PROCEEDINGS of the
24th International Congress on Acoustics

October 24 to 28, 2022 in Gyeongju, Korea

ABS-0267

PRTF notch enhancement by selective modification of pinna surface reflection coefficient

Parham MOKHTARI[1]; Yutaro HIROTA[2]; Daisuke MORIKAWA[3]

Toyama Prefectural University, Japan

ABSTRACT

Auditory cues to sound localization in the median plane are the monaural spectral cues, i.e., the peaks and notches of pinna-related transfer functions (PRTFs). While a typical or stylized peak-notch pattern has been identified in previous studies, actual patterns exhibit considerable individuality and notches are not always clearly defined at all elevation angles, their trajectories appearing sometimes broken or discontinuous. Here we investigate the possibility of enhancing PRTF notch trajectories in the median plane, simply by adjusting physical properties (the acoustic reflection coefficient) at selected parts of the pinna surface, without altering pinna geometry. This is done with the aid of acoustic simulation, and a combination of sensitivity and clustering analyses. Results show that notches throughout the median plane can be deepened and their trajectories made more continuous, by lowering the reflection coefficient at certain parts of the pinna surface.

Keywords: PRTF, pinna, notch, median plane, reflection coefficient

1. INTRODUCTION

While inter-aural time and level differences (ITD and ILD) are the main auditory cues for sound localization along the left-right spatial dimension (e.g., in the horizontal plane), in the median plane such differences are minimal and we rely instead on monaural spectral cues, i.e., the peaks and notches of head-related transfer functions (HRTFs). As most of these peaks and notches are generated by reflections and associated resonances within the cavities of the outer ear (pinna), they are essentially the same as the peaks and notches of pinna-related transfer functions (PRTFs).

Typically in the median plane, PRTF peak frequencies are nearly invariant while the notches vary in frequency with elevation angle (1, 2). The peaks are the normal modes (resonances) of the pinna cavities (3, 4), and empirical formulae have been proposed to estimate their frequency and amplitude from individual pinna anthropometry (5, 6). While the notches are known to be generated by sound-wave cancellation (i.e., destructive interference) within and around the pinna cavities, the precise physical mechanisms are not yet fully clarified, although a number of compelling hypotheses have been proposed: e.g., reflection model for sound sources in the front hemifield (7, 1, 8, 9); tripartite generation mechanism covering all directions (2); dipole generation mechanism (10).

As pinna morphology is unique to each individual, the patterns of PRTF peaks and notches in the median plane also display considerable individuality (11); hence the need for personalization in 3D audio applications. One aspect of individuality is the apparent discontinuities in notch trajectories: while in a stylized pattern the first two or three notches typically rise and then fall in frequency with increasing polar angle, reaching a highest frequency typically a few degrees behind directly above (depending on individual pinna geometry) before decreasing again at lower elevations at rear (2), in practice the notches often do not form well-defined and continuous trajectories. Rather, notch trajectories are often broken by jumps, locations where they appear to be shallow or missing, or other discontinuities. This can make it difficult to trace a continuous path for a given notch or even to label a notch consistently across a wide range of polar angles.

For the front location, previous research has shown that broadband sounds delivered by headphones or tube-phones can be correctly externalized in front, provided the listener's peaks and notches are

[1] parham@pu-toyama.ac.jp

[3] dmorikawa@pu-toyama.ac.jp

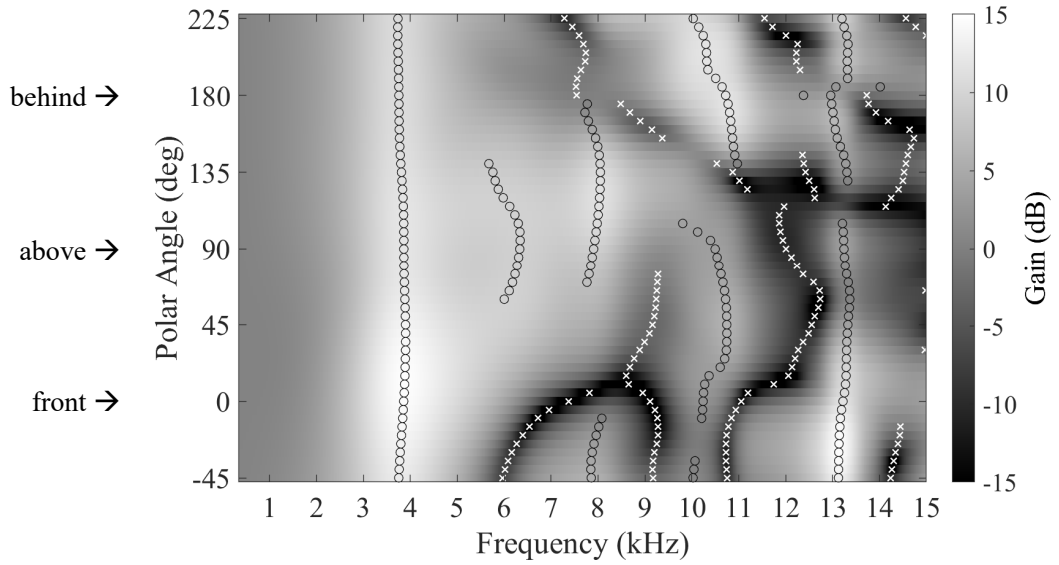

Figure 1 – Median-plane baseline PRTFs calculated by FDTD simulation, with peaks (o) and notches (x) overlayed.

not overly smoothed (12), or a typical mean depth of the first two notches is either maintained or deepened (13). While notch discontinuities are presumably a characteristic of individual pinna geometry, an intriguing question is whether a listener's sound localization performance in the median plane, or more generally a listener's 3D audio experience including externalization of sound sources (real or virtual), might be enhanced by somehow deepening notches and filling in the locations where notches were missing along their broken trajectories. While this can be done purely with signal processing in the context of a virtual audio system, this approach would not yield any insight regarding notch generation mechanisms, nor would it be useful in a natural (unassisted) listening situation.

Here we use a 3D voxelated model of a human pinna together with an acoustics simulator, to assess whether, and to what extent, manipulation of only the acoustic reflection coefficient at judiciously selected parts of the pinna surface can help to deepen PRTF notches and render their trajectories in the median plane more continuous. The present study is therefore concerned primarily with physiological acoustics, and evaluation of psychoacoustic performance with enhanced PRTFs is left for future work.

2. BASELINE PRTFs

The human pinna used in this study was previously measured by magnetic resonance imaging (14) and voxelated on a 2 mm isovoxel 3D grid (15). Sound speed c and medium density ρ at each voxel were set to values corresponding to either air (for air voxels) or water (for head/pinna voxels); the baseline reflection coefficient at the air-skin interface was thus $k_r = 0.999$. Employing the reciprocity principle in 3D finite-difference time domain (FDTD) acoustic simulation, an air voxel adjacent to the center of the blocked entrance of the ear-canal was excited with a broadband Gaussian pulse, and the Kirchhoff-Helmholtz method (16) was used to project the evolving sound field around the pinna to 55 farfield observation points a distance 1 m from the head center: i.e., to 55 locations in the median plane, with polar angles ranging from $-45°$ (below front) through $90°$ (directly above) to $225°$ (below back) in $5°$ steps. PRTFs were then calculated by free-field normalization with respect to the transfer functions obtained using the same source and receiver positions with the head/pinna absent.

The baseline PRTFs computed in this way are shown in Fig. 1, with extracted peaks and notches overlayed. As noted in the Introduction, while a typical rise-and-fall in frequency is qualitatively observed in the notch patterns, it is difficult to trace a continuous, unbroken path. For example, in the front hemifield the trajectory of the first notch rises rapidly in frequency and appears to merge with the second notch at polar angle $10°$ (just under 9 kHz); a very shallow notch at around 9 kHz then fades away at higher elevations. Meanwhile, what started out as the third notch around 11 kHz at low front, also rapidly rises in frequency and appears to become the second or even first notch at higher

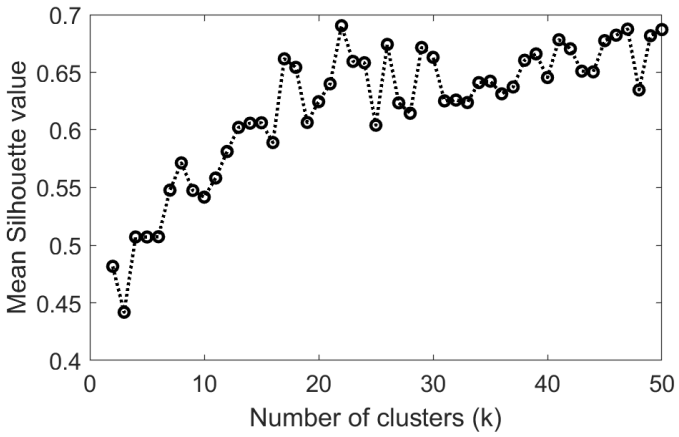

Figure 2 – Evaluation of clustering quality as a function of the assumed number of clusters.

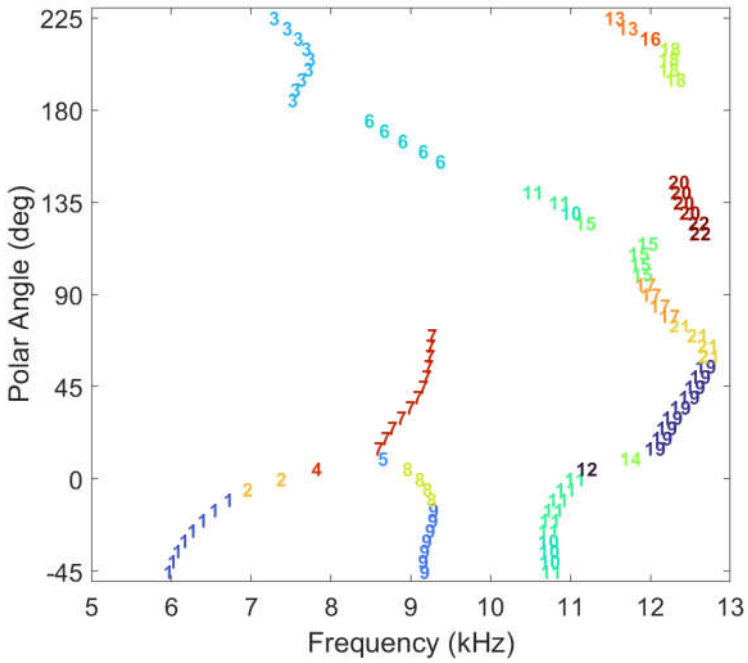

Figure 3 – Results of clustering the notch amplitude sensitivity maps, at the optimum $k = 22$ clusters. Each notch is depicted by its cluster number (and a unique color), in order of increasing frequency.

elevations. Attempting to follow its trajectory at increasing polar angles in the rear hemifield as it descends in frequency, the trajectory appears to be broken up or weakened in many places.

In this paper our aim is to explore whether notch trajectories in the median plane can be deepened and thus made more continuous, simply by adjusting the reflection coefficient at certain parts of the pinna surface. To find the required combination of surface voxels, as described in the following section we first use perturbation analysis to calculate pinna sensitivity maps, that summarize the acoustic effects of modifying the reflection coefficient at only one voxel at a time.

3. SINGLE-VOXEL PERTURBATIONS

Without changing the pinna shape, c and ρ at each of the 1786 surface-voxels (covering the surface of the pinna and a small adjoining patch of the side of the head) were modified one voxel at a time such that $k_r = 0.5$ at only that part of the surface. For each such perturbation, the 55 PRTFs in the median plane were re-calculated by FDTD simulation, peak and notch frequencies F (in Hz) and amplitudes A (in dB) were extracted from the PRTFs, and the corresponding baseline values were subtracted to obtain ΔF and ΔA for every peak and notch. By color-coding these delta values on the

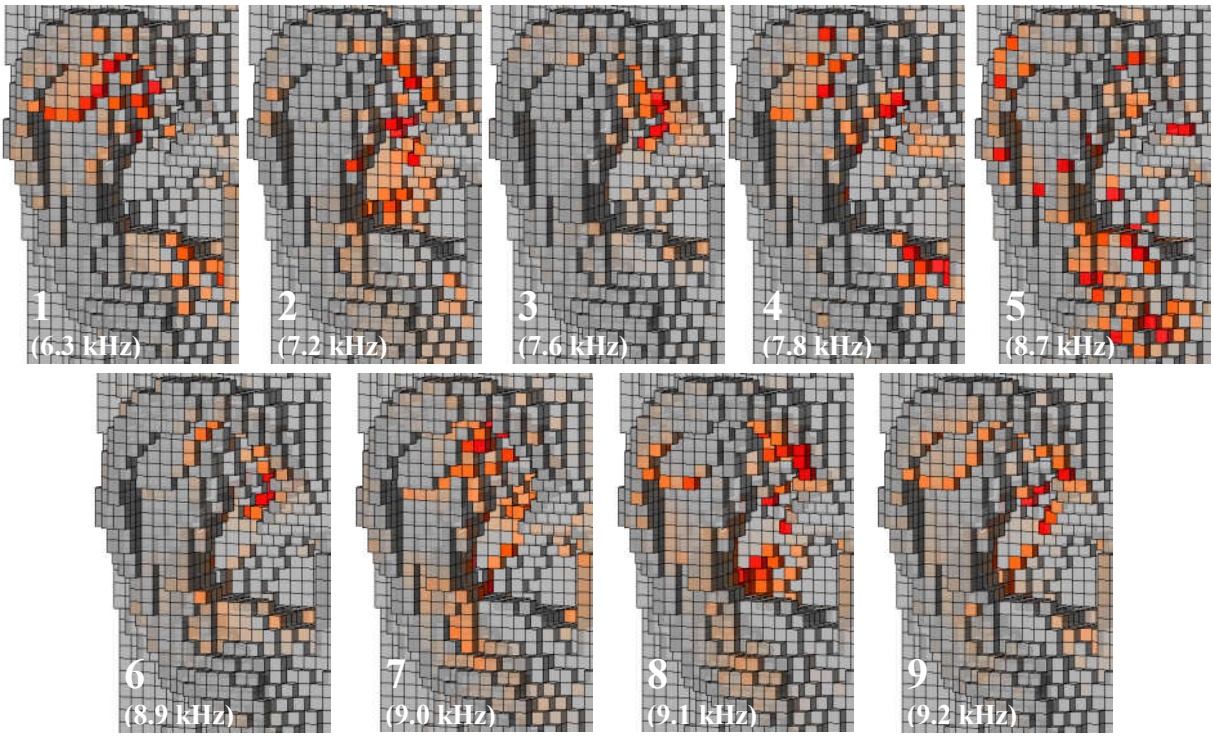

Figure 4 – The first 9 **cluster centroids** (mean sensitivity maps for notch amplitude), covering all notches up to 10 kHz as shown in Fig. 3. For visual clarity, only negative shifts in amplitude, i.e. deepening of notches, are shown in color (positive shifts, i.e. shallowing of notches, are ignored). Gray: $\Delta A = 0$ dB. Most saturated color: $\Delta A = -0.0727$ dB. The mean frequency of the notches in each cluster is shown in parentheses.

pinna surface, a frequency or amplitude *sensitivity map* was created for each peak or notch. Here we focus on only *amplitude* sensitivity maps for *notches*.

We had previously shown a small selection of amplitude sensitivity maps (15), for what appeared to be the first notch at five different elevation angles in the front hemifield. Here we extend that sensitivity analysis to all the notches depicted in Fig. 1. Visualization of these sensitivity maps (not shown here) confirmed our prior results (15), in that the parts of the pinna surface that most strongly affect notch amplitude (i.e., depth) depend to a large extent on the notch frequency and spatial location. Our current analysis revealed that while notches close in both frequency and polar angle often had somewhat similar sensitivity maps, when considering all the notches throughout the median plane a large variety of sensitivity maps were obtained. Therefore, we next applied clustering analysis to identify the main types of sensitivity map, as described in the following sub-section.

3.1 Clustering analysis of notch amplitude sensitivity maps

We applied *k*-means clustering to the 99 amplitude sensitivity maps obtained for the notches below 13 kHz (a notch was defined simply as any local minimum with a gain below 0 dB). For this purpose, each sensitivity map was represented by a ΔA vector of length 1786 (the total number of surface voxels). The 99 vectors were subjected to *k*-means clustering using *correlation* as a measure of similarity (i.e., 1 – *correlation* as a distance measure), and 100 randomly-seeded trials (replicates) were run at each number of clusters from $k = 2$ to $k = 50$. To decide the optimum number of clusters, we used the mean silhouette value, which is a measure of how well separated the clusters are. As shown in Fig. 2, this objective criterion increased to a maximum at $k = 22$ clusters, beyond which there was no further improvement.

At the optimum $k = 22$ clusters, the clustering membership of the 99 notches is shown in Fig. 3. These results confirm that the amplitude sensitivity maps of notches close in proximity in terms of both frequency and polar angle, resemble each other to the extent that they form local clusters (e.g., clusters 1, 3, 6, 7, 8, 9). On the other hand, the results also indicate that a number of sensitivity maps are so dissimilar to all others that they form single-member clusters: these often occur at relatively unstable locations along a notch trajectory, where the notch frequency rapidly changes with polar

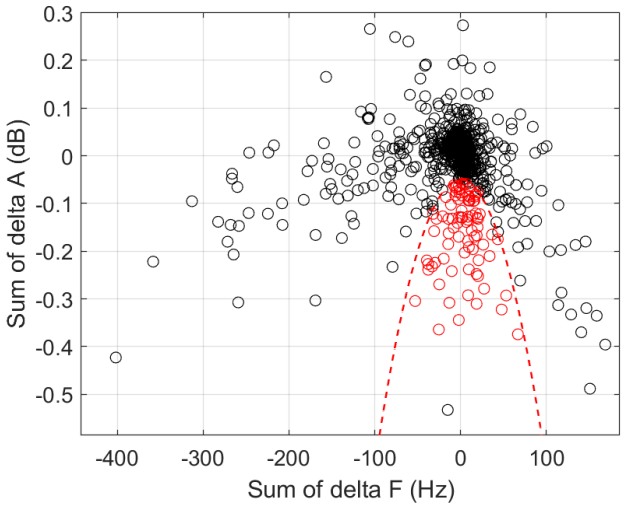

Figure 5 – Scatterplot of ΔF_{Sum} vs ΔA_{Sum}, the combined effects of the 9 cluster centroids (Fig. 4) on notch frequency and amplitude. The 84 red points were selected by defining a parabola (red dashed curve) as described in section 4.

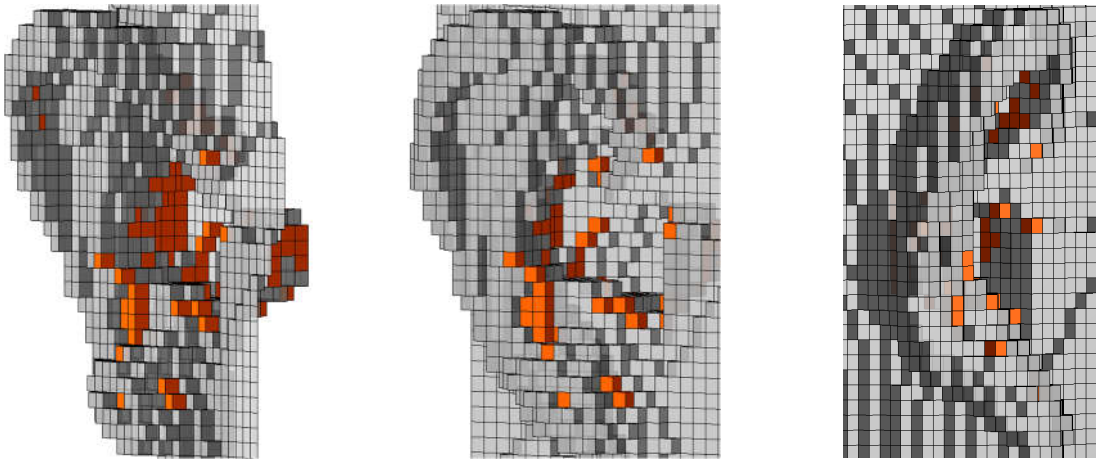

Figure 6 – Three views of the right pinna, highlighting the 84 surface-voxels selected in Fig. 5. Left panel: frontal view. Middle panel: side-frontal view. Right panel: rear view.

angle (e.g., clusters 4 and 12 at 5°, and clusters 5 and 14 at 10°).

Fig. 4 shows mean sensitivity maps representing the first nine cluster centroids, thus summarizing the data for all notches under 10 kHz. In these maps, colored voxels imply that *reducing* the reflection coefficient at that part of the pinna surface resulted in *deeper* notches; the higher the color saturation, the larger the effect. In other words, normal amounts of reflection from the colored parts tend to impede or obstruct the formation of deep notches.

While the sensitivity to each surface-voxel perturbation appears to be quite small (only −0.07 dB shift in notch amplitude at the most saturated color), the combined effect of reducing the reflection coefficient at several voxels simultaneously is expected to be more substantial. In the next section we explore this possibility.

4. MULTI-VOXEL PERTURBATIONS

Despite the diversity of patterns in Fig. 4, we aimed to find one particular combination of surface-voxels, for which reducing their reflection coefficients together would give the following result: maximally deepen as many of the notches as possible, while minimally influencing the notch frequencies.

In (17) we reported on how we achieved these aims for the first two notches of the PRTF at front (polar angle 0°) only. Indeed, the sensitivity map for cluster centroid 2 shown here in Fig. 4, which

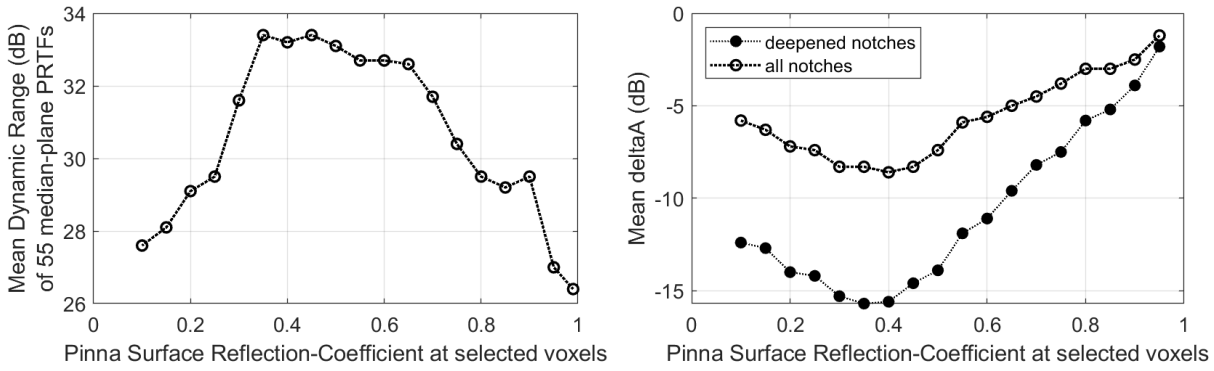

Figure 7 – PRTF dynamic range (left panel) and shift in notch amplitude (right panel) as a function of the value of k_r, the reduced reflection coefficient at the 84 selected pinna surface-voxels.

is a combination of notch sensitivity patterns at $0°$ and $-5°$, closely resembles our earlier results (17: Fig. 1(b)). Here, we extend those results to the entire median plane.

To achieve this, we combined the sensitivity patterns for the first 9 clusters shown in Fig. 4, simply by summing the numerical values across the 9 cluster centroids: this yielded a combined amplitude sensitivity map ΔA_{Sum}. Similarly, by summing across the 9 centroid patterns of shifts in notch frequency, we obtained a combined frequency sensitivity map ΔF_{Sum}.

A scatterplot of the 1786 pairs of $\{\Delta F_{Sum}, \Delta A_{Sum}\}$ is shown in Fig. 5. On the basis of this scatterplot, we selected a subset of surface-voxels that would help achieve our aim of deepening notches while minimally affecting their frequencies. This was done heuristically by defining a parabola $\Delta A_{Sum} = -0.00006(\Delta F_{Sum})^2 - 0.05$ (shown by the red dashed curve in Fig. 5) and choosing all voxels lying below the curve (the red points in Fig. 5).

As a side issue, we noticed that one of the points lying below the curve corresponded to the surface voxel directly adjacent to the acoustic source, i.e. at the blocked entrance of the ear-canal. This voxel was excluded from the selected list, because reducing its reflection coefficient tended to simply lower the amplitude of PRTFs not only at the notches but at all frequencies. This explains the lone black point under the parabola in Fig. 5.

The 84 surface-voxels thus selected are colored orange in 3 views of the pinna in Fig. 6. The selected voxels are located mainly across a large part of the concha's back wall, in smaller parts of the front walls of the concha and cymba, along a patch extending down from the antihelix partway toward the lobule, and in a small patch between the antitragus and the intertragic notch. Further research is needed to assess the acoustic-physiological significance of the selected parts of the pinna surface, for example by visualizing the 3D patterns of acoustic pressure in and around the pinna cavities, specifically at notch frequencies; such a line of investigation could help clarify the notch generation mechanism, and is left for future work.

Next, we recalculated all 55 PRTFs by FDTD simulation with reduced reflection coefficient at only the 84 selected voxels. The reflection coefficient was reduced to a value from $k_r = 0.95$ to 0.10 in steps of 0.05, and at each step we extracted the mean dynamic range of the PRTFs (the difference between the maximum and minimum gain in dB) as well as the change in the amplitudes of notches relative to baseline. The extracted values are shown in the two panels of Fig. 7. These data indicate an optimum value for the reduced reflection coefficient: at $k_r = 0.35$, the mean dynamic range attained its highest value of 33.4 dB (i.e., 7 dB greater than the baseline dynamic range of 26.4 dB), and the deepened notches shifted in amplitude by the largest amount (-15.7 dB on average).

Lastly, Fig. 8 provides a visual comparison of the baseline median-plane PRTFs (at top) with those obtained after reducing the reflection coefficient at only the 84 selected voxels to an intermediate step $k_r = 0.65$ (middle) and further down to $k_r = 0.35$ (bottom). These images clearly show the deepening of most (though not all) notches as k_r at the selected voxels was reduced. We would like to draw particular attention to notch trajectories in the rear hemifield (polar angles from $135°$ to $225°$) which were originally weak or broken in several places; at $k_r = 0.65$ those notches were clearly transformed to two continuous, unbroken trajectories.

While the analysis results in Fig. 7 indicated an optimum at $k_r = 0.35$ both in terms of notch deepening and PRTF dynamic range, the corresponding image at the bottom of Fig. 8 reveals several

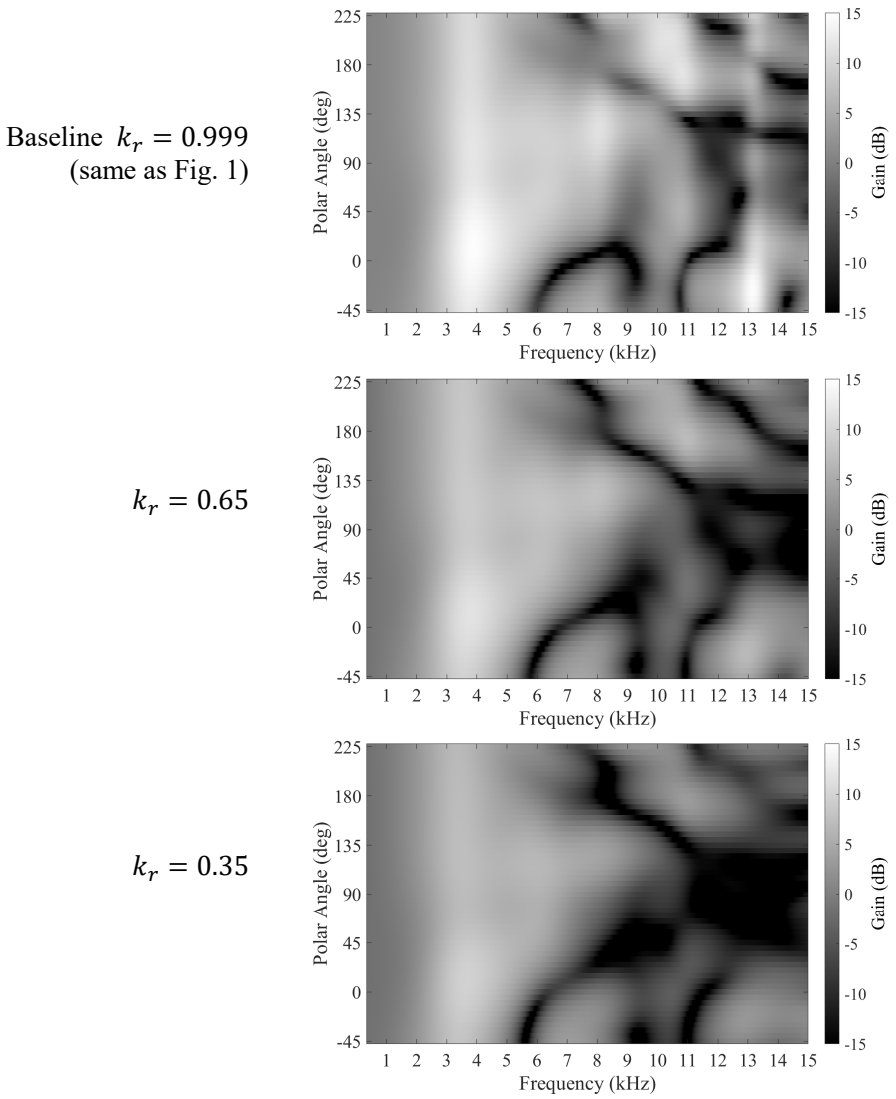

Figure 8 – Effect of reducing pinna surface reflection coefficient k_r at the 84 selected voxels (Fig. 6), on PRTFs in the median plane.

locations where notches were unintentionally weakened, such as the second notch just above 9 kHz at front-low, or the second notch trajectory at higher frequencies in the rear hemifield. Furthermore, while the 84 voxels were intentionally selected on the basis of minimally influencing the notch frequencies, the mean absolute shift in notch frequency monotonically increased as the reflection coefficient was lowered, reaching 207 Hz at $k_r = 0.65$, and 359 Hz at $k_r = 0.35$. In light of these observations, $k_r = 0.35$ may be regarded as too extreme in terms of notch distortion, and a more moderate value such as $k_r = 0.65$ may be preferable auditorily. In any case, it is necessary to supplement the present study with auditory evaluations to assess whether such modifications to a listener's PRTFs would help to enhance their spatial hearing.

5. CONCLUSIONS

In this study we used FDTD acoustic simulation as a tool to investigate the possibility of deepening PRTF notches, and thus enhancing the dynamic range of PRTFs (and by extension HRTFs) in the median plane. This was achieved by sensitivity analysis whereby, retaining the original pinna geometry, only the reflection coefficient at one surface-voxel at a time was modified and the consequences of such modification on the PRTFs were quantified. Due to the large number of sensitivity maps thus generated, clustering analysis was applied to reduce the patterns to a few basic types. Lastly, a subset of pinna surface-voxels was selected after combining the cluster centroids, and

it was shown that by reducing the reflection coefficient at only those parts of the pinna surface (only 84 of the 1786 surface-voxels), notches throughout the median plane were deepened and the PRTF dynamic range was enhanced by 7 dB on average.

As our study focussed purely on physiological acoustics, it is left for future work to assess the psychoacoustic response to such PRTF enhancements. In addition to applications in virtual audio, our results may be useful for investigating the effects on natural spatial hearing, of physically modifying a listener's pinna surface reflection, for example by painting parts of the pinna with an acoustically absorptive material. We also hope that such investigations can lead to an even better understanding of how notches are generated, and how they contribute to sound externalization and localization.

ACKNOWLEDGEMENTS

This work was partially supported by JSPS KAKENHI Grant Number JP17K00260.

REFERENCES

1. Raykar VC, Duraiswami R, Yegnanarayana B. Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. J Acoust Soc Am. 2005;118(1):364-374.
2. Takemoto H, Mokhtari P, Kato H, Nishimura R, Iida K. Mechanism for generating peaks and notches of head-related transfer functions in the median plane. J Acoust Soc Am. 2012;132(6):3832-3841.
3. Shaw EAG. Acoustical features of the human external ear. In: Gilkey RH, Anderson TR, editors. Binaural and Spatial Hearing in Real and Virtual Environments. New Jersey, USA: Lawrence Erlbaum Associates; 1997. p. 25-47.
4. Kahana Y, Nelson PA. Numerical modelling of the spatial acoustic response of the human pinna. J Sound and Vibration. 2006;292:148-178.
5. Mokhtari P, Takemoto H, Nishimura R, Kato H. Frequency and amplitude estimation of the first peak of head-related transfer functions from pinna anthropometry. J Acoust Soc Am. 2015;137(2):690-701.
6. Mokhtari P, Takemoto H, Nishimura R, Kato H. Vertical normal modes of human ears: individual variation and frequency estimation from pinna anthropometry. J Acoust Soc Am. 2016;140(2):814-831.
7. Hebrank J, Wright D. Spectral cues used in the localization of sound sources on the median plane. J Acoust Soc Am. 1974;56(6):1829-1834.
8. Mokhtari P, Takemoto H, Nishimura R, Kato H. Pinna sensitivity patterns reveal reflecting and diffracting surfaces that generate the first spectral notch in the front median plane. Proc IEEE-ICASSP; 22-27 May 2011; Prague, Czech Republic 2011. p. 2408-2411.
9. Spagnol S, Geronazzo M, Avanzini F. On the relation between pinna reflection patterns and head-related transfer function features. IEEE Trans ASLP. 2013;21(3):508-519.
10. Takemoto H, Mokhtari P, Kato H, Nishimura R, Iida K. Elevation-frequency relationship for the first spectral notch of pinna-related transfer functions in the median plane. Proc Acoust Soc Japan Autumn Mtg; 25-27 September 2013; Toyohashi, Japan. Paper 3-1-11; p. 843-846 (in Japanese).
11. Mokhtari P, Takemoto H, Nishimura R, Kato H. On individual differences in pinna-related transfer functions calculated by numerical simulation. Proc ICSV; 8-12 July 2012; Vilnius, Lithuania. Paper 120; p. 1-8.
12. Kulkarni A, Colburn HS. Role of spectral detail in sound-source localization. Nature. 1998;396:747-749.
13. Nishiyama O, Iida K. Generation of individual parametric head-related transfer functions of the front direction -- Validation of the fixed values of levels and Q factors for the notches and peaks. Proc Acoust Soc Japan Autumn Mtg; 7-9 September 2021; online. Paper 1-4-10; p. 661-664 (in Japanese).
14. Hirahara T, Otani M, Toshima I. Issues on HRTF measurements and binaural reproduction. Fundamentals Review. 2009;2(4):68-85.
15. Hirota Y, Mokhtari P, Morikawa D. An investigation of the effect of local pinna-surface reflection on the pinna-related transfer function by sensitivity analysis. Proc Acoust Soc Japan Autumn Mtg; 7-9 September 2021; online. Paper 1-4-7; p. 651-652 (in Japanese).
16. Mokhtari P, Takemoto H, Nishimura R, Kato H. Efficient computation of HRTFs at any distance by FDTD simulation with near to far field transformation. Proc Acoust Soc Japan Autumn Mtg; 10-12 September 2008; Fukuoka, Japan. Paper 1-8-12; p. 611-614.
17. Mokhtari P, Hirota Y, Morikawa D. Deepening of PRTF notches by selective modification of pinna surface reflection coefficient. Proc Acoust Soc Japan Spring Mtg; 9-11 March 2022; online. Paper 2-4-8; p. 649-650.