Frequency and amplitude estimation of the first peak of head-related transfer functions from individual 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 September 2014; revised 3 December 2014; accepted 8 December 2014)

The first (lowest) peak of head-related transfer functions (HRTFs) is known to be a concha depth resonance and a spectral cue in human sound localization. However, there is still no established model to estimate its center-frequency F_1 and amplitude A_1 from pinna anthropometry. Here, with geometries of 38 pinnae measured and their median-plane HRTFs calculated by numerical simulation, linear regression models were evaluated in estimating F_1 and A_1 from 25 concha depth and aperture measurements. F_1 was best estimated (correlation coefficient $r = 0.84$, mean absolute error MAE = 118 Hz) by lateral distances from the base of the posterior cavum concha to the outer surface of the antitragus and antihelix (longest measures of concha depth). A_1 was best estimated ($r = 0.83$, MAE = 0.84 dB) by the lateral distance from the ear-canal entrance to the side of the cheek near the anterior notch (shortest measure of concha depth) and by the equivalent diameter of the concha aperture. These results suggest that the first resonance's quarter-wavelength corresponds to the longest lateral extent of the concha and that its energy lost to the surrounding air depends on the concha aperture and the cavum concha's shortest lateral depth.
 [http://dx.doi.org/10.1121/1.4906160]

[BLM] Pages: 690–701

I. INTRODUCTION

Humans can localize sounds in three-dimensional (3D) space by using head-related transfer functions (HRTFs), which are characterized by spectral peaks and notches (Shaw and Teranishi, 1968; Blauert, 1997). As the peaks and notches depend on individual head and pinna (external ear) geometry, a listener's own HRTFs are necessary for externalization and accurate localization of sounds presented in virtual auditory displays (Wenzel *et al.*, 1993). Physically, the first peak is generated by the first normal mode of the pinna with a center-frequency around 4–5 kHz depending on the shape and size of the individual pinna. Acoustic measurements (Shaw and Teranishi, 1968), computer simulations (Kahana and Nelson, 2006; Takemoto *et al.*, 2012), and sensitivity analyses of pinna geometry (Mokhtari *et al.*, 2010b; Mokhtari *et al.*, 2013) have shown that the first normal mode resonates with the entire base of the concha (including cavum concha and cymba) supporting a pressure anti-node (anatomical nomenclature is indicated schematically in Fig. 1). As the concha shape roughly resembles a shallow acoustic tube closed at its base and open laterally, it has been presumed that about one-quarter of a wavelength extends outward to the concha rim and surrounding air. The first normal mode, and by association the first peak, has therefore been appropriately called a "concha depth" resonance (Shaw and Teranishi, 1968).

As the first normal mode has an approximately omnidirectional resonance pattern in the farfield, it alone does not provide localization cues. However, results of psychophysical experiments have suggested that the first peak (of HRTFs measured with blocked meatus) could act as an anchor in reference to which the human auditory system analyzes other spectral cues (specifically the first two notches) to determine sound source elevation angle in the median plane (Iida *et al.*, 2007). In this sense, personalization of the center-frequency and amplitude of the first peak could be a critical first step in any applications involving 3D audio (including, e.g., entertainment, navigation, and healthcare systems).

Despite this basic knowledge on the psychophysical importance of the first peak in sound localization and the physical mechanism by which it is generated, there is still no established method or model to estimate its center-frequency F_1 and amplitude A_1 (or indeed, the center frequency and amplitude of any HRTF peak) from individual pinna geometry. In a recent study (Ishii and Iida, 2013) using acoustics and anthropometry of 46 ears, F_1 was extracted from HRTFs measured directly in front of each listener at a distance 1.2 m from the ear canal, and multiple linear regression models were trained to estimate F_1 from various subsets of 11 anthropometric parameters. By backward elimination, they found the best three-parameter estimation model to include concha depth, width, and length, with a correlation coefficient of 0.56 and a mean absolute error of 160 Hz in F_1. Although no details were given in that study concerning their specific definition of the concha depth measurement, it was subsequently revealed (Iida *et al.*, 2014) to be the same as distance "d_8" suggested in the documentation accompanying the CIPIC HRTF database (Algazi *et al.*, 2001, Fig. 3): Briefly, "d_8" is the concha depth measured perpendicular to a straight line joining the lateral surface of the tragus and the

[a]Author to whom correspondence should be addressed. Electronic mail: parham@nict.go.jp

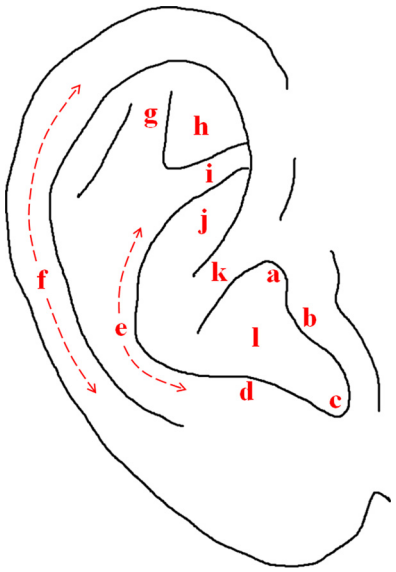

FIG. 1. (Color online) Anatomical nomenclature illustrated on a schematic of a right pinna. Key to annotation: (a) Supratragic (or anterior) notch, (b) tragus, (c) intertragic notch, (d) antitragus, (e) antihelix, (f) helix, (g) upper crus of antihelix, (h) triangular fossa, (i) lower crus of antihelix, (j) cymba, (k) crus of helix, (l) cavum concha. The ear-canal entrance is in the anterior part of (l), hidden behind (b). The "entire concha" refers to (l), (k), and (j) combined. The "concha aperture" refers to the outline of the entire concha along its rim, joining (a), (b), (c), (d), (e), and (i).

antihelix (an example is shown later in Fig. 3). From this and the studies cited earlier, it would seem that F_1 is related primarily with concha depth and perhaps also partly influenced by concha aperture. However, due to both the complicated shape of the pinna and inter-individual differences in concha geometry, it is still not known what kind of concha depth and aperture measurements are most closely associated with F_1 across a range of individuals. Furthermore, although elementary acoustical theory has shown an approximately linear relation between the pressure gain and the ratio of the depth and radius of a cylindrical resonator (e.g., Teranishi and Shaw, 1968, Fig. 3), there is still no model for estimation of A_1 from pinna anthropometry nor indeed an account of how A_1 varies with human pinna geometry.

Here we aim to derive more accurate and practical models to estimate both F_1 and, for the first time, A_1, from individual pinna anthropometry. For this purpose, head and pinna geometries of 19 adults (i.e., 38 pinnae) were measured by magnetic resonance imaging (MRI). From these morphological data, farfield transfer functions were calculated by numerical simulation with an acoustic finite-difference time domain (FDTD) method that was previously shown to yield HRTFs (Mokhtari *et al.*, 2011) and vocal-tract transfer functions (Takemoto *et al.*, 2010) that matched well with acoustic measurements. As the first normal mode is essentially independent of source incident direction (e.g., Kahana and Nelson, 2006; Mokhtari *et al.*, 2010b, 2011), representative values for F_1 and A_1 of each pinna were extracted from the mean of transfer functions in the median plane. Motivated by the hypothesis that both F_1 and A_1 depend mainly on pinna dimensions related to concha depth and aperture, a set of anatomically consistent landmarks were manually identified on each pinna, including 4 landmarks in the concha base and 12 landmarks around the concha rim. Linear regression models were then trained and evaluated for their accuracy in estimating F_1 and A_1 from each of 25 pinna measurements (including CIPIC "d_8") that were derived from various combinations of the anatomical landmarks. Multiple linear regression was then used in an attempt to improve estimation performance by recruiting additional parameters.

Section II describes the materials and methods, Sec. III presents the results, and Sec. IV concludes the paper with a discussion including an interpretation of the best models in terms of the physical determinants of F_1 and A_1.

II. MATERIALS AND METHODS

A. Head and pinna geometries

Head geometries of 19 adults (5 women and 14 men) were measured by MRI with either a Siemens 3.0 T Magnetom Trio (18 participants) or a Shimadzu-Marconi 1.5 T Magnex Eclipse (1 participant) installed at the Brain Activity Imaging Center, ATR-Promotions Inc. (Kyoto, Japan). The spatial resolution of the MRI data (i.e., the size of each voxel) was either 1.0 mm (11 participants) or 1.2 mm (8 participants). Each head volume was processed in several stages with in-house software developed in MATLAB (The Mathworks, Inc.): (a) The grayscale data were binarized by thresholding to distinguish air and non-air voxels, (b) the ear canals were occluded at their entrance to ensure blocked-meatus conditions, and (c) a contiguous head volume was segmented from the surrounding air by a 3D region-growing algorithm, after occluding all remaining inlets (e.g., at the tracheal, esophageal, oral, and nasal airways). These data at either 1.0 or 1.2 mm spatial resolution were used for rendering and visualizing the 3D heads and pinnae and for manually defining pinna-surface landmarks to be used for anthropometric measurements (Sec. II B). Each head volume was also downsampled to an isovoxel resolution 2.0 mm for the purpose of acoustic simulations (Sec. II C) to ease computation load while ensuring sufficient accuracy up to 14 kHz (Mokhtari *et al.*, 2010a, 2011).

B. Pinna anthropometry

As introduced earlier, it is known that the first peak of HRTFs is essentially a lateral resonance that depends mainly on concha depth and perhaps also concha aperture. However, due to both the complicated shape of the pinna and inter-individual differences in concha geometry, it is not known what type of depth or aperture measurements are most suitable for estimation of F_1 and A_1 across a range of individuals. To address this question, anatomically consistent landmarks were identified manually on the surface of a 3D rendered model of each pinna. The models used for this purpose were rendered from the processed volumetric data at the original resolution (1.0 or 1.2 mm) to mimic as closely as

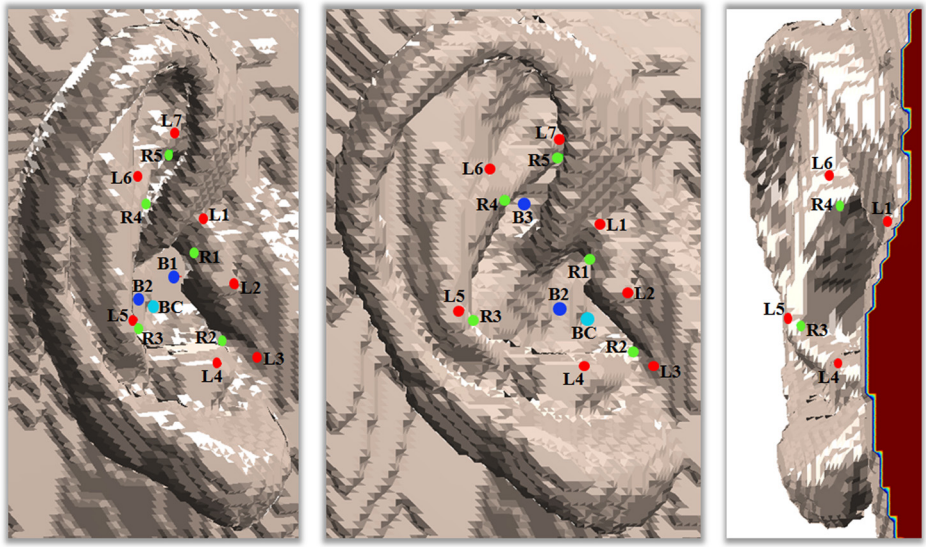

FIG. 2. (Color online) Pinna landmarks used in this study, shown here in three views of a right pinna (right rear view, right side view, and front view): Base landmarks B1–B3 and BC, rim landmarks R1–R5, and lateral landmarks L1–L7. Anatomical definitions are given in the text (Sec. II B) and in reference to Fig. 1.

possible the task of identifying landmarks and making measurements on real human pinnae.

As exemplified in Fig. 2 and listed in Table I, a total of 16 landmarks were identified on each pinna (refer to Fig. 1 for anatomical nomenclature) with the aim of providing a variety of concha depth and aperture measurements that are reasonably easy to identify consistently on all the pinnae. Four landmarks served to define the concha base (B1–B3, and BC), and 12 landmarks were located on or near the concha rim (R1–R5, and L1–L7) as described in the following text.

Base landmark B1 coincided with the plane of the ear-canal entrance, which was always the deepest part of the cavum concha when regarded laterally (i.e., parallel to the interaural axis). More specifically, B1 was located at the center of the (blocked) ear-canal entrance. As an alternative landmark that would be easier to define on human pinnae with an unobstructed ear canal, B2 was located on the most posterior part of the cavum concha base, posterior to the ear-canal entrance and inferior to the crus of helix. More specifically, B2 was located immediately anterior to the largest change in curvature rising toward the ridge formed by the lower part of the crus of helix. Thus due to the generally concave shape of the cavum concha, B2 coincided with the most lateral part of the cavum concha base (i.e., its lateral position was either equal to or more lateral than that of B1). Base landmark BC was also located in the posterior part of the cavum concha but at a position conforming with the CIPIC "d_8" measurement: At the deepest part of the cavum concha when viewed normal to the concha aperture (i.e., normal to the line joining landmarks L2 and L5, as described in the following text and illustrated in the top panel of Fig. 3). As the pressure anti-node associated with the first normal mode has previously been shown to cover the base of the entire concha including the cymba (e.g., Kahana and Nelson, 2006, Fig. 18; Takemoto *et al.*, 2012, Fig. 8a), landmark B3 was located at the deepest part of the cymba.

The concha depth varies along its perimeter, not only due to the varying protrusion of its base, but also because of the varying amount by which the concha rim extends laterally. To facilitate measurements of concha depth according to various possible definitions of its lateral extent, seven "lateral" (L) landmarks were located: On the side of the cheek close to the anterior notch (L1), on the most lateral surface of the tragus (L2), on the side of the jaw close to the intertragic notch (L3), on the most lateral surface of the antitragus (L4), on the most lateral surface of the antihelix (L5), at the posterior corner of the triangular fossa (L6), and at the least lateral point along the lower crus of antihelix (L7). Although the triangular fossa sometimes appeared as a rather shallow cavity with an ill-defined posterior corner, in such cases, L6 was identified as the intersection of two curves approximating its inferior and posterior/superior boundaries. Twenty-one candidate measurements of concha depth were then defined as the lateral distance from each of the three base landmarks B1–B3, to each of the seven lateral landmarks L1–L7. Thus for example, the bottom panel of Fig. 3 illustrates the lateral distance from B2 to L4, which will be denoted as $d_{B2\text{-}L4}$.

In approximate conformance with the CIPIC "d_8" measurement, d_{CIPIC} was calculated as the distance from base landmark BC to a straight line running through lateral landmarks L2 and L5 (cf. top panel of Fig. 3). As the original schematic for CIPIC "d_8" (Algazi *et al.*, 2001, Fig. 3) appeared to suggest a two-dimensional measurement in a horizontal plane, for completeness $d_{CIPIC\text{-}2D}$ was also calculated by ignoring any differences in the height of the three landmarks. However, as the two distances d_{CIPIC} and $d_{CIPIC\text{-}2D}$ were found to be highly correlated ($r = 0.98$), only the full 3D version d_{CIPIC} was retained here.

To enable measurements of concha aperture, five "rim" (R) landmarks were located: At the anterior-most point along the rim of the supratragic (or anterior) notch (R1), at the rim of the intertragic notch (R2), at the most posterior-inferior corner along the rim of the antihelix (R3), at the rim of the antihelix close to the posterior corner of the triangular fossa (R4), and at the superior rim of the cymba (R5).

The concha aperture was represented first in standard terms by its overall width and height. Concha width W was measured as the distance between landmarks R1 and R3.

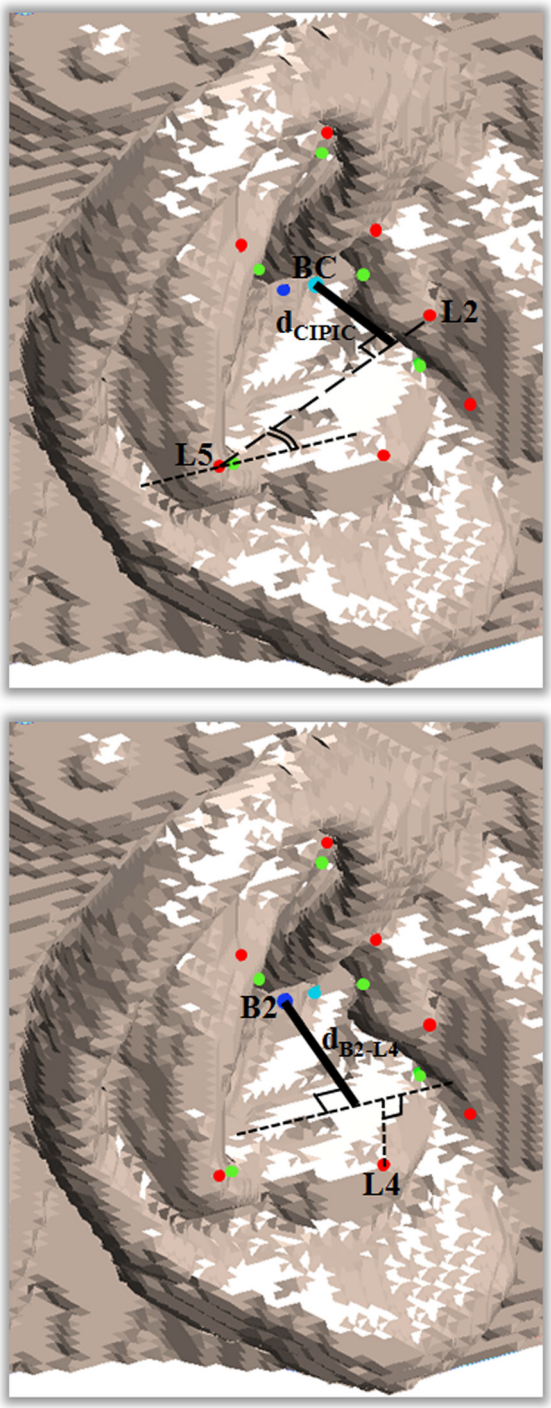

FIG. 3. (Color online) Two examples of concha depth measurements (thick solid lines) as seen from an elevated viewing angle on the same pinna shown in Fig. 2. Top panel: d_{CIPIC} is the distance from landmark BC to a straight line (long-dashed line) joining landmarks L2 and L5. For reference, a short-dashed line is drawn parallel to the parasagittal plane through L5, and the angle between the short-dashed and long-dashed lines is therefore the pinna's angle of tilt. Bottom panel: $d_{B2\text{-}L4}$ is the lateral distance (parallel with the interaural axis) between landmarks B2 and L4. For reference, two mutually orthogonal, short-dashed lines are drawn parallel to the parasagittal plane through L4, to indicate the separation between B2 and L4 along the up-down and front-back axes.

Concha height H was measured as the distance between landmarks R2 and R5.

The concha aperture was also represented in more detail by a six-sided polygon with vertices at landmarks {R1, R2, L4, R3, R4, R5}. The total area $A_{aperture}$ of the polygon was calculated as the sum of the areas of the four constituent triangles with vertices {R1, R2, L4}, {R1, L4, R3}, {R1, R3, R4}, and {R1, R4, R5}. With a view toward practical implementation, the aperture area was also approximated as might be measured on a side-view photograph of the pinna: Area $A_{aperture\text{-}2D}$ was calculated by ignoring any differences in the coordinates of these six landmarks parallel to the interaural axis (i.e., by setting to a common value their coordinates along the lateral dimension). Interestingly, the two sets of areas $A_{aperture}$ and $A_{aperture\text{-}2D}$ were found to be highly correlated ($r = 0.99$), suggesting that concha aperture area can be sufficiently accurately approximated on a lateral-view photograph of the pinna; therefore, only the more practical measurement $A_{aperture\text{-}2D}$ was retained here. For consistency with the units of all other pinna measurements, aperture area was transformed to the diameter $D = 2 \times \text{sqrt}\{A_{aperture\text{-}2D}/\pi\}$ of an equivalent circular aperture.

In summary, 25 anthropometric distances were considered in this study: Concha depth d_{CIPIC}, the 21 concha depths $\{d_{B1\text{-}L1} — d_{B1\text{-}L7}\}$, $\{d_{B2\text{-}L1} — d_{B2\text{-}L7}\}$, and $\{d_{B3\text{-}L1} — d_{B3\text{-}L7}\}$, concha aperture width W, height H, and equivalent diameter D. Among the 38 pinnae in this study, there was no case where a landmark was not able to be identified by following the descriptions detailed here and summarized in Table I.

C. Acoustic simulation

With the head volumes downsampled to resolution 2.0 mm, acoustic pressure responses to a broadband Gaussian pulse were calculated for each pinna by an FDTD method (Takemoto *et al.*, 2010; Mokhtari *et al.*, 2011) incorporating an optimal perfectly matched layer (Mokhtari *et al.*, 2010a) designed to minimize artifactual reflections from the computation domain boundaries. Conventionally, the response at each ear would be recorded with the source at one farfield position at a time; instead for computational efficiency, the acoustic reciprocity theorem (e.g., Pierce, 1989) was employed by placing the source adjacent to the center of the (blocked) ear-canal entrance and calculating the response at all farfield observation points in a single run. Moreover, rather than extend the 3D spatial grid all the way to the farfield, for greater efficiency, the computation domain was restricted to just 3 cm from the head on five sides (the neck being partially immersed within the perfectly matched layer on the lower side), and 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 50 spatial locations in the median plane: In accord with the CIPIC database (Algazi *et al.*, 2001), from 45° below the horizontal plane at front, to 50.625° below the horizontal plane behind, in steps of 5.625°. Free-field responses at all 50 locations were also calculated with the source positioned at the center of the (absent) head. Left and right pinna HRTFs were then obtained at each spatial location from these pressure waveforms by Fourier transformation followed by free-field normalization.

TABLE I. List of pinna landmarks and their anatomical definitions.

Pinna landmark	Anatomical description
B1	Entrance plane of ear-canal
B2	Most posterior part of cavum concha base (posterior to ear-canal entrance and inferior to crus of helix)
B3	Base of cymba
BC	Deepest part of cavum concha when viewed normal to concha aperture, i.e., normal to line joining L2 and L5 (the "C" of BC refers to the CIPIC database, where this type of landmark was proposed)
L1	Side of cheek close to supratragic (or anterior) notch
L2	Most lateral surface of tragus
L3	Side of jaw close to intertragic notch
L4	Most lateral surface of antitragus
L5	Most lateral surface of antihelix
L6	Posterior corner of triangular fossa
L7	Least lateral point along lower crus of antihelix
R1	Anterior-most point along rim of supratragic (or anterior) notch
R2	Rim of intertragic notch
R3	Most posterior-inferior corner along rim of antihelix
R4	Rim of antihelix close to posterior corner of triangular fossa (L6)
R5	Superior rim of cymba

These simulation methods were previously evaluated favorably against independent acoustic measurements of a manikin's HRTFs; within the target frequency range 0.5–14 kHz a good spectral match was obtained (Mokhtari *et al.*, 2011). Moreover basically the same simulation methods were used by Takemoto *et al.* (2010), who reported a good spectral match with measured acoustic transfer functions of the human vocal tract up to 10 kHz. The focus of the present study—the first resonance of the human pinna, which is normally in the range 4–5 kHz—is therefore well within the frequency range validated in those two earlier studies.

D. Extraction of individual F_1 and A_1

As HRTF peaks are relatively constant in frequency while the notch frequencies vary with elevation angle in the median plane (Takemoto *et al.*, 2012), a representative value for the peaks of each pinna can be obtained on the mean transfer function, calculated by averaging the HRTFs across all 50 spatial locations in the median plane. In this way, the center-frequency F_1 and amplitude A_1 of the first peak were measured on each pinna's mean transfer function (an example is shown in Fig. 4). Compared with the alternative of measuring the peak at only one particular spatial location (e.g., the front direction as measured by Ishii and Iida, 2013), the spatial averaging method provides a more robust measurement of the first peak, and a degree of immunity against acoustic variations that may be caused by inter-individual inconsistencies in head alignment (i.e., pitch angle). Henceforth in this paper, F_1 and A_1 extracted from simulated HRTFs will be referred to as *measured* acoustic parameters (not to be confused with acoustically measured HRTFs) and compared with their values *estimated* from pinna anthropometry.

E. Linear regression modeling

The principal aim of this study was to find a practical solution for accurate estimation of F_1 and A_1 from pinna anthropometry. For this purpose, linear regression was used as a convenient computational tool for learning to map from the physiological to the acoustical domain. Due to the emphasis on practicality, only the best-performing *single* parameter, or at most the best *pair* of parameters, was sought with simple and multiple regression models, respectively (although the improvement gained by including a third parameter was also briefly evaluated). In view of this emphasis, and also due to the relatively small sample size (38 pinnae) compared with the total number of parameters (25 anthropometric distances), more sophisticated procedures such as the backward elimination or the stepwise method were abandoned in favor of the simpler *forward selection* of parameters. Also in this vein, the criteria used to evaluate regression models and to select or recruit parameters were simply the coefficient of correlation (r) and the mean absolute error (MAE) between original and estimated values of the dependent variable (either F_1 or A_1).

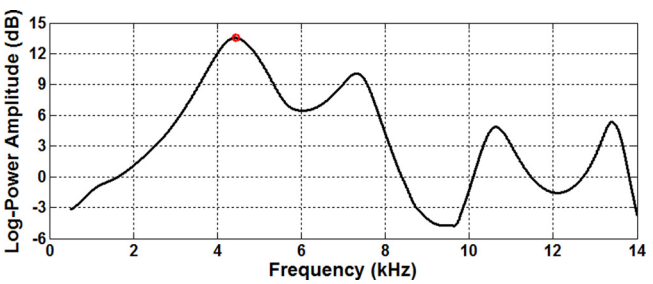

FIG. 4. (Color online) An example of a mean transfer function (averaged over 50 HRTFs in the median plane) with the center-frequency and amplitude of the first peak as indicated (small circle at $F_1 = 4.4$ kHz and $A_1 = 13.5$ dB).

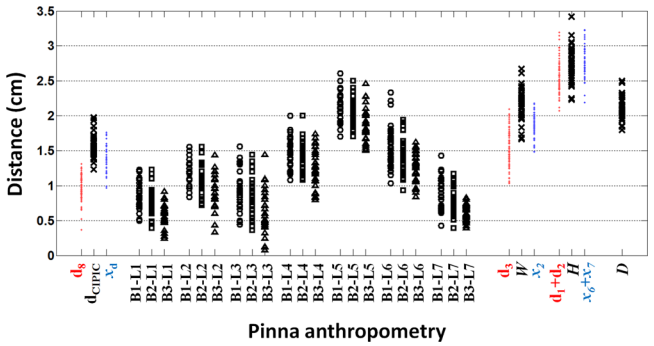

FIG. 5. (Color online) Anthropometric distances measured on all 38 pinnae (large symbols) and a selection of distances from two other databases (small dots). The abscissa indicates the type of distance (e.g., concha depth d_{CIPIC}, concha aperture width W, height H, and equivalent diameter D) as described in Sec. II B. For each lateral landmark L1–L7, the three adjacent vertical columns show lateral distances from base landmark B1 (circles on the left), B2 (squares in the middle), and B3 (triangles on the right), e.g., "L1–B1" refers to distance $d_{B1\text{-}L1}$; refer back to Fig. 2 for the typical position of all anatomical landmarks. For d_{CIPIC}, W, and H (our data shown by crosses), equivalent distances available in two other databases are shown for comparison (small dots): Respectively, "d_8," "d_3," and "d_1"+"d_2" for 74 pinnae of the CIPIC database (http://interface.cipic.ucdavis.edu/sound/hrtf.html) are shown by dots on the left side of our data; and, respectively, "x_d," "x_2," and "x_6"+"x_7" for 54 pinnae listed in Iida *et al.* (2014) are shown by dots on the right side of our data.

III. RESULTS

A. Anthropometric measurements

In Fig. 5, the large symbols show the 25 sets of anthropometric distances measured on all 38 pinnae: 22 concha depths d_{CIPIC}, $\{d_{B1\text{-}L1}–d_{B1\text{-}L7}\}$, $\{d_{B2\text{-}L1}–d_{B2\text{-}L7}\}$, and $\{d_{B3\text{-}L1}–d_{B3\text{-}L7}\}$, concha aperture width W, height H, and equivalent diameter D, as detailed earlier in Sec. II B. On average, each set of distances had a range of about 1 cm, which reflects mainly inter-individual differences in pinna geometry. The longest individual concha depth (2.61 cm) was measured by the distance $d_{B1\text{-}L5}$, while the shortest measure of concha depth (0.07 cm) was given by $d_{B3\text{-}L3}$. As exemplified by the latter case, which represents the lateral distance from the deepest part of the cymba (B3) to the side of the jaw close to the intertragic notch (L3), the 22 types of depth measurement considered here may not all be suitable or even practical. However, for the sake of completeness, they will all be evaluated with regression models to find the best type of distance for estimating F_1 and A_1.

For three of the sets of distances as shown by crosses in Fig. 5 (d_{CIPIC}, W, and H), corresponding pinna anthropometry available in the CIPIC database (Algazi *et al.*, 2001; http://interface.cipic.ucdavis.edu/sound/hrtf.html) are shown by dots on the left side of our data (respectively, "d_8," "d_3," and "d_1"+"d_2"), and those listed in Iida *et al.* (2014) are shown by dots on the right side (respectively, "x_d," "x_2," and "x_6"+"x_7"). The range, mean, and standard deviation of each of these sets of measurements are listed in Table II. Compared with each of those two previous datasets, as a group the participants in the present study showed on average a deeper concha as evaluated by d_{CIPIC} (60% deeper and 13% deeper, respectively), a wider concha as evaluated by W (39% wider and 15% wider, respectively), and a comparable concha height as evaluated by H (7% longer and 1% shorter, respectively). The greater similarity with the pinna measurements of Iida *et al.* (2014) may possibly reflect the common inclusion of mostly Japanese participants, but the small sample sizes preclude any further discussion on racial tendencies in pinna geometry.

B. Distributions of F_1 and A_1

A scatter-plot of center-frequency F_1 (in kHz) versus log-power amplitude A_1 (in dB) of the first peak extracted from the mean transfer functions of all 38 pinnae, is shown in Fig. 6. For these data, F_1 ranged over 1 kHz (3.9–4.9 kHz or 0.33 octave, mean 4.5 kHz, s.d. 274 Hz), and A_1 ranged over 7 dB (7.3–14.3 dB, mean 10.9 dB, s.d. 1.8 dB). While these ranges of variation were mostly due to individual differences between participants (inter-person s.d. 254 Hz in F_1 and 1.7 dB in A_1), the data also revealed that differences between the left and right pinnae of each participant were not small: Mean inter-pinna absolute difference of 175 Hz or 0.06 octave (about [3/4] of a semitone) in F_1, and 1.1 dB in A_1; statistically significant as indicated by paired-samples Student's $t = 5.77$ at $p < 0.001$ for F_1, and $t = 7.79$ at $p < 0.001$ for A_1.

Interestingly, across all 38 pinnae F_1 and A_1 were found to be independent ($r = 0.02$), which already suggests that the regression models best suited to estimate F_1 and A_1 likely involve two independent sets of pinna measurements.

C. Estimation of F_1 from pinna anthropometry

A comparison of the performance of the 25 regression models is shown in Fig. 7, where the height of each bar

TABLE II. The range, mean, and standard deviation of three sets of pinna anthropometric distances as measured on 38 pinnae in the present study, 74 pinnae in the CIPIC database (Algazi *et al.*, 2001; http://interface.cipic.ucdavis.edu/sound/hrtf.html), and 54 pinnae in Iida *et al.* (2014).

		Range (cm)	Mean (cm)	s.d. (cm)
This study	d_{CIPIC}	1.23–1.98	1.57	0.17
	W	1.67–2.67	2.15	0.22
	H	2.23–3.42	2.75	0.24
CIPIC data (Algazi *et al.*, 2001)	d_8	0.37–1.31	0.98	0.18
	d_3	1.04–2.10	1.55	0.26
	$d_1 + d_2$	2.07–3.19	2.57	0.23
Iida *et al.* (2014)	x_d	0.97–1.76	1.39	0.18
	x_2	1.48–2.18	1.87	0.18
	$x_6 + x_7$	2.19–3.23	2.79	0.22

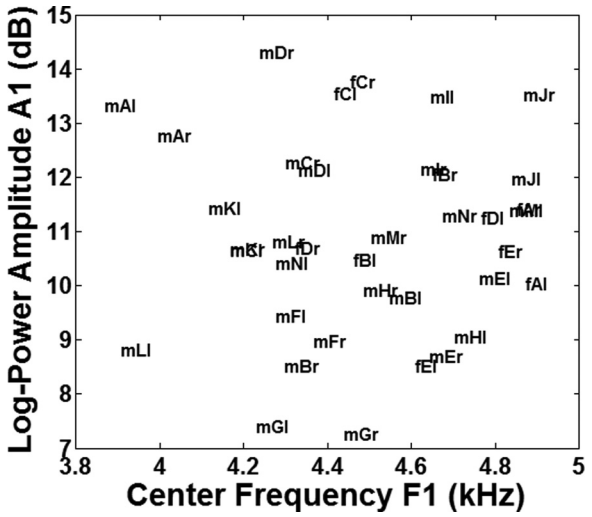

FIG. 6. Scatter-plot of center-frequency F_1 versus log-power amplitude A_1 extracted from the mean transfer functions of all 38 pinnae. Each pinna is represented by a three-letter label. First letter: m, male; f, female. Second letter: Individual code A, B,… separately for males and females. Third letter: l, left pinna; r, right pinna.

indicates the coefficient of correlation (r) between original and estimated values of F_1. As the mean absolute error (MAE) in F_1 revealed essentially the same comparison among models, only the correlation results are shown here.

As d_{CIPIC} is the only type of concha depth measurement used previously for F_1 estimation, it is regarded here as a baseline and shown in Fig. 7 by the first bar on the left and a heavy dashed line for ease of comparison with other model results. This baseline model yielded the following relation (estimated F_1 in Hz and d_{CIPIC} in cm):

$$\hat{F}_1 = 5894 - 887 d_{CIPIC}, \quad (1)$$

with $r = 0.55$ and MAE = 175 Hz (equal to the mean inter-pinna difference in F_1). These results are remarkably similar to those reported by Ishii and Iida (2013), who used the

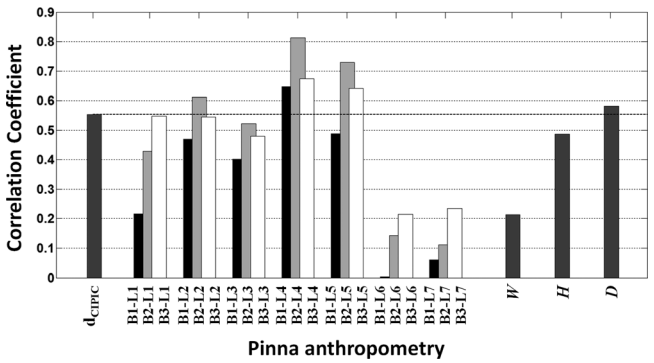

FIG. 7. Performance evaluation of 25 simple linear regression models in estimating F_1 from pinna anthropometry. The ordinate shows correlation coefficient (r) between original and estimated values of F_1. The abscissa indicates the anthropometric parameter used in each model: The baseline model (gray bar at left, and heavy dashed horizontal line) refers to distance d_{CIPIC}; for each lateral landmark L1–L7, the three models refer to lateral distances from base landmark B1 (black bar), B2 (light-gray bar), and B3 (white bar), therefore the three bars in each group correspond to the three columns within each group of distances in Fig. 5; the last three models (gray bars at right) refer to concha aperture width W, height H, and diameter D.

same type of concha depth measurement and two additional parameters, as reviewed in Sec. I ($r = 0.56$, $MAE = 160$ Hz). The negative sign on the coefficient on d_{CIPIC} in Eq. (1) confirms the basic physical principle that a greater concha depth supports a longer resonance wavelength and therefore a lower resonance frequency.

Of the remaining 24 anthropometric parameters considered here, seven yielded models that were either comparable to or outperformed the baseline for F_1 estimation. Interestingly, the concha aperture diameter D (in cm) performed comparably with the baseline, yielding the relation

$$\hat{F}_1 = 6498 - 953D, \quad (2)$$

with a marginally better correlation coefficient ($r = 0.58$) but a larger spread (MAE = 186 Hz). The negative sign on the coefficient on D in Eq. (2) implies that a larger concha aperture leads to a lower resonance frequency. This is consistent with the principle of open-end correction (e.g., Teranishi and Shaw, 1968), which states that the acoustical depth of a resonant tube or cavity is longer than its physical depth by an amount that is proportional to its aperture radius (or diameter) with the constant of proportionality being dependent on the effective amount of baffle surrounding the cavity rim. These results suggest that for the purpose of F_1 estimation, baseline concha depth d_{CIPIC} may be replaced by concha aperture diameter D which in principle can be obtained more easily (e.g., from a side-view photograph of the pinna, as shown later in Fig. 11), with comparable performance. However, as discussed next, both D and d_{CIPIC} were outperformed by alternative measurements of concha depth.

As shown in Fig. 7, for the lateral landmarks L2–L5, which were distributed around the lower parts of the concha (i.e., surrounding the cavum concha; see Fig. 2), distances to base landmark B2 (light-gray bars) consistently outperformed those to base landmark B1 (black bars) or B3 (white bars). In contrast, for lateral landmarks L1, L6, and L7, which were located more superiorly (i.e., closer to the cymba; see Fig. 2), distances to the superior base landmark B3 (cymba) performed the best. These results indicate that for F_1 estimation, lateral distances from the cavum concha or cymba base are best defined relative to lateral landmarks that are close to, or surround, each cavity; i.e., when the cavum concha base (B2) is paired with lower lateral landmarks (L2–L5) and when the cymba base (B3) is paired with higher lateral landmarks (L1, L6, and L7).

However, among all seven lateral landmarks, the best performance was obtained consistently with distances to L4, the landmark defining the lateral surface of the antitragus. Overall, the best single-parameter model for estimating F_1 was obtained with distance $d_{B2\text{-}L4}$ (in cm):

$$\hat{F}_1 = 5939 - 1030 d_{B2-L4}, \quad (3)$$

which yielded $r = 0.81$ and MAE = 125 Hz (a reduction of 29% compared with the baseline error and less than half of 1 s.d. in the original distribution of F_1). As in Eq. (1), the coefficient on the depth parameter in Eq. (3) has a negative sign, owing to the inverse relation between wavelength and

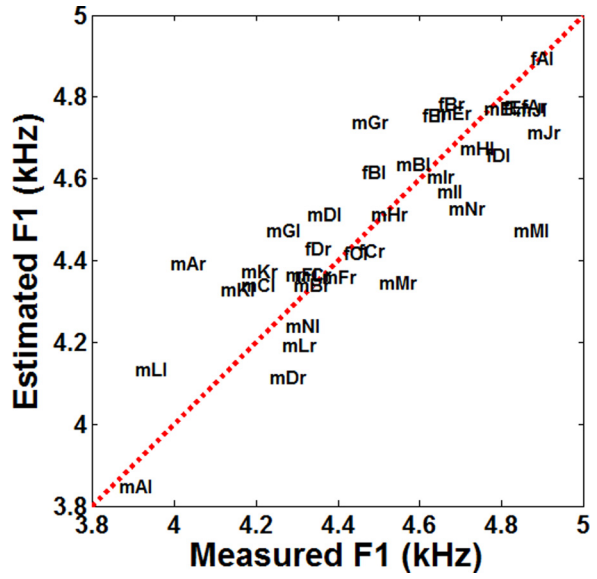

FIG. 8. (Color online) Scatter-plot of measured versus estimated F_1, for the linear regression model with two parameters $d_{B2\text{-}L4}$ and $d_{B2\text{-}L5}$, as given in Eq. (4) that yielded $r = 0.84$ and MAE = 118 Hz. Each pinna is denoted by a three-letter label as described in the caption of Fig. 6. The dashed line represents the ideal performance.

frequency of resonance. These results indicate that the posterior part of the cavum concha base (B2) and the outer surface of the antitragus (L4) provide anatomical landmarks that are most effective for F_1 estimation. In practical terms, it is fortuitous that distance $d_{B2\text{-}L4}$ is quite easy to measure on real human pinnae, as discussed later in Sec. IV.

It is reasonable to investigate next which of the remaining anthropometric parameters might improve the model in Eq. (3). To answer this question, multiple linear regression models were trained and evaluated on two parameters at a time: $d_{B2\text{-}L4}$ and each of the remaining 24 parameters in turn. Consistent with the results in Fig. 7 where $d_{B2\text{-}L5}$ was the second-best performer among the single-parameter models, the largest improvement was obtained by including distance $d_{B2\text{-}L5}$ (in cm):

$$\hat{F}_1 = 6461 - 758 d_{B2-L4} - 439 d_{B2-L5}, \quad (4)$$

with $r = 0.84$ and MAE = 118 Hz (i.e., a reduction of 6% compared with the Eq. (3) model error). These results indicate a modest improvement in F_1 estimation by a weighted combination of the lateral distances from B2 to the outer surface of the antitragus (L4) and to the most lateral surface of the antihelix (L5). It is noteworthy that across the 38 pinnae in the present study, these two distances were found to be mildly co-dependent ($r = 0.72$), which is perhaps not surprising in view of the anatomical proximity and connection between landmarks L4 and L5, from antihelix to antitragus (cf. Fig. 2); nevertheless, the improvement gained by the two-parameter model indicates that the two distances also carried some complementary and independent information. The fact that in Eq. (4) the distance $d_{B2\text{-}L4}$ is weighted preferentially over $d_{B2\text{-}L5}$ by a ratio approximately 1.7:1 indicates that the antitragus landmark is about twice as important or influential for F_1 estimation. A scatter-plot comparing the measured values of F_1 with those estimated by Eq. (4) is shown in Fig. 8.

Based on the open-end correction principle, it may be reasonable to expect that the F_1 estimation models in Eqs. (3) and (4) would be improved by recruiting aperture diameter D as an additional parameter. However, in both cases, only small improvements were obtained: The two-parameter model for $d_{B2\text{-}L4}$ and D yielded $r = 0.83$ and MAE = 122 Hz, and the three-parameter model for $d_{B2\text{-}L4}$, $d_{B2\text{-}L5}$ and D yielded $r = 0.85$ and MAE = 116 Hz. The lower than expected improvements may possibly be due to pinna-to-pinna variations in the effective amount of baffle surrounding each individual concha, which would have caused large uncertainty in the constant factor on D; or, more simply, there may have been insufficient accuracy in approximating the concha aperture by a six-sided polygon as described in Sec. II.

Recruitment of additional parameters will not be detailed here, as further improvements to F_1 estimation performance as indicated by r and *MAE* were contradictory (i.e., highest $r = 0.85$ for the three-parameter model which included $d_{B1\text{-}L2}$, or lowest MAE = 107 Hz for the three-parameter model that included $d_{B1\text{-}L6}$), and the practical complexity of additional measurements was therefore not justified.

D. Estimation of A_1 from pinna anthropometry

While pinna anthropometric relations with F_1 were well-motivated on the basis of previous studies, this is not the case with amplitude A_1 the relations of which with pinna geometry are largely undocumented. Linear regression models were therefore evaluated for estimating A_1 from each of the 25 anthropometric parameters. A comparison of the performance of the 25 models in terms of the coefficient of correlation (r) between original and estimated values of A_1, is shown in Fig. 9. As the MAE in A_1 revealed essentially the same comparison among models, only the correlation results are shown here.

As already foreshadowed by the independence of F_1 and A_1 (Sec. III B), in contrast with the results for F_1, the A_1 estimation model for d_{CIPIC} performed poorly ($r = 0.08$) and could therefore not be regarded as a baseline. Furthermore, the models based on distances from base landmark B2 (light-gray bars in Fig. 9) were never the best; indeed, for A_1 estimation, distances from B1 (black bars) consistently outperformed the other two base landmarks. These results suggest that lateral distances measured from the deepest part of the cavum concha at the plane of the ear-canal entrance, rather than the posterior base of the cavum concha or the base of the cymba, are more relevant to A_1.

Also in contrast with the results for F_1, Fig. 9 indicates that for A_1 estimation with distances to base landmark B1 (black bars) and B2 (light-gray bars), lateral landmarks L4 and L5 were the least successful, while L1 yielded the best results. Overall, the best single-parameter model for estimating A_1 (in dB) was obtained with distance $d_{B1\text{-}L1}$ (in cm):

$$\hat{A}_1 = 4.32 + 7.50 d_{B1-L1}, \quad (5)$$

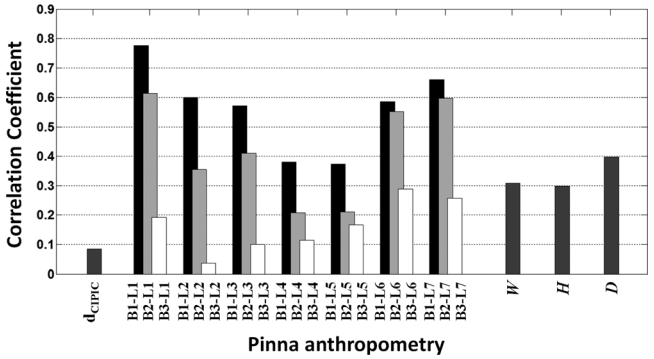

FIG. 9. Performance evaluation of 25 simple linear regression models in estimating A_1 from pinna anthropometry. The ordinate shows correlation coefficient (r) between original and estimated values of A_1. The abscissa indicates the anthropometric parameter used in each model as explained in the caption of Fig. 7.

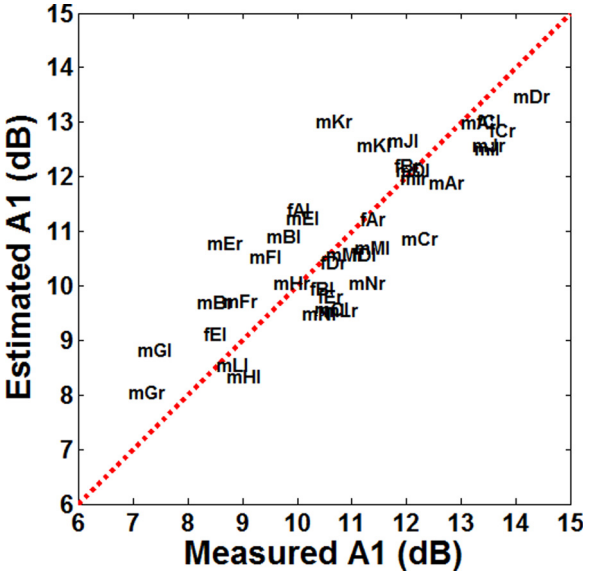

FIG. 10. (Color online) Scatter-plot of measured versus estimated A_1, for the linear regression model with two parameters $d_{B1\text{-}L1}$ and D, as given in Eq. (6), which yielded $r = 0.83$ and MAE = 0.84 dB. Each pinna is denoted by a three-letter label as described in the caption of Fig. 6. The dashed line represents the ideal performance.

which yielded $r = 0.78$ and MAE = 0.94 dB (smaller than the mean inter-pinna difference in A_1). Thus the best estimate of the log-power amplitude of the first peak was provided by the lateral distance between the side of the cheek close to the anterior notch (L1) and the plane of the ear-canal entrance (B1). The coefficient on $d_{B1\text{-}L1}$ in Eq. (5) has a positive sign, meaning that A_1 increases with this measure of concha depth; implications of this positive relation will be discussed in Sec. IV.

Although concha aperture diameter D by itself did not yield a good model for A_1 estimation ($r = 0.40$, cf. the rightmost bar in Fig. 9), inclusion of D (in cm) as a second parameter in multiple regression yielded the largest improvement over the model in Eq. (5),

$$\hat{A}_1 = 11.35 + 7.12 d_{B1-L1} - 3.19D, \quad (6)$$

with $r = 0.83$ and MAE = 0.84 dB (i.e., less than half of 1 s.d. in the original distribution of A_1). In Eq. (6), the magnitude of the coefficient on $d_{B1\text{-}L1}$ is about twice that on D; this implies that in this two-parameter model for A_1 estimation, the concha depth is approximately as important as the concha aperture radius. Interestingly, across all 38 pinnae, the two parameters $d_{B1\text{-}L1}$ and D were independent ($r = -0.13$). The negative sign on the D term in Eq. (6) implies an inverse relation between A_1 and concha aperture; this will be discussed further in Sec. IV. A scatter-plot comparing the measured values of A_1 with those estimated by Eq. (6) is shown in Fig. 10.

Recruitment of additional parameters will not be detailed here as further improvements to A_1 estimation performance as indicated by r and MAE were contradictory (i.e., highest $r = 0.87$ for the three-parameter model which included $d_{B3\text{-}L5}$ or lowest MAE = 0.71 dB for the three-parameter model which included $d_{B1\text{-}L4}$), and the practical difficulty of additional measurements was therefore not justified.

IV. DISCUSSION

A. Interpretation of models for F_1 estimation

In this study, linear regression was used to evaluate several candidate measurements of concha aperture and depth (including one measure of concha depth as suggested in previous literature), all presumed to be related with the *center-frequency* F_1 of the first peak of HRTFs. As shown in Fig. 7 and exemplified in the bottom panel of Fig. 3, F_1 was best estimated by concha depth as measured laterally from the posterior part of the cavum concha base (landmark B2) to the outer surface of the antitragus (landmark L4); F_1 estimation accuracy was somewhat improved by including the lateral distance from B2 to the most lateral point of the antihelix (landmark L5). Thus if it can be assumed that approximately one-quarter of a wavelength extends outward from the concha base to the concha rim and surrounding air as presumed in previous studies, then the present results indicated that pinna landmarks L4 and L5 were the best for defining the physical outer limit of the concha commonly across several individual pinnae in terms of the extent to which it supports an F_1 quarter-wavelength resonance. Indeed, as shown in Figs. 2 and 5, these best landmarks were among the laterally furthest from the concha base. Our results therefore suggest that the F_1 quarter-wavelength is likely supported physically at least as far as the *longest* lateral extent of the concha, as represented by landmarks L4 and L5 (and possibly even further, due to end-correction effects).

Moreover, the results in Fig. 7 indicated that the best models for F_1 estimation involved depth measurements from base landmark B2. A possible explanation of the preference for B2 over B1 and B3 may be found in previous studies that have shown that F_1 is maximally sensitive to physical perturbations of the concha base (Mokhtari *et al.*, 2010b; Mokhtari *et al.*, 2013): Owing to the pressure anti-node throughout the concha base, any protrusion of the base surface (even a local protrusion by only one voxel element of side 0.2 cm) tends to shift F_1 upward in frequency. Already, compared with a

simplified acoustic tube having a flat base, the concha base varies remarkably in lateral position over its entire surface, especially considering the lower part of the crus of helix that forms a raised ridge that separates cavum concha from cymba. The greater accuracy in F_1 estimation afforded by B2 may therefore suggest that this landmark best represented a kind of average lateral position over the entire base surface, across a group of individuals. This interpretation will need to be tested in future work, perhaps with the help of sensitivity analyses (Mokhtari *et al.*, 2013) and visualizations of pressure and velocity anti-nodes (Mokhtari *et al.*, 2014) associated with the first normal mode.

B. Interpretation of models for A_1 estimation

Regression models were also used to evaluate the same set of candidate measurements in estimation of the log-power *amplitude* A_1 of the first peak of HRTFs. Consistent with the observation that A_1 and F_1 were independent, the results in Fig. 9 indicated that A_1 was best estimated by a different set of pinna measurements: Concha depth as measured from the plane of the ear-canal entrance (landmark B1) to the side of the cheek (landmark L1), and concha aperture as approximated by diameter D. As shown in Figs. 2 and 5, pinna landmark L1 defined the concha baffle at a location that was among the laterally closest to the ear-canal entrance and hence the position of the acoustic point source. Therefore whereas the *longest* depth of the concha was found to be the main predictor of resonance frequency F_1, the main predictors of resonance amplitude A_1 appeared to be (a) the *shortest* depth of the cavum concha and (b) the concha aperture.

Both these factors may be understood in terms of basic physical principles involving the relative storage and loss of acoustic resonance energy. Generally, the greater the amount of energy stored at each resonance cycle, the stronger the resonance and the higher its amplitude; inversely, the greater the proportion of energy that leaks out of a resonator at each cycle, the weaker the resonance and the lower its amplitude. In the case of the human pinna, our results therefore suggest the following explanations: (a) While approximately a quarter of a wavelength is supported at least as far as the antitragus and antihelix, the side of the cheek close to the anterior notch likely represents the opening or passage of escape of acoustic energy to the surrounding air, nearest to the pressure anti-node in the cavum concha base; therefore the portion of the cavum concha that is indicated by this shortest depth likely represents the part of the cavity that is available for storing energy at every cycle of resonance, so that a shorter depth results in a lower A_1 due to less energy retained and more energy lost to the surrounding air at each cycle; and (b) the concha aperture likely represents the total surface over which the first pinna resonance is coupled to the surrounding air, so that a larger aperture results in a lower A_1 due to greater loss of energy at each resonance cycle. In fact, our results concerning A_1 are also in agreement with elementary acoustical theory of a simple cylindrical resonator for which there is an approximately linear relation between the peak pressure gain and the *ratio* of tube depth and tube radius (e.g., Teranishi and Shaw, 1968, Fig. 3); indeed, we confirmed that the performance of the model in Eq. (6) was equally attained by a simple linear regression on the ratio $d_{B1\text{-}L1} / D$.

C. Independence of best anthropometric parameters

Individual differences in pinna geometry notwithstanding, it is reasonable to question whether human physiological constraints impose a degree of dependence between the two pairs of anthropometric measurements that were found to be the best in estimating F_1 ($d_{B2\text{-}L4}$ and $d_{B2\text{-}L5}$) and A_1 ($d_{B1\text{-}L1}$ and D). As indicated in Table III, cross-correlations among these parameters did not reveal any strong dependence. The two largest values ($r = 0.56$ and $r = 0.52$) indicated a moderate co-variation between concha aperture diameter D and F_1-related concha depth; this result is not surprising, for as discussed earlier in connection with Fig. 7, D performed comparably with the baseline in F_1 estimation. The remaining lower values in Table III ($r = 0.33$ and $r = 0.31$) indicated a positive but only weak dependence between the two F_1-related measures of concha depth and the A_1-related measure of cavum concha depth. These low correlations suggest that the present data must have included individual pinnae with various combinations of independently varying (either longer or shorter) F_1- and A_1-related depths; indeed, e.g., the ratio $d_{B2\text{-}L5}/d_{B1\text{-}L1}$ ranged widely, from 1.8 to 3.9 in the present data. Hence these two sets of parameters can be considered sufficiently independent for the purposes of estimating F_1 and A_1, notwithstanding the weak correlations which probably reflect natural constraints on human pinna physiology.

D. Superiority of lateral vs tilted concha depths

All 21 concha depth measurements considered here as alternatives to the baseline d_{CIPIC} were exclusively *lateral* depths, i.e., measured parallel to the interaural axis. On the other hand, across the 38 pinnae the angle of concha aperture tilt (as defined by the horizontal angle between a sagittal plane and a line joining landmarks L2 and L5) ranged from 17° to 35° (mean 23.5°, s.d. 4.3°); for comparison in the CIPIC database, pinna flare angle (relative to the side of the head) had a mean of 28.5° and a s.d. of 6.7° (Algazi *et al.*, 2001). If the concha could be regarded simply as a flat-based acoustic tube rotated about the vertical axis according to individual angle of tilt, then theoretically, depths measured perpendicular to the concha aperture (and therefore not necessarily purely lateral) should be physically and acoustically more appropriate.

TABLE III. Coefficients of correlation (r), calculated across the 38 pinnae, between pairs of anthropometry parameters in two sets: Those that were found most suitable for F_1 estimation (two rows) and those most suitable for A_1 estimation (two columns).

	$d_{B1\text{-}L1}$	D
$d_{B2\text{-}L4}$	0.33	0.56
$d_{B2\text{-}L5}$	0.31	0.52

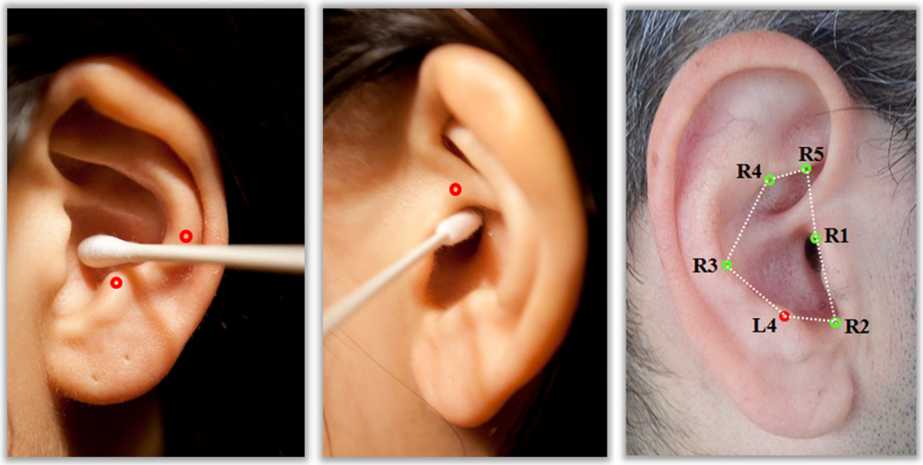

FIG. 11. (Color online) Practical measurements of concha depth and aperture suggested here for F_1 and A_1 estimation. Left panel: F_1-related concha depths $d_{B2\text{-}L4}$ and $d_{B2\text{-}L5}$ (lateral landmarks L4 and L5 indicated by small circles). Middle panel: A_1-related cavum concha depth $d_{B1\text{-}L1}$ (lateral landmark L1 indicated by a small circle). Right panel: Approximation of concha aperture area by a six-sided polygon marked on a side-view photograph.

To investigate this issue, all concha depths were re-measured perpendicular to individual tilt angle, and all the regression models were re-evaluated with these *perpendicular* (or *tilted* with respect to the interaural axis) depths. In summary, the main trends of comparative results reported in Figs. 7 and 9 were replicated but with overall lower levels of performance: e.g., for F_1, the best model was yielded by the perpendicular version of $d_{B2\text{-}L4}$, with $r = 0.66$ and MAE = 162 Hz [i.e., worse than the Eq. (3) model]; and for A_1, the best model was yielded by the perpendicular version of $d_{B1\text{-}L1}$, with $r = 0.60$ and MAE = 1.2 dB [i.e., worse than the Eq. (5) model]. We thereby confirmed that the proposed *lateral* depth measurements were superior, for both F_1 and A_1 estimation from pinna anthropometry. However, this should not be taken to imply that the first normal mode of every pinna resonates strictly laterally (i.e., parallel to the interaural axis)—in view of the complicated shape of the human concha, this would be a naive misinterpretation. The acoustical and physical reasons for the superiority of lateral depths should be clarified in future work.

E. Practical simplicity of pinna measurements

Fortunately, lateral concha depths are also relatively simple to measure on human pinnae. The practical simplicity of the suggested pinna anthropometry is shown by the photographs in Fig. 11. A soft but rigid tool such as a cotton bud can be used in both cases: To measure the lateral distances from the cavum concha base to the antitragus and antihelix for F_1 estimation, by resting the tip lightly against the concha base (left panel of Fig. 11), and to measure the lateral distance from the ear-canal entrance to the side of the cheek for A_1 estimation, by resting the tool lightly against the anterior notch and extending it inward as far as the ear-canal entrance (middle panel of Fig. 11). In both cases, the lateral position can be marked on the tool with a pen, and a small ruler can then be used to measure the distance from the tool tip to the mark. Concha aperture can be approximated on a lateral-view photograph of the pinna: In the particular case shown in the right panel of Fig. 11 (angle of concha tilt 22°), measurements on the 3D rendered model yielded $A_{aperture} = 4.36\,cm^2$ (i.e., an effective diameter 2.36 cm), while the photographic measurements yielded $A_{aperture\text{-}2D} = 4.12\,cm^2$ and $D = 2.29$ cm, a difference of only 3% in diameter compared with the 3D measurement.

A remaining question in regard to practical measurement of the concha aperture is whether D can be replaced by a simpler approximation such as sqrt$\{W \times H\}$, which assumes either a rectangular or, to within a constant factor, an elliptical aperture, and which therefore requires measurement of only the total width and height of the entire concha. In line with expectations, regression models evaluated with the simpler measure always yielded poorer performance than with D: For F_1 estimation, sqrt$\{W \times H\}$ yielded $r = 0.40$ and MAE = 206 Hz (compared with $r = 0.58$ and MAE = 186 Hz for D); for A_1 estimation with only one parameter, sqrt$\{W \times H\}$ yielded $r = 0.37$ and MAE = 1.33 dB (compared with $r = 0.40$ and MAE = 1.27 dB for D); and for A_1 estimation with two parameters including $d_{B1\text{-}L1}$, sqrt$\{W \times H\}$ yielded $r = 0.81$ and MAE = 0.87 dB (compared with $r = 0.83$ and MAE = 0.84 dB for D). These results support the efficacy of D and the extra effort required to measure the six-sided polygon.

F. Expected errors in measurement and estimation

A simplified error analysis can be carried out on the suggested regression formulae by assuming that the distances $d_{B2\text{-}L4}$, $d_{B2\text{-}L5}$, $d_{B1\text{-}L1}$, and D can be measured in practice to the nearest millimeter (i.e., to within ±0.5 mm). Then by Eqs. (4) and (6), the largest expected measurement-related error in F_1 and in A_1 would be only 60 Hz and 0.5 dB, respectively. Conversely, if measurement-related errors are not to exceed the modeling errors (MAE = 118 Hz and MAE = 0.84 dB, respectively), then, assuming equal contributions to the estimation error from each term in Eqs. (4) and (6), $d_{B2\text{-}L4}$ would need to be measured to within ±0.8 mm, $d_{B2\text{-}L5}$ to within ±1.3 mm, $d_{B1\text{-}L1}$ to within ±0.6 mm, and D to within ±1.3 mm. A more complete evaluation of the estimation models with measurements on human pinnae, including measurements by more than one person to test consistency and repeatability, would be desirable in future work.

G. Concluding remarks

As stated earlier, the first peak of individual HRTFs has been implicated as a potential anchor in analyzing spectral cues for sound source localization (Iida *et al.*, 2007). It is also known that human auditory sensitivity to sound intensity (as measured in equal loudness contours) is largest in approximately the frequency range 2–5 kHz (e.g., Blauert, 1997, p. 120), i.e., the frequency range encompassing the first peak; thus, in addition to localization, the first peak may also affect the perceived naturalness of 3D sounds. Therefore accurate estimation of F_1 and A_1 from a few simple pinna measurements as suggested here, should be an important first step in either selecting or synthesizing a set of personalized HRTFs for 3D audio applications targeting individual listeners.

As the first normal mode is essentially a monopole with an approximately omni-directional directivity pattern, a single estimated value for F_1 and A_1 of each pinna as provided here based on the mean of median-plane transfer functions may be adequate for basic sound localization. Nevertheless, it remains for future work to determine (e.g., by psychoacoustic experiments) whether localization and perceived naturalness might be enhanced by additionally accounting for the natural variations in F_1 and A_1 across the full range of spatial locations around a listener's head, especially for A_1, which tends to have lower values on the contralateral side owing mainly to head-shadowing.

Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (**2001**). "The CIPIC HRTF database," in *Proceedings of the 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.

Ishii, Y., and Iida, K. (**2013**). "Estimation of spectral notch frequencies of the individual head-related transfer function from anthropometry of listener's pinna," in *Proceedings of the 21st International Congress on Acoustics (ICA)*, June 2–7, Montreal, Canada, pp. 1–7.

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

Mokhtari, P., Takemoto, H., Adachi, S., Nishimura, R., and Kato, H. (**2013**). "Three-dimensional acoustic sensitivity analysis of pinna geometry at peaks of head-related transfer functions," in *Proceedings of the Autumn Meeting of the Acoustical Society of Japan*, Toyohashi, Japan, Paper 3-1-17, pp. 861–864.

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 Lang. Process. **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.

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

Shaw, E. A. G., and Teranishi, R. (**1968**). "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," J. Acoust. Soc. Am. **44**(1), 240–249.

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., 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.

Teranishi, R., and Shaw, E. A. G. (**1968**). "External-ear acoustic models with simple geometry," J. Acoust. Soc. Am. **44**(1), 257–263.

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.