2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems

A Head-Related Transfer Function Model for Real-Time Customized 3-D Sound Rendering

Michele Geronazzo* Simone Spagnol† Federico Avanzini‡

Department of Information Engineering
Università degli Studi di Padova

ABSTRACT

This paper addresses the problem of modeling head-related transfer functions (HRTFs) for 3-D audio rendering in the front hemisphere. Following a structural approach, we build a model for real-time HRTF synthesis which allows to control separately the evolution of different acoustic phenomena such as head diffraction, ear resonances, and reflections through the design of distinct filter blocks. Parameters to be fed to the model are both derived from mean spectral features in a collection of measured HRTFs and anthropometric features of the specific subject (taken from a photograph of his/her outer ear), hence allowing model customization. Visual analysis of the synthesized HRTFs reveals a convincing correspondence between original and reconstructed spectral features in the chosen spatial range. Furthermore, a possible experimental setup for dynamic psychoacoustical evaluation of such model is depicted.

Index Terms: H.5.1 [Information Interfaces and Presentation (e.g., HCI)]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.5 [Information Interfaces and Presentation (e.g., HCI)]: Sound and Music Computing—Modeling

1 INTRODUCTION

Head-Related Transfer Functions (HRTFs) capture the transformations undergone by a sound wave in its path from the source to the eardrum, typically due to diffraction and reflections on the torso, head, shoulders and pinnae of the listener. Such characterization allows virtual positioning of sound sources in the surrounding space by filtering the corresponding signals through a pair of HRTFs, thus creating left and right ear signals to be delivered by headphones [6]. In this way, three-dimensional sound fields with a high immersion sense can be simulated and integrated in augmented and/or virtual reality contexts.

Alongside critical dependence on the relative position between listener and sound source, anthropometric features of the human body have a key role in HRTF characterization. While non-individualized HRTFs represent a cheap and straightforward mean of providing 3-D perception in headphone reproduction, listening to non-individualized spatialized sounds may likely result in evident sound localization errors such as incorrect perception of the source elevation, front-back reversals, and lack of externalization [8], especially in static conditions. On the other hand, individual HRTF measurements on a significant number of subjects may be both time- and resource-expensive.

Structural modeling of HRTFs ultimately represents an attractive solution to these shortcomings. As a matter of fact, if one isolates the contributions of the listener's head, pinnae, ear canals, shoulders, and torso to the HRTF in different subcomponents - each accounting for some well-defined physical phenomenon - then, thanks to linearity, he can reconstruct the global HRTF from a proper combination of all the considered effects. Relating each subcomponent's temporal and/or spectral features (in the form of digital filter parameters) to the corresponding anthropometric quantities would then yield a HRTF model which is both economical and individualizeable [5].

This paper focuses on an extension of one such model [11] that can be employed for immersive sound reproduction. The proposed approach allows for an interesting form of content adaptation and customization, since it includes parameters related to the user's anthropometry in addition to the spatial ones. Our approach has also implications in terms of delivery, since it operates by processing a monophonic signal exclusively at the receiver side (e.g., on a terminal or mobile device) by means of low-order filters, allowing for reduced computational costs. Thanks to its low complexity, the model can be used to render scenes with multiple audiovisual objects in a number of contexts such as computer games, cinema, edutainment, and any other scenario where realistic sound spatialization and personalized sound reproduction is a major requirement.

Following Section 2, in which we propose a possible parametrization of pinna-related HRTF features based on anthropometry, Section 3 includes a complete description of our structural model. An example of real-time system where the implemented model can be validated and merged into VR/AR contexts is sketched in Section 4.

2 PINNA-BASED CUSTOMIZATION

There is no doubt that, if we fix the source direction with respect to the listener, the greatest dissimilarities among different people's HRTFs are due to the massive subject-to-subject pinna shape variation. As a matter of fact, the pinna plays a primary role in determining the frequency content of the HRTF thanks to two primary acoustic phenomena,

1. reflections over pinna edges. According to Batteau [4], sound waves are typically reflected by the outer ear, as long as their wavelength is small enough compared to the pinna dimensions, and the interference between the direct and reflected waves causes sharp notches to appear in the high-frequency side of the received signal's spectrum;

2. resonant modes in pinna cavities. As Shaw argued [12], since the concha acts as a resonator some frequency bands of both the direct and reflected sound waves are significantly enhanced, depending on the elevation of the source.

Consequently, the part of the HRTF due to the pinna's contribution (commonly known as Pinna-Related Transfer Function, PRTF [1]) presents a sequence of peaks and notches in its magnitude. In order to isolate the spectral modifications due to reflections from those due to resonances, we implemented an algorithm (details of which can be found in [7]) that iteratively compensates the PRTF magnitude spectrum with an approximate multi-notch filter until no significant notches are left, so that, once convergence is reached at

*e-mail:geronazzo@dei.unipd.it
†e-mail:spagnols@dei.unipd.it
‡e-mail:avanzini@dei.unipd.it

978-0-7695-4635-3/11 $26.00 © 2011 IEEE
DOI 10.1109/SITIS.2011.21

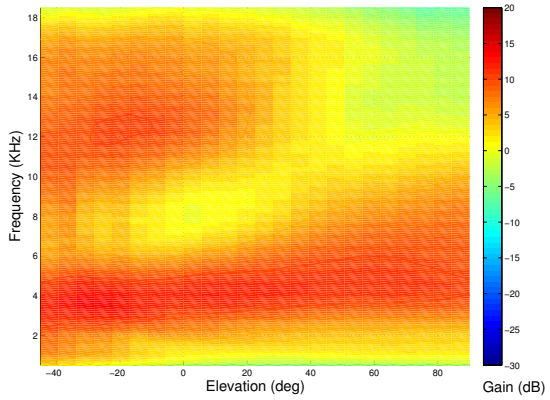


Figure 1: *Mean magnitude spectrum of the pinna resonant component, averaged on all* 45 *CIPIC subjects' left-ear responses.*

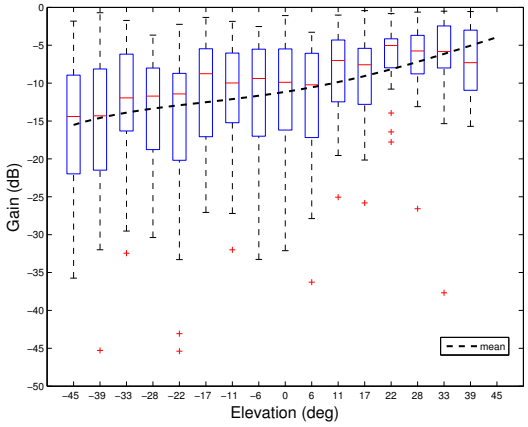


Figure 2: *Box plot and mean of the first notch track's gain values among CIPIC subjects.*

iteration n, the PRTF spectrum contains the resonant component, while a combination of the n multi-notch filters provides the reflective component. Such an algorithm allows separate analysis of the two phenomena's influence on HRTFs.

To this purpose, we ran the algorithm on all median-plane left HRTFs [1] of subjects in the CIPIC database [2]. Results showed that, while the resonant component is in broad terms similar among different subjects, the reflective component comes along critically subject-dependent. We now move to a more detailed analysis of resonance and notch patterns.

2.1 The resonant component

Every subject's PRTF presents two main resonance areas along the frequency axis. The first one, centered around 4 kHz, appears to be very similar amongst subjects since it spans all elevations; also, this area includes Shaw's omnidirectional mode 1. The resonance's bandwidth looks like increasing with elevation; however, knowledge of pinna modes implies that a second resonance is likely to interfere within this frequency range, specifically Shaw's mode 2 (centered around 7 kHz with a magnitude of 10 dB). On the other hand the second resonance area, although differing both in shape and magnitude amongst subjects, is most prominent at low elevations between 12 and 18 kHz (a frequency range which is in general agreement with Shaw's horizontal modes 4, 5, and 6) and smoothly dissolves as the elevation angle increases up above the horizontal plane.

Given that resonances have a similar behaviour in all of the analyzed PRTFs, customization of this component for the model may be overlooked. The mean magnitude spectrum (shown in Figure 1) was instead calculated and analyzed for resynthesis. More in detail, we applied a naïve procedure that extracts for every available elevation angle the two maxima of the mean magnitude spectrum, which outputs the gain G_p^i and central frequency CF_p^i of each resonance peak, $i = 1,2$, and the corresponding 3dB bandwidth BW_p^i. Then, a fifth-order polynomial (with the elevation ϕ as independent variable) was fitted to each of the former three parameters, yielding the functions $G_p^i(\phi)$, $CF_p^i(\phi)$, and $BW_p^i(\phi)$, $i = 1,2$. These functions will be used in the model to continuously control the evolution of the resonant component when the sound source is moving along elevation.

[1] Considering the interaural polar coordinate system, responses for azimuth $\theta = 0^\circ$ and elevations ranging from $\phi = -45^\circ$ to $\phi = 90^\circ$ at 5.625-degree steps were considered. Each HRTF was adequately windowed in order to eliminate reflections coming from shoulders or torso, hence roughly yielding the corresponding PRTF.

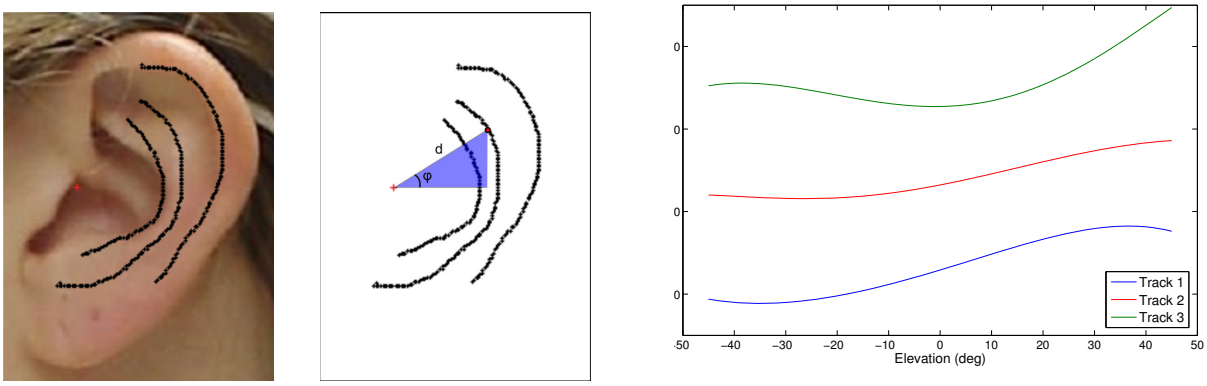
2.2 The reflective component

Analysis of the reflective part revealed that while PRTFs generally exhibit poor notch structures when the source is above the head, as soon as elevation decreases the number, spectral location, and depth of frequency notches grows to an extent that differs from subject to subject. Nonetheless, tracking notches along elevation angles highlighted the presence, in the vast majority of subjects, of three main (and apparently continuous) notch tracks between 5 and 14 kHz, whose evolution could be directly related to the location of reflection points over pinna surfaces. As a matter of fact, assuming that the coefficient of all reflections occurring inside the pinna is negative, the extra distance travelled by the reflected wave with respect to the direct wave must be equal to half a wavelength in order for destructive interference (i.e. a notch) to occur, which translates into a notch frequency that is inversely proportional to such distance. Hence, under the simplification that the reflection surface is always perpendicular to the soundwave, we consider the mapping function

$$d(\phi) = \frac{c}{2CF_n}, \tag{1}$$

where $d(\phi)$ is the distance of the hypothetical reflection point from the ear canal at elevation ϕ, CF_n is the notch central frequency, and c is the speed of sound. Direct translation of all the notch frequencies along elevations to distances through the above mapping function showed an encouraging correspondence between computed reflection points and pinna contours seen from a side view of the head [13].

This result allows us to perform the inverse procedure, sketched in Figure 3, in order to extract notch frequencies from a representation of the pinna contours. Specifically, first a picture of the left pinna of a CIPIC subject is resized to match the real dimensions according to his/her anthropometric data. Then, since automatic contour extraction is beyond the scope of this paper and, additionally, is not straightforward because of both the low pixel resolution of pinna photographs and the presence of "hidden" contours, the three most prominent and relevant contours of the pinna are manually traced with the help of a pen tablet and stored as a sequence of pixels. These are translated into a couple of polar coordinates (d, ϕ),



(a) *Contour tracing* (b) *Mapping to polar coords* (c) *Notch track extraction and approximation*

Figure 3: *Notch frequency extraction from a picture of the pinna (Subject* 048*).*

with respect to the point where the microphone lied[2], through simple trigonometric computations. Finally, the notch frequency CF_n is derived just by reversing Equation 1 and, similarly to the case of resonances, the sequence of points (CF_n^j, ϕ) for each of the three notch tracks, $j = 1,2,3$, is linearly approximated through a fifth-order polynomial $CF_n^j(\phi)$.

For what concerns the other two parameters defining a notch, i.e. gain G_n and 3dB bandwidth BW_n, there is still no evidence of correspondence with anthropometric quantities. We performed a first-order statistical analysis on the depth and bandwidth of notches, subdivided by notch track and elevation, among CIPIC subjects. This analysis reveals a high variance within each track and elevation, and mean values which lie approximately constant apart from a slight decrease in notch depth and a slight increase in bandwidth as the elevation increases (as an example, see Figure 2). In absence of clear elevation-dependent patterns, the mean of both gains and bandwidths for all tracks and elevations ϕ among all subjects is computed, and again a fifth-order polynomial dependent on elevation is fitted to each of these sequences of points, yielding functions $G_n^j(\phi)$ and $BW_n^j(\phi)$, $j = 1,2,3$. In the following Section we explain how to use this parametrization of PRTF features in synthesis phase, and discuss the objective effectiveness of all the introduced approximations.

3 The Structural Model

In the proposed model we introduce a fundamental assumption, i.e. elevation and azimuth cues are handled orthogonally and the corresponding contributions are thus separated in two distinct parts. The vertical control is associated with the acoustic effects relative to the pinna and the horizontal one is delegated to head diffraction. We establish this approximation following an informal inspection of different HRTF sets that revealed how median-plane reflection and resonance patterns generally vary very slowly when the azimuth's absolute value is increased, especially up to about 30°. In this way we are able to define customized elevation and azimuth cues that maintain their average behaviour throughout the front hemisphere.

3.1 Filter Model

Figure 4 depicts the global view of the model fed with the input parameters obtained through the analysis procedure described in Section 2. Examining its structure from left to right, we first use the simple spherical model that approximates head shadowing and diffraction described in [5], where the head radius parameter a is defined by a weighted sum of the subject's head dimensions using the optimal weights obtained in [3] through a regression on the anthropometric data for all CIPIC subjects. Then, a "resonances-plus-reflections" block approximating the pinna effects described in the previous Section allows elevation control. Our work's focus is on the latter part, knowing that several works on Interaural Time Difference (ITD) and Interaural Level Difference (ILD) estimation are available in literature and can be easily merged with our pinna model.

The only independent parameter used by the pinna block is the source elevation ϕ, which drives the evaluation of the polynomial functions describing resonances' center frequency $CF_p^i(\phi)$, 3db bandwidth $BW_p^i(\phi)$, and gain $G_p^i(\phi)$, $i = 1,2$, and the corresponding notch parameters $(CF_n^j(\phi), BW_n^j(\phi), G_n^j(\phi), j = 1,2,3)$. Only the center frequency CF_n is customized on the individual pinna shape, hence the corresponding polynomial must be computed offline previous to the rendering process.

The resonant part is modeled with a parallel of two different second-order peak filters. The first peak ($i = 1$) is of the form [15]

$$H_{res}^{(1)}(z) = \frac{1+(1+k)\frac{H_0}{2}+l(1-k)z^{-1}+(-k-(1+k)\frac{H_0}{2})z^{-2}}{1+l(1-k)z^{-1}-kz^{-2}}, \quad (2)$$

where

$$k = \frac{\tan\left(\pi\frac{BW_p^1(\phi)}{f_s}\right)-1}{\tan\left(\pi\frac{BW_p^1(\phi)}{f_s}\right)+1}, \quad (3)$$

$$l = -\cos\left(2\pi\frac{CF_p^1(\phi)}{f_s}\right), \quad (4)$$

$$V_0 = 10^{\frac{G_p^1(\phi)}{20}}, \quad (5)$$

$$H_0 = V_0 - 1, \quad (6)$$

and f_s is the sampling frequency. The second peak ($i = 2$) is implemented in the following form [9],

$$H_{res}^{(2)}(z) = \frac{V_0(1-h)(1-z^{-2})}{1+2lhz^{-1}+(2h-1)z^{-2}}, \quad (7)$$

[2]The position of the microphone during the CIPIC database measurements is not documented. This was optimized by taking the point to whose respect the mean square distance between hypothetical reflection points and pinna contours is minimal.

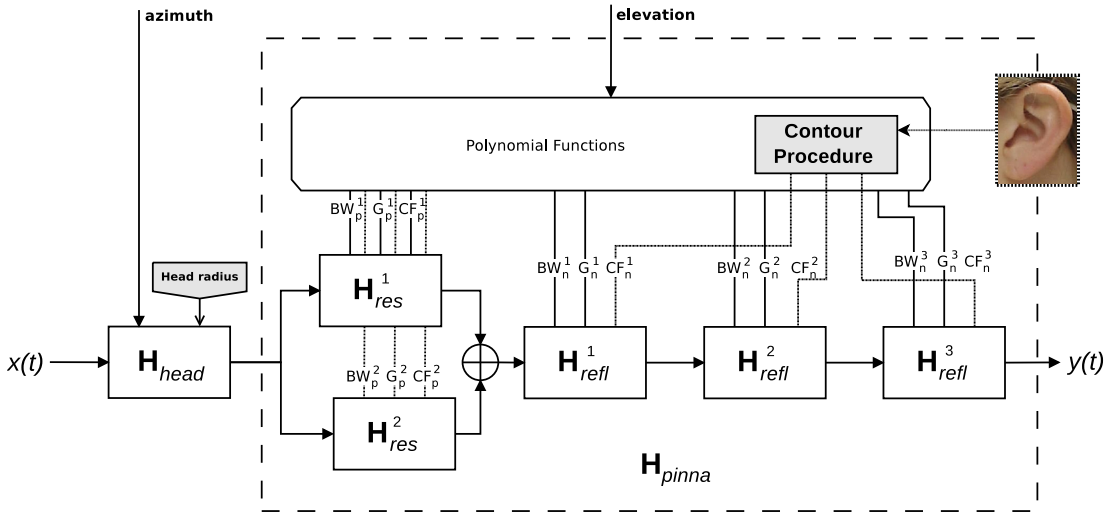


Figure 4: *The structural customized HRTF model.*

$$h = \frac{1}{1+\tan\left(\pi\frac{BW_p^2(\phi)}{f_s}\right)}, \tag{8}$$

while l and V_0 are defined as in Eqs. (4) and (5), where the index of the polynomial is $i = 2$. The reason for this distinction lays on the exact low-frequency behaviour we want to model. The former implementation has unitary gain at low frequencies so as to preserve such characteristic in the parallel filter structure, while the latter has a negative dB magnitude in this frequency range. In this way, the overall pinna filter does not alter the spectrum of the signal outputted by the spherical head filter.

The notch filter implementation is of the same form as the first peak filter with the only differences in the parameters' description. In order to keep the notation correct, the polynomials $\mathscr{P}_p^1$ must be substituted by the corresponding notch counterpart $\mathscr{P}_n^j$, $j = 1,2,3$, and parameter k defined in Eq. (3) is replaced by its cut version

$$k = \frac{\tan\left(\pi\frac{BW_n^j(\phi)}{f_s}\right) - V_0}{\tan\left(\pi\frac{BW_n^j(\phi)}{f_s}\right) + V_0}. \tag{9}$$

The three notch filters are cascaded, resulting in a global 6-th order multi-notch filter.

3.2 Results and Discussion

The above model was tested on different CIPIC subjects; in the following we present the results for two of them, Subject 020 (see Figure 5) and Subject 048 (see Figure 3a). In both cases for the tracing procedure we chose as contours the rim border, concha wall and border, and antihelix. Since the concha back wall is not fully visible from the picture's lateral view of the head, a tentative contour for this surface was drawn. This complication would suggest as future work the use of a 3-D representation of the pinna that allows to investigate its horizontal section, also because in most cases the pinna structure does not lie on a parallel plane with respect to the head's median plane, especially in subjects with protruding ears. Still, beside the unavalaibility of such kind of reconstruction for CIPIC subjects, our original aim is to keep the contour extraction procedure as low-cost and accessible as possible.

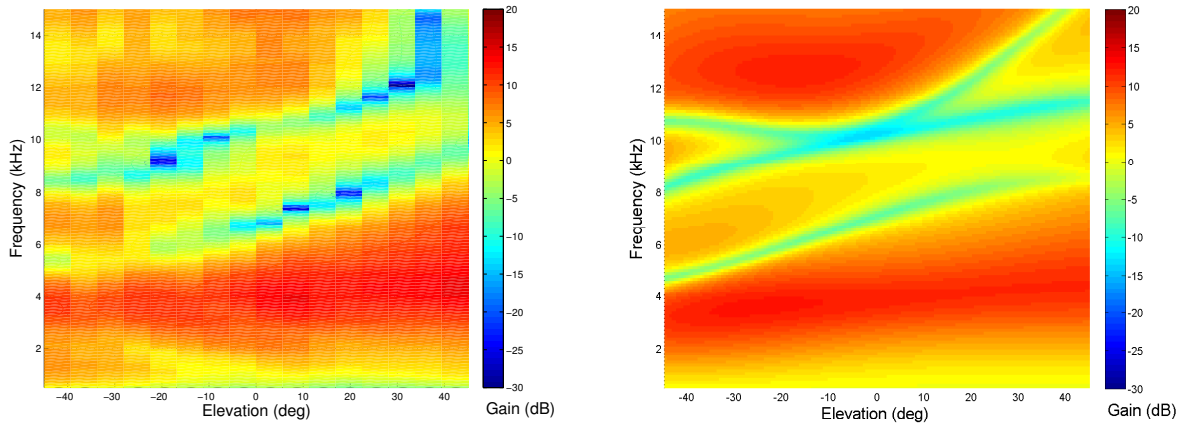
Having fed the model with all polynomial functions evaluated at half-degree elevation step, we are now ready to compare the original versus synthesized HRTF magnitude plots, shown in Figures 6-7. We focus on the frequency range up to 15 kHz where all the relevant informations are included, spanning elevations between -45 and 45 degrees in the median plane.

Figure 5: *Contour extraction on Subject* 020*'s pinna.*

Besides the different elevation resolution in the original and synthetic HRTF plots, similar features can be observed:

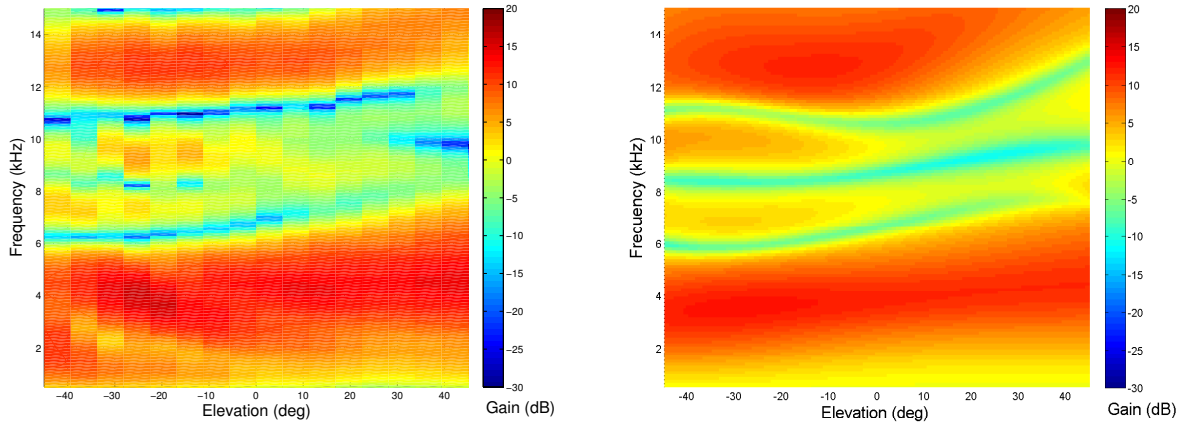
1. The first resonance, being omnidirectional and having an almost common behaviour in all subjects, is well approximated in both cases;
2. The extracted notch tracks, although much smoother than the original ones, closely follow the measured patterns, attesting fitness of the contour extraction and mapping procedure;
3. Gains, even in the intermediate frequency areas between notches and resonances, are overall preserved.

Coming to subtler differences, Subject 020 originally exhibits a wide dip around $\phi = 40°$ in the highest frequency range which is not correctly reproduced; this may be due to the superposition of two or more notches that cannot be detected when tracing the pinna contours. As for Subject 048, comparing his picture with the original HRTF plots we can note a relationship between the shorter antihelix and concha wall reflection surfaces and two distinct notch tracks, the first located around 8 kHz at negative elevation and the second around 10 kHz at positive elevation. Since we have chosen



(a) *Original plot.* (b) *Synthetic plot.*

Figure 6: *Original and synthetic HRTF magnitude plots for Subject* 020.



(a) *Original plot.* (b) *Synthetic plot.*

Figure 7: *Original and synthetic HRTF magnitude plots for Subject* 048.

to model three contours only, these two notches are collapsed in one continuous track, see Fig. 7b. A further notch appears around 15 kHz, yet it is likely associated with a mild pinna contour.

Finally, in each of the two subjects the second resonance is clearly overestimated and its shape doesn't find a strong visual correspondence. Such mismatch highlights a complex spectrum evolution due the presence of two or more resonances interacting in the upper frequency range for elevations in proximity of the horizontal plane [12]. However, following our choice of limiting the number of resonances to two, and assuming the first resonance to be omnipresent, the second synthetic resonance has to cover multiple contributions.

Further analysis is required toward a detailed model that takes into account the individual differences among subjects and their psychoacoustical relevance besides the observed objective dissimilarities. Synthetic notches bear a smoother magnitude and bandwidth evolution compared to the original ones; in particular, magnitude irregolarities in the original notches could arise from superposition of multiple reflections and, in addition, from a strong sensitivity of the subject's spatial position during the HRTF recording session. Psychoacoustical evaluations into virtual environments are definitely needed to reveal the appreciation degree of our approach together with the real perceived weight of such homogeneous notch and peak shapes.

4 The Real-Time System

Our model was designed so as to avoid expensive computational and temporal steps such as HRTF interpolation on different spatial locations, best fitting non-individual HRTFs, or the addition of further artificial localization cues, allowing implementation and evaluation in a real-time environment such as the one we have realized and which is now described.

The setup is schematically represented in Figure 8. The user is wearing a pair of wireless headphones with three LED markers positioned on them, one on the left earphone (green marker), one on the right (red), and one on top of the headphones (blue). A specific area is delimited by 8 high-resolution cameras coordinated by the PhaseSpace Impulse motion tracking system, featuring a data capture rate of 480 Hz (frames/s). The square walking area is 3.0 m $\times$ 3.5 m, while the eight cameras are placed on 3 horizontal bars (par-

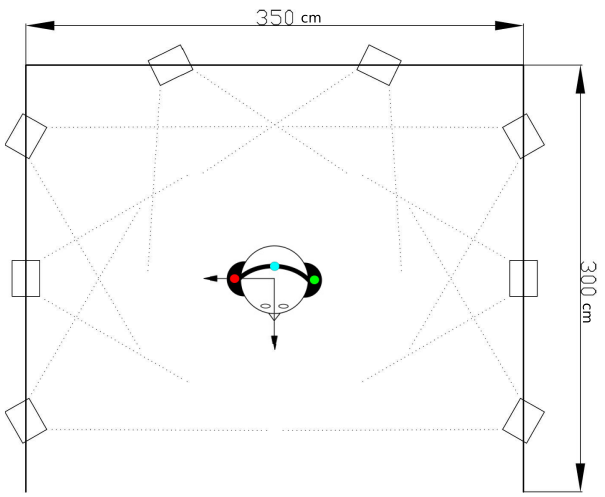


Figure 8: *The real-time experimental setup.*

allel to the square perimeter) placed 2.5 m high. Such redundant camera placement easily allows tracking of the three markers even in certain occlusion cases.

Two informations are required to properly compute the two audio channels: the listener's head position in the tracked region and the relative direction of the head with respect to one or more simulated sound sources. In this case we focus on a single source scenario because it's the simplest experiment not involving any acoustic effects due the mutual interaction of possible multiple sound sources. The sound source movements could be defined in different ways, depending on the evaluation techniques involved. In sake of simplicity, we assume the sound source to move in front of the listener following precomputed trajectories.

The PhaseSpace Impulse routes each marker position via the OSC (Open Sound Control) protocol towards another workstation responsible for computationally carrying out the sound rendering task. Our HRTF model is implemented in Pure Data [10], a real-time audio processing environment. A C-developed external module realizes some simple three-dimensional geometric calculations to convert the markers' positions in the interaural polar coordinates of the sound source with respect to the user's head, and the head position is continuously kept updated. Azimuth, elevation and distance of the sound source become the input parameters on each of the two audio channels, left and right, processed in a Pure Data patch. A single audio channel includes our model anteceded by a delay block taking into account for the ITD cue,

$$ITD \sim \frac{a}{c}(\theta + sin\theta) \tag{10}$$

where a is the listener's head radius, c is the speed of sound and θ is the source azimuth.

Furthermore, distance of the source is simulated by varying the signal's loudness proportionally to the inverse of the square distance. Air absorption and reverberation could be also added as auxiliary distance cues. These last considerations strictly depend on the applications and tasks where our model is employed and on how the experimental protocol is designed.

5 Conclusions and Future Work

In this paper we presented a customized structural model of the HRTF that can be used in real-time environments for 3-D audio rendering. While having verified the objective fitness of the model to real, measured HRTFs in the considered spatial range, subjective evaluations are required in order to attest its effectiveness in binaural hearing, with the aid of the setup described in Section 4. Also, ongoing and future work includes automatic pinna contour extraction from 2-D or 3-D representations through guided edge detection or the use of depth cameras to detect hidden contours, respectively, and extension of the model to source positions behind, above, and below the listener. To the latter end, analysis of a newly collected database of PRTFs [14] through the same tools and algorithms that were used for the CIPIC data could provide understanding of the behaviour of the pinna in such cases. Finally, in order to have a full, surrounding binaural experience, inclusion of the shoulders and torso contribution to the HRTF in the model would add further reflection patterns and shadowing effects when the source is below the listener.

Acknowledgements

The authors wish to acknowledge Professor Ralph Algazi for the kind provision of publicly unavailable data from the CIPIC database.

References

[1] R. V. Algazi, R. O. Duda, R. P. Morrison, and D. M. Thompson. Structural composition and decomposition of HRTFs. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 103–106, New Paltz, New York, USA, 2001.

[2] R. V. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano. The CIPIC HRTF database. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 1–4, New Paltz, New York, USA, 2001.

[3] V. R. Algazi, C. Avendano, and R. O. Duda. Estimation of a spherical-head model from anthropometry. *J. Audio Eng. Soc.*, 49(6):472–479, 2001.

[4] D. W. Batteau. The role of the pinna in human localization. *Proc. R. Soc. London. Series B, Biological Sciences*, 168(1011):158–180, August 1967.

[5] C. P. Brown and R. O. Duda. A structural model for binaural sound synthesis. *IEEE Transactions on Speech and Audio Processing*, 6(5):476–488, 1998.

[6] C. I. Cheng and G. H. Wakefield. Introduction to head-related transfer functions (HRTFs): Representations of hrtfs in time, frequency, and space. *J. Audio Eng. Soc.*, 49(4):231–249, April 2001.

[7] M. Geronazzo, S. Spagnol, and F. Avanzini. Estimation and modeling of pinna-related transfer functions. In *Proc. of the 13th Int. Conference on Digital Audio Effects (DAFx-10)*, Graz, Austria, September 2010.

[8] H. Møller, M. F. Sørensen, C. B. Jensen, and D. Hammershøi. Binaural technique: Do we need individual recordings? *J. Audio Eng. Soc.*, 44(6):451–469, 1996.

[9] S. J. Orfanidis, editor. *Introduction To Signal Processing*. Prentice Hall, 1996.

[10] M. Puckette. Pure Data: another integrated computer music environment. In *Proc. International Computer Music Conference*, pages 37–41, 1996.

[11] P. Satarzadeh, R. V. Algazi, and R. O. Duda. Physical and filter pinna models based on anthropometry. In *Proc. 122nd Convention of the Audio Engineering Society*, Vienna, Austria, May 5-8 2007.

[12] E. A. G. Shaw. *Binaural and Spatial Hearing in Real and Virtual Environments*, chapter Acoustical features of human ear, pages 25–47. R. H. Gilkey and T. R. Anderson, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1997.

[13] S. Spagnol, M. Geronazzo, and F. Avanzini. Fitting pinna-related transfer functions to anthropometry for binaural sound rendering. In *IEEE International Workshop on Multimedia Signal Processing*, pages 194–199, Saint-Malo, France, October 2010.

[14] S. Spagnol, M. Hiipakka, and V. Pulkki. A pinna-related transfer function database. In *Proc. of the 14th Int. Conference on Digital Audio Effects (DAFx-11)*, Paris, France, September 2011.

[15] U. Zölzer, editor. *Digital Audio Effects*. J. Wiley & Sons, New York, NY, USA, 2002.