



In vivo functional imaging of the human middle ear with a hand-held optical coherence tomography device

CHRISTOPHER G. LUI,^{1,3} WIHAN KIM,^{1,3}  JAMES B. DEWEY,¹
FRANK D. MACÍAS-ESCRIVÁ,¹ KUMARA RATNAYAKE,¹ JOHN S.
OGHALAI,¹ AND BRIAN E. APPLGATE^{1,2,*} 

¹Department of Otolaryngology - Head and Neck Surgery, Keck School of Medicine, University of Southern California, 1450 San Pablo Street, Suite 5708, Los Angeles, CA 90033, USA

²Department of Biomedical Engineering, University of Southern California, 1042 Downey Way, Denney Research Center (DRB) 140, Los Angeles, CA 90089, USA

³These authors contributed equally to this work

*brianapp@usc.edu

Abstract: We describe an optical coherence tomography and vibrometry system designed for portable hand-held usage in the otology clinic on awake patients. The system provides clinically relevant point-of-care morphological imaging with 14-44 μm resolution and functional vibratory measures with sub-nanometer sensitivity. We evaluated various new approaches for extracting functional information including a multi-tone stimulus, a continuous chirp stimulus, and alternating air and bone stimulus. We also explored the vibratory response over an area of the tympanic membrane (TM) and generated TM thickness maps. Our results suggest that the system can provide real-time *in vivo* imaging and vibrometry of the ear and could prove useful for investigating otologic pathology in the clinic setting.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Optical coherence tomography (OCT) is a rapid, noninvasive, and non-contact imaging modality originally developed for, and widely used in ocular imaging [1]. The range of applications are growing with clinical devices for imaging the coronary arteries and esophagus as well as preclinical devices for imaging the skin [2,3], lower gastrointestinal tract [4,5], and the ear [6]. Relatively low cost, compact systems can enable point-of-care imaging with high-resolution and can readily be integrated into the clinical workflow.

While the application of OCT vibrometry in otology has primarily been as a tool for hearing research in laboratory animals [7–11], it has several key advantages that make it well suited for use in the clinic. OCT is able to image the entire depth of the middle ear space, allowing detailed imaging and cross-sectional views of the tympanic membrane (TM) and ossicular chain. Compare this to the otoscope, the principal tool for examining the ear used by physicians in the clinic. Otoscopy only allows for visualization of the superficial surface of the TM. Its ability to see into the middle ear is extremely limited due to the limited translucency of the TM especially in diseased ears where the TM can be more opaque due to thickened middle-ear mucosa and formation of bacterial biofilms lining the TM [12]. OCT provides imaging in real-time that can be performed in the clinic by the physician, unlike other imaging modalities such as computed tomography (CT) or magnetic resonance imaging (MRI) which are slow and expensive [13]. Neither CT or MRI scans are routinely conducted in the clinic and are infrequently used as a primary diagnostic tool for middle ear disorders due to their cost and scheduling obstacles that conflict with the rapid workflow of the clinic.

In addition to volumetric imaging of middle ear morphology, OCT is capable of vibratory imaging of the middle ear structures to directly measure middle ear function [14]. This offers additional information beyond the typical functional diagnostic information that is normally derived from an audiological evaluation, which commonly includes air and bone conduction thresholds, speech audiometry, acoustic reflex testing, and impedance audiometry [15]. Other common imaging modalities such as CT and MRI are also unable to provide functional information about vibrations within the middle ear. Compared to other traditional methods of measuring vibrations such as laser doppler vibrometry, OCT is not limited to surface measurements but can make non-invasive vibratory measurements in structures deep to the TM, as well as providing imaging of those structures as well.

Recently, other groups have recognized the value of developing an OCT system designed specifically for usage in human patients in the otology clinic. Cross-sectional imaging of the middle ear using OCT was first demonstrated in 2001 [16], and the technique has since been used for the diagnosis of otitis media [17,18]. Pathologies such as tympanosclerosis, cholesteatoma, dimeric tympanic membrane, and hyperkeratosis can also be reliably visualized [19]. More recently, vibratory measurements of the middle ear have been made in humans [20] and used to discriminate otosclerotic stapes fixation from normal ears in a patient population [21] and to evaluate the success of tympanoplasty [22]. OCT has the potential to be a standard tool for otologists in the clinic, with the ability to both visualize anatomy and diagnose pathology via imaging and vibrometry in real-time.

For an OCT system to be suitable for use in the clinic, the device must be easy to use, capable of both 3D imaging and vibrometry, and the entire system must be small enough to fit into a typical exam room. Here we present a hand-held OCT (HHOCT) system that was developed to assess clinical needs that are currently unmet. Compared to systems previously described including our own [14,23], the current system is designed for rapid non-invasive imaging and functional measurements in awake human patients in a form factor suitable for the clinic setting. The system has a significantly smaller physical footprint to allow for portable hand-held usage that is intuitive for clinicians to use in a manner similar to an otoscope, and that would allow for convenient transportation between clinic rooms as needed within the existing workflow. Unlike other devices, the described HHOCT system functions as a stand-alone system, without requiring attachment to a surgical microscope [14] or other specialized equipment. It is capable of volumetric imaging of the middle ear space including functional vibratory imaging. The sensitivity to vibration has been optimized, allowing for vibration measurements from the tympanic membrane, ossicles, and cochlear promontory with both air and bone sound conduction. All the necessary components for OCT imaging and vibrometry are housed inside the device, including a camera for real-time video otoscopy, and speakers used when measuring sound-evoked vibrations. This technology will allow for a wide range of quantitative measures with rapid visualization of the middle ear space for clinical diagnosis and monitoring in a device designed specifically for the clinic setting.

2. Methods

The components of the HHOCT system are shown in Fig. 1(a). The OCT light source was a swept laser with center wavelength at 1310 nm and 95 nm bandwidth with a 100 kHz sweep rate (Insight Photonic Solution, Inc.) providing 13.7 μm (in tissue, $n=1.3$) axial resolution for OCT imaging. The laser sweep was highly repeatable and linear in wavenumber, making it particularly suitable for phase-sensitive measurements such as conducted here [24]. The laser was directed into a 90:10 optical fiber coupler (fC). 10% of the laser light was directed into an optical delay line (ODL) and then onto one side of a balanced photodiode (BPD). The remaining 90% was directed to a circulator (CIRC) and on to the OCT interferometer housed in the HHOCT unit, shown schematically in Fig. 1(a).

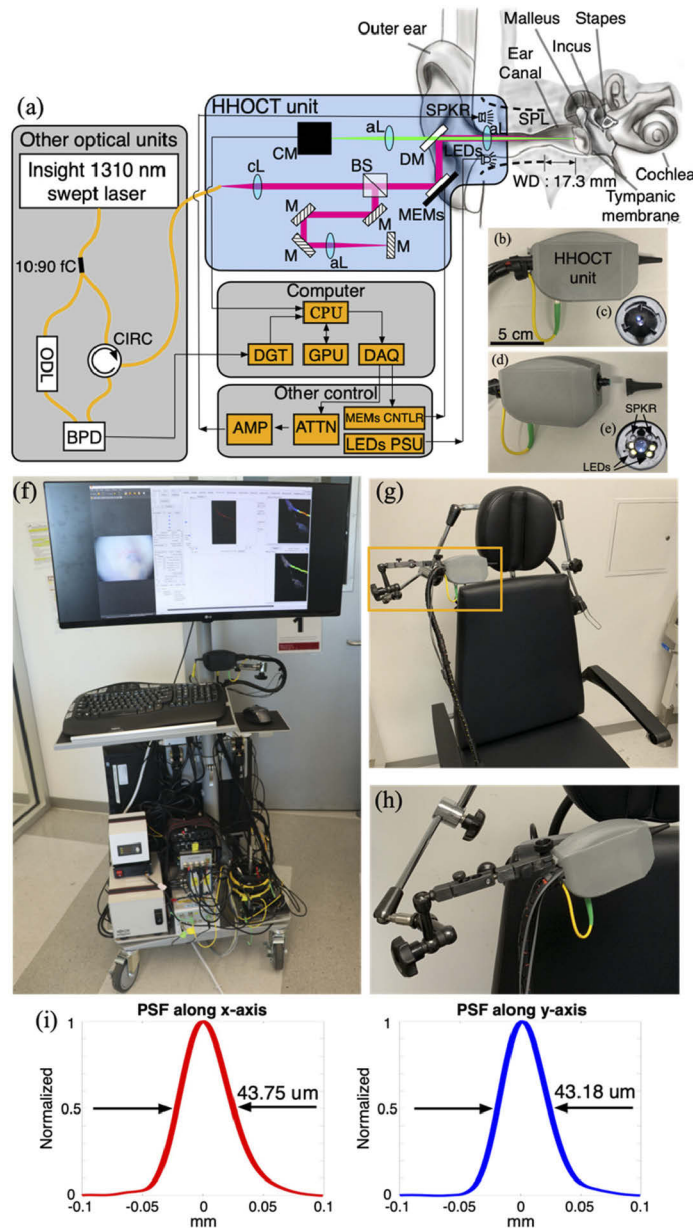


Fig. 1. (a) Schematic diagram of the entire HHOCT system. Abbreviations: 10:90 fC, 10X90 fiber optic coupler; ODL, optical delay line; Circ, fiber optic circulator; BPD, balanced photo-detector; CM, video camera; aL, achromatic lens; DM, dichroic mirror; cL, collimating lens; BS, beamsplitter; MEMs, MEMs scan mirror; M, mirror; DGT, digitizer; AMP, amplifier; ATTN, attenuator; MEMs CNTLR, MEMs scan mirror controller; LEDs PSU, light-emitting diode power supply unit; SPKR, speaker; LEDs, light-emitting diode; SPL (dashed line), speculum; WD, working distance. Close-up image of (b) the HHOCT device and (c) attached otoscopic speculum. (d) Speculum installation and (e) the front of the device without an attached speculum is pictured to show the orientation of speaker outputs and LEDs. (f) The entire HHOCT system is shown with handheld device visible in the center in a black housing. (g) ENT examination chair with HHOCT device attached. (h) Close-up of the attachment accessory to mount the HHOCT device. (i) Lateral point-spread function (PSF) along the X and Y-axes of OCT imaging using the HHOCT device.

The entire Michelson interferometer is contained within the HHOCT unit in order to maintain good interferometric phase stability even in a hand-held device [25]. Light exiting the single mode optical fiber is collimated by a reflective lens (cL) before passing through a 50:50 beamsplitter (BS). The reflected beam from the beamsplitter was folded with a set of three mirrors (M) before passing through an achromat lens (aL) to be focused on the final reference reflector of the Michelson interferometer. The light passing through the 50:50 BS was reflected by a dual axes MEMs mirror (Mirrorcle Tech, Inc.) used to scan the OCT beam across the sample. This light was reflected by a short-pass dichroic mirror (DM, 950 nm cutoff, DMSP950R, Thorlabs) and focused onto the sample using an achromatic scan lens (aL) generating a 43 μm (FWHM) beam spot size (see Fig. 1(f)). The light from the sample was back reflected through the optical system, through the CIRC and onto the other side of the BPD. The ODL was tuned so that the optical paths after the 90:10 coupler were identical. This enabled cancellation of common mode noise and the DC term of the interference signal [25]. The power of the 1310 nm light on the sample was 7.5 mW, well below the ANSI limit for both the eye and skin [26].

The sample was also illuminated with a set of visible light-emitting diodes (LEDs) arranged around the scan lens (aL). Visible light back scattered by the sample was imaged onto a 2D CMOS sensor-based camera (MU9PC-MH, XIMEA). A USB interface provided a real-time digital video otoscopic view to improve positioning of the device while collecting functional and morphological OCT images from the ear. The final component of the HHOCT unit was a pair of speakers used for acoustic stimulus. The speakers consisted of stereo earbuds (Klipsch, Inc., R6 II) attached to rubber tubes fed into the space around the aL also occupied by the LEDs.

All of the electronic components were positioned on a cart as described earlier in [14]. The signal from the BPD was digitized (ATS9373, Alazar Tech, Inc.) and then processed using custom software written in Python, C++, and CUDA with a companion high-end GPU (GeForce GTX-1080, Nvidia). We will only briefly describe the processing of the OCT signal here, however a detailed description of our signal processing steps may be found in [27]. In order to measure vibrations, the interferometric signal is consecutively recorded during the presentation of a sound, to generate what is called an M-scan. After processing of the spectral interferogram to generate the complex OCT signal, the interferometric phase is extracted and unwrapped along the time-dimension. The slow drift of the interferometric phase is removed by fitting a 3rd order polynomial to the phase and subtracting the fit. For experiments requiring high phase-sensitivity, we collected signal over multiple 50 ms time segments (trials) and averaged the trials in the time-domain. As we have shown before [14], this approach works well even in the presence of some unintentional patient movement during the measurement due to respiration or other physiological sources of motion such as spontaneous swallow reflex or slight head movements. Following any temporal averaging and phase drift subtraction, a fast Fourier transform was computed along the time-dimension. The magnitude of the Fourier transform was converted from radians to nanometers by applying the scaling factor, $\lambda/(4\pi n)$, where λ is the center wavelength of the light source, and n is the refractive index of the sample. The amplitude and phase of the Fourier transform then correspond to the amplitude and phase of vibratory motion as a function of frequency.

For the otoscopic examination of a human ear, a commercial disposable ear speculum (Welch Allyn, Inc.) attaches to the HHOCT unit and can be easily detached from the custom designed connection shown in Fig. 1(c) and 1(d). The dimensions of this device are 130 \times 65 \times 35 ($L \times H \times D$) mm with a speculum attached. Based on the configuration in Fig. 1(a), the working distance from the speculum tip to the image plane is 17.3 mm. The maximum imaging depth based on sampling (Nyquist depth) is 11 mm, hence the instrument is capable of imaging the entire depth of the middle ear by inserting the device into the ear canal in a similar manner as a routine otoscopic ear exam. In order to make more stable measurements over longer periods of time, the HHOCT device can be attached to the articulating arm of a customized otolaryngology

examination chair as shown in [14], and visualized in Fig. 1(g,h). As shown from the magnified image in Fig. 1(h), a clinician can use the device in a hand-held manner or mounted to the chair with the articulating arm using locking knobs. Usage of the articulating arm allows for stable measurements from any position within the ear canal of a patient and may be more suitable than hand-held usage for measurements requiring longer acquisition times to minimize noise introduced by spontaneous user motion.

We used a pair of stereo earbuds (Klipsch, Inc., R6 II) for acoustic stimulation. Using two separate speakers was important for generating a flat sound output in the lower (~ 0.25 -1 kHz) and upper (>1 kHz) frequency ranges. Each earbud was inserted into a port that directed the sound to the speculum tip. The speakers were calibrated separately to recreate a flat response across all frequencies using rubber tubing to approximate the length and diameter of the ear canal with a microphone positioned at the approximate distance the tympanic membrane would be located. Placement of the microphone will also introduce random error above 4 kHz. In practice, a patient's individual anatomy will influence the delivered acoustic levels. Hence, there is some inherent systematic error due to individual ear canal anatomy. The artificial ear canal was used to help reduce this error by recapitulating some of the morphological features of the ear canal.

In one set of experiments, we collected air and bone stimulation data in serial. The patient was fitted with a set of bone conduction headphones (SPORTZ Titanium, AFTERSHOKZ). These were interfaced with the HHOCT system so that a pure tone could be played via the embedded speakers in the HHOCT device, followed immediately (5 ms time interval) by stimulus via bone conduction excited only on the ipsilateral side. The bone stimulus was first calibrated so that the volunteer perceived the same level of loudness for both bone and air conduction. We then found the differential calibration, air +10 dB SPL, which provided similar vibration of the cochlear promontory as air stimulus at a specific sound pressure level.

In order to ensure that the bone conducting headphones were not also generating air conducted sound, we performed the following set of experiments. After calibration as noted above, the bone conducting headphones were set up in free space with a microphone at approximately the distance to the tympanic membrane. An 800 Hz sound was played through the headphones at a level corresponding to 65 dB SPL when placed on the volunteer. No sound was measured by the microphone. Considering the microphone's noise level this implies that any air conducted sound generated by the bone conduction headphones was <4 dB SPL. Since the bone conduction speaker will perform differently under load and the volunteer's skull can act as a speaker, we also performed a second test. We placed the microphone in the volunteer's ear canal and played a sound through the bone conduction headphones that corresponded to 65 dB SPL. The microphone measured 39 dB SPL at 800 Hz. This implies a differential of at least 26 dB. To the extent that this sound is produced by the vibrating skull, this is common to any bone conduction experiment, no matter the type of transducer. If the sound is produced by the transducer housing vibrating, this sound would be further attenuated because the speculum obstructs the ear canal during our measurements. Given the measurements with the microphone and the fact that in the human measurements the ear canal is obstructed, we are confident that bone conduction measurements are not significantly contaminated by air conducted sound.

Figure 2 provides some quantification of the optical system performance. A latex drum was constructed by stretching a piece of latex over a cylindrical tube to mimic the TM [28]. The reflectivity of the latex is similar to what we find on the living TM. The HHOCT device was attached to an articulating arm while the interferometric signal from the latex drum was collected as a function of time for 50 ms. Figure 2 shows the log-scaled vibrational amplitude as a function of frequency from 0.25 to 25 kHz following the signal processing noted above. The noise mean (μ) and standard deviation (σ) was calculated along the frequency dimension with moving 100 Hz windows. We define our system sensitivity using $\mu + 3\sigma$ as a threshold, corresponding to the 98% confidence interval of the Rayleigh distribution [27]. Any signal below the sensitivity

is considered as noise and rejected. For this test sample, with an SNR of 45.2 dB, the mean and standard deviation of the noise was 20.2 pm and 11.3 pm in the 2.5-8 kHz range. With the exception of the two noise peaks at 8.5 and 16.5 kHz, the noise statistics are similar in the entire 2.5-25 kHz range. This implies a sensitivity of 54.2 pm above 2.5 kHz, which is off by about a factor of 2.5 from the expected theoretical sensitivity of 21.7 pm calculated from eq. 15 in [27]. At frequencies below 2.5 kHz, the noise has a nonlinear frequency dependence as shown in the inset of Fig. 2. Frequencies of particular relevance because of their common use in pure tone audiometry are 250 Hz, 500 Hz, 1 kHz, and 2 kHz with measured sensitivity of 9.49 nm, 1.60 nm, 352 pm, and 339 pm, respectively. Lower frequency sound stimulation produces relatively larger amplitude motion, hence in spite of poorer sensitivity we were capable of measuring the vibratory response in the middle ear with 60-80 dB SPL. For reference, normal conversational speech is ~60 dB SPL. As we have shown before [14] with temporal averaging, it is possible to measure at even lower sound pressure levels, down to the threshold of human hearing.

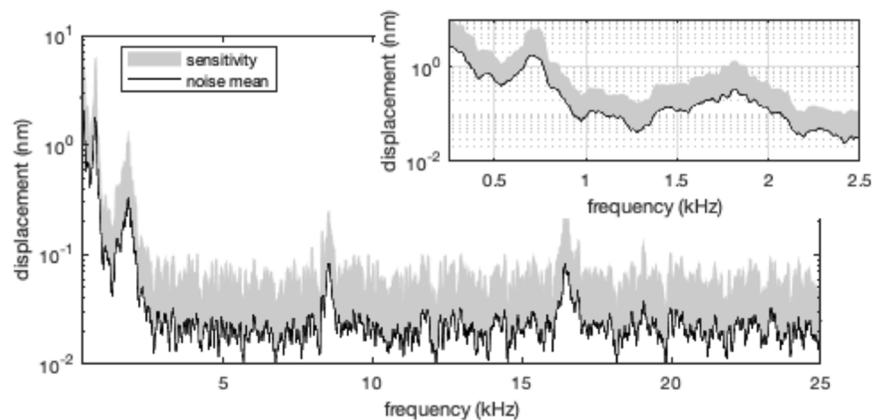


Fig. 2. System performance. Noise mean and system sensitivity as measured from a latex drum. The noise mean (μ) and standard deviation (σ) were calculated with a sliding window of 100 Hz.

3. Results

We used the HH OCT device to image the middle ear of two healthy volunteers. A number of stimulus approaches were evaluated. A summary of stimulus and imaging parameters are summarized in Table 1. The device was inserted into the right ear canal using the attached speculum and positioned to provide a view of the middle ear through the camera as shown in Fig. 3(a). Cross-sectional OCT B-scans of the subject's TM at the umbo and cone of light are shown in Fig. 3(b) and (c) with corresponding red and blue dotted lines in 3(a) identifying the path where the B-scans were collected, respectively. Acquisition time for each B-scan including post-imaging computational processing was around 140 ms. The arrows on the B-scans denote the location where vibratory tuning curves were measured. While vibratory data was collected over the entire axial range, only data from the strongest axial reflections (TM surface) are plotted in Fig. 3(d) and (e). The frequencies at which the magnitude of the curve was below the limit of detection (sensitivity) are indicated by the thin dotted portions of Fig. 3(d) and 3(e). For the tuning curves, we recorded 2 trials of 50 ms duration resulting in a total acquisition time of approximately 19 seconds in one of the volunteers. This acquisition time was sufficient to record signals at 80 dB SPL for the majority of stimulus frequencies at both the umbo and cone of light. We found that using a relatively shorter acquisition time helped to minimize unintentional volunteer movement during measurements resulting in reduced motion artifact. Normalizing

the displacement magnitudes to the eliciting stimulus pressure revealed an approximately linear relationship between increasing displacement and increasing stimulus pressure level is observed, as expected for the middle ear.

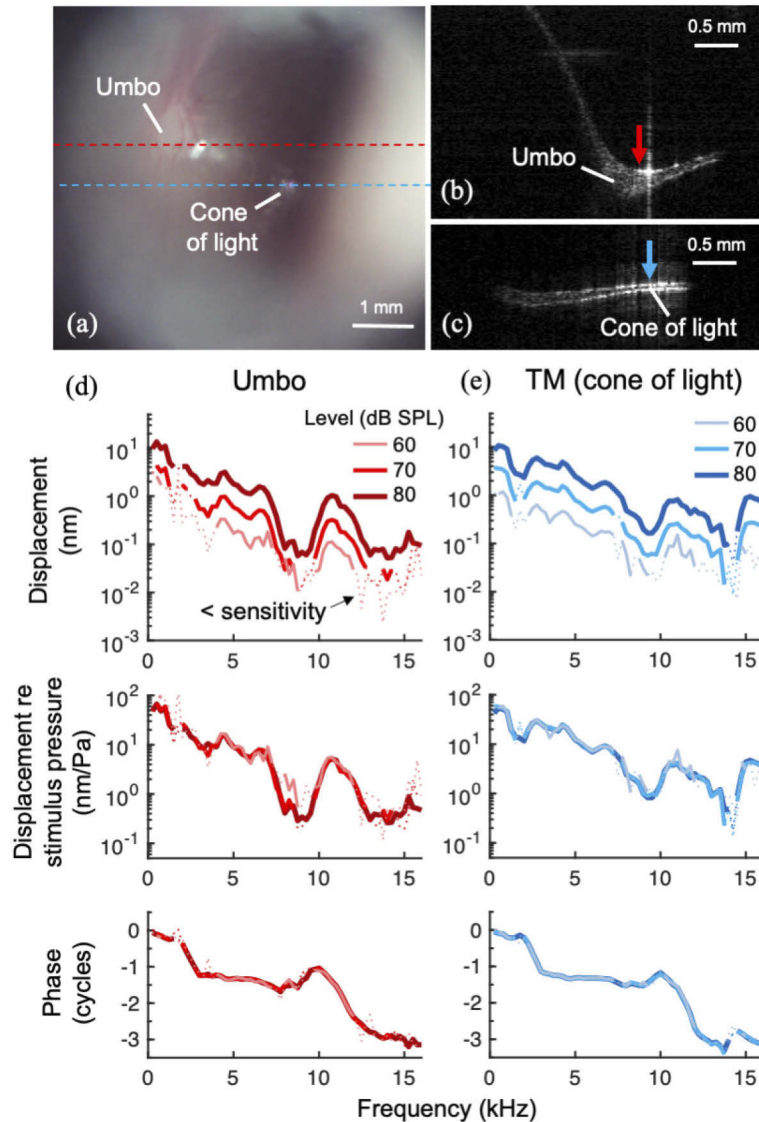


Fig. 3. (a) Image of the right TM taken with the probe's camera in one human subject. Dashed lines indicate the path over which the light source was scanned to obtain subsequent B-scans. Scale bar = 1 mm (b)-(c). B-scans were taken using the hand-held device with the light source scanned over the paths shown in (a). Arrows indicate locations on the (b) umbo and (c) TM near the cone of light where vibration measurements were obtained. Scale bars = 0.5 mm. (d)-(e) Displacement magnitude and phase of the (d) umbo and (e) TM near the cone of light in response to tones presented over the frequency range of 250 Hz - 16 kHz at 250 Hz intervals with a range of sound pressure levels from 60-80 dB at 10 dB intervals. Middle panels show the displacements normalized to the stimulus pressure, showing that the displacements grow linearly with stimulus level. Thin dotted portions of the curves indicate where the displacement magnitudes were lower than the system sensitivity.

Table 1. Summary of vibrometry experiment parameters for different stimulus types

Stimulus Type	Single Tone	Chirp	Air/Bone	Multi-Tone Complex
Figure	3	4	5	6
Conduction type	Air	Air	Air/Bone	Air
dB SPLs	60, 70, 80	~78	80/70	70
Frequencies	.25-16 kHz (.25 kHz intervals)	200 Hz-24 kHz (Swept logarithmically)	800 Hz	0.25, 0.50, 1.0, 2.0, 4.0, 6.0, 7.8 kHz (Tone complex)
Spatial sampling	1	1	1 × 66	22 × 25
Stimulus time	50 ms (2 trials)	100 ms (100 trials)	50/50 ms	50 ms
Total data acquisition time	9.6 s	10 s	6.6 s	27.5 s

We then recorded the middle ear vibratory response of the volunteer in response to a chirp stimulus. Imaging was conducted as above in the right ear. The area near the cone of light on the TM was selected as the point of interest. The chirp stimulus was designed to sweep from 200 Hz to 24 kHz in a logarithmic manner over the course of 100 ms at 78 dB SPL in order to rapidly test responses at all frequencies. The stimulus waveform is shown in Fig. 4(a). The stimulus was repeated for 100 trials, with the resulting responses averaged and high-pass filtered above 80 Hz to eliminate motion artifact that manifests primarily at the low frequency range. The resultant displacement waveform after averaging and filtering is shown in Fig. 4(b). We used a least-squares-fit technique [29] to estimate the magnitude and phase of the responses as a function of time/frequency to normalize the displacement and phase in relation to the stimulus pressure as shown in Fig. 4(c) and Fig. 4(d) respectively. The portion of the curves that contain thin dotted lines indicate measurement points that were below the system sensitivity. For these measurements, the sensitivity was quantified by taking all of the individual response waveforms and inverting the polarity of every other waveform (multiplying by -1) prior to averaging and performing the least-squares fit analysis. This yields an estimate of the variance in the response.

We then recorded tuning curves using bone conduction headphones. The subject's right ear was imaged with the HHOCT device in the same manner as above. A cross-sectional image with vibratory response was recorded. The cross-section bisected the tympanic membrane, a portion of the cochlear promontory, the incus near the lenticular process, and the stapes. At each lateral position a 50 ms pure tone at 800 Hz was presented first via one of the embedded speakers (70 dB SPL) and then via bone conduction after a 5 ms interstimulus interval. The results are shown in Fig. 5. Figure 5(a) and 5(b) are overlays of the vibratory response (color scale) on top of the morphological image (gray scale). The morphological image was generated by taking the time average at each lateral position. The vibratory response was only plotted if the data was above the system sensitivity, $\mu + 3\sigma$. The noise was measured simultaneously with vibratory response at 800 Hz by extracting the mean and standard deviation in the 825-875 Hz range. Figure 5(c) shows three regions of interest (ROI) where we have calculated the statistics of the vibratory response. The ROIs include areas of the TM (in all ROIs), the cochlear promontory (ROI 1), the lenticular process of the incus (ROI 2) and a portion of the long crus of the incus (ROI 3). The response at the tympanic membrane and boney areas were computed separately in each ROI. After removal of outliers, (value that is more than three scaled median absolute deviations), the mean, median, and standard deviation was calculated. The values for the median and mean were very similar, hence for reporting here we only use the mean. Figure 5(d) is the mean ± 3 std displacement for each ROI including both air and bone stimulation. Figure 5(e) shows the same results relative to the vibratory response at the cochlear promontory.

A volume M-scan of a segment of the TM was collected using a multi-tone complex stimulus consisting of a sum of pure tones at 0.25, 0.50, 1.0, 2.0, 4.0, 6.0, and 7.8 kHz calibrated to

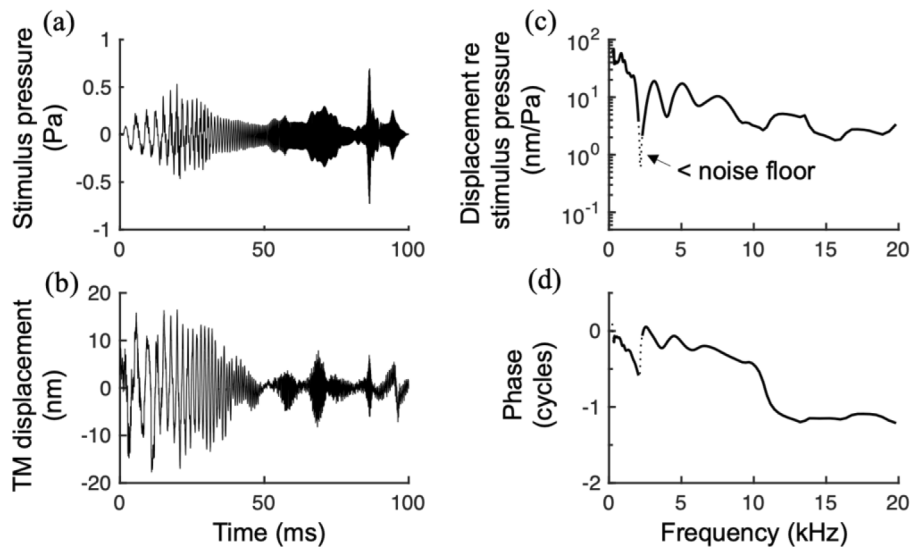


Fig. 4. (a) Waveform of a ~ 78 dB SPL chirp swept logarithmically in frequency from 200 Hz to 24 kHz in 100 ms with resultant (b) displacement waveform obtained from the TM (near the cone of light) using the handheld device. Responses were averaged over 100 stimulus repetitions and high-pass filtered above 80 Hz to eliminate low-frequency motion artifact. (c) Displacement magnitude and (d) phase of the TM normalized to the estimated chirp stimulus pressure in the ear canal using a least-squares-fit technique. Thin dotted portions of the curves indicate where the displacement magnitudes were below the measurement noise floor.

70 dB SPL per tone. These frequencies were selected as they are often used during pure tone audiometry testing. We used both earbuds, one for low frequencies (≤ 1 kHz) and one for high frequencies (> 1 kHz). This was necessary in order to achieve 70 dB SPL at all frequencies.

Compared to pure tone stimuli, a tone complex stimulus allows for more rapid acquisition than tones presented sequentially and may also more accurately reflect the response to wideband sounds encountered outside of the experimental setting. The stimulus was presented for 50 ms at each of the 22×25 spatial positions which covered a field of view of $2.21 \times 2.48 \text{ mm}^2$. Total data collection time is nominally 27.5 s, but when accounting for mirror flyback and other overhead the actual collection time was roughly 1.3x longer. In order to generate a volume image, the data set was averaged along the time dimension. The surface of the TM viewed from the ear canal is shown in Fig. 6(a). Shadowing from the speculum is seen at the bottom left of this image. The two orthogonal dotted lines point to cross-sectional (B-scan) images from the volume taken at the position indicated by the dotted lines. From these cross-sections, it is clear that the brightest area in the surface image corresponds to the thickening of the TM near its interface with the malleus. The vibrational analysis was completed as described previously in [27]. The frequency domain signal from the brightest pixel on the TM is shown in Fig. 6(b) and indicated in Fig. 6(a) by a magenta x. The peaks corresponding to the stimulus frequencies are labeled in red. The green symbols indicate frequencies 100 Hz above the stimulus frequencies which were used for subsequent calculations of the sensitivity. The volume M-scan was masked such that any signal falling below the sensitivity was set to zero. From the remaining signals, the median through the thickness of the TM was calculated at each of the 22×25 spatial positions. The median amplitude and phase are plotted in Fig. 6(c) for 0.5 and 6 kHz. These are oriented the same as the TM surface view in Fig. 6(a), i.e. the view is down the ear canal. For the amplitude maps, the

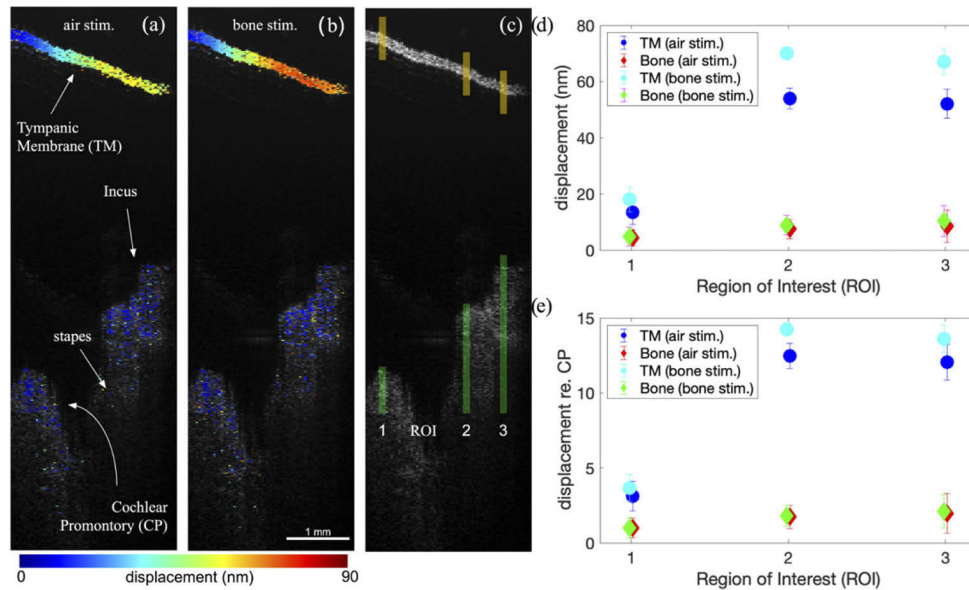


Fig. 5. (a)–(b) Overlay of vibratory response (color scale) on the morphological image (gray scale) of the middle ear obtained using the device mounted to an articulating arm. (a) Pure tone stimulus was at 800 Hz and 70 dB SPL. (b) Bone stimulus calibrated to give similar vibratory response at the cochlear promontory. (c) Morphological image showing the 3 regions of interest (ROI) used to generate the plots in (d) and (e). (d) Plot of the $\mu + 3\sigma$ displacement for the ROI at the TM and bone (Cochlear Promontory or Incus) resulting from air and bone stimulation. (e) Displacement data from (d) plotted relative to the vibratory response at the cochlear promontory. For the computation of the statistics, the number of pixels used varied from 23–94.

overall range (min-max) is given at the top of the image and correspond to the min/max in the color bar on the right. The phase images vary from $-\pi$ to π radians for both.

For this healthy volunteer, there were many similarities among the vibrational and phase maps from the lowest to highest frequency. Therefore, 0.5 and 6 kHz were chosen as representative of the seven collected. The low frequency range is characterized by generally higher amplitude vibration. Note the near order of magnitude difference between 0.5 and 6 kHz. The phase of the vibration was fairly homogenous at low frequencies but started to show phase shifts going to higher frequency. That is exemplified by the transition from approximately 2 radians to -1 radian in the phase map for 6 kHz stimulus.

We then used the HHOCT device to collect three OCT scans over a square field of view of $5 \times 5 \text{ mm}^2$ from the healthy volunteers. The corresponding B-scans are shown in Fig. 7(a)–(c). The thickness maps of the TM were computed for these scans and are depicted in Fig. 7(d)–(f) both in 2D and 3D representations. The location of the umbo is marked by a red arrow in Fig. 7, except for the scan in the first column where the umbo is outside the field of view. It can be seen from the 2D and 3D maps that the imaged membranes have similar thickness profile albeit some differences can be observed.

Computing the thickness of the TM is challenging because of its cone-shaped structure. The thickness measurements at each voxel on the membrane must be taken orthogonal to the surface of the membrane. In addition, there may be some contribution from refraction at the angled surfaces, following Snell's law. In this work, we adopted a computationally efficient technique developed by Aganj et al. [30]. The approach is based on computing the minimum

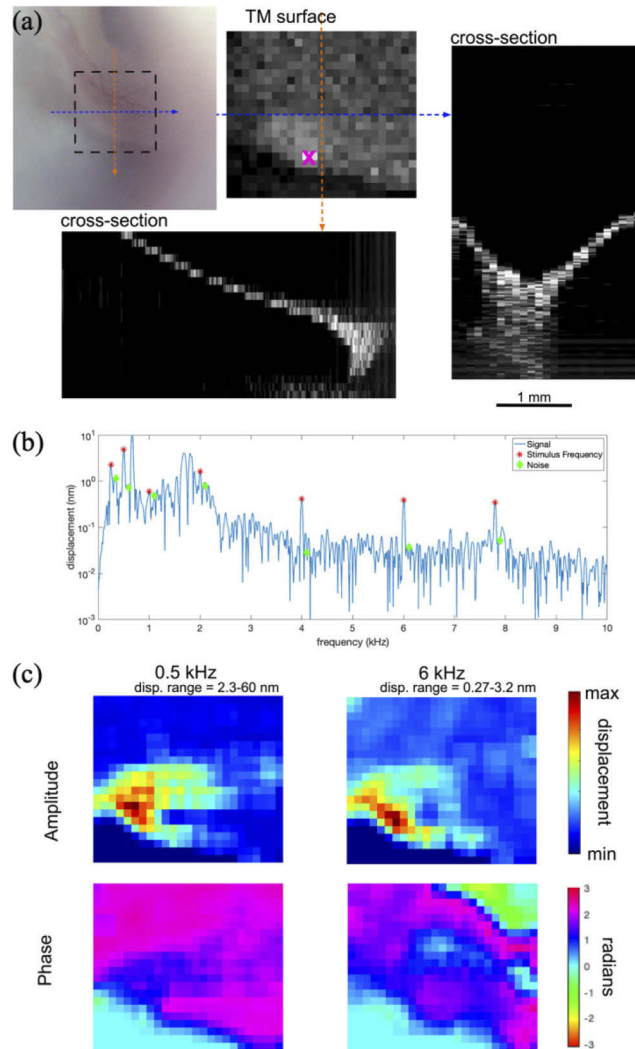


Fig. 6. Volume M-scan (x,y,z,t) using a tone complex acoustic stimulus using the device mounted to an articulating arm. (a) Otoloscopic image of the surface of the tympanic membrane (TM) and volume image (time averaged) as viewed from the ear canal. Cross-sectional B-scan from approximately the middle of the volume stack. (b) Frequency response of pixel marked with magenta x in TM surface image. Signals at each stimulus frequency (red) can be seen along with the frequency used to estimate noise (green). (c) Amplitude and phase of vibration along the TM at 22×25 spatial positions.

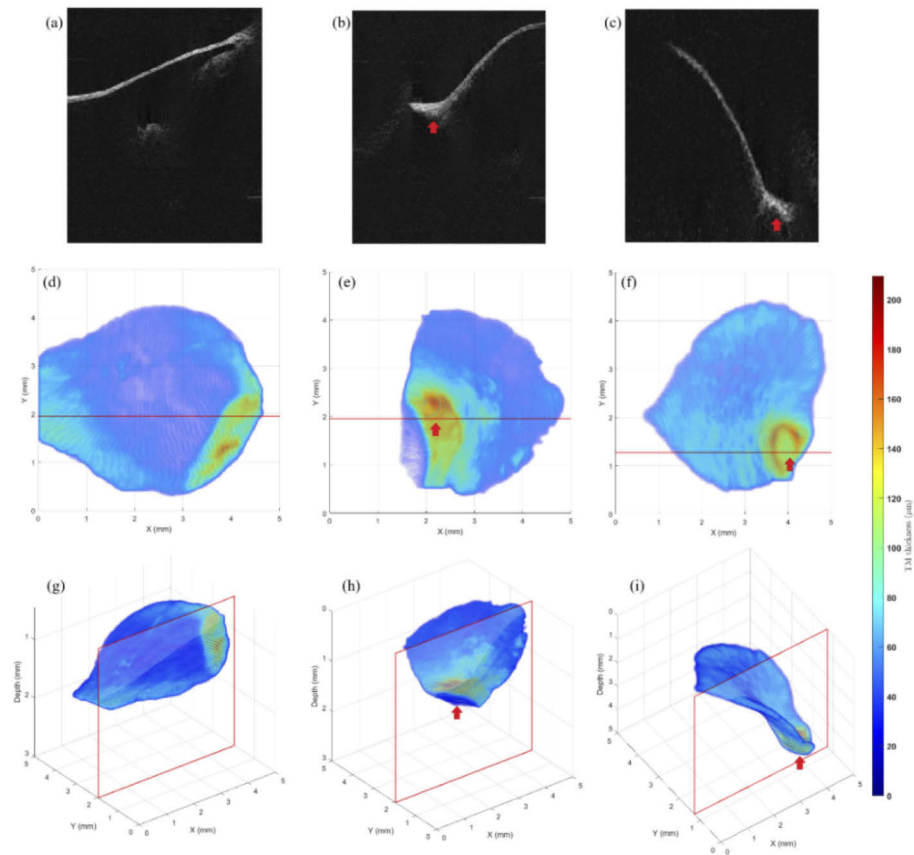


Fig. 7. Thickness maps of three healthy human tympanic membranes. (a)-(c) shows the cross sectional (B-scans) images acquired using the hand-held device. (d)-(f) is the 2D (XY) projection of the corresponding 3D thickness maps shown in (g)-(i). The solid red arrows indicate the location of the umbo. The B-scans are the slices enclosed by the red borders in the 3D thickness maps and are acquired along the red lines shown in the 2D projection subfigures.

line integrals on a probabilistic segmentation mask of the TM. We constructed the probabilistic segmentation mask by first applying a 3D median filter and increasing the contrast in the OCT volume, and then thresholding with the Otsu method [31]. Prior to that, the OCT volume was preprocessed to remove background noise and saturated A-lines. A more detailed description of the pre-processing is described in our previous work [14].

The current algorithm is semi-automated consisting of two main stages. First, a set of coarse segmentation masks and a denoised volume are generated from the original B-scan volume. The segmentation technique currently employed is highly sensitive to noise, contrast, occlusion, and motion inherently present in the acquired scans. Thus, a set of segmentation masks are saved at different stages of the segmentation pipeline, so that an optimal mask can be subjectively selected for manual refinements. Finally, a refined mask along with the denoised volume are used as input to generate the TM thickness map. Due to the necessary manual intervention, generating the TM thickness of a typical volume currently takes several minutes.

4. Discussion

Various pathologies of the middle ear present with conductive hearing loss due to alterations in sound vibration conduction [32–34]. Recent studies have shown that OCT vibratory measurements can help discriminate between normal controls and patients with otosclerosis [21]. However, given the limited use of the approach in the clinic so far, it is not clear what information is most valuable. Nevertheless, a tool that can rapidly diagnose disorders of conductive hearing loss like otosclerosis, tympanosclerosis, and post-surgical scar formation by measuring sound-evoked vibrations in the clinic would represent a significant step forward in the quality of care and decrease the latency period for patients between symptom onset, diagnosis, and treatment. In this work, we tried various approaches for extracting functional information on the middle ear using our HHOT device that may prove valuable for diagnosing various pathologies.

In the first approach presented here in Fig. 3, we followed the procedure we have commonly used for animal studies of the inner ear (see for instance [7,35]). We presented pure tones at various frequencies and sound pressure levels. The results show a trend toward lower displacement with higher stimulus frequency. There was a significant notch recorded in the 7.5–10 kHz range. The notch was more pronounced at the umbo compared to the cone-of-light measurement, but otherwise both measurements showed similar features. When normalized to the stimulus pressure, measurements at 3 different sound pressure levels fall on each other, as we would expect for a linear system. Similarly, the phase for all sound pressure levels is the same, with corresponding features in the umbo and cone-of-light measurements. Like the amplitude displacement, the phase trends down with increasing frequency and has a peak corresponding to the high frequency edge of the notch noted in the displacement amplitude.

These data can also be compared to laser Doppler vibrometry (LDV) of the tympanic membrane. Our data from this subject fall within the range observed by Whittemore et al. [36], with velocity re stimulus pressure being ~ 0.1 at low frequencies and reaching ~ 0.3 to 0.4 at higher frequencies (Whittemore et al. only went up to 6 kHz). They also had a few subjects with phases that accumulated > 1 cycle from 0.3–6 kHz, similar to what we observed in this ear over the same frequency range.

The three stimulus levels appear to map similar features in the vibratory response; hence the measurements appear to be largely redundant. Likewise, the response in this healthy volunteer looks similar at both the umbo and cone-of-light. Based on these observations, in the second approach, Fig. 4, we used a single chirped stimulus at one level, ~ 78 dB SPL, on the same healthy volunteer at the cone-of-light. The advantage of the chirp is that we can measure the vibratory response at a much higher frequency resolution, hence more clearly mapping the features in the vibratory response. This is particularly evident at frequencies below 5 kHz where the response appears to oscillate more than was evident in the discrete tone measurement, i.e. compare Fig. 3(e) second row to Fig. 4(b). Qualitatively, larger features are reproduced as well, although there appears to be small differences, perhaps due to the fact that the measurements were done on different days and likely at slightly different positions on the TM.

The umbo and cone-of-light are two spatial areas on the TM that are readily identified in patients and could be used as points for longitudinal monitoring of middle ear vibratory response and health. Nevertheless, under some circumstances, it may be advantageous to image the vibratory response along a cross-section or volume.

As an example, we measured the middle ear vibratory response with bone conduction stimulation sequential with air stimulation. This offers several potential clinical advantages. Bone conduction measurements are typically used to make the initial diagnosis of conductive hearing loss and as a quantitative measure of the success of various clinical treatments [37]. For instance, the air-bone gap is used post-ossiculoplasty as a measure of the success of the surgical intervention [38]. However, current diagnostic tests rely on precise calibration of the transducer used for bone stimulation. Maintaining this calibration can be problematic in the harsh clinical

environment [39], potentially introducing measurement error. Likewise, the fit of the device on the patient as well as variability in head geometry also contribute to measurement error [37].

As shown in Fig. 5, we can measure the vibratory response in a cross-section that includes the tympanic membrane, the cochlear promontory, and part of the ossicular chain. We can see that in a healthy human subject, the TM vibrates similarly under bone or air stimulus. It vibrates the least on the left side of the image, nearest the wall of the ear canal. The vibrations get progressively larger in amplitude moving to the right in the image, toward the center of the TM. Likewise, the response at the cochlear promontory and incus is similar with both air and bone stimulation. This is quantified in Fig. 5(d) where we have plotted the mean \pm 3 standard deviations in the regions of interest identified in Fig. 5(c). The bone conducting headphones were calibrated to give similar vibratory response at the cochlear promontory with both air and bone stimulation, nevertheless there is a small 0.6 nm difference in the mean in ROI 1. We observe similar small differences at the two ROIs on the incus (1.4 and 1.9 nm, respectively). The cochlear promontory response should provide a quantitative measure of the bone stimulation at the cochlea that is insensitive to transducer calibration or patient fit and head geometry. Measures of hearing based on this metric could thus provide a more precise measure of conductive hearing loss than is currently possible, also providing a means of standardization across clinics.

LDV has also been used to measure the vibration in the middle ear via bone conduction. Stenfelt et al. [40] measured the vibratory response in 26 temporal bones and one human volunteer. Particularly relevant is their measurement of the ratio of the velocity of the stapes footplate and cochlear promontory. This ratio can be compared directly to the ratios we plot in Fig. 5(e). In ROI 2, we make measurements from the incus at the lenticular process and near the incudostapedial joint. Motion of this bone should be similar to the stapes footplate, although there are probably small losses at the joint between the two bones and there may be differences in the angle of motion. In any case, we estimate the ratio from Fig. 5(a) from Stenfelt et al. as 1.14 [40], compared to our measure of 1.8 ± 0.7 . The good agreement between our measurement and the average of 26 temporal bones provides some confidence in our measurement.

As noted above, the middle ear is generally well represented by a linear system, hence we propose normalizing the vibratory response to cochlear promontory vibratory response to provide a measure of the relative motion of the elements of the middle ear. This is shown in Fig. 5(e) for the 3 ROIs. In the pathological cases we expect the relative motion to be altered. For instance, if the stapes is fused, the ossicles should move less, and the tympanic membrane would likely be affected as well. Identifying differences among the elements of the middle ear in conductive hearing loss could help to refine the diagnosis and improve therapeutic planning.

We are also able to map the vibratory response over a volumetric image. Practically, it would be very difficult to collect the vibratory response with the same frequency resolution as in Figs. 3 and 4. It is not likely that a patient could remain still long enough to collect such data without substantial motion artifact. As a compromise, we chose to present a complex of pure tones at frequencies commonly used for pure tone audiometry. Tone complexes, where multiple tones are presented as stimulus at the same time are commonly used in animal research on hearing [41]. This allowed us to record the vibratory response over a range of frequencies with a single 50 ms stimulus at each lateral position. For a factor of 2-3 increase in total acquisition time we were able measure the vibratory response at 550 lateral positions, compared to the single lateral position shown in Figs. 3 and 4. The response at the TM is shown in Fig. 5. Limitations in the speaker output only allowed us to stimulate with 70 dB SPL per tone. If we were able to stimulate at 80 dB SPL per tone, we could have used a shorter stimulus time, enabling a larger field-of-view or simply a faster acquisition time.

The measurements of the vibratory response, as a function of lateral position or frequency and sound pressure level, took seconds to record. As a consequence, motion artifact may play a role in the measurement. Since the vibratory response does not vary dramatically with small lateral

motions, these will not significantly contribute to errors in amplitude or phase, but do add noise to the system as we discussed previously [14]. The longer the measurement, the more likely the patient is to move. Large motion results in a large increase in phase noise, hence it would be apparent in the data.

Finally, we explored the measure of TM thickness. There is increasing evidence that certain pathologies of the TM can cause altered thickness of the membrane. In otitis media (OM), studies have shown the presence of a bacterial biofilm that develops on the TM can lead to increased thickness [42,43]. This has been exploited to differentiate normal TM, and acute and chronic OM [17,18]. Other pathologies such as chronic myringitis [44] and tympanosclerosis [19,45] have also been measured using OCT. A recent study used an OCT-based otoscope to monitor the healing of grafts inserted during tympanoplasty and to assess the junction between the native TM and the graft, as well as to monitor healing of tympanic perforations [6]. A tool that can quickly and reliably produce TM thickness maps from OCT volume images could rapidly become an important diagnostic tool in the otolaryngology clinic.

Here we continued to develop an approach first used in MRI [30] for measuring cortical thickness, but recently adopted by our group for TM thickness [14]. While it is still a stand-alone program, we have reduced the run time and the amount of user intervention. Nevertheless, this manual intervention can still restrict the usability of the current software particularly in a clinical setting. Our future work includes exploring Artificial Intelligence (AI) approaches to improve both the accuracy and execution time of the segmentation. With the current software, we expect to generate a large number of manually segmented masks which will eventually form a labeled dataset for training the AI model. Adopting an AI technique would not only enable computing the TM thickness map near real-time but also minimize, if not completely eliminate, the current manual intervention.

We report various vibrometric and topographical approaches towards extracting functional information about the middle ear using the HHOCT device including measuring response to pure tone stimuli along with newer techniques such as multi-tone stimulus, continuous chirp stimulus, and alternating air and bone stimulus. In our vibrometry experiments, subject motion did not significantly interfere with measurements in response to stimuli at higher sound pressure levels. In the routine clinic setting, certain patient populations may not be capable of remaining as motionless as the normal subjects that were tested in this study, and subject motion may be more of a concern for this system across a wider range of measurement parameters. We have not yet explored using various patient positioning devices, such as head holders or restraints, to reduce unintentional motion, but we propose that such solutions may improve the sensitivity of our system and improve its effectiveness in the clinic. While the data in this report was acquired from healthy subjects, these approaches should prove useful towards acquiring pathologic data in patients that will aid clinicians in the diagnosis of middle-ear diseases such as tympanosclerosis, cholesteatoma, and otosclerosis, among others. The clinical utility of this device will require validation in large patient populations and our future work will include exploring the HHOCT's ability to provide diagnostic data that will help differentiate between various pathologic middle-ear conditions, as well as the ease of integration within the routine workflow in an otolaryngology clinic.

5. Conclusions

In conclusion, we have developed a hand-held OCT device capable of functional and morphological imaging of the middle ear optimized for usage in a clinic setting. It essentially incorporates all of the functionality of the microscope-based system we previously described [14] but in a much smaller stand-alone footprint that allows for intuitive hand-held usage and convenient transport of the system as needed within the workflow of a busy otology clinic. By comparison, the two systems can both measure vibratory response with sub-nanometer scale sensitivity. The hand-held

device had better performance at low frequency, which allowed us to measure TM vibrations below 2 kHz, while the microscope system had a larger amount of noise at low frequencies. We attributed this difference to the long boom arm which holds the surgical microscope system, which is likely susceptible to low frequency motion. Nevertheless, the microscope-based system was easier to use for measurements where significant temporal averaging with long acquisition times was essential, such as measures of distortion product otoacoustic emissions. The microscope-based system is more likely appropriate for otologic surgery or for large institutional hearing clinics that already regularly use surgical microscopes for middle ear imaging. In comparison, the development of a hand-held OCT device capable of both rapid point-of-care middle ear imaging and functional diagnostic vibrometry in a stand-alone system that can be readily adopted into the clinic workflow may provide otologists and audiologists with valuable clinical information that is not readily available in current practice. While we investigated various approaches to extracting functional information, follow-up studies will be necessary to show the effectiveness of using OCT in the clinic to detect and diagnose middle ear pathology.

Funding. National Institute of Biomedical Imaging and Bioengineering (1R01EB027113); National Institute on Deafness and Other Communication Disorders (1R01DC013774, 1R01DC01445, 1R01DC017741).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. 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," *Science* **254**(5035), 1178–1181 (1991).
2. K. König, M. Speicher, R. Bückle, J. Reckfort, G. McKenzie, J. Welzel, M. J. Koehler, P. Elsner, and M. Kaatz, "Clinical optical coherence tomography combined with multiphoton tomography of patients with skin diseases," *J. Biophoton.* **2**(6-7), 389–397 (2009).
3. L. Ferrante di Ruffano, J. Dinnes, J. J. Deeks, N. Chuchu, S. E. Bayliss, C. Davenport, Y. Takwoingi, K. Godfrey, C. O'Sullivan, R. N. Martin, H. Tehrani, and H. C. Williams, and C. S. C. D. T. A. Group, "Optical coherence tomography for diagnosing skin cancer in adults," *Cochrane Database Syst. Rev.* **12**, CD013189 (2018).
4. Y. Li, Z. Zhu, J. J. Chen, J. C. Jing, C. H. Sun, S. Kim, P. S. Chung, and Z. Chen, "Multimodal endoscopy for colorectal cancer detection by optical coherence tomography and near-infrared fluorescence imaging," *Biomed. Opt. Express* **10**(5), 2419–2429 (2019).
5. Y. Zeng, S. Xu, W. C. Chapman, S. Li, Z. Alipour, H. Abdelal, D. Chatterjee, M. Mutch, and Q. Zhu, "Real-time colorectal cancer diagnosis using PR-OCT with deep learning," *Theranostics* **10**(6), 2587–2596 (2020).
6. K. Park, N. H. Cho, M. Jeon, S. H. Lee, J. H. Jang, S. A. Boppart, W. Jung, and J. Kim, "Optical assessment of the in vivo tympanic membrane status using a handheld optical coherence tomography-based otoscope," *Acta Oto-Laryngol.* **138**(4), 367–374 (2018).
7. 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).
8. N. Choudhury, G. Song, F. Chen, S. Matthews, T. Tschinkel, J. Zheng, S. L. Jacques, and A. L. Nuttall, "Low coherence interferometry of the cochlear partition," *Hear. Res.* **220**(1-2), 1–9 (2006).
9. Y. Tona, T. Sakamoto, T. Nakagawa, T. Adachi, M. Taniguchi, H. Torii, K. Hamaguchi, S. Kitajiri, and J. Ito, "In vivo imaging of mouse cochlea by optical coherence tomography," *Otol. Neurotol.* **35**(2), e84–e89 (2014).
10. R. K. Wang and A. L. Nuttall, "Phase-sensitive optical coherence tomography imaging of the tissue motion within the organ of Corti at a subnanometer scale: a preliminary study," *J. Biomed. Opt.* **15**(5), 056005 (2010).
11. A. Ramier, J. T. Cheng, M. E. Ravicz, J. J. Rosowski, and S.-H. Yun, "Mapping the phase and amplitude of ossicular chain motion using sound-synchronous optical coherence vibrometry," *Biomed. Opt. Express* **9**(11), 5489–5502 (2018).
12. 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.* **109**(24), 9529–9534 (2012).
13. 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).
14. 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), 4395–4410 (2019).

15. J. E. Isaacson and N. M. Vora, "Differential diagnosis and treatment of hearing loss," *Am Fam Physician* **68**(6), 1125–1132 (2003).
16. 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).
17. 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," *Quant. Imaging Med. Surg.* **5**(1), 69–77 (2015).
18. 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).
19. H. R. Djalilian, J. Ridgway, M. Tam, A. Sepehr, Z. Chen, and B. J. Wong, "Imaging the human tympanic membrane using optical coherence tomography in vivo," *Otol. Neurotol.* **29**(8), 1091–1094 (2008).
20. 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," *Biomed. Opt. Express* **7**(11), 4621–4635 (2016).
21. 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," *Otol. Neurotol.* **40**(4), e349–e355 (2019).
22. 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," *Otol. Neurotol.* **41**(7), e901–e905 (2020).
23. H. Y. Lee, P. D. Raphael, A. Xia, J. Kim, N. Grillet, B. E. Applegate, A. K. Ellerbee Bowden, and J. S. Oghalai, "Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of corti," *J. Neurosci.* **36**(31), 8160–8173 (2016).
24. 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).
25. W. Kim, S. Kim, J. S. Oghalai, and B. E. Applegate, "Endoscopic optical coherence tomography enables morphological and subnanometer vibratory imaging of the porcine cochlea through the round window," *Opt. Lett.* **43**(9), 1966–1969 (2018).
26. L. I. o. America, "American National Standard for Safe Use of Lasers," in *ANSI Z136.1-2014* (American National Standards Institute, 2014).
27. S. Kim, J. S. Oghalai, and B. E. Applegate, "Noise and sensitivity in optical coherence tomography based vibrometry," *Opt. Express* **27**(23), 33333–33350 (2019).
28. D. Seong, J. Kwon, D. Jeon, R. E. Wijesinghe, J. Lee, N. K. Ravichandran, S. Han, J. Lee, P. Kim, M. Jeon, and J. Kim, "In Situ Characterization of Micro-Vibration in Natural Latex Membrane Resembling Tympanic Membrane Functionally Using Optical Doppler Tomography," *Sensors* **20**(1), 64 (2019).
29. G. R. Long and C. L. Talmadge, "Spontaneous otoacoustic emission frequency is modulated by heartbeat," *J. Acoust. Soc. Am.* **102**(5), 2831–2848 (1997).
30. I. Aganj, G. Sapiro, N. Parikshak, S. K. Madsen, and P. M. Thompson, "Measurement of cortical thickness from MRI by minimum line integrals on soft-classified tissue," *Hum. Brain Mapp.* **30**(10), 3188–3199 (2009).
31. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.* **9**(1), 62–66 (1979).
32. A. Burkhardt, L. Kirsten, M. Bornitz, T. Zahnert, and E. Koch, "Investigation of the human tympanic membrane oscillation ex vivo by Doppler optical coherence tomography," *J. Biophotonics* **7**(6), 434–441 (2014).
33. E. W. Chang, J. T. Cheng, C. Rösli, 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).
34. 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).
35. J. B. Dewey, B. E. Applegate, and J. S. Oghalai, "Amplification and suppression of traveling waves along the mouse organ of corti: evidence for spatial variation in the longitudinal coupling of outer hair cell-generated forces," *J. Neurosci.* **39**(10), 1805–1816 (2019).
36. K. R. Whittemore, S. N. Merchant, B. B. Poon, and J. J. Rosowski, "A normative study of tympanic membrane motion in humans using a laser Doppler vibrometer (LDV)," *Hear. Res.* **187**(1-2), 85–104 (2004).
37. R. Carhart, "Clinical application of bone conduction audiometry," *Arch. Otolaryngol., Head Neck Surg.* **51**(6), 798–808 (1950).
38. S. Iurato, G. Marioni, and M. Onofri, "Hearing results of ossiculoplasty in Austin-Kartush group A patients," *Otol. Neurotol.* **22**(2), 140–144 (2001).
39. R. E. Roach and R. Carhart, "A clinical method for calibrating the bone-conduction audiometer," *Arch. Otolaryngol., Head Neck Surg.* **63**(3), 270–278 (1956).
40. S. Stenfelt, N. Hato, and R. L. Goode, "Factors contributing to bone conduction: the middle ear," *J. Acoust. Soc. Am.* **111**(2), 947–959 (2002).

41. S. W. Meenderink and M. van der Heijden, "Distortion product otoacoustic emissions evoked by tone complexes," *J. Assoc. Res. Otolaryngol.* **12**(1), 29–44 (2011).
42. C. T. Nguyen, S. R. Robinson, W. Jung, M. A. Novak, S. A. Boppart, and J. B. Allen, "Investigation of bacterial biofilm in the human middle ear using optical coherence tomography and acoustic measurements," *Hear. Res.* **301**, 193–200 (2013).
43. G. L. Monroy, W. Hong, P. Khampang, R. G. Porter, M. A. Novak, D. R. Spillman, R. Barkalifa, E. J. Chaney, J. E. Kerschner, and S. A. Boppart, "Direct analysis of pathogenic structures affixed to the tympanic membrane during chronic otitis media," *Otolaryngol.–Head Neck Surg.* **159**(1), 117–126 (2018).
44. E. Guder, E. Lankenau, F. Fleischhauer, H. Schulz-Hildebrandt, G. Hüttmann, H. W. Pau, and T. Just, "Microanatomy of the tympanic membrane in chronic myringitis obtained with optical coherence tomography," *Eur. Arch. Otorhinolaryngol.* **272**(11), 3217–3223 (2015).
45. P. Pande, R. L. Shelton, G. L. Monroy, R. M. Nolan, and S. A. Boppart, "A mosaicking approach for in vivo thickness mapping of the human tympanic membrane using low coherence interferometry," *J. Assoc. Res. Otolaryngol.* **17**(5), 403–416 (2016).