Vertical normal modes of human ears: Individual variation and frequency estimation from pinna anthropometry

Parham Mokhtari,[a] Hironori Takemoto, Ryouichi Nishimura, and Hiroaki Kato
National Institute of Information and Communications Technology, 3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0289, Japan

(Received 11 November 2015; revised 7 April 2016; accepted 11 July 2016; published online 5 August 2016)

Beyond the first peak of head-related transfer functions or pinna-related transfer functions (PRTFs) human pinnae are known to have two normal modes with "vertical" resonance patterns, involving two or three pressure anti-nodes in cavum, cymba, and fossa. However, little is known about individual variations in these modes, and there is no established model for estimating their center-frequencies from anthropometry. Here, with geometries of 38 pinnae measured, PRTFs were calculated and vertical modes visualized by numerical simulation. Most pinnae were found to have both Cavum-Fossa and Cavum-Cymba modes, with opposite-phase anti-nodes in cavum and either fossa or cymba, respectively. Nevertheless in both modes, fossa involvement varied substantially across pinnae, dependent on scaphoid fossa depth and cymba shallowness. Linear regression models were evaluated in mode frequency estimation, with 3322 measures derived from 31 pinna landmarks. The Cavum-Fossa normal mode frequency was best estimated [correlation coefficient $r = 0.89$, mean absolute error (MAE) = 257 Hz or 4.4%] by the distance from canal entrance to helix rim, and cymba horizontal depth. The Cavum-Cymba normal mode frequency was best estimated ($r = 0.92$, MAE = 247 Hz or 3.2%) by the sagittal-plane distance from concha floor to cymba anterior wall, and cavum horizontal depth. © *2016 Acoustical Society of America*.
[http://dx.doi.org/10.1121/1.4960481]

[BLM] Pages: 814–831

I. INTRODUCTION

It is well established that humans localize sound sources along the left-right dimension parallel to the interaural axis using binaural acoustic cues, whereas sources along so-called cones of confusion characterized by nearly constant interaural differences (including the entire median plane) are localized using monaural cues imparted mainly by the external ear or pinna (e.g., Morimoto and Aokata, 1984; Blauert, 1997). Each pinna generates its own, individual set of spectral peaks and notches in head-related transfer functions (HRTFs) above about 3 kHz, where the wavelength of sound becomes comparable to pinna dimensions. As the peaks and notches depend on individual pinna geometry, externalization and accurate localization of sounds presented in virtual auditory displays rely on using a listener's own HRTFs (Wenzel *et al.*, 1993). Furthermore, essentially the same set of peaks and notches are found in pinna-related transfer functions (PRTFs) where an individual pinna, baffled by a plane or an adjoining patch from the side of the head, is considered in isolation (e.g., Takemoto *et al.*, 2012, Fig. 6; Mokhtari *et al.*, 2010b, Fig. 3).

It is therefore of practical interest to find efficient ways to determine individualized HRTFs or PRTFs, as previously attempted for example, by estimating the following acoustical parameters from anthropometry: the optimal frequency-scaling of transfer functions (Middlebrooks, 1999), the parameters of a simplified structural model of transfer functions (Satarzadeh *et al.*, 2007), the weights on transfer-function principal components (Jin *et al.*, 2000; Nishino *et al.*, 2007), or more directly the center-frequencies and amplitudes of peaks and notches (Iida *et al.*, 2014; Mokhtari *et al.*, 2015; Spagnol and Avanzini, 2015). Indeed, as a step towards individualized HRTFs we recently reported on estimation from pinna anthropometry, of the frequency and amplitude of the first peak (Mokhtari *et al.*, 2015) which is associated with the first normal mode, having characteristics similar to an acoustic monopole. However, there is still no established method for estimating any of the upper peaks that are associated with the second or higher normal modes.

Our received knowledge of pinna normal modes is based almost entirely on physical measurements (as summarized by Shaw, 1997) and computer simulations (Kahana and Nelson, 2006) of sound pressure distributions around the pinna. Thanks to these seminal studies, it is known that above the first normal mode which is a "concha depth" resonance, typically there are two "vertical" resonances: (i) the second normal mode involves two pressure anti-nodes of opposite phase, one located in the cavum concha, the other in the fossa or in an area partly overlapping the cymba and fossa; and (ii) the third normal mode involves three anti-nodes, located in the cavum concha, cymba, and fossa, with the cymba anti-node having opposite phase relative to the other two (anatomical nomenclature is indicated schematically in Fig. 1). However, despite this basic knowledge there is still no established model for estimating the relevant (presumably the second and third) peaks of HRTFs or PRTFs from pinna anthropometry. Moreover, as Shaw (1997, Fig. 7) presented

[a]Electronic mail: parham@nict.go.jp

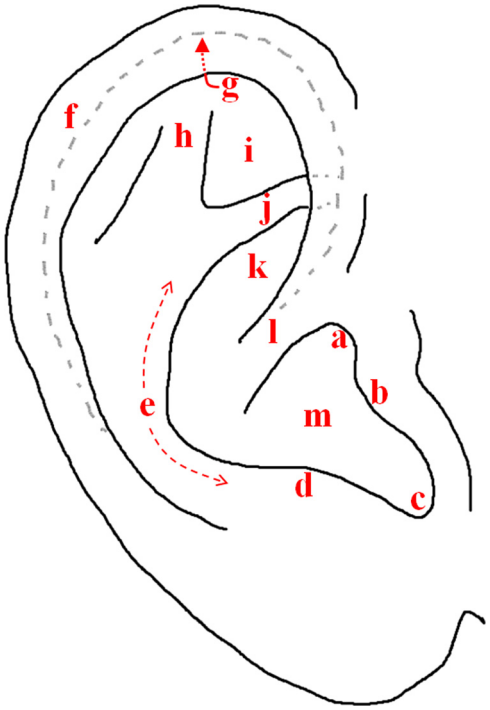

FIG. 1. (Color online) Anatomical nomenclature illustrated on a schematic of a right pinna. Light dashed lines indicate surfaces hidden behind helix. Key to annotation: (a) supratragic (or anterior) notch, (b) tragus, (c) intertragic notch, (d) antitragus, (e) antihelix, (f) helix, (g) scaphoid fossa, which is under the helix rim and may extend across posterior, superior (indicated by arrow), and anterior parts, (h) upper crus of antihelix, (i) triangular fossa, (j) lower crus of antihelix, (k) cymba, (l) crus of helix, (m) cavum concha (or just cavum). The ear-canal entrance is in the anterior part of (m), hidden behind (b). The "entire concha" refers to (k), (l), and (m) combined. The "concha aperture" refers to the outline of the entire concha along its rim, joining (a), (b), (c), (d), (e), and (j). The "fossa" refers to the upper pinna cavities (g) and (i).

only one, averaged set of normal modes (the mean of measurements on ten human pinnae), and Kahana and Nelson (2006, Figs. 18–21) presented normal modes of four manikins' pinnae and only one human pinna, currently little is known about idiosyncrasies and individual variations in pinna normal modes.

In this regard, results presented by Kahana and Nelson (2006) already indicated the possibility of individual deviations from averaged patterns: contrary to their four manikins' resonance patterns, their human pinna's second mode (first vertical mode) appeared to involve pressure anti-nodes in cavum and *cymba*, rather than fossa (Kahana and Nelson, 2006, Fig. 19); also, their human pinna was missing the expected third mode (second vertical mode) that typically involves cavum, cymba, and fossa. In normal human pinnae, the fossa (including triangular and scaphoid fossae) is generally a shallower and less enclosed cavity than the cavum concha or even the cymba; it may therefore be expected that the fossa's involvement in supporting a vertical mode would highly depend on its individual geometry, and thus would vary from pinna to pinna. However, from previous studies it is not clear whether any pinna cavity's involvement is categorical (i.e., either fully, or not at all, supporting an anti-node) or, perhaps more likely, quantifiable on a continuous scale.

Either way, the limited evidence in the literature suggests that the upper (opposite-phase) anti-node of the second normal mode can be supported by either fossa or cymba, depending on individual pinna geometry. Furthermore as described above, in either type of vertical mode it may be hypothesized that the relative strength of the fossa anti-node would vary, depending on pinna geometry. If such individual variations are confirmed across a population of human pinnae, then for the purpose of deriving models for estimation of transfer function peaks from pinna anthropometry it may be crucial to identify and label the peaks not simply in terms of their order of appearance along the frequency axis, nor in terms of the total number of pressure anti-nodes of their associated resonance pattern, but instead in terms of the anti-nodes' anatomical locations (i.e., either a Cavum-Fossa or a Cavum-Cymba, vertical normal mode). For this purpose, it would be essential to properly visualize the resonance pattern associated with each normal mode or transfer function peak being considered, in order to guarantee correct labeling and thus facilitate reasonable accuracy in estimating mode frequencies from pinna measurements.

Here we aim to quantify individual variations in the physical mechanisms of the so-called vertical (nominally the second and third) normal modes, and to derive accurate and practical models for estimation of the PRTF peak frequencies associated with those modes. For this purpose, head and pinna geometries of 19 adults were measured by magnetic resonance imaging (MRI), and computer simulation with the finite-difference time domain (FDTD) method was used to calculate farfield PRTFs and pinna normal mode patterns. Crucially, pressure and velocity anti-nodes in and around the pinna cavities were visualized in order to correctly identify the vertical normal modes and to investigate their variations across individual pinnae. Linear regression was then used to investigate and rank-order the accuracy of pinna anthropometry (among 3322 unique measures derived from 31 anatomical landmarks on each pinna) in estimating the center-frequency of the peak associated with each of the two main types of vertical normal mode: frequency $F_{\text{cav-fos}}$ of the Cavum-Fossa normal mode, and frequency $F_{\text{cav-cym}}$ of the Cavum-Cymba normal mode.

II. MATERIALS AND METHODS

A. Head and pinna geometries

The head and pinna geometries used in this study were identical to the data described in Mokhtari *et al.* (2015). Briefly, head geometries of 19 adults (5 women and 14 men) were measured by MRI at a spatial resolution of either 1.0 mm (11 participants) or 1.2 mm (8 participants), and each head volume was processed in several stages with in-house software to obtain a contiguous volume segmented from the surrounding air, with ear canals occluded at their entrance (blocked meatus). For acoustic simulations, the head volumes were downsampled to isovoxel resolution 2.0 mm, to ease computation load and reduce run-time while ensuring sufficient accuracy up to 14 kHz (Mokhtari *et al.*, 2010b, 2011). As the HRTF peaks above about 3 kHz are known to be generated by the pinna (e.g., Takemoto *et al.*, 2012,

Fig. 6; Mokhtari *et al.*, 2010b, Fig. 3), two smaller volumes each containing only the left or the right pinna and an adjoining patch from the side of the head were extracted from each head volume, thus facilitating efficient computation of PRTFs (Sec. II C) and pinna normal mode patterns (Sec. II E). As in Mokhtari *et al.* (2015), hereafter each pinna is referred to by a three-letter code, e.g., "fBr" refers to the right pinna of female B.

B. Pinna anthropometry

To investigate the relationship between mode frequency and pinna geometry, it is necessary to robustly define and locate a consistent set of anatomical landmarks on the pinna surface. As the vertical modes are known to involve the cavum concha, cymba, and fossa (which includes the triangular fossa and scaphoid fossa), the set of landmarks must adequately delineate at least the main dimensions of, and distances between, these pinna cavities; thus, for example, each cavity could be represented by at least one landmark at its base, and by two or more landmarks around its rim. However, each landmark must also be anatomically well-defined to form a set of corresponding points across a range of pinnae. As listed in Table I and exemplified in Fig. 2, our chosen set comprises 31 landmarks, 13 of which were implicitly defined in the documentation accompanying the well-known CIPIC database (Center for Image Processing and Integrated Computing; Algazi *et al.*, 2001), and 15 of which were used in our earlier work (Mokhtari *et al.*, 2015). All landmarks were located manually, on a three-dimensional (3D) surface rendered using volume data at the original resolution (1.0 or 1.2 mm) to mimic as closely as possible the task of identifying landmarks and making measurements on real human pinnae.

Nine measurements were made according to the CIPIC guidelines, specifically with reference to the pinna schematic in Fig. 3 of Algazi *et al.* (2001). First, while viewing the pinna approximately normal to the concha aperture, a reference straight line was drawn connecting the intertragic notch (landmark LC2) and the furthest point along the inner rim of the helix (LC5); then, landmarks LC3 and LC4 were determined by locating the points of intersection on the crus of helix and antihelix, respectively, with a plane passing through the reference line and parallel with the line of sight (i.e., approximately normal to the concha aperture). The reference line was thereby divided into three contiguous segments, the lengths of which provided measurements of cavum height, cymba height, and fossa height. Cavum width was measured as the distance between the supratragic notch (LC6) and the posterior-inferior rim of the antihelix (LC7).

TABLE I. List of 31 pinna landmarks and their anatomical definitions. Preceding "C" denotes a CIPIC-defined landmark. In the text, all landmark labels are prefixed with "L," thus L1–L18 and LC1–LC13.

Pinna landmark	Anatomical description
1	Center of ear-canal at its (blocked) entrance
2	Posterior-most part of cavum concha base (posterior to ear-canal entrance and inferior to crus of helix)
3	Anterior-most point on cavum concha wall
4	Inferior-most point on cavum concha floor
5	Posterior-most point on concha wall
6	Least lateral point along crus of helix (thus its base)
7	Least lateral point in base of cymba
8	Side of cheek close to supratragic (or anterior) notch
9	Most lateral surface of antitragus
10	Rim of antihelix close to posterior corner of triangular fossa (L15)
11	Superior rim of cymba
12	Anterior-most point on cymba wall
13	Least lateral point along lower crus of antihelix
14	Least lateral point in base of triangular fossa
15	Posterior corner of triangular fossa
16	Superior corner of triangular fossa
17	Superior-most point along lower rim of helix
18	Superior-most point in scaphoid fossa (thus its base)
C1	Deepest part of cavum concha when viewed normal to the line joining C12 and C13
C2	Rim of intertragic notch
C3	Intersection of crus helix with line joining C2 and C5, when viewed normal to concha aperture
C4	Intersection of antihelix rim with line joining C2 and C5, when viewed normal to concha aperture
C5	Point on lower rim of helix furthest from intertragic notch (C2)
C6	Anterior-most point along rim of supratragic (or anterior) notch
C7	Most posterior-inferior corner along rim of antihelix
C8	Point on ear-lobe, lowest along pinna's major axis
C9	Point on helix furthest from the C8–C10 axis
C10	Point on helix, highest along pinna's major axis
C11	Point on inner rim of helix, furthest from the C8–C10 axis
C12	Most lateral surface of tragus
C13	Most lateral surface of antihelix

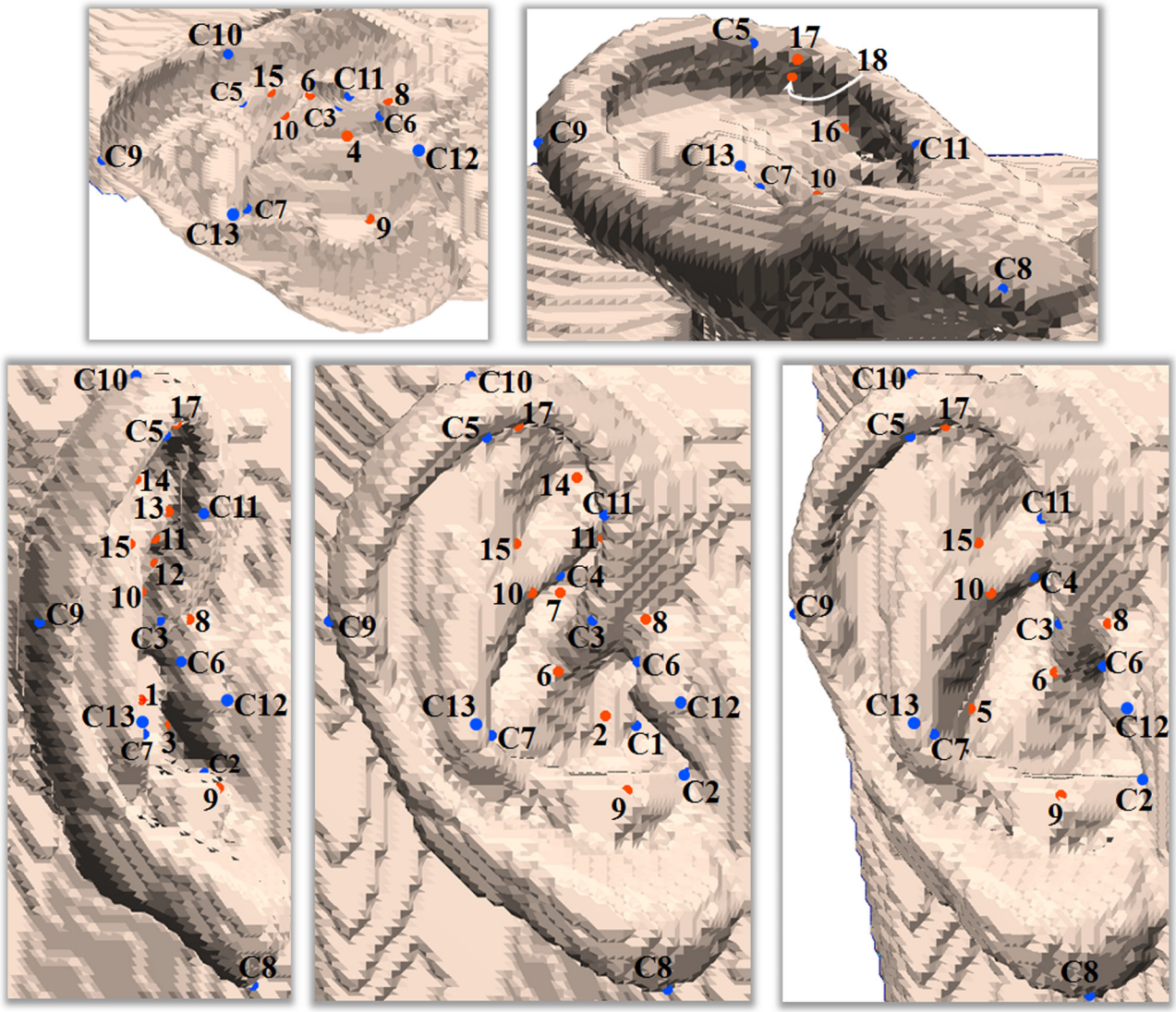

FIG. 2. (Color online) Pinna landmarks used in this study, shown here in five views of a right pinna (upper two panels: views from above and below; lower three panels: views from behind, the side, and a more frontal side-view). Preceding "C" denotes a CIPIC-defined landmark. In the text, all landmark labels are prefixed with "L," thus L1–L18 and LC1–LC13. Anatomical definitions are given in Table I and in reference to Fig. 1.

Pinna height was measured as the length of a straight line connecting the pinna bottom (LC8) and pinna top (LC10), which also defined the pinna's overall major axis. Pinna width was then measured as the distance between two points, that were calculated by projecting landmarks LC9 (pinna back) and LC11 (pinna front) onto a common plane perpendicular to the major axis. CIPIC-defined concha depth was measured (cf. Mokhtari *et al.*, 2015) as the distance from a line joining the most lateral points on the tragus (LC12) and antihelix (LC13), to the deepest point LC1 in the cavum concha. Pinna rotation angle was calculated as the backward tilt relative to the vertical, of the pinna's major axis; and pinna flare angle as the laterally outward tilt of the line joining LC12 and LC13, relative to a parasagittal plane (rather than the more ambiguous, planar approximation to the curved side of the head as implied in the CIPIC documents). In this way, nine of the ten CIPIC-defined pinna measurements (seven distances and two angles) were included in the present study—all but the intertragal incisure width, which we found difficult to define robustly in most cases.

In addition to the CIPIC-defined measurements described above which derived from 13 landmarks, a further 18 landmarks were defined (see Table I and Fig. 2) in order to more comprehensively quantify the sizes of, and distances between, the cavum [L1: canal entrance, L2: base as defined in Mokhtari *et al.* (2015), L3: anterior wall, L4: floor], cymba (L7: base, L10: superior-posterior rim, L11: superior rim, L12: anterior wall), triangular fossa (L13: anterior corner at base of crus antihelix, L14: base, L15: posterior corner, L16: superior corner), scaphoid fossa (L17: superior rim, L18: superior base), and other structures (L5: concha posterior wall, L6: crus helix base, L8: cheek, L9: antitragus). This enabled a larger set of measurements, by calculating the distance between every pair of all 31 landmarks (the number of unique pairs being ${}^{31}C_2 = 465$). Moreover, to investigate a wider range of possible anthropometric relations, the distance between each pair of landmarks was calculated in seven distinct ways: (i–iii) along each of the three, global coordinate axes separately, i.e., 1D distances along only the front-back, left-right, or up-down (vertical) axis;

(iv–vi) in each of the three planes defined by unique pairs of coordinate axes, i.e., 2D distances calculated using only two coordinates at a time, thus in a horizontal (or transverse) plane, a coronal (or frontal) plane, or a sagittal plane (the latter yielding the same distance as would be measured on a side-view photograph of the pinna); and (vii) using all three coordinates, i.e., the full 3D distance. In total, we thus measured $465 \times 7 = 3255$ unique distances among the set of 31 landmarks for each pinna, only two of which (pinna height and concha width) were already calculated among the nine CIPIC measurements.

The single-distances described so far measure only along a straight line between pairs of landmarks. In contrast, as shown in Sec. III, acoustic oscillations between opposite-phase anti-nodes usually traverse a curved path in the space between the pinna cavities. Therefore, with the intention of better approximating such curved paths, 60 supplementary measurements each involving three landmarks were also made: 18 distances hypothetically related with the normal mode involving anti-nodes in the cavum and fossa (the sum of two distances, involving three landmarks taken from all combinations of {L1 or L3}, {L6 or LC3 or LC6}, and {L17 or LC5 or L18}), 12 distances hypothetically related with the normal mode involving anti-nodes in the cavum and cymba (the sum of two distances, involving three landmarks taken from all combinations of {L1 or L3}, {L6 or LC3 or LC6}, and {L7 or L12}), and all these 30 distances calculated in a sagittal plane, as if on a side-view photograph of the pinna (i.e., disregarding lateral coordinates). Although the weighted sum of two or more distances would be evaluated using multiple linear regression models with the forward selection method (Sec. II F), these particular distances were explicitly evaluated because of their possibly direct relations with vertical mode frequencies, and because the forward selection method is not exhaustive.

Thus, a grand total of 3322 anthropometric measurements (3320 distances and two angles) were considered: 465 single-distances $\times$ 7 types of measurement (including 2 of the CIPIC-defined measures), plus 7 CIPIC-defined measures not already included among the single-distances, plus 60 supplementary distances involving three landmarks each.

C. Acoustic simulation of farfield PRTFs

With the head volumes downsampled to resolution 2.0 mm and each pinna and an adjoining patch from the side of the head represented in isolation, farfield pressure responses to a broadband Gaussian pulse (placed directly adjacent to the center of the blocked ear-canal entrance) were calculated by FDTD acoustic simulation (Takemoto *et al.*, 2010; Mokhtari *et al.*, 2011, 2015). In this way, responses at all farfield locations were calculated in a single run thanks to acoustic reciprocity (Pierce, 1989; Zotkin *et al.*, 2006). Artifactual reflections from the computation domain boundaries were minimized by an optimal perfectly matched layer (Mokhtari *et al.*, 2010a). With the computation domain restricted to about 1.4 cm from the pinna on five sides (the head-patch being partially immersed within the perfectly matched layer on the medial side), Kirchoff-Helmholtz integration was used to calculate farfield responses (Mokhtari *et al.*, 2008). In this way, acoustic pressure waveforms of duration 5 ms were calculated a distance 1 m from the head center, at 1250 spatial locations in accord with the CIPIC data (Algazi *et al.*, 2001): in a head-centered interaural polar coordinate system, at all combinations of 25 azimuth (or lateral) angles {−80°, −65°, −55°, −45°,…,5° steps,…, +45°, +55°, +65°, +80°} measured with respect to the median plane from left to right, and 50 elevation (or polar) angles {−45°,…,5.625° steps,…, +230.625°} measured with respect to the front horizontal plane from below-front to below-back along each cone of confusion. Free-field responses were also calculated at all 1250 spatial locations, with the source at the same position and the pinna absent. PRTFs were then obtained by Fourier transformation of the pressure waveforms followed by free-field normalization; the PRTF magnitude therefore indicates the acoustic pressure gain over free-field conditions, due to the presence of the pinna and adjoining head-patch.

These simulation methods were previously evaluated against acoustic measurements and shown to provide a good spectral match up to 14 kHz for HRTFs of a manikin (Mokhtari *et al.*, 2011).

D. Extraction of normal modes by PRTF peaks histogram

To identify each pinna's vertical mode(s), as detailed below we first identified all candidate modes by extraction of PRTF magnitude peaks up to 14 kHz, and then selected the vertical subset by visualization of pinna normal-mode patterns (Sec. II E). Candidate normal-modes for each pinna were identified in the following way. First, all spectral peaks within 2–14 kHz were extracted by simple peak-picking of the PRTF magnitude at every location. Then, a histogram was prepared (bin size 50 Hz) of all the peak center-frequencies extracted at all 1250 locations (cf. Fig. 3); the histogram was smoothed by mean filtering (window size five bins) to reduce noise and thus avoid detection of spurious histogram-peaks. Each peak of the smoothed histogram, including the surrounding bins as far as the local minimum on either side, was taken to represent a single normal-mode, i.e., a group of PRTF peaks across a range of (farfield) spatial locations that shared a similar center-frequency and therefore essentially the same resonance mechanism, assuming no overlapping or degenerate modes. Among all the PRTF peaks that were grouped in each mode, a representative peak was then selected as the one with the highest amplitude. While a peak in the histogram is better represented statistically by the median value of its underlying distribution, instead the PRTF peak with the highest amplitude was selected in order to find the center-frequency and source location corresponding to the strongest (most energetic) and therefore the most efficient and typical manifestation of each normal mode.

An example of a histogram and the PRTFs corresponding to the representative, second, and third modes is shown in Fig. 3. This example also shows evidence that the peaks extracted from a transfer function at any one spatial location do not necessarily include all the modes: in the two transfer

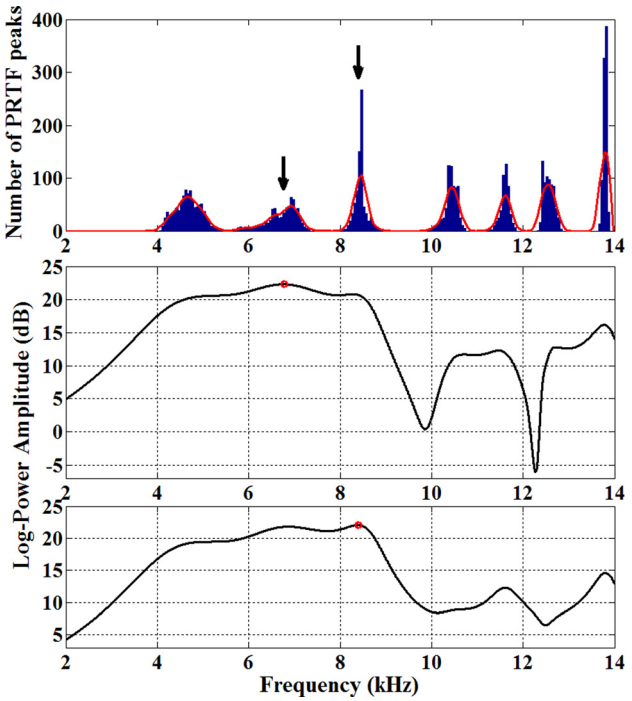

FIG. 3. (Color online) An example of a histogram of one right pinna's PRTF peaks extracted at all 1250 farfield locations. Bar-graph: raw histogram with 50 Hz bins. Superimposed solid line: histogram smoothed with a sliding, five-bin mean filter. Arrows mark the second (6769 Hz) and third (8400 Hz) mode frequencies, each selected as the PRTF peak with the highest amplitude underlying the corresponding peak of the smoothed histogram. The two lower panels show the PRTF (solid line) and its peak (small circle) corresponding to the selected second (azimuth +65°, elevation +51°) and third (azimuth +55°, elevation +73°) mode, respectively.

functions shown, the first mode (with median value at about 4.7 kHz) is visually apparent as an inflection but not as a distinct peak; and in the PRTF with the highest third-mode amplitude (lower panel), the fourth and sixth modes (at about 10.4 kHz and 12.6 kHz) also do not appear as clear peaks. In other words, owing to individual radiation patterns and associated phenomena such as acoustic cancellation of a peak in the vicinity of a spectral notch, each mode usually does not appear as a PRTF peak at every spatial location. Therefore, compared with extraction of peaks on single or averaged transfer functions, the histogram method promises more robust identification of all normal modes of a pinna.

The necessity of the histogram method can be further appreciated by the fact that, as shown by the spread of frequency values around each histogram-peak in Fig. 3, different farfield source locations generate slightly different (but essentially the same) resonance modes at the pinna. Such small-scale variations in mode pattern and corresponding mode frequency are an inevitable consequence of the relatively open structure of the human pinna and its high degree of acoustic coupling with the surrounding air. However, it would be impractical to estimate from pinna geometry, all possible resonance frequencies of each mode at every possible farfield location. Instead, here we focus on estimating typical resonance frequencies associated with typical normal modes, as identified here by selection of one representative (strongest) PRTF peak underlying each histogram-peak.

E. Visualization of pinna normal-mode patterns

The preceding methods yielded for each pinna, a set of candidate mode center-frequencies paired with their spatial locations, at which the farfield responses were found to be the strongest. Each normal-mode pattern was then visualized by calculating acoustic pressure and velocity near the pinna during steady-state resonance, in response to a sinusoidal source at the corresponding frequency and farfield spatial location. This was achieved without having to extend the FDTD grid to 1 m, by the total-field/scattered-field (TF/SF) method which is well known in electromagnetics FDTD simulations (Taflove and Hagness, 2005, Chap. 5) and which was adapted to acoustic FDTD by Xiao and Liu (2003, Sec. II B). Briefly, within the computation domain the pinna was surrounded by a so-called virtual interface, at every point on which acoustic pressure and velocity were injected by free-field illumination from the specified source in the farfield (i.e., by appropriately amplitude-scaling and time-delaying the source waveform). The FDTD grid interior to the virtual interface therefore supported the total (incoming plus scattered) field, while the grid exterior to the interface (including the perfectly matched layer) saw only the scattered field. Thus, the TF/SF method provided a computationally efficient way to illuminate the scattering object of interest (here one pinna and adjoining head-patch) by an acoustic source at any position and distance outside the FDTD computation domain.

In particular, the external source was a sinusoid at a candidate mode frequency and total duration 3 ms, including an initial 0.75 ms fade-in by a raised-cosine window. In this way, normal-mode oscillations at the pinna were allowed to reach a steady-state, and the pressure and velocity in the interior (total) sound field during only the last five periods were stored for further analysis. As summarized below, normal modes were then visualized using a variant of the methods described in Mokhtari *et al.* (2014).

In this study, we visualized normal modes by using color to depict the magnitude and phase of pressure anti-nodes on the pinna surface, and line segments to depict the principal vectors of velocity anti-nodes in the intervening space. Pressure anti-nodes were visualized by computing potential energy density (PED) which is proportional to the square of pressure, then averaging over the five steady-state periods to obtain time-averaged $\langle \mathrm{PED} \rangle$. Crucially, relative phase information lost in the squaring and time-averaging process at every point in the sound field was reinstated, by computing the inverse cosine of the coefficient of correlation between the pressure waveform at that point and the pressure waveform at the point of $\langle \mathrm{PED} \rangle_{\max}$ in the cavum concha anti-node (two finite-length sinusoidal signals with matching frequency will have a correlation coefficient equal to $+1$, 0, or -1 when they are perfectly in-phase, out-of-phase, or opposite-phase, respectively). In-phase (phase difference $<45°$) and opposite-phase (phase difference $>135°$) anti-nodes were then depicted by warm and cool colors, respectively, while color saturation depicted the magnitude of $\langle \mathrm{PED} \rangle$.

Velocity anti-nodes were visualized by computing at each point in the sound field, the first principal component

(or eigenvector) of the time-course of the velocity vector across the five steady-state periods. As acoustic resonance is associated with reciprocal movements of air, velocity vectors at anti-nodes tend to oscillate back and forth along mainly one (principal) axis, which can therefore be found by principal component analysis (PCA); moreover, the magnitude of the first eigenvalue is generally larger at a velocity anti-node compared with other points in the sound field. The eigenvector fields associated with velocity anti-nodes were therefore visualized by drawing a line-segment centered at each grid point and oriented in the direction of the principal component, with the length of the line being proportional to σ_1 (the standard-deviation along the first principal component, equal to the square-root of the first eigenvalue). To reduce clutter and highlight mainly velocity *anti-nodes*, only vectors with σ_1 larger than a threshold (e.g., 25% of the largest value in that entire sound field) were depicted; furthermore, a line-segment was drawn only if the velocity-vector oscillations were sufficiently uniaxial (i.e., if σ_1 accounted for more than 95% of the total variance at that point). The relative phase at every point of the velocity field was computed as the inverse cosine of the absolute value of the coefficient of correlation between the principal velocity waveform at that point (i.e., the time-course of the velocity vector projected onto its principal component) and the largest pressure waveform in the cavum concha (i.e., the same reference waveform used earlier in calculating the phase of pressure anti-nodes). As velocity anti-nodes in a resonating sound field are normally out of phase with pressure anti-nodes (phase difference about 90°), a principal velocity-vector was drawn only if its phase difference relative to the pressure anti-node in the cavum was within a certain threshold around 90° (typically within a margin ±10°–30°).

To quantify inter-individual differences, key features were extracted from every normal mode: $\langle \mathrm{PED} \rangle_{\max}$ within each pressure anti-node, and the phase of the pressure waveform at that point relative to the reference point ($\langle \mathrm{PED} \rangle_{\max}$ in cavum concha). Furthermore, the relative degree of fossa or cymba involvement (i.e., the symmetry of oscillations between cavum and either fossa or cymba) was quantified by the ratio of $\langle \mathrm{PED} \rangle_{\max}$ in the fossa or cymba to that in the cavum. For the ideal case of a closed, uniform tube with perfectly symmetric anti-nodes at both ends, this ratio would be unity, or 0 dB. In contrast, for the open structure and non-uniform geometry of the human pinna this ratio is expected to vary across individual pinnae; in particular, as the fossa is typically a shallower and less enclosed cavity compared with the cavum, the fossa-to-cavum ratio of $\langle \mathrm{PED} \rangle_{\max}$ is expected to be lower than 0 dB for most pinnae.

Vertical modes, the focus of this study, were identified by the presence of either only two, or only three, pressure anti-nodes in the main pinna cavities (i.e., in cavum, cymba, and fossa); and by the presence of a velocity anti-node in the space between adjacent pairs of opposite-phase, pressure anti-nodes. In particular, we considered all normal modes that involved either (i) opposite-phase pressure anti-nodes in cavum and fossa (henceforth referred to as a Cavum-Fossa mode), or (ii) opposite-phase pressure anti-nodes in cavum and cymba, regardless of fossa involvement (henceforth referred to as a Cavum-Cymba mode). Any mode patterns with the dominant pressure anti-node (i.e., overall $\langle \mathrm{PED} \rangle_{\max}$) lying outside the main pinna cavities (e.g., in cavities formed by the back surface of the pinna and the side of the head) were rejected from further consideration. Also rejected were any mode patterns with an anti-node whose $\langle \mathrm{PED} \rangle_{\max}$ was located outside the corresponding cavity (i.e., clearly away from the pinna surface).

F. Linear regression modeling and nearest-neighbor classification

Beyond investigating individual variations in pinna vertical modes, we aimed to find a practical solution for accurate estimation of the PRTF peaks associated with those modes, from individual pinna anthropometry. For the purpose of learning to map from the physiological to the acoustical domain, we used linear regression. In view of the emphasis on practicality, and the relatively small sample size (38 pinnae) compared with the total number of parameters (3322 anthropometric measurements), only *forward selection* of the best performing two or three parameters was considered. Regression models were evaluated and parameters were selected or recruited, on the basis of the coefficient of correlation (r) and the mean absolute error (MAE, in both Hz and percent) between original and estimated values of the dependent variable (peak center-frequency). Cross-validation (leaving one person out at a time) was used to evaluate the robustness of the regression models.

As discussed in Sec. III, the 38 pinnae were divided in two groups based on their normal mode (i.e., *acoustical*) characteristics. To quantify the groups' defining *physiological* features, the anthropometric measures were evaluated with (i) a nearest-neighbor classifier, using the leave-one-out method where each pinna was classified after training the classifier on all the remaining pinnae, and (ii) the unequal variance Student's t test, which measures the significance of the difference of the means of two distributions having in general significantly different variances.

III. RESULTS

A. Cavum-Fossa normal modes

The majority of pinnae (33 out of 38) were found to have a Cavum-Fossa normal mode, i.e., pressure anti-nodes of opposite phase in cavum and fossa. Among the five pinnae rejected from consideration for this mode, two had $\langle \mathrm{PED} \rangle_{\max}$ of the opposite-phase anti-node located well away from the pinna surface ("fEl" and "fEr"), one had the dominant anti-node located behind the pinna ("fCr"), and two had both of these conditions ("fAl" and "mMr"). Among the 33 Cavum-Fossa modes, the fossa anti-node generally covered an area including the triangular fossa and a substantial part of the scaphoid fossa and helix rim; however, the exact location of $\langle \mathrm{PED} \rangle_{\max}$ in the fossa anti-node varied from pinna to pinna, and could be broadly described in one of four categories: in the posterior scaphoid fossa (five pinnae), in the superior-posterior scaphoid fossa (ten pinnae), in the superior scaphoid fossa (ten pinnae), or in the triangular fossa

(eight pinnae). Nevertheless, the extent to which our 33 Cavum-Fossa modes resembled the second normal mode of Shaw (1997) or Kahana and Nelson (2006), and in particular, the degree to which the fossa was involved in the resonance, varied substantially.

First, as anticipated, among the 33 pinnae there was wide variation in the degree to which the fossa anti-node participated in oscillatory storage and exchange of acoustic energy with the cavum. The fossa-to-cavum ratio of $\langle \mathrm{PED} \rangle_{\max}$ ranged from −15.1 dB to −1.7 dB for 32 pinnae (fossa weaker than cavum anti-node), and was +6.0 dB for one pinna ("fAr," where a particularly deep and well-enclosed scaphoid fossa caused a doubling of the maximum pressure there compared with the cavum). Two typical examples of a clear second mode with strong fossa involvement (ratio +6.0 dB and −3.7 dB, respectively) are shown in Figs. 4(a) and 4(b). As indicated by the velocity vectors, in both cases the pressure anti-nodes exchanged acoustic energy along a curved path around the crus helix and over the cymba; also in both cases, velocity vectors projected outwards from the pressure anti-nodes (particularly from the fossa), indicating cyclic energy loss to the surrounding air. In contrast, Figs. 4(c) and 4(d) show two typical examples of weak fossa involvement (ratio −10.6 dB and −12.1 dB, respectively). In both these cases, the cavum anti-node was dominant and covered a wider part of the concha base (partly including the cymba); and velocity vectors radiated mainly from this dominant anti-node, not only up towards the fossa anti-node but also outwards to the surrounding air.

To investigate the physiological basis of a strong versus weak fossa involvement, the pinnae were divided in two groups by a threshold on the fossa-to-cavum ratio of $\langle \mathrm{PED} \rangle_{\max}$. Assigned to group 1 (Cavum-Fossa mode with strong fossa involvement) were the 18 pinnae for which $\langle \mathrm{PED} \rangle_{\max}$ in the fossa anti-node was no more than 9 dB weaker than that in the cavum [as in the examples from Figs. 4(a) and 4(b)]. Assigned to group 2 were the remaining 15 Cavum-Fossa modes with fossa involvement lower than threshold [as in the examples from Figs. 4(c) and 4(d)], and the five pinnae with no Cavum-Fossa mode. The physiological features that best characterized

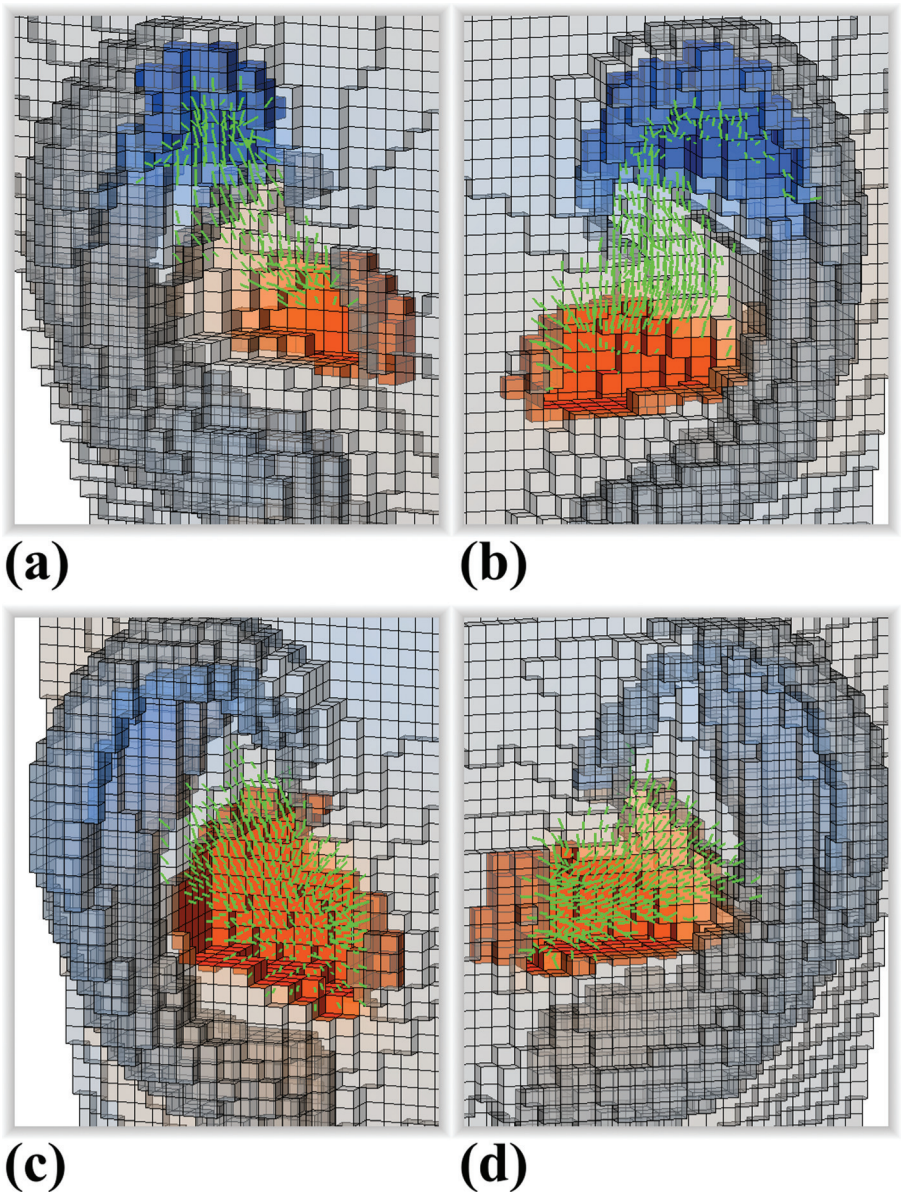

FIG. 4. Cavum-Fossa normal modes of four pinnae, showing examples of typical patterns and individual differences. Spatial resolution 0.2 cm. Warm/cool colors on pinna surface indicate in-phase/opposite-phase pressure anti-nodes. Green line-segments represent the principal velocity-vector field. (a) and (b) Two examples of a clear second mode with strong fossa involvement (pinnae "fAr" and "fBl"). (c) and (d) Two examples of a second mode with weak fossa involvement (pinnae "mHr" and "mEl"). Note on color scale: in each panel, the color saturation limit was set to 20% of the maximum value in the cavum anti-node.

the two groups were then identified with a two-category classification task, on the premise that the most characteristic features should clearly discriminate the two groups and therefore yield the highest classification accuracy.

Remarkably, of all 3322 anthropometric measurements evaluated, the highest classification accuracy (87%, or 33/38 pinnae) was attained by the difference in height between pinna landmarks L17 and L18, i.e., the *vertical depth of the scaphoid fossa* d_{scaph}. The largest improvement in classification accuracy (to 92%, or 35/38 pinnae) was then attained by recruiting the horizontal-plane distance between landmarks L7 and L13, i.e., the *horizontal depth of the cymba relative to the lower crus antihelix* d_{cym} (by itself, this measurement yielded 71% accuracy, or 27/38 pinnae). A scatterplot of all 38 pinnae on the plane defined by these best two parameters is shown in Fig. 5. Mean values of d_{scaph} (0.39 cm for group 1 and 0.15 cm for group 2) were significantly different (unequal variance Student's $t = 5.82$ at $p < 0.001$), and mean values of d_{cym} (0.61 cm for group 1 and 0.77 cm for group 2) were also significantly different ($t = 4.02$ at $p < 0.001$). These results indicate that a pinna with a sufficiently deep scaphoid fossa and shallow cymba (toward the lower right in Fig. 5) is more likely to have a clear and nearly symmetrical, Cavum-Fossa normal mode with opposite-phase pressure anti-nodes in cavum and fossa; and in contrast, this normal mode is more likely to not exist or have only weak fossa involvement (a more asymmetric resonance pattern) for a pinna with a shallow fossa and deep cymba (toward the upper left in Fig. 5). Quantitatively, the perpendicular bisector of the line joining the two group means (dotted line in Fig. 5) can be used to classify a pinna by entering these two anthropometric measurements (d_{scaph} and d_{cym}, in cm) in the expression

$$0.29 + 1.50 d_{scaph} - d_{cym}, \quad (1)$$

where a positive or a negative result would suggest group 1 or group 2 membership, respectively.

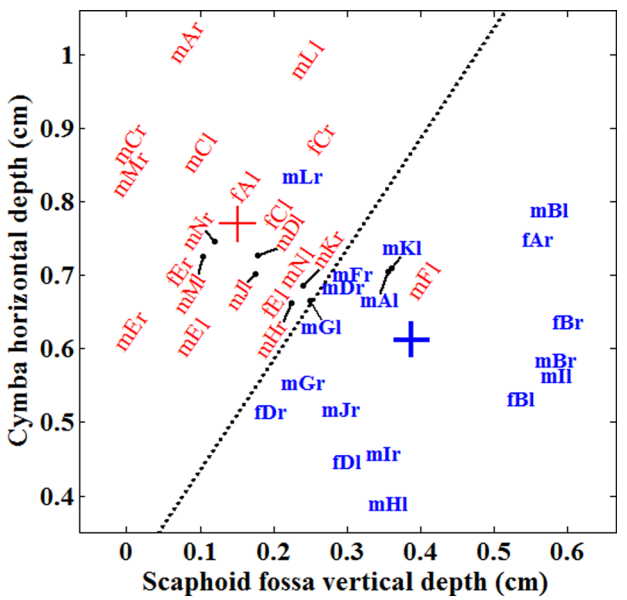

FIG. 5. (Color online) Scatterplot of all 38 pinnae on the plane defined by the two anthropometric measures that yielded the highest classification accuracy (92%) with respect to group 1 (*bold text*: Cavum-Fossa mode with strong fossa involvement) versus group 2 (*rotated text*: weak fossa involvement or no Cavum-Fossa mode). The two group means are indicated by the "+" signs. The classification boundary is depicted by the dotted line [cf. Eq. (1)], which is the perpendicular bisector of the line joining the two group means. To avoid overlapping text labels and thus improve clarity, coordinates of nine pinnae are indicated by small filled circles, and short line segments point to their respective labels.

B. Estimation of $F_{cav\text{-}fos}$ from pinna anthropometry

Irrespective of whether a pinna belonged to group 1 or 2, all 33 Cavum-Fossa normal modes were derived on the basis of a PRTF peak and shared the acoustic-physiological characteristic of opposite-phase anti-nodes in cavum and fossa. Therefore, all 33 pinnae were used to evaluate regression models for estimation of $F_{cav\text{-}fos}$ (center-frequency of the PRTF peak associated with the Cavum-Fossa mode regardless of the degree of fossa involvement) from anthropometry.

Table II lists the 10 best models, ranked by correlation coefficient r which declines gradually from 0.80 to 0.77, indicating that these models all yielded comparable performance. Also listed (in parentheses) are the results of cross-validation models, where data of one *person* at a time were left out of the training set and used for evaluation; as expected, the cross-validation method always yielded lower r and higher MAE, but remarkably the degradations were small and the list of top performers remained unchanged. Thus, the cross-validation tests confirmed the robustness of the top regression models and the acoustical significance of the top anthropometric measures; henceforth, only the models using all the available data will be discussed.

In Table II, the negative sign on every coefficient m conforms with the inverse relation between resonance wavelength and frequency (i.e., the longer the measured distance, the lower the estimated frequency), and suggests that each of these distances can be regarded as a (partial) measurement of resonance wavelength. As anticipated (in Sec. II B), the

TABLE II. The top ten pinna anthropometric measurements which yielded the best linear regression models for estimation of the Cavum-Fossa normal mode frequency $F_{cav\text{-}fos}$. The models were ranked according to the coefficient of correlation (r) between original and estimated frequencies. The mean absolute error (MAE) is also listed in both Hz and %, but these do not necessarily follow the same ranking. Model coefficients are b (intercept) and m (slope), as in $F_{cav\text{-}fos} = b + m\ d_{1\text{-}C3\text{-}17}$. Figures in parentheses for r and MAE pertain to results of cross-validation, where data of *one person at a time* were left out of the training set and used for evaluation.

	Distance landmarks	r	MAE (Hz)	MAE (%)	b (Hz)	m (Hz/cm)
1	L1 - LC3 - L17	0.80 (0.78)	320 (343)	5.6 (5.9)	12,229	−1551
2	L1 - LC3 - LC5	0.80 (0.76)	342 (369)	5.9 (6.4)	12,658	−1603
3	L1 - L17	0.79 (0.76)	332 (356)	5.8 (6.2)	12,441	−1647
4	L6 - LC5	0.78 (0.75)	340 (365)	6.0 (6.4)	11,446	−1818
5	L1 - LC5	0.78 (0.74)	356 (384)	6.2 (6.6)	12,740	−1648
6	L2 - LC5	0.78 (0.75)	355 (381)	6.1 (6.6)	12,617	−1692
7	L1 - L6 - LC5	0.78 (0.74)	350 (373)	6.2 (6.5)	12,818	−1568
8	L2 - L17	0.77 (0.74)	353 (377)	6.1 (6.5)	12,041	−1599
9	L7 - L17	0.77 (0.74)	346 (376)	6.0 (6.5)	10,122	−1628
10	L3 - LC3 - L17	0.77 (0.73)	360 (392)	6.2 (6.7)	12,014	−1389

overall best performance was obtained by one of the supplementary measurements involving the sum of two distances; in particular, the 3D distance $d_{1\text{-}C3\text{-}17}$ (in cm) from the canal entrance (L1) to the CIPIC-defined crus helix (LC3) and from there to the helix rim (L17),

$$\hat{F}_{\text{cav-fos}} = 12229 - 1551 d_{1\text{-}C3\text{-}17}, \quad (2)$$

with estimated $F_{\text{cav-fos}}$ in Hz, $r = 0.80$, and MAE = 320 Hz or 5.6%. Although the typically curved path of oscillations between cavum and fossa anti-nodes [as indicated by the stream of velocity vectors most clearly visible in Fig. 4(b)] is only crudely represented by the two straight-line distances, it is reasonable and encouraging that this approximation of the path's length should be a top performer. Nevertheless, as listed in Table II, three alternative, related measures performed comparably: the helix rim landmark (L17) could be replaced with the CIPIC-defined version (LC5), and subsequently the CIPIC-defined crus helix landmark (LC3) could be replaced with the crus helix base (L6), without much degradation in performance ($d_{1\text{-}C3\text{-}C5}$ yielded $r = 0.80$ and MAE = 342 Hz or 5.9%, and $d_{1\text{-}6\text{-}C5}$ yielded $r = 0.78$ and MAE = 350 Hz or 6.2%). Similarly, the canal entrance landmark (L1) could be replaced with the concha anterior wall (L3), with only a small degradation in performance ($d_{3\text{-}C3\text{-}17}$ yielded $r = 0.77$ and MAE = 360 Hz or 6.2%).

Remarkably, comparable accuracy was also attained by the best-performing, single distance $d_{1\text{-}17}$ (in cm) between the canal entrance (L1) and the helix rim (L17):

$$\hat{F}_{\text{cav-fos}} = 12441 - 1647 d_{1\text{-}17}, \quad (3)$$

with $r = 0.79$ and MAE = 332 Hz or 5.8%. A scatter-plot comparing measured values of $F_{\text{cav-fos}}$ with those estimated by Eq. (3) is shown in the upper panel of Fig. 6. This model is practically advantageous over the Eq. (2) model, as it attains similar estimation accuracy while requiring measurement of only one, rather than two, distances. It is indeed remarkable that such accuracy was attained by perhaps the simplest approximation of the total path length between cavum and fossa anti-nodes, i.e., a straight line from canal entrance to helix rim. Interestingly, while $d_{1\text{-}17}$ was a top performer, comparable performance was attained by five other, single-distance models (ranked 4th-6th, 8th, and 9th in Table II). Also, while not listed in Table II, comparable performance was attained by a model involving the 2D distance (measured in a sagittal plane and therefore amenable for measurement on a side-view photograph of the pinna) between cymba superior rim (L11) and helix rim (L17), with $r = 0.77$, MAE = 343 Hz or 6.0%, $b = 8292$ Hz, and $m = -1662$ Hz/cm.

Retaining either $d_{1\text{-}C3\text{-}17}$ or $d_{1\text{-}17}$, multiple linear regression models were trained and evaluated on two parameters at a time, in search of the next best, complementary measurement. In both cases, the largest improvement over the one-parameter model was obtained by recruiting the distance $d_{7\text{-}11(\text{horiz})}$ (in cm) measured in a horizontal plane, between the cymba base (L7) and the superior rim of the cymba (L11). Thus the model in Eq. (2) was best improved by

$$\hat{F}_{\text{cav-fos}} = 12461 - 1260 d_{1\text{-}C3\text{-}17} - 2194 d_{7\text{-}11(\text{horiz})}, \quad (4)$$

with $r = 0.89$ and MAE = 240 Hz or 4.1%; and the model in Eq. (3) was best improved by

$$\hat{F}_{\text{cav-fos}} = 12646 - 1333 d_{1\text{-}17} - 2239 d_{7\text{-}11(\text{horiz})}, \quad (5)$$

with $r = 0.89$ and MAE = 257 Hz or 4.4%. This complementary distance is a measure of cymba depth, similar (but not identical) to d_{cym} discussed in Sec. III A. While these two models performed nearly equally well, the model in Eq. (5) has the practical advantage of requiring measurement of only two, rather than three, distances. A scatterplot comparing measured values of $F_{\text{cav-fos}}$ with those estimated by Eq. (5) is shown in the lower panel of Fig. 6. Any further improvements over the Eq. (5) model will not be discussed, as the next best parameter was not significantly related with the dependent variable $F_{\text{cav-fos}}$.

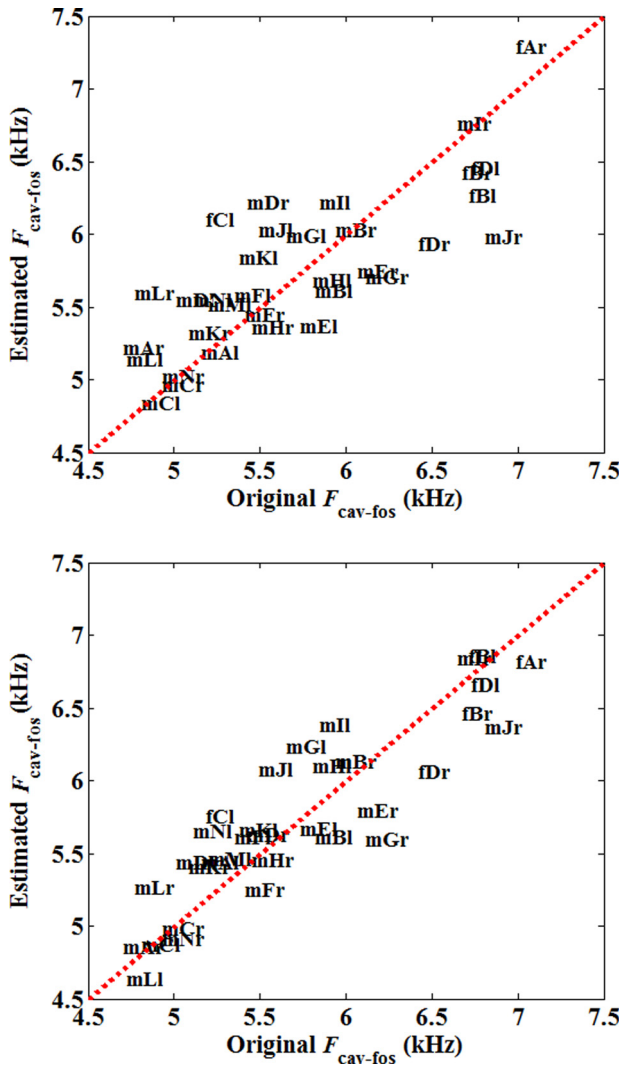

FIG. 6. (Color online) Scatter-plots of original versus estimated, normal mode frequency $F_{\text{cav-fos}}$, for the 33 pinnae that were found to have a Cavum-Fossa mode. Upper panel: linear regression model of Eq. (3) with $d_{1\text{-}17}$ ($r = 0.79$ and MAE = 332 Hz or 5.8%). Lower panel: linear regression model of Eq. (5) with $d_{1\text{-}17}$ and $d_{7\text{-}11(\text{horiz})}$ ($r = 0.89$ and MAE = 257 Hz or 4.4%). The dashed line represents the ideal performance.

C. Cavum-Cymba normal modes

Similar to the results in Sec. III A, nearly all the pinnae (37 out of 38, all except pinna "fDl" whose modes are discussed further below in Sec. III E) were found to have a Cavum-Cymba normal mode, i.e., pressure anti-nodes of opposite phase in cavum and cymba. However, the extent to which these mode patterns resembled the third normal mode of Shaw (1997) or Kahana and Nelson (2006), and in particular, the degree to which the fossa also took part in the resonance, varied substantially.

First and most typically, 29 of our Cavum-Cymba modes also had an anti-node in the fossa (in-phase with cavum, and opposite-phase with cymba), so that these patterns could be referred to as *third* normal modes. However, the degree of fossa involvement varied substantially: the fossa-to-cavum $\langle \text{PED} \rangle_{\max}$ ratio ranged from -15.3 dB to -0.8 dB (fossa always weaker than cavum anti-node). Using the same criterion described in Sec. III A to divide these 29 pinnae in two groups (i.e., setting a threshold of -9 dB on the fossa-to-cavum $\langle \text{PED} \rangle_{\max}$ ratio), resulted in 16 pinnae with a clear, third normal mode, and 13 pinnae with relatively weak fossa involvement. Two examples of clear modes with strong fossa involvement (fossa-to-cavum ratio -5.9 dB and -0.8 dB, respectively) are shown in Figs. 7(a) and 7(b). In both cases, the opposite-phase anti-node in the cymba was energetically connected to its neighboring, in-phase anti-nodes in cavum and fossa. In contrast, Figs. 7(c) and 7(d) show two examples of Cavum-Cymba normal modes with only weak fossa involvement (fossa-to-cavum ratio -10.7 dB and -10.8 dB, respectively).

As expected, a majority of the pinnae with Cavum-Cymba modes having strong fossa involvement (14 out of 16) had similarly been found in Sec. III A to have higher than threshold fossa involvement in their Cavum-Fossa mode (i.e., group 1). Also, a majority of the pinnae with Cavum-Cymba modes having weak fossa involvement (11 out of 13) had similarly been found to have lower than threshold fossa involvement in their Cavum-Fossa mode (i.e., group 2). Therefore, despite the arbitrary value of the threshold which was set only to divide the pinnae in roughly equal groups, in most cases (25 out of 29 pinnae) the relative level of fossa involvement was consistent between the two vertical modes. This general consistency is emphasized by the fact that the four individual pinnae shown in Figs. 4(a)–4(d) are the same as those in Figs. 7(a)–7(d), so that these two figures taken together show the most typical patterns across the two modes.

Second, eight of our Cavum-Cymba modes were found to have no third anti-node (i.e., no anti-node in the fossa), so that these patterns could be referred to as *second* normal modes. For these eight pinnae, the ratio of $\langle \text{PED} \rangle_{\max}$ in the cymba to that in the cavum ranged only from -4.6 dB to $+4.3$ dB, implying a relatively symmetric resonance pattern with comparable involvement of cavum and cymba. Two examples are shown in Figs. 7(e) and 7(f). In both cases, the two pressure anti-nodes of opposite phase were connected by a velocity anti-node that traversed a tightly curved path around the crus helix. As expected, a majority of these pinnae with Cavum-Cymba modes having no fossa involvement (seven out of eight) had earlier in Sec. III A been classified in group 2: four had no Cavum-Fossa mode ("fAl," "fCr," "fEl," and "fEr"), and three had a Cavum-Fossa mode with weaker than threshold fossa involvement ("fCl," "mAr," and "mFl"). However, even for the single exception that had earlier been classified in group 1 (pinna "mAl"), the fossa-to-cavum $\langle \text{PED} \rangle_{\max}$ ratio was -8.3 dB, i.e., just above the arbitrary threshold of -9 dB and among the weakest in group 1; also in terms of anthropometry, this pinna was situated close to the classification boundary in Fig. 5, with both a fossa depth and a cymba depth close to the overall average.

D. Estimation of $F_{\text{cav-cym}}$ from pinna anthropometry

Although the 37 Cavum-Cymba normal modes had a wide range of fossa involvement to the extent that eight were nominally second, rather than third, modes, they all shared the acoustic-physiological characteristic of opposite-phase anti-nodes in cavum and cymba. Therefore all 37 pinnae were used in evaluating regression models for estimation of $F_{\text{cav-cym}}$ (center-frequency of the PRTF peak associated with the Cavum-Cymba mode regardless of fossa involvement) from anthropometry.

Table III lists the six best models (ranked by correlation coefficient r). As with the results in Table II, figures in parentheses pertain to cross-validation; as this resulted in the same set of top six distances, henceforth only the models obtained using all the available data will be discussed. In Table III, the negative sign on every coefficient m again suggests that each of these distances can be regarded as a (partial) measurement of the resonance wavelength. Indeed, as anticipated in Sec. II B, among the top three performers was one of the supplementary measures approximating the path length from cavum to cymba; in particular, the 3D distance $d_{3\text{-C6-12}}$ (in cm) from the concha anterior wall (L3) to the supratragic rim (LC6) and from there to the anterior wall of the cymba (L12),

$$\hat{F}_{\text{cav-cym}} = 14304 - 2757 d_{3\text{-C6-12}}, \tag{6}$$

with estimated $F_{\text{cav-cym}}$ in Hz, $r = 0.82$, and MAE $= 341$ Hz or 4.5%. Although yielding lower performance, the measures ranked fourth to sixth in Table III also approximated the curved path between cavum and cymba anti-nodes: all three involved the sum of two distances, with the cavum end-point being either the canal entrance (L1) or the concha anterior wall (L3), the intermediate landmark being the CIPIC-defined crus helix (LC3), and the cymba end-point being either the cymba base (L7) or the cymba anterior wall (L12).

However, remarkably the overall best performance was attained by the single distance $d_{4\text{-12(sag)}}$ (in cm) measured in a sagittal plane between the lowest point on the concha floor (L4) and the anterior wall of the cymba (L12) as follows:

$$\hat{F}_{\text{cav-cym}} = 15341 - 3142 d_{4\text{-12(sag)}}, \tag{7}$$

with $r = 0.83$ and MAE $= 321$ Hz or 4.3%. A scatter-plot comparing the measured values of $F_{\text{cav-cym}}$ with those estimated by Eq. (7) is shown in the upper panel of Fig. 8.

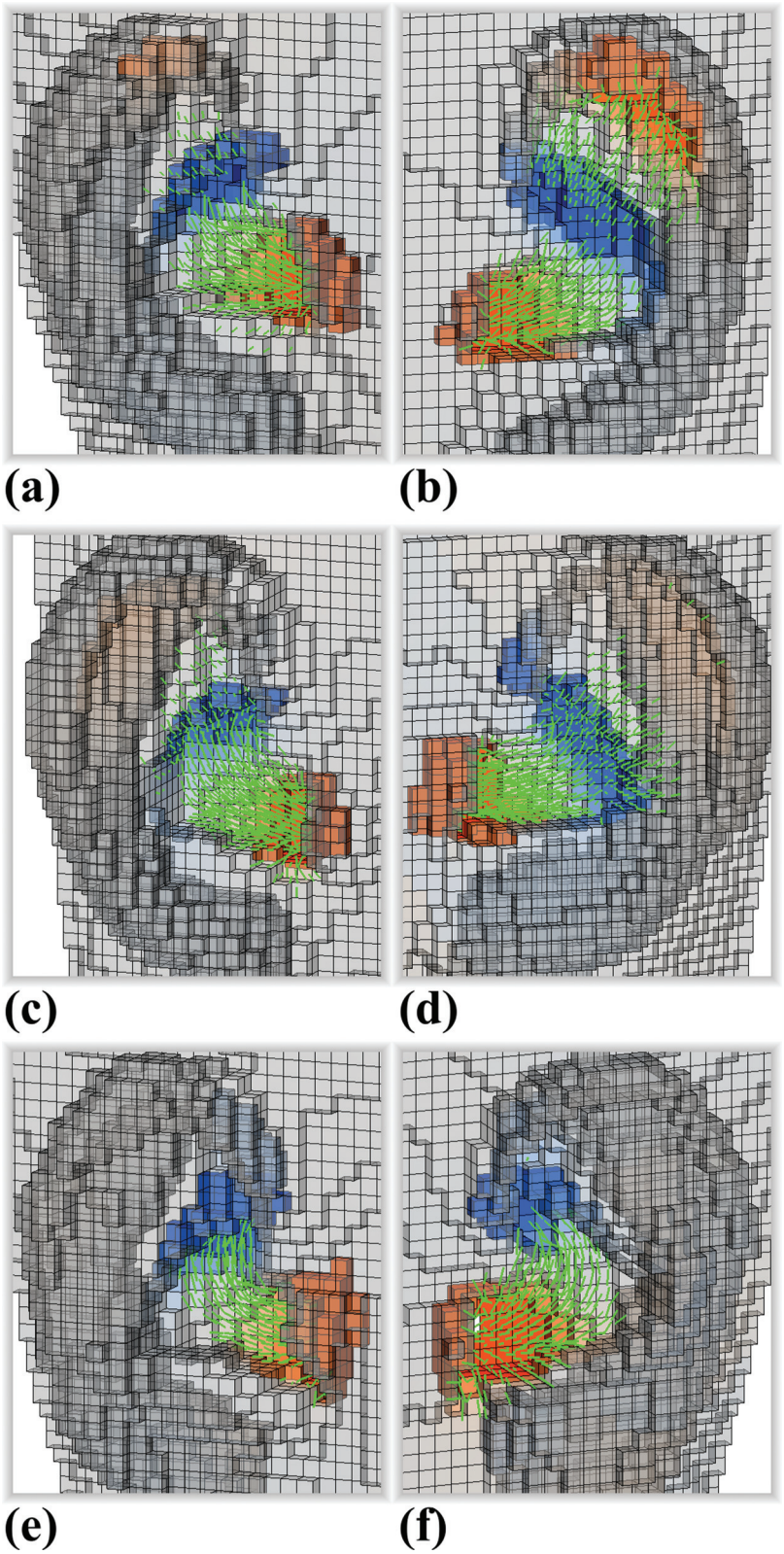

FIG. 7. Cavum-Cymba normal modes of the same four pinnae as in Fig. 4 and two additional pinnae, showing examples of typical patterns and individual differences. Spatial resolution 0.2 cm. Warm/cool colors on pinna surface indicate in-phase/opposite-phase pressure anti-nodes. Green line-segments represent the principal velocity-vector field. (a) and (b) Two examples of a clear third mode with strong fossa involvement (pinnae "fAr" and "fBl"). (c) and (d) Two examples of a third mode with weak fossa involvement (pinnae "mHr" and "mEl"). (e) and (f) Two examples of a clear second mode with no fossa involvement (pinnae "mAr" and "fCl"). Note on color scale: in panels (a)–(d), the color saturation limit was set to 33% of the maximum value in the cavum anti-node; in panels (e) and (f), this limit was raised to 67%.

Indeed, the second-best performance among the single-distance measures was attained by the vertical distance $d_{4\text{–}12(\text{vert})}$ (in cm) between those same two landmarks,

$$\hat{F}_{\text{cav-cym}} = 15631 - 3298 d_{4\text{-}12(\text{vert})}, \quad (8)$$

with $r = 0.82$ and MAE = 328 Hz or 4.4%. This result indicates that even in the best single-distance model of Eq. (7), the *vertical* component of $d_{4\text{-}12(\text{sag})}$ contributed the most, while the front-back component of the distance contributed only a minor improvement.

Retaining $d_{4\text{–}12(\text{sag})}$ as the single best measure for $F_{\text{cav-cym}}$ estimation, multiple linear regression models were trained and evaluated on two parameters at a time to find the next best, complementary measurement. The largest improvement over the model in Eq. (7) was obtained by recruiting the

TABLE III. The top six pinna anthropometric measurements which yielded the best linear regression models for estimation of the Cavum-Cymba normal mode frequency $F_{\text{cav-cym}}$. The models were ranked according to the coefficient of correlation (r) between original and estimated frequencies. The mean absolute error (MAE) is also listed in both Hz and %, but these do not necessarily follow the same ranking. Model coefficients are b (intercept) and m (slope), as in $F_{\text{cav-cym}} = b + m\ \text{d}_{\text{4-12(sag)}}$. Figures in parentheses for r and MAE pertain to results of cross-validation, where data of *one person at a time* were left out of the training set and used for evaluation.

	Distance landmarks	r	MAE (Hz)	MAE (%)	b (Hz)	m (Hz/cm)
1	L4 - L12 (sagittal)	0.83 (0.75)	321 (363)	4.3 (4.9)	15341	−3142
2	L3 - LC6 - L12	0.82 (0.75)	341 (382)	4.5 (5.1)	14304	−2757
3	L4 - L12 (vertical)	0.82 (0.74)	328 (370)	4.4 (5.0)	15631	−3298
4	L3 - LC3 - L7	0.77 (0.66)	386 (442)	5.0 (5.9)	13661	−2308
5	L1 - LC3 - L12	0.77 (0.66)	357 (404)	4.7 (5.4)	13027	−2155
6	L1 - LC3 - L7	0.75 (0.62)	400 (461)	5.1 (6.0)	12619	−2192

distance $\text{d}_{\text{3-C3(horiz)}}$ (in cm) measured in a horizontal plane between the concha anterior wall (L3) and the CIPIC-defined crus helix (LC3),

$$\hat{F}_{\text{cav-cym}} = 16445 - 2911\text{d}_{\text{4-12(sag)}} - 1235\text{d}_{\text{3-C3(horiz)}}, \tag{9}$$

with $r = 0.92$ and MAE = 247 Hz or 3.2%. A scatterplot comparing the measured values of $F_{\text{cav-cym}}$ with those estimated by Eq. (9) is shown in the lower panel of Fig. 8. Any further improvements over the Eq. (9) model will not be discussed, as the next best parameter was not significantly related with the dependent variable $F_{\text{cav-cym}}$.

E. Atypical normal modes

In addition to the typical mode patterns presented so far, a small number of idiosyncratic patterns were also found. Figure 9 presents three such cases that involved Cavum-Fossa modes with $\langle \text{PED} \rangle_{\max}$ located in the triangular fossa (left panels of Fig. 9).

Although the first case [Figs. 9(a) and 9(b), "mIr"] was another example of a pinna with two clear, vertical normal modes, it was atypical in that the Cavum-Cymba mode's upper-most anti-node was also in the triangular fossa [Fig. 9(b)]. In fact, across our 29 Cavum-Cymba (third) modes, the fossa anti-node's $\langle \text{PED} \rangle_{\max}$ was far more typically located in the scaphoid fossa, as in Figs. 7(a)–7(d); it was found to be located in the triangular fossa for only two pinnae (left and right pinnae of the same participant: "mIl" and "mIr"). This may be due to the fact that, among all 38 pinnae, "mIl" and "mIr" had the deepest triangular fossa, as measured by the lateral distance between its base (L14) and posterior corner (L15): range 0.34 cm to 1.06 cm across all pinnae, 1.06 cm for "mIl", and 1.03 cm for "mIr".

The second case [Figs. 9(c) and 9(d), "mJr"] shows that the scaphoid fossa was indeed the preferred location of the Cavum-Cymba mode's upper anti-node, even for a pinna whose Cavum-Fossa mode's upper anti-node was mainly in the triangular fossa. For this pinna, the fossa anti-node's $\langle \text{PED} \rangle_{\max}$ location was therefore displaced from triangular

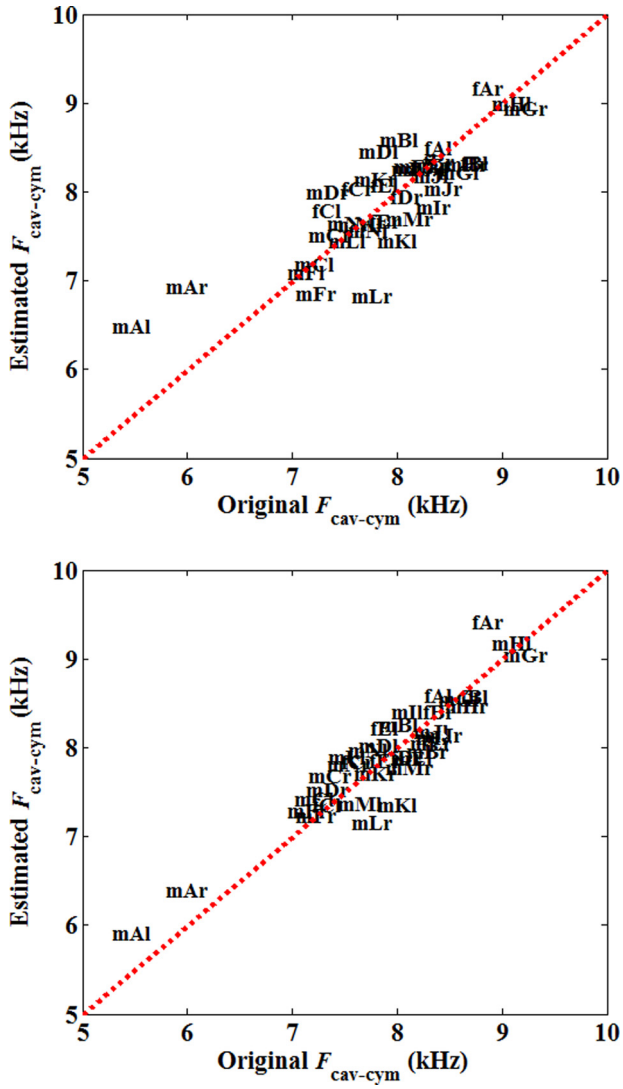

FIG. 8. (Color online) Scatterplots of original versus estimated, normal mode frequency $F_{\text{cav-cym}}$, for the 37 pinnae that were found to have a Cavum-Cymba mode. Upper panel: linear regression model of Eq. (7) with $\text{d}_{\text{4-12(sag)}}$ ($r = 0.83$ and MAE = 321 Hz or 4.3%). Lower panel: linear regression model of Eq. (9) with $\text{d}_{\text{4-12(sag)}}$ and $\text{d}_{\text{3-C3(horiz)}}$ ($r = 0.92$ and MAE = 247 Hz or 3.2%). The dashed line represents the ideal performance.

fossa (Cavum-Fossa mode) to superior scaphoid fossa (Cavum-Cymba mode). While such a clear displacement was atypical in our data (appearing for only four pinnae), in general it would not be unexpected, as the exact locations of anti-nodes probably depend on an intricate balance among the relative sizes and proximities of cavities that might potentially support a pressure anti-node; all the more so for extended cavities such as the fossa, where the triangular and scaphoid fossae are distinguishable anatomically but can be joined acoustically. Such dependence would be even stronger for the third and higher modes, where three or more anti-nodes co-exist within pinna geometrical constraints.

A different form of this dependence is shown in the third case [Figs. 9(e) and 9(f), "fDl"], where a particularly shallow cymba (cf. Fig. 5) in combination with a large and deep triangular fossa, caused this pinna to be the only one with a missing Cavum-Cymba mode (as noted in Sec. III C). Bypassing the cymba, this pinna's third normal mode involved anti-nodes in

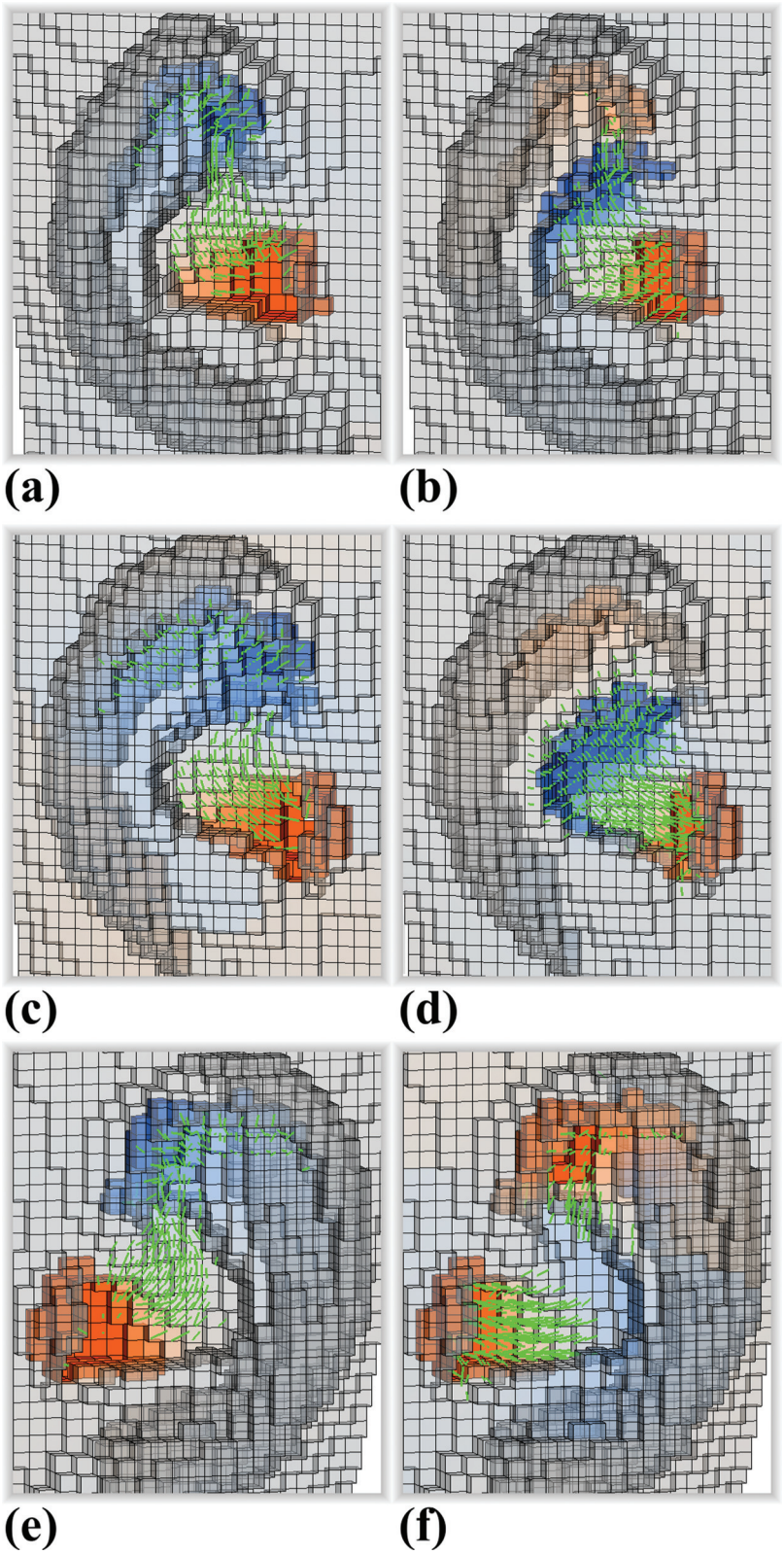

FIG. 9. Atypical pairs of normal modes, of three individual pinnae. Spatial resolution 0.2 cm. Warm/cool colors on pinna surface indicate in-phase/opposite-phase pressure anti-nodes. Green line-segments represent the principal velocity-vector field. (a) and (b) Cavum-fossa (second) and Cavum-Cymba (third) normal modes of "mIr," where in both cases the fossa anti-node was located primarily in the triangular fossa. (c) and (d) Cavum-Fossa (second) and Cavum-Cymba (third) normal modes of "mJr," where the fossa anti-node shifted from triangular fossa to superior scaphoid fossa. (e) and (f) Cavum-Fossa (second) and "horizontal" (third) normal modes of "fDl," where the triangular fossa dominated and thus precluded a Cavum-Cymba mode. Note on color scale: for clarity, the color saturation limit as a percentage of the maximum value in the cavum anti-node was set to 60% in (a) and (b), 25% in (c) and (d), and 67% in (e) and (f).

the posterior concha and triangular fossa. For this reason, the resonance pattern in Fig. 9(f) could not be considered a vertical mode, but instead the first of this pinna's so-called "horizontal" normal modes [resembling Shaw's (1997) fourth mode].

Although atypical in our small sample of 38 pinnae, these cases highlight three possibilities in the distribution of pressure anti-nodes for the third mode (right panels in Fig. 9), despite all three pinnae sharing a common resonance pattern for the second (Cavum-Fossa) mode (left panels in Fig. 9).

IV. CONCLUDING DISCUSSION

A. Vertical normal modes

The first aim of this study was to quantify individual variations in the vertical, pinna normal modes.

1. Physiological bases for individual variations in fossa involvement

While our results on the basic patterns of these modes were in general agreement with previous studies, our multi-subject data revealed that the degree of involvement of the fossa varied over a wide range. This resonance asymmetry was quantified on a continuous scale, by the ratio of the maximum power of the anti-node in the fossa to that in the cavum. Furthermore, our anthropometry data revealed that the fossa's relative involvement depended on individual pinna geometry. As quantified in Eq. (1), pinnae with a deeper scaphoid fossa and a shallower cymba were more likely to have a relatively strong, fossa anti-node, and thus have clear and nearly symmetrical Cavum-Fossa (second) and Cavum-Cymba (third) vertical modes [as with 14 of our pinnae, exemplified in the three pairs of Figs. 4(a)-7(a), 4(b)-7(b), and 9(a)-9(b)]. In contrast, pinnae with a shallower scaphoid fossa and deeper cymba were more likely to have either a weak anti-node in the fossa for both modes [as with ten of our pinnae, exemplified in the two pairs of Figs. 4(c)-7(c) and 4(d)-7(d)], or no fossa involvement in one or both modes [as with nine of our pinnae, exemplified in Figs. 7(e) and 7(f) for the Cavum-Cymba mode].

While it is reasonable that a deeper and better enclosed fossa would support a stronger pressure anti-node, the role of a shallower cymba in promoting the Cavum-Fossa mode is less clear. One interpretation that warrants further investigation, is that a deeper cymba may be more likely to acoustically interfere with the Cavum-Fossa resonance pattern, because this normal mode relies on oscillatory exchange of energy between cavum and fossa, along a path that passes laterally over the cymba (cf. velocity vectors in Fig. 4 and left panels of Fig. 9).

2. Physiologically descriptive labels for pinna normal-modes

In the case of no fossa involvement, the Cavum-Cymba mode is reduced to a nominally *second* mode [cf. Figs. 7(e) and 7(f)]. Thus, in the conventional way of labeling normal modes by their total number of pressure anti-nodes, one pinna's Cavum-Cymba (second) mode may have the same label as another pinna's Cavum-Fossa (second) mode. Naturally, this raises the question of how best to classify these two types of second mode which share the same nominal label, but are morphologically distinct. We have already seen this discrepancy implicitly in previous studies, where the human pinna analyzed by Kahana and Nelson (2006, Fig. 19) appeared to have a Cavum-Cymba (second) mode, in contrast with their manikins' Cavum-Fossa (second) modes. Clearly, if pinna normal modes are to be classified or labeled for the purpose of constructing models for estimation of mode frequencies from anthropometry, it would be erroneous to group these two types of second mode in the same category. By the same token, it would be erroneous to group a third (horizontal) normal mode having anti-nodes in the posterior concha and triangular fossa [cf. Fig. 9(f)], in the same category as a third (vertical) mode involving cymba and fossa. For these reasons, we here proposed the *physiologically descriptive* labels Cavum-Fossa and Cavum-Cymba for the vertical normal modes, which then allowed a correct evaluation of anthropometric measurements in estimation models.

In a similar vein, in this study we avoided labeling transfer function peaks as "P2," "P3," and so on. In the research literature on HRTFs/PRTFs, such generic terms have understandably been used to label peaks in increasing order of center-frequency, but the methods and assumptions used to extract the peaks have varied: some studies extracted peaks on only one transfer function (e.g., in the frontal direction), other studies extracted peaks on a mean transfer function (e.g., averaged over elevation angles in the median plane), and yet other studies extracted each peak by approximating a nearly constant-frequency ridge on 2D images of transfer functions across a band of spatial locations (e.g., in the median plane). Thus, there does not appear to be an established method for extracting and correctly labeling transfer function peaks across a wide range of spatial locations. A fundamental issue in this regard, as mentioned in Sec. II D, is that each normal mode of an individual pinna has its own, characteristic resonance pattern (nearfield) and radiation pattern (farfield), with the second and higher modes including spatial directions at which acoustic wave interferences cause nulls, or spectral notches (Takemoto *et al.*, 2012, 2013). Consequently, not all normal modes appear as transfer function peaks at all spatial locations. Therefore, a "P2" or a "P3" measured in one study will not necessarily be the same normal mode referred to by similarly labeled peaks in a different study. For these reasons, we here proposed a method based on the histogram of transfer function peaks, to identify every possible, candidate normal mode across a wide range of spatial locations. These methods, together with the labeling of each mode according to its visualized resonance pattern rather than simply the order in which the peaks appeared along the frequency axis, also ensured that our anthropometric data would be evaluated correctly in frequency estimation models.

Our combined statistical methods (histogram-based detection of all potential normal modes) and analytical methods (visualization-based identification of normal mode type) provide useful tools for determining the physical basis of observed peaks in HRTFs or PRTFs. In this study, these methods were critical in clarifying the physical mechanisms that generated the second and third peaks, and the mechanism by which a second normal mode may appear weak or non-existent.

3. Current limitations and future work

Nevertheless, our method had the limitation of assuming that all the PRTF peaks underlying a smoothed-histogram peak corresponded to the same normal mode; while in reality, depending on pinna geometry, the center-frequencies of different normal modes may overlap. Though computationally prohibitive in the present study, a brute force way to detect and separate such degenerate modes would be to visualize and classify the resonance patterns corresponding to *every* PRTF peak, i.e., at hundreds of spatial locations, rather than only the one location corresponding to the highest-amplitude

peak as chosen here. It would be desirable to find more efficient ways to overcome this limitation.

Also, as noted in Sec. III A, five candidate mode patterns were rejected from consideration as Cavum-Fossa normal modes, because either the dominant pressure anti-node lay outside the main pinna cavities or an anti-node's $\langle \mathrm{PED} \rangle_{\max}$ was located far from the pinna surface. Though rejected as vertical normal modes, the resonance mechanisms of such cases, including the role of cavities behind the pinna, remain to be clarified.

B. Frequency estimation from pinna anthropometry

The second aim of this study was to derive accurate and practical models for estimation of the PRTF peaks associated with vertical modes, from pinna anthropometry.

1. Estimation of $F_{\text{cav-fos}}$

The best estimate of $F_{\text{cav-fos}}$ requiring only two pinna measurements, was provided by the 3D distance $d_{1\text{-}17}$ between canal entrance and helix rim, and the horizontal-plane distance $d_{7\text{-}11(\text{horiz})}$ between the cymba's base and superior rim [$r = 0.89$, MAE = 257 Hz or 4.4%, cf. Eq. (5) and lower panel of Fig. 6]. The weighted sum of these two, geometrically separate distances [as specified in the Eq. (5) model] outperformed the straight sum of two, conjoint distances [i.e., the Eq. (2) model] which was intended to approximate the path length between cavum and fossa anti-nodes. Indeed, $d_{1\text{-}17}$ which was the best-performing single measurement (cf. Table II), was also the simplest approximation to the distance between cavum and fossa. A common factor among the top ten measures listed in Table II, was the representation of the fossa end-point by either the helix rim (L17) or the CIPIC-defined helix rim (LC5). Either way, the results in Table II underscore the importance of the helix rim as a robust anchor for measurements related with the Cavum-Fossa normal mode. Meanwhile, these results also indicate that the measurement need not traverse the entire distance to the cavum anti-node: landmarks lying approximately along the fossa-to-cavum path (cymba base L7, crus helix base L6, concha base L2, canal entrance L1, and concha anterior wall L3) all yielded close to the best performance. Indeed, across the 33 pinnae, the top 10 measures were found to be closely related to each other: correlation coefficients between either the best performer $d_{1\text{-}C3\text{-}17}$ or the best single-distance $d_{1\text{-}17}$, and each of the other nine measures, always exceeded 0.89. Therefore, the comparable performances are likely due to anatomical constraints on human pinna geometry, and the foregoing results suggest that a good estimation model requires anatomical distances that measure a sufficiently similar fraction of the resonance wavelength across different pinnae, regardless of the fraction itself. Thus, although the fractions of a wavelength measured by the distances using those five lower landmarks were different (e.g., $d_{1\text{-}17}$ measured the distance approximately between anti-nodes, or a half-wavelength, while $d_{6\text{-}C5}$ measured the distance approximately from node to anti-node, or a quarter-wavelength), their comparable performances suggest that the distances, and their associated landmarks, must have been comparably robust across pinnae.

Regarding the best complementary parameter $d_{7\text{-}11(\text{horiz})}$, we note that in both Eqs. (4) and (5) the coefficient on that parameter had a negative sign, meaning that a greater cymba depth yielded a lower $F_{\text{cav-fos}}$, as would be expected if $d_{7\text{-}11(\text{horiz})}$ were contributing to an estimate of the resonance wavelength. In this context, it may be relevant to recall that a deeper cymba was found to predict weaker fossa involvement [cf. Fig. 5], and that Cavum-Fossa normal modes with weaker fossa involvement typically had a cavum anti-node that covered a larger area of the concha base, including part of the cymba [cf. Figs. 4(c) and 4(d)]. From elementary acoustic theory it is well known that for a closed-closed resonant tube, increasing the cross-sectional area at a closed end (pressure anti-node) will lower the resonance frequency, or effectively increase the resonance wavelength. Although the open structure of the human pinna makes analogies with a closed resonant tube at best hypothetical, it is nevertheless possible that the wider area of the concha base that supported the cavum anti-node for pinnae with a deeper cymba, contributed to a longer resonance wavelength (and a lower $F_{\text{cav-fos}}$) by a similar mechanism. This hypothesis remains to be tested.

2. Estimation of $F_{\text{cav-cym}}$

The best two-parameter estimate of $F_{\text{cav-cym}}$ was provided by the sagittal-plane distance $d_{4\text{-}12(\text{sag})}$ between the lowest point on the concha floor and the anterior wall of the cymba, and the horizontal-plane distance $d_{3\text{-}C3(\text{horiz})}$ between the concha anterior wall and the CIPIC-defined crus helix [$r = 0.92$, MAE = 247 Hz or 3.2%, cf. Eq. (9) and lower panel in Fig. 8]. Remarkably, the weighted sum of these two, geometrically disjoint distances [i.e., the Eq. (9) model] outperformed the straight sum of two, conjoint distances [i.e., the Eq. (6) model] which was intended to approximate the path length between cavum and cymba anti-nodes. Even though the best two measurements did not share an intermediate landmark and therefore cannot be regarded as measuring a longer, composite distance, in a loose sense they can be considered to measure complementary (or spatially, nearly orthogonal) distances between the anti-nodes: $d_{3\text{-}C3(\text{horiz})}$ is by definition a horizontal measure of cavum depth, while, as noted in Sec. III D, $d_{4\text{-}12(\text{sag})}$ is primarily a vertical measure of entire-concha height.

It is not immediately clear why $d_{4\text{–}12(\text{sag})}$ alone outperformed even the best supplementary measure, as it appears to be only a crude approximation of the cavum-to-cymba path. However, as noted earlier, a likely explanation is that the best estimation models do not necessarily require a complete measure of the resonance wavelength (or the full length of the path between anti-nodes), but instead require anatomical landmarks whose relative positions across different pinnae robustly scale in proportion to the wavelength; i.e., any fraction of a wavelength will suffice, provided the same (or similar) fraction is measured on all, or most, pinnae used to train the model. It would therefore appear that in relation to the Cavum-Cymba mode, the concha floor (L4) and cymba anterior wall (L12) provide such robust anatomical anchors.

3. Current limitations and future work

While these results were the best among 3322 anthropometric measures, our list of 31 pinna landmarks (cf. Table I and Fig. 2) from which the measures were derived, was not comprehensive. Thus in future research, special emphasis should be placed on improving or extending the list, and on augmenting these measures with other types that may be even better descriptors of acoustically related, pinna geometry.

Though the main application of our estimation models is in HRTF personalization, they may also provide a way to predict the physical origin (i.e., either Cavum-Fossa or Cavum-Cymba resonance) of the second-lowest peak observed in an acoustically measured transfer function. The resonance mode could for example, be inferred according to whether the measured peak is closer in frequency to either $F_{\text{cav-fos}}$ or $F_{\text{cav-cym}}$ estimated from that person's pinna anthropometry.

The present study considered only the normal modes' center-frequencies. For a more complete specification of an individual pinna's transfer functions, it may also be important to investigate individual variations in, and estimation of, the modes' bandwidths or amplitudes. In this context, our numerical results showed that the cymba anti-node is usually more energetic than the fossa anti-node, which agrees with the general observation that the cymba is the deeper and better enclosed cavity. Thus in general, the transfer function peak associated with the Cavum-Cymba mode ought to be more prominent (narrower bandwidth, higher amplitude, and less variable center-frequency across spatial locations as partially exemplified in Fig. 3) than the peak associated with the Cavum-Fossa mode. These relations, and estimation of the peak amplitudes from pinna anthropometry, remain to be investigated.

C. Practical considerations

Among our 38 pinnae, only five were missing a Cavum-Fossa mode ("fAl," "fCr," "fEl," "fEr," and "mMr") and only one was missing a Cavum-Cymba mode ["fDl," cf. Figs. 9(e) and 9(f)]. In contrast, disregarding individual variations in the degree of fossa involvement, the majority of pinnae (32/38, or 84%) were found to have *both* types of vertical mode: a Cavum-Fossa mode (as in Fig. 4) and a Cavum-Cymba mode (as in Fig. 7). Thus, with transfer function peaks numbered by considering all spatial locations (i.e., according to the histogram method of Sec. II D), for the majority of pinnae $F_{\text{cav-fos}}$ corresponded to the second peak "P2" and $F_{\text{cav-cym}}$ corresponded to the third peak "P3." For personalization of HRTFs for each pinna of a new listener, it is therefore advisable to estimate *both* mode center-frequencies $F_{\text{cav-fos}}$ and $F_{\text{cav-cym}}$.

In practice, the most efficient way would be to use the best estimation models that required only one measurement each. Thus, $F_{\text{cav-fos}}$ would be estimated with Eq. (3) which requires only the distance between canal entrance and helix rim (left panel of Fig. 10); and $F_{\text{cav-cym}}$ would be estimated with either Eq. (7) or Eq. (8), which require only the sagittal-plane or vertical distance, respectively, between concha floor and cymba anterior wall, i.e., approximately the entire-concha height (right panel of Fig. 10). As depicted in Fig. 10, these distances can be measured simply with two cotton buds attached to the pointer tips of a pair of calipers. Estimation accuracy may then be improved by using Eq. (5) for $F_{\text{cav-fos}}$ and Eq. (9) for $F_{\text{cav-cym}}$, if convenient methods can also be found for measuring the two additional distances required in those models.

FIG. 10. (Color online) Practical measurement of pinna anthropometric distances for estimation of Cavum-Fossa and Cavum-Cymba normal mode frequencies, using cotton buds attached to a pair of calipers, on pinna "fAr" [the same pinna depicted in Figs. 4(a) and 7(a)]. Left panel: measurement of the distance $d_{1\text{-}17}$ between ear-canal entrance and helix rim, for estimation of $F_{\text{cav-fos}}$ using Eq. (3). Right panel: measurement of the vertical distance $d_{4\text{-}12(\text{vert})}$ between concha floor and cymba anterior wall, for estimation of $F_{\text{cav-cym}}$ using Eq. (8).

Psychophysical experiments have suggested that for determination of sound source elevation angle in the median plane, the most important cues are the first peak and the first two notches (Iida *et al.*, 2007). Our earlier efforts yielded models for estimation of the first peak from individual pinna anthropometry (Mokhtari *et al.*, 2015). Hereafter, our proposed models for estimation of the two vertical modes should pave the way for estimating the trajectory of the first notch in the median plane according to the method of Takemoto *et al.* (2013): this method is based on the radiation pattern of an essentially vertical acoustic dipole, and assumes prior knowledge of the center-frequency of the relevant normal mode. While our estimation models would therefore be a critical first step towards estimating the first notch frequency, it remains to be clarified whether the Cavum-Fossa or the Cavum-Cymba normal mode is the more appropriate vertical dipole required in that method.

It is hoped that these methods together will be useful for spatial audio systems having the flexibility of HRTF personalization. In a practical system, a set of simplified HRTFs might be calculated on the basis of these peak and notch estimates. The listener might then be given the opportunity to adjust the estimated values, and to perceptually evaluate the inclusion or exclusion of the second peak [especially if Eq. (1) were to predict only weak fossa involvement for that listener]. This may be achieved by using a virtual auditory display with flexible specification of transfer function peaks (as, e.g., with the parametric HRTFs of Iida *et al.*, 2007).

The perceptual efficacy of our estimation models, both in regard to sound source externalization and localization, remain to be evaluated by psychophysical experiments.

Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (**2001**). "The CIPIC HRTF database," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 99–102.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization*, revised ed. (The MIT Press, Cambridge, MA), pp. 36–200.

Iida, K., Ishii, Y., and Nishioka, S. (**2014**). "Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae," J. Acoust. Soc. Am. **136**(1), 317–333.

Iida, K., Itoh, M., Itagaki, A., and Morimoto, M. (**2007**). "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," Appl. Acoust. **68**, 835–850.

Jin, C., Leong, P., Leung, J., Corderoy, A., and Carlile, S. (**2000**). "Enabling individualized virtual auditory space using morphological measurements," in *Proceedings of 1st IEEE Pacific-Rim Conference on Multimedia (International Symposium on Multimedia Information Processing)*, Sydney, Australia, pp. 235–238.

Kahana, Y., and Nelson, P. A. (**2006**). "Numerical modelling of the spatial acoustic response of the human pinna," J. Sound Vib. **292**, 148–178.

Middlebrooks, J. C. (**1999**). "Individual differences in external-ear transfer functions reduced by scaling in frequency," J. Acoust. Soc. Am. **106**(3), 1480–1492.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2008**). "Efficient computation of HRTFs at any distance by FDTD simulation with near to far field transformation," in *Proceedings of the Autumn Meeting of the Acoustical Society of Japan*, Fukuoka, Japan, Paper 1-8-12, pp. 611–614.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2010a**). "Optimum loss factor for a perfectly matched layer in finite-difference time-domain acoustic simulation," IEEE Trans. Audio Speech Language Processing **18**(5), 1068–1071.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2010b**). "Acoustic sensitivity to micro-perturbations of KEMAR's pinna surface geometry," in *Proceedings of the 20th International Congress on Acoustics (ICA)*, Sydney, Australia, Paper 790, pp. 1–8.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2011**). "Computer simulation of KEMAR's head-related transfer functions: Verification with measurements and acoustic effects of modifying head shape and pinna concavity," in *Principles and Applications of Spatial Hearing*, edited by Y. Suzuki, D. Brungart, Y. Iwaya, K. Iida, D. Cabrera, and H. Kato (World Scientific, Singapore), pp. 205–215.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2014**). "Visualization of acoustic pressure and velocity patterns with phase information in the pinna cavities at normal modes," in *Proceedings of the 39th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Firenze, Italy, pp. 8267–8271.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2015**). "Frequency and amplitude estimation of the first peak of head-related transfer functions from pinna anthropometry," J. Acoust. Soc. Am. **137**(2), 690–701.

Morimoto, M., and Aokata, H. (**1984**). "Localization cues of sound sources in the upper hemisphere," J. Acoust. Soc. Jpn. **5**(3), 165–173.

Nishino, T., Inoue, N., Takeda, K., and Itakura, F. (**2007**). "Estimation of HRTFs on the horizontal plane using physical features," Appl. Acoust. **68**, 897–908.

Pierce, A. D. (**1989**). *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, Melville, NY), pp. 180–183.

Satarzadeh, P., Algazi, V. R., and Duda, R. O. (**2007**). "Physical and filter pinna models based on anthropometry," in *Proceedings of the 122nd Convention of the Audio Engineering Society*, Vienna, Austria, Paper 7098, pp. 1–20.

Shaw, E. A. G. (**1997**). "Acoustical features of the human external ear," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 25–47.

Spagnol, S., and Avanzini, F. (**2015**). "Frequency estimation of the first pinna notch in head-related transfer functions with a linear anthropometric model," in *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx)*, Trondheim, Norway, pp. 1–6.

Taflove, A., and Hagness, S. C. (**2005**). *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. (Artech House, Norwood, MA), Chap. 5, pp. 169–228.

Takemoto, H., Mokhtari, P., Kato, H., Nishimura, R., and Iida, K. (**2012**). "Mechanism for generating peaks and notches of head-related transfer functions in the median plane," J. Acoust. Soc. Am. **132**(6), 3832–3841.

Takemoto, H., Mokhtari, P., Kato, H., Nishimura, R., and Iida, K. (**2013**). "Elevation-frequency relationship for the first spectral notch of pinna-related transfer functions in the median plane," in *Proceedings of the Autumn Meeting of the Acoustical Society of Japan*, Toyohashi, Japan, Paper 3-1-11, pp. 843–846 (in Japanese).

Takemoto, H., Mokhtari, P., and Kitamura, T. (**2010**). "Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method," J. Acoust. Soc. Am. **128**(6), 3724–3738.

Wenzel, E., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). "Localization using nonindividualized head-related transfer functions," J. Acoust. Soc. Am. **94**(1), 111–123.

Xiao, T., and Liu, Q. H. (**2003**). "Finite difference computation of head-related transfer function for human hearing," J. Acoust. Soc. Am. **113**(5), 2434–2441.

Zotkin, D. N., Duraiswami, R., Grassi, E., and Gumerov, N. A. (**2006**). "Fast head-related transfer function measurement via reciprocity," J. Acoust. Soc. Am. **120**(4), 2202–2215.