Investigation of middle ear anatomy and function with combined video otoscopy-phase sensitive OCT

Jesung Park,[1] Jeffrey T. Cheng,[2] Daniel Ferguson,[1] Gopi Maguluri,[1] Ernest W. Chang,[1] Caitlin Clancy,[1] Daniel J. Lee,[2] and Nicusor Iftimia[1,*]

[1]*Physical Sciences Inc., Andover, MA 01810, USA*
[2]*Eaton-Peabody Laboratories, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA*
[*]*iftimia@psicorp.com*

Abstract: We report the development of a novel otoscopy probe for assessing middle ear anatomy and function. Video imaging and phase-sensitive optical coherence tomography are combined within the same optical path. A sound stimuli channel is incorporated as well to study middle ear function. Thus, besides visualizing the morphology of the middle ear, the vibration amplitude and frequency of the eardrum and ossicles are retrieved as well. Preliminary testing on cadaveric human temporal bone models has demonstrated the capability of this instrument for retrieving middle ear anatomy with micron scale resolution, as well as the vibration of the tympanic membrane and ossicles with sub-nm resolution.

References and links

1. H. Kurokawa and R. L. Goode, "Sound pressure gain produced by the human middle ear," Otolaryngol. Head Neck Surg. **113**(4), 349–355 (1995).
2. S. N. Merchant and J. J. Rosowski, "Acoustic and mechanics of the middle ear," in *Glasscock-Shambaugh Surgery of the Ear*, A.J. Gulya, L.B. Minor, and D. Poe, eds. (People's Medical Publishing House, 2010), pp. 49–72.
3. M. C. Holley, "Keynote review: The auditory system, hearing loss and potential targets for drug development," Drug Discov. Today **10**(19), 1269–1282 (2005).
4. D. W. Chakeres and M. A. Augustyn, "Temporal bone imaging," in *Head and Neck Imaging*, 4th ed., S. A. Sorrentino and B. Gorek, eds. (Mosby's Textbook for Long-Term Care Assistants, 2003), pp. 1093–1108.
5. R. H. Margolis and D. E. Morgan, "Automated pure-tone audiometry: an analysis of capacity, need, and benefit," Am. J. Audiol. **17**(2), 109–113 (2008).
6. J. J. Walker, L. M. Cleveland, J. L. Davis, and J. S. Seales, "Audiometry screening and interpretation," Am. Fam. Physician **87**(1), 41–47 (2013).
7. E. Onusko, "Tympanometry," Am. Fam. Physician **70**(9), 1713–1720 (2004).
8. F. Waridel, "[Use of tympanometry in children]," Rev. Med. Suisse **2**(91), 2881–2883 (2006).
9. J. Shanks and J. Schohet, "Tympanometry in Clinical Practice," in *Handbook of Clinical Audiology,* J. Katz, ed. (Lippincott Williams & Wilkins, 2009), pp. 157–188.
10. H. H. Nakajima, D. V. Pisano, C. Roosli, M. A. Hamade, G. R. Merchant, L. Mahfoud, C. F. Halpin, J. J. Rosowski, and S. N. Merchant, "Comparison of ear-canal reflectance and umbo velocity in patients with conductive hearing loss: a preliminary study," Ear Hear. **33**(1), 35–43 (2012).
11. J. J. Rosowski, H. H. Nakajima, and S. N. Merchant, "Clinical utility of laser-Doppler vibrometer measurements in live normal and pathologic human ears," Ear Hear. **29**(1), 3–19 (2008).
12. J. T. Cheng, A. A. Aarnisalo, E. Harrington, M. S. Hernandez-Montes, C. Furlong, S. N. Merchant, and J. J. Rosowski, "Motion of the surface of the human tympanic membrane measured with stroboscopic holography," Hear. Res. **263**(1-2), 66–77 (2010).
13. L. S. Folio, G. Wollstein, and J. S. Schuman, "Optical coherence tomography: future trends for imaging in glaucoma," Optom. Vis. Sci. **89**(5), E554–E562 (2012).
14. L. Liu, J. A. Gardecki, S. K. Nadkarni, J. D. Toussaint, Y. Yagi, B. E. Bouma, and G. J. Tearney, "Imaging the subcellular structure of human coronary atherosclerosis using micro-optical coherence tomography," Nat. Med. **17**(8), 1010–1014 (2011).
15. A. Sepehr, H. R. Djalilian, J. E. Chang, Z. Chen, and B. J. Wong, "Optical coherence tomography of the cochlea in the porcine model," Laryngoscope **118**(8), 1449–1451 (2008).

16. H. M. Subhash, V. Davila, H. Sun, A. T. Nguyen-Huynh, A. L. Nuttall, and R. K. Wang, "Volumetric in vivo imaging of intracochlear microstructures in mice by high-speed spectral domain optical coherence tomography," J. Biomed. Opt. **15**(3), 036024 (2010).
17. N. H. Cho, J. H. Jang, W. Jung, and J. Kim, "In vivo imaging of middle-ear and inner-ear microstructures of a mouse guided by SD-OCT combined with a surgical microscope," Opt. Express **22**(8), 8985–8995 (2014).
18. 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," Proc. Natl. Acad. Sci. U.S.A. **109**(24), 9529–9534 (2012).
19. 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," Laryngoscope **125**(8), E276–E282 (2015).
20. E. W. Chang, J. T. Cheng, C. Röösli, J. B. Kobler, J. J. Rosowski, and S. H. Yun, "Simultaneous 3D imaging of sound-induced motions of the tympanic membrane and middle ear ossicles," Hear. Res. **304**, 49–56 (2013).
21. H. Y. Lee, P. D. Raphael, J. Park, A. K. Ellerbee, B. E. Applegate, and J. S. Oghalai, "Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea," Proc. Natl. Acad. Sci. U.S.A. **112**(10), 3128–3133 (2015).
22. H. M. Subhash, A. Nguyen-Huynh, R. K. Wang, S. L. Jacques, N. Choudhury, and A. L. Nuttall, "Feasibility of spectral-domain phase-sensitive optical coherence tomography for middle ear vibrometry," J. Biomed. Opt. **17**(6), 060505 (2012).
23. M. A. Choma, A. K. Ellerbee, C. Yang, T. L. Creazzo, and J. A. Izatt, "Spectral-domain phase microscopy," Opt. Lett. **30**(10), 1162–1164 (2005).
24. S. S. Gao, P. D. Raphael, R. Wang, J. Park, A. Xia, B. E. Applegate, and J. S. Oghalai, "In vivo vibrometry inside the apex of the mouse cochlea using spectral domain optical coherence tomography," Biomed. Opt. Express **4**(2), 230–240 (2013).
25. S. Van der Jeught, J. J. Dirckx, J. R. Aerts, A. Bradu, A. G. Podoleanu, and J. A. Buytaert, "Full-field thickness distribution of human tympanic membrane obtained with optical coherence tomography," J. Assoc. Res. Otolaryngol. **14**(4), 483–494 (2013).
26. J. A. Beyea, S. A. Rohani, H. M. Ladak, and S. K. Agrawal, "Laser Doppler vibrometry measurements of human cadaveric tympanic membrane vibration," J. Otolaryngol. Head Neck Surg. **42**(1), 17 (2013).
27. R. K. Manapuram, V. G. R. Manne, and K. V. Larin, "Phase-sensitive swept source optical coherence tomography for imaging and quantifying of microbubbles in clear and scattering media," J. Appl. Phys. **105**(10), 102040 (2009).
28. J. Park, E. F. Carbajal, X. Chen, J. S. Oghalai, and B. E. Applegate, "Phase-sensitive optical coherence tomography using an Vernier-tuned distributed Bragg reflector swept laser in the mouse middle ear," Opt. Lett. **39**(21), 6233–6236 (2014).
29. W. Choi, B. Potsaid, V. Jayaraman, B. Baumann, I. Grulkowski, J. J. Liu, C. D. Lu, A. E. Cable, D. Huang, J. S. Duker, and J. G. Fujimoto, "Phase-sensitive swept-source optical coherence tomography imaging of the human retina with a vertical cavity surface-emitting laser light source," Opt. Lett. **38**(3), 338–340 (2013).

1. Introduction

Over approximately 12 million adults worldwide suffer from moderate to severe conductive hearing loss [1–3]. The diagnosis of hearing loss causes is somewhat problematic since there is no direct, nonsurgical access, to examine the middle ear pathology. Transcanal otoscopic examination with visible light illumination is conventionally used at the primary care offices to determine if there is any potential infection or perforation of the tympanic membrane (TM). Unfortunately, although the TM is semi-transparent, the middle ear ossicles (malleus, incus and stapes) behind the intact TM cannot be properly observed to determine if they are intact and if the motion from the TM is transmitted to every single ossicle. Computed tomography (CT) and magnetic resonance imaging (MRI) can be used to determine eventual morphological changes in the middle ear and diagnose some diseases, such as semicircular canal dehiscence and cholesteatoma. However, they do not provide sufficient resolution to observe minor changes in the ossicular pathology [4]. Furthermore, CT use is limited due the risk of radiation overdose. In addition, both CT and MR exams are expensive.

Audiometry and tympanometry are the most commonly used clinical tests to quantify hearing loss, but are of limited diagnostic value. These examinations measure the reflection of sound from the middle ear. Audiometry can evaluate the degree of conductive hearing loss based on air-bone conduction. Differences within sound pressure levels (SPLs) on the order of about 10 dB are detectable [5, 6]. However, the audiogram interpretation is somewhat subjective and highly dependent upon the cooperation of the subject.

Tympanometry, is another techniques used for the diagnosis of conductive hearing loss. It allows for testing the TM and the conduction of the energy by the middle ear bones by measuring variations of air pressure in the ear canal. Using the tympanometry, the

audiologists can diagnose some middle ear diseases including otitis media, perforation of the TM and problems with the Eustachian tube [7,8]. However, tympanometry has limited sensitivity in detecting various middle ear diseases (i.e., many patients with middle ear diseases have normal tympanogram) [9,10].

Laser Doppler Vibrometry (LDV) and laser holography are used as well to quantify the functionality of the TM and ossicles. However, since they require direct access to the middle ear bones, a surgical incision in the TM is needed for the examination. Both LDV and laser holography are usually combined with audiometry to enable a more complete assessment of the hearing sensitivity [11, 12].

More recently, optical coherence tomography (OCT), which is a noninvasive optical imaging modality capable of imaging sub-surface biological tissues with a high axial resolution (~5-15 um), has been investigated for diagnosing middle ear pathology. OCT has been clinically used for early diagnosis of various diseases including ophthalmological and coronary artery diseases [13, 14]. In recent years there has been an increased interest in using OCT for imaging middle ear morphology and function. Conventional OCT imaging has been investigated for imaging structures of middle and/or inner ear on animal models [15–17]. A pilot study with a handheld OCT scanner has been recently reported [18], where the thickness of the TM was investigated as a diagnostic indicator for otitis media [19]. However, the functional characteristics of the middle ear structures cannot be analyzed with conventional OCT. A newer variant of OCT, called phase-sensitive OCT, has been used by some research groups both in animal models, such as mice and chinchillas, to measure the vibration of the middle ear TM and the ossicles [20], as well as to assess the functionality of the inner ear (e.g. the cochlea) [21]. The vibrational amplitudes of the malleus and the incus through the TM in cadaveric human ears have also been measured with phase sensitive when a 0.5 kHz sound stimulus was used [22]. Although quite preliminary, these studies suggest the feasibility of using phase-sensitive OCT for clinical diagnosis of hearing loss. However, a multifunctional clinical otoscope with integrated phase-sensitive OCT, video imaging, and sound stimulus channel has not been yet reported.

Here we report the development of a novel instrument that might be used to investigate middle ear morphology and function. The imaging probe of this instrument combines video imaging with phase-sensitive OCT (PsOCT) within the same optical path to simultaneously acquire *enface* video images of the TM and micron-scale cross-sectional structural images of the TM and ossicles. It can also be used in the vibrometry mode to measure the vibration of the TM and ossicles with sub-nanometer resolution. The instrument was calibrated on phantoms and preliminarily evaluated on cadaveric human temporal bone models. Most of the middle ear structures were visualized by OCT, and their vibratory response to a sound stimulus was tested, demonstrating instrument capability to detect very small vibrations with sub-nanometer resolution. These initial results indicate the potential use of this instrument for monitoring middle ear morphology and function.

2. Materials and methods

2.1. Instrument description

A dual modality imaging instrument using a miniaturized probe, which integrates PsOCT and high speed video-otscopy (VO) capabilities within the same optical path was developed. The simplified schematic of the instrument, as well as the Solidworks design of the probe are shown in Fig. 1. The instrument contains 3 subsystems: PsOCT unit, PsOCT/VO probe, and data acquisition and processing unit. The PsOCT subsystem is based on the spectral domain (SD) approach. A 1310 nm superluminescent diode (IPSDD1303 InPhenix, Inc.) with a 3dB bandwidth of 65 nm and an average power of 7 mW serves as the OCT light source. This bandwidth enables an axial resolution of about 12 μm. The near-infrared light is sent through a fiber circulator to a Michelson interferometer based on a single mode coupler with a 90:10 splitting ratio. Thus, 90% of the light goes to the imaging probe and 10% to an optical delay line (ODL). The light beam going to the imaging probe is scanned by two galvanometers

(6200H, Cambridge Technology) to generate cross-sectional and volumetric OCT images. The interference fringes generated by the reflected light from both arms of the fiber interferometer are sent to a commercial spectrometer (Cobra, Wasatch Photonics) that uses a high-speed line-scan camera (2048L, Sensors Unlimited) with a 90 kHz line-scan rate and 2048 pixel elements. This spectrometer provides an imaging range of about 4.5 mm.
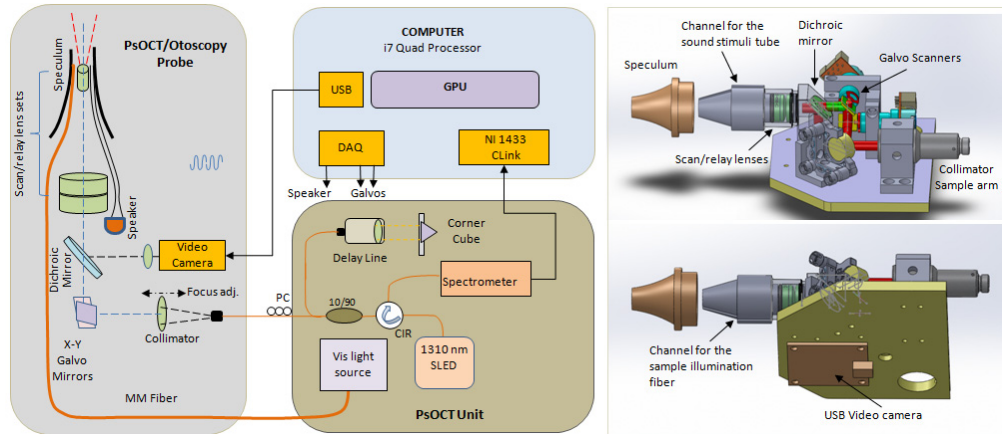

Fig. 1. Instrument simplified schematic and probe Solidworks design.

A frame grabber digitizes and transfers the interference fringes to a graphical processing unit (GPU), where the typical OCT processing steps (background subtraction, interpolation, dispersion compensation, FFT) are performed. Depth-resolved structural images are generated in real-time and displayed with a rate up to 90 frames/second for 1000 A-lines per frame.

The imaging probe was designed to enable generation of high resolution OCT images over an extended imaging range, up to 8 mm. This range is needed to accommodate for the average distance between the TM and the stapes footplate, which is about 6-8 mm. The extended imaging range was possible by using a custom designed optical system consisting of a set of relay lenses and an objective with a long depth of focus (see Zemax design in Fig. 2).
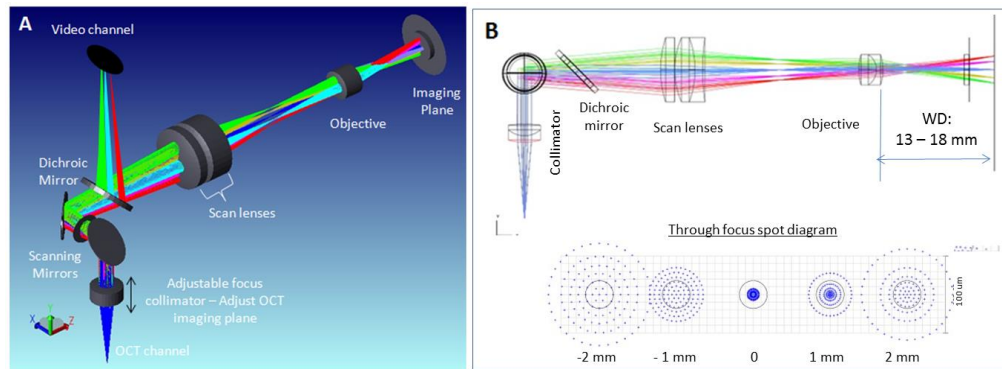

Fig. 2. Zemax design of the hand-held probe. (**A**): 3D representation of both OCT and video channels; (**B**): Zemax design of the OCT channel showing the through focus imaging performance over 4 mm imaging range.

As observed (see Fig. 2(A)), the position of the imaging plane is simply adjusted by modifying the collimation parameters of the OCT sample arm imaging beam. For a given collimation, a lateral resolution of about 15 μm is obtained in the focal plane, and it slightly degrades towards the ends of the 4.5 mm imaging range (see Fig. 2(B)). Then, by changing the collimation parameters in the sample arm of the instrument, the imaging plane can be

moved up to 4 mm away, without affecting the lateral resolution. However, since the spectrometer has an imaging range of only 4.5 mm, the reference arm of the interferometer has to be adjusted as well to keep the image of the sample within the imaging range. A longer imaging range spectrometer could be designed if a camera with a larger number of pixels would be available. However, being limited by the 2048 pixels currently available cameras, a longer imaging range can be obtained if a narrower spectrum light source is used. Unfortunately, this approach would negatively impact the axial resolution of the instrument. Therefore, our approach of refocusing the probe has the benefit that both the axial and the lateral imaging resolution are kept tight.

The video channel shares most of the optical elements with the OCT channel. A dichroic mirror is used to separate the two optical paths at the entrance of the OCT scanning lens system, and the image is focused on a CMOS array (Model UI-1410SE 480x 640 pixels, 1st Vision) using an achromatic lens. Sample illumination is provided by a MM fiber (200 μm core), which is connected to a halogen light source. A sound stimulus cannel is incorporated as well within the probe. A miniature speaker is excited by a computer generated waveform (1-8 kHz) and a flexible micro-capillary tube transports the sound to the probe distal end.

Photographs of the imaging instrument and imaging probe are shown in Fig. 3. The instrumentation is enclosed within a 19" rack, while the probe is enclosed within a 3D printed case, about 1.5"x 3"x2.5" in size. As observed, the probe is designed to use a disposable speculum. Further miniaturization of this probe is possible by replacing the Cambridge Technology scanners with a dual axis MEMS scanner.
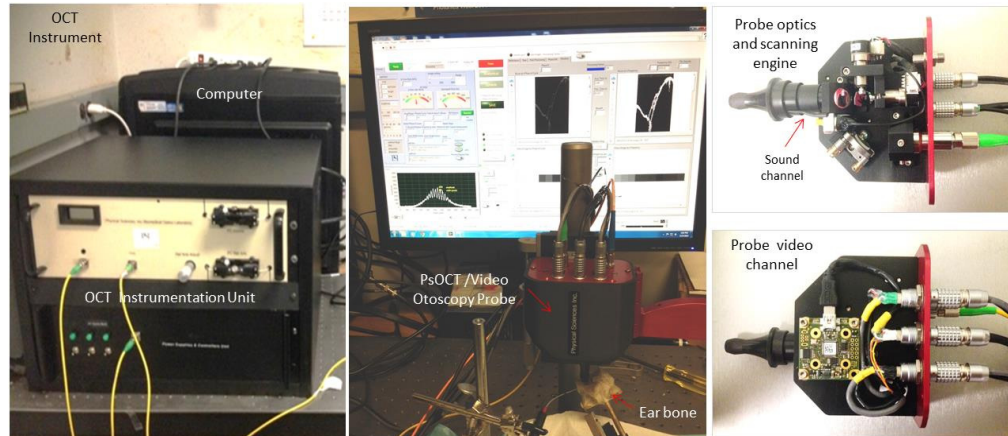

Fig. 3. Photographs of the instrumentation unit, ear bone testing setup, and imaging probe.

2.2 Data processing

The theoretical derivation and practical considerations for the OCT parameters required for vibration measurements using a phase-sensitive scheme have been previously described [23, 24]. The vibration of a sample at a point of interest can be measured by recording OCT M-scans (multiple A-scans recorded at the same physical location as a function of time), while simultaneously providing a pure-tone sound stimulation. The FFT of the OCT signal enables the extraction of the structural reflectivity of the sample, while the FFT of the M-scan phase signal enables the extraction of the vibration magnitude and phase data. The vibration magnitude is obtained in radians (rad) and then is converted to nanometers (nm) by multiplying the magnitude data with $\lambda_o/(4n\pi)$, where λ_o is the center wavelength of the laser light source and n is the sample refractive index. The refractive index of the tympanic membrane was assumed to be 1.44 [25].

2.3 Instrument optimization

The instrument was optimized to enable nm-scale measurements of the TM and middle ear ossicles vibrations. The optimization of phase signal-to noise-ratio (SNR) was the main focus of this task. As expected, phase SNR increases with the number of recorded A-lines within a M-scan. However, the increase in the size of the M-scan also increases the measurement time, which is a critical parameter in a clinical setting. Therefore, a balance was identified to fulfill both requirements. An increase in the camera speed can reduce the dwell time/voxel, but also reduces the number of photons that can be collected. We optimized the instrument to provide over 30 dB phase SNR when using 250 Alines/M-scan and an imaging speed of 75 Klines/sec, which corresponds to 3 ms measurement time per voxel. Thus, a 2D vibrogram of 200 points can be generated in 0.7 sec.

The optimal frequency range for the sound stimuli was evaluated as well. The frequency response of the speaker was measured through the 0 to 8 kHz frequency range, which is typical for audiology tests. The goal was to characterize the relationship between the intensity of the excitation signal and the sound pressure level (SPL) generated by the speaker for multiple frequencies. A circular film with a thickness of 0.05 mm and a diameter of 1/2" was prepared to mimic the size of the adult TM, which has a thickness on the order of 0.05 to 0.12 mm and diameter of on the order of 8 to 10 mm [25]. As the surface of the film is covered with the acrylate adhesive, the film was self-adhesively tightened on a half inch-diameter optical mount. Continuous single tones at 0.5, 1, 2, 4, 6, and 8 kHz and four different sound intensities were produced by a speaker to induce the vibration of the sample. A pre-calibrated microphone (ER 7C Probe Microphone Series B, Etymotic Research) was used to monitor the SPL near the surface of the sample. The tip of the microphone probe was positioned near the edge of the sample. A time-dependent waveform of the sound pressure was captured for each stimulus frequency and intensity, and a fast Fourier transform was used to determine the SPL. Next, vibration measurements of the film sample were recorded varying the M-scan size (between 100 and 10000 A-lines) and using pure-tone sinusoidal sound stimuli between 0.5 and 8 kHz, for sound intensities of 50, 60 and 70 dB SPL, with the goal of determining the optimal parameters of the instrument to be used as a reference for the ear bone study. The vibration peak was determined from the averaged M-scan data for each stimulus frequency and SPL. The signal-to-noise ratio (SNR) of each measurement was determined as a ratio between the vibration peak and the mean of the noise over the entire frequency range (0 to 20 kHz). The phase noise figure (phase noise amplitude as a function of frequency) was then calculated. The same measurements were performed for a swept source based setup, previously built by us using a commercially available swept source (SS) laser (Axsun Technologies, Billerica, MA), with the goal of determining if a longer imaging range SS approach can be used instead of the spectrometer-based approach. The Axsun laser has a center wavelength of 1310 nm and a sweeping frequency of 40 kHz. The phase noise was computed in identical conditions for both instruments using 250 A-lines/M-scan, which were acquired from a reflective stationary target (mirror). As observed from Figs. 4(A) and 4(B), the phase noise of the SDOCT setup was two orders of magnitude lower than that of the SSOCT setup. The phase noise spectrum of the swept-source OCT also showed high random spikes, suggesting the additional contribution of the electronic noise. Therefore, we concluded that the spectral domain (SD) approach is currently more suitable for this application.
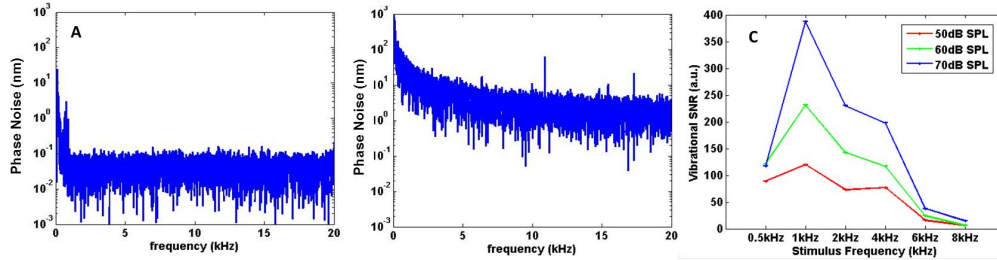
Fig. 4. Phase noise and SNR measurements on phantom consisting of a pellicle of thin cellulose acetate film. A,B: Comparison of the phase noise figures as a function of frequency for the SDOCT setup and for the SSOCT setup. C: SDOCT vibrational SNR for 3 different sound pressures within the 0-8 kHz frequency excitation range and 250Alines/M scan.

We also measured the variation of the SNR of the SDOCT setup with the stimulus frequency for several sound pressures. The summary of the measurements is shown in the graphs from Fig. 4(C). As observed, the phase noise was over 20 dB at very low frequencies (0-500Hz) for the SD setup and over 60dB for the SS setup. This noise is mainly caused by the background sources (computer fans, vibrations in the room, etc). Therefore, for practical reasons the excitation frequency of the sound stimuli in this study was chosen outside of this range. On the other hand, as observed from the roll off of the SNR, higher frequencies of the stimulus signal result in lower SNR. This is the result of the convolution between the speaker frequency response and the fringe wash of the OCT signal due to the reduced dwell time on a stable structure on the sample. As a result, we used the 1 to 6 kHz range for the ear bone study. Higher frequencies are measurable, but more averaging (higher number of A-lines/M scan) will be needed. The measured sensitivity of the instrument was 0.5nm (calculated for an SNR = 10). This correlate well with sensitivities reported by other groups: 4.8 mrad, reported by Chang et al. [20] and 7.5 mrad, reported by Subhash et al. [22].

3. Results

The instrument was first tested on a thin film phantom (cellulose acetate film). The impact of the potentially induced motion artifacts (operator handshaking in this case) was preliminarily investigated. Measurements were first performed with the imaging probe sitting on the optical table, and then with the probe being held by the operator (see photographs in Fig. 5).

The 1st set of experiments were performed with a stationary beam (short duration point measurements), while using a stimulus signal of 5kHz with a sound pressure level (SPL) of 50. These values for SPL and excitation frequency chosen with purpose of testing probe behavior for a suboptimal response of the speaker, when the vibration SNR is fairly low (see Fig. 4(C)). M-scans of 500 A-lines were acquired and processed to extract the phase signal. The amplitude of the signal relative to the noise floor was measured in both cases: stationary probe sitting on the optical table and hand-held probe. As observed from Fig. 6, the phase noise floor was basically the same for both cases within the 1.5 to 20 kHz range and significantly higher within the 0 to 1.5 kHz range for the hand-held case, while the amplitude of the vibration was basically the same in both cases. This confirms the fact that the low frequency artifacts induced by hand shaking will not affect the measurements when stimuli signals within the 1.5 kHz to 20 kHz range will be used.

The 2nd set of measurements was performed with a scanning beam to collect 200 M-scans across sample surface. Thus, 200 points were investigated, spanning a distance of 6 mm (30 microns separation between M-scans). The stimulus signal was first turned off to record a phase noise vibrogram, and then turned on. Again, a 5 kHz stimulus with a 50 dB SPL was used in this experiment. The results are of this experiment are presented in Fig. 7.
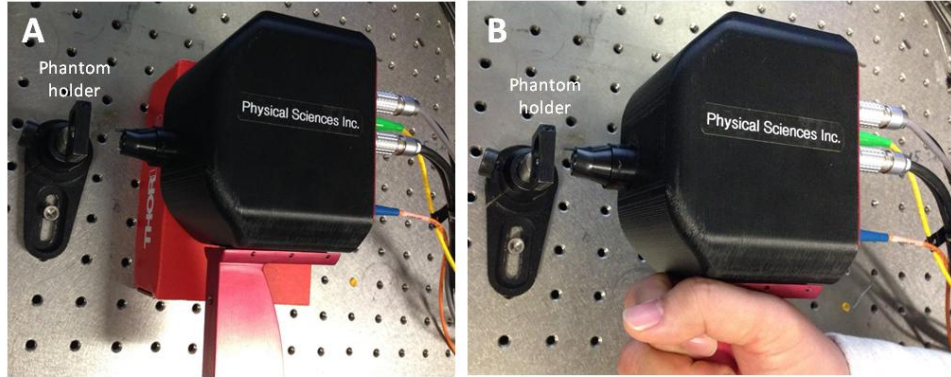
Fig. 5. Photographs showing the preliminary testing of the probe on a ½" pellicle of cellulose acetate film, 0.05 mm in thickness. A- Probe resting on the optical table; B: Probe held by a study investigator.
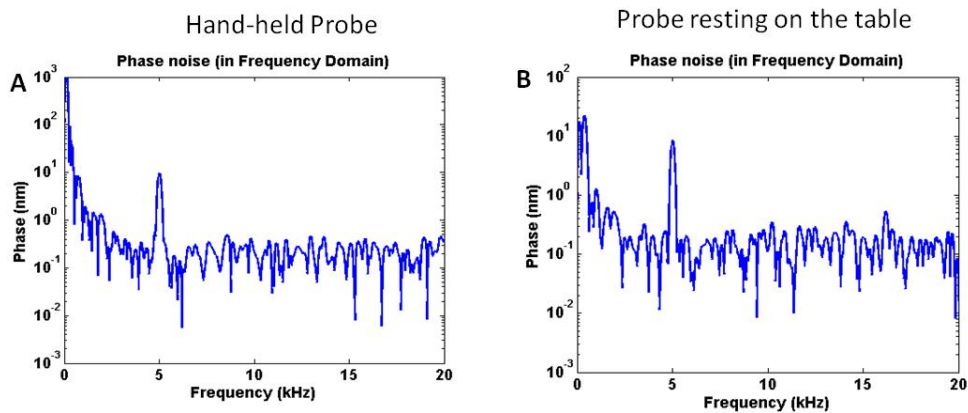

Fig. 6. Phase measurements on a ½" pellicle of cellulose acetate film, 0.05 mm in thickness when a 50dB sound stimulus with a frequency of 5 kHz was applied. A- hand held; B: Probe resting on the table.

As it can be observed from Fig. 7, there were no differences between the phase noise levels for the two cases. Similarly, the vibrograms of the phantom pellicle showed very similar amplitude levels (within 5 to 7 nm scale). Since the measurement took ~1.4 sec for a 500 A-lines M scan (larger M-scan was used on purpose to determine the impact of the handshaking), slightly higher oscillations in the shape of the vibrogram are observed in the hand-held case than in the stationary case. However, as expected, these low frequency oscillations did not impact the measurement of the vibration amplitude.
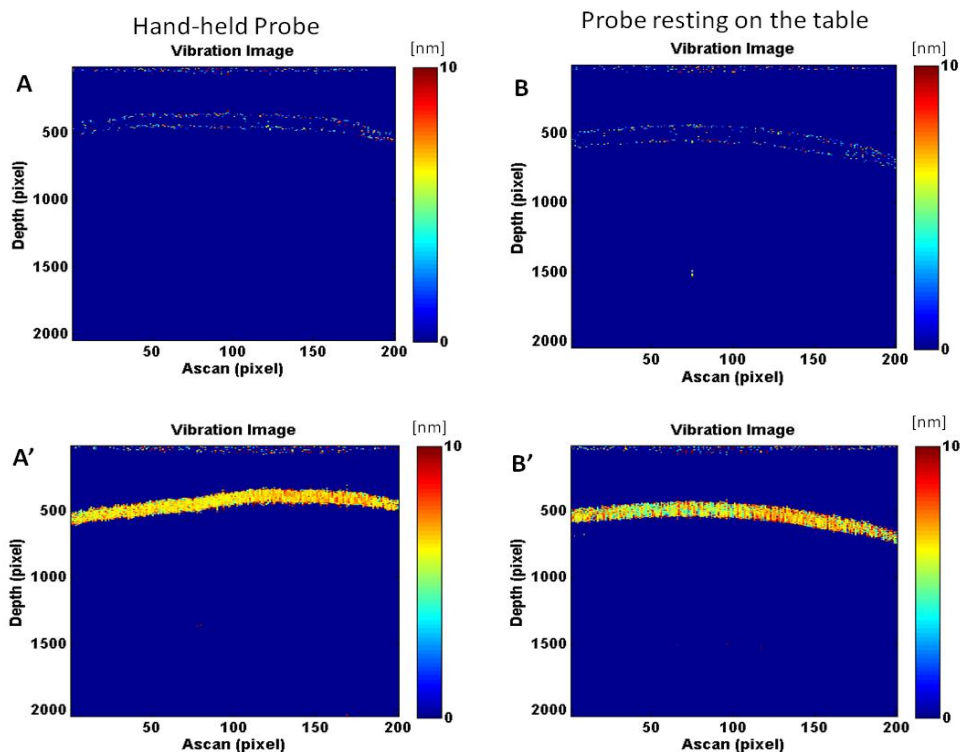
Fig. 7. Vibrational images of a ½" pellicle of cellulose acetate film, 0.05 mm in thickness. A: Hand-held case without sound stimuli applied; A': Hand –held case with a 50dB/5 kHz sound stimulus applied. B: Probe resting case without sound stimuli applied; B': Probe resting case with a 50dB/5 kHz sound stimulus applied.

Instrument testing continued with pre-clinical investigation on 5 fresh human temporal bones that were obtained from Massachusetts Eye and Ear Infirmary (MEEI). The specimens were excised from donors with a clear history of no otologic diseases. The tympanic membranes were intact and still fairly transparent. These temporal bones had the cartilaginous and bony ear canal removed to expose the majority of the TM surface for combined PsOCT/VO view of the TM, as well as a facial recess opening of the middle ear to check the normality of the ossicular bones. The temporal bone was tightly attached to a rod (see Fig. 2) with the rod's long axis roughly perpendicular to the plane of the tympanic ring. The rod was then held by a 3-axis manipulator and the temporal bone was positioned with the TM surface perpendicular to the OCT laser beam.

VO enabled proper orientation of the specimen, such that the TM and the ossicles can be reached by the OCT beam. Since the imaging range of the OCT spectrometer is about 4.5 mm and the distance between the TM and the stapes is usually over 5 mm, the anatomy of all three ossicles could not be resolved without readjusting the focal position of the imaging probe. Therefore, a first set of images was collected with the focal plan positioned in the close proximity of the TM, and then a 2nd set of images was collected after pushing the focal plane ~3 mm deeper, such that the Stapes-Incus joint was brought within the imaging range. The bone was also rotated to enable the imaging of the stapes. The images collected from two different focal planes and two different angular orientations of the probe were manually stitched together (translation + rotation) to display the anatomy of all three ossicles, in a similar manner with that shown in the cartoon from Fig. 8(A), which serves for better understanding of the OCT visualized anatomy. The stitched OCT images are shown in Fig. 8(B). As observed, the head of the malleus and the body of the incus were not well recovered. This is in part caused by the obscuration of the incus body by the malleus head by the ear

bone, and by the limited OCT penetration depth. Although OCT penetration depth is typically 1.5 mm, these bonny structures are highly scattering and thus very few photons can reach the body of the incus. Fortunately, the Incus-Stapes joint and footplate of the stapes, which are important in the evaluation of middle ear transfer function as an input to the cochlea, were clearly seen by the OCT.
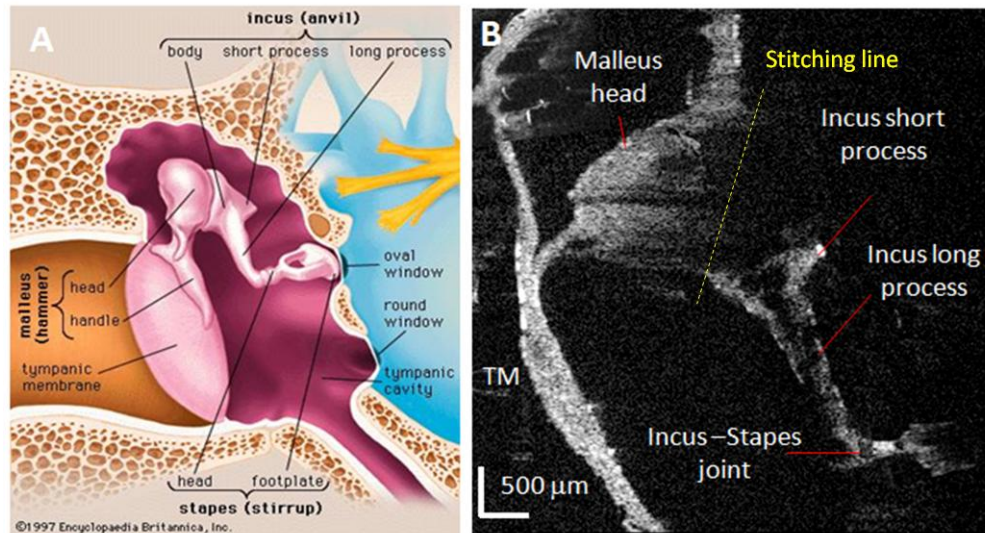

Fig. 8. A: Cartoon showing the anatomy of the middle ear ossicles [Source: http://www.britannica.com/science/middle-ear]. B: Cross-sectional OCT image of the middle ear. Two images obtained for two focal positions and slightly different angular orientations were stitched to display the morphology of all three ossciles within the same image.

In addition to micron-scale structural imaging, the advantage of using OCT for diagnosing middle ear pathology is that it also enables functional imaging. To test this capability, we turned on the sound excitation channel of the probe and switched the software in the PsOCT mode. Instead of collecting cross-sectional B-scans, the beam was successively moved over a number of points across the TM to acquire an M-scan of 250 A-lines for each location. The M-scans were then processed to build a vibrogram, which is a 2-dimmensional representation of the vibration amplitude of each pixel from a sparse cross-sectional image. The measurements were repeated with different sound stimuli with known frequencies (1, 2, 4 and 6 kHz) and sound pressure levels (SPLs) of 50, 60, 70 and 80 dB. Cross-sectional vibration data sets of the middle ear were acquired with 40 μm lateral sampling (200 points over a 8 mm scan). The total acquisition time was 0.7 seconds for a cross-sectional scan. The processing of the vibrational data has required depth-dependent calculations on all depth pixels, which became computationally intensive.

A representative example of the vibrograms recorded by us from the 5 examined middle-ear bones, as well as a summary of the vibration responses for various sound pressures and excitation frequencies are presented in Fig. 9. A cartoon showing the directions of the OCT imaging beams with respect to the anatomy of the middle ear is oriented in Fig. 9(A), while a video frame of the TM is presented in Fig. 9(B) to indicate the angular orientations of the OCT scan with respect to the main body of the malleus, which can be partially visualized through the TM.
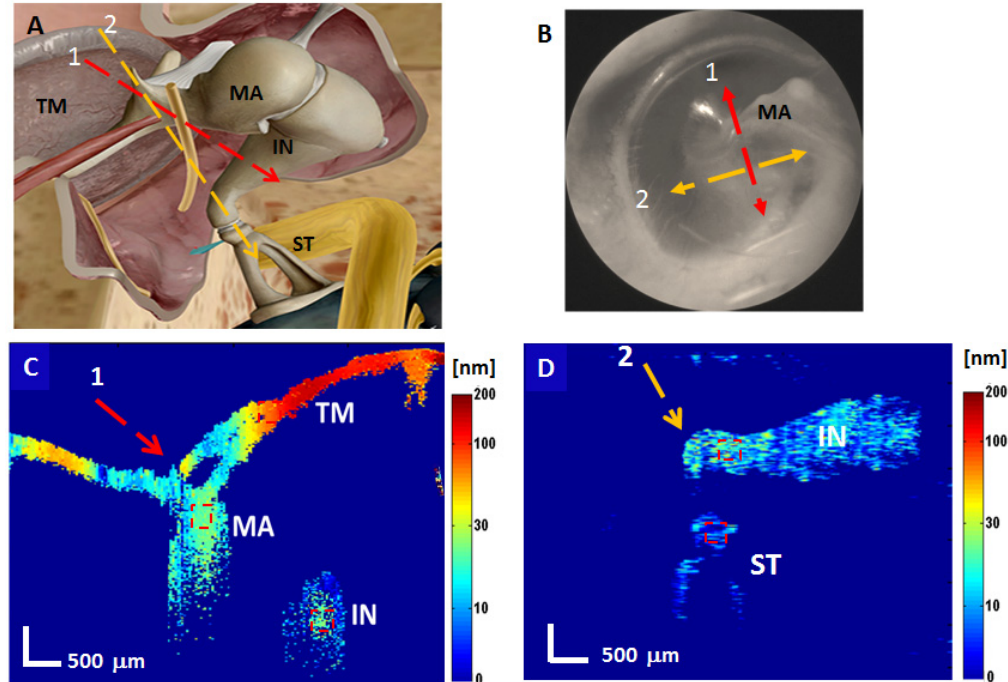
Fig. 9. Vibrational images of the human middle ear. (A): Cartoon showing the scanning direction with respect to the position of the TM and ossicles [Source: http://info.visiblebody.com/bid/323583/Five-Cool-Facts-about-the-Middle-and-Inner-Ear]. (B): Video image frame of the TM- showing OCT scanning directions. (C,D): Vibrograms acquired at two different locations for a SP of 70 dB and a stimulus frequency of 4 kHz.;. Legend: TM-typmanic membrane; MA- malleus; IN- incus, ST- stapes.

Vibrograms of the middle ear corresponding to the two imaging sequences are presented in Figs. 9(C) and 9(D). The vibrations of the structural areas at the stimulus frequency were visualized as pseudo-color images. Interpolation was performed to smooth the images. As shown in Fig. 9(A), it is almost impossible to get all three ossicles into a cross-sectional vibrogram due to the complex three-dimensional anatomy of the middle ear. The location of the malleus limits the visualization of the incus, while the deep location of the stapes makes it difficult to record its vibration when the imaging range is limited to a few mm. As shown in Fig. 9(C), the 1st imaging location allowed the visualization of the TM and a fraction (the "long process") of the incus. In order to visualize the incus/stapes joint, the imaging plane was slightly shifted to a higher depth (see Fig. 9(D)) and the scan direction was slightly shifted and rotated with almost 90 deg. The reference arm of the spectrometer was re-adjusted as well to bring the image within the range of the spectrometer.

The vibrations of the TM and the ossicles were also quantitatively analyzed to observe the frequency response of the TM and the ossicles, as well as the dependence between the vibration amplitude and the sound pressure. These measurements were performed on all 5 specimens, while successively modifying either the sound pressure or the frequency of the sound stimuli. Regions of interest (20x20 pixels) from the central area of each ossicles (see red-dotted areas in Fig. 9), as well of the TM, were used and the average value of the vibration was considered for each ROI. Then the averaged values for all 5 specimens were plotted as either a function of the frequency of the sound stimuli for a given sound pressure (see Fig. 10(A)), or as a function of the sound pressure for a given frequency of the sound stimuli (see Fig. 10(B)). As it can be observed, the standard deviation (SD) for the incus and stapes was quite large at lower frequencies and sound pressures, suggesting that the SNR was not optimal for these cases. Furthermore, the vibration amplitude of the stapes was close to 1nm in the 1 to 2 KHz range at sound pressures lower than 60 dB, which is very close to the

measurement resolution (0.5 nm); therefore, this range was excluded for the measurements on stapes.
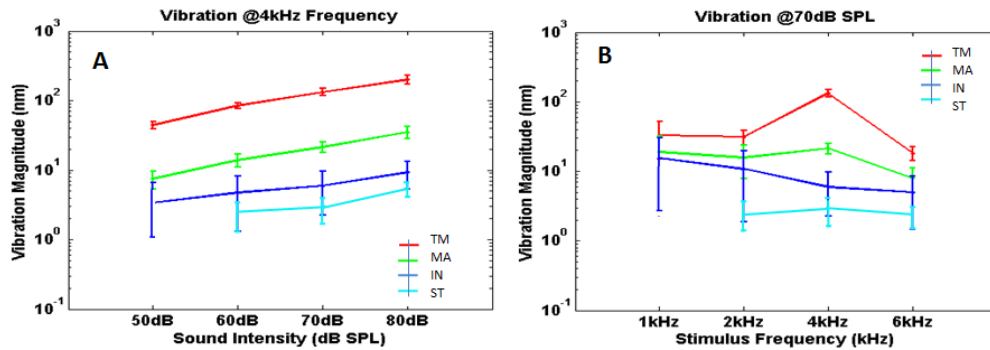

Fig. 10. Summary of the ear bone measurements. (A): The magnitude of the vibration as a function of the SPL for a 4 kHz sound stimuli; (B): The magnitude of the vibration as a function of the frequency of the sound stimuli for a 70 dB SPL.

Another important observation, is that the best behavior was obtained in general for a 4kHz frequency of the sound stimuli, which might be favored by a potential enhanced response of the mechanical structure of the ossicles in our specimens. Therefore, in our analysis for the vibration response to various sound pressures, a 4kHz sound stimulus was used (see Fig. 10(A)). The vibration magnitude of the TM was 132.6 ± 15.9 nm, higher than those of the malleus (21.4 ± 3.8 nm), incus (6.0 ± 3.8 nm), and stapes 2.1 nm ± 0.8 nm. Although the frequency responses were generated for only four stimulus frequencies, the patterns of the frequency response in the TM, malleus and incus were in good agreement with another study performed using laser Doppler vibrometry [26]. The study reported that the frequency responses of the TM and the malleus in the human cadaveric middle ear has a valley between 1 kHz and 2 kHz, and gradually increases towards 4- 5 kHz and then decreases quite rapidly with the increase of the frequency. The vibration magnitudes of the TM, malleus and incus increased linear with the sound intensity for a stimulus frequency of 4 kHz.

4. Discussion and conclusion

We have demonstrated a novel otoscopy probe that can provide simultaneous PsOCT and video measurements of the middle ear. Besides visualizing the TM, video otoscopy enables proper placement of the probe within the ear canal, such that the middle ear ossicles can be properly investigated with PsOCT. Our preliminary study on ear bones has clearly demonstrated the feasibility of this dual modality imaging approach for monitoring middle ear morphology and function.

This study also showed some shortcomings of our current technology, which need to be addressed to facilitate clinical transition. One of the limitations is the difficulty in assessing the head of the malleus and the malleus-incus joint. These anatomical parts are partially obscured by the ear bone, and thus a forward looking probe cannot reach them. Therefore, we are already considering the development of a probe with a specially designed objective that can tilt the imaging plane with about 10-15 deg., making it possible to visualize a larger part of the malleus and incus.

Another limitation is the relatively short imaging range, compared to the anatomy that has to be resolved. A longer imaging range is possible by using a swept-source (SS) approach instead of the SD approach. However, vibration measurements on nm scale are challenging and therefore require the use of an instrument with a very low phase noise. Although the SS-based OCT has the advantage of providing a long imaging range, its phase instability is currently a major limiting factor in achieving accurate measurements of the ossicle vibrations. Swept source/Fourier domain OCT instruments suffers from sampling and A-scan trigger

jitter, which makes phase-sensitive measurements more challenging. A phase noise up to π radian was measured in such instruments [27]. Various approaches (e.g. use of Mach-Zehnder interferometer (MZI) and fiber Bragg grating) have been developed to calibrate the sweeping nonlinearity and inter-sweep variability and thus to improve phase noise. However, they require complicated schemes and thus significantly increase the cost of the instrument [28]. More recently, newer swept-source lasers including vertical external-cavity surface-emitting laser (VECSEL) and Vernier-tuned distributed Bragg reflector laser (VT-DBRL) have shown improved phase stability. Unfortunately, their cost is still prohibitive [28, 29]. Our approach to sequentially choose different imaging planes addresses this problem to some extent. The current inconvenience is that the reference arm of the interferometer has to be adjusted as well after adjusting the collimation of the OCT sample arm beam. Therefore, our new approach will be to integrate the reference arm of the OCT interferometer within the probe and use a micromotor that simultaneously adjusts the delay line and the imaging plane. A better approach will be to increase spectrometer imaging range. However, currently available InGaAs line cameras have only 2048 pixels, and thus the increase of the imaging range will come up with a penalty on the axial imaging resolution. When larger arrays will become available, this will no longer be an issue.

Probe miniaturization to enable easy handling and use of off the shelf specula is another issue that has to be addressed. We are planning to change the bulky scanners with a dual axis MEMS scanner, which will significantly reduce probe size and will enable easier handling. Further optimization of the optical design should also enable the use of smaller optical components.

The time needed to record/display a vibrogram is another aspect of the technology that has to be investigated in more detail. Although recording a vibrogram takes less than 1 sec., approximately 2 minutes are currently needed in Matlab to generate a cross-sectional vibration image. Since the clinician needs immediate feedback to determine if the vibrogram was recorded from the intended area of interest and if contains the desirable information, it is crucial to significantly reduce the processing time and enable real-time display of the results. While the imaging speed can be relatively easily improved using the newly available faster InGaAs linear arrays (GL2048 R boosts the speed of SD-OCT imaging to >147 klps via Medium Camera Link® interfaces), the processing speed will require the use of a parallel graphical processing (GP) architecture. We are already taking preliminary steps for implementing our data processing algorithm in a GEForce GTX 970 GP module, which with 1663 CUDA cores. It is expected that the parallel processing of each M -scan of the vibrogram will significantly reduce the processing time. Furthermore, we also use the cross-sectional OCT images recorded from the exactly same location to segment the main features of interest (TM and ossicles) and eliminate the non-structural areas (e.g. the tympanic cavity) and only process the depth points where a structure was present. This reduces the number of points/A-line that has to be processed, and as a result the processing time was decreased by a factor of two.

In conclusion, although the current technology has some limitations, we believe that it can be substantially improved and potentially used for the diagnosis of middle ear diseases. Its clinical use might help narrowing the differential diagnosis of middle ear pathology by providing clinicians with volumetric anatomical information of the middle ear at rest to assess ossicular chain integrity and to rule out middle ear fluid accumulation (e.g. otitis media), as well as with functional information by offering the capability for monitoring sound induced vibrations of the middle ear ossicles through the intact TM.

Acknowledgments