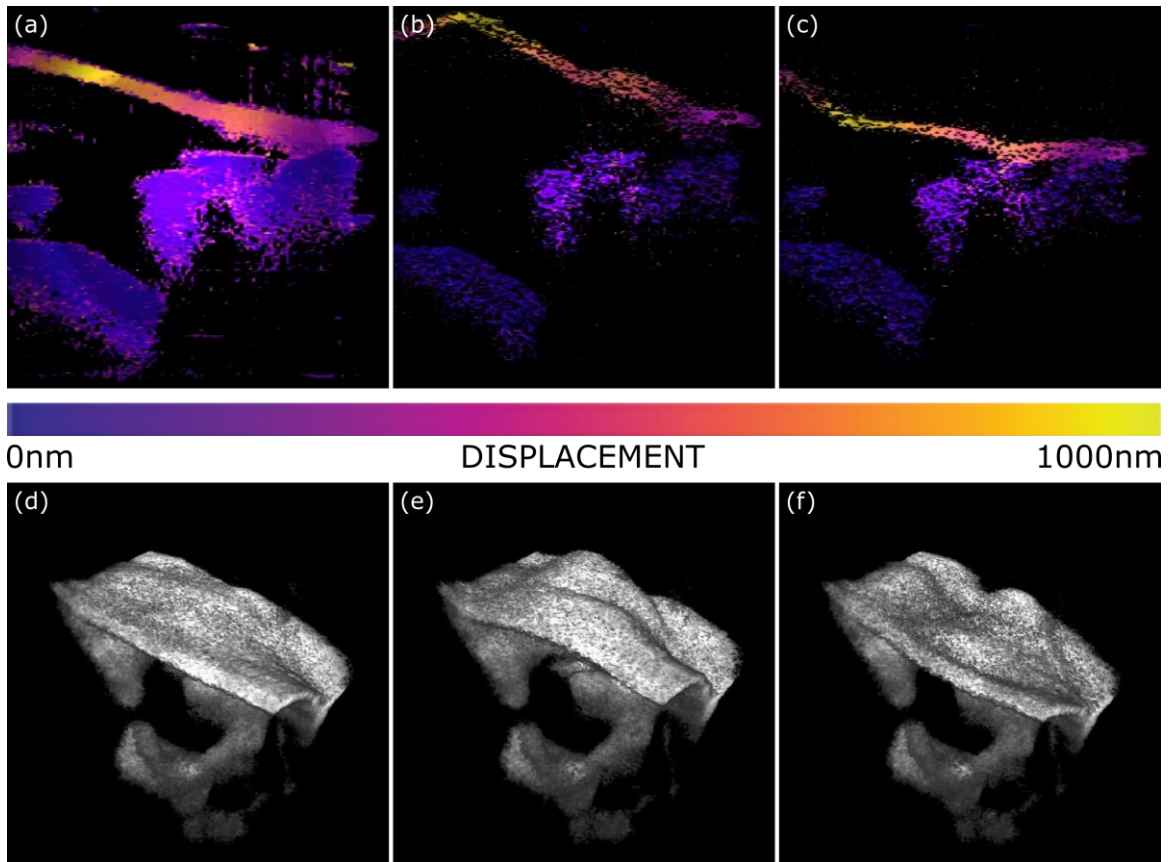
Fig. 28: Cross-sectional and Volumetric animation of a ME-OCT Doppler vibrometry dataset of an ex-vivo middle ear cadaver model acquired at a stimulus frequency of 500 Hz with a nominal pressure of 100 dB SPL. (a) Cross-sectional view of a central slice from the ME-OCT Doppler vibrometry dataset with a false colour associated with vibrational amplitude displacement overlaid. (b) Snapshot of the cross-sectional animation when the TM vibrational amplitude is at its positive peak. (c) Snapshot of the cross-sectional animation when the TM vibrational amplitude is at its negative peak. (d) Volumetric render of the ME-OCT Doppler vibrometry dataset before animation. (e) Snapshot of the volumetric animation when the TM vibrational amplitude is at its positive peak. (f) Snapshot of the volumetric animation when the TM vibrational amplitude is at its negative peak. From [1]. Videos of both the cross-sectional and volumetric Doppler animation can be found in Supplementary File C and Supplementary File D respectively. Adapted with permission from [1] © Optical Publishing Group.

3.5 Discussion

The rendering engine produced in this study has been extensively used by our group and collaborators in several research studies requiring real-time 3D visualization [1,24,111,166,167]. The volumetric Doppler animation feature was used by a collaboration with Golabbakhsh et al. [165] comparing measured middle ear response in human cadaveric temporal bones to the response obtained from finite-element-models of the human middle ear built using micro-CT data acquired for the same cadaveric ears.

3.5.1 Digital Tympanotomy Limitations and Improvements

While digital tympanotomy produces striking results when applied to certain datasets, its current implementation limits its usage to intact TMs and relatively noise free images.

Digital tympanotomy is applied uniformly across the lateral FOV with each ray being segmented independently of other rays. Thus, in situations with a perforated TM, the surfaces of underlying structures exposed by the perforation will be erroneously removed. Furthermore, in the presence of point-spread function artefacts at locations where the TM surface is normal to the OCT imaging axis [167], the surface of the TM is obscured (Fig. 29) and digital tympanotomy fails. A solution to both issues would be to apply a more sophisticated segmentation approach that treats the TM as an extended, connected contour. Such an approach could also avoid the requirement for the user to specify the depth of the TM by automatically determining the axial extent of the TM.
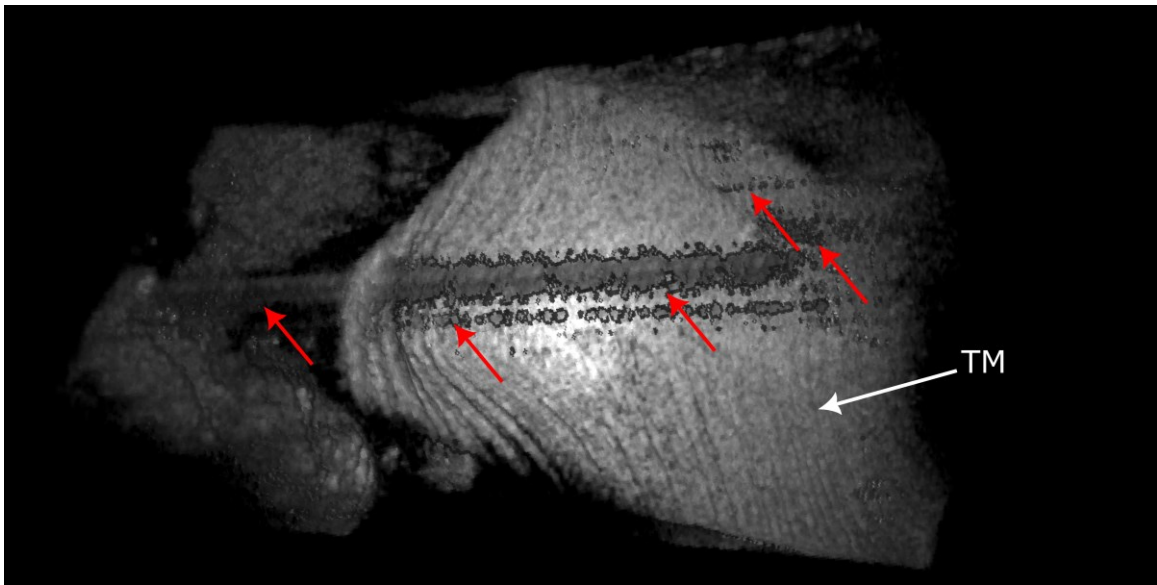
Fig. 29:   Point spread function (PSF) artefact, red arrows, in a ME-OCT volume obscuring the surfaces of the TM and underlying structures.TM, tympanic membrane.

In post-surgical ears where a cartilage graft was implanted (i.e., tympanoplasty) the effectiveness of digital tympanotomy to visualize underlying structures is restricted by the ability of the ME-OCT system to image through these grafts. However, optical clearing agents, such as glycerol, have been demonstrated to reduce two-way signal loss through cartilage grafts by $13.6\,dB$ [24] which may enable digital tympanotomy to visualize underlying structures in these post-surgical ears.

3.5.2   Geometrical Accuracy of Doppler Animations

ME-OCT vibrometry measures the projection of the velocity of each voxel onto the optical imaging axis which differs from velocity magnitude according to:

$$v_{ax} = |v| \cos \theta \qquad\qquad 10$$

Where θ is the angle between the velocity vector direction and the optical axis. While the full velocity vector can be obtained in a multi-beam OCT configuration [168], such a scheme is not practical within the limited aperture available in the ear canal. However, a reasonable estimate of the velocity magnitude of the tympanic membrane can be obtained by recognizing that the tympanic membrane motion normal to the surface is $10 - 20\ dB$ larger than the motion parallel to the surface [169]. If we make the approximation that the velocity of the TM is always perfectly normal to the surface, then the surface normal can be calculated from structural OCT measurements, the angle θ can be calculated and the measured velocity divided by $\cos \theta$ to obtain an estimate of $|v|$.

Accurate calculation of the surface normal direction requires that the optical distortions, as outlined in Section 2.5, be fully compensated so that the volumetric dataset is a faithful, undistorted representation of the true anatomy [111]. While these corrections were not applied at the time that the rendering engine described in this chapter was first developed, they are addressed in Chapter 5 later in this thesis.

3.6   Conclusion

We have demonstrated techniques for real-time rendering of volumetric ME-OCT datasets and two novel data visualization techniques, digital tympanotomy and exaggerated animated cross-sectional and volumetric Doppler vibrometry visualizations. Single-pass ray-casting was shown to produce faithful renders of ME-OCT volumes for a variety of in-vivo applications at rendering rates compatible with real-time manipulation of the datasets. A simple technique for digital tympanotomy showed promise in producing structural images like those obtained through conventional otomicroscopy during exploratory surgery. Finally, the capability of the rendering engine to visualize the acoustic response of the middle ear to single tone stimuli was demonstrated using a ME-OCT Doppler vibrometry dataset acquired in a cadaveric temporal bone.

Developing and implementing a reliable rendering engine capable of real-time volumetric display of ME-OCT datasets is a key component of later work presented in this thesis including real-time volumetric ME-OCT angiography and geometrically accurate 4D ME-OCT imaging.

Chapter 4

Development of a Real-time Software Architecture for Clinical ME-OCT Imaging

4.1 Abstract

This chapter covers the design, verification, and validation of a heterogeneous software architecture that combines CPU multi-threading and GPU acceleration to achieve real-time ME-OCT suitable for clinical imaging at the point-of-care. The developed software architecture is generic and sufficiently flexible to support easy extension in later thesis work to include more advanced ME-OCT imaging modalities including real-time OCTA and geometrically accurate live, continuous, 3D ME-OCT imaging in-vivo.

4.2 Introduction

To deploy a ME-OCT system in a clinic, consideration must be given not only to the imaging performance but also to workflow. Clinicians expect an integrated solution capable of providing real-time guidance and producing real-time visualization of pathologies that is easy to align and makes it easy to acquire and analyze imaging data. The system must be sufficiently easy to use such that a single clinician is able to hold the handpiece in a patient's ear while controlling the software. While the rendering and visualization techniques described in Chapter 3 are an important component of a clinical imaging software solution, the solution must also incorporate highly automated data collection, a UI that can be navigated by the clinician while imaging a patient, capabilities for reviewing and exporting data and features for managing patient data records.

The real-time volumetric rendering engine described in Chapter 3 was originally developed as one of four different standalone software applications needed to acquire and visualize ME-OCT data. As shown in Fig. 30, each of these applications had to be manually synchronized by the user during an imaging session. This proved impossible to do for a single operator as the software required nearly the same level of attention as the patient being imaged. Furthermore, effective use of the software required an intimate knowledge

of its design and implementation. Thus, an imaging session could not be carried out by a clinician without close engineering support. With this system as a starting point, I set out to design an integrated application that could control all hardware and user interactions, including the volumetric rendering engine described in Chapter 3, under one unified, responsive software framework that would be suitable for use by a single clinician operator with minimal training.

Endoscopic camera application

Tone controller application

3D renderer application

Imaging application
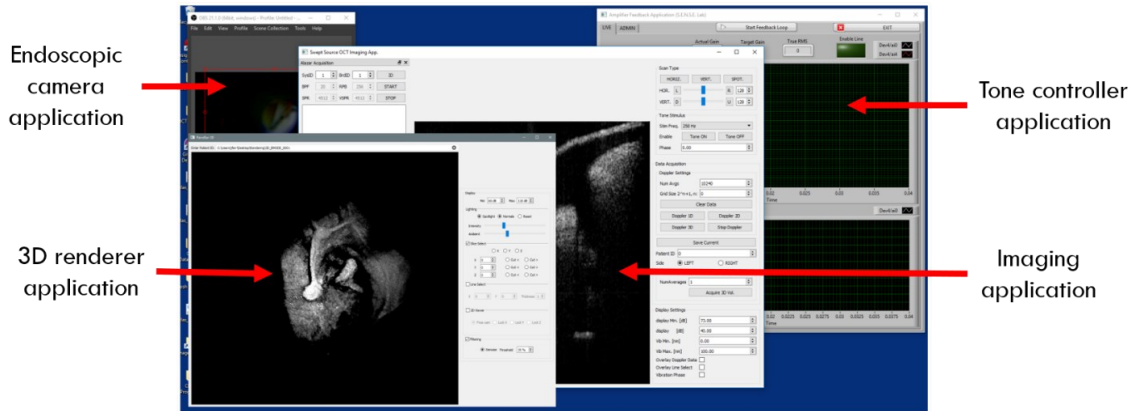
Fig. 30: Original software imaging suite which consisted of multiple stand-alone applications that had to be manually synchronized to facilitate imaging.

4.2.1 Requirements for a Real-time ME-OCT Software Framework for Clinical Point-of-Care Imaging

Based on our experience with the original imaging suite (Fig. 30), I defined five general requirements the new software framework had to fulfill.

1. The software framework must acquire ME-OCT imaging mode data at real-time rates ($\geq 20\ FPS$) and must allow simultaneous imaging from multiple modalities (camera, B-mode OCT, Doppler OCT, 3D OCT) and user interaction without lagging or frame rate reduction.

2. The software framework must be usable by a single operator holding the imaging handpiece in a patient ear while controlling the software. This extends to an outward simplification of implementation complexities to lower the learning curve of the new technology.

3. The software must be sufficiently reliable to not crash during an imaging session and/or be otherwise unavailable when needed. Unstable software erodes confidence in the technology, increasing operator frustration and slowing clinical adoption.

4. The software must integrate well with the ENT clinician's existing workflow. It should feel natural to use, taking advantage of familiar procedures, and guide the clinician through imaging sessions.

5. The software must be easily extendible to support new features and imaging modalities as they are developed. Over recent years, ME-OCT has evolved rapidly [12] and so being able to quickly incorporate new methods and discoveries into the software framework will ensure longevity of the ME-OCT system and improve clinical adoption of the technology. Extension of this software for more advanced imaging modalities and capabilities is the focus of later thesis work and is covered in Chapters 5 and 6.

4.2.2 Trends in Real-time Volumetric Processing Software Frameworks

Modern real-time clinical imaging software frameworks have focused primarily on a GPU oriented approach to reduce imaging latency, increase throughput, and present results to clinicians in real-time [170–172]. This trend has been reflected in the development of OCT imaging systems [1,15,19,173]. While there have been previous efforts to provide GPU accelerated architectures for OCT imaging [174] these efforts were focused primarily on processing performance and not extendibility or usability. As discussed back in Chapter 3, open-source medical imaging toolkits such as 3D slicer [151,152], ITK [153], and VTK [154] had minimal native support for GPU acceleration at the time of this work, making them ill-suited for real-time imaging applications at the point-of-care [157,158].

Huang et al. [175] recently demonstrated a software framework that works for both SS-OCT and SD-OCT. Their framework successfully abstracted away hardware differences between SS-OCT and SD-OCT processing through a clever hardware implementation and judicious use of software interface polymorphism and design principles. However, performance of their software framework was limited due to the decision to use CPU acceleration through OpenMP [176] and not to support GPU acceleration. The extendibility to more complex functional OCT imaging modalities such as Doppler vibrometry or OCTA was not demonstrated, and the framework's extendibility was limited due to the low-level multi-threading strategy used.

4.2.3　ME-OCT Imaging Pipelines

The top of Fig. 31 depicts a software abstraction known as a pipeline which is a chain of processing steps used to generate and display ME-OCT imaging data to users. Pipelines are composed of three main parts: data sources, data transforms, and data sinks. Data sources represent inputs to the pipeline such as A-lines, endoscopic video, or audio waveforms. Data transforms represent the sequence of processing steps applied to the given input data from the data source(s) to produce the desired output data such as the processing steps described earlier in Section 2.8. Each data transform in the pipeline is dependent only on the data at its input and is independent of the exact processing details in other pipeline transforms. The black-box nature of the transforms can be leveraged to facilitate concurrent processing of multiple pipelines, representing multiple data streams, with the transforms' input dependencies used to combine and synchronize the various data streams. Finally, data sinks represent outputs from the pipeline and are used to display the output data produced by transforms to the user.

Fig. 31 summarizes the various pipelines for ME-OCT imaging that transform raw OCT interferograms into structural 2D and 3D B-mode images, calculate Doppler vibrometry measurements and process the endoscopic camera imaging stream. As can be seen in Fig. 31, the various ME-OCT imaging modalities share a common A-line data source where the underlying processing details of the A-lines do not impact downstream processing. A well-designed software architecture will abstract away the imaging hardware details such as the OCT topology used to generate these A-lines [175] or even the imaging modality altogether. This abstraction makes the architecture reusable across various imaging systems and technologies and allows for interfacing to a simulated hardware environment for testing purposes. Since each pipeline shares common data sources as input, execution of each pipeline must be coordinated to avoid possible hardware resource contention. This can occur in situations where different pipelines use a common data source but require a different configuration of the underlying data acquisition hardware parameters. For example, Doppler vibrometry imaging may need A-lines from a single location of the sample being imaged whereas 3D requires A-lines from locations covering the full FOV across the sample.

A well-designed software architecture must therefore provide a means to encapsulate and select the desired underlying imaging hardware, multi-threaded support for concurrent real-time pipeline execution, and a generic means of coordinating system resource utilization while providing a simplified, intuitive interface to clinicians such that the desired information for a patient's report can be efficiently generated. Additionally, it must be extendable to support new imaging modalities as they are developed without impacting existing imaging functionality.
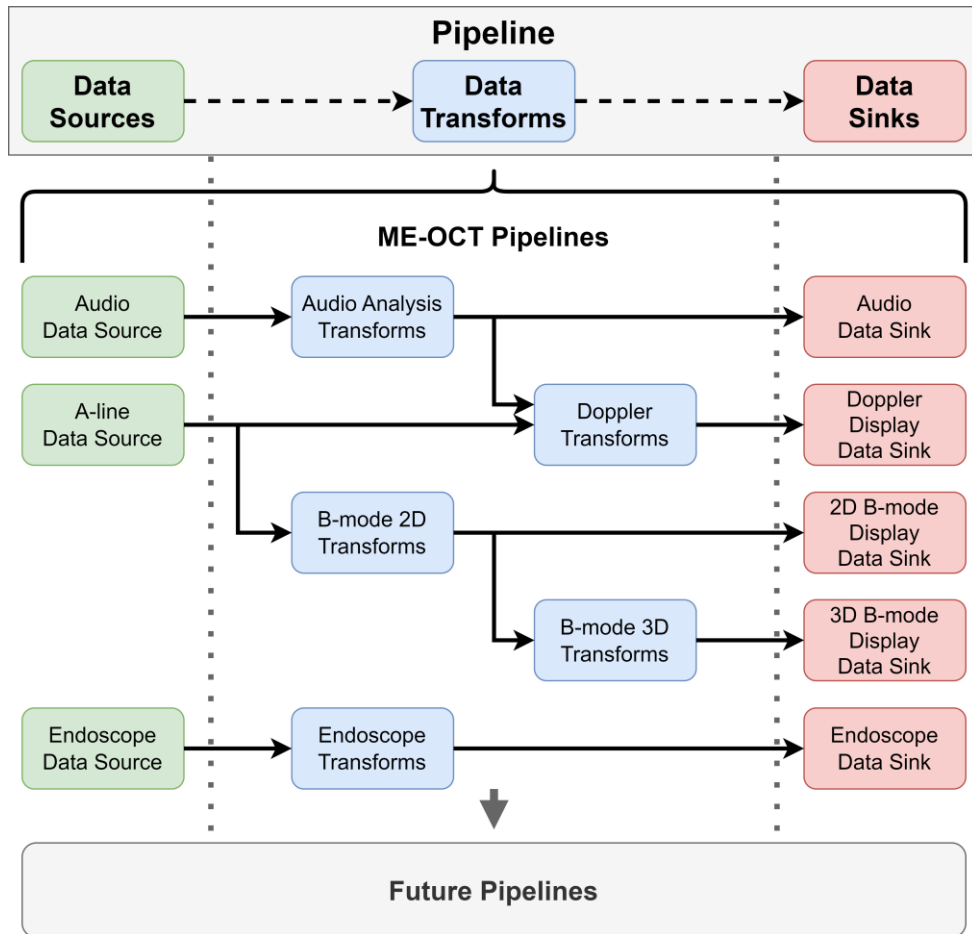
Fig. 31: *Processing pipelines responsible for transforming OCT interferograms into structural 2D and 3D B-mode images, and Doppler vibrometry measurements along with endoscopic imaging. Pipelines consist of three parts: data sources, data transforms, and data sinks. Data sources represent inputs to the pipeline such as A-lines, endoscopic video, or audio waveforms. Data transforms represent the sequence of processing steps applied to the given input data from the data source(s) to produce the desired output data. Data sinks represent outputs to the pipeline and are used to display the output data produced by transforms to the user. Additional pipelines can be added to support new imaging modalities as they are added.*

4.3 Methods

As shown in Fig. 32, the software architecture is comprised of three libraries: the UI library referred to as the *frontend*, the real-time image data generating library referred to as the

backend, and the *data access library* comprising the patient database and data storage aspects of the software. The frontend is responsible for user interaction, managing workflow, collecting patient data, and displaying imaging data generated by the backend. The backend generates imaging data by controlling system hardware, data processing, and data exporting capabilities. The patient database and data storage components are responsible for storing collected patient data and generated imaging data, respectively. Data can be stored either on local disk or remotely depending on the type of data access library used and depends on the clinical environment in which the software is run. For example, the data access library could interface with a hospital's picture archiving and communication system (PACS) to store and enable review of imaging data collected by the system stored as digital imaging and communications in medicine (DICOM) datasets. However, this is outside the scope of this thesis work and will not be discussed further. Both the frontend and backend libraries can be run independently of each other for greater software flexibility and ease of testing. For example, the frontend library can be run without the backend if the clinician just wishes to view previously collected data. Likewise, the backend can be run without the UI for automated testing purposes.
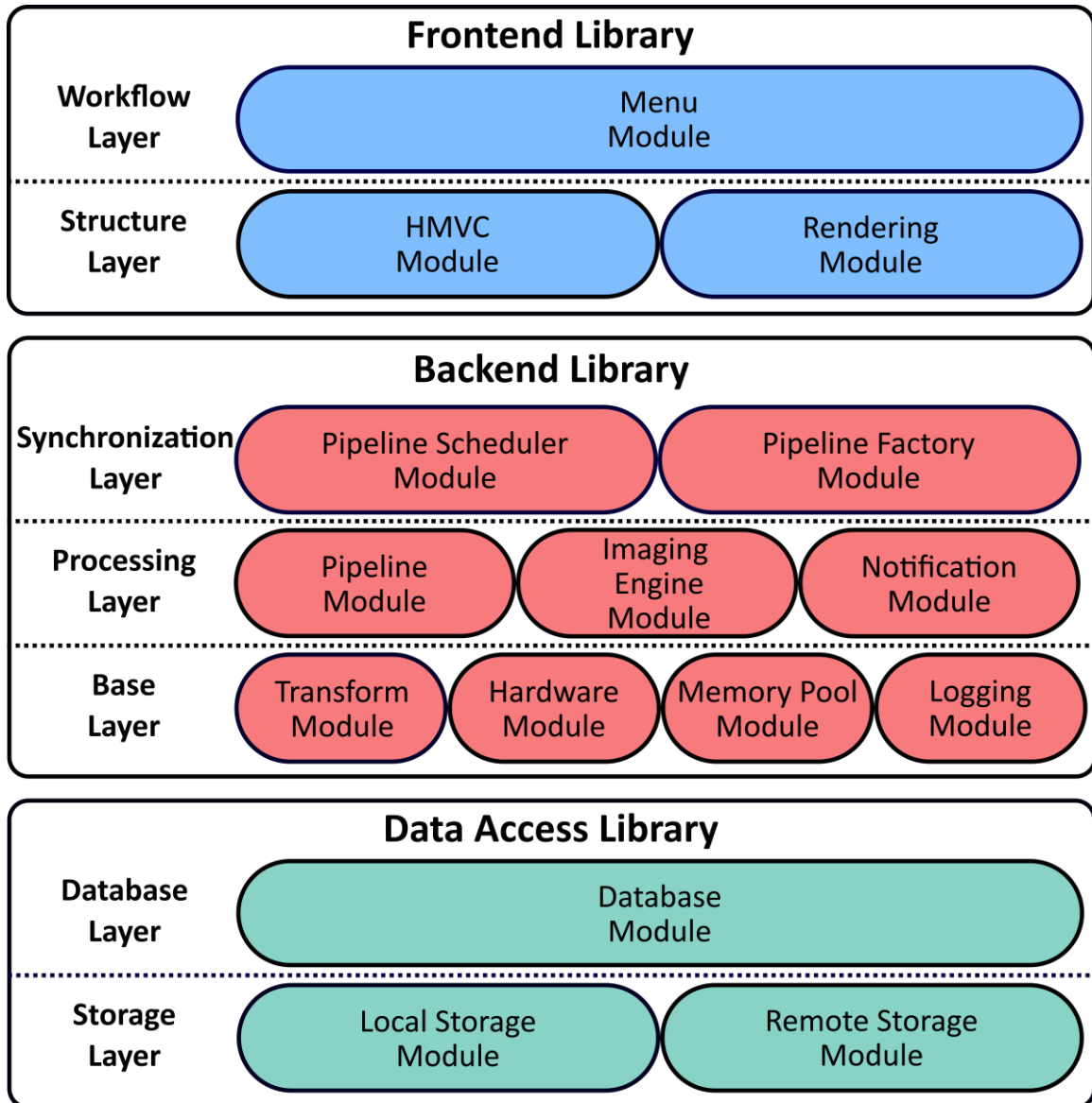
Fig. 32: Software architecture overview. The architecture is comprised of three libraries: the UI library referred to as the frontend, the real-time image data generating library referred to as the backend, and the data access library comprising the patient database and data storage aspects of the software.

4.3.1   Frontend

To encapsulate and define the clinical workflow we created a software abstraction called a *report* that is broken up into a set of *report components*. This report structure is illustrated in Fig. 33. A *report component* is a set of diagnostic measurements that would be used by a clinician to help distinguish specific middle ear pathologies. Each report component is associated with a set of UI steps required to step the clinician through the collection of the diagnostic information needed for the report component. Each report component is also

associated with a set of outputs such as grouping of images and measurements viewable within the UI and a review screen to convey diagnostic information to the clinician.
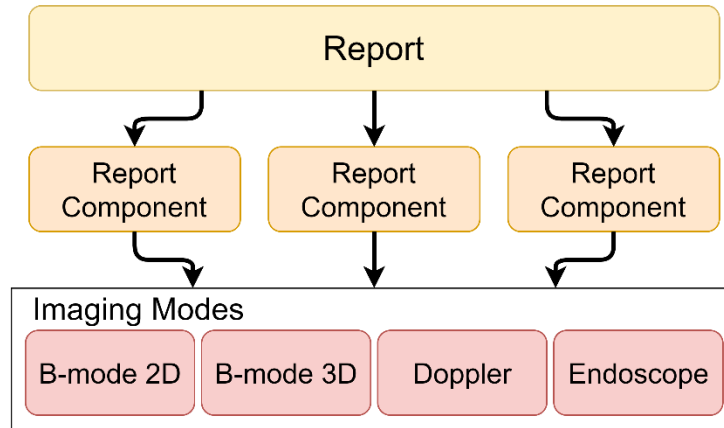
Fig. 33:   Reports encapsulate desired clinical workflow and are broken up into one or more report components. Each report component is associated with an imaging modality and a set of instructions to help guide users through the imaging session.

Facilitating this clinical workflow model is the frontend library, topmost of Fig. 32, which is comprised of two layers: the *workflow* layer, and the *structure* layer. The frontend serves as the main UI to control, store, and display imaging data generated by the backend library. The frontend is responsible for loading in user configuration settings from a data access library, instantiating and configuring the backend library, and maintaining a connection to the patient database.

Structure Layer

The structural layer contains the core framework and components required to construct the frontend UI. The structure layer contains two modules: the hierarchical model view controller (HMVC) module and the rendering module.

The HMVC module makes use of a version of the model view controller (MVC) framework originally proposed by Krasner et al. [177]. The MVC framework is an architectural design pattern that dictates that UI logic should be broken down into three components: model, view, and controller. The model encapsulates the UI's core data structures and defines the functional rules/behavior of the UI (often referred to as the UI's 'business logic'). The view presents data to the end user and handles user interaction. Finally, the controller coordinates the flow of data between the model and view.
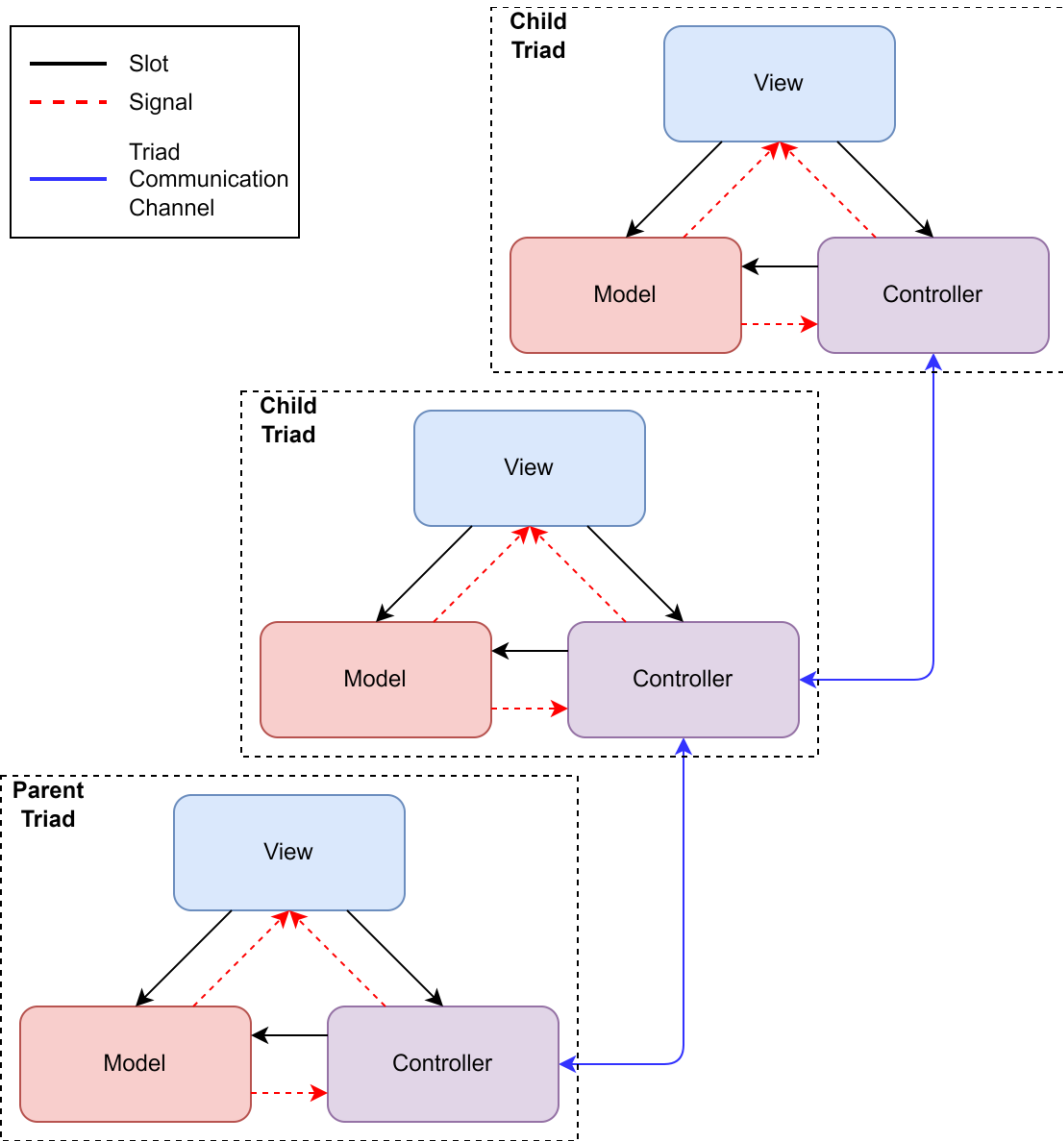
Fig. 34: The Hierarchical Model View Controller (HMVC) is an extension of the classic MVC pattern that allows UIs to be constructed form a hierarchy of MVC groupings called triads. Triads communicate internally using Qt's signal and slots framework [178] whereas triads communicate through a dedicated parent-child communication channel.

As shown in Fig. 34, the HMVC is an extension to the classic MVC pattern that allows one to construct the UI from a hierarchy of MVC groupings [179]. Here a single MVC group is called a triad (referring to the three components of the MVC). The UI can be dynamically presented as a hierarchy leading to a more intuitive way to implement clinical workflows. Like the classic MVC pattern the view is responsible for displaying data to the end user, and handling user interaction. The model is unchanged from the classic MVC pattern. The controller is responsible for maintaining the triad hierarchy through a parent-child relationship. This relationship acts as a communication channel between

triads, where data in a parent triad can propagate to all the child triads in the system. Through this channel a child triad can also request information, send information, and trigger the creation of new triads in other parts of the hierarchy.

The HMVC module was implemented using Qt [178], an open-source API used to construct graphical user interfaces (GUIs) through the composition of visual UI elements called widgets. Widgets handle both the visual display of data and user interaction, and can also act as containers, or windows, for other widgets to be grouped together to form more complex UIs. Widgets can communicate with other widgets asynchronously through call-back functions provided by Qt's built in signal and slots framework. Here *signals* refer to the call-back functions exposed by a widget's interface that other widgets can listen for using signal handlers called *slots*. Signals and slots are used for communication between the elements of a triad whereas the HMVC parent-child communication channel is reserved for communication between triads only.
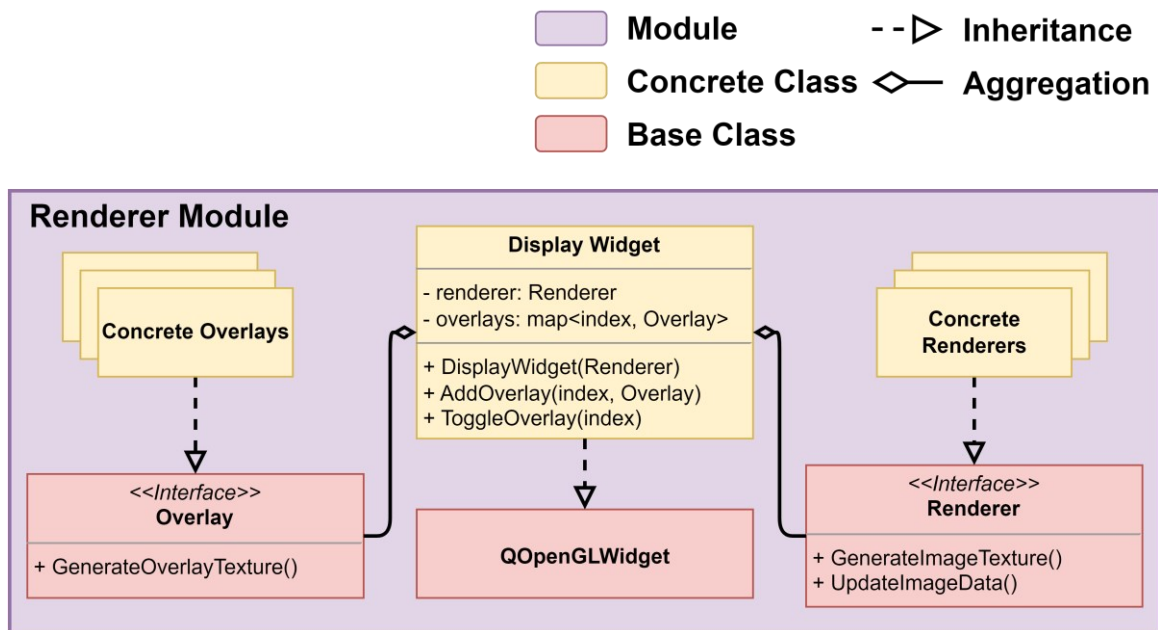
Fig. 35: The rendering module utilized by the frontend library to provide an integrated rendering target for the developed rendering engine to display imaging data generated through the backend library.

The structure of the rendering module is shown in Fig. 35. The rendering module can be used anywhere in the frontend that imaging data generated by the backend needs to be displayed to the user. The rendering module supports real-time display of 2D, 3D and endoscopic image data as well as different graphical overlays present for each report component. To provide this flexibility, the rendering module consists of a display widget

75

acting as a rendering target, and two distinct groups of objects: renderers and overlays. Renderers are designed to convert image data generated within the backend into displayable images. For 3D datasets, the volumetric ray-casting-based rendering engine described in Section 3.3.1 is used for direct real-time visualization [160]. Overlays are designed to display imaging related metadata to the user on top of the renderer. Examples of metadata displayed using an overlay include image timestamps, quality readouts, current ear being imaged, etc. The display widget is built on top of Qt's widget library and designed to be integrated within the UI to provide the means to display the generated images. For compatibility with the other widgets, the display widget is a concrete implementation of Qt's QOpenGLWidget [178] allowing for seamless integration with the rest of Qt's signal and slot framework. Within each display widget is a reference to a single renderer object and potentially multiple overlay objects. Overlay objects can be layered on top of each other to facilitate a modular approach to metadata display. This allows overlays to be constructed from simple layers combined appropriately. Like the rendering engine (Section 3.3.1), the overlay objects are implemented using OpenGL which allows overlay images to be generated and displayed in-place on the GPU at real-time rates without transferring backend imaging data from the GPU to the CPU first. This prevents display lag incurred from the GPU-CPU memory copy and simplifies implementation by offloading the interpolation step required to display the generated images to the GPU's built-in interpolation hardware. The renderers and overlays used within the display widget vary depending on the currently active report component within the imaging menu node.

Workflow Layer

As shown in Fig. 36, the frontend utilizes the HMVC module to organize the UI into a hierarchy of triads stemming from a singular parent triad called the *application*. The application contains the main widget window of the UI for which all other UI widgets are placed and maintains a connection to the patient database, backend, system configuration file, and system log. Child triads can interact with these contained elements in the application triad through the parent-child triad communication channel provided by the HMVC module. This pre-existing, data-driven communication channel improves the ease with which new functional UI elements can be added or extended.
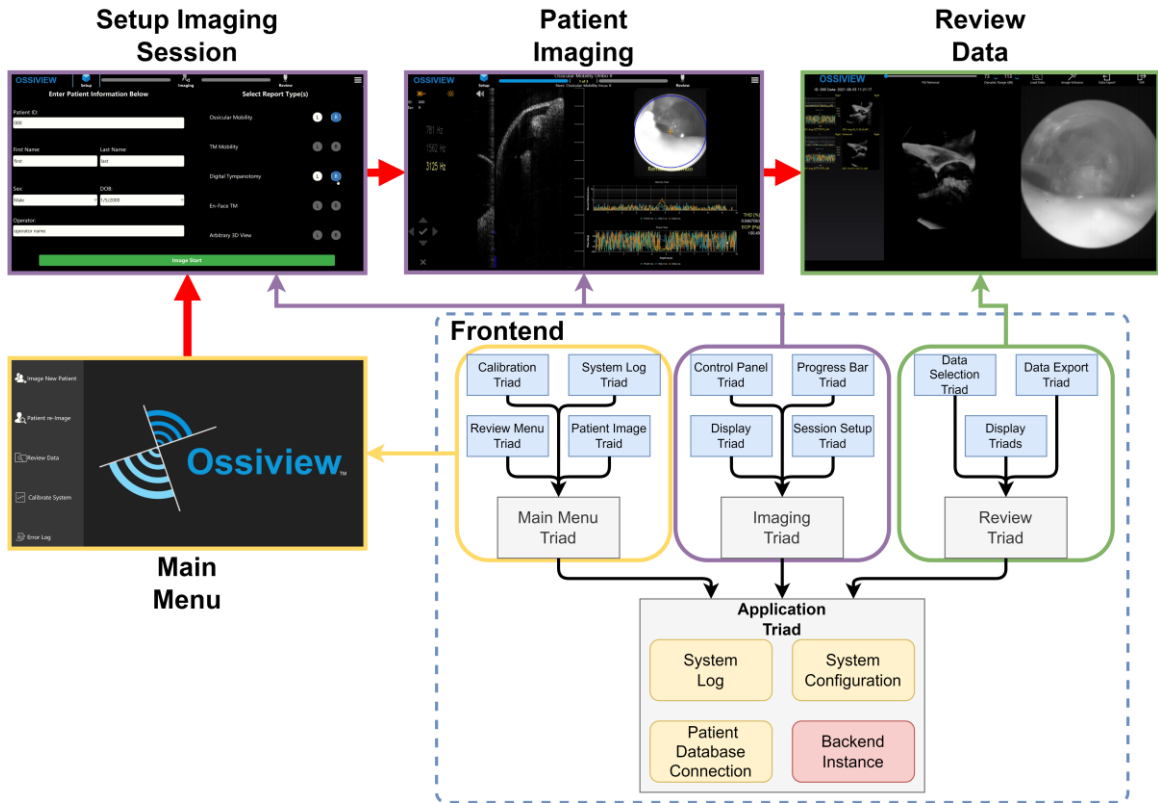
Fig. 36: Frontend library UI triad hierarchy.

The application triad has three child triads: the main menu triad, imaging triad, and review triad. Each of these triads represents a different set of menus users interact with and are guided through. For example, after the UI finishes initializing, the clinician is brought to the main menu where they can review previously collected data, view system logs, calibrate the swept-source laser or initiate an imaging session by entering a new patient or searching for an existing entry. Following patient selection, the frontend switches over to imaging where the clinician is navigated through the imaging session by the UI. After the imaging session is complete the UI switches over to the review screen where all the collected imaging information is displayed allowing for quick export of imaging session results.

Fig. 37 shows the imaging triad. The imaging triad coordinates the data collection efforts of an imaging session and serves as the main imaging window in which all other UI widgets are contained. Attached to the imaging triad are four child triads including the control panel, display, progress bar and session setup that handle user input, display imaging mode data, inform the user of the completion status of the imaging session and collect patient information for the imaging session respectively.

The Imaging triad's controller contains dedicated communication handlers for each child triad to facilitate data requests to/from the application triad through the triad communication channel and to ensure all the child triads are synchronized according to the current imaging session. In addition to serving as the main window widget for the imaging session, the imaging triad's view handles the options menu, logo, indicators, and patient info widgets that persist between imaging screens. The child triads each have a one-to-one visual representation where the control panel triad provides a control panel widget located in the bottom right corner of the imaging screen, the progress bar triad provides the progress bar widget at the top, the session setup triad displays a setup widget before imaging starts, and the display triad which acts as a render target for backend imaging data. The display triad itself is comprised of three display widgets where image data is displayed to the user including the OCT image on the left, enface endoscopic video on the top right, and a preview of the collected data on the lower right.
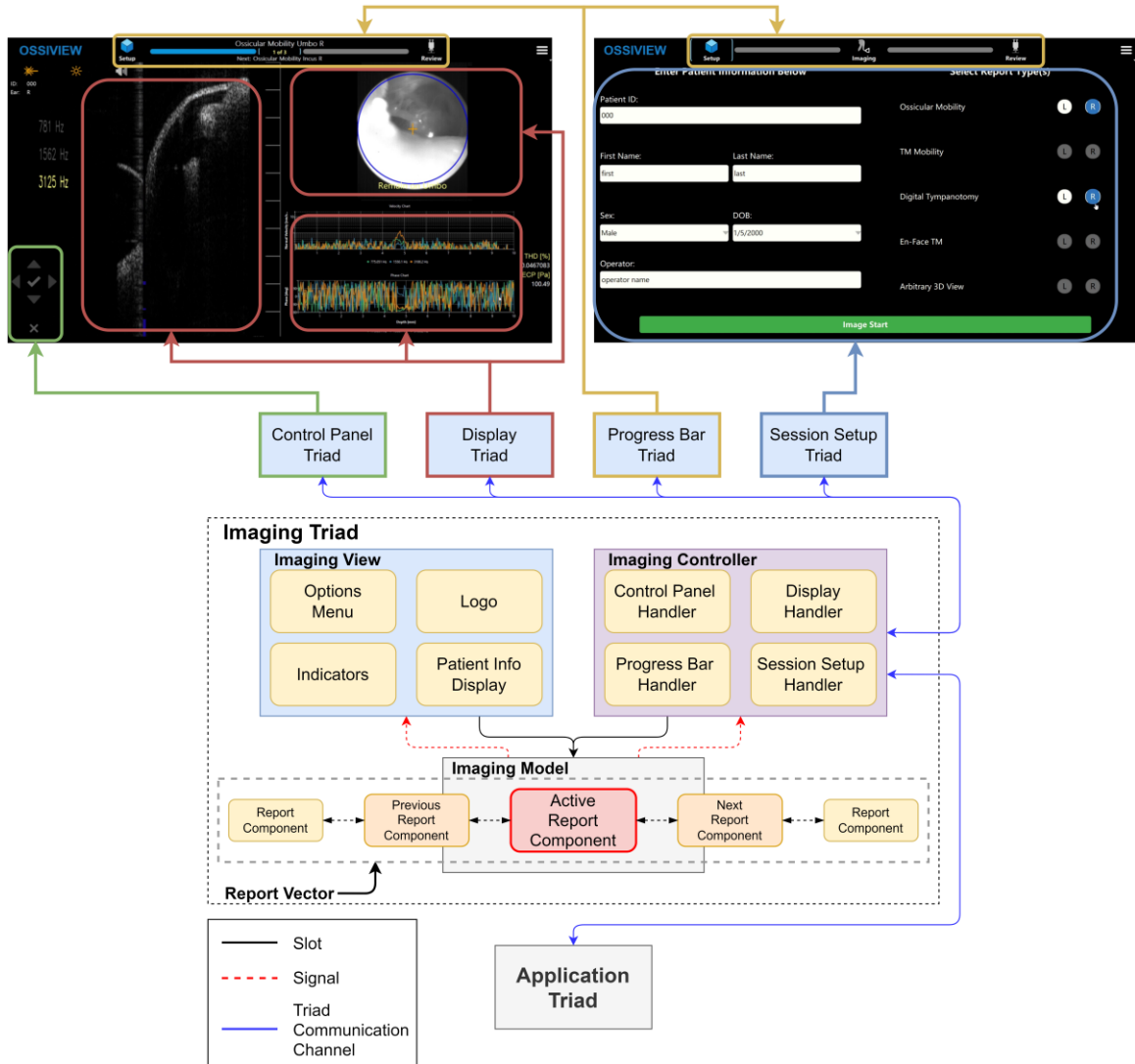
Fig. 37: Overview of the imaging triad hierarchy supporting frontend clinical imaging.

To coordinate the data collection efforts of the imaging session, the imaging triad's model is implemented as a state machine where the states are the report component objects in Fig. 33. By switching the active report component object, the behavior of the imaging menu changes as the user is guided through the workflow defined within each report component. The behavior of the child triads in turn are driven by the active report component and help facilitate the desired workflow. For example, the display triad may be configured to display extra data overlays appropriate to the data being collected for different imaging modalities like Doppler (Fig. 38a) or 3D (Fig. 38b). Contained within the imaging triad's model is a vector of report components, which collectively represent the final set of results the clinician wishes to see. The progress bar keeps track of the overall

progress of the report being generated while the control panel provides a mechanism to quickly navigate through the report components, start/stop imaging and save or reject imaging results as they are collected.
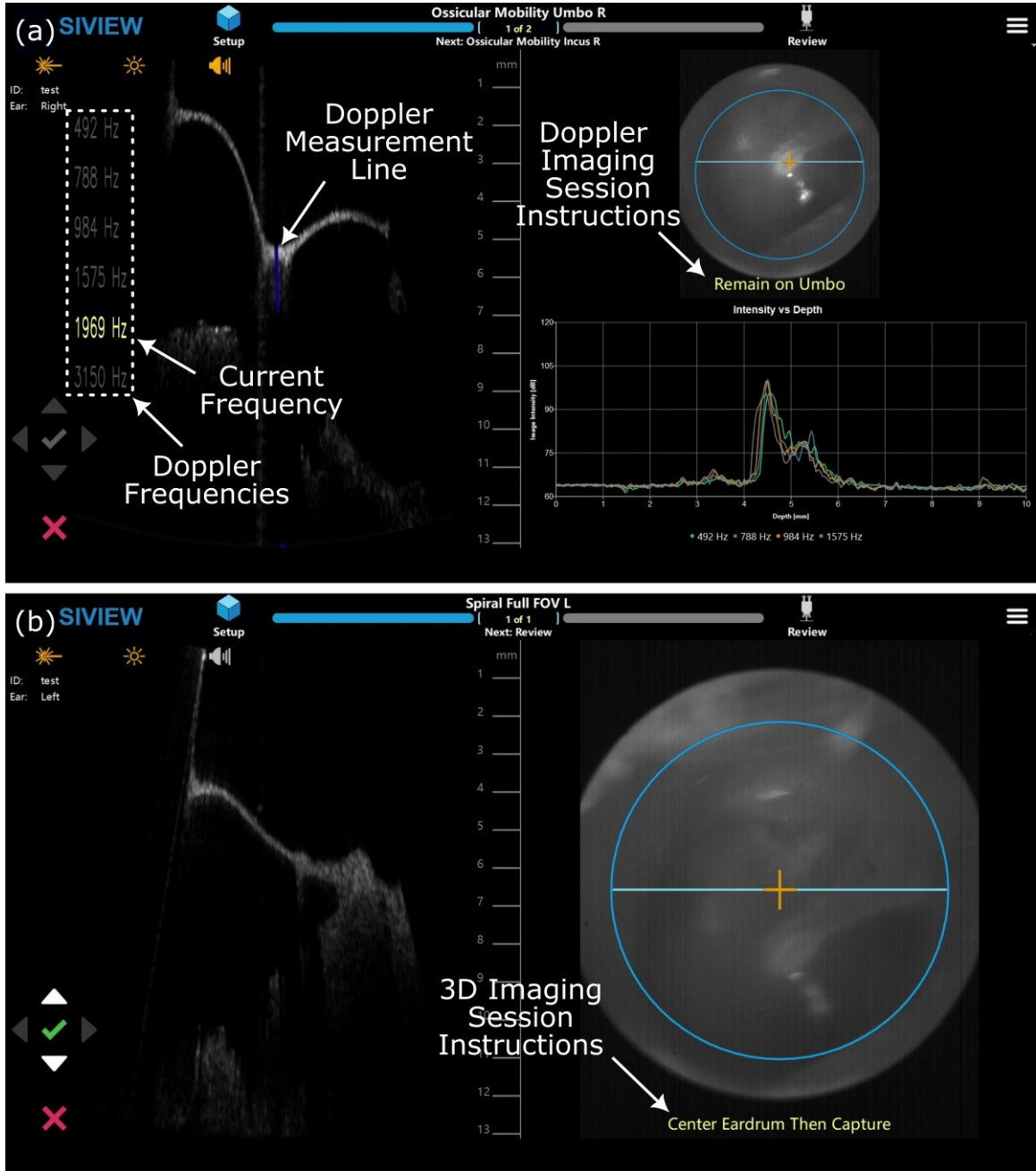
Fig. 38: Report component driven workflow for (a) Doppler imaging session and (b) 3D imaging session.

4.3.2 Backend Library

To support simultaneous execution of multiple imaging pipelines, the backend library (Fig. 32) must provide multithreaded functionality, coordination of concurrent work, and

automatic resource management including system memory and encapsulated imaging hardware. For this purpose, we designed the backend with three layers: the *synchronization layer*, the *processing layer*, and the *base layer* which are broken down into one, or more, software modules. A detailed unified modelling language (UML) diagram of the backend library can be found below in Fig. 39. The following sections will explore each part of this UML diagram in detail and how each module interacts to achieve our goals.
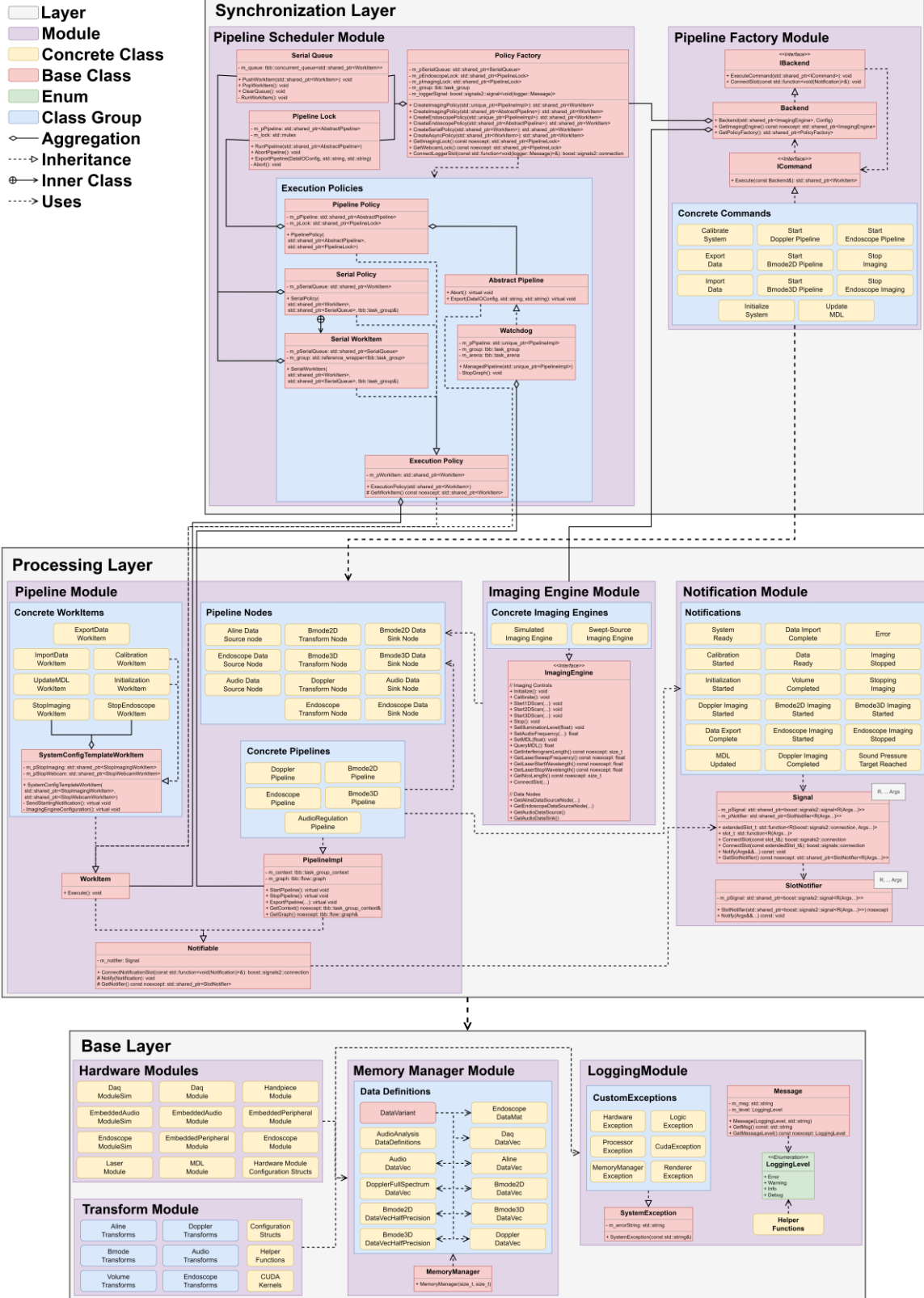
Fig. 39: *Detailed Unified Modelling Language (UML) diagram of the backend library.*

Base Layer

As shown in Fig. 40, the base layer of the backend library contains the hardware module, memory manager module, transform module and logging module. These modules provide the core functionality required by the rest of the backend library and form the foundation on which all other layers are built upon.
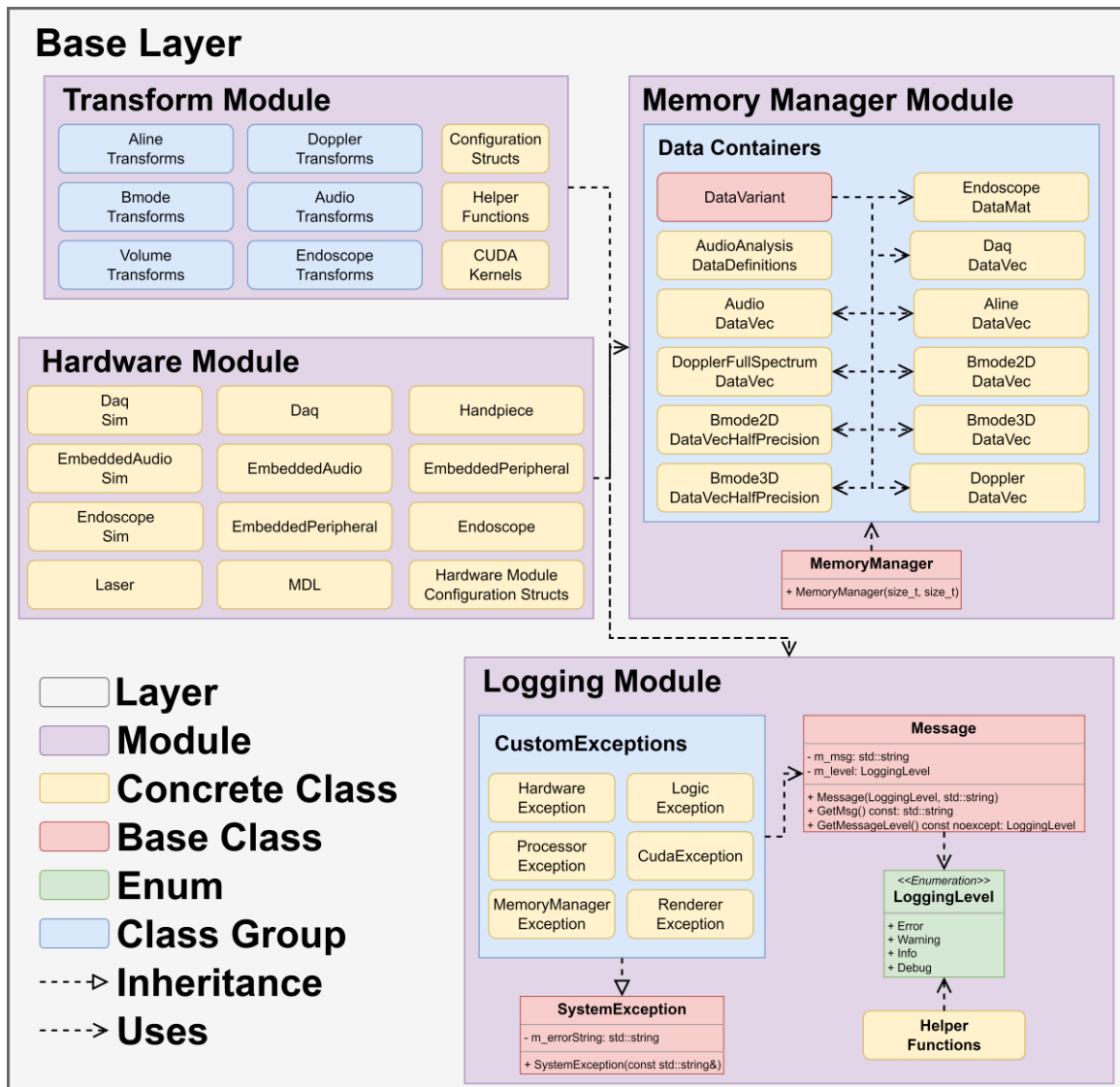
Fig. 40: Detailed Unified Modelling Language (UML) diagram of the Base layer of the backend library.

All hardware functionality of the ME-OCT imaging system is contained within the hardware module. The hardware module is a collection of hardware abstraction layers (HAL) for both SS-OCT and simulated hardware components. HALs serve as a logical division between software and physical hardware. Exposing simplified control interfaces

of that hardware to the rest of the software architecture. A singular piece of hardware could have one or more HALs depending on the type of interface one wishes to expose. For example, the embedded controller of the ME-OCT system has two HALs, one for controlling the system's audio and the other for controlling the scanning optics. Simulated hardware components through HALs allow simplified testing of the architecture without requiring access to the physical ME-OCT imaging system.

The logging module contains all the software definitions required for data logging and error handling throughout the software architecture. Errors, called exceptions, are thrown when erroneous situations occur within software objects and propagate outwards until caught and handled whether inside the object that threw it or in the objects that use it. If not handled, the program will exit abruptly without notice to users besides a generic Windows error message. To handle exceptions appropriately, a descriptive type and message is provided by the logging module for each error type. For example, if a hardware error was encountered within a HAL, a hardware exception object would be instantiated and thrown. The exception would contain a description of the hardware error and would propagate up through the call stack until reaching any error handling code designed to handle hardware exceptions. Besides errors, informative messages can also be sent out at any point to help notify developers on the current status of the software components without interrupting program flow. The logging module provides the message definition required for notifications used throughout the architecture.

The transform module contains all the data transforms used within the various pipelines of the backend. Each data transform has an input and an output data connection. As the implementation details of each transform object are hidden from the outside, they can either be implemented using CPU or GPU algorithms that best suit the type of processing being done. All OCT imaging modalities were implemented using CUDA. As discussed in Section 2.7.1, CUDA uses functions called 'kernels' to execute millions of processing threads simultaneously on the GPU. This paradigm works best in situations where a small number of repetitive, independent calculations need to be done on large quantities of data. In such cases, GPU-based processing can be multiple orders of magnitude faster than CPU-based sequential processing. For example, GPU-based processing is adept at processing raw OCT interferograms into A-lines since the processing

steps are the same across all interferograms. All endoscopic image processing was done using the Open-Source Computer Vision Library (OpenCV), a highly optimized real-time computer vision library that runs on the CPU [180]. The flexibility to use different open-source libraries facilitates rapid incorporation of new processing algorithms and techniques which improves the stability, maintainability, and extendibility of the software.

The memory manager module is comprised of a collection of data containers and the memory manager. Data containers define all the types of image data passed throughout the architecture and serve to provide a common description and uniform access to that data no matter where it is accessed within the architecture. This increases the modularity and maintainability of the architecture as image data definitions are self contained and consistent. Thus, modules do not need to depend on the implementation details of other modules to determine the structure of given image data. All image data passed into and out of transform and hardware modules use these data containers.

The memory manager makes use of the open-source RAPIDS Memory Manager (RMM) [181] library to control memory resource utilization, reduce memory fragmentation, and avoid dynamic memory allocation to facilitate real-time data rates. RMM improves pipeline throughput by allowing transform objects to dynamically allocate and propagate new data containers to other transforms without worrying about dynamic memory allocation overhead or memory fragmentation. As all memory in the pool is allocated during instantiation of the backend library, there is no memory being dynamically allocated. Instead, there is only a small overhead related to fetching the block of memory from the pool. For the case of OCT imaging data where many interferograms are typically processed in large chunks on the GPU (several megabytes per frame) this overhead can be safely ignored.

Processing Layer

Shown in Fig. 41, the processing layer of the backend is comprised of the pipeline module, imaging engine module, and notification module. Utilizing the functionality provided by the base layer, these modules provide all the software required to construct the imaging pipelines supported by the ME-OCT system and the notification system used by the backend to help communicate data generation efforts to the frontend.
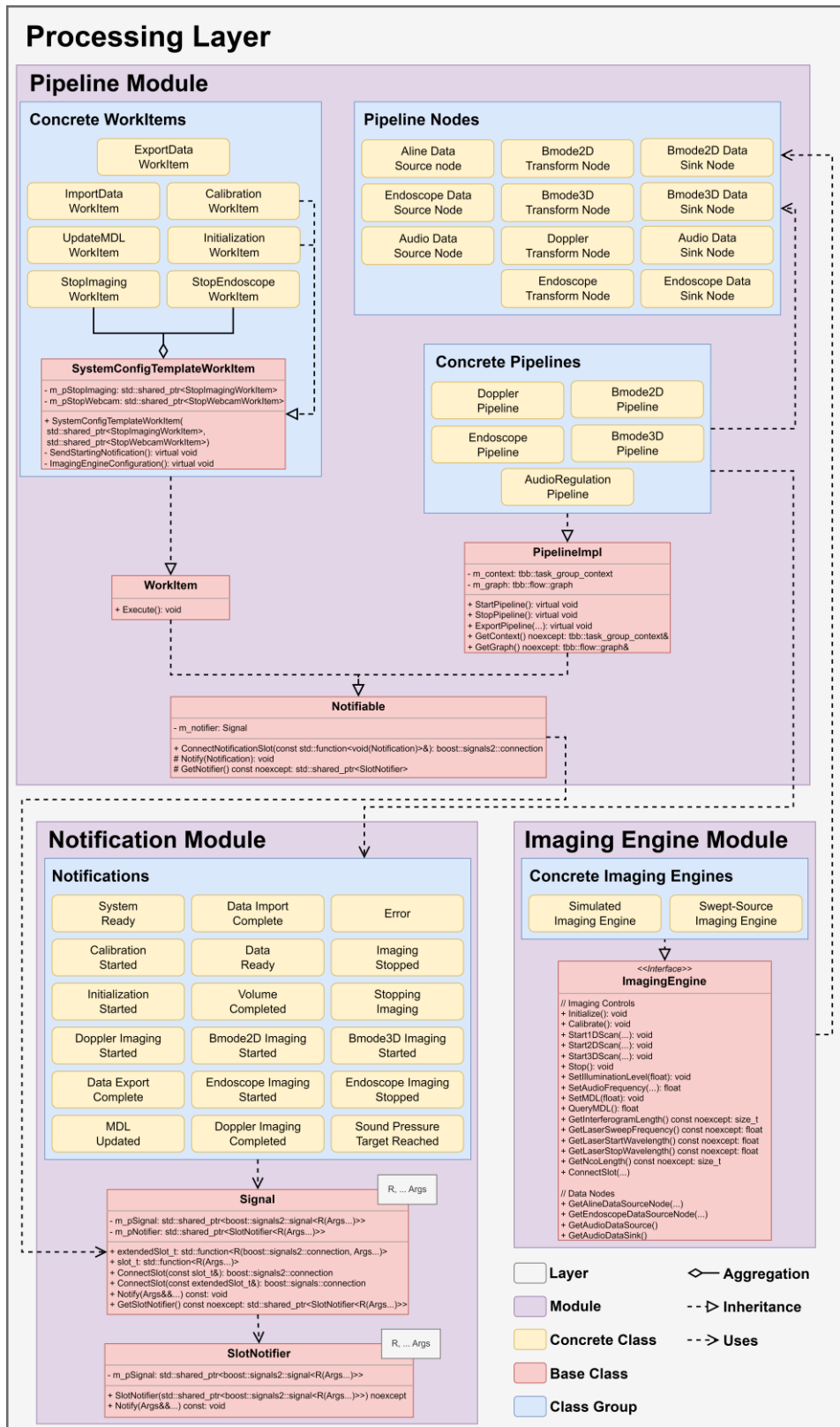
Fig. 41: *Detailed Unified Modelling Language (UML) diagram of the Processing layer of the backend library.*

The imaging engine module incorporates all the HALs, within the hardware module of the base layer, for OCT and endoscopic imaging into a standardized control and data acquisition interface for use by the imaging pipelines. As discussed by Huang et al. [175], regardless of the specific type of OCT system used they all share similar hardware control layouts and only have minor differences to how acquired interferograms are processed into A-lines after which all processing is the same. By defining the appropriate interface, the imaging engine could use different OCT technologies without changing downstream processing. Thus, from the perspective of software, the same imaging pipelines can be used regardless of the specifics of the OCT system used reducing development time and improving extendibility and stability of the software. Fig. 42 shows the layout of the imaging engine module for the case of our ME-OCT system [166], along with the imaging engine's interface. The interface itself is comprised of two main sections: data nodes and imaging controls. Data nodes serve as generic data sources or sinks for pipelines with their implementation again dependent on the underlying hardware of the imaging engine. For example, the A-line source node may process raw interferograms differently depending on whether a MEMS tunable vertical cavity surface-emitting laser [182] (VCSEL) was used instead of the VT-DBR akinetic swept laser as we use currently. Regardless of the underlying hardware however, the output of the source nodes remains the same. The imaging controls interface directly with the hardware which in turn directly controls all aspects of OCT and endoscopic imaging. The imaging hardware itself is semi-autonomous in that it ensures proper synchronization between components via triggers and clocks cycles such that raw data is acquired correctly. However, through the imaging engine interface, the software directs the hardware when to start imaging, which scanning waveforms to use and what audio tones to output during scanning. This also includes any other hardware peripherals that may be available such as illumination for endoscopic imaging or LED indicators to signal to users when certain aspects of the hardware are in use such as tone generation or the laser.

Currently the backend supports imaging engines for both the SS-OCT system and simulated hardware, Fig. 41. Simulated hardware allows imaging pipelines to be tested without requiring access to the actual imaging hardware.
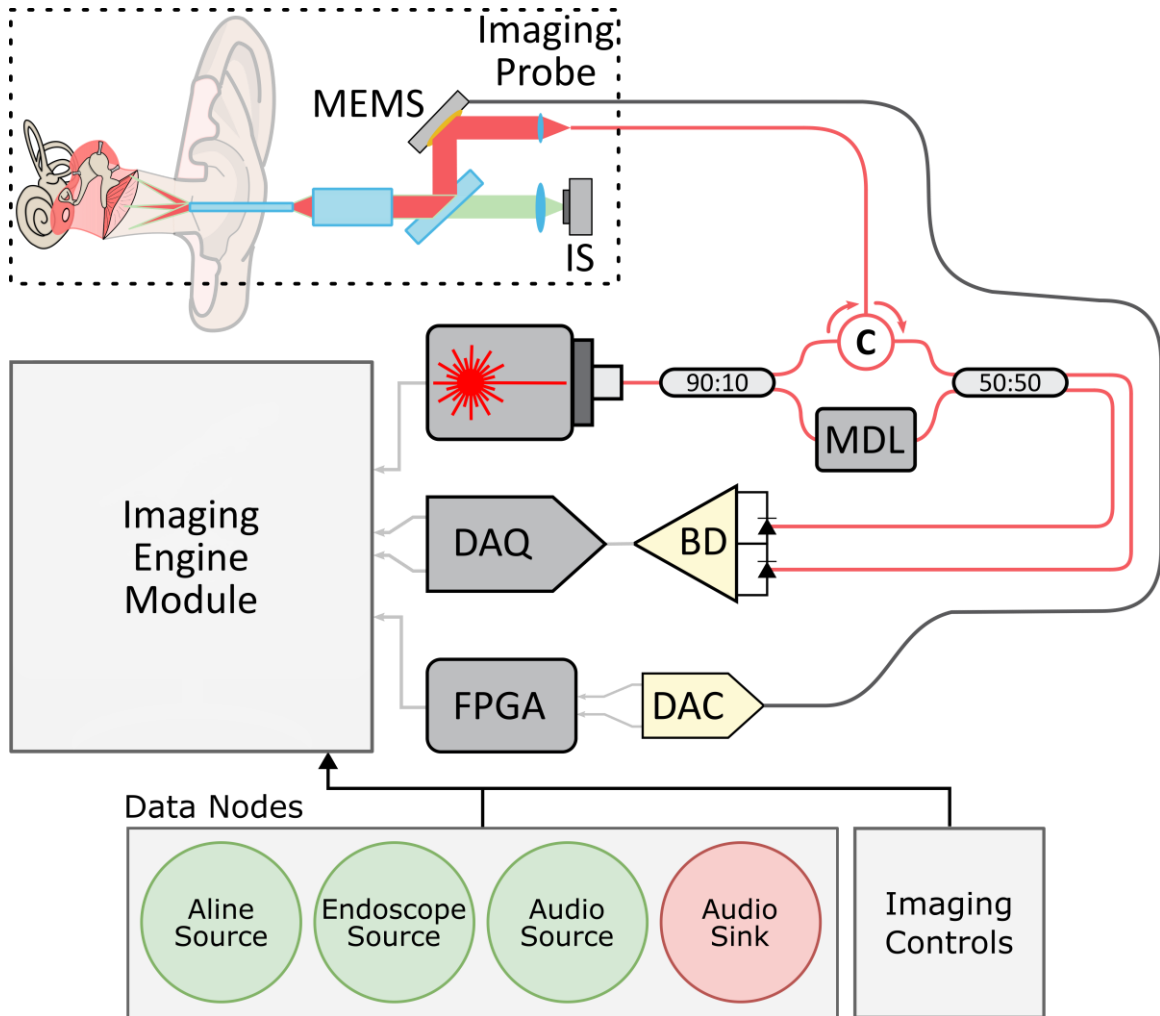
Fig. 42: The Imaging Engine Module encapsulates all the hardware required for OCT and endoscopic imaging and provides a standardized interface for hardware control and data acquisition to the rest of the software system. The imaging engine interface is comprised of two main sections: data nodes and imaging controls. Data nodes serve as generic data generating inputs for downstream imaging pipelines with their implementation again dependent on the underlying hardware of the imaging engine. The imaging controls interface directly with the hardware which in turn directly controls all aspects of OCT and endoscopic imaging.

The pipeline module, Fig. 41, comprises all the imaging pipelines supported by the ME-OCT system as well as other work items required for system initialization, calibration, data import/export, and starting/stopping processing. Every pipeline is implemented using Intel's thread building blocks (TBB) library. Intel's TBB is an API providing high-level thread abstractions and data structures that automate low-level thread management [183]. Shown in Fig. 43a, each data transform within an image pipeline is mapped directly to a node representing a single unit of work or 'task'. Tasks are executed from left to right according to the data dependency links between them. i.e., tasks will not execute until given data from upstream tasks. At the start of each pipeline is one, or more, data source nodes

provided through the interface of the imaging engine module that streams in data from underlying imaging hardware. Data source nodes can be combined with data transform nodes to form one large composite node as shown for the A-line data source node in Fig. 43a. Each node within the composite node will still be treated as a single task, however, it makes wiring up the pipeline easier as one only needs to deal with the interface of the single composite node rather than the multiple sub-nodes. At the end of each image pipeline is one or more data sink nodes that are used to stream generated image data to disk and/or display to the frontend via a display widget in the UI. The imaging engine also has an audio sink node to monitor and regulate stimulus pressure during Doppler imaging.

Each task within the pipeline is automatically mapped to available CPU hardware threads by TBB's built in task scheduler to balance workload across the CPU hardware and maximize processing performance. For example, GPU kernel execution is asynchronous with respect to CPU execution, thus, in cases where a single CPU thread is used to manage kernel execution CPU processing is blocked until the kernel execution is completed. Using tasks, the CPU thread can manage several GPU kernels concurrently, allowing kernel execution to be interleaved. This prevents blocking of CPU threads and makes better use of available processing bandwidth. The thread scheduler is configured to not map any task to the main UI thread to avoid interfering with the frontend and prevent perceived slowdown or lag. Decoupling of low-level thread management allows developers to quickly add new transforms and pipelines, without having to worry about low-level multi-threading related errors, greatly improving system extendibility and stability. Furthermore, as new CPU hardware is introduced, pipelines will automatically scale to use the new threading resources, increasing system robustness and future-proofing performance. In contrast, if threads were manually managed (Fig. 43b) the pipeline would have to be rewritten to make full use of any additional threading resources. Or if the number of threading resources were reduced, the pipeline would fail to execute without changes to the software.
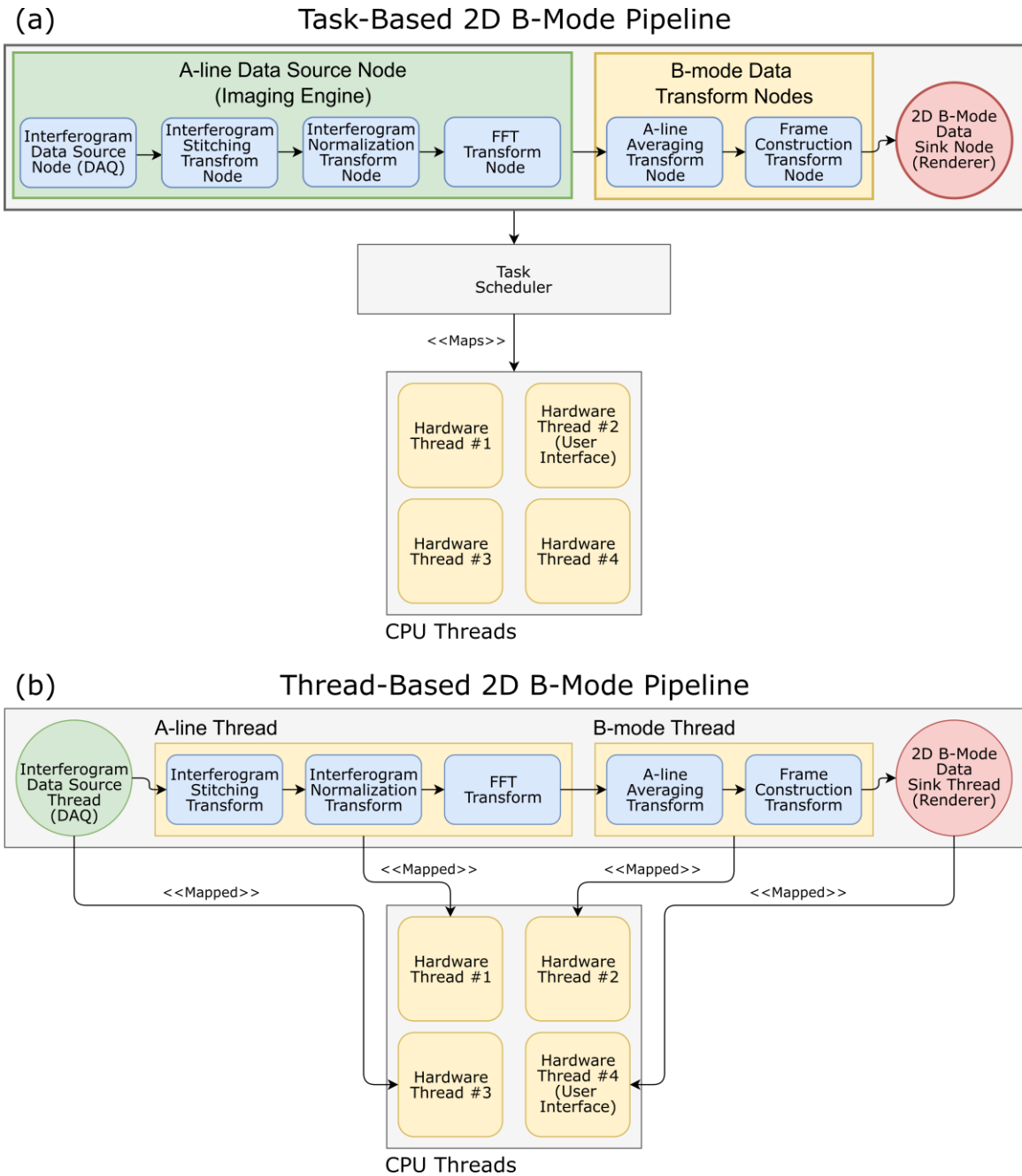
Fig. 43: 2D B-Mode Pipelines configured to use two different multi-threaded management strategies where (a) is Task-based using Intel's thread building blocks (TBB) library and (b) is thread-based using low-level, manually assigned threads.

Each imaging pipeline and work item has access to a collection of notifications contained within the notification module. These notifications are sent out at one or more different locations to help communicate to the frontend of the current state of backend processing. For example, it may notify to the frontend end that an imaging pipeline has started/stopped, if data has finished importing/exporting, or whether the system is ready

after being initialized and calibrated. Most importantly however, it notifies the frontend if the backend encountered and handled any errors that may impact the user experience. While exceptions will propagate within a single thread, they will not propagate between threads. Therefore, notifications are used to inform the frontend of any errors on asynchronously running (i.e., on threads other than the UI thread) imaging pipelines within the backend.

Synchronization Layer

The synchronization layer, shown in Fig. 44, comprises the pipeline scheduler module and pipeline factory module. Both modules work together to execute and synchronize the imaging pipelines and work items provided within the processing layer. This functionality is exposed to the frontend through a data-driven command-notification interface.
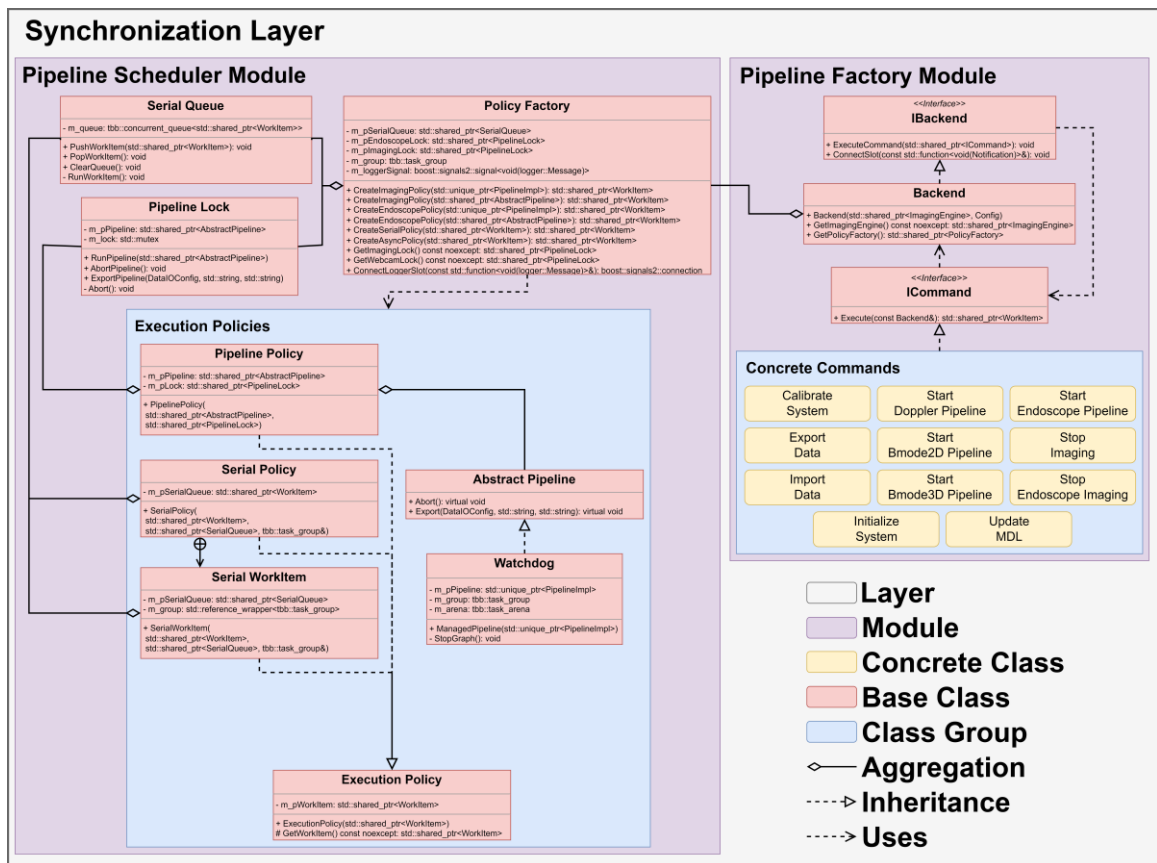
Fig. 44:   Detailed Unified Modelling Language (UML) diagram of the Synchronization layer of the backend library.

The pipeline scheduler module is responsible for synchronizing imaging pipelines and work items to avoid contention of hardware resources, enforcing expected execution order for asynchronously running processes and error handling. To achieve this, the

91

pipeline scheduler module makes use of a stack of wrapper objects, called *execution policies*, that encapsulate a pipeline or work item to coordinate execution flow. Executed from the top down in Fig. 44, policies provide fine-grained control over execution behavior. The watchdog policy handles exceptions detected during asynchronous execution of imaging pipelines. Safely shutting down the pipeline and notifying the frontend of the error, thereby improving system reliability. Under normal operation, the watchdog thread is repurposed by the pipeline, by leveraging TBB's built in concurrency composability, to avoid wasting threading resources. The serialization policy maintains expected execution order by pushing each pipeline or work item onto a lock-free queue where the front of the queue is serviced only after the previously pushed pipelines or work item have completed or the queue was previously empty. For example, a user command to start a pipeline followed by a command to stop the pipeline are executed in the order in which the user issued the commands. Since pipelines are started and executed asynchronously, it is important to ensure that stop commands are issued after a pipeline start command to avoid unexpected behavior. Finally, the pipeline policy provides a handle to the currently executing imaging pipelines and ensure only one pipeline can use the corresponding imaging hardware at a time to avoid resource contention. For example, pipelines for endoscopic imaging can run concurrently with pipelines for OCT imaging. However, multiple OCT imaging pipelines cannot be run concurrently since they require conflicting configurations of the same hardware. If other pipelines require use of the same imaging hardware as the pipeline currently executing, the pipeline policy explicitly stops the current pipeline, and waits for the hardware to be free before starting the new pipeline. This process can be made lock-free when used in conjunction with the serialization policy. Creation of execution policies are simplified through the policy factory. Given a work item or pipeline, the policy factory will automatically construct the required execution policies, in the correct order, for subsequent execution.

Execution Policies

Execution Order

Watchdog

Serialization Policy
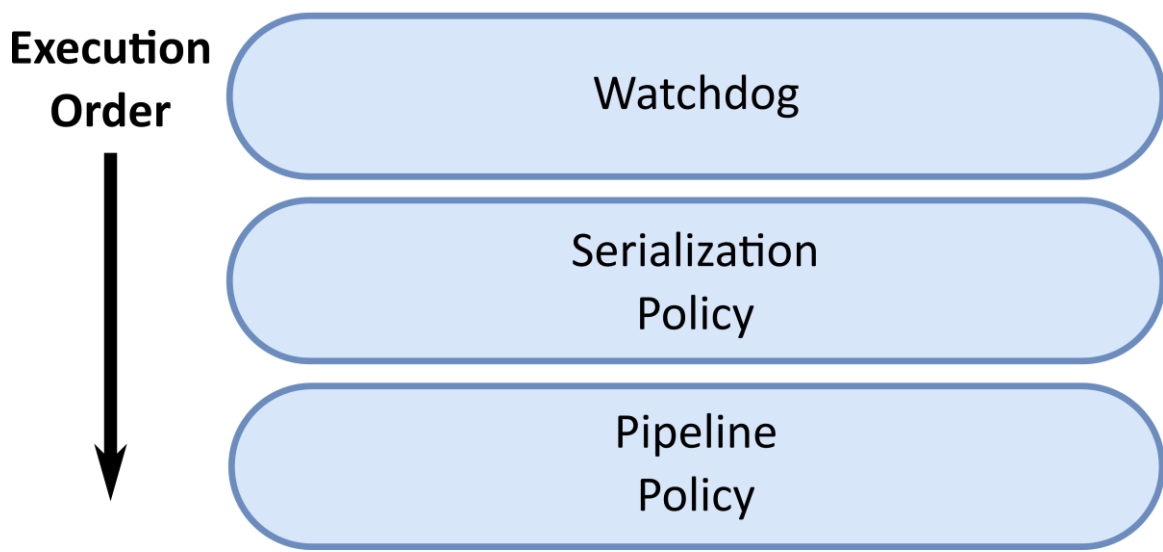
Pipeline Policy

Fig. 45: Execution policies used by the synchronization layer to control and synchronize system resources. Execution policies are a stack of wrapper objects that execute from the top down to ensure pipeline synchronizations and stability. The topmost execution policy is the watchdog, ensures that any errors thrown in long running asynchronous pipelines are handled properly and that the system is reset into a known stable state. The serialization policy maintains expected execution order of issued commands from the frontend. The OCT and endoscope lock policies provide a handle to imaging hardware resources and avoid contention by ensuring that only one pipeline can hold the lock at any given time.

The pipeline factory module, Fig. 44, is responsible for constructing imaging pipelines or work items via commands exposed to the frontend. Commands share a common interface with a singular execution method that takes in backend dependencies and returns an executable pipeline or work item. Each command is a factory that constructs a singular corresponding concrete pipeline or work item. i.e., one concrete command for each concrete pipeline or work item. Concrete commands are instantiated by the frontend with the required frontend data dependencies and display widget references (to be used by pipeline data sinks) and given to the backend interface to be executed. Once executed, each command instantiates its underlying concrete pipeline or work item, passes it to the policy factory to be wrapped in the appropriate execution policies, and returns it to the backend for execution. The returned concrete pipeline or work item is executed in the backend via its public interface, thereby hiding the implementation and synchronization details of the execution policies from the rest of the backend. This makes it easy to add new pipelines to the system without modifying the rest of the framework, greatly improving software stability, maintainability, and extendibility. Additionally, this delayed instantiation does not consume system resources until the pipeline or work item is ready to be executed by the backend. All pipelines run asynchronously. Notifications are used to communicate on the

status of data generation efforts to the frontend while data is streamed directly into the connected display widgets. This data-driven command-notification backend interface integrates easily with the HMVC framework of the frontend allowing any frontend triad to easily use backend functionality.

4.4   Results

Fig. 46 shows a series of snapshots taken from a demo video showcasing the developed software architecture being used at the point-of-care on a healthy adult volunteer. The demo video can be found in Supplementary File E. As can be seen in Fig. 46a, the simplified user interface, in comparison to the original imaging suite in Fig. 30, allows the ME-OCT imaging system to be used by a single operator holding the handpiece while imaging the volunteer. This simplification allowed the ME-OCT imaging system to be better integrated into clinical workflows and examination procedures. As demonstrated in Fig. 46b where B-mode imaging was used to visualize the volunteer's TM while they pressurized their middle ear through a routine Valsalva maneuver. The Valsalva maneuver consists of pinching the nose and attempting to exhale through the nose to pressurize the middle ear space through the Eustachian tube. Patients are often asked to perform the Valsalva maneuver during clinical ear examinations to qualitatively assess tympanic membrane compliance and middle ear ventilation. Fig. 46b−f demonstrates the software architecture's ability to support multiple diverse ME-OCT imaging modalities including B-mode, Doppler vibrometry, 3D structural imaging, and 4D structural imaging which will be discussed further in Chapter 5. Collected imaging data can be reviewed immediately after the imaging session as shown in Fig. 46d. The developed software architecture provided smooth and reliable ME-OCT imaging at the point-of-care with no noticeable lag or latency in the user experience.
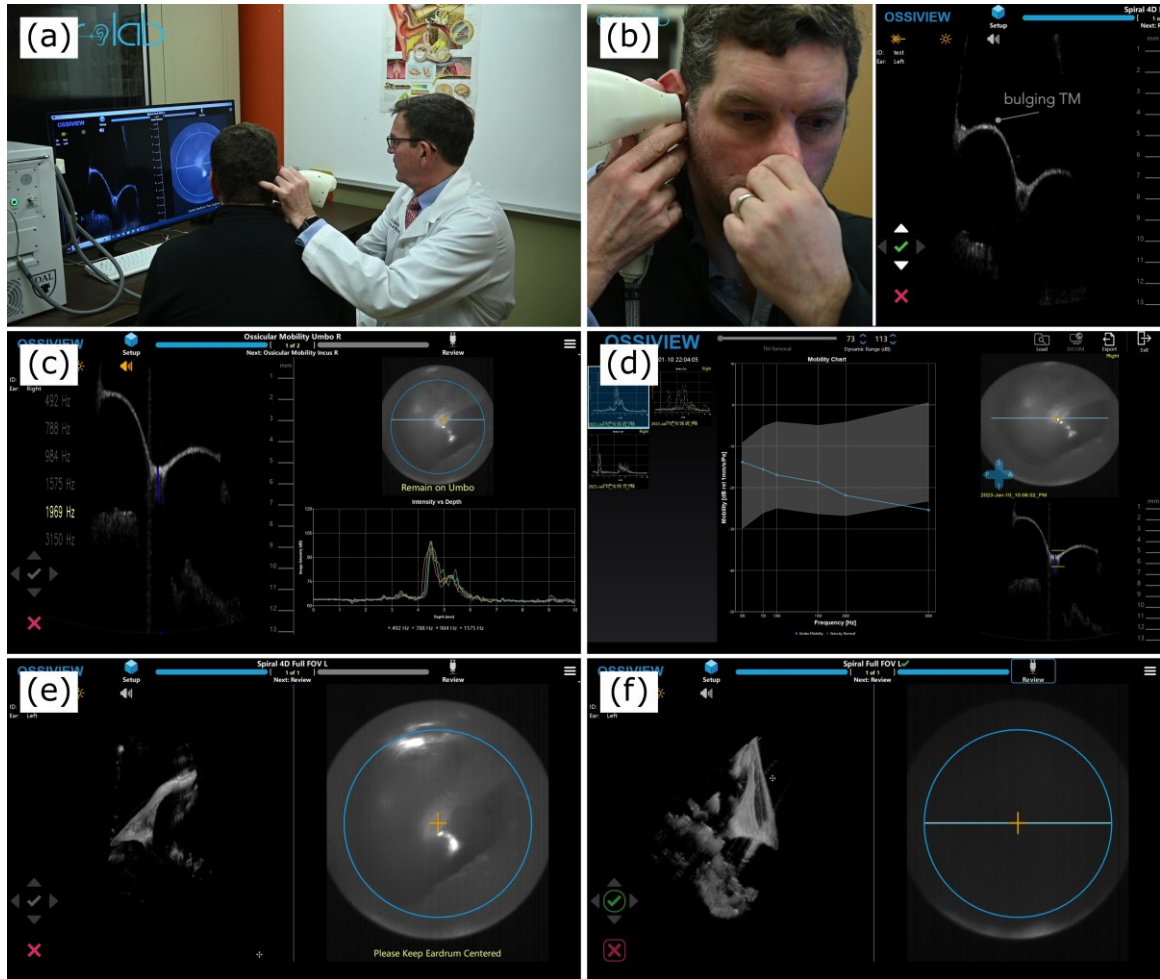
Fig. 46: Demo of OCT software built with the architecture described in this chapter being used at the point-of-care on a healthy adult volunteer. (a) Simplified user interface allowing the software to be used by a single operator holding the handpiece while imaging a healthy adult volunteer. (b) In-vivo visualization of a healthy adult volunteer's TM while they perform a Valsalva procedure using 2D B-mode imaging. (c) Demo of Doppler vibrometry imaging acquiring a full suite of audiometric frequencies centered on the umbo. (d) Review screen showcasing collected Doppler vibrometry results. (e) 3D structural imaging. (f) 4D structural imaging. A video of the software demo can be found in Supplementary File E.

4.4.1 Software Architecture Performance Profiling

Processing performance of the developed software architecture was verified through profiling of the 2D B-mode pipeline. A description of the pipeline's processing step can be found in Section 2.8. Profiling consisted of three pipeline configurations including a single threaded pipeline, and two multi-threaded pipelines (Fig. 43) but utilizing different CPU threading strategies. The first multi-threaded pipeline utilized thread-based multi-threading (Fig. 43b) in which CPU threads were manually assigned to each transform node in the pipeline. This corresponds to a low-level thread management strategy as discussed in Section 4.3.2. The second multi-thread pipeline utilized a task-based multi-threading (Fig.

43a) using TBB. All pipelines were GPU accelerated. Profiling results comparing the throughput of these pipelines can be found in Table 1.

Table 1: *Profiling results for different pipeline threading configurations. Mb/s, megabits per second; Gb/s. gigabits per second.*

Pipeline Configuration	Single Thread	Thread-based	Task-Based	Real-time Threshold
Pipeline Throughput	22.08 Gb/s	27.76 Gb/s	38.48 Gb/s	6.25 Gb/s
Frames Per Second (FPS)	72.46	91	126.1	20

Each pipeline configuration was treated as a black-box with an open-ended output, i.e., no renderer, to prevent rendering performance from impacting results, and a fixed input providing a constant stream of raw interferogram data. Throughput was calculated as the amount of data streamed through the pipeline in one second. The real-time throughput threshold is based on the data acquisition rate of the DAQ card used by our ME-OCT imaging system as described in Section 1.1.1. Any pipeline throughput below which would not be fast enough to keep up with OCT interferogram data as it is acquired. Thus, for real-time imaging, pipeline throughput needs to be at least fast enough to service the DAQ card.

As can be seen in Table 1, all pipeline configurations were more than adequate for real-time imaging rates. Compared to the single threaded pipeline, a low-level thread-based management strategy increased performance by a modest 5.68 Gb/s (18.54 FPS) or 25.7%. In comparison, the task-based thread management strategy added an additional 16.4 Gb/s (53.64 FPS) of throughput resulting a performance increase of 74.3% . Profiling was done using a Microsoft Windows 10 PC with 16 GB of RAM, a Nvidia Titan XP GPU, and an Intel i7-6800K processor.

4.5   Discussion

In this chapter we have introduced the design, and verification of a real-time heterogeneous software architecture enabling ME-OCT imaging at the point-of-care in the clinic. As shown in Fig. 46 and Supplementary File E, the software architecture was demonstrated to adequately fulfill the requirements laid out in Section 4.2.1. However, a full user-experience/usability survey must still be conducted to fully validate the system against these requirements.

The architecture was shown to have processing rates far exceeding the minimal 20 FPS target. The extra processing throughput helps facilitate running of multiple imaging pipelines in parallel alongside rendering operations without impacting real-time performance or introducing latency or lag to the user experience. Furthermore, the extra throughput afforded by task-based threading of pipelines allows the software framework to support extension to more advanced imaging modalities such as 4D imaging [111] and real-time ME-OCTA and as will be discussed in Chapter 5 and Chapter 6 respectively.

In addition to enabling later thesis work, the developed software architecture has facilitated several recent clinical studies within our lab [24,25] exploring the applications of ME-OCT to post-surgical follow-up and intraoperative cochlear implants imaging.

4.5.1 Simplifications to the Software Architecture

After completion of this work, and with the benefit of hindsight, there are a few key areas where the presented software architecture could be simplified. Starting with the frontend, the use of the HMVC, while providing a structured approach to UI design, is difficult to extend and maintain from a technical perspective. Removal of the HMVC framework would greatly simplify the architecture and could be potentially replaced with a basic object-oriented hierarchy of Qt widgets instead. For the backend, export of imaging data could be removed from the pipelines and placed into the frontend. This would improve encapsulation of data objects within each pipeline and enforce a more data-orientated approach to data flow from the backend to the frontend. Furthermore, in less data intensive ME-OCT imaging modalities like 2D B-mode or Doppler, the extra throughput provided by the task-based threading model is unlikely to provide real-world benefit that justifies the added complexity. More fined grained control over the use of threading may make it easier for future developers to extend and maintain the software architecture without having to worry about multi-threading in areas where it is not needed.

4.6 Conclusion

We have demonstrated a heterogeneous software architecture purpose-built for real-time ME-OCT at the point-of-care. Many of the design principles and techniques presented in this chapter have addressed issues and constraints of the original imaging suite. The result

is an extensible, reliable, and intuitive software framework suitable for use by a single operator and a reduced barrier for future clinical adoption of ME-OCT.

Chapter 5

Geometrically Accurate Real-time Volumetric Visualization of the Middle Ear using Optical Coherence Tomography

Joshua D. Farrell[1], Junzhe Wang[1], Dan MacDougall[6],
Xiaojie Yang[1], Kimberly Brewer[1,2], Floor Couvreur[3,4],
Nael Shoman[3], David P Morris[3], Robert B.A. Adamson[1,5]

[1]School of Biomedical Engineering, Dalhousie University, NS B3H 4R2, Canada
[2]Department of Diagnostic Radiology, Microbiology & Immunology, Physics & Atmospheric Science, Dalhousie University, Halifax, NS B3H 4R2, Canada
[3]Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, Dalhousie University, Halifax, NS B3H 4R2, Canada
[4]Department of Otorhinolaryngology, Head and Neck Surgery, AZ Sint-Jan's Hospital, Ruddershove 10, 8000 Bruges, Belgium
[5]Electrical and Computer Engineering Department, Dalhousie University, Halifax, NS B3H 4R2, Canada
[6]Audioptics Medical Incorporated, 1344 Summer St, Halifax, NS, B3H 0A8, Canada

5.1   Author Contribution Statement

J. Farrell, J. Wang, and R. Adamson collectively contributed to the design, methods, and experiments of this study. R. Adamson provided additional support in drafting the manuscript, providing relevant clinical context to the study, and helping with in-vivo data collection efforts. J. Wang was instrumental in setting up experiments, developing calibration targets, providing technical support for setting up the software for in-vivo data collection, and further software support to help offline analysis of collected data.

F. Couvreur, N. Shoman, and D. Morris provided key clinical insight into the clinical relevancy of this work and provided constructive feedback around results discussion helping shape the manuscript. D. Morris provided additional support in collecting in-vivo imaging data and voice-over for supplementary material.

X. Yang prepared cadaver specimens for the imaging sessions and crucial support for cadaver imaging. K. Brewer provided gave access to the micro-CT imaging system

used and performed micro-CT cadaver imaging. D. MacDougal designed the OCT imaging system used in this study and provided insights into optical design and distortion.

J. Farrell was responsible for conducting the experiments, developing the methodologies and real-time software required for the experiments, data analysis and renders, and served as the lead author for journal submission.

5.2 Preamble

This chapter of my thesis is based on the manuscript for an article that was recently published in the journal Biomedical Optics Express (BOE) [111]. While the content of this manuscript remains largely unaltered, some sections have been omitted or revised to avoid repetition of any previously discussed material found in earlier chapters. Additionally, minor adjustments were made to ensure consistency between the material presented in this chapter and the rest of the thesis.

5.3 Abstract

We introduce a novel system for geometrically accurate, continuous, live, volumetric middle ear optical coherence tomography imaging over a $10.9\ mm \times 30^{\circ} \times 30^{\circ}$ FOV from a handheld imaging probe. The system employs a discretized spiral scanning (DC-SC) pattern to rapidly collect volumetric data and applies real-time scan conversion and lateral angular distortion correction to reduce geometric inaccuracies to below the system's lateral resolution over 92% of the FOV. We validate the geometric accuracy of the resulting images through comparison with co-registered micro-computed tomography (micro-CT) volumes of a phantom target and a cadaveric middle ear. The system's real-time volumetric imaging capabilities are assessed by imaging the ear of a healthy subject while performing dynamic pressurization of the middle ear in a Valsalva maneuver.

5.4 Introduction

Apart from its use in measuring TM thickness [19,21,173,184], ME-OCT imaging has so far only been used as a qualitative tool to assess the presence and relative location of middle ear structures without reference to or validation against an absolute coordinate space. For many metrology-based applications such as prosthesis sizing, monitoring anatomical changes in healing or in progressive disease, assessing congenital malformations of the middle ear, measurement of TM perforations and quantifying the degree of prosthesis

migration, an ability to produce images with high geometric fidelity to the patient anatomy would enhance ME-OCT's diagnostic value.

However, obtaining geometrically accurate OCT images of the middle ear is challenging. Recall from Section 1.1.1, that because the middle ear lies at the end of a 3 cm long ear canal that is often narrow and curved, middle ear OCT uses an entocentric imaging geometry where a wide angular FOV is imaged from a small aperture [1,19,21]. As previously discussed in Section 2.5, such an imaging geometry produces images exhibiting fan-beam distortion [105] when OCT data is displayed without correcting for the fact that the data was collected along non-parallel scan lines. The geometric accuracy of OCT images may also be affected by distortions introduced from imperfect imaging optics or a non-linear relationship between scanning mirror drive signals and scanning mirror angle. If geometrically accurate images of anatomy are to be displayed to clinicians, all these distortions must be corrected so that image data is displayed on an equally spaced Cartesian grid of points and the images represent a scaled version of the true anatomy.

In point-of-care applications, it is desirable to have imaging data collected, processed, and rendered in real time in both 2D and 3D, including any transformations or corrections applied to mitigate distortion effects. Continuous, live 3D middle ear imaging, sometimes referred to as 4D imaging (3D over time) [185], could enable intuitive clinical exploration of the middle ear space and the capturing of diagnostically relevant dynamic processes such as middle ear muscle contraction, middle ear conformation changes in response to pressurization and the insertion of surgical instruments into the middle ear space. 4D visualization necessitates a capability to process and render OCT data at real-time rates as fast as it can be produced from the OCT imaging hardware including the application of any image transformations or corrections applied to achieve a geometrically accurate image. It also entails a data collection strategy capable of rapid scanning of the OCT beam across the FOV so as to make optimal use of the available scanning mirror bandwidth and avoid dead time where the mirrors are scanning but not collecting useful imaging data [185].

Volumetric OCT imaging has previously been reported for ophthalmic OCT imaging systems employing a constant linear velocity spiral scanning (CVL-SC) pattern to drive a galvanometric scanner [185–188]. Scanning in a spiral pattern allows continuous,

uninterrupted data collection and optimizes the available mirror bandwidth by scanning both mirror axes at similar speeds, resulting in efficient interrogation of a volume. In previous studies employing spiral scanning, data collected along the spiral beam trajectory was interpolated onto a uniform Cartesian grid for display either in post-processing using bi-linear interpolation [185] or in real time using nearest-neighbor interpolation [188].

Geometric distortion in OCT images is well documented, and many approaches to correcting for it have been reported [189–194]. However, to our knowledge, these algorithms have never been previously integrated into a single system capable of acquiring and rendering geometrically accurate OCT volumes, in real-time, in-vivo. Nor have they been applied to image the middle ear [194] or assessed for geometric accuracy against other gold standard volumetric imaging modalities such as computed tomography (CT).

In this chapter we present novel techniques for real-time, geometrically accurate 4D ME-OCT imaging over a $10.9\ mm \times 30^{\circ} \times 30^{\circ}$ sector from a handheld, entocentric probe. OCT images were acquired using a two-axis, gimbal-less MEMS mirror driven to scan in a discretized spiral pattern, scan converted to correct fan-beam distortion, corrected for optical and mirror-related distortions, and displayed to the user in real-time using a custom-volumetric ray-cast renderer as discussed in Chapter 3. We verify the system by quantitative measurement of geometric accuracy against benchtop phantom imaging targets and validate its use for quantitative imaging of the middle ear by comparing 3D OCT images obtained in human cadaveric middle ears to micro-CT images of the same ear. A cadaver model is also used to demonstrate 4D OCT visualization of a middle ear implant in a simulated diagnostic assessment scenario. Finally, we demonstrate the real-time imaging capabilities of the system by performing continuous, live, in-vivo imaging of a subject who performs a Valsalva maneuver that causes dynamic pressurization and conformational changes to the middle ear which are captured in real time 4D OCT.

5.5 Methods

Imaging in this chapter was done using our ME-OCT system as described in more detail in Section 1.1.1. The rendering engine and software architecture described in Chapter 3 and Chapter 4 respectively was extended to facilitate real-time processing and display of geometrically accurate, live, continuous 4D ME-OCT images.

5.5.1 Discretized Spiral Scanning

Given the cylindrical symmetry of the imaging optics and the limited bandwidth of the MEMS mirror, a spiral scanning technique was developed to maximize the volumetric scanning rate and to match the extent of scanning to the exit pupil dimensions. This ensured that no scanning time was wasted imaging outside the system's aperture or during times when no data collection occurred.

Two-axis, gimbal-less MEMS mirrors [195] are appealing for handheld OCT systems due to their small size and weight. While they lack position sensing capabilities, their position is highly repeatable when driven open-loop [196]. With an open-loop driving strategy, a one-time calibration must be used to map drive signals to correct angular locations. The high repeatability of the mirrors then ensures that this calibration remains valid over time. In this section we present our approach to using a spiral scan pattern to drive a two-axis MEMS mirror to collect volumetric data.
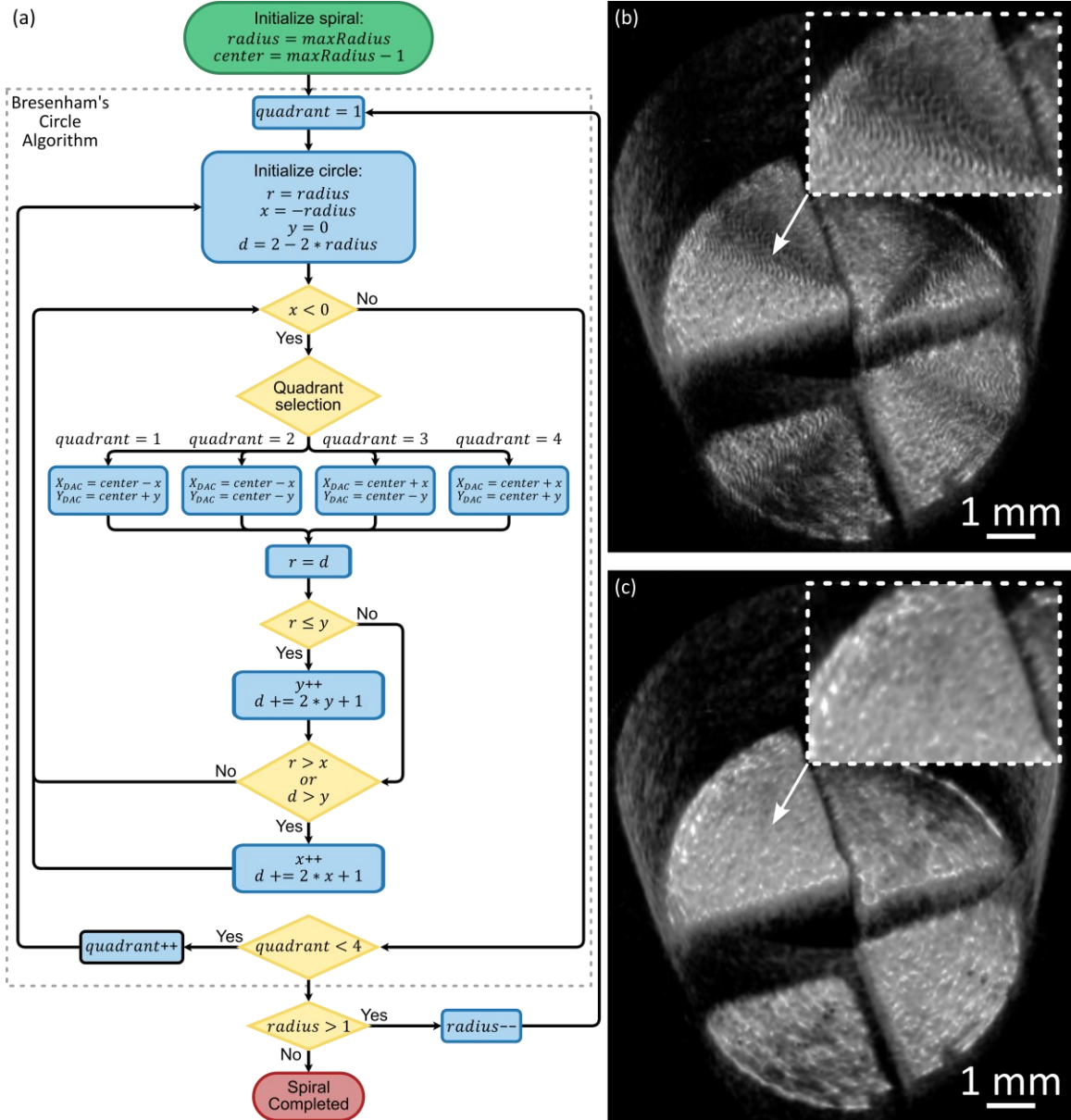
Fig. 47: Discretized spiral scanning (DC-SC) algorithm and OCT volume smoothing. (a) Flowchart to generate a discretized spiral scan pattern based on the Bresenham's circle drawing algorithm. Discretized spiral OCT volume before (b) and after (c) median filtering. $\mathrm{maxRadius}$, control variable specified by the user to control the FOV of the spiral scan; radius, index variable tracking the spiral radius currently being generated and determine when the spiral is finished; center, center point of the spiral scan along the discretized grid; $\mathrm{quadrant}$, index variable of the current circle quadrant being generated; x and y, circle indices along the x-axis and y-axis respectively; d, decision variable used to increment x and y such that circular error is minimized along the discretized grid; r, temporary variable to store the radius of the circle being generated; X_{DAC} and Y_{DAC}, DAC output signals driving the MEMS mirror.

Our spiral scan pattern is a discretized approximation to a constant linear velocity (CLV) spiral scan [185] as shown in Fig. 48a and Fig. 48b where the discretized spiral was generated using Bresenham's algorithm [197]. The flowchart in Fig. 47a depicts the proposed DC-SC algorithm. Starting on the outer edge of the image FOV, we generate a discrete approximation of a circle of a fixed radius one quadrant at a time. After completion

of the circular scan, the radius is decremented and scanning continues along a new circle with a radius one pixel smaller than the previous circle. This process continues until the center of the FOV is reached. The radius is then reset to acquire the next spiral volume. Each location of the discretized spiral scan (Fig. 48b) lies on a grid of uniformly spaced angles allowing the data to be organized into a discrete (r, θ, ϕ) array. However, due to the spiral discretization, not all lateral locations in the volume are covered by scan lines resulting in the final constructed volume having "holes" where the density of scan lines is lower than in surrounding areas (e.g. along the diagonals in Fig. 47b and Fig. 48b). This effect can be addressed by applying a 2D median filter laterally across (θ, ϕ) at each depth to homogenize the density of image lines. Fig. 47c shows the result of applying such a median filter on the volume of Fig. 47b.
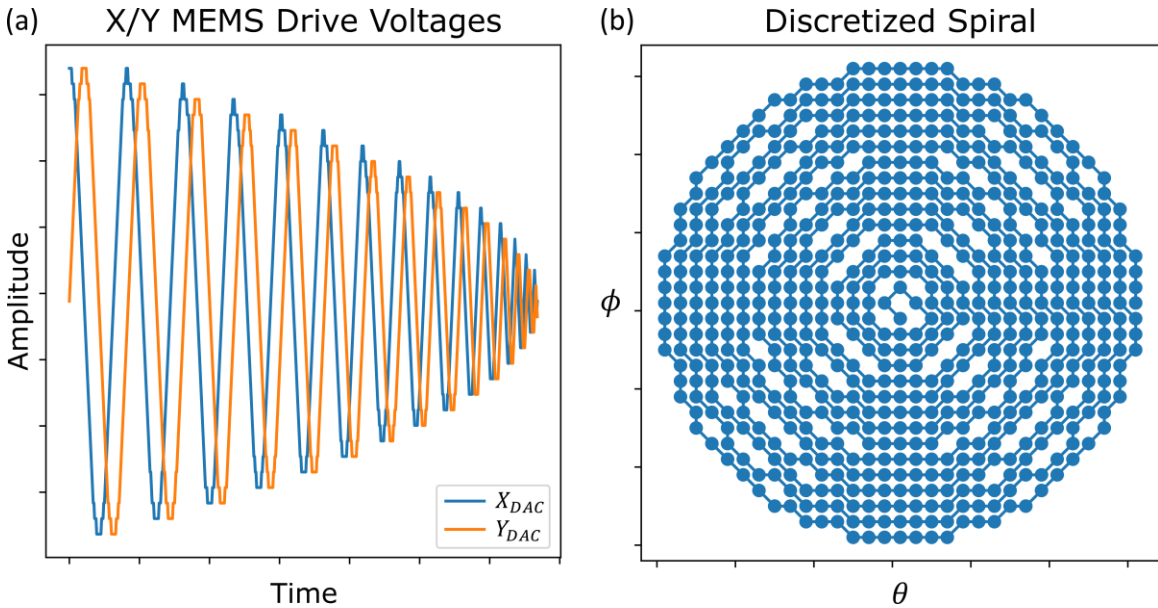
Fig. 48:   *Discretized spiral scanning (DC-SC) pattern and waveforms. (a) MEMS mirror driving voltage waveforms for spiral scan. (b) En-face view of the discretized spiral scan over the angular FOV. For illustration, only a fraction of the spiral scan's FOV was used to produce the waveforms (a) and (b). X_{DAC} and Y_{DAC}, FPGA's 10-bit DAC signals.*

As the Bresenham's circle algorithm only makes use of integer math, it can be implemented using economical hardware without the need for lookup tables (LUTs). In our system we implemented the algorithm on a field programmable gate array (MachXO2, Lattice Semiconductor; USA) driving a 10-bit DAC to generate volumes containing 512×512 image lines corresponding to a $30° \times 30°$ FOV in 4.5 seconds (0.22 Hz). A speed-up factor of 4 can be achieved for every factor of 2 decrease in acquired FOV diameter. For example, halving the FOV to $15° \times 15°$ we can acquire volumes containing

256×256 image lines in 1.13 seconds ($0.89\ Hz$) and quartering the FOV to $7.5° \times 7.5°$ we can acquire volumes containing 128×128 image lines in 0.28 seconds ($3.56\ Hz$). For our MEMS mirror, the manufacturer's maximum recommended drive signal bandwidth was $200\ Hz$. With this bandwidth limitation we collect $N = 5$ A-lines per image line for all acquired volumes and average them together to form the line.

A limitation of our spiral scanning technique is that as the beam gets closer to the center of the spiral (i.e., to the center of the FOV) and the scanned radius gets smaller, the frequency of the drive signal increases as can be seen in Fig. 48a. At some point, the frequency of the drive signal will approach the $200\ Hz$ bandwidth limit of the drive electronics, resulting in an inability of the MEMS mirror to follow the spiral path. This results in an artefact at the center of the spiral. The severity of the artefact increases as the volume scanning rate increases and the artefact takes up an increasing fraction of the image volume.

5.5.2 Real-time Volumetric Scan Conversion using GPU Texture Memory

Because the optical system shown in Fig. 1d and Fig. 1e is designed to image the MEMS mirror rotation axis to the exit pupil, all OCT beamlines intersect at the pupil. Volumes are constructed by scanning the angles (θ, ϕ) that the beamlines make about this common origin and OCT data is collected in spherical coordinates (r, θ, ϕ) where r is the radial distance along the A-line and (θ, ϕ) are the spherical angles about the exit pupil center. To display data with geometrically correct proportions, a transformation must be made from the spherical coordinates in which it was measured (Fig. 49a) to the Cartesian coordinates in which it is displayed (Fig. 49b). This type of image conversion from spherical to Cartesian coordinates is commonplace in phased array ultrasound imaging where it is called scan conversion [198], and we will use that terminology here.
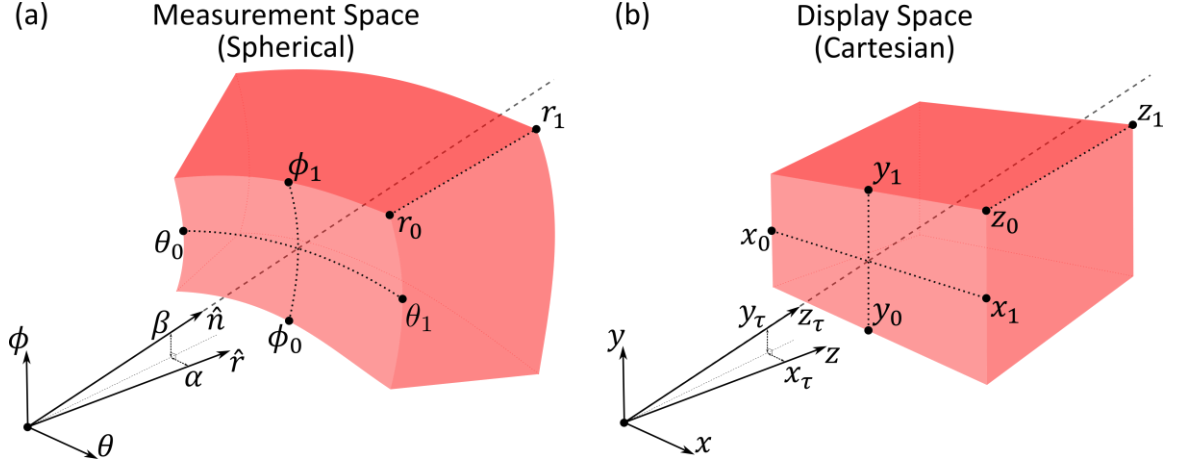
Fig. 49: Spherical measurement space (a) in which OCT volume data is acquired and the Cartesian display space (b) in which it is displayed. $\hat{n}$ MEMS mirror surface normal vector; $\hat{r}$ handpiece's optical imaging axis; α, β tilt angles between $\hat{n}$ and $\hat{r}$ along θ and ϕ respectively; r_0, r_1 radial extent of the spherical measurement space; θ_0, θ_1 and ϕ_0, ϕ_1 are the lateral angular extents of the spherical measurement space; x_τ, y_τ, z_τ tilted Cartesian axis; x_0, x_1 and y_0, y_1 the lateral extents of the Cartesian display space; z_0, z_1 depth extent of the Cartesian display space.

To perform real-time volumetric scan conversion for ME-OCT we extended the real-time 2D ultrasound scan conversion kernels in the open-source FAST toolkit [199]. FAST provides a set of common medical image processing algorithms including 2D scan conversion implemented as GPU kernels to achieve very fast processing times by exploiting the massive parallelism of GPU architectures. GPUs provide dedicated hardware for interpolation between voxel values stored in texture memory resulting in a significant increase in processing throughput as compared to CPU-based approaches.

In our scan conversion method, the extent of the spherical measurement space $\{[r_0, r_1], [\theta_0, \theta_1], [\phi_0, \phi_1]\}$ is converted into a Cartesian cuboid of extents $\{[x_0, x_1], [y_0, y_1], [z_0, z_1]\}$ that bounds the displayed data, where:

$$[x_0, x_1] = [r_1 \sin(\theta_0), r_1 \sin(\theta_1)]$$
$$[y_0, y_1] = [r_1 \sin(\phi_0), r_1 \sin(\phi_1)] \tag{11}$$
$$[z_0, z_1] = [r_0 \cos(\theta_0), r_1]$$

The display space bounded by these Cartesian extents is divided into an integer number of voxels $D_x \times D_y \times D_z$ where D_x, D_y and D_z are the number of voxels along each axis. The integer voxel coordinates (x, y, z) of the display space are then normalized to the range $[0,1]$ to form the normalized Cartesian coordinates (x_f, y_f, z_f):

$$x_f = \frac{x(x_1 - x_0)}{(D_x - 1)} + x_0$$

$$y_f = \frac{y(y_1 - y_0)}{(D_y - 1)} + y_0 \tag{12}$$

$$z_f = \frac{z(z_1 - z_0)}{(D_z - 1)} + z_0$$

Due to imperfect optical alignment, there may be a non-zero tilt between the normal of the MEMS mirror when the mirror drive signals are set to zero and the image axis. To account for this, we introduce a set of tilted floating-point coordinates (x_τ, y_τ, z_τ) given by:

$$x_\tau = x_f \cos \alpha + z_f \sin \alpha$$

$$y_\tau = x_f \sin \beta \sin \alpha + y_f \cos \beta - z_f \sin \beta \cos \alpha \tag{13}$$

$$z_\tau = -x_f \cos \beta \sin \alpha + y_f \sin \beta + z_f \cos \beta \cos \alpha$$

Where (α, β) represent the relative angles between the MEMS mirror's normal $\hat{n}$ and the image axis $\hat{r}$ along $(\theta = 0, \phi = 0)$. The spherical coordinates (r, θ, ϕ) corresponding to each point (x_τ, y_τ, z_τ) are then calculated as:

$$r = \sqrt{x_\tau{}^2 + y_\tau{}^2 + z_\tau{}^2}$$

$$\theta = \arctan\left(\frac{x_\tau}{z_\tau}\right) \tag{14}$$

$$\phi = \arctan\left(\frac{y_\tau}{z_\tau}\right)$$

Finally, the measured data is interpolated onto the point in spherical space (r, θ, ϕ) that corresponds to each set of voxel coordinates (x, y, z) of the display space [194]. All these steps are performed continuously in real-time using GPU textures at a rate of $52.8\ Hz$ ($18.9\ ms$/vol) for an OCT volume containing $D_x \times D_y \times D_z = 512 \times 512 \times 330$ voxels, more than enough to enable real-time 4D imaging.

A motorized delay line in the reference arm of our OCT engine was used to adjust the image depth to accommodate different ear canal lengths (see Fig. 1a). Changes to the reference arm delay effectively change the radial extent of the spherical measurement space from $[r_0, r_1]$ to $[r_0 + \Delta/2, r_1 + \Delta/2]$ where Δ is the amount of additional reference arm delay. The factor of $\frac{1}{2}$ arises because in our interferometer the reference arm delay is a one-

way delay while the sample arm delay is two-way. Our scan conversion algorithm is set up to automatically adjust for these changes by scaling the apparent width and radius of curvature of the proximal and distal boundaries of the image window while keeping it centered on the screen.

5.5.3 Geometric Calibration Through Minimization of Residual Surface Fitting Error

To apply the scan conversion method described in the previous section, we first calibrated the system to obtain estimates of the extent of the spherical measurement space, i.e., to determine $\{[r_0, r_1], [\theta_0, \theta_1], [\phi_0, \phi_1]\}$ for a particular reference arm delay. We did this by imaging a flat, white acrylic plate placed near the distal end of the imaging window. A 3D image of the plate was then acquired, and the proximal plate surface was segmented within each image line using a peak-finding algorithm followed by filtering to smooth out the extracted surface in the lateral dimensions. The output from the segmentation were the radial indices S_{mn} of the surface location for each discrete angle (m, n) in the spiral scan grid of Fig. 48b. Points near the center of the image that were affected by the bandwidth limit artefact were excluded from the calibration volume. The radial extent of the imaging window $\Delta R = (r_1 - r_0)$ was measured by translating the reflector from the proximal to distal end of the window using a graduated translation stage until the image of the plate surface intersected the proximal and distal bounds of the image window along the central line. ΔR was found to be $10.89 \; mm$ in our system.

Defining the value of S_{mn} along $\hat{r}$ to be S_0, we construct a plane in Cartesian space perpendicular to the z-axis at a depth $z = S_0$. A given a set of parameters $\rho = \{[r_0, r_1 = r_0 + \Delta R], [\theta_0, \theta_1], [\phi_0, \phi_1], [\alpha, \beta]\}$ describes a transformation of this plane into spherical coordinates. Applying this transformation according to Eq. (12) – (14) of the previous section, we calculate the intersection of this plane in spherical coordinates with each scan line $(\theta_m, \phi_n) = (m\Delta\theta, n\Delta\phi)$ in the spiral scan. Here the angular grid spacing is $\Delta\theta = \frac{\theta_1 - \theta_0}{M}$ and $\Delta\phi = \frac{\phi_1 - \phi_0}{N}$ where M and N are the number of grid points in the θ and ϕ directions. The radial location at which the plane intersects each scan line is $I_{mn}(\rho)$. We then perform a least-squares minimization to determine the set of parameters ρ that

minimizes the error between $I_{mn}(\rho)$ and S_{mn} over the scan to solve the minimization problem in Eq. (15).

$$\underset{\rho}{\mathrm{argmin}} \sum_{m,n} \|I_{mn}(\rho) - S_{mn}\|^2 \qquad (15)$$

This minimization process is illustrated in Fig. 50. The minimization was performed using the minimize function in the Python SciPy optimization toolbox [200] using the Sequential Least Squares Programming (SLSQP) method with default option values. The parameters ρ obtained from this minimization were stored and applied to all subsequent real-time scan conversions performed in this study. The calibration was found to remain stable across the six-month period of the study and did not require updating or modification.
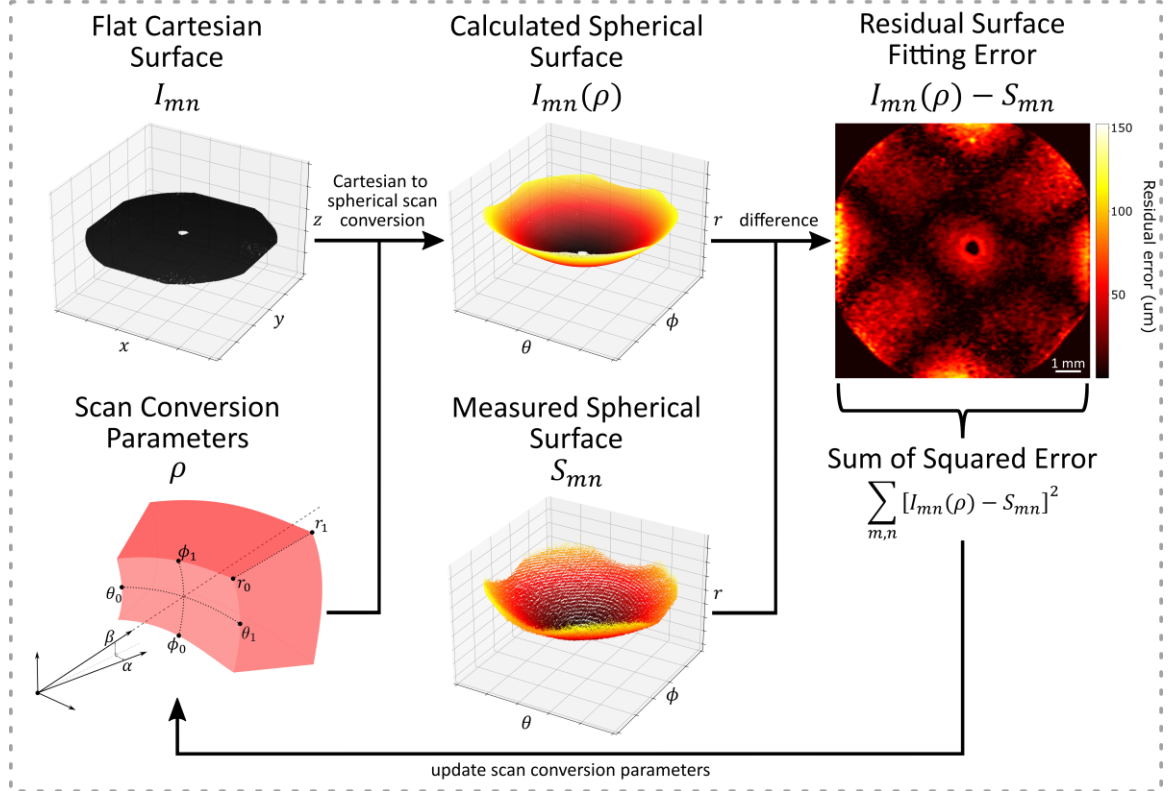
Fig. 50: *Diagram describing the geometric calibration procedure for finding the scan conversion parameters $\rho = \{[r_0, r_1], [\theta_0, \theta_1], [\phi_0, \phi_1], [\alpha, \beta]\}$ that optimize the fidelity of the OCT volumes with the known geometry of a flat surface. I is an ideal flat Cartesian surface reference; $I(\rho)$ is the calculated spherical surface after Cartesian-to-spherical coordinate transformation using parameters ρ; S is the measured spherical surface segmented and extracted from an OCT volume of a flat white acrylic plate; $\sum_{m,n} \|S_{mn} - I_{mn}(\rho)\|^2$ is the sum of squared error between the surface model and measured surface.*

5.5.4   Lateral Angular Distortion Correction

MEMS mirrors are known to exhibit inter-axis crosstalk in which the sensitivity of mirror tilt angle to applied voltage in the x-axis depends on the tilt angle in the y-axis and vice versa [34–36]. Using a similar MEMS scanner to the one in our system, Kim et al. determined angular error during dual axis scanning due to crosstalk to be as high as 6.13% depending on MEMS scanner alignment [203]. Furthermore, the dependence of MEMS mirror angle on drive voltage is non-linear with mirror angle tending to saturate at the ends of the drive voltage range. The effect of mirror crosstalk and nonlinear distortion must be corrected to obtain geometrically accurate OCT volumes.

Our method for correcting distortion due to MEMS scanner nonlinearity and crosstalk is illustrated in Fig. 51. Taking a similar approach to Izatt et al. [185], we captured an *en-face* image of a uniform dot grid. Our dot grid consisted of an ultraviolet (UV)-

printed black acrylic plate with a rectangular grid of circular dots (UJF-6042 Mk II, Mimaki Engineering, Japan) with a center-to-center spacing of 0.5 mm and a dot diameter of 0.25 mm printed at a resolution of 21 μm $\times$ 21 μm. The dot grid was imaged near the distal edge of the imaging window to maximize the number of dots visible in the image and an *en-face* image was generated from a scan-converted OCT volume. The *en-face* image of the dot grid captures the observed 2D map of scan angles corresponding to the driving voltages of the spiral scan pattern after undergoing lateral distortion effects caused by the MEMS mirror inter-axis crosstalk and nonlinearity. For each dot in the *en-face* image, we extracted its apparent angular coordinate on the distorted grid using the Discorpy library [204]. With the number of dots in the grid known, a uniform dot grid consisting of the same number of dots was then created. A one-to-one relationship between the coordinates of the extracted and uniform grid was established to form an array of uncorrected/corrected coordinate pairs. Finally, a 2D correction map was created by interpolating this coordinate pair array across a normalized grid containing the same number of pixels as the original *en-face* image. With the lateral correction map stored as a GPU texture, each pair of corrected angular values is read from the correction map using its corresponding pair of drive voltage values $[\theta(X_{DAC}, Y_{DAC}), \phi(X_{DAC}, Y_{DAC})]$ (Fig. 48) as texture coordinates and used during volumetric scan conversion to correct lateral distortion in real-time. The angular calibration only needed to be performed at a single depth to correct lateral distortions at all depths. Due to the excellent open-loop repeatability of the MEMS mirrors, this calibration only needs to be done once, although, it must be repeated if the scanning optics are physically modified or adjusted.

| MEMS Axial Crosstalk | Observed Dot Grid | Corrected Dot Grid |

lateral distortion

lateral correction

Grid Segmentation · Grid Extraction · Uniform Grid
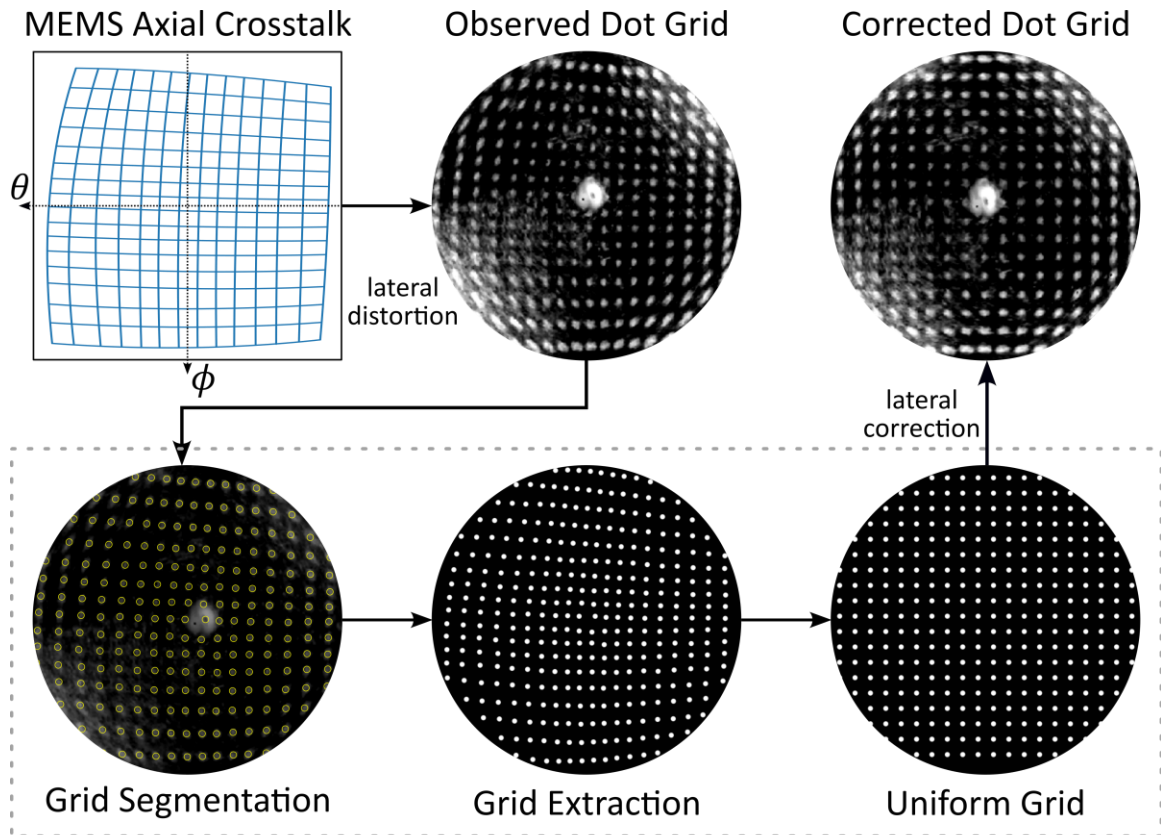
Correction Map Creation

Fig. 51: Correction of lateral angular distortion caused by MEMS nonlinearity and interaxis crosstalk. An en-face image of a reference dot grid exhibiting lateral distortion was generated from a scan converted OCT volume. The dots were segmented, extracted, and matched to a uniform grid creating a correction map for correction of the observed lateral distortions in real-time during scan conversion. (Center-to-center dot spacing of 0.5 mm, dot diameter of 0.25 mm)

5.5.5 Geometrical Comparison of ME-OCT and co-registered micro-CT volumes

The geometric accuracy of corrected OCT volumetric data was compared to nominal design dimensions of a 3D printed phantom target and to co-registered micro-CT images of a cadaveric middle ear. Phantoms were designed in Fusion 360 and printed in photopolymer resin with a UV stereolithographic 3D printer (Form 3, Formlabs, USA) using a $50\,\mu m$ layer thickness. The phantom consisted of a tube with holes in the circumferential wall and a four-step spiral staircase pattern at the distal end as shown in Fig. 53a.

For cadaveric middle ear imaging, a fresh-frozen human cadaveric temporal bone, the right ear from a 30-year-old female was cut *en-bloc,* thawed, mounted, and prepared with the otic capsule, middle ear, ear canal, tympanic membrane and soft tissues of the ear canal and pinna left intact. The tympanic membrane was debrided and wax, hair, and other

debris were removed from the ear canal to provide a clear view of the tympanic membrane. The OCT view of the tympanic membrane in this preparation is typical of that obtained in live subjects. During the experiment and immediately before micro-CT imaging, the ear canal was irrigated with saline and drained to maintain hydration of the tympanic membrane. Procedures used in this study were approved by the Dalhousie University Research Ethics Board under protocol 2021-5486.

Ground truth anatomy for the temporal bone and phantoms was obtained using a Triumph LabPET4/CT (Trifoil Imaging, USA) micro-CT imaging system (volume shape: $512 \times 512 \times 512$) with a slice thickness and FOV of 54 μm and 27.65 mm, respectively for the phantom target and 155 μm and 79.6 mm for the cadaveric middle ear. Both micro-CT and OCT volumetric data was exported to DICOM format files for offline co-registration in 3DSlicer [152]. Artefacts were removed from the OCT volume using a previously described convolutional dictionary learning filter [167] and the micro-CT volume was cropped to a region of interest (ROI) that overlapped with the OCT volume. The two datasets were aligned using semi-manual, rigid, fiducial registration using eight boney anatomical landmarks visible in both datasets. The geometric accuracy of the OCT volume was evaluated using root mean squared error (RMSE) by taking the micro-CT image as ground truth for the cadaver and the nominal design specifications as ground truth for the phantom.

5.6   Results

5.6.1   Verification of Geometric and Lateral Angular Distortion Correction

The calibration process of Section 5.5.3 gave spherical imaging extents for our system of $\Delta R = (r_1 - r_0) = 10.89$ mm , $(\theta_1 - \theta_0) = 30.6°$, $(\phi_1 - \phi_0) = 31.9°$ and tilt angles $(\alpha, \beta) = (-0.4°, -0.4°)$. To assess the accuracy of the correction, it was applied to a set of OCT volumes of the flat acrylic plate translated to 19 different depths covering the axial range of the image window from r_0 to r_1. After scan conversion, the RMSE surface flatness across all depths was 69 μm. When the lateral distortion correction described in Section 5.5.4 was also applied, as shown in Fig. 52e, the RMSE surface flatness improved to 42 μm. The improvement in RMSE error from lateral distortion correction arises because of coupling between angular and axial errors towards the edges of the FOV.

Fig. 52a and Fig. 52c show scan-converted OCT volumes of the reference dot grid acquired using our discretized spiral scanning algorithm before and after lateral distortion correction. In addition to the distortion effects that we correct for, Fig. 52a and Fig. 52c also show clear evidence of the astigmatism of the GRIN rod objective in the diagonal smearing of points into streaks near the corners of the FOV. We did not attempt to correct for this effect.

Fig. 52b and Fig. 52d show heat maps of the absolute error between the measured dot grid before and after lateral distortion correction. Prior to correction, angular error was highest around the corners of the FOV where the effects of both mirror crosstalk and non-linearity are strongest. Distortion was also high at the center of the image due to the mirror bandwidth limitation artefact. The RMSE angular error before correction was found to be $1.15°$ ($383 \, \mu m$ at the distal end of the image). After lateral distortion correction, as Fig. 52d shows, the RMSE residual angular error was found to be $0.036°$ ($12 \, \mu$m at the distal end of the image).

Total residual angular and radial error shown in Fig. 52f, was largest around the center and the diagonal corners of the FOV. These residual errors arise mainly from inaccuracies in correct identification of dot coordinates in the image due to astigmatic blurring and the presence of the central artefact. Even with these error sources, the worst-case total error after scan conversion and lateral distortion correction was $70 \, \mu m$ as shown in Fig. 52f. 92% of pixels had an accuracy of $40 \, \mu m$ or better, indicating that the residual geometric error was less than the lateral resolution over most of the FOV. Fig. 52c and Fig. 52d also show that after lateral distortion correction the FOV of the OCT volume is no longer circular since the crosstalk introduces anisotropic angular range limits with a larger angular range along the diagonals than along the mirror axes. A summary of the ME-OCT imaging system's performance for various FOVs after geometrical and lateral distortion correction can be found in Table 2.
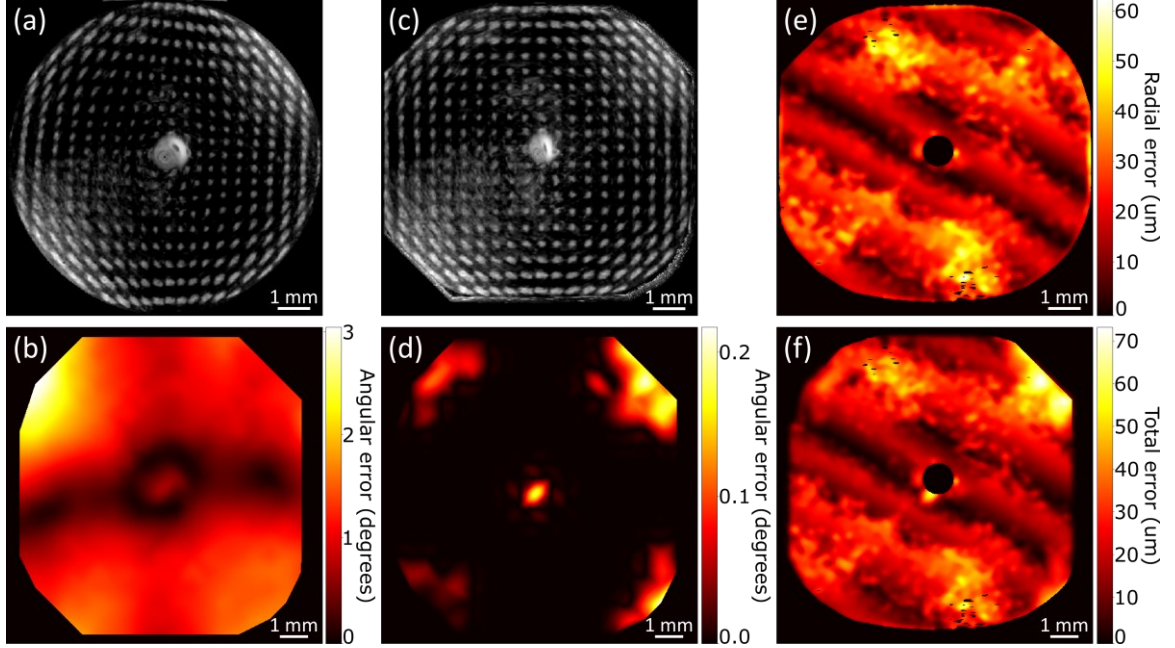
Fig. 52: (a) En-face image of a reference dot grid (0.5 mm center-to-center spacing, dot diameter of 0.25 mm) generated from a scan-converted discretized spiral OCT volume before lateral distortion correction. (b) Heatmap of angular error (degrees) between the measured grid points and the true grid point locations. (c) En-face image of the dot grid after lateral distortion correction. (d) Residual angular error (degrees) between the measured grid points in (c) and the true grid point locations. (e) Residual radial error (μm) after scan conversion and lateral distortion correction. (f) Residual total lateral and radial error $\delta = \sqrt{\Delta r^2 + r\Delta\theta^2 + r\Delta\phi^2}$ (μm) after scan conversion and lateral distortion correction.

Table 2: Benchmarks of the ME-OCT imaging system for various field of views (FOV) including FOV dimension, volumetric acquisition rate, 3-dB lateral and axial resolution in air, and residual error after geometric and lateral distortion correction. RMSE, root mean squared error.

Imaging FOV	Voxels per volume (x, y, z)	Volumetric acquisition rate (vol/s)	Lateral resolution (μm)	Axial resolution (μm)	Residual lateral angular RMSE (μm)	Residual axial RMSE (μm)
$30° \times 30°$ $\times 10.9\ mm$	512×512 $\times 330$	0.22				
$15° \times 15°$ $\times 10.9\ mm$	256×256 $\times 330$	0.89	40	40	12	42
$7.5° \times 7.5°$ $\times 10.9\ mm$	128×128 $\times 330$	3.56				

We verified the geometric accuracy of scan-converted and lateral distortion corrected OCT volumes by comparison against nominal design dimensions of a 3D printed phantom target as shown in Fig. 53 and by the similarity of the OCT volume to a co-registered CT volume. Fig. 53a shows overlaid OCT and micro-CT images of the phantom.

116

Fig. 53b–d show measurements made between various landmarks within the OCT volume in each view plane which were compared to the nominal dimensions in the design file used to 3D print the phantom. The comparison was made against the phantom's design file instead of the micro-CT volume because the 3D printer manufacturing tolerance of 50 μm was less than the resolution and segmentation error of the micro-CT images. The design dimensions and the corresponding measurement obtained from the OCT volume are reported in Table 3. The dimensions derived from the corrected OCT images agree with the design dimensions to better than the 40 μm lateral resolution near the central FOV of the system indicating that scan conversion and lateral distortion correction have effectively eliminated technical sources of geometric inaccuracy within the region of interest. The mean and standard deviation given in Table 3 are calculated across multiple independent attempts by a single assessor to identify dimensional boundaries in the OCT images to calculate the measured dimensions.

Table 3: Measured dimensions in the co-registered OCT images of the phantom targe compared against nominal dimensions taken from the phantom's design file. RMSE: root mean squared error.

Measurement	Image plane	Nominal dimension (*mm*)	Mean (*mm*)	Standard deviation (*μm*)	RMSE (*μm*)
Diameter	Axial (Fig. 53b)	8	7.988	31	33
Staircase step	Coronal (Fig. 53c)	1	1.019	12	22
Staircase step	Sagittal (Fig. 53d)	1	1.002	12	11
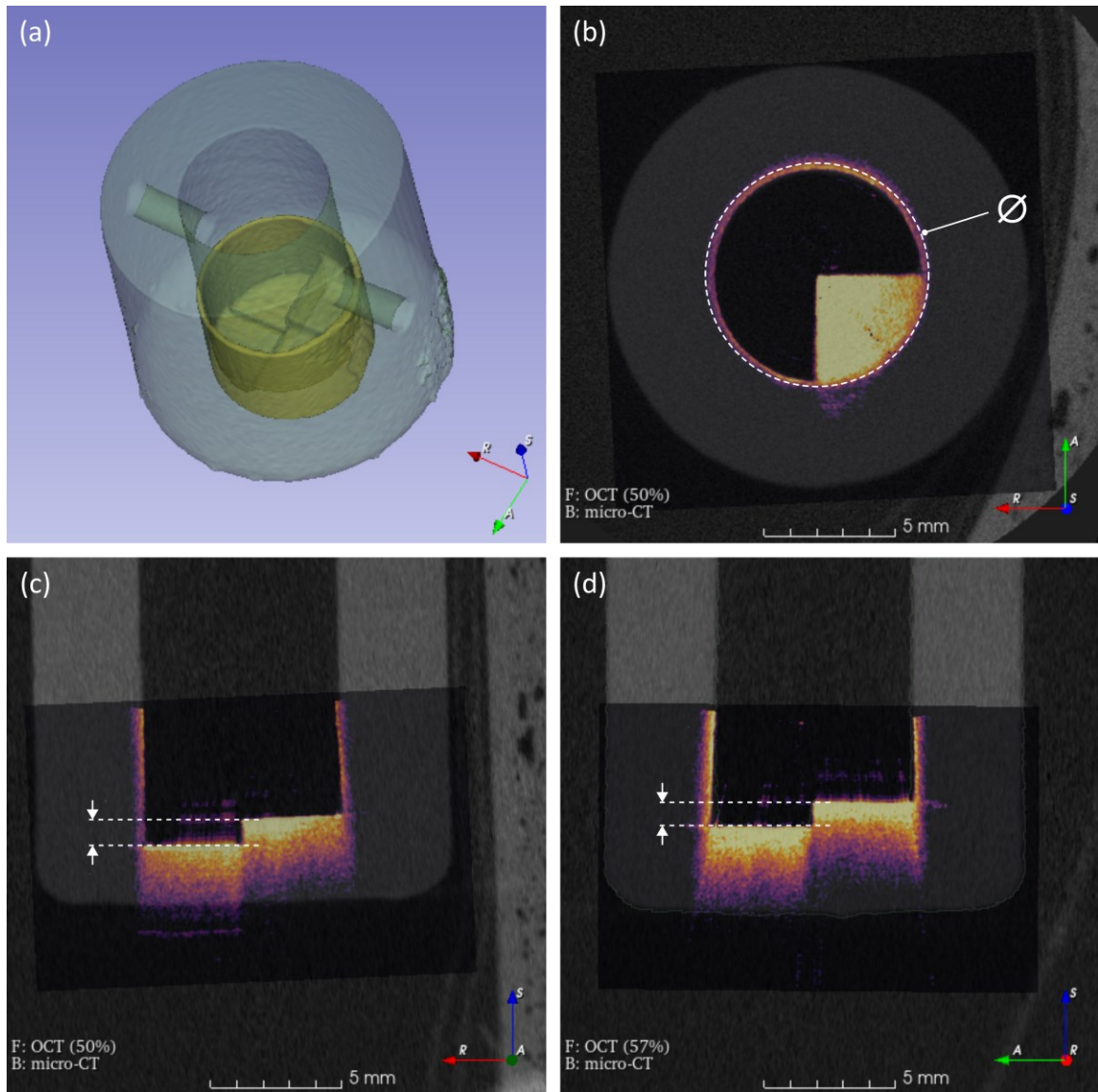
Fig. 53: Micro-CT (greyscale) and OCT (color overlay) of a 3D printed phantom target viewed in 3DSlicer. (a) 3D render of micro-CT (green) and OCT (yellow) co-registered volume segmented contours. (b), (c), and (d) show three orthogonal cross-sectional views (labeled axial, coronal and sagittal) of the co-registered volume shown in (a). White arrows and outlines are the dimensional measurements in Table 3 made to assess the geometric accuracy of the OCT volume. A flythrough video of the co-registered volume's image stack, seen from the coronal plane (c), can be found in Supplementary File F.

5.6.2 Validation of Geometrical Accuracy in a Cadaveric Middle Ear

The ability of our scan conversion and lateral distortion correction procedure to produce geometrically accurate OCT volumetric images of the middle ear was validated by comparison of OCT and micro-CT volumetric images in the temporal bone preparation described in Section 5.5.5. Fig. 54 shows slices through the OCT volume overlaid on corresponding micro-CT slices. Four anatomical dimensions, measured from the ISJ to the umbo, the ISJ to the promontory, the umbo of the malleus to the promontory, and the ISJ

to the TM along the lateral-medial axis were measured on the OCT and micro-CT volumes separately as shown in Fig. 54c and are compared in Table 4. These dimensions are clinically relevant because they quantify the capacity of the middle ear to accommodate middle ear micro manipulations during surgery such as prothesis placement. Each measurement was taken independently five times in each modality by the same assessor. The mean and standard deviation reported in Table 4 were calculated across these five independent measurements on the OCT dataset. The RMSE was calculated between OCT and micro-CT datasets across the five independent measurements. The comparison shows that scan-converted and distortion-corrected OCT can achieve geometric accuracy exceeding the micro-CT resolution of 155 μm. The lack of ISJ to TM measurement for micro-CT is due to the very poor soft tissue contrast of micro-CT which made it impossible to accurately identify the TM surface in the micro-CT data. By comparison, OCT has excellent soft tissue contrast that makes the TM highly visible in the fused images. A flythrough video of the fused OCT and micro-CT image stack is included as supplementary material (Supplementary File G).

Table 4: Clinically relevant dimensions obtained from measurements on the micro-CT and OCT volumes of the cadaveric middle ear. ISJ, incudostapedial joint; TM, tympanic membrane; RMSE, root mean squared error.

Measurement	ISJ to Umbo		ISJ to Promontory		Umbo to Promontory		ISJ to TM	
Modality	OCT	Micro-CT	OCT	Micro-CT	OCT	Micro-CT	OCT	Micro-CT
Mean (mm)	2.813	2.798	2.007	1.999	1.610	1.577	1.869	NA
Standard deviation (μm)	36	51	44	64	38	79	27	NA
RMSE (μm)	60		52		75		NA	
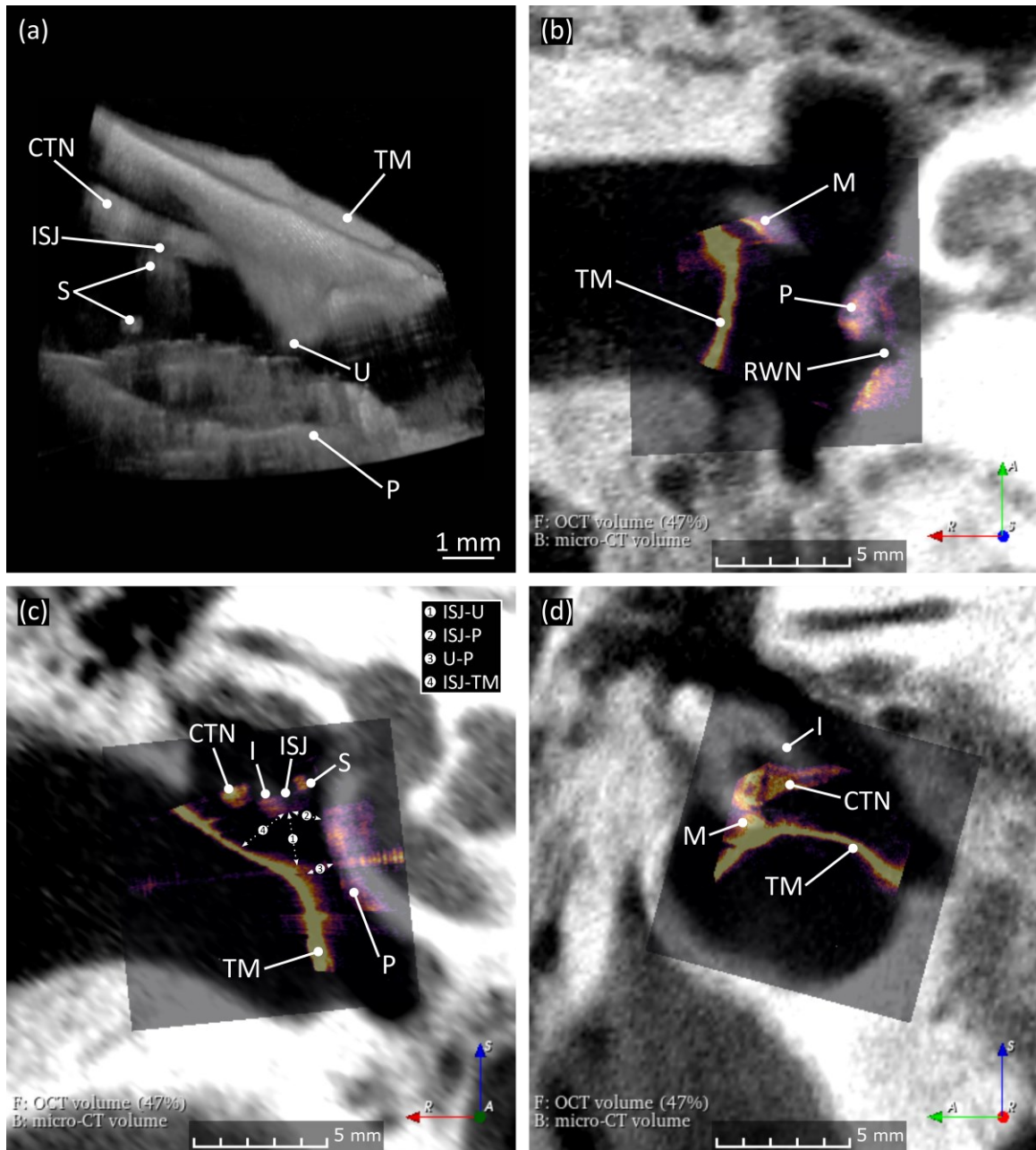
Fig. 54: Selected slices from a fused micro-CT (greyscale) and OCT (color overlay) volume of a cadaveric human middle ear. (a) 3D render of the OCT volume. (b) Axial plane view. (c) Coronal plane view with measurement locations 1–4 marked. (d) Sagittal plane view. CTN, chorda tympani nerve; TM, tympanic membrane; ISJ, incudostapedial joint; S, stapes; P, promontory; I, incus; U, umbo; M, malleus. Measured dimensions: (1) ISJ to umbo, (2) ISJ to promontory, (3) Umbo to promontory, and (4) ISJ to TM. A flythrough video of the co-registered volume's image stack, seen from the coronal plane (c), can be found in Supplementary File G.

The ability of corrected OCT to produce geometrically accurate measurements was further validated using a second cadaveric right middle ear from a 68-year-old male. The facial recess in the bone was opened through a mastoidectomy, the incus was removed and a 2.5 mm long Dresden clip PORP (Kurz, Germany) with a 2.6 *mm* diameter head was

implanted between the stapes superstructure and a cartilage graft placed against the medial side of the TM, simulating an ossiculoplasty, a common surgical treatment for conductive hearing loss. The cartilage was optically cleared using the technique described in [24] and imaged with OCT. Fig. 55 shows an OCT volume of the prepared ear imaged transtympanically following scan conversion and lateral distortion correction. The PORP's diameter as measured in the OCT volume using the distance measurement tool in 3DSlicer [152] was found to be 2.61 mm $\pm$ 34 μm, consistent with the specified prosthesis head diameter of 2.6 mm to within the system's optical resolution. OCT imaging of prostheses is of interest to verify coupling and placement of the prosthesis following middle ear surgery, especially in cases where hearing results are suboptimal [15].
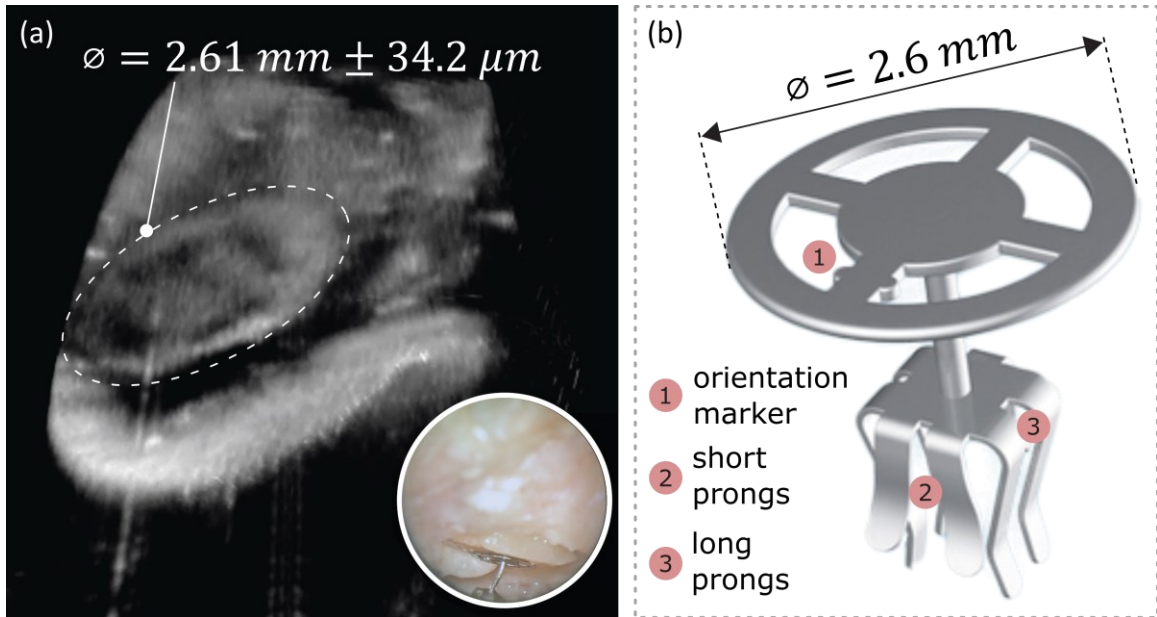
Fig. 55: Cadaveric middle ear containing a 2.6 mm diameter Dresden clip PORP crimped onto the stapes head in a simulated ossiculoplasty. (a) OCT volume of the PORP obtained transtympanically through an optically cleared cartilage graft. Annotations show the measured PORP diameter mean and standard deviation (b) Manufacturer's diagram of the PORP specifying the nominal head diameter. PORP, partial ossicular replacement prosthesis.

To validate the ability of scan conversion and lateral geometric correction to be applied to 4D OCT imaging in real-time, the right middle ear of a healthy 44-year-old male volunteer was imaged while the volunteer performed a Valsalva maneuver. The Valsalva maneuver consists of pinching the nose and attempting to exhale through the nose to pressurize the middle ear space through the Eustachian tube. Patients are often asked to perform the Valsalva maneuver during clinical ear examinations to qualitatively assess tympanic membrane compliance and middle ear ventilation. 4D datasets of the Valsalva

maneuver were acquired at full, half and quarter FOVs corresponding to three different volume rates. Videos of the acquired 4D datasets (Supplementary File H) can be found in the supplementary materials. In all three acquisitions, the three-dimensional displacement of the TM caused by the Valsalva maneuver is clearly visualized in real time.

5.7 Discussion

We have introduced a novel system for live, continuous, and geometrically accurate 3D ME-OCT imaging at the point-of-care using the novel DC-SC spiral scanning technique, real-time volumetric scan conversion and real-time lateral angular distortion correction. DC-SC offers a promising method for scanning over large volumes like the middle ear, by making efficient use of available scanning bandwidth and enabling higher volume rates than is possible with sequential B-mode scanning. The imaging technique is computationally efficient and on standard hardware can run many times faster than data is acquired in even the fastest reported 4D OCT systems [188].

For our system, the combination of scan conversion and 2D lateral distortion correction was adequate to achieve resolution-limited geometric accuracy over 92% of a $30° \times 30° \times 10.9$ mm FOV. Geometric distortion along the edges of the FOV arises due to the difficulty in correctly segmenting calibration grid points in the presence of optical astigmatism. Methods for correcting astigmatism in OCT volumes have previously been reported [205–207] and their combination with the techniques of this study could potentially improve image accuracy at the edges of the FOV. Accuracy was also low in the central few pixels of our image due to mirror bandwidth limitation and central artefacts from internal reflections in our handpiece optics. The bandwidth-related artefact could be improved by slowing down the mirror scan at the expense of a reduced volume rate and the artefact from internal reflections could likely be reduced through optical design improvements in which case the center of the image volume should have the same accuracy as other nearby locations.

Through co-registration of OCT and micro-CT volumes we have demonstrated that the geometric accuracy of our corrected OCT volumes for measuring anatomical distances is $52 \, \mu$m to $75 \, \mu$m. This estimate of OCT's accuracy is likely an underestimate since the $155 \, \mu$m resolution of the micro-CT system likely contributed significantly to the measurement RMSE between OCT and micro-CT images. Even a $52 \, \mu m$ to $75 \, \mu m$

accuracy is substantially better than the resolution of typical clinical temporal bone CT ($400\ \mu m$) or MRI ($300\ \mu m$) systems [12] and similar to the resolution of binocular surgical stereomicroscopes ($\sim 35\ \mu m$) [208]. The limited FOV of the ME-OCT and its inability to see through bone, means that OCT cannot replace clinical CT for middle ear visualization. However, this study shows that OCT can offer images of the middle ear areas visible through the TM with excellent geometric fidelity. We also note that the techniques we used to compare OCT and micro-CT could also be applied to fuse corrected OCT images with clinical CT images. The higher resolution and soft tissue contrast of OCT could potentially enhance CT images or reduce the number of CT images needed to track longitudinal changes to the middle ear over time.

Geometrically accurate 4D OCT imaging is an appealing capability for the middle ear whose complex 3D structure cannot be fully appreciated with 2D B-mode cross-sectional imaging alone. In addition to clinical diagnostics, there may be a role for 4D ME-OCT in surgical guidance including robotic surgeries [209–211]. Surgical guidance applications have been the primary motivator of recent developments in ophthalmic 4D OCT imaging technology [188].

Using DC-SC, we have demonstrated applying corrected 4D ME-OCT in vivo for visualization of conformational changes in a healthy normal ear during a Valsalva maneuver at three different FOVs (Supplementary File H), achieving volume rates of 0.22 vol/s over the full $30^{\circ} \times 30^{\circ} \times 10.9\ mm$ FOV and 3.56 vol/s over a limited $7.5^{\circ} \times 7.5^{\circ} \times 10.9\ mm$ FOV. The achievable volume rate with DC-SC is limited by the bandwidth of our scanning mirror. In our system design we use a large mirror with a large scanning range to achieve a large number of resolvable spots at the expense of scanning rate, but alternative designs could employ a smaller mirror or reduced scan range to achieve higher volume rates. More advanced driving waveforms could also potentially increase volume rates by operating the MEMS mirror closer to its resonance frequency ($450\ Hz$ for the A8L1.1). The processing volume rate of $52.8\ Hz$ that we demonstrated for our technique provides substantial headroom to support faster systems than the one presented here.

5.8 Conclusion

We have demonstrated techniques for real-time spiral scanning, scan conversion and correction of lateral distortion and crosstalk from a MEMS scanning mirror in a handheld, entocentric OCT imaging system designed for imaging the middle ear. Analysis of residual errors in corrected OCT images of grid targets, 3D phantom targets and cadaveric temporal bones consistently show geometric accuracy to be limited only by lateral resolution over most of the FOV. The techniques presented are simple and of low computational complexity, making them suitable for application in real-time 4D imaging. We demonstrate this 4D capability by performing real-time volumetric imaging on the ear of a live subject during a dynamic pressurization maneuver.

Having geometrically accurate images is important for many applications in surgical planning, multimodal image fusion, objective diagnostics, and longitudinal monitoring of disease progression. The techniques presented in this study may serve to make OCT a competitive imaging modality for many of these applications.

Chapter 6

Real-time In-vivo Volumetric Middle Ear Optical Coherence Tomography Angiography

Joshua D. Farrell[1], Junzhe Wang[1], Floor Couvreur[2,3], Nael Shoman[2], David P Morris[2], Robert B.A. Adamson[1,4]

[1]School of Biomedical Engineering, Dalhousie University, NS B3H 4R2, Canada
[2]Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, Dalhousie University, Halifax, NS B3H 4R2, Canada
[3]Department of Otorhinolaryngology, Head and Neck Surgery, AZ Sint-Jan's Hospital, Ruddershove 10, 8000 Bruges, Belgium
[4]Electrical and Computer Engineering Department, Dalhousie University, Halifax, NS B3H 4R2, Canada

6.1   Author Contribution Statement

J. Farrell, J. Wang, and R. Adamson collectively contributed to the design, methods, and experiments of this study. R. Adamson provided additional support in helping draft the manuscript, providing relevant clinical context to the study, and a remarkably steady hand for in-vivo data collection efforts. J. Wang was instrumental in setting up experiments, providing technical support for setting up the software for in-vivo data collection, and further software support to help offline analysis of collected data. F. Couvreur, N. Shoman, and D. Morris provided key clinical insight into the clinical relevancy of this work and provided constructive feedback around results discussion helping shape the manuscript draft. J. Farrell was responsible for conducting the experiments, developing the methodologies and real-time software required for the experiments, data analysis and renders, and served as the lead author in making sure this manuscript draft is completed and submitted for peer-review.

6.2 Preamble

This chapter of my thesis is based on a manuscript being prepared for submission to the Journal of Biomedical Optics (JBO). Some sections of the manuscript have been omitted or revised to avoid repetition of any previously discussed material found in Chapter 1 and Chapter 2. Additionally, minor adjustments were made to ensure consistency between the material presented in this chapter and the rest of the thesis.

6.3 Abstract

We present a point-of-care system for real-time, volumetric OCTA imaging of the human middle ear and demonstrate its use in visualizing the vasculature of the TM and middle ear mucosa in an in-vivo measurement. We also demonstrate the use of OCTA-like phase variance methods for direct visualization of diagnostically relevant dynamic middle ear processes through real-time imaging of the contralateral stapedius reflex.

6.4 Introduction

Over the past 30 years, OCT has become a standard-of-care diagnostic technology in ophthalmology, providing real-time structural measurements of the retina [2] and anterior segments of the eye. OCTA is an important functional extension of OCT that is now widely used for imaging retinal vasculature [113,114], particularly in the diagnosis of diabetic retinopathy and glaucoma [115–117]. Recall from Section 2.6 that OCTA works by taking a sequential time series of images and using the speckle variance of the images to obtain contrast on vasculature. Because blood vessels contain moving erythrocytes, vasculature exhibits more rapid changes in OCT speckle than surrounding tissue, allowing vessels to be visualized with high contrast against background tissues.

While technical developments around ME-OCT have focused on structural and vibrometric imaging, one study also applied amplitude variance OCTA post-processing to middle ear images and successfully demonstrated visualization of blood vessels in the TM, cochlear promontory and cochlear duct [19]. The middle ear OCTA image presented in that study was a depth-resolved 2D image in which the vessels appear as bright points in the variance image. In ophthalmic OCT, much of the diagnostic value of angiography was unlocked when OCTA was combined with rapid 3D OCT imaging techniques that enabled

the visualization of vascular networks presented as 2D enface angiograms or 3D volumetric renders of vasculature.

In this chapter, we apply phase variance OCTA with bulk motion phase noise correction to achieve real-time, in-vivo, 3D OCTA visualization of the vascular networks in the TM and middle ear for the first time. We also demonstrate the application of OCTA techniques for direct visualization of the contraction of the stapedius muscle during activation of the contralateral stapedius reflex whose presence or absence is an important diagnostic indicator of middle ear and acoustic nerve pathology [212].

6.4.1 OCTA for the Evaluation of Middle Ear Pathologies

The middle ear and medial side of the TM are lined by well-vascularized mucosa which is responsible for cleaning, immunological and wound healing responses, and aeration of the mastoid and middle ear spaces [29,213]. As discussed in Section 2.3, many common middle ear pathologies are associated with disruption of this mucosal function leading to complications such as chronic otitis media (COM) or ossicular necrosis from inadequate vascularization [213]. This makes the vascularization of the mucosa a potentially useful diagnostic signal that can currently only be evaluated by qualitative assessment of the colour of the TM. Adequate vascularization of the TM is also an important factor in determining the likelihood of TM grafts to integrate effectively, making visualization of the TM vasculature potentially useful for surgical planning for myringoplasty and tympanoplasty procedures [214–216]. Finally, the techniques of OCTA can be used to detect motions other than those induced by blood flow. Reflexive contraction of stapedius muscle in response to contralateral sound stimulation is a very widely used diagnostic for assessing sensorineural hearing loss, Section 2.2.2, and in identifying ossicular fixation [46,217]. Though not examined as part of this chapter, micromotions of middle ear structures are also associated with other pathologies including patulous Eustachian tube, middle ear myoclonus, and various dehiscences. The diagnosis of these pathologies might also benefit from the techniques we present.

6.5 Methods

Imaging was performed using the ME-OCT system described in Section 1.1.1. A commercial color video otoscope (Welch Allyn Macroview) was used to capture otoscopic

images of the ear. The rendering engine and software architecture described in Chapter 3 and Chapter 4 respectively were extended to facilitate real-time processing and display of phase variance ME-OCTA data with bulk motion correction.

For ME-OCTA imaging, the handpiece's speculum was inserted into a subject's external ear canal and a set of N sequential B-mode frames were collected at 20 FPS for each generated angiography image frame where N was a software-configurable parameter. This resulted in an effective OCTA acquisition rate of $20/N$ FPS. Each image line within each B-mode image was formed by complex averaging of 5 A-lines. For volumetric angiographic imaging, a set of frames at 512 different vertical slice angles in the image volume were generated.

6.5.1 Phase-Variance Angiography

We constructed 2D angiographic images at a set of positions along the slow lateral scanning direction y, by taking a set of sequentially acquired complex image frames $s_i(x, z)$, where x, z represent the fast lateral (horizontal) scanning direction and axial (depth) dimension respectively and i represents the frame index in time. The phase difference between sequentially acquired frames is then calculated as:

$$\Delta\varphi_i(x, z) = \arg[s_{i+1}(x, z)] - \arg[s_i(x, z)] \qquad (16)$$

where arg [∗] is the angle function.

To provide contrast on locally moving structures, a correction must be made for the "bulk motion" of the local structure. Recall from Section 2.6.1, that bulk motion is defined as the common-mode motion of the pixels within each A-line. In ME-OCT, bulk motion arises due to hand motion on the part of the operator holding the probe and to physiological motion of the patient such as that due to breathing. In ophthalmic OCTA, the bulk motion is typically calculated across the entire depth of the each image line as the bulk motion of the retina is assumed to be independent of depth [122,218]. However, this assumption does not hold for the middle ear as each image line can contain multiple independent structures, each with its own unique bulk motion. For example, the bulk motion of the tympanic membrane may be different from the bulk motion of the cochlear promontory since forces exerted on the ear canal wall by the probe may cause motion of the TM but not of the

promontory. Thus, bulk motion must be calculated separately for each independent structure contained within the image line.

Common methods for bulk motion correction in ophthalmic OCTA include intensity weighted phase averaging [122,123], histogram estimation [125,219], and edge detection [120,121]. However, none of these methods accurately estimates the bulk motion when the phase step between frames is large [120,220] as we have found it to be in our datasets. The application of principal component analysis (PCA), which separates blood flow from static tissue using differences in spatio-temporal correlations, was also found to be easily skewed and sensitive to this bulk motion as well [221]. Robust PCA (RPCA) was demonstrated by Wang et al. to overcome this bulk motion sensitivity, however, further development and optimization is needed to achieve real-time processing rates [221].

To compensate for large interframe phase changes from bulk motion, we developed a variant of weighted phase averaging [122,123] in which we calculate an amplitude-weighted sum over the phasors $exp[j\Delta\varphi_i(x,k)]$ whose phase angle is the interframe phase difference. The phase is extracted for the resultant phasor calculated within a sliding window of length K pixels according to:

$$\Delta\varphi_{i,bulk}(x,z) = \arg\left\{ \sum_{k=z-(K-1)/2}^{z+(K-1)/2} |s_i(x,k)| \exp[j\Delta\varphi_i(x,k)] \right\} \qquad (17)$$

The value of K defines a fixed-width sliding window that is chosen to be larger than the depth extent in pixels of the smallest independently moving structure but smaller than the space between independently moving structures. For convenience, we assume that Z is an odd number. Zero-padding at the beginning and end of the A-line allows $\Delta\varphi_{i,bulk}(x,z)$ to extend over the full depth. In this study, a window size of $K = 21$ was chosen which is large enough to capture the bulk motion of the tympanic membrane, the thinnest structure of interest within our images, but not so large as to mix TM bulk motion with the bulk motion of the boney structures of the middle ear. This phasor-based windowed bulk motion compensation approach offers a simple and computationally efficient method of separating the bulk motion for different structures in depth without requiring phase-unwrapping.

The bulk motion corrected interframe phase difference is then calculated as:

$$\Delta\varphi_{i,corr}(x,z) = \Delta\varphi_i(x,z) - \Delta\varphi_{i,bulk}(x,z) \qquad (18)$$

Following bulk motion correction, $\Delta\varphi_{i,corr}(x,z)$ is phase-wrapped to limit phase changes to $(-\pi, \pi]$ [122] to avoid phase step artefacts.

Finally, the phase-variance angiographic image $\varphi_{phase}{}^2(x,z)$ is constructed as the average of the squared phase deviation between each corrected phase difference frame $\Delta\varphi_{i,corr}(x,z)$, $i \in [1, N-1]$ and the mean of the phase difference frames $\overline{\Delta\varphi_{corr}(x,z)}$ defined as:

$$\overline{\Delta\varphi_{corr}(x,z)} = \frac{1}{N-1}\sum_{i=1}^{N-1}[\Delta\varphi_{i,corr}(x,z)] \qquad (19)$$

$$\varphi_{phase}{}^2(x,z) = \frac{1}{N-1}\sum_{i=1}^{N-1}[\Delta\varphi_{i,corr}(x,z) - \overline{\Delta\varphi_{corr}(x,z)}]^2 \qquad (20)$$

$\varphi_{phase}{}^2(x,z)$ is the quantity plotted in the angiographic images shown in the next section. For the imaging results presented in this study, we selected $N = 4$ sequential image frames which provides a good tradeoff between angiographic image quality and imaging speed.

6.6 Results

Fig. 56 shows an *enface* structural image (Fig. 56c) and corresponding OCT angiogram (Fig. 56d) of the TM generated from a ME-OCTA volume (Fig. 56b) of a healthy volunteer's left ear collected in-vivo. The frustrum-shaped volume contains $512 \times 512 \times 330$ voxels covering a ROI of $\sim 30° \times 30° \times 10.9$ mm and took 51 seconds to acquire. Fig. 56a presents the corresponding otoscopic view of the subject's TM with the OCTA volume's ROI highlighted by the dotted circle. *En-face* renders were produced by first identifying and segmenting the region of the volume corresponding to the TM. Maximum intensity projection of the segmented sub-volume was then used to construct the *en-face* structural (Fig. 56c) image while median intensity projection was used to construct the *en-face* phase-variance (Fig. 56d) angiogram. Contrast-limited adaptive histogram equalization (CLAHE) [222] was applied to the angiographic data to normalize its display

range before overlaying it onto the structural image of Fig. 56c to produce the final angiogram shown in Fig. 56d.
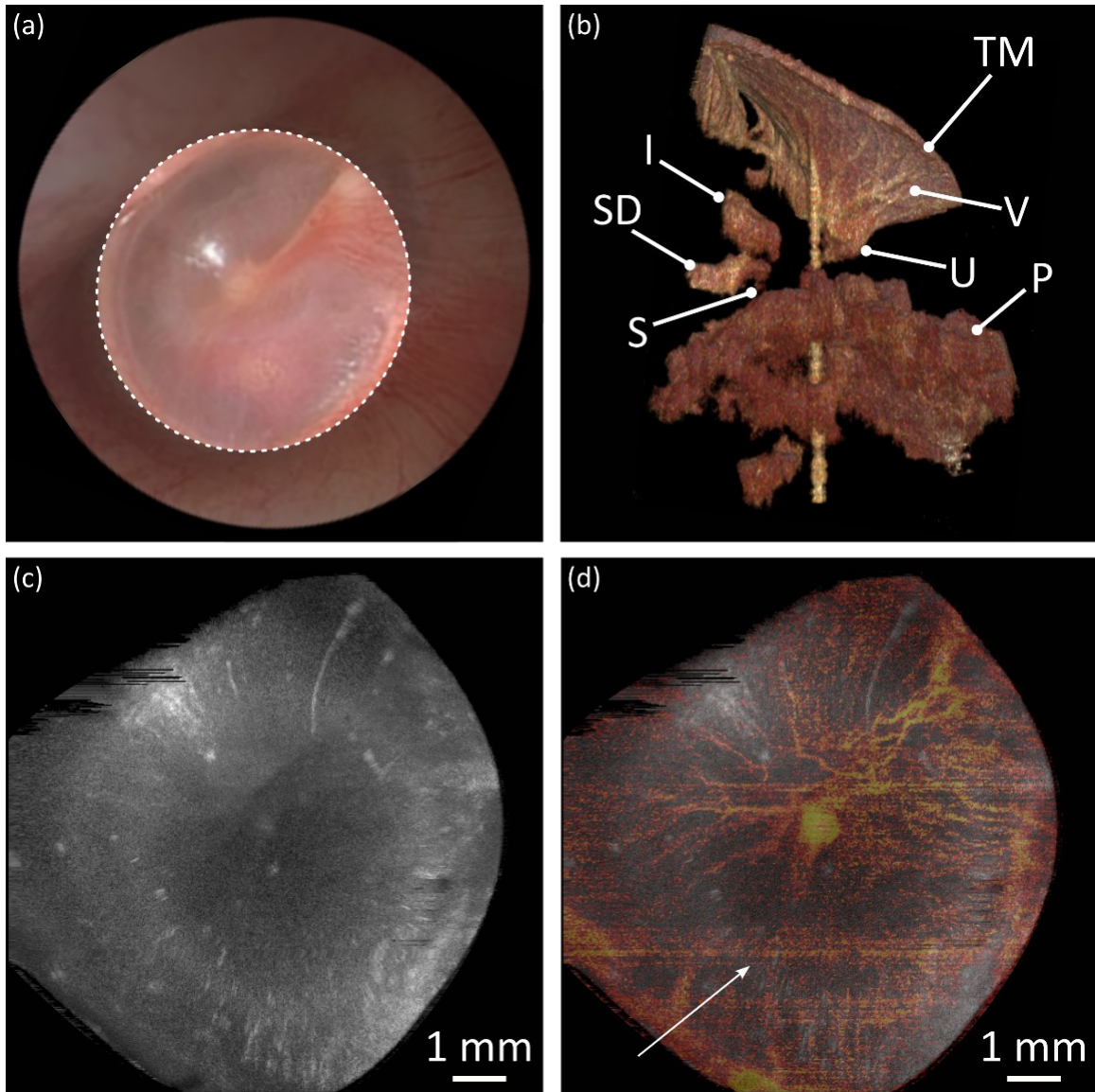
Fig. 56:   En-face structural image and corresponding angiogram of the TM generated from a ME-OCTA volume of a healthy volunteer acquired in-vivo. (a) Otoscopic view of the healthy adult volunteer's TM. OCT volume ROI highlighted by dotted circle. (b) Volumetric render of the ME-OCTA volume used to generate the en-face images in (c) and (d). (c) En-face maximum intensity projection structural image of the TM. (d) En-face phase-variance angiogram of the TM generated using median intensity projection. White arrow denotes residual bulk motion artefacts after correction. TM, tympanic membrane; U, umbo; I, incus; SD, stapedius; S, stapes; P, promontory; V, vasculature.

Compared to the structural image of Fig. 56c and to the endoscopic photograph of Fig. 56a, the phase-variance angiogram of Fig. 56d provides higher contrast and enables depth-resolved visualization of more of the tympanic vasculature [223]. As compared to surgical microscopes, where high-resolution images of surface vasculature can be obtained,

phase-variance angiograms such as Fig. 56d can provide additional vasculature information on the underside of the TM or the layers of the TM itself. For example, as discussed in Section 2.3 for the case of OM, phase-variance angiograms has the potential to visualize biofilm formation present on the underside of the TM, thickening of the TM or the formation of granulation tissue all indicating the progression of AOM to CSOM earlier as compared to traditional otomicroscopy. In Fig. 56d some residual artefacts of bulk motion remain even after correction (white arrow). In both Fig. 56c and Fig. 56d the upper left quadrant of the TM is slightly obscured by the anterior scutum of the volunteer's ear canal (see Fig. 56a) which blocks the OCT sample beam in this region during imaging.
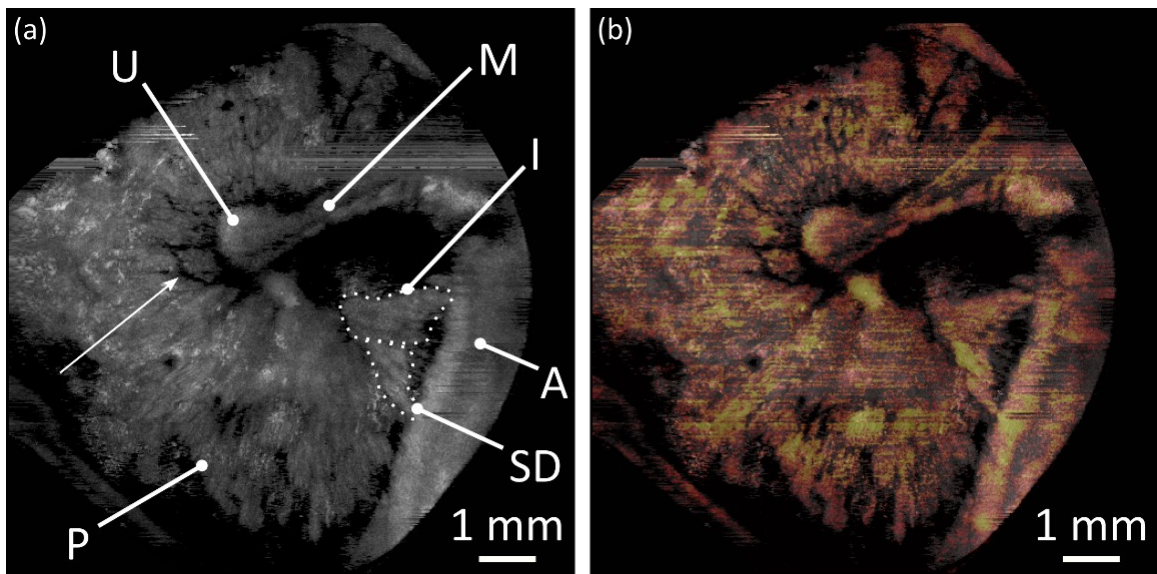
Fig. 57: En-face structural image and corresponding angiogram of the middle ear space with the TM removed generated from a ME-OCTA volume of a healthy volunteer acquired in-vivo. (a) En-face maximum intensity projection structural image of the middle ear space. Incus and stapedius ROI highlighted by dotted paths. White arrow denotes shadow artefacts caused by the vasculature of the TM. (b) En-face phase variance angiogram generated using median intensity projection. U, umbo; M, malleus; I, incus; SD, stapedius; A, annulus; P, promontory.

Fig. 57 shows an *enface* structural image and corresponding angiogram of the middle ear space with digital tympanotomy applied generated from the same OCTA volume (Fig. 56c) used for the *enface* TM images in Fig. 56. All the *enface* images and angiograms in Fig. 57 were generated in the same way as their Fig. 56 counterparts. The angiogram of the middle ear space, Fig. 57b, shows the vascularization of the malleus arm, incus, stapedius, and annulus as well as areas of high blood flow on the mucosal lining of the promontory. Shadow artefacts can be seen around the umbo in Fig. 57a (white arrow) caused by the vasculature of the TM [224]. The vasculature in the promontory is not as

evident as that observed in the ossicles or TM possibly due to a combination of shadowing from the overlying TM and a lower signal-to-noise ratio (SNR) at this imaging depth within the volume.

As evidence that not all the observed contrast is due to vasculature, we found that there was significant angiographic signal deep in the cortical bone in the promontory. While vasculature is expected in the mucosa at the surface of the promontory, the promontory bone should be essentially avascular in a healthy ear. The presence of angiographic signal at depth in the promontory therefore points to the distal parts of the OCTA signal being dominated by phase noise, at least over some of the field of view.

To confirm that OCT is capable of detecting vasculature on the mucosa of the promontory we repeated scanning in the same ear with of $N = 8$ repeated frames. Fig. 58 shows a cross-sectional slice through several adjacent frames of this new volume. The phase noise in this image was significantly reduced as compared to the slices from the original acquisition. There appears to be a strong angiographic signal present at the locations highlighted by the white arrows that is persistent across adjacent slices.
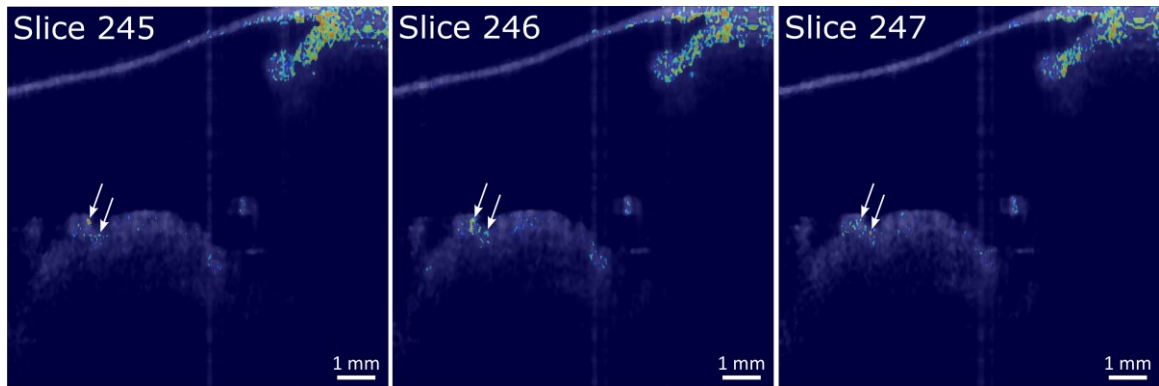
Fig. 58: Adjacent cross-sectional ME-OCTA frames acquired in-vivo from a healthy adult volunteer using an interframe repetition of N = 8. The white arrows show likely blood vessels on the surface of the promontory.
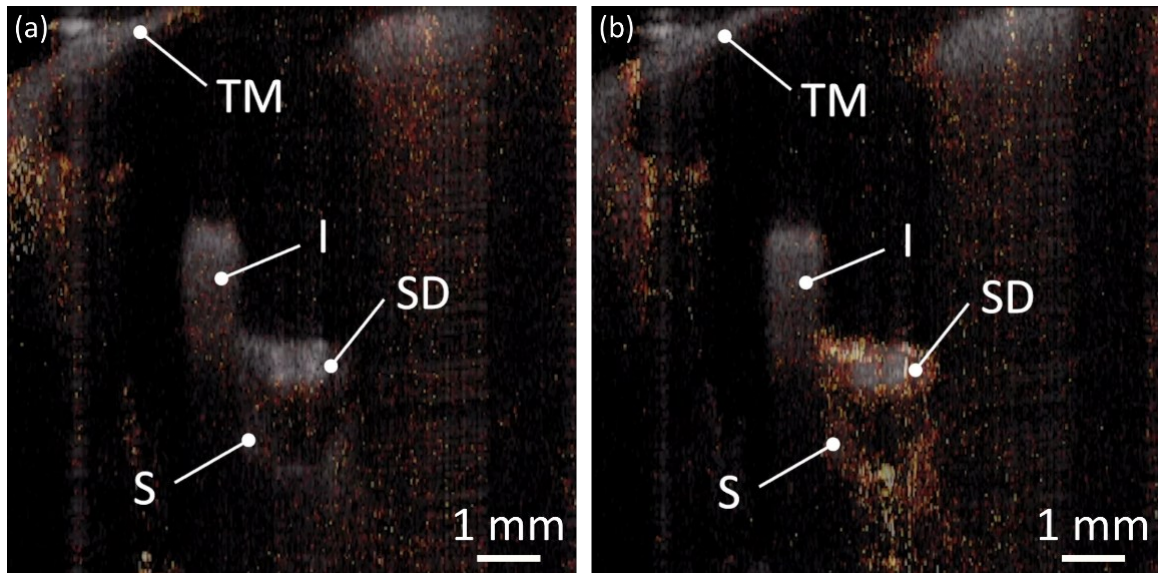
Fig. 59: Real-time, in-vivo, cross-sectional 2D angiography images of a healthy volunteer's right ear taken while a 100 dB-SPL 1-second 1 kHz tone is played in the contralateral ear to trigger their stapedius reflex. (a) Angiographic image of the incus, stapes, and stapedius before the tone is applied. (b) Angiographic image of the incus, stapes and stapedius after the tone is applied. TM, tympanic membrane; I, incus; S, stapes; SD, stapedius. A video of the angiography imaging session can be found in Supplementary File I.

Fig. 59 shows 2D cross-sectional angiography images of the stapedius tendon, stapes and incus during a stapedius muscle contraction caused by the stapedius reflex acquired in-vivo, in real-time, from the right ear of the same healthy volunteer whose TM is shown in Fig. 56. To trigger the reflex, a 100 dB-SPL 1-second 1 kHz tone was played every four seconds into the subject's left (contralateral) ear through a headphone while performing 2D phase-variance angiography imaging of the subject's right ear. The stapedius reflex occurs when a loud sound in either ear triggers a contraction of the stapedius muscles in both ears, pulling on the stapedius tendon attached to the posterior crus of the stapes. Supplementary File I shows that the stapedius tendon produces a strong phase-variance signal at the onset of the tone burst (Fig. 59b) and no phase-variance signal between tone bursts (Fig. 59a) consistent with the tone activating the stapedius reflex. The contraction of the stapedius muscle can be seen to detectable motions of the stapedius tendon and posterior crus of the stapes, but not the incus. This is consistent with previous studies of the biomechanics of the stapedius reflex in cats which found that contraction of the stapedius muscle translates the head of the stapes but leaves the incus stationary [225].

6.7 Discussion

Obtaining clear visualization of vasculature in the ear requires that a different bulk motion factor $\Delta\phi_{i,bulk}$ be applied to each independently moving structure along the A-line depth. In this study this approach to bulk motion correction was implemented using the window-based approach described by Eq. (17). Applying a single bulk motion factor $\Delta\phi_{i,bulk}$ to the whole image line (i.e. by setting K to be the length of the whole A-line) resulted in the images of Fig. 60a (*en-face*) and Fig. 60b (cross-section). In these images, regions where bulk motion is out of phase between the TM and the promontory (Fig. 60c) result in OCTA artefacts in one or both structures. The application of the depth-dependent localized bulk motion correction of Eq. (17) significantly reduces these artefacts as shown in Fig. 60d.
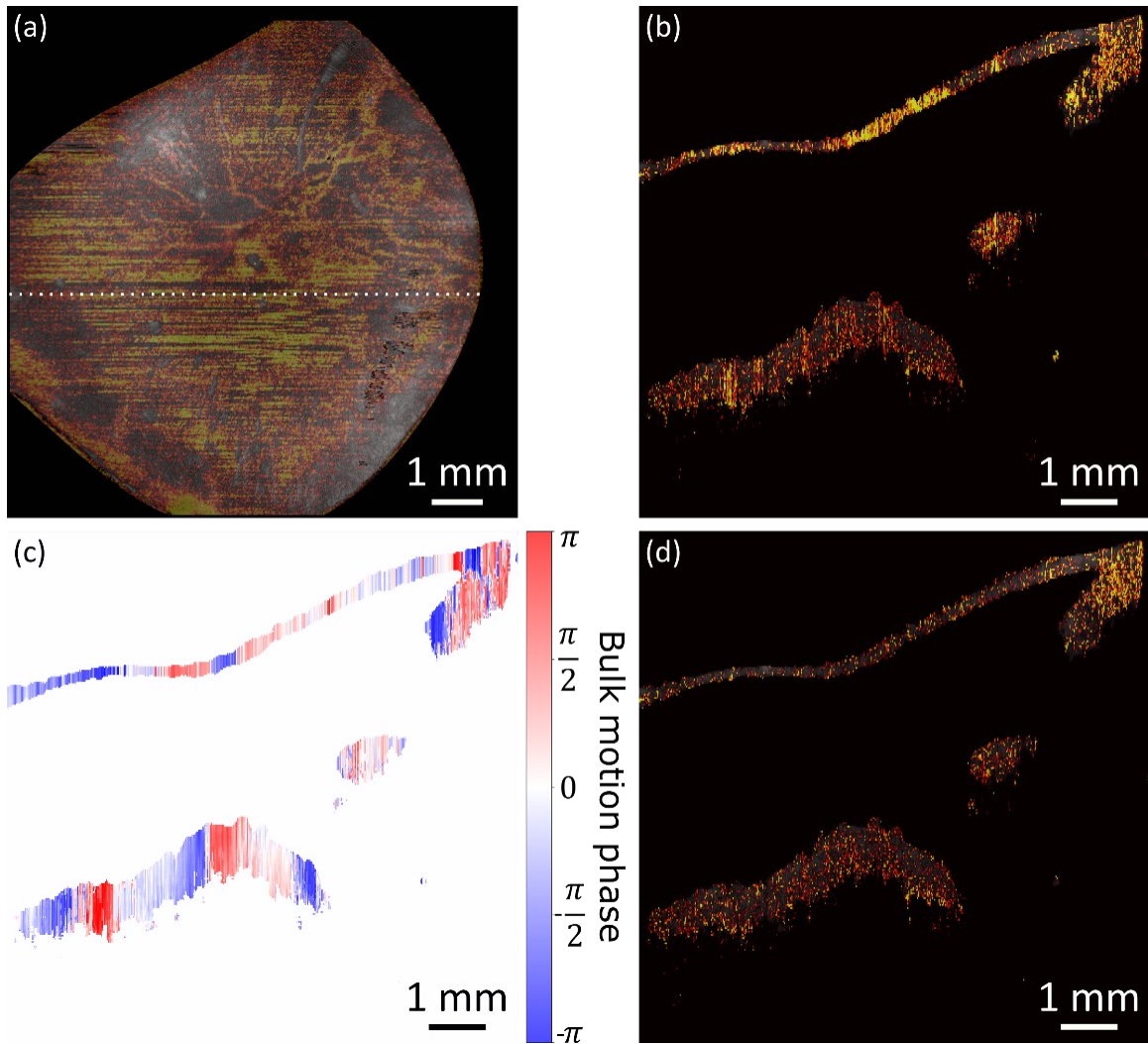
Fig. 60: Bulk motion artefact analysis. (a) En-face phase-variance angiogram of the TM generated from a ME-OCTA volume of a healthy adult volunteer acquired in-vivo exhibiting the bulk motion artefact. (b) Cross-sectional OCTA image frame, dotted line in (a), exhibiting the artefact. (c) Bulk motion correction factor for each pixel calculated within a local window K = 21 overlaid onto (b). (d) Cross-sectional OCTA image frame calculated using the bulk motion correction factors in (c).

The differential bulk motion between the TM and underlying structures may arise, for example, from changes to the quasi-static ear canal pressure or to changes in stresses on the ear canal wall which affect the TM but not the promontory [226,227].

Currently it takes 51 seconds to capture a full ME-OCTA volume consisting of 512 frames covering a FOV of $10\ mm \times 10\ mm$. While it is possible for a trained subject to remain still for this long, the duration of the measurement will make it challenging to obtain high quality images in patients. By comparison, commercially available OCTA ophthalmology systems have demonstrated volumetric acquisition rates of 3.6 seconds covering a FOV of $2.4\ mm \times 2.4\ mm$ [228]. ME-OCTA volumetric acquisition rates could

be improved by restricting the imaged FOV to a similarly small region of interest. A move from a raster-scanning to a spiral-based scanning approach could also improve the scan rate by making better use of the available bandwidth of the scanning mirror in our system and eliminating dead time when the mirror is scanning but no image lines are being acquired [185]. Utilizing spiral scanning we have previously demonstrated volumetric structural imaging at a rate of 3.55 volumes per second for a FOV of $2.5 \; mm \times 2.5 \; mm$ [111]. Assuming we continue to use $N = 4$ sequentially acquired lines to estimate the phase variance, we could potentially achieve a volumetric OCTA acquisition rate of ~ 1 volume per second over a similar area. Spiral-based OCTA has already been successfully demonstrated for dermatological, gastrointestinal, and ophthalmological applications [187,229,230].

There are numerous potential clinical applications for ME-OCTA imaging that we intend to explore in future work. For example, grafts are often used to repair perforations or retracted segments of the TM, but their integration depends on having adequate vascularization [214]. Visualization of the microvasculature of the TM through OCTA may allow clinicians to estimate local vascularization around a damaged portion of the TM and help decide which portion of the TM needs to be removed to refreshen the edges to ensure good blood supply at the margins of the perforation and promote graft integration. Transtympanic visualization of the vascular supply to the ossicles may also predict the likelihood of developing ossicular necrosis [231]. This may provide an additional source of information during pre-surgical planning on the type of ossicular reconstruction to be considered by surgeons when they are faced with managing a disease. For example, a surgeon may be directed towards a malleo-stapedotomy instead of an incudo-stapedotomy if the vascularization feeding the incus is compromised, since the latter intervention is likely to fail if the necrosis of the incus progresses [232].

The direct visualization of the stapedius reflex that we demonstrated may also have clinical utility in patients with conductive hearing loss, a TM perforation or in other situations where traditional measurement of acoustic reflex is challenging.

6.8   Conclusion

We have demonstrated for the first time, the application of real-time, phase-sensitive 2D and 3D OCTA imaging to the human middle ear in-vivo and investigated its ability to image

the vascular networks of the tympanic membrane, to detect blood vessels in the middle ear mucosa and to directly visualize the stapedius reflex. As middle ear OCT becomes more widespread, OCTA may provide useful diagnostic information to otolaryngologists to help in managing a variety of middle ear diseases.

Chapter 7

Conclusion

With the limitations and imprecision of current pre-surgical middle ear diagnostics and radiological imaging, ear surgeons often face situations where a confident diagnosis can only be reached through exploratory tympanotomy [57]. However, there is a risk that the pathology discovered in exploration is outside the scope of informed consent or not fixable through surgical intervention. For example, with a prevalence of 0.48% [233], a persistent stapedial artery is a rare congenital vascular anomaly which contraindicates for stapes surgery. If encountered intraoperatively, surgery must be terminated without addressing the pathology, at significant cost to the healthcare system and substantial risk and inconvenience to the patient.

ME-OCT's potential as a diagnostic tool for clinical otology presents tremendous opportunity to impact pre-surgical decision making and improve surgical outcomes. Clinical interest and growth in the field of ME-OCT has been significant over recent years with several groups making rapid progress on potential clinical applications [1,9,19,234]. However, to fully realize the technology's clinical potential there is a need for systems-level engineering to deliver usable solutions for clinicians that allow them to image and visualize structural and diagnostic information quickly at the point-of-care and without requiring engineering support to be able to incorporate ME-OCT into their clinical workflow.

Pathologies of the middle ear often manifest over multiple regions within the complex 3D dynamic structure of the middle ear making it difficult to appreciate its extent in 2D cross-sectional images alone. Therefore, for a more accurate diagnosis of pathology, there exists a need to visualize 3D structural and functional imaging data as it is collected at the point-of-care. This need motivated the development of the custom rendering engine capable of displaying acquired volumetric ME-OCT datasets at real-time interactive rates that was the focus of Chapter 3. This work introduced a single-pass volumetric ray-casting technique implemented using OpenGL and CUDA to produce faithful renders for a variety

of ME-OCT datasets. Additionally, two novel visualization techniques built on this engine were demonstrated: digital tympanotomy and Doppler animation. Digital tympanotomy shows promise in producing enface views of the subtympanic middle ear structures like those obtained through exploratory tympanotomy by digitally removing the TM from volumetric ME-OCT datasets collected non-invasively in the clinic. Doppler animation simulates the acoustic response of the middle ear in response to an applied pure tone stimulus using ME-OCT Doppler vibrometry in both 2D cross-sectional and volumetric datasets. This provides an intuitive visualization of the functional mechanics of pathological ears to both clinicians and patients which has the benefit of educating patients on their pathology, potentially improving communication of the pathology to patients and improving informed consent prior to surgery. This work has benefitted numerous research studies lead by our group [1,24,111,166,167] and our collaborators [165] in understanding the 3D dynamics of healthy and pathological middle ears.

A crucial component of a ME-OCT system, or, indeed, of any modern medical imaging system, is the system software, but software systems to support real-time data acquisition, data processing and user interaction for clinical ME-OCT systems have not yet been described in literature and remain underdeveloped. While Chapter 3 focused on the volumetric visualization of ME-OCT datasets, Chapter 4 set out to integrate this work into a complete software architecture suitable for use in the clinic by clinicians at the point-of-care. The goal of the software was to allow clinicians to independently acquire ME-OCT images and to lower the bar for future clinical adoption of the technology. The software architecture we developed provided a turn-key imaging experience that allowed it to be integrated into the clinical workflow without close engineering support and it provided a reliable and extendable framework for later research work. Running on modest PC and GPU hardware, the architecture demonstrated data processing at rates as high as 38.48 Gb/s which was 6 times faster than the minimum required rate for real-time processing.

While qualitative visualization of ME-OCT datasets can provide crucial context to understanding the complex 3D manifestation of middle ear pathologies, in many situations it is required to have geometrically accurate measurements of middle ear structures. Furthermore, to fully appreciate and explore the 3D structure of the middle ear it is desirable to have a system capable of collecting and rendering middle ear volumes at real-

time rates. The work of Chapter 5 [111] extended the software architecture of Chapter 4 to provide geometrically accurate, continuous, live 4D ME-OCT in-vivo imaging.

Using the methods outlined in this work, geometrical inaccuracies of the ME-OCT imaging system were reduced below the system's lateral resolution over 92% of the FOV. Corrected OCT volumes were validated through co-registration with micro-CT volumes of a 3D-printed phantom and a human cadaveric middle ear. The study demonstrated that the geometric accuracy of corrected OCT volumes was between $52\ \mu$m and $75\ \mu$m which exceeded the $155\ \mu$m resolution of the micro-CT system to which the OCT system was compared. Additionally, a novel spiral scanning technique, DC-SC, was used to demonstrate real-time, geometrically accurate, volumetric ME-OCT imaging of in-vivo middle ear dynamics for the first time, achieving volume rates of $0.22\ \text{vol}/s$ over the full $30° \times 30° \times 10.9\ mm$ FOV and $3.56\ \text{vol}/s$ over a reduced $7.5° \times 7.5° \times 10.9\ mm$ FOV.

Many of the most common pathologies arise due to disruption of the mucosal lining of the middle ear space. Therefore, the vascularization of the mucosa is a potentially powerful diagnostic signal that cannot currently be accessed in clinic due to the limitations of current pre-surgical non-invasive imaging methods. By calculating the phase variance across sequentially acquired image frames, OCT can extract an angiographic signal, enabling the construction of 3D, depth-resolved maps of vasculature non-invasively in the clinic. Applying these phase variance angiographic functional imaging techniques to the middle ear was the focus of the work described in Chapter 6. Through development of a novel bulk motion correction technique, enface ME-OCTA angiograms of both the TM and middle ear mucosa were generated for the first time in-vivo. Furthermore, real-time 2D cross-sectional ME-OCTA in-vivo imaging was demonstrated for the first time and used to visualize the contralateral stapedius reflex of a healthy adult volunteer.

Although the results of this thesis have contributed to better visualization and imaging of middle ear structure and dynamics at the point-of-care, there are many aspects that will require future refinement and improvement. These improvements include enhancements to the digital tympanotomy technique to enable its application to more diverse and pathological ears, refinements to the Doppler animation technique to account for the imaging geometry of the ME-OCT system, simplification of the software

architecture, and validation of vascular imaging techniques introduced in this thesis in patients.

The goal of the work presented in Chapters 3—6 was to improve the visualization of middle ear pathologies using ME-OCT, develop software necessary to enable its use by clinicians at the point-of-care and to lower the barriers for the clinical adoption of ME-OCT. There are several technical directions this work could potentially be extended including: using geometrically accurate 4D imaging to map out the middle ear space in real-time through mosaicking of collected volumes, incorporation of spiral-scanning with angiography to achieve 4D ME-OCTA, and incorporation of 4D ME-OCT imaging into surgical CI robots to bring ME-OCT out of the clinic and into the OR. Ultimately however, the adoption of ME-OCT into clinical practice will depend on whether it can deliver conclusive diagnostic benefits to clinicians. By enabling in-vivo ME-OCT imaging at the point-of-care, this work provides a solid foundation for further clinically focused research studies to investigate the clinical utility of the technology which are needed to make ME-OCT a viable diagnostic tool for clinical otology.

One of the most promising extensions of this research, beyond ME-OCT, is its potential for real-time robotic surgical navigation. Current robotic systems for cochlear implantation, like HEARO [235], rely on pre-surgery HRCT and MRI images for path planning. However, with the advancements highlighted in this thesis, 4D OCT imaging could offer real-time intraoperative feedback, enhancing surgical navigation by continuously comparing OCT volumes to preplanned HRCT paths. This can minimize patient risks and enhance outcomes, especially during complications that deviate from the initial surgical route. This thesis lays a robust groundwork for advancing OCT-guided robotic navigation research, broadening the applications of real-time ME-OCT imaging beyond clinics.

References

1. D. MacDougall, J. Farrell, J. Brown, M. Bance, and R. Adamson, "Long-range, wide-field swept-source optical coherence tomography with GPU accelerated digital lock-in Doppler vibrography for real-time, in vivo middle ear diagnostics," Biomed. Opt. Express **7**(11), 4621 (2016).

2. J. Fujimoto and E. Swanson, "The Development, Commercialization, and Impact of Optical Coherence Tomography," 13.

3. M. Rubinstein, E. L. Fine, A. Sepehr, W. B. Armstrong, L. Crumley, J. H. Kim, Z. Chen, and B. J. F. Wong, "Optical Coherence Tomography of the Larynx Using the Niris System," 16 (2010).

4. E. Sattler, R. Kästle, and J. Welzel, "Optical coherence tomography in dermatology," J. Biomed. Opt **18**(6), 061224 (2013).

5. B. E. Bouma, G. J. Tearney, C. C. Compton, and N. S. Nishioka, "High-resolution imaging of the human esophagus and stomach in vivo using optical coherence tomography," Gastrointest Endosc **51**(4 Pt 1), 467–474 (2000).

6. "Advances in Optical Coherence Tomography and Confocal Laser Endomicroscopy in Pulmonary Diseases," 16.

7. A. Meller, M. Shakhova, Y. Rilkin, A. Novozhilov, M. Kirillin, and A. Shakhov, "Optical coherence tomography in diagnosing inflammatory diseases of ENT,".

8. C. Pitris, K. T. Saunders, J. G. Fujimoto, and M. E. Brezinski, "High-resolution imaging of the middle ear with optical coherence tomography: a feasibility study," Arch. Otolaryngol. Head Neck Surg. **127**(6), 637–642 (2001).

9. D. MacDougall, J. Rainsbury, J. Brown, M. Bance, and R. Adamson, "Optical coherence tomography system requirements for clinical diagnostic middle ear imaging," J. Biomed. Opt **20**(5), 056008 (2015).

10. M. Bonesi, M. P. Minneman, J. Ensher, B. Zabihian, H. Sattmann, P. Boschert, E. Hoover, R. A. Leitgeb, M. Crawford, and W. Drexler, "Akinetic all-semiconductor programmable swept-source at 1550 nm and 1310 nm with centimeters coherence length," Opt. Express **22**(3), 2632 (2014).

11. T. J. Matthews and R. Adamson, "Optical coherence tomography: current and future clinical applications in otology," Curr Opin Otolaryngol Head Neck Surg **28**(5), 296–301 (2020).

12. H. E. I. Tan, P. L. Santa Maria, P. Wijesinghe, B. Francis Kennedy, B. J. Allardyce, R. H. Eikelboom, M. D. Atlas, and R. J. Dilley, "Optical Coherence Tomography of the Tympanic Membrane and Middle Ear: A Review," Otolaryngol Head Neck Surg **159**(3), 424–438 (2018).

13. G. L. Monroy, J. Won, J. Shi, M. C. Hill, R. G. Porter, M. A. Novak, W. Hong, P. Khampang, J. E. Kerschner, D. R. Spillman, and S. A. Boppart, "Automated classification of otitis media with OCT: augmenting pediatric image datasets with gold-standard animal model data," Biomed Opt Express **13**(6), 3601–3614 (2022).

14. D. Preciado, R. M. Nolan, R. Joshi, G. M. Krakovsky, A. Zhang, N. A. Pudik, N. K. Kumar, R. L. Shelton, S. A. Boppart, and N. M. Bauman, "Otitis Media Middle Ear Effusion Identification and Characterization Using an Optical Coherence Tomography Otoscope," Otolaryngol Head Neck Surg **162**(3), 367–374 (2020).

15. J. Morgenstern, M. Schindler, L. Kirsten, J. Golde, M. Bornitz, M. Kemper, E. Koch, T. Zahnert, and M. Neudert, "Endoscopic Optical Coherence Tomography for Evaluation of Success of Tympanoplasty," Otology & Neurotology **41**(7), e901–e905 (2020).

16. S. Van der Jeught, J. J. J. Dirckx, J. R. M. Aerts, A. Bradu, A. G. H. Podoleanu, and J. A. N. Buytaert, "Full-field thickness distribution of human tympanic membrane obtained with optical coherence tomography," J Assoc Res Otolaryngol **14**(4), 483–494 (2013).

17. H. R. Djalilian, M. Rubinstein, E. C. Wu, K. Naemi, S. Zardouz, K. Karimi, and B. J. F. Wong, "Optical Coherence Tomography of Cholesteatoma:," Otology & Neurotology **31**(6), 932–935 (2010).

18. D. MacDougall, L. Morrison, C. Morrison, D. P. Morris, M. Bance, and R. B. A. Adamson, "Optical Coherence Tomography Doppler Vibrometry Measurement of Stapes Vibration in Patients With Stapes Fixation and Normal Controls:," Otology & Neurotology **40**(4), e349–e355 (2019).

19. W. Kim, S. Kim, S. Huang, J. S. Oghalai, and B. E. Applegate, "Picometer scale vibrometry in the human middle ear using a surgical microscope based optical coherence tomography and vibrometry system," Biomed. Opt. Express **10**(9), 4395 (2019).

20. M. Golabbakhsh, X. Wang, D. MacDougall, J. Farrell, T. Landry, W. R. J. Funnell, and R. Adamson, "Finite-element modelling based on optical coherence tomography and corresponding X-ray microCT data for two human middle ears,".

21. C. G. Lui, W. Kim, J. B. Dewey, F. D. Macías-Escrivá, K. Ratnayake, J. S. Oghalai, and B. E. Applegate, "In vivo functional imaging of the human middle ear with a hand-held optical coherence tomography device," Biomed. Opt. Express **12**(8), 5196 (2021).

22. V. Milanovic, "Linearized Gimbal-less Two-Axis MEMS Mirrors," in *Optical Fiber Communication Conference and National Fiber Optic Engineers Conference* (OSA, 2009), p. JThA19.

23. MirrorcleTech, "A8L1.1-3000AL datasheet," (2016).

24. J. Wang, G. Chawdhary, X. Yang, F. Morin, M. Khalid-Raja, J. Farrell, D. MacDougall, F. Chen, D. P. Morris, and R. B. A. Adamson, "Optical Clearing Agents for Optical Imaging Through Cartilage Tympanoplasties: A Preclinical Feasibility Study," Otology & Neurotology **43**(4), e467–e474 (2022).

25. J. Wang, G. Chawdhary, J. Farrell, X. Yang, M. Farrell, D. MacDougall, M. Trudel, N. Shoman, D. P. Morris, and R. B. A. Adamson, "Transtympanic Visualization of Cochlear Implant Placement With Optical Coherence Tomography: A Pilot Study," Otol Neurotol **43**(8), e824–e828 (2022).

26. S. Mansour, J. Magnan, H. Haidar, K. Nicolas, and S. Louryan, *Comprehensive and Clinical Anatomy of the Middle Ear* (Springer Berlin Heidelberg, 2013).

27. C. B. Ruah, P. A. Schachern, D. Zelterman, M. M. Paparella, and T. H. Yoon, "Age-Related Morphologic Changes in the Human Tympanic Membrane: A Light and Electron Microscopic Study," Archives of Otolaryngology - Head and Neck Surgery **117**(6), 627–634 (1991).

28. T. George and B. Bordoni, "Anatomy, Head and Neck, Ear Ossicles," https://www.ncbi.nlm.nih.gov/books/NBK570549/.

29. B. M. Carlson, "Special Senses—Vision and Hearing," in *The Human Body* (Elsevier, 2019), pp. 177–207.

30. S. Mukerji, A. M. Windsor, and D. J. Lee, "Auditory Brainstem Circuits That Mediate the Middle Ear Muscle Reflex," Trends in Amplification **14**(3), 170–191 (2010).

31. R. J. Tanna, J. W. Lin, and O. De Jesus, "Sensorineural Hearing Loss," in *StatPearls* (StatPearls Publishing, 2023).

32. K. J. Cruickshanks, T. L. Wiley, T. S. Tweed, B. E. K. Klein, R. Klein, J. A. Mares-Perlman, and D. M. Nondahl, "Prevalence of Hearing Loss in Older Adults in Beaver Dam, Wisconsin: The Epidemiology of Hearing Loss Study," American Journal of Epidemiology **148**(9), 879–886 (1998).

33. K. J. Cruickshanks, T. S. Tweed, T. L. Wiley, B. E. K. Klein, R. Klein, R. Chappell, D. M. Nondahl, and D. S. Dalton, "The 5-Year Incidence and Progression of Hearing Loss: The Epidemiology of Hearing Loss Study," Arch Otolaryngol Head Neck Surg **129**(10), 1041 (2003).

34. L. L. Cunningham and D. L. Tucci, "Hearing Loss in Adults," N Engl J Med **377**(25), 2465–2473 (2017).

35. B. Schick and J. Dlugaiczyk, "Surgery of the ear and the lateral skull base: pitfalls and complications," GMS Current Topics in Otorhinolaryngology - Head and Neck Surgery; 12:Doc05; ISSN 1865-1011 (2013).

36. N. Pollak, "Endoscopic and minimally-invasive ear surgery: A path to better outcomes," World j. otorhinolaryngol.-head neck surg. **3**(3), 129–135 (2017).

37. P. Chang and S. Kim, "Cholesteatoma - diagnosing the unsafe ear," (n.d.).

38. J. C. Goddard and J. N. Fayad, "Otoscopic Findings in Otosclerosis," Ear Nose Throat J **91**(4), E30–E30 (2012).

39. S. R. Falkson and P. Tadi, "Otoscopy," in *StatPearls* (StatPearls Publishing, 2023).

40. D.-H. Lee and S.-W. Yeo, "Clinical Diagnostic Accuracy of Otitis Media with Effusion in Children, and Significance of Myringotomy: Diagnostic or Therapeutic?," J Korean Med Sci **19**(5), 739 (2004).

41. L. Ma and B. Fei, "Comprehensive review of surgical microscopes: technology development and medical applications," J. Biomed. Opt. **26**(01), (2021).

42. A. C. Moberly, M. Zhang, L. Yu, M. Gurcan, C. Senaras, T. N. Teknos, C. A. Elmaraghy, N. Taj-Schaal, and G. F. Essig, "Digital otoscopy versus microscopy: How correct and confident are ear experts in their diagnoses?," J Telemed Telecare **24**(7), 453–459 (2018).

43. D. E. Young, W.-J. F. Ten Cate, Z. Ahmad, and R. P. Morton, "The accuracy of otomicroscopy for the diagnosis of paediatric middle ear effusions," International Journal of Pediatric Otorhinolaryngology **73**(6), 825–828 (2009).

44. A. Garcia, S. E. Ridge, J. M. Garcia, M. Cohen, and D. J. Lee, "New perspectives in office-based otoendoscopy and endoscopic ear surgery," Operative Techniques in Otolaryngology-Head and Neck Surgery **32**(2), 68–78 (2021).

45. M. Preis, "Otoendoscopy in the Office and Operating Room," Otolaryngologic Clinics of North America **54**(1), 59–64 (2021).

46. S. Hoth and I. Baljić, "Current audiological diagnostics," GMS Curr Top Otorhinolaryngol Head Neck Surg **16**, Doc09 (2017).

47. M. Sliwinska-kowalska, "Hearing," in *Handbook of Clinical Neurology* (Elsevier, 2015), **131**, pp. 341–363.

48. J. T. Jacobson, "The Role of Immittance Audiometry in Detecting Middle Ear Disease," **27**, 7 (1981).

49. T. Frank and D. R. Petersen, "Accuracy of a 40 dB HL Audioscope™ and Audiometer Screening for Adults:," Ear and Hearing **8**(3), 180–183 (1987).

50. F. E. Musiek, J. Shinn, G. D. Chermak, and D.-E. Bamiou, "Perspectives on the Pure-Tone Audiogram," J Am Acad Audiol **28**(07), 655–671 (2017).

51. R. Carhart, "CLINICAL APPLICATION OF BONE CONDUCTION AUDIOMETRY," Archives of Otolaryngology - Head and Neck Surgery **51**(6), 798–808 (1950).

52. K. J. Van Camp, R. H. Margolis, R. H. Wilson, and W. L. Creten, "Principles of Tympanometry," American Speech-Language-Hearing Association 1–88 (1986).

53. J. J. Rosowski, H. H. Nakajima, M. A. Hamade, L. Mahfoud, G. R. Merchant, C. F. Halpin, and S. N. Merchant, "Ear-Canal Reflectance, Umbo Velocity, and Tympanometry in Normal-Hearing Adults:," Ear and Hearing **33**(1), 19–34 (2012).

54. C. A. Sanford, T. Schooling, and T. Frymark, "Determining the presence or absence of middle ear disorders: an evidence-based systematic review on the diagnostic accuracy of selected assessment instruments," Am J Audiol **21**(2), 251–268 (2012).

55. J. C. Benson, M. L. Carlson, and J. I. Lane, "MRI of the Internal Auditory Canal, Labyrinth, and Middle Ear: How We Do It," Radiology **297**(2), 252–265 (2020).

56. C. L. Thukral, "Role of High Resolution Computed Tomography in Evaluation of Pathologies of Temporal Bone," JCDR (2015).

57. B. Schick and J. Dlugaiczyk, "Surgery of the ear and the lateral skull base: pitfalls and complications," GMS Current Topics in Otorhinolaryngology - Head and Neck Surgery; 12:Doc05; ISSN 1865-1011 (2013).

58. G. Chodick, N. Bekiroglu, M. Hauptmann, B. H. Alexander, D. M. Freedman, M. M. Doody, L. C. Cheung, S. L. Simon, R. M. Weinstock, A. Bouville, and A. J. Sigurdson, "Risk of cataract after exposure to low doses of ionizing radiation: a 20-year prospective cohort study among US radiologic technologists," Am. J. Epidemiol. **168**(6), 620–631 (2008).

59. J. L. Dornhoffer and M. B. Gluth, eds., *The Chronic Ear* (Thieme, 2016).

60. A. Trojanowska, A. Drop, P. Trojanowski, K. Rosińska-Bogusiewicz, J. Klatka, and B. Bobek-Billewicz, "External and middle ear diseases: radiological diagnosis based on clinical signs and symptoms," Insights Imaging **3**(1), 33–48 (2012).

61. M. Hori, A. Hagiwara, M. Goto, A. Wada, and S. Aoki, "Low-Field Magnetic Resonance Imaging: Its History and Renaissance," Invest Radiol **56**(11), 669–679 (2021).

62. C. S. Ebert and H. C. Pillsbury, "Otitis Media: Background and Science," in *Managing the Allergic Patient* (Elsevier, 2008), pp. 175–191.

63. J. Sade and A. Ar, "Middle ear and auditory tube: Middle ear clearance, gas exchange, and pressure regulation," Otolaryngology - Head and Neck Surgery **116**(4), 499–524 (1997).

64. A. G. M. Schilder, M. F. Bhutta, C. C. Butler, C. Holy, L. H. Levine, K. J. Kvaerner, G. Norman, R. J. Pennings, D. Poe, J. T. Silvola, H. Sudhoff, and V. J. Lund, "Eustachian tube dysfunction: consensus statement on definition, types, clinical presentation and diagnosis," Clinical Otolaryngology **40**(5), 407–411 (2015).

65. J. C. Luers and K.-B. Hüttenbrink, "Surgical anatomy and pathology of the middle ear," J. Anat. **228**(2), 338–353 (2016).

66. E. M. Rash, "Recognize Cholesteatomas Early:," The Nurse Practitioner **29**(2), 24–27 (2004).

67. B. A. Jennings, P. Prinsley, C. Philpott, G. Willis, and M. F. Bhutta, "The genetics of cholesteatoma. A systematic review using narrative synthesis," Clin Otolaryngol **43**(1), 55–67 (2018).

68. D. J. Lim and W. H. Saunders, "Acquired Cholesteatoma: Light and Electron Microscopic Observations," Ann Otol Rhinol Laryngol **81**(1), 2–12 (1972).

69. A. L. Alves, C. S. B. Pereira, F. de A. Q. Ribeiro, and J. H. T. G. Fregnani, "Analysis of histopathological aspects in acquired middle ear cholesteatoma," Brazilian Journal of Otorhinolaryngology **74**(6), 835–841 (2008).

70. L. Podoshin, A. Margalit, M. Fradis, A. Tamir, Y. Ben-David, and L. Epstein, "Cholesteatoma: An Epidemiologic Study among Members of Kibbutzim in Northern Israel," Ann Otol Rhinol Laryngol **95**(4), 365–368 (1986).

71. M. Verhoeff, E. L. van der Veen, M. M. Rovers, E. A. M. Sanders, and A. G. M. Schilder, "Chronic suppurative otitis media: A review," International Journal of Pediatric Otorhinolaryngology **70**(1), 1–12 (2006).

72. P. S. Roland, "The Formation and Management of Middle Ear Granulation Tissue in Chronic Ear Disease," Ear, Nose & Throat Journal **83**(1_suppl), 5–8 (2004).

73. M. Burmølle, T. R. Thomsen, M. Fazli, I. Dige, L. Christensen, P. Homøe, M. Tvede, B. Nyvad, T. Tolker-Nielsen, M. Givskov, C. Moser, K. Kirketerp-Møller, H. K. Johansen, N. Høiby, P. Ø. Jensen, S. J. Sørensen, and T. Bjarnsholt, "Biofilms in chronic infections – a matter of opportunity – monospecies biofilms in multispecies infections," FEMS Immunol Med Microbiol **59**(3), 324–336 (2010).

74. M. R. Lee, K. S. Pawlowski, A. Luong, A. D. Furze, and P. S. Roland, "Biofilm presence in humans with chronic suppurative otitis media," Otolaryngol Head Neck Surg **141**(5), 567–571 (2009).

75. L. O. Bakaletz, "Bacterial Biofilms in Otitis Media: Evidence and Relevance," The Pediatric Infectious Disease Journal **26**(Supplement), S17–S19 (2007).

76. S. Mansour, J. Magnan, K. Nicolas, and H. Haidar, *Middle Ear Diseases: Advances in Diagnosis and Management* (Springer International Publishing, 2018).

77. D. Trune and A. Nguyen-Huynh, "Vascular Pathophysiology in Hearing Disorders," Semin Hear **33**(03), 242–250 (2012).

78. P. S. Roland, B. B. Isaacson, and J. W. Kutz, "Office Management of Tympanic Membrane Perforation and the Draining Ear," in *Otologic Surgery* (Elsevier, 2010), pp. 107–118.

79. S. Asiri, A. Hasham, F. A. Anazy, S. Zakzouk, and A. Banjar, "Tympanosclerosis: review of literature and incidence among patients with middle-ear infection," J. Laryngol. Otol. **113**(12), 1076–1080 (1999).

80. D. Cohen and D. Tamir, "The Prevalenc of Middle Ear Pathologies in Jerusalem School Children:," Otology & Neurotology **10**(6), 456–459 (1989).

81. WHO, *Chronic Suppurative Otitis Media: Burden of Illness and Management Options* (WHO, 2004).

82. L. Monasta, L. Ronfani, F. Marchetti, M. Montico, L. Vecchi Brumatti, A. Bavcar, D. Grasso, C. Barbiero, and G. Tamburlini, "Burden of Disease Caused by Otitis Media: Systematic Review and Global Estimates," PLoS ONE **7**(4), e36226 (2012).

83. A. G. M. Schilder, T. Chonmaitree, A. W. Cripps, R. M. Rosenfeld, M. L. Casselbrant, M. P. Haggard, and R. P. Venekamp, "Otitis media," Nat Rev Dis Primers **2**, 16063 (2016).

84. R. W. Hinerman, R. J. Amdur, C. G. Morris, J. Kirwan, and W. M. Mendenhall, "Definitive radiotherapy in the management of paragangliomas arising in the head and neck: A 35-year experience," Head Neck **30**(11), 1431–1438 (2008).

85. D. A. Moffat and D. G. Hardy, "Surgical management of large glomus jugulare tumours: infra-and trans-temporal approach," J. Laryngol. Otol. **103**(12), 1167–1180 (1989).

86. W. F. Young, "Paragangliomas: Clinical Overview," Annals of the New York Academy of Sciences **1073**(1), 21–29 (2006).

87. V. R. Appannan and M. K. Md Daud, "Glomus tympanicum," Malays Fam Physician **13**(1), 45–48 (2018).

88. J. Thelen and A. A. Bhatt, "Multimodality imaging of paragangliomas of the head and neck," Insights Imaging **10**(1), 29 (2019).

89. S. Kösling, M. Omenzetter, and S. Bartel-Friedrich, "Congenital malformations of the external and middle ear," European Journal of Radiology **69**(2), 269–279 (2009).

90. S. Quesnel, T. Benchaa, S. Bernard, F. Martine, P. Viala, T. Van Den Abbeele, and N. Teissier, "Congenital Middle Ear Anomalies: Anatomical and Functional Results of Surgery," Audiol Neurotol **20**(4), 237–242 (2015).

91. S. Bartel-Friedrich and C. Wulke, "Classification and diagnosis of ear malformations," GMS Curr Top Otorhinolaryngol Head Neck Surg **6**, Doc05 (2007).

92. M. Rudic, I. Keogh, R. Wagner, E. Wilkinson, N. Kiros, E. Ferrary, O. Sterkers, A. Bozorg Grayeli, K. Zarkovic, and N. Zarkovic, "The pathophysiology of otosclerosis: Review of current research," Hear. Res. **330**(Pt A), 51–56 (2015).

93. B. Purohit, R. Hermans, and K. Op De Beeck, "Imaging in otosclerosis: A pictorial review," Insights Imaging **5**(2), 245–252 (2014).

94. "Creative Commons — Attribution 4.0 International — CC BY 4.0," https://creativecommons.org/licenses/by/4.0/legalcode.

95. N. Dolhi and A. D. Weimer, "Tympanic Membrane Perforations," https://www.ncbi.nlm.nih.gov/books/NBK557887/.

96. L. Podoshin, M. Fradis, S. Malatskey, and J. Ben-David, "Tympanoplasty in Adults: A Five-Year Survey," Ear Nose Throat J **75**(3), 149–156 (1996).

97. S. Delrue, N. Verhaert, J. V. Dinther, A. Zarowski, T. Somers, C. Desloovere, and E. Offeciers, "Surgical Management and Hearing Outcome of Traumatic Ossicular Injuries," Int Adv Otol **12**(3), 231–236 (2016).

98. M. Abhishek, R. Kaleeswaran, and K. Srinivasan, "Assessment of Hearing Loss in Temporal Bone Fractures," Indian Journal of Otology **27**(3), (2021).

99. G. Lee, Y. Kim, and B. J. Kim, "Multiple Ossicular Dislocation Including Stapediovestibular Dislocation Presenting with Conductive Hearing Loss," J Audiol Otol **25**(3), 159–162 (2021).

100. P. Meriot, F. Veillon, J. F. Garcia, M. Nonent, J. Jezequel, P. Bourjat, and M. Bellet, "CT appearances of ossicular injuries.," RadioGraphics **17**(6), 1445–1454 (1997).

101. B. C. A. van Stekelenburg and M. C. J. Aarts, "Determinants influencing success rates of myringoplasty in daily practice: a retrospective analysis," Eur Arch Otorhinolaryngol **276**(11), 3081–3087 (2019).

102. U. Fisch, G. Ö. Acar, and A. M. Huber, "Malleostapedotomy in Revision Surgery for Otosclerosis:," Otology & Neurotology **22**(6), 776–785 (2001).

103. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, and C. A. Puliafito, "Optical Coherence Tomography," 12 (2015).

104. J. F. Bille, ed., *High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics* (Springer International Publishing, 2019).

105. D. Wang, P. Liang, S. Samuelson, H. Jia, J. Ma, and H. Xie, "Correction of image distortions in endoscopic optical coherence tomography based on two-axis scanning MEMS mirrors," Biomed. Opt. Express **4**(10), 2066 (2013).

106. A. Podoleanu, I. Charalambous, L. Plesea, A. Dogariu, and R. Rosen, "Correction of distortions in optical coherence tomography imaging of the eye," Phys. Med. Biol. **49**(7), 1277–1294 (2004).

107. E. Hecht, *Optics*, 5 ed (Pearson Education, Inc, 2017).

108. F. L. Pedrotti, L. M. Pedrotti, and L. S. Pedrotti, *Introduction to Optics*, 3rd ed (Cambridge university press, 2018).

109. W. M. Lee and S. H. Yun, "Adaptive aberration correction of GRIN lenses for confocal endomicroscopy," Opt. Lett. **36**(23), 4608 (2011).

110. V. Nguyen, S. Larouche, N. Landy, J. S. Lee, and D. R. Smith, "Quantitative comparison of gradient index and refractive lenses," J. Opt. Soc. Am. A **29**(11), 2479 (2012).

111. R. Adamson, J. Farrell, J. Wang, X. Yang, K. Brewer, D. Morris, F. Couvreur, N. Shoman, and D. MacDougall, *Geometrically Accurate Real-Time Volumetric Visualization of the Middle Ear Using Optical Coherence Tomography* (2023).

112. M. J. Nadeau, J. J. Miceli, and D. T. Moore, "Image analysis of curved gradient-index rods," Appl. Opt. **25**(11), 1780 (1986).

113. J. C. Wang and J. B. Miller, "Optical Coherence Tomography Angiography: Review of Current Technical Aspects and Applications in Chorioretinal Disease," Seminars in Ophthalmology **34**(4), 211–217 (2019).

114. T. E. de Carlo, A. Romano, N. K. Waheed, and J. S. Duker, "A review of optical coherence tomography angiography (OCTA)," Int J Retin Vitr **1**(1), 5 (2015).

115. L. Van Melkebeke, J. Barbosa-Breda, M. Huygens, and I. Stalmans, "Optical Coherence Tomography Angiography in Glaucoma: A Review," Ophthalmic Res **60**(3), 139–151 (2018).

116. Z. Sun, D. Yang, Z. Tang, D. S. Ng, and C. Y. Cheung, "Optical coherence tomography angiography in diabetic retinopathy: an updated review," Eye **35**(1), 149–161 (2021).

117. K. Y. Tey, K. Teo, A. C. S. Tan, K. Devarajan, B. Tan, J. Tan, L. Schmetterer, and M. Ang, "Optical coherence tomography angiography in diabetic retinopathy: a review of current applications," Eye and Vis **6**(1), 37 (2019).

118. S. S. Gao, Y. Jia, M. Zhang, J. P. Su, G. Liu, T. S. Hwang, S. T. Bailey, and D. Huang, "Optical Coherence Tomography Angiography," Invest. Ophthalmol. Vis. Sci. **57**(9), OCT27 (2016).

119. R. F. Spaide, J. G. Fujimoto, N. K. Waheed, S. R. Sadda, and G. Staurenghi, "Optical coherence tomography angiography," Progress in Retinal and Eye Research **64**, 1–55 (2018).

120. B. Baumann, B. Potsaid, M. F. Kraus, J. J. Liu, D. Huang, J. Hornegger, A. E. Cable, J. S. Duker, and J. G. Fujimoto, "Total retinal blood flow measurement with ultrahigh speed swept source/Fourier domain OCT," Biomed. Opt. Express **2**(6), 1539 (2011).

121. R. K. Wang and L. An, "Doppler optical micro-angiography for volumetric imaging of vascular perfusion in vivo," Opt. Express **17**(11), 8926 (2009).

122. J. Fingler, D. Schwartz, C. Yang, and S. E. Fraser, "Mobility and transverse flow visualization using phase variance contrast with spectral domain optical coherence tomography," Opt. Express **15**(20), 12636 (2007).

123. B. R. White, M. C. Pierce, N. Nassif, B. Cense, B. H. Park, G. J. Tearney, B. E. Bouma, T. C. Chen, and J. F. de Boer, "In vivo dynamic human retinal blood flow imaging using ultra-high-speed spectral domain optical Doppler tomography," Opt. Express **11**(25), 3490 (2003).

124. V. X. D. Yang, M. L. Gordon, A. Mok, Y. Zhao, Z. Chen, R. S. C. Cobbold, B. C. Wilson, and I. Alex Vitkin, "Improved phase-resolved optical Doppler tomography using the Kasai velocity estimator and histogram segmentation," Optics Communications **208**(4–6), 209–214 (2002).

125. L. An, "High-resolution wide-field imaging of retinal and choroidal blood perfusion with optical microangiography," J. Biomed. Opt **15**(2), 026011 (2010).

126. Z. Chen, T. E. Milner, S. Srinivas, X. Wang, A. Malekafzali, M. J. C. van Gemert, and J. S. Nelson, "Noninvasive imaging of in vivo blood flow velocity using optical Doppler tomography," Opt. Lett. **22**(14), 1119 (1997).

127. J. A. Izatt, M. D. Kulkarni, S. Yazdanfar, J. K. Barton, and A. J. Welch, "In vivo bidirectional color Doppler flow imaging of picoliter blood volumes using optical coherence tomography," Opt. Lett. **22**(18), 1439 (1997).

128. J. Zhang and Z. Chen, "In vivo blood flow imaging by a swept laser source based Fourier domain optical Doppler tomography," Opt. Express **13**(19), 7449 (2005).

129. L. Wang, Y. Wang, S. Guo, J. Zhang, M. Bachman, G. P. Li, and Z. Chen, "Frequency domain phase-resolved optical Doppler and Doppler variance tomography," Optics Communications **242**(4–6), 345–350 (2004).

130. R. K. Wang, S. L. Jacques, Z. Ma, S. Hurst, S. R. Hanson, and A. Gruber, "Three dimensional optical angiography," Opt. Express **15**(7), 4083 (2007).

131. G. Liu, L. Chou, W. Jia, W. Qi, B. Choi, and Z. Chen, "Intensity-based modified Doppler variance algorithm: application to phase instable and phase stable optical coherence tomography systems," Opt. Express **19**(12), 11429 (2011).

132. A. Zhang, Q. Zhang, C.-L. Chen, and R. K. Wang, "Methods and algorithms for optical coherence tomography-based angiography: a review and comparison," J. Biomed. Opt **20**(10), 100901 (2015).

133. Y. Jia, O. Tan, J. Tokayer, B. Potsaid, Y. Wang, J. J. Liu, M. F. Kraus, H. Subhash, J. G. Fujimoto, J. Hornegger, and D. Huang, "Split-spectrum amplitude-decorrelation angiography with optical coherence tomography," Opt. Express **20**(4), 4710 (2012).

134. J. Xu, S. Song, Y. Li, and R. K. Wang, "Complex-based OCT angiography algorithm recovers microvascular information better than amplitude- or phase-based algorithms in phase-stable systems," Phys. Med. Biol. **63**(1), 015023 (2017).

135. S. Mill, "The Free Lunch Is Over: A Fundamental Turn Toward Concurrency in Software," 7.

136. D. Kirk and W. Hwu, *Programming Massively Parallel Processors: A Hands-on Approach*, Third edition (Elsevier, 2017).

137. J. Cheng, *Professional Cuda C Programming*, Wrox Programmer to Programmer (John Wiley and Sons, Inc, 2014).

138. Nvidia, *NVIDIA CUDA Compute Unified Device Architecture, Programming Guide, Version 1.0* (2007).

139. M. J. Flynn, "Some Computer Organizations and Their Effectiveness," IEEE Trans. Comput. **C–21**(9), 948–960 (1972).

140. J. L. Hennessy, "Computer Architecture: A Quantitative Approach," 1527.

141. "History of OpenGL - OpenGL Wiki," https://www.khronos.org/opengl/wiki/History_of_OpenGL#OpenGL_1.0_(1992).

142. "OpenGL 4.6 (Core Profile) - May 5, 2022".

143. "CUDA C Best Practices Guide," 85.

144. S. Cook, *CUDA Programming: A Developer's Guide to Parallel Computing with GPUs* (Elsevier, MK, 2013).

145. "Tuning CUDA Applications for Ampere," 10.

146. G. L. Monroy, R. L. Shelton, R. M. Nolan, C. T. Nguyen, M. A. Novak, M. C. Hill, D. T. McCormick, and S. A. Boppart, "Noninvasive depth-resolved optical measurements of the tympanic membrane and middle ear for differentiating otitis media: Differentiation of Otitis Media Using OCT," The Laryngoscope **125**(8), E276–E282 (2015).

147. C. T. Nguyen, W. Jung, J. Kim, E. J. Chaney, M. Novak, C. N. Stewart, and S. A. Boppart, "Noninvasive in vivo optical detection of biofilm in the human middle ear," Proceedings of the National Academy of Sciences **109**(24), 9529–9534 (2012).

148. K. Park, N. H. Cho, J. H. Jang, S. H. Lee, P. Kim, M. Jeon, S. A. Boppart, J. Kim, and W. Jung, "In vivo 3D imaging of the human tympanic membrane using a wide-field diagonal-scanning optical coherence tomography probe," Appl. Opt. **56**(9), D115 (2017).

149. O. M. Carrasco-Zevallos, B. Keller, C. Viehland, L. Shen, G. Waterman, B. Todorich, C. Shieh, P. Hahn, S. Farsiu, A. N. Kuo, C. A. Toth, and J. A. Izatt, "Live volumetric (4D) visualization and guidance of in vivo human ophthalmic surgery with intraoperative optical coherence tomography," Sci Rep **6**(1), 31689 (2016).

150. C. Viehland, B. Keller, O. M. Carrasco-Zevallos, D. Nankivil, L. Shen, S. Mangalesh, D. T. Viet, A. N. Kuo, C. A. Toth, and J. A. Izatt, "Enhanced volumetric visualization for real time 4D intraoperative ophthalmic swept-source OCT," Biomed Opt Express **7**(5), 1815–1829 (2016).

151. "3D Slicer image computing platform," https://slicer.org/.

152. A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. V. Miller, S. Pieper, and R. Kikinis, "3D Slicer as an image computing platform for the Quantitative Imaging Network," Magnetic Resonance Imaging **30**(9), 1323–1341 (2012).

153. M. McCormick, X. Liu, J. Jomier, C. Marion, and L. Ibanez, "ITK: enabling reproducible research and open science," Front. Neuroinform. **8**, (2014).

154. W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit (4th Ed.)* (Kitware, 2006).

155. W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics ; [Visualize Data in 3D - Medical, Engineering or Scientific ; Build Your Own Applications with C++, Tcl, Java or Python ; Includes Source Code for VTK (Supports Unix, Windows and Mac)*, 4. ed (Kitware, Inc, 2006).

156. M. McCormick, X. Liu, J. Jomier, C. Marion, and L. Ibanez, "ITK: enabling reproducible research and open science," Front. Neuroinform. **8**, (2014).

157. W.-K. Jeong, H. Pfister, and M. Fatica, "Medical Image Processing Using GPU-Accelerated ITK Image Filters," in *GPU Computing Gems Emerald Edition* (Elsevier, 2011), pp. 737–749.

158. M. Nunes, M. Schlachter, and K. Bühler, "Visualization Tools for Radiotherapy - A Survey," Innovative imaging to improve radiotherapy treatments 21 (n.d.).

159. D. MacDougall, "Optical Coherence Tomography for Clinical Otology," Dalhousie University (2021).

160. S. Stegmaier, M. Strengert, T. Klein, and T. Ertl, "A simple and flexible volume rendering framework for graphics-hardware-based raycasting," in *Fourth International Workshop on Volume Graphics, 2005.* (IEEE, 2005), pp. 187–241.

161. M. Levoy, "Display of surfaces from volume data," IEEE Comput. Grap. Appl. **8**(3), 29–37 (1988).

162. L. Kay and T. Kajiya, "Ray Tracing Complex Scenes," **20**(4), 10 (1986).

163. Q. Zhang, R. Eagleson, and T. M. Peters, "Volume Visualization: A Technical Overview with a Focus on Medical Applications," J Digit Imaging **24**(4), 640–664 (2011).

164. J. F. Blinn, "Models of Light Reflection for Computer Synthesized Pictures,".

165. M. Golabbakhsh, X. Wang, D. MacDougall, J. Farrell, T. Landry, W. R. J. Funnell, and R. Adamson, "Finite-Element Modelling Based on Optical Coherence Tomography and Corresponding X-ray MicroCT Data for Three Human Middle Ears," JARO (2023).

166. D. MacDougall, J. Farrell, J. Brown, and R. Adamson, "Swept-source optical coherence tomography for real-time in vivo otology imaging with GPU accelerated Doppler vibrography," 12.

167. J. Wang, B. Wohlberg, and R. B. A. Adamson, "Convolutional dictionary learning for blind deconvolution of optical coherence tomography images," Biomed. Opt. Express **13**(4), 1834 (2022).

168. W. Kim, D. Liu, S. Kim, K. Ratnayake, F. Macias-Escriva, S. Mattison, J. S. Oghalai, and B. E. Applegate, "Vector of motion measurements in the living cochlea using a 3D OCT vibrometry system," Biomed. Opt. Express **13**(4), 2542 (2022).

169. M. Khaleghi, J. T. Cheng, C. Furlong, and J. J. Rosowski, "In-plane and out-of-plane motions of the human tympanic membrane," J Acoust Soc Am **139**(1), 104–117 (2016).

170. A. Eklund, P. Dufort, D. Forsberg, and S. M. LaConte, "Medical image processing on the GPU – Past, present and future," Medical Image Analysis **17**(8), 1073–1094 (2013).

171. H. L. Khor, S.-C. Liew, and J. M. Zain, "A review on parallel medical image processing on GPU," in *2015 4th International Conference on Software Engineering and Computer Systems (ICSECS)* (IEEE, 2015), pp. 45–48.

172. T. Kalaiselvi, P. Sriramakrishnan, and K. Somasundaram, "Survey of using GPU CUDA programming model in medical image analysis," Informatics in Medicine Unlocked **9**, 133–144 (2017).

173. Z. Hubler, N. D. Shemonski, R. L. Shelton, G. L. Monroy, R. M. Nolan, and S. A. Boppart, "Real-time automated thickness measurement of the in vivo human tympanic membrane using optical coherence tomography," Quantitative Imaging in Medicine and Surgery **5**(1), 9 (2015).

174. Y. Jian, K. Wong, and M. V. Sarunic, "Graphics processing unit accelerated optical coherence tomography processing at megahertz axial scan rate and high resolution video rate volumetric rendering," J. Biomed. Opt. **18**(02), 1 (2013).

175. Y.-P. Huang, T.-Y. Tsai, T.-H. Chen, C.-B. Chueh, Y.-N. Tsai, Y.-W. Chang, Y.-C. Wu, H.-W. Chen, M.-T. Tsai, Y.-P. Hung, and H.-C. Lee, "A Generic Framework for Fourier-Domain Optical Coherence Tomography Imaging: Software Architecture and Hardware Implementations," IEEE Access **8**, 191726–191736 (2020).

176. R. Chandra, L. Dagum, D. Kohr, R. Menon, D. Maydan, and J. McDonald, *Parallel Programming in OpenMP* (Morgan kaufmann, 2001).

177. G. E. Krasner, S. T. Pope, and P. Systems, "A Description of the Model-View-Controller User Interface Paradigm in the Smalltalk-80 System,".

178. "Qt - One framework. One codebase. Any platform.," https://www.qt.io/.

179. A. Karagkasidis, "Developing GUI Applications. Architectural Patterns Revisited,".

180. G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools (2000).

181. "RMM: RAPIDS Memory Manager," (2023).

182. V. Jayaraman, J. Jiang, B. Potsaid, M. Robertson, P. J. S. Heim, C. Burgner, D. John, G. D. Cole, I. Grulkowski, J. G. Fujimoto, A. M. Davis, and A. E. Cable, "VCSEL Swept Light Sources," in *Optical Coherence Tomography*, W. Drexler and J. G. Fujimoto, eds. (Springer International Publishing, 2015), pp. 659–686.

183. M. Voss, R. Asenjo, and J. Reinders, *Pro TBB: C++ Parallel Programming with Threading Building Blocks* (Apress, 2019).

184. H. R. Djalilian, J. Ridgway, M. Tam, A. Sepehr, Z. Chen, and B. J. F. Wong, "Imaging the human tympanic membrane using optical coherence tomography in vivo," Otol Neurotol **29**(8), 1091–1094 (2008).

185. O. M. Carrasco-Zevallos, C. Viehland, B. Keller, R. P. McNabb, A. N. Kuo, and J. A. Izatt, "Constant linear velocity spiral scanning for near video rate 4D OCT ophthalmic and surgical imaging with isotropic transverse sampling," Biomed. Opt. Express **9**(10), 5052 (2018).

186. S. Ni, T.-T. P. Nguyen, R. Ng, S. Khan, S. Ostmo, Y. Jia, M. F. Chiang, D. Huang, J. Peter Campbell, and Y. Jian, "105° field of view non-contact handheld swept-source optical coherence tomography," Opt. Lett. **46**(23), 5878 (2021).

187. M. Niederleithner, M. Salas, L. Ginner, R. Leitgeb, W. Drexler, and T. Schmoll, "Spiral Scanning OCT Angiography," presented at ARVO Imaging in the Eye Conference (August 2019).

188. A. Britten, P. Matten, J. Weiss, M. Niederleithner, H. Roodaki, B. Sorg, N. Hecker-Denschlag, W. Drexler, R. A. Leitgeb, and T. Schmoll, "Surgical microscope integrated MHz SS-OCT with live volumetric visualization," Biomed. Opt. Express **14**(2), 846 (2023).

189. M. Ourak, B. Tamadazte, G. J. Laurent, and N. Andreff, "Geometric Calibration of an OCT Imaging System," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2018), pp. 3993–3999.

190. A. Podoleanu, I. Charalambous, L. Plesea, A. Dogariu, and R. Rosen, "Correction of distortions in optical coherence tomography imaging of the eye," Phys. Med. Biol. **49**(7), 1277–1294 (2004).

191. M. Chen, J. C. Gee, J. L. Prince, and G. K. Aguirre, "2D Modeling and Correction of Fan-Beam Scan Geometry in OCT," in *Computational Pathology and Ophthalmic Medical Image Analysis*, D. Stoyanov, Z. Taylor, F. Ciompi, Y. Xu, A. Martel, L. Maier-Hein, N. Rajpoot, J. van der Laak, M. Veta, S. McKenna, D. Snead, E. Trucco, M. K. Garvin, X. J. Chen, and H. Bogunovic, eds., Lecture Notes in Computer Science (Springer International Publishing, 2018), **11039**, pp. 328–335.

192. V. Westphal, A. Rollins, S. Radhakrishnan, and J. Izatt, "Correction of geometric and refractive image distortions in optical coherence tomography applying Fermat's principle," Opt. Express **10**(9), 397 (2002).

193. R. J. Zawadzki, A. R. Fuller, S. S. Choi, D. F. Wiley, B. Hamann, and J. S. Werner, "Correction of motion artifacts and scanning beam distortions in 3D ophthalmic optical coherence tomography imaging," in F. Manns, P. G. Soederberg, A. Ho, B. E. Stuck, and M. Belkin, eds. (2007), p. 642607.

194. S. Van der Jeught, J. A. N. Buytaert, A. Bradu, A. G. Podoleanu, and J. J. J. Dirckx, "Real-time correction of geometric distortion artefacts in large-volume optical coherence tomography," Meas. Sci. Technol. **24**(5), 057001 (2013).

195. V. Milanovic, G. A. Matus, and D. T. McCormick, "Gimbal-Less Monolithic Silicon Actuators for Tip–Tilt–Piston Micromirror Applications," IEEE J. Select. Topics Quantum Electron. **10**(3), 462–471 (2004).

196. V. Milanović, A. Kasturi, J. Yang, Y. R. Su, and F. Hu, "Novel packaging approaches for increased robustness and overall performance of gimbal-less MEMS mirrors," in W. Piyawattanametha and Y.-H. Park, eds. (2017), p. 1011607.

197. J. Bresenham, "A linear algorithm for incremental digital display of circular arcs," Commun. ACM **20**(2), 100–106 (1977).

198. D. Robinson, "Interpolation scan conversion in pulse-echo ultrasound," Ultrasonic Imaging **4**(4), 297–310 (1982).

199. E. Smistad, M. Bozorgi, and F. Lindseth, "FAST: framework for heterogeneous medical image computing and visualization," Int J CARS **10**(11), 1811–1822 (2015).

200. P. Virtanen, R. Gommers, T. E. Oliphant, et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python," Nat Methods **17**(3), 261–272 (2020).

201. S. T. S. Holmstrom, U. Baran, and H. Urey, "MEMS Laser Scanners: A Review," J. Microelectromech. Syst. **23**(2), 259–275 (2014).

202. C. Gorecki and S. Bargiel, "MEMS Scanning Mirrors for Optical Coherence Tomography," Photonics **8**(1), 6 (2020).

203. K. Kim, S. Moon, J. Kim, Y. Park, and J.-H. Lee, "Two-axis crosstalk analysis of gimbal-less MEMS scanners with consideration of rotational alignment," Measurement **171**, 108785 (2021).

204. N. T. Vo, R. C. Atwood, and M. Drakopoulos, "Radial lens distortion correction with sub-pixel accuracy for X-ray micro-tomography," Opt. Express **23**(25), 32859 (2015).

205. D. Zhu, R. Wang, M. Žurauskas, P. Pande, J. Bi, Q. Yuan, L. Wang, Z. Gao, and S. A. Boppart, "Automated fast computational adaptive optics for optical coherence tomography based on a stochastic parallel gradient descent algorithm," Opt. Express **28**(16), 23306 (2020).

206. S. G. Adie, B. W. Graf, A. Ahmad, P. S. Carney, and S. A. Boppart, "Computational adaptive optics for broadband optical interferometric tomography of biological tissue," Proc. Natl. Acad. Sci. U.S.A. **109**(19), 7175–7180 (2012).

207. J. Wang, E. J. Chaney, E. Aksamitiene, M. Marjanovic, and S. A. Boppart, "Computational adaptive optics for polarization-sensitive optical coherence tomography," Opt. Lett. **46**(9), 2071 (2021).

208. L. Ma and B. Fei, "Comprehensive review of surgical microscopes: technology development and medical applications," J Biomed Opt **26**(1), 010901 (2021).

209. A. B. Auinger, V. Dahm, R. Liepins, D. Riss, W.-D. Baumgartner, and C. Arnoldner, "Robotic Cochlear Implant Surgery: Imaging-Based Evaluation of Feasibility in Clinical Routine," Front. Surg. **8**, 742219 (2021).

210. N. M. Shoman, "Robotics and cochlear implant surgery: goals and developments," Current Opinion in Otolaryngology & Head & Neck Surgery **30**(5), 314–319 (2022).

211. S. Appachi, S. Schwartz, S. Ishman, and S. Anne, "Utility of intraoperative imaging in cochlear implantation: A systematic review," The Laryngoscope **128**(8), 1914–1921 (2018).

212. R. H. Margolis, "Detection of Hearing Impairment with the Acoustic Stapedius Reflex:," Ear and Hearing **14**(1), 3–10 (1993).

213. H. M. Massa, D. J. Lim, Y. Kurono, and A. W. Cripps, "Middle Ear and Eustachian Tube Mucosal Immunology," in *Mucosal Immunology* (Elsevier, 2015), pp. 1923–1942.

214. A. Ismail, "Surgical Outcomes of Revision Myringoplasty," J Otolaryngol Rhinol **7**(2), (2021).

215. J. S. Han, J. J. Han, J. M. Park, J.-H. Seo, and K. H. Park, "The Long-Term Stability of Fat-Graft Myringoplasty in the Closure of Tympanic Membrane Perforations and Hearing Restoration," ORL **83**(2), 85–92 (2021).

216. H. Noh and D.-H. Lee, "Vascularisation of myringo-/tympanoplastic grafts in active and inactive chronic mucosal otitis media: a prospective cohort study: *Vascularisation of myringo-/tympanoplastic grafts*," Clinical Otolaryngology **37**(5), 355–361 (2012).

217. K. Terkildsen, P. Osterhammel, and P. Bretlau, "Acoustic Middle Ear Muscle Reflexes in Patients With Otosclerosis," Archives of Otolaryngology - Head and Neck Surgery **98**(3), 152–155 (1973).

218. L. An and R. K. Wang, "In vivo volumetric imaging of vascular perfusion within human retina and choroids with optical micro-angiography," Opt. Express **16**(15), 11438 (2008).

219. S. Makita, Y. Hong, M. Yamanari, T. Yatagai, and Y. Yasuno, "Optical coherence angiography," Opt. Express **14**(17), 7821 (2006).

220. L. An, T. T. Shen, and R. K. Wang, "Using ultrahigh sensitive optical microangiography to achieve comprehensive depth resolved microvasculature mapping for human retina," J. Biomed. Opt. **16**(10), 106013 (2011).

221. N. Le, S. Song, Q. Zhang, and R. K. Wang, "Robust principal component analysis in optical micro-angiography," Quant. Imaging Med. Surg. **7**(6), 654–667 (2017).

222. S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," Computer Vision, Graphics, and Image Processing **39**(3), 355–368 (1987).

223. R. J. Triana, J. Prazma, V. N. Carrasco, and H. C. Pillsbury, "Analysis of blood flow in the tympanic membrane: the use of intravital fluorescence microscopy," Am J Otol **11**(4), 266–271 (1990).

224. R. F. Spaide, J. G. Fujimoto, and N. K. Waheed, "Image Artifacts in Optical Coherence Tomography Angiography".

225. X. D. Pang and W. T. Peake, "How Do Contractions of the Stapedius Muscle Alter the Acoustic Properties of the Ear?," in *Peripheral Auditory Mechanisms*, J. B. Allen, J. L. Hall, A. E. Hubbard, S. T. Neely, and A. Tubis, eds., Lecture Notes in Biomathematics (Springer Berlin Heidelberg, 1986), **64**, pp. 36–43.

226. L. Caminos, J. Garcia-Manrique, A. Lima-Rodriguez, and A. Gonzalez-Herrera, "Analysis of the Mechanical Properties of the Human Tympanic Membrane and Its Influence on the Dynamic Behaviour of the Human Hearing System," Applied Bionics and Biomechanics **2018**, 1–12 (2018).

227. L. C. Lobato, S. Paul, and J. A. Cordioli, "Statistical analysis of the human middle ear mechanical properties," The Journal of the Acoustical Society of America **151**(3), 2043–2054 (2022).

228. Q. Zhang, Y. Huang, T. Zhang, S. Kubach, L. An, M. Laron, U. Sharma, and R. K. Wang, "Wide-field imaging of retinal vasculature using optical coherence tomography-based microangiography provided by motion tracking," J. Biomed. Opt **20**(6), 066008 (2015).

229. L. M. Wurster, R. N. Shah, F. Placzek, S. Kretschmer, M. Niederleithner, L. Ginner, J. Ensher, M. P. Minneman, E. E. Hoover, H. Zappe, W. Drexler, R. A. Leitgeb, and Ç. Ataman, "Endoscopic optical coherence tomography angiography using a forward imaging piezo scanner probe," J. Biophotonics **12**(4), e201800382 (2019).

230. L. Yao, Y. Zhou, K. Liu, X. Yin, X. Deng, Z. Ding, and P. Li, "Endoscopic OCT Angiography Using Clinical Proximal-End Scanning Catheters," Photonics **9**(5), 329 (2022).

231. V. Vital, G. Psillas, I. Vital, A. Printza, S. Triaridis, and J. Constantinidis, "Ossicular necrosis following head injury," B-ENT **3**(3), 131–134 (2007).

232. G. G. Babighian and S. Albu, "Failures in Stapedotomy for Otosclerosis," Otolaryngol.--head neck surg. **141**(3), 395–400 (2009).

233. E. H. Moreano, M. M. Paparella, D. Zelterman, and M. V. Goycoolea, "Prevalence of facial canal dehiscence and of persistent stapedial artery in the human middle ear: a report of 1000 temporal bones," Laryngoscope **104**(3 Pt 1), 309–320 (1994).

234. H. E. I. Tan, P. L. Santa Maria, P. Wijesinghe, B. Francis Kennedy, B. J. Allardyce, R. H. Eikelboom, M. D. Atlas, and R. J. Dilley, "Optical Coherence Tomography of the Tympanic Membrane and Middle Ear: A Review," Otolaryngol Head Neck Surg **159**(3), 424–438 (2018).

235. "HEARO – Robotic Cochlear Implantation," https://www.cascination.com/en/hearo.

Appendix A   Permissions of Reproduction

This appendix contains permissions for reproduction of the content published in the journal article on which Chapter 5 is based. Additionally, permission for reproduction of borrowed materials found in Fig. 2, Fig. 3, Fig. 4, Fig. 19, Fig. 26, and Fig. 28 are provided.

Appendix A.1   Copyright Permission for Chapter 5

RE: Permission to Reproduce Material for PhD Thesis Chapter

Optica Publishing Group Copyright <copyright@optica.org>
Wed 2023-06-14 11:28 AM
To:Josh Farrell <Josh.Farrell@dal.ca>;Optica Publishing Group Copyright <copyright@optica.org>

CAUTION: The Sender of this email is not from within Dalhousie.

Dear Josh Farrell,

Thank you for contacting Optica Publishing Group.
For the use of material from Joshua D. Farrell, Junzhe Wang, Dan MacDougall, Xiaojie Yang, Kimberly Brewer, Floor Couvreur, Nael Shoman, David P Morris, and Robert B. A. Adamson, "Geometrically accurate real-time volumetric visualization of the middle ear using optical coherence tomography," Biomed. Opt. Express 14, 3152-3171 (2023):

    Because you are the author of the source paper from which you wish to reproduce material, Optica Publishing Group considers your requested use of its copyrighted materials to be permissible within the author rights granted in the Copyright Transfer and Open Access Publishing Agreement submitted by the requester on acceptance for publication of his/her manuscript. It is requested that a complete citation of the original material be included in any publication. This permission assumes that the material was not reproduced from another source when published in the original publication.

    If the entire article is being included, it is permissible to use the **version of record**.

While your publisher should be able to provide additional guidance, we prefer the below citation formats:

    For citations in figure captions:

    [Reprinted/Adapted] with permission from [ref #] © Optica Publishing Group. (Please include the full citation in your reference list)

    For images without captions:

    Journal Vol. #, first page (year published) An example: Opt. Express 19, 2720 (2011)

Please let me know if you have any questions.

Kind Regards,
Hannah Greenwood

Hannah Greenwood
June 14, 2023
Authorized Agent, Optica Publishing Group

OPTICA
PUBLISHING GROUP | Formerly **OSA**

From: Josh Farrell <Josh.Farrell@dal.ca>
Sent: Friday, June 9, 2023 12:22 PM
To: Optica Publishing Group Copyright <copyright@optica.org>
Subject: Permission to Reproduce Material for PhD Thesis Chapter

To whom it may concern,

I am preparing my PhD thesis for submission to the Faculty of Graduate Studies at Dalhousie University, Halifax, Nova Scotia, Canada. I would like your permission to include a manuscript version of the following as a chapter in the thesis:

- **Joshua D. Farrell**, Junzhe Wang, Dan MacDougall, Xiaojie Yang, Kimberly Brewer, Floor Couvreur, Nael Shoman, David P Morris, and Robert B. A. Adamson, "Geometrically accurate real-time volumetric visualization of the middle ear using optical coherence tomography," Biomed. Opt. Express 14, 3152-3171 (2023)

Canadian graduate theses are reproduced by the Library and Archives of Canada through a non-exclusive, world-wide license to reproduce, loan, distribute, or sell theses. I am also seeking your permission for the material described above to be reproduced and distributed by the LAC(NLC). Further details about the LAC(NLC) thesis program are available on the LAC(NLC) website (www.nlc-bnc.ca

> ## Bienvenue au site Web Bibliothèque et Archives Canada / Welcome to the Library and Archives Canada website
>
> Access online resources, services and the collection of Library and Archives Canada: Canadian documentary heritage (publications and records) providing knowledge about Canadian society and serving as the collective memory of the Government of Canada.
>
> www.nlc-bnc.ca

).

Full publication details and a copy of this permission letter will be included my thesis.

Thank you,

Josh Farrell

Appendix A.2   Copyright Permission for Fig. 2

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 31, 2023

This Agreement between Mr. Joshua Farrell ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5559270826069
License date	May 31, 2023
Licensed Content Publisher	Elsevier
Licensed Content Publication	Elsevier Books
Licensed Content Title	The Human Body
Licensed Content Author	Bruce M. Carlson
Licensed Content Date	Jan 1, 2019
Licensed Content Pages	31
Start Page	177
End Page	207
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic

164

Are you the author of this Elsevier chapter? No

Will you be translating?   No

Title   Software Framework for Middle Ear Optical
                                            Coherence Tomography

Institution name   Dalhousie University

Expected presentation date   Jul 2023

Portions   Figure 7.16, page 196

                                            Mr. Joshua Farrell

Requestor Location

Publisher Tax ID   GB 494 6272 12

Total   0.00 USD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection
with completing this licensing transaction, you agree that the following terms and conditions
apply to this transaction (along with the Billing and Payment terms and conditions
established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your
Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to
the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has
appeared in our publication with credit or acknowledgement to another source, permission
must also be sought from that source. If such permission is not obtained then that material
may not be included in your publication/copies. Suitable acknowledgement to the source
must be made, either as a footnote or in a reference list at the end of your publication, as
follows:

165

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk here). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those

166

established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you

168

scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

<u>Elsevier Open Access Terms and Conditions</u>

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0.

169

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.10

Questions? customercare@copyright.com.

170

Appendix A.3   Copyright Permission for Fig. 3 and Fig. 19

Re: Obtain permission request - Journal (1358037) [230531-019125]

Rights and Permissions (ELS) <Permissions@elsevier.com>
Thu 2023-06-01 8:11 AM
To:Josh Farrell <js478190@dal.ca>

CAUTION: The Sender of this email is not from within Dalhousie.

Dear Joshua Farrell,

We hereby grant you permission to reproduce the material detailed below in print and electronic format at no charge subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.

2. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication as follows:

"This article was published in Publication title, Vol number, Author(s), Title of article, Page Nos, Copyright Elsevier (or appropriate Society name) (Year)."

3. This permission is granted for non-exclusive world rights in all languages.

4. Reproduction of this material is granted for the purpose for which permission is hereby given, and includes use in any future editions.

Kind regards,

Kaveri Thakuria
Senior Copyrights Specialist
ELSEVIER | HCM - Health Content Management

Visit Elsevier Permissions

From: Administrator
Date: Wednesday, May 31, 2023 10:56 AM GMT

Dear ,

Thank you for contacting the Permissions Granting Team.

We acknowledge the receipt of your request and we aim to respond within seven business days. Your unique reference number is 230531-019125.

Please avoid changing the subject line of this email when replying to prevent a delay with your query.

Regards,

Permission Granting Team

From:
Date: Wednesday, May 31, 2023 10:56 AM GMT

Submission ID: 1358037
Date: 31 May 2023 11:56am

Name: Mr. Joshua Farrell
Institute/company: Dalhousie University

Email: js478190@dal.ca

Type of Publication: Journal

Title: World Journal of Otorhinolaryngology - Head and Neck Surgery
Auhtors: Pollak, Natasha
Year: 2017
From page: 129
To page: 135
ISSN: 2095-8811, 2589-1081
Volume: 3
Issue: 3
Article title: Endoscopic and minimally-invasive ear surgery: A path to better outcomes

I would like to use: Figure(s)

Quantity of material: 1
I am the author of the Elsevier material: No
Elsevier author is involved in my project: No

In what format will you use the material: Print and Electronic
Translation: No

Proposed use: Reuse in a thesis/dissertation

Material can be extracted: No

Additional Comments / Information: Would like to use Fig. 2 on page 131 for inclusion in my thesis

Appendix A.4　　　Copyright Permission for Fig. 4

June 1, 2023

Dear Joshua Farrell:

The Radiological Society of North America (RSNA®) is pleased to grant you permission to reproduce the following <u>figure in print and electronic format for reuse in thesis,</u> provided you give full credit to the authors of the original publication.

> Figure 5
> Benson J C, Carlson M L, Lane J I. MRI of the Internal Auditory Canal, Labyrinth, and Middle Ear: How We Do It. *Radiology* 2020;297:252-265.

This permission is a one-time, non-exclusive grant for English-language use and is exclusively limited to the usage stated and underlined above. The requestor guarantees to reproduce the material as originally published. Permission is granted under the condition that a full credit line is prominently placed (i.e. author name(s), journal name, copyright year, volume #, inclusive pages and copyright holder).

This permission becomes effective upon receipt of this signed contract. Please sign a copy of this agreement, return a signed copy to me and retain a copy for your files. Thank you for your interest in our publication.

[Print Name]:　　___Joshua Farrell_____

SIGNATURE:　　___██████████_____　Date:　June 2, 2023_____

Sincerely,
LaShundra Carson
Coordinator, Journal Business
Publications

Phone: 630-571-7833
Fax: 630-590-7724
E-mail: *permissions@rsna.org*

Appendix A.4   Copyright Permission for Fig. 26, and Fig. 28

RE: Permission to Reproduce Material for PhD Thesis

Optica Publishing Group Copyright <copyright@optica.org>
Thu 2023-06-08 12:16 PM
To:Josh Farrell <Josh.Farrell@dal.ca>;Optica Publishing Group Copyright <copyright@optica.org>

Dear Josh Farrell,

Thank you for contacting Optica Publishing Group.

For the use of material from Dan MacDougall, Joshua Farrell, Jeremy Brown, Manohar Bance, and Robert Adamson, "Long-range, wide-field swept-source optical coherence tomography with GPU accelerated digital lock-in Doppler vibrography for real-time, in vivo middle ear diagnostics," Biomed. Opt. Express 7, 4621-4635 (2016):

> Because you are the author of the source paper from which you wish to reproduce material, Optica Publishing Group considers your requested use of its copyrighted materials to be permissible within the author rights granted in the Copyright Transfer Agreement submitted by the requester on acceptance for publication of his/her manuscript. If the entire article is being included, it is requested that the **Author Accepted Manuscript** (or preprint) version be the version included within the thesis and that a complete citation of the original material be included in any publication. This permission assumes that the material was not reproduced from another source when published in the original publication.

> The **Author Accepted Manuscript** version is the preprint version of the article that was accepted for publication but not yet prepared and/or formatted by Optica Publishing Group or its vendors.

While your publisher should be able to provide additional guidance, we prefer the below citation formats:

> For citations in figure captions:

> [Reprinted/Adapted] with permission from [ref #] © Optica Publishing Group. (Please include the full citation in your reference list)

> For images without captions:

> Journal Vol. #, first page (year published) An example: Opt. Express 19, 2720 (2011)

Please let me know if you have any questions.

Kind Regards,
Hannah Greenwood

Hannah Greenwood
June 8, 2023
Authorized Agent, Optica Publishing Group

OPTICA
PUBLISHING GROUP | Formerly OSA

To whom it may concern,

I am preparing my PhD thesis for submission to the Faculty of Graduate Studies at Dalhousie University, Halifax, Nova Scotia, Canada. I would like your permission to include the supplementary material present in the following paper as part of my thesis:

- MacDougall, D., **Farrell, J.**, Brown, J., Bance, M. & Adamson, R. Long-range, wide-field swept-source optical coherence tomography with GPU accelerated digital lock-in Doppler vibrography for real-time, in vivo middle ear diagnostics. Biomed Opt Express 7, 4621–4635 (2016).

Canadian graduate theses are reproduced by the Library and Archives of Canada through a non-exclusive, world-wide license to reproduce, loan, distribute, or sell theses. I am also seeking your permission for the material described above to be reproduced and distributed by the LAC(NLC). Further details about the LAC(NLC) thesis program are available on the LAC(NLC) website (www.nlc-bnc.ca).

Full publication details and a copy of this permission letter will be included my thesis.

Thank you,
Josh Farrell

175