Full-field transient vibrometry of the human tympanic membrane by local phase correlation and high-speed holography

Ivo Dobrev
Cosme Furlong
Jeffrey T. Cheng
John J. Rosowski

Full-field transient vibrometry of the human tympanic membrane by local phase correlation and high-speed holography

Ivo Dobrev,[a,*] Cosme Furlong,[a,b,c] Jeffrey T. Cheng,[b,c] and John J. Rosowski[b,c]
[a]Worcester Polytechnic Institute, Center for Holographic Studies and Laser Micro-MechaTronics, Department of Mechanical Engineering, Worcester, Massachusetts 01609, United States
[b]Massachusetts Eye and Ear Infirmary, Eaton-Peabody Laboratory, Boston, Massachusetts 02114, United States
[c]Harvard Medical School, Department of Otology and Laryngology, Boston, Massachusetts 02114, United States

Abstract. Understanding the human hearing process would be helped by quantification of the transient mechanical response of the human ear, including the human tympanic membrane (TM or eardrum). We propose a new hybrid high-speed holographic system (HHS) for acquisition and quantification of the full-field nanometer transient (i.e., >10 kHz) displacement of the human TM. We have optimized and implemented a $2 + 1$ frame local correlation (LC) based phase sampling method in combination with a high-speed (i.e., >40 K fps) camera acquisition system. To our knowledge, there is currently no existing system that provides such capabilities for the study of the human TM. The LC sampling method has a displacement difference of <11 nm relative to measurements obtained by a four-phase step algorithm. Comparisons between our high-speed acquisition system and a laser Doppler vibrometer indicate differences of <10 μs. The high temporal (i.e., >40 kHz) and spatial (i.e., >100 k data points) resolution of our HHS enables parallel measurements of all points on the surface of the TM, which allows quantification of spatially dependent motion parameters, such as modal frequencies and acoustic delays. Such capabilities could allow inferring local material properties across the surface of the TM. © The Authors.

1 Introduction

Ongoing hearing research efforts[1–5] to understand the mechanics of the human hearing process are mainly focused on the response of the ear to tonal acoustic stimulation. However, studying the transient response of the human ear, and of the human tympanic membrane (TM) in particular,[1] could expand the understanding of the processes by which acoustical energy is transformed to mechanical energy and transmitted to the ossicular chain.[1]

Measurements of transient phenomena on the surface of the TM and the quantification of corresponding motion parameters, such as traveling wave speeds,[4] standing wave ratios,[1,4] acoustic absorbance and immittance,[6] damping,[7] stiffness, and modal frequencies, could help, for instance, validate and improve physical models,[4,5] identify and diagnose pathologies, and improve the design of hearing aids.

Current state-of-the-art methods to measure the acousto-mechanical response of the TM rely on averaged acoustic[6] information or a local displacement response (at 1 to 50 points on the membrane) using single-point or scanning laser Doppler vibrometry[8,9] and/or capacitive probes.[10] The average or sparsely sampled points are not sufficient for the full description of the complex patterns unfolding across the full surface of the

TM. Therefore, there is a need for a full-field high-speed measurement method that can quantify the spatiotemporal complexity of the acoustically induced transient displacement of the TM.

Current state-of-the-art holographic methods for full-field measurements of transient events use phase sampling methods that constrain the maximum sampling speed[11] or require complex experimental setups,[12,13] such as custom apertures[14] or spatial filters,[12] that further increase the required illumination source power and/or cost of the system. Recent developments of hybrid spatiotemporal phase sampling methods[15,16] allow for single frame measurements of the deformation state of the object without any hardware modifications or resolution constraints on the camera.[11]

In this paper, we report on the development and implementation of a novel full-field high-speed holographic system (HHS) that utilizes an optimized hybrid $2 + 1$ frame local correlation (LC) phase sampling method.[17] HHS utilizes the temporal resolution of a high-speed camera [high-speed camera (HSC)] without imposing constraints on its spatial resolution[11] and without the need of specialized optical setups.[12–14] Automatic execution and synchronization of high-speed measurements is achieved by a modular control system. The high temporal (>40 kHz) and spatial (>100 k data points) measuring resolutions of the HHS enable the investigation of the complex spatiotemporal behavior of the TM at a sufficient level of detail to expand our knowledge of TM function and hearing mechanics.

*Address all correspondence to: Ivo Dobrev, E-mail: ivo_d@wpi.edu

2 Design Constraints for Transient Acquisition

The human ear is most sensitive to sounds of frequencies between 0.2 and 8 kHz.[18,19] To allow sufficient temporal resolution for the measurement of the instantaneous magnitude and phase[2,3] of the acoustically induced motion of the TM as well as its total harmonic distortion,[20] the acquisition system needs to capture at least 5 to 10 samples per cycle.[5] Thus, a sampling rate of 40 kHz is required to capture the motion produced by sounds with frequencies at the high end of this sensitive range.

If acoustic energy is propagated across the surface of the TM, such a propagation depends on wave speed and direction. Existing methods[5,21] based on steady-state measurements estimate a surface wave speed in the 5 to 70 m/s range, resulting in acoustic delays across the 8-mm-diameter membrane of 110 to 1600 μs. In order to reliably quantify the speed and direction of acoustic energy propagation across the surface of the TM, an acquisition method should allow the capture of at least three instances[22] of the spatiotemporal evolution of the surface waves within the duration of the acoustical delay. Considering a more conservative estimate for acoustic delay (i.e., ~100 μs), capturing three instances of the propagation of the surface waves would require <33 μs interframe time resulting in minimal sampling rate of 30 kHz.

The typical duration of an acoustical click response of the human TM is <5 ms.[23] At a minimal sampling rate of 30 kHz, the full duration of the transient will be represented within ~150 frames. The duration of the click response defines a minimal frequency resolution of >200 Hz for a single measurement. The frequency resolution of the acquisition method could be further improved by longer sampling or by capturing several consecutive click responses of the TM.

In order to establish a suitable exposure time to capture the transient response of the human TM, we use an analogy with previous stroboscopic holographic acquisition methods,[1–3] where the illumination duty cycle relative to a steady-state excitation frequency defines the effective exposure time per cycle. Using this analogy, for frequencies within the range of highest sensitivity of human hearing[18] (i.e., up to 8 kHz), a typical stroboscopic duty cycle[1–3] of 5 to 10% would correspond to an equivalent single frame exposure time of ~6 to 12 μs.

Table 1 summarizes the human hearing characteristics and the corresponding sampling parameters implemented in the HHS to achieve temporal resolutions >7 μs at a sampling rates >40 kHz and a measurement duration of 5 ms to provide a frequency resolution of 200 Hz.

3 Methods

3.1 Phase Sampling Based on 2 + 1 Frame Local Correlation

The 2 + 1 frame LC phase sampling method[17] is based on hybrid spatiotemporal phase sampling methods[15,16] that quantify changes in the state of deformation of an object throughout the duration of a transient event. In the LC method, two sets of frames are acquired. The first set consists of two temporally phase shifted reference frames gathered just before object deformation and the second set consists of a series of rapidly acquired frames gathered during object deformation. The method quantifies the double-exposure phase change at each instance by correlating the deformed and reference states. The correlation of individual frames relies on the assumption that the object is sufficiently stable during the acquisition of the two phase shifted reference frames, which we will demonstrate is the case in our measurements.

Consider two individually recorded camera frames at a reference, I_{ref}, and at a deformed, I_{def}, state with intensities

$$I_{\text{ref}} = I_{\text{r}} + I_{\text{o}} + 2\sqrt{I_{\text{r}} I_{\text{o}}} \cos(\theta), \tag{1}$$

$$I_{\text{def}} = I_{\text{r}} + I_{\text{o}} + 2\sqrt{I_{\text{r}} I_{\text{o}}} \cos(\theta + \phi), \tag{2}$$

where I_{r} and I_{o} are the intensities of the reference and object beams, respectively, θ is the initial random phase difference between the two interfering beams, and ϕ is the phase change corresponding to the object deformation. By assuming that only ϕ varies, the correlation function, $\rho(I_{\text{ref}}, I_{\text{def}})$, can be expressed as[24]

$$\rho(I_{\text{ref}}, I_{\text{def}}) = \frac{1 + r^2 + 2r \cos(\phi)}{(1 + r)^2}, \tag{3}$$

with $r = I_{\text{r}}/I_{\text{o}}$. The $I_{\text{r}} + I_{\text{o}}$ terms, estimated by temporally averaging each measurement point and corresponding to a constant background illumination (DC), can be subtracted from Eqs. (1) and (2) to yield an equation equivalent to Eq. (3):

Table 1 Identification of the high-speed holographic system sampling parameters to investigate the transient behavior of the human tympanic membrane (TM).

Human hearing parameters	Value	Acquisition design constraints	Value
Frequency range	0.2 to 8 kHz[18]	Sampling rate	1 to 40 kHz[a]
		Exposure time	<6 to 12 μs (Refs. 1 to 3)[a]
TM's acoustical delay	40 μs (Ref. 21) to 1.3 ms (Ref. 5)	Minimum sampling rate	>30 kHz
Click response duration	<5 ms settling time[23]	Number of samples	~150 frames at 30 kHz
		Frequency resolution	200 Hz

[a]With five samples per cycle and 5 to 10% duty cycle.

$$\rho(I_{\text{ref}}, I_{\text{def}}) \approx \cos(\phi), \tag{4}$$

which contains information about the deformation of the object. Furthermore, acquiring a second reference state with a $\pi/2$ phase shift,

$$I_{\text{ref}+\pi/2} = I_{\text{r}} + I_{\text{o}} + 2\sqrt{I_{\text{r}}I_{\text{o}}} \cos\left(\theta + \frac{\pi}{2}\right), \tag{5}$$

and correlating it with a deformed state represented by Eq. (2) leads to

$$\rho(I_{\text{ref}+\pi/2}, I_{\text{def}}) \approx \cos\left(\phi + \frac{\pi}{2}\right) = \sin(\phi). \tag{6}$$

By assuming a constant beam ratio, ϕ, the phase change of interest, ϕ, in space, (m, n), corresponding to the deformation of the object at a specific time instance, ϕ, can be computed by the combination of Eqs. (4) and (6) as

$$\phi(m, n, t) = \tan^{-1}\left[\frac{\rho(I_{\text{ref}+\pi/2}, I_{\text{def}})}{\rho(I_{\text{ref}}, I_{\text{def}})}\right]. \tag{7}$$

The computational process to recover $\phi(m, n, t)$, schematically illustrated in Fig. 1, utilizes data sets from two phase shifted reference images, I_{ref} and $I_{\text{ref}+\pi/2}$, recorded at times t_{ref} and $t_{\text{ref}+\pi/2}$ before object excitation and from deformed images, I_{def}, at time instances, t_{def}, during and after object excitation. A temporally estimated DC is subtracted from the intensity values of I_{ref}, $I_{\text{ref}+\pi/2}$, and I_{def}. The DC-compensated frames I_{ref} and I_{def} are then used for the evaluation of the correlation coefficient, $\rho(I_{\text{ref}}, I_{\text{def}})$, at every measurement point on the frame. The numerical calculation of $\rho(I_{\text{ref}}, I_{\text{def}})$ is implemented based on a computationally efficient Pearson product-moment correlation coefficient method[17,25] that uses sets of intensity values from spatial kernels (e.g., 3×3 pixels) around each measurement point of frames I_{ref} and I_{def}. An analogous procedure is applied to the evaluation of the Pearson correlation
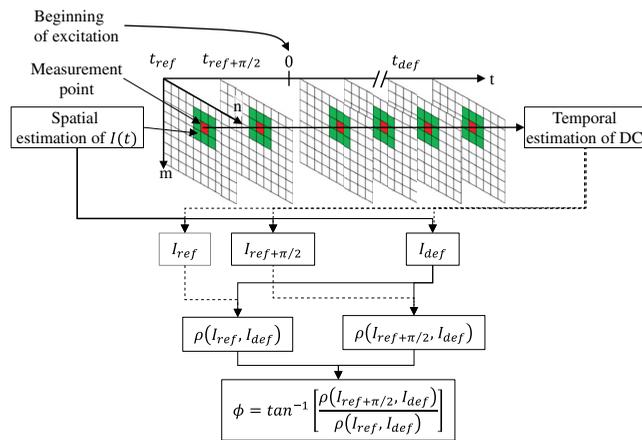

Fig. 1 Flow chart illustrating the pointwise implementation of the $2+1$ frame local correlation (LC) phase sampling algorithm to recover the phase change, $\phi(m, n, t)$, corresponding to transient deformations of an object. The correlation coefficient, ρ, at a measurement point between any two consecutive frames is based on the evaluation of intensities defined by a kernel within the proximity of the measurement point. The zero of the time axis corresponds to the initiation of the acoustic excitation and signifies the first of the set of deformed frames.

coefficient for $\rho(I_{\text{ref}+\pi/2}, I_{\text{def}})$. By computing $\rho(I_{\text{ref}}, I_{\text{def}})$ and $\rho(I_{\text{ref}+\pi/2}, I_{\text{def}})$ for each postdeformation measured time point using Eq. (7), $\phi(m, n, t)$ is computed for modulus 2π, which requires further processing to obtain a continuous phase distribution by application of spatiotemporal phase unwrapping algorithms.[2,26]

The pointwise evaluation of the correlation coefficients assumes that $\phi(m, n, t)$ varies slower in space (m, n) than all of the other parameters in Eqs. (1) and (2) and is sufficiently constant within a spatial kernel[16] (i.e., small spatial neighborhood of measurement points). This assumption is adequate for the acoustically induced response of the human TM under typical loading conditions since the resulting fringe density within the hologram is small (<4 fringes/mm) compared to the spatial resolution (>25 lp/mm) of a typical HSC.[25]

The use of a spatial kernel to represent each measurement point in the LC method is equivalent to applying a spatial band-pass filter. In addition, the intensity values within the kernel can be individually weighted to obtain the frequency response of specific spatial filters, such as mean or Gaussian.[16] Variation of the size of the spatial kernel also allows for control of the spatial frequency range of filtration.

The LC algorithm we describe is independently executed at every measurement point, which allows for the simultaneous evaluation of ϕ at all measurement points across each deformed frame, I_{def}, in parallel. By using this computational parallelism, we have developed multithreaded and GPU accelerated software that improves computational speed and efficiency.

3.2 High-Speed $2+N$ Frame Acquisition

To measure the transient deformations of the TM, a high-speed $2+N$ frame acquisition approach based on the $2+1$ frame LC phase sampling method has been developed and implemented. In this approach, two reference frames, I_{ref} and $I_{\text{ref}+\pi/2}$, and N consecutive frames, $(I_{\text{def}})_{i,i\in 1,2\ldots N}$, are recorded before and throughout the evolution of an event.

Figure 2 shows the timing diagrams of the events that occur during the high-speed $2+N$ acquisition, including camera exposure, acoustic excitation, phase shifting, and the TM's response. According to this diagram, the two reference frames, I_{ref} and $I_{\text{ref}+\pi/2}$, are recorded with a temporal separation that is longer than the piezoelectric transducer (PZT) settling time after the introduction of the initial $\pi/2$ phase shift.[27]

For efficient acquisition of $(I_{\text{def}})_i$ frames, the PZT is kept at its final position after phase shifting, which leads to the $(I_{\text{def}})_i$ frames containing a constant $\pi/2$ phase offset expressed as $(I_{\text{def}+\pi/2})_i$. Therefore, the optical phase, ϕ, within any $(I_{\text{def}+\pi/2})_i$ frame, and corresponding to the deformation of the TM at a specific instance, can be determined from Eq. (7) with

$$\phi(m, n, t) = \tan^{-1}\left[\frac{\rho(I_{\text{ref}}, I_{\text{def}+\pi/2})}{\rho(I_{\text{ref}+\pi/2}, I_{\text{def}+\pi/2})}\right]. \tag{8}$$

All deformed frames, $(I_{\text{def}+\pi/2})_{i,i\in 1,2\ldots N}$, before and throughout the evolution of the transient response of the TM, are relative to the same reference frames, I_{ref} and $I_{\text{ref}+\pi/2}$, before acoustic excitation.

The camera continuously records frames during the full duration of the measurement, including during the settling time of the PZT, with the maximum sampling rate of acquisition constrained by the frame rate of the camera. During a typical
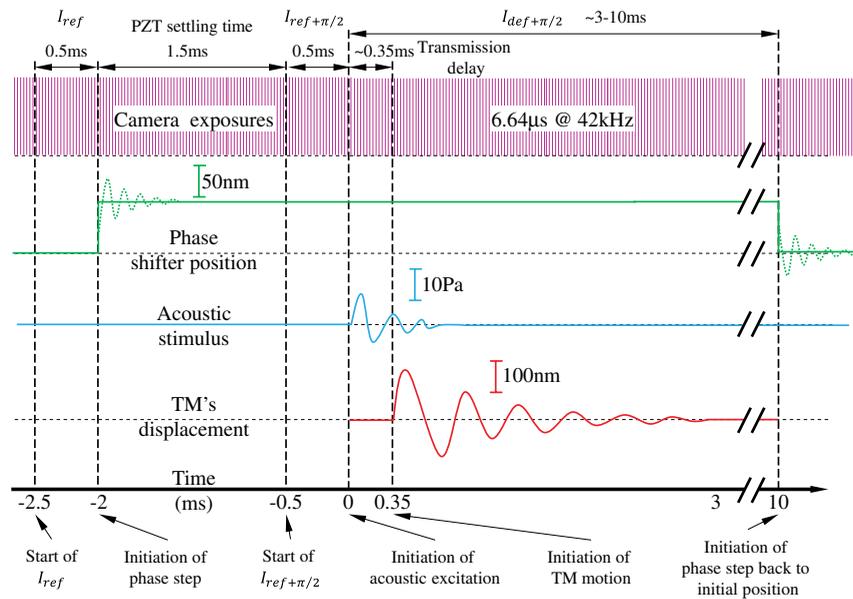
Fig. 2 Timing diagrams of the events occurring during high-speed $2 + N$ frame acquisition with the LC phase sampling method. N deformed frames are acquired at the maximum frame rate of the camera and referenced relative to frames captured during two periods prior to acoustic excitation. Acquisition takes into account the settling time of the PZT phase shifter, while minimizing the overall acquisition time. PZT is returned to its original position at the end of the acquisition. Due to the distance between the sound source and the sample, there is an acoustic transmission delay. Representative values of major parameters are indicated.

acquisition, multiple frames (i.e., ~20 frames at 42 k fps) at each reference position are captured and temporally averaged in order to compensate for potential external disturbances. In addition, and by confirmation with laser Doppler vibrometer (LDV) measurements, 1.5 ms are allocated for the settling time of the PZT and the frames recorded over this time are discarded.

3.3 Hardware and Software Implementation of the Acquisition

The hardware implementation of the high-speed acquisition involves the synchronization of the procedures for acoustic excitation of the sample, temporal phase shifting, and frame
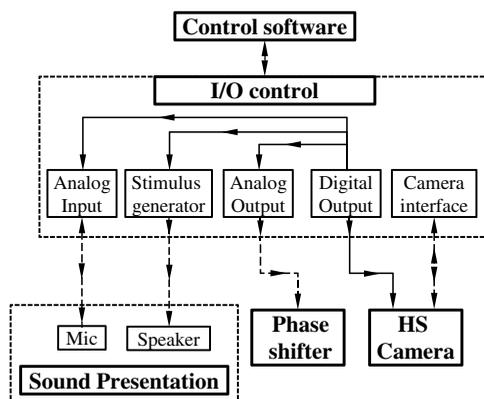

Fig. 3 Schematic showing the major modules of the high-speed acquisition hardware and software. The components within the I/O control module communicate with the sound presentation and measurement, phase shifter, and HS camera (HSC) modules that are synchronized by the digital output component serving as the master clock.

acquisition. This is achieved by the development and implementation of a control architecture that consists of four major hardware modules, which include electronic I/O control (I/O), sound presentation and measurement (SP), phase shifter (PS), and HSC, as shown in Fig. 3.[17,25,26] Hardware modules are managed by a unified control software with a user interface for setting and customizing acquisition and synchronization parameters.[28]

To perform high-speed transient measurements of a TM sample, the user is required to specify the following parameters:

- Frame acquisition: starting time, frame rate, spatial resolution, exposure time, and number of frames to be acquired.

- Temporal phase shifting: starting time.

- Acoustic excitation: starting time, transient excitation type (i.e., click, chirp, sine), duration, and signal input level.

During a measurement, the control software utilizes the digital output component of the I/O module to automatically trigger the execution of all procedures through correlated timing signals with a temporal accuracy >1 μs. The applied sound pressure level (SPL) and frequency content of the acoustic excitation is automatically quantified with a calibrated microphone within the SP module interfaced to an analog input port of the I/O module.[5,29]

4 Validation of the Measurement System

In order to reliably use the HHS for the characterization of TM samples, its measuring capabilities are validated using other methodologies that include LDV and holographic methods based on temporal phase shifting. In particular, we quantify the performance characteristics of the developed phase sampling and acquisition system using different test metrics.
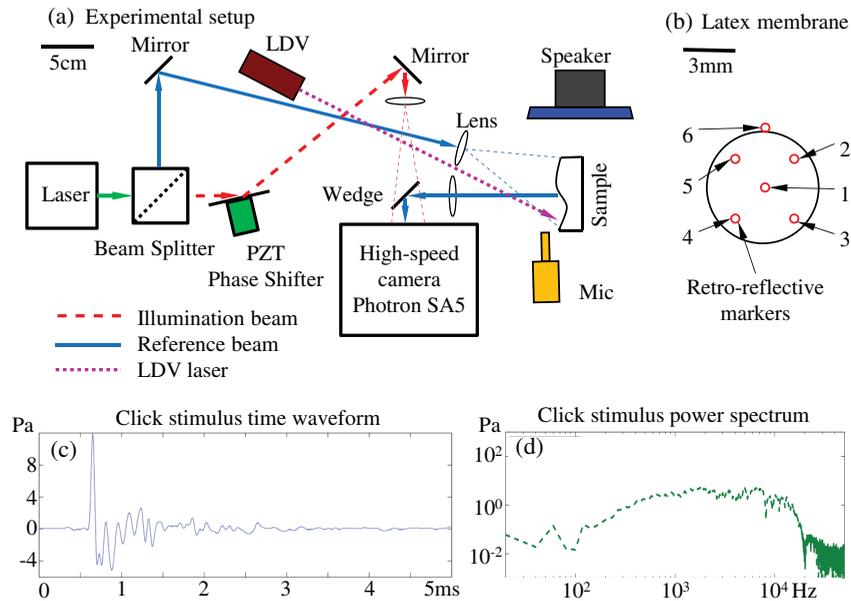
Fig. 4 Experimental setup and artificial sample: (a) schematic of the HHS setup that includes sound presentation, laser delivery, and high-speed camera modules; (b) latex membrane with six retroreflective markers; (c) and (d) time waveform and power spectrum of an average of 20 acoustic clicks indicating 10 Pa or 115 dB peak sound pressure level (SPL) based on a calibrated microphone. The wedge in the HSC module redirects 95% of the object beam and 5% of the reference beam power to the sensor of the HSC. A laser Doppler vibrometer (LDV) is incorporated for characterization of the HHS measuring capabilities.

- 2 + 1 frame LC phase sampling

  - Noise floor—quantification of the displacement of a static object with no external excitation.

  - Displacement accuracy—comparison with four-frame temporal phase sampling measurements of the steady-state displacement of a statically loaded membrane.

- High-speed acquisition

  - Phase stepping at high speed—quantification of the time constant and settling time of the response of the PS during the phase stepping procedure as described in Sec. 3.3.

  - Temporal accuracy—comparison of HHS and LDV measurements of displacement and velocity time-waveforms in a latex membrane sample excited by an acoustic click.

4.1 Experimental Setup

The components of the different modules of the HHS are shown in Fig. 4. The sound presentation module consists of a calibrated microphone (Etymotic Research ER-7C, Elk Grove Village, Illinois) and a speaker (SB Acoustics SB29RDC-C000-4, Brookfield, Wisconsin).[29] The laser delivery module includes a continuous wave laser (Oxxius SLIM-532, 50 mW, Lannion, France), variable ratio beam splitter, and a PZT PS. The HSC module is a Photron SA5 1000k (Tokyo, Japan) (42 K fps at 384^2 pixels). For validation of the transient

measurement capabilities of the HHS, an LDV is incorporated within the HHS experimental setup.

A schematic of the 10-mm-diameter circular latex membrane used for transient measurements is shown in Fig. 4(b). Retroreflective markers were applied at several predefined points on the surface of the sample in order to improve the signal quality for LDV measurements. Five markers were distributed equidistantly to allow for even coverage of the latex surface (points 1 to 5), as shown in Fig. 4(b). In order to monitor any rigid body motion, a sixth marker was placed at the top on the locking ring that clamped the latex membrane to its cylindrical support (point 6).

Acoustic excitation for all dynamic measurements was based on a 50-μs acoustic click produced by the signal generator in the I/O module and introduced to the sample by the speaker in the SP module. The time waveform and power spectrum of an average of 20 clicks produced by the speaker are recorded by the calibrated microphone and are shown in Figs. 4(c) to 4(d).

4.2 Validation of the 2 + 1 Frame LC Phase Sampling Method

The noise floor and accuracy of the 2 + 1 frame LC phase sampling method were quantified independent of the high-speed acquisition method through a set of experiments performed in static conditions.

4.2.1 Noise floor

In order to quantify the noise floor of the 2 + 1 frame LC phase sampling method, we measured the displacement of a solid object (i.e., 12.7-mm-thick aluminum plate) under static conditions and no external excitation. We assumed that the changes of
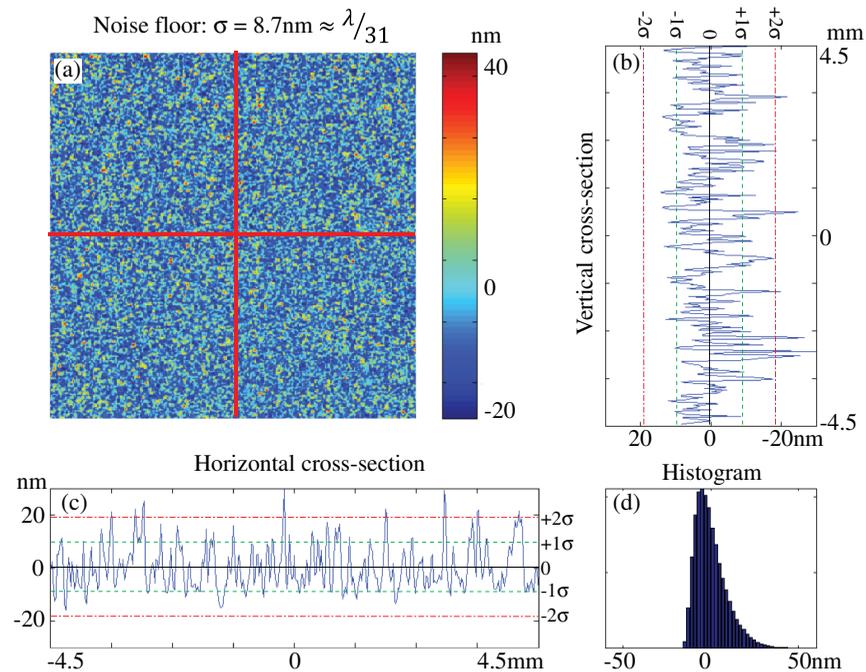
Fig. 5 Representative results of the spatial distribution of the noise floor of the 2 + 1 frame LC phase sampling method under static conditions and no external excitation of the object: (a) map of the spatial distribution of the noise signal; (b) horizontal and (c) vertical cross-sections of (a); and (d) histogram of (a). The ±1σ (dashed line) and the ±2σ (dash dot line) of the noise measurement map cross-sections are highlighted.

the sample during the measurements are insignificant and any detected deformations were associated with the noise floor of the phase sampling method. During the noise floor measurements, all optical and sampling conditions were kept identical to ones used with the high-speed acquisition. The spatial distribution of the noise floor, as shown in Fig. 5, indicates a standard deviation (STD) of 8.7 nm or $\lambda/31$.

4.2.2 Displacement accuracy

To assess the accuracy of the 2 + 1 frame LC phase sampling method, we compared the displacement measurements of a statically deformed sample against a four-frame temporal phase sampling method.[2,3] In order to maintain identical conditions for both phase sampling methods, we used two phase shifted references and one of the deformed frames of the four-frame data set and applied it to the 2 + 1 frame LC phase sampling method. Because of the static loading conditions and the longer (>100 ms) measurement time of the four-frame acquisition method, the camera's sampling rate was reduced to 60 Hz (the minimum for Photron SA5).

The sample was a 12-mm-diameter aluminum membrane under static loading with a constant force applied normal to the back surface. Aside from the sampling rate, the optical setup and illumination conditions for the accuracy measurements were identical to the high-speed acquisition settings.

The comparison of the modulation and phase maps of the four-frame double exposure and 2 + 1 frame LC phase sampling methods is shown in Fig. 6. Based on the phase maps, the corresponding object displacements are calculated and their differences are shown in Figs. 6(e) and 6(f). The STD of the

displacement differences, which can be used to assess the accuracy of the LC method, is within 11 nm or $\lambda/25$.

4.3 Validation of the High-Speed Acquisition

4.3.1 Phase stepping at high speed

In order to optimize the timing of the high-speed acquisition method, we characterized the dynamic response of the PS (measured at the center of the mirror) by measuring its velocity time waveform with an LDV, as shown in Fig. 7. Analysis of the data indicates a 0.5 ms time constant, with <1.5 ms settling time with 6% residual fluctuations, and a noise floor of <4% of the maximum response.

4.3.2 Temporal accuracy

We characterized the temporal accuracy of the high-speed acquisition method relative to an LDV by correlating coincident LDV and HHS measurements of controlled transient acoustic stimulation of the latex sample. The LDV sampled the motion at the six locations described in Sec. 4.1. In order to characterize the accuracy in position and velocity, and due to the HHS and LDV measuring domains, time waveforms were either differentiated or integrated.

The HHS was set at a sampling rate of 42 kHz (corresponding to 23.8 μs interframe time) with 6.62 μs exposure time (corresponding to a shutter speed of 151 kHz), in order to be consistent with the sampling parameters specified in Table 1. With these settings, the HSC has a spatial resolution of 384 × 384 pixels corresponding to a spatial resolution of >25 lp/mm at the object plane. The LDV was sampled at 84 kHz through a 16-bit analog input at the I/O module. Representative HHS and
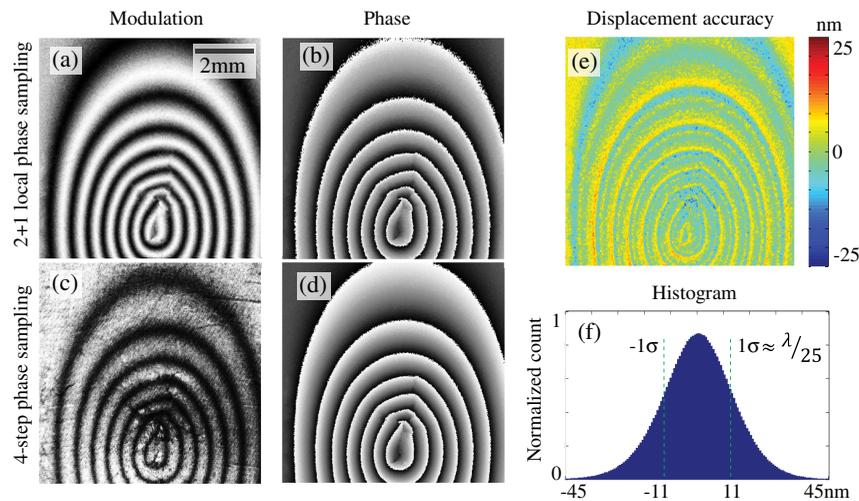
Fig. 6 Characterization of the accuracy of $2 + 1$ LC relative to the four-frame phase sampling method: (a) modulation and (b) phase of the $2 + 1$ LC phase sampling method; (c) modulation and (d) phase of the four-frame phase sampling method; (e) displacement differences between the two methods; and (f) histogram of (e). The difference in displacement has a mean of 0 with a standard deviation of 11 nm or $\lambda/25$.
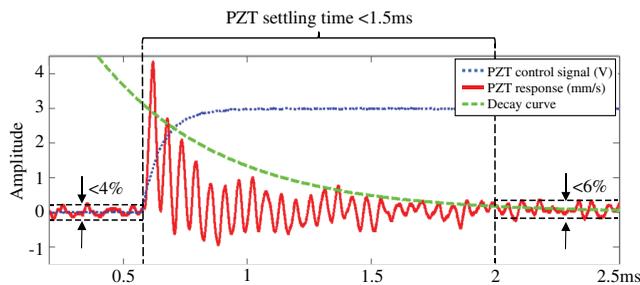

Fig. 7 Representative results of the velocity time-waveform of the response of the PZT phase shifter (PS) indicating 0.5 ms time constant and <1.5 ms settling time with <6% residual vibrations relative to the maximum response. The PZT control input (dash dot line), PZT velocity response (solid line), and fitted decay curve (dashed line) are shown.

LDV displacement and velocity results, depicted in Fig. 8, show the time waveforms at two of the discrete locations of the sample, as identified in Sec. 4.1, which correspond to points with the minimum (point 5) and maximum responses (point 1). The acoustical excitation was a 50-μs click with a peak SPL of 104 dB.

Correlation between the time waveforms is, on average, >95% for both displacements and velocities at all points (1 to 5) on the surface of the sample, as specified in Sec. 4.1. The average differences between the time waveforms are <5% relative to the p-p response of the sample. Differences between temporal locations of the local maxima and minima in the time waveforms of the HHS and LDV are within 10 μs.

5 Applications

Having validated the phase sampling and the high-speed acquisition methods, we utilized HHS to quantify the transient response of the surface of a human TM sample excited by click-like acoustic stimuli. We present preliminary results on the quantification of several temporal acousto-mechanical properties as well as their spatial dependency.

The human TM sample used in our measurements is part of a temporal bone from a 46-year-old female donor. The sample was prepared in accordance with previously established procedures.[1,5] The surface of the TM was coated with a solution of ZnO to improve the surface reflectivity and reduce required camera exposure times, resulting in better temporal resolutions, while having minimal change[5] (-0.41 ± 4.4 dB) in the acoustic
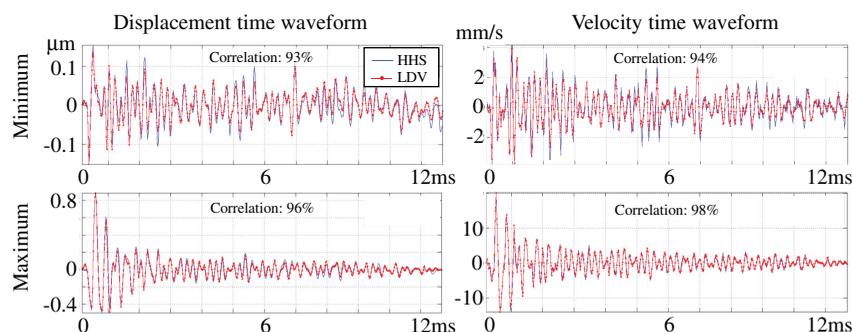

Fig. 8 Representative results of the temporal accuracy of the high-speed acquisition relative to LDV based on the minimum (point 5) and maximum (point 1) responses of the latex membrane shown in the top and bottom rows of graphs, respectively. Acoustic excitation was a 50-μs click with a 104 dB maximum SPL. The time axis is relative to the beginning of the transient acoustic stimuli. Correlation between the time waveforms of each method is on average >95% for both displacement and velocity.
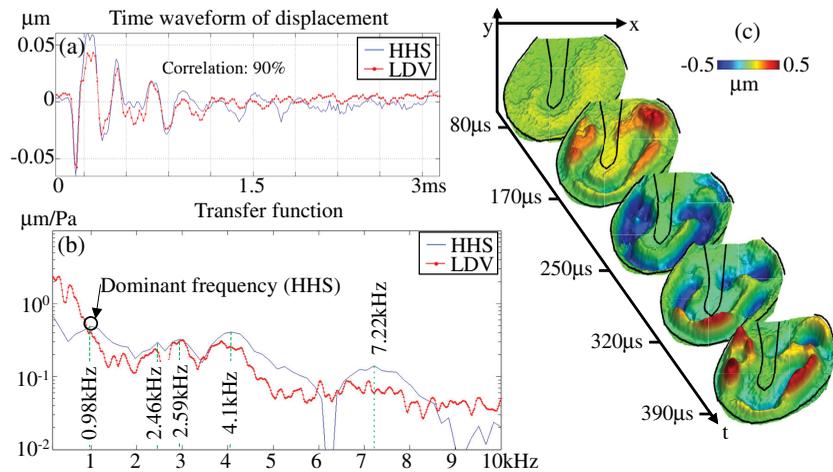
Fig. 9 Representative measurement of the spatial, temporal, and frequency measurement capabilities of the HHS based on the acoustically induced transient response of a human TM sample: (a) displacement versus time waveform of the umbo measured with HHS and LDV indicating 90% correlation of the time waveforms; (b) power spectrum of the displacement transfer function at the umbo measured with HSS (solid line) and LDV (dotted line); and (c) full-field HHS displacement maps at five instances 80 to 390 μs after the arrival of the acoustic pulse indicating a 0.83 μm p-p maximum along the periphery of a human TM. The excitation was a 50-μs click with a peak SPL of 115 dB. Dashed lines in (b) refer to automatically determined modal frequencies from the HHS measurement. With the exception of the lowest modal frequency (0.98 kHz, which was not identifiable in the LDV), there was a <5% difference between HHS and LDV determined modal frequencies. Outlines of the outer boundary of the membrane and the manubrium (the handle of the malleus that is attached to the TM) in (c) are indicated with solid lines.

response (volume displacement) of the TM. Complimentary to each HHS measurement, we conducted LDV measurements of the TM sample at the approximate center of the membrane near the umbo (a connection point between the TM and the inferior tip of the malleus—the most lateral ossicle) and in the anterior half of the TM. In order to monitor the rigid body motions of the temporal bone, a third marker was placed on the temporal bone close to the boundary of the TM. Retroreflective markers were placed at each LDV measurement location in order to improve the signal quality of the LDV measurement.

The HHS experimental setup was identical to the one used for the latex membrane measurements described in Sec. 4.1. The speaker used for acoustic click stimuli was positioned ~12 cm away from the sample, resulting in a 350 μs acoustical transmission delay between the speaker and the TM surface. Because of that, the beginnings of all time axes in this section are referenced to the estimated time of arrival of the sound pressure front to the sample surface. The acoustic excitation used for all human TM measurements was a 50-μs click with a peak SPL of 115 dB.

LDV measurements of the TM at retroreflective markers at the umbo (Fig. 10) and on the anterior TM, indicate a total duration of the transient event of <3 ms. Frequency analysis of the response indicates significant frequency content up to 11 kHz. The delay between the acoustical excitation and the response of the sample was within the estimated transmission delay between the sample and the speaker, as indicated in Sec. 4.1. The LDV was sampled as described in Sec. 4.3.2.

In order to adequately sample the full-field transient response of the human TM, we adjusted the HHS camera's sampling rate, exposure time, and recording duration in accordance with acquisition design constraints, specified in Sec. 2, as well as with the preliminary LDV measurements. The HHS was set at a sampling rate of 42 kHz (corresponding to 24 μs interframe time) with a 6.62 μs exposure time (corresponding to a shutter speed of 151 kHz), in order to be consistent with the sampling

parameters specified in Table 1. The duration of every measurement was set to 1000 frames corresponding to a period from −2.5 to ~20 ms relative to the beginning of the acoustic excitation. While the settling time of the human TM is <5 ms, as described in Table 1, more frames are collected after the settling of the sample to allow for better DC estimation, as described in Sec. 3.1. The frames acquired before the acoustic excitation account for the effects of external disturbances on the reference images (Sec. 3.2) and the PZT settling time (Sec. 4.3.1).

5.1 Full-Field-of-View Transient Displacement of the Human TM

The HHS provides high temporal (>42 kHz) and spatial (>100 k data points) resolutions that enable the simultaneous measurements of the spatial distribution and the temporal and frequency domain of the displacement of transient response of the human TM. Representative measurements of the spatial, temporal, and frequency measurement capabilities of the HHS, based on the acoustically induced transient response of a human TM, are shown in Fig. 9. Displacement time waveforms of the response of the umbo, as shown in Fig. 9(a), indicate >90% correlation between the HHS and LDV measurements above the noise floor (<11 nm) of the HHS. The average correlation of the time waveforms of the HHS and LDV at points (1 and 2) on the TM's surface was >94%. According to Figs. 8 and 9(a), it is observed that the magnitude of the correlation coefficient is related to the magnitude of the response. Lower correlation levels (i.e., ~90%) are associated with displacement measurements closer to the noise floor (<11 nm) of the HHS, as shown in the waveform corresponding to the minimum response of the human TM in Fig. 9(b).

Based on the frequency domain of the HHS and the calibrated microphone measurements of the displacement and acoustical pressure time waveforms, respectively, the transfer
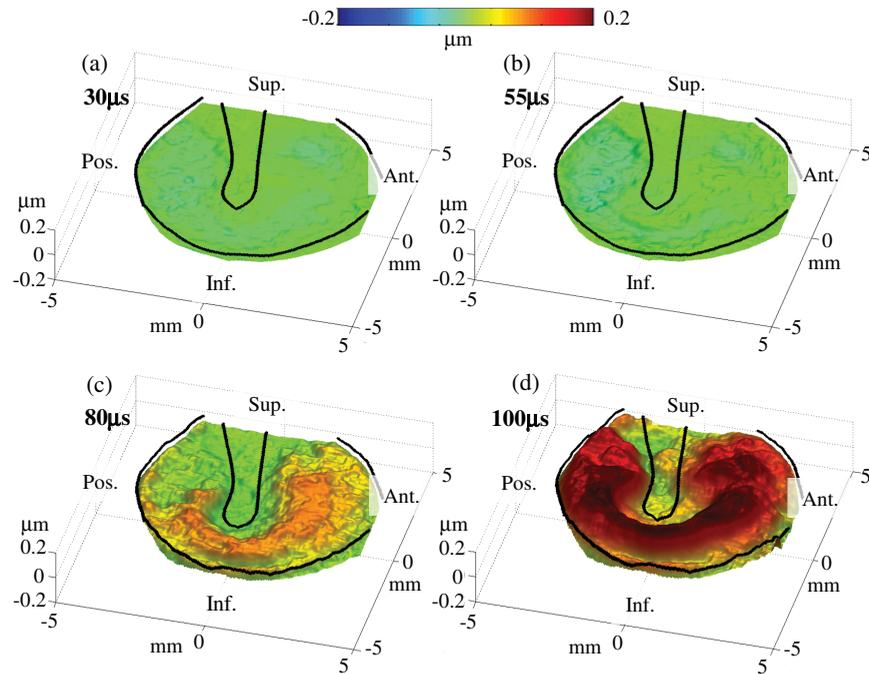
Fig. 10 Representative results of the 30 to 100 μs of the initial stages of the transient response of the TM indicating mostly in-phase displacement across the full surface. The acoustic excitation was a 50-μs click with a 115 dB maximum SPL. Outlines of the outer boundary of the membrane and the manubrium are indicated with solid lines.

function (TF) of each point across the surface of the TM can be calculated,[5] and the frequencies of maximum motion associated with the modal frequencies can be identified by automatic quantification of the local maxima of the TF, as shown in Fig. 9(b). Due to the short duration of the transient event (i.e., <5 ms), frequency components <200 Hz (i.e., >5 ms period) are disregarded and are not considered in the modal frequencies' identification. The detected modal frequencies at the umbo, shown in Fig. 9(b), based on the HHS and LDV differ by <5%. The difference between magnitude of the TF measured with HHS and LDV in the frequency range of 1 to 8 kHz is within 5 dB. Based on measurements on a latex membrane, described in Sec. 4.3.2, the noise floor of the TF was estimated as ~20 dB below the average signal level for frequencies <8 kHz. The difference between the HHS and LDV at ~6 kHz could be associated with the acoustical properties of the experimental setup as well as the response of the speaker used. In particular, the speaker could have band gaps in its response, which could result in local higher noise floor (lower signal-to-noise ratio) for the HHS, while the LDV data have been averaged over multiple runs (i.e., >10) and exhibit fewer effects from that phenomenon. Future work will be focused on specifying an acoustic source with a flatter click response.

Based on the spatiotemporal evolution of the HHS displacement measurements, as shown in Fig. 9(c), it can be seen that the transient response of the human TM undergoes two distinct stages—global initiation of the surface motion and local surface wave propagation. The first 30 to 100 μs of the initial stages of the transient displacement of the human TM, shown in Fig. 10, indicate motion that is approximately in-phase across the full surface of the visible TM. It can be seen that there is <30 μs delay between the arrival of the acoustic pressure front at the

surface of the TM and the beginning of its transient response. This delay can be contributed to the temporal resolution of the HSS system (24 μs interframe time) and the reaction time of the speaker.

The further spatiotemporal evolution of the transient response of the TM indicates circumferentially traveling surface waves propagating symmetrically relative to the manubrium and radially from the inferior to the superior parts of the TM, as shown in Fig. 11. The local phase velocity of the surface waves can be estimated by automatically identifying the shift of the spatial location of the local minima and maxima of the displacement maps between successive frames. Preliminary surface wave speed estimations indicate 24 m/s, which is in agreement with previous research.[5,21]

5.2 *Acoustic Delay and Dominant Modal Frequency Maps*

The HHS provides the time waveforms of the transient response of all points across the surface of the human TM sample, allowing for estimation of the spatial dependence of motion parameters, such as acoustic delays and dominant modal frequencies, as shown in Fig. 12.

The acoustical delay map, as shown in Fig. 12(a), is calculated by automatically identifying the peak time of the first local extrema of the time waveform of every point across the surface of the TM and referencing it to the peak time of the umbo, indicated with $\oplus$ in Fig. 12(a). This quantifies the spatial distribution of the delay of the time-domain response of each point relative to the umbo. The range of the measured acoustical delays across the surface of the human TM, as shown in Fig. 12(b), is within previously reported data.[5,21]
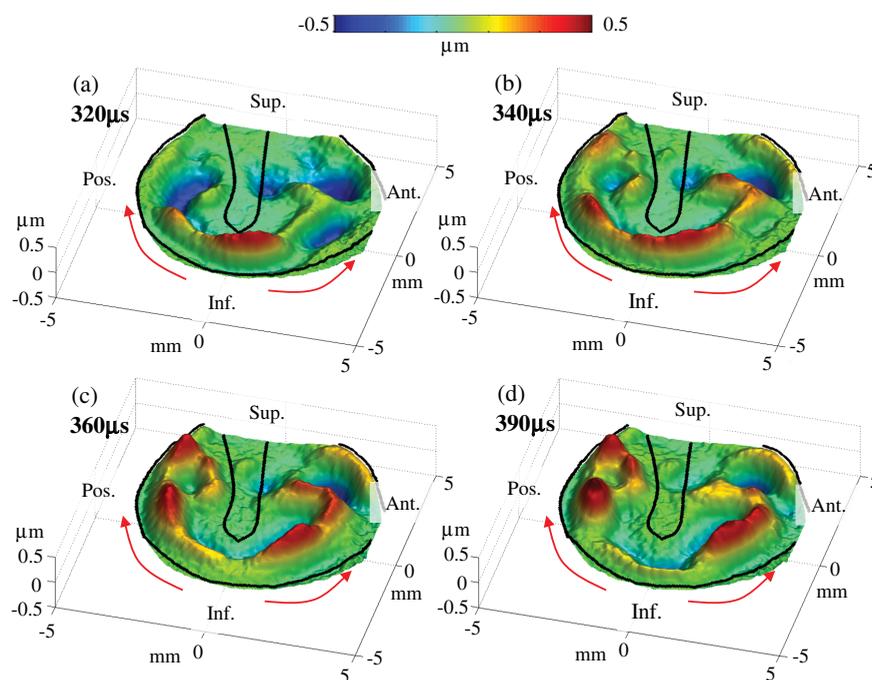
Fig. 11 Representative results at 320 to 390 μs of the transient response of the surface of the TM indicating circumferentially traveling surface waves propagating at 24 m/s symmetrically relative to the manubrium and radially from the inferior to the superior parts of the TM as indicated by solid arrows. The acoustic excitation is a 50-μs click with a peak SPL of 115 dB. Outlines of the outer boundary of the membrane and the manubrium are indicated with solid lines.

The acoustic delay map and histogram indicate that the peak motion response of >50% of the surface is within ±50 μs of the peak motion of the umbo, which supports the observations of predominantly in-phase motion during the initial response of the surface of the TM, as shown in Fig. 10. The acoustic delay data, shown in Fig. 12(a), also indicate that the surface of the TM at the interior and posterior boundary moves with a −25 μs acoustical delay relative to the umbo. This suggests that the acousto-mechanical response of the TM reaches its
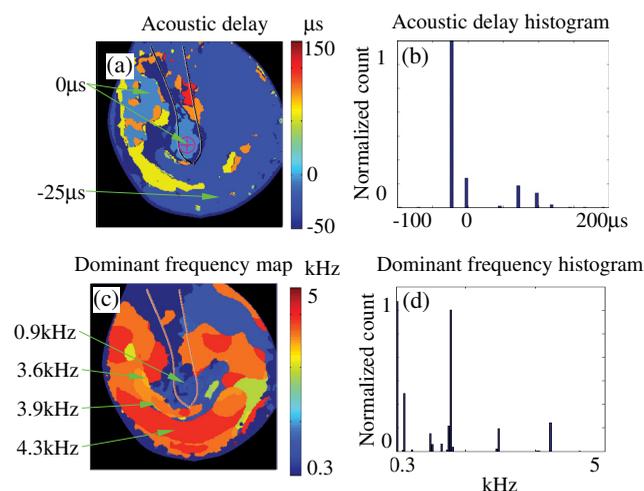

Fig. 12 Spatial dependence of the (a) acoustic delay of each point on the TM relative to the peak time of the center of the umbo (marked with ⊕); histogram of (b); spatial distribution map of the dominant frequency at each point on the surface of the TM; and (d) histogram of (c). Outline of the manubrium is indicated with a solid line.

first extrema in the region between the TM boundary and the umbo, which agrees with theories of acousto-mechanical energy transfer from the TM periphery to the umbo.[4] However, since our measure of delay depends on the time to the first temporal extrema, it includes the time associated with responses to different natural frequencies.

Based on automatic identification of the local maxima of the TF, shown in Fig. 9(b), at every point across the TM, we can determine the dominant modal frequency and its spatial distribution, as shown in Fig. 12(c). The dominant frequency map indicates noticeable differences (three- to fourfold) between the regions of the TM near the manubrium and the central region midway between the manubrium and the TM boundary. Assuming that the spatial distribution of the dominant frequency is representative of the local variations of stiffness and thickness of the TM, our observations can be related to previous studies indicating more than a threefold increase in the local thickness of the TM near the manubrium relative to the central region of the TM.[30]

In order to relate the data in Figs. 12(a) and 12(c), we assume that the initial displacement (i.e., <100 us) of the click response of the TM, as shown in Fig. 10, exhibits a single tone decayed response based on a dominant frequency spatially varying across the TM as shown in Fig. 12(c). Based on this assumption and the approximately in-phase start of the motion of all points suggested by Fig. 10, the acoustical delay (i.e., the temporal location of the first local extrema) between any two points should be within a ¼ of the difference of the periods of oscillation of the points. Figure 12 indicates dominant frequencies of ~1 kHz (~250 μs for ¼ cycle) at the umbo and an average of 3.5 kHz (~70 μs for ¼ cycle) across the central region midway between the manubrium and the TM boundary. This suggests ~180 μs of maximum acoustic delay that can be associated

with the difference in the natural frequency of the response between the umbo and the region midway between the umbo and the rim. Such a delay is in agreement with the range of the measured delay to the first extrema (-25 to 75 μs) indicated in Fig. 12(b).

6 Conclusions and Future Work

In this paper, a new method is proposed to quantify the full-field transient dynamics of the TM using an HHS and a hybrid $2 + 1$ frame LC phase sampling algorithm that utilizes the temporal resolution of an HSC without imposing constraints on its spatial resolution. We have also developed a customizable modular control system for high-speed acquisition within the HSS.

The HHS provides simultaneous high-speed (i.e., >40 kHz) measurement of the motion of >100 k data points on the surface of the TM allowing for a $>10^3$-fold decrease in measurement time compared to existing stroboscopic holographic measurement methods.[2,3,5] This reduces the effects of the environmental variations on the acoustic response of the samples and allows for applications *in vivo*.[17]

Analysis of the transient response of every point across the surface of the TM in the time and frequency domains allows for observations of spatially dependent motion parameters, such as modal frequencies and acoustical delays. These observations can then be used to infer local material properties across the surface of the TM. The observations of the transient dynamics of TM surface motion from this study could further the understanding of the sound-receiving function of the TM and how it couples acousto-mechanical energy to the ossicular chain and inner ear. The HSS could provide a new tool for the investigation of the auditory system with applications in research, medical diagnosis and hearing aid design.

Future work should be focused on the analysis and interpretation of the measured transient displacement time waveform to extract medically meaningful information on the TM's health condition. Further research is also needed to explain the initial transient dynamics of the TM and its relationship to the energy transfer into the middle-ear, as well as its connection to previous steady-state dynamics research. Improvements of the HHS should include optimization of the optical design and spatial resolution, as well as its packaging for *in vivo* applications and for medical research.

References

1. J. J. Rosowski et al., "Computer-assisted time-averaged holograms of the motion of the surface of the mammalian tympanic membrane with sound stimuli of 0.4–25 kHz," *Hear. Res.* **253**(1), 83–96 (2009).

2. J. J. Rosowski et al., "Measurements of three-dimensional shape and sound-induced motion of the chinchilla tympanic membrane," *Hear. Res.* **301**, 44–52 (2013).

3. M. Khaleghi et al., "Digital holographic measurements of shape and 3D sound-induced displacements of tympanic membrane," *Opt. Eng.* **52**(10), 101916 (2013).

4. S. Puria and J. B. Allen, "Measurements and model of the cat middle ear: evidence of tympanic membrane acoustic delay," *J. Acoust. Soc. Am.* **104**(6), 3463 (1998)..

5. J. T. Cheng et al., "Wave motion on the surface of the human tympanic membrane: holographic measurement and modeling analysis," *J. Acoust. Soc. Am.* **133**(2), 918 (2013).

6. J. J. Rosowski, S. Stenfelt, and D. Lilly, "An overview of wideband immittance measurements techniques and terminology: you say absorbance, I say reflectance," *Ear Hear.* **34**, 9s–16s (2013).

7. X. Zhang and R. Z. Gan, "Dynamic properties of human tympanic membrane—experimental measurement and modeling analysis," *Int. J. Exp. Comput. Biomech.* **1**(3), 252–268 (2010).

8. S. Puria, "Measurements of human middle ear forward and reverse acoustics: implications for otoacoustic emissions," *J. Acoust. Soc. Am.* **113**(5), 2773–2789 (2003).

9. W. Decraemer, S. M. Khanna, and W. R. J. Funnell, "Vibrations at a fine grid of points on the cat tympanic membrane measured with a heterodyne interferometer," presented at *EOS/SPIE Int. Symp. on Industrial Lasers and Inspection, Conf. on Biomedical Laser and Metrology and Applications*, EOS/SPIE, Munchen, Germany (June 1999).

10. J. P. Wilson and J. R. Johnstone, "Basilar-membrane and middle-ear vibration in guinea pig measured by capacitive probe," *J. Acoust. Soc. Am.* **57**(3), 705–723 (1975).

11. G. Pedrini, W. Osten, and M. E. Gusev, "High-speed digital holographic interferometry for vibration measurement," *Appl. Opt.* **45**(15), 3456–3462 (2006).

12. M. Novak et al., "Analysis of a micropolarizer array-based simultaneous phase-shifting interferometer," *Appl. Opt.* **44**(32), 6861–6868 (2005).

13. T. Y. Chen and C. H. Chen, "An instantaneous phase shifting ESPI system for dynamic deformation measurement," in *Proc. of the Society for Experimental Mechanics*, Vol. 5, pp. 279–283, Springer, New York (2011).

14. H. O. Saldner, N. E. Molin, and K. A. Stetson, "Fourier-transform evaluation of phase data in spatially phase-biased TV holograms," *Appl. Opt.* **35**(2), 332–336 (1996).

15. D. R. Schmitt and R. W. Hunt, "Optimization of fringe pattern calculation with direct correlations in speckle interferometry," *Appl. Opt.* **36**(34), 8848–8857 (1997).

16. P. J. Georgas and G. S. Schajer, "Modulo-2pi phase determination from individual ESPI images," *Opt. Laser Eng.* **50**(8), 1030–1035 (2012).

17. I. Dobrev et al., "Implementation and evaluation of single frame recording techniques for holographic measurements of the tympanic membrane in-vivo," in *Proc. of the Society for Experimental Mechanics*, Vol. 3, pp. 85–95, Springer International Publishing (2014).

18. Y. Suzuki and H. Takeshima, "Equal-loudness-level contours for pure tones," *J. Acoust. Soc. Am.* **116**(2), 918–933 (2004).

19. R. Crochiere, S. Webber, and J. Flanagan, "Digital coding of speech in sub-bands," *Bell Syst. Tech. J.* **55**(8), 1069–1085 (1976).

20. J. T. Cheng et al., "Motion of the surface of the human tympanic membrane measured with stroboscopic holography," *Hear. Res.* **263**(1), 66–77 (2010).

21. K. N. O'Connor and S. Puria, "Middle-ear circuit model parameters based on a population of human ears," *J. Acoust. Soc. Am.* **123**(1), 197 (2008).

22. G. Pedrini et al., "Transient vibration measurements using multi-pulse digital holography," *Opt. Laser Technol.* **29**(8), 505–511 (1998).

23. D. T. Kemp, "Stimulated acoustic emissions from within the human auditory system," *J. Acoust. Soc. Am.* **64**(5), 1386 (1978).

24. R. Jones and C. Wykes, *Holographic and Speckle Interferometry: A Discussion of the Theory, Practice and Application of the Techniques*, Appendix E, Cambridge University Press, Cambridge, UK (1983).

25. I. Dobrev et al., "High-speed digital holographic methods to characterize the transient acousto-mechanical response of human TM," presented at *21st DYMAT Technical Meeting, High-Speed Imaging for Dynamic*

Testing of Materials and Structures, London, UK, Institute of Physics (18–20 November 2013).

26. M. Khaleghi et al., "Long term effects of cyclic environmental conditions on painting in museum exhibition by laser shearography," in *Proc. of the Society for Experimental Mechanics*, Vol. 3, pp. 283–288 (2014).

27. Physik Instrumente, "Application notes," 15 January 2014, http://www.pi-usa.us/pdf/Piezo-Actuators_Ceramics-www.pdf (15 January 2014).

28. E. Harrington et al., "Automatic acquisition and processing of large sets of holographic measurements in medical research," in *Proc. of the Society for Experimental Mechanics*, Vol. 5, pp. 219–228 (2011).

29. N. D. Bapat, "Development of sound presentation system (SPS) for characterization of sound induced displacements in tympanic membranes," MS Thesis, Worcester Polytechnic Institute, Worcester, MA (2011).

30. S. Van der Jeught et al., "Full-field thickness distribution of human tympanic membrane obtained with optical coherence tomography," *JARO* **14**(4), 483–494 (2013).

Ivo Dobrev received his PhD degree in mechanical engineering from Worcester Polytechnic Institute in 2014. He is currently a postdoctoral researcher at the Universitätsspital Zürich working on bone conduction and middle-ear mechanics. His work includes high-speed optical metrology in field environment and *in-vivo*, robotics and mechatronics, system design and integration.

Cosme Furlong is an associate professor of mechanical engineering, working in mechanics, optical metrology, and nanoengineering, science and technology.

Jeffrey T. Cheng is an instructor in otology and laryngology at Harvard Medical School, working on middle-ear mechanics and tissue biomechanics through measurements and modeling.

John J. Rosowski is a professor of otology and laryngology, and also of health sciences and technology, working in acoustics and mechanics of the external, middle, and inner ear, as well as the comparative middle and external ear structure and function.