



HAL
open science

Inter-frequency band correlations in auditory filtered median plane HRTFs

Yukio Iwaya, Brian F. G. Katz, Tetsu Magariyachi, Yôiti Suzuki

► **To cite this version:**

Yukio Iwaya, Brian F. G. Katz, Tetsu Magariyachi, Yôiti Suzuki. Inter-frequency band correlations in auditory filtered median plane HRTFs. EAA Spatial Audio Signal Processing Symposium, Sep 2019, Paris, France. pp.179-182, 10.25836/sasp.2019.33 . hal-02275195

HAL Id: hal-02275195

<https://hal.science/hal-02275195>

Submitted on 30 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INTER-FREQUENCY BAND CORRELATIONS IN AUDITORY FILTERED MEDIAN PLANE HRTFS

Yukio Iwaya¹Brian FG Katz²Tetsu Magariyachi³Yôiti Suzuki⁴¹ Tohoku Gakuin University, Japan & Sorbonne Université, France² Sorbonne Université, CNRS, Institut Jean Le Rond d'Alembert, France³ Tohoku University, Japan (Currently SONY Corp., Japan)⁴ Tohoku University, Japan

iwaya@ieee.org

ABSTRACT

In this paper, head-related transfer functions (HRTFs) in the median plane were analyzed to investigate details of spectral cues of sound localization in the median plane. Head-related impulse responses (HRIRs), which are representation of HRTFs in time domain, were filtered with auditory filters based on equivalent rectangular bands (ERB). Level changes according to median plane angle were calculated for each band and inter-band correlations were analyzed. Results showed some ERB bands, which include characteristic notches, are negatively correlated to other bands. Furthermore, whole frequency bands were grouped into three aggregated bands and their level changes were analyzed to clarify causes of negative correlations.

1. INTRODUCTION

Spectral cues in head-related transfer functions (HRTF), such as peaks and notches occurring above 4 kHz, are important for sound localization in the median plane [1–5]. However, it may be complicated for the auditory system to detect absolute frequency and level peaks and notches, mapping them to three-dimensional positions. In contrast, it may be more reasonable that comparisons are made of the relative level differences between frequency bands due to various peaks and notches. With this approach, it is not necessary to detect peaks and notches directly, only comparisons in levels across frequency bands are needed. In this paper, we analyze level changes of median plane HRTFs in narrow frequency bands using auditory filters and inter-band correlations. These changes are investigated to clarify effects of peaks and notches on comprehensive level changes in the corresponding HRTFs.

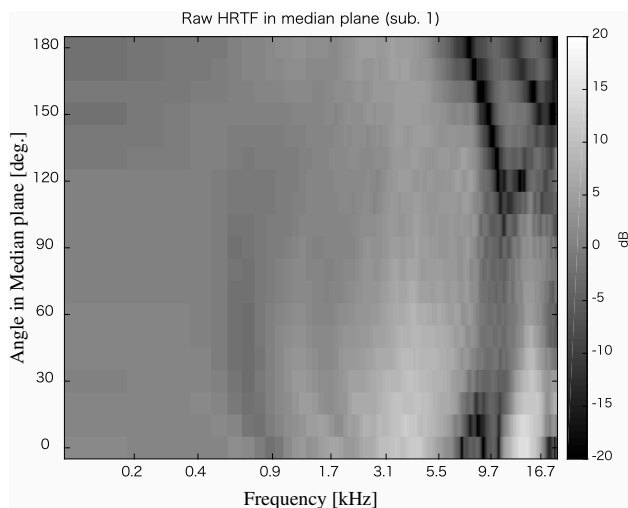


Figure 1. An example of HRTFs in median plane (HRTF 1).

2. OVERVIEW OF HRTFS FILTERED BY AUDITORY FILTERS

We analyzed 105 HRTF sets from the RIEC (Research Institution of Electrical Communication, Tohoku University) database, available in the SOFA [6] format standard. HRTFs were measured using a spherical loudspeaker array at RIEC for individual listeners. Head-related impulse responses (HRIRs) were acquired in the upper median plane from front (0°) to rear (180°) in 10° -steps. Fig. 1 shows an example median plane HRTF. We can observe deep notches from 4 kHz that shift with changed in median plane angle. These notches are considered to be due to acoustic phenomena of the pinnae. Furthermore, shallow notches can be seen at low frequencies. These notches are associated to the head being a rigid obstacle.

Each HRIR was then filtered by a band limited auditory filter. A Gammatone filter bank [7] was employed in this analysis, with 40 equivalent rectangular bandwidth (ERB) [8] over the full audible frequency range (up to 20 kHz). Output power level of the filtered HRIRs for the 19 median plane angles was calculated, resulting in 760 values ($19 \text{ angles} \times 40 \text{ bands}$) for each listener. Gammatone filter on the time domain can be expressed as follows:



© Yukio Iwaya, Brian FG Katz, Tetsu Magariyachi, and Yôiti Suzuki. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** Yukio Iwaya, Brian FG Katz, Tetsu Magariyachi, and Yôiti Suzuki. "Inter-frequency band correlations in auditory filtered median plane HRTFs", 1st EAA Spatial Audio Signal Processing Symposium, Paris, France, 2019.

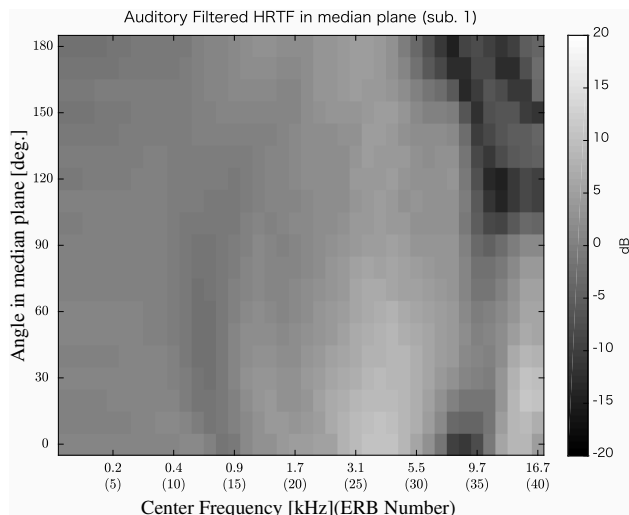


Figure 2. An example filtered median plane HRTF (HRTF 1). Numbers in parentheses indicate ERB frequency band (ID#), numbered from 1 to 40 from the lowest frequency.

$$g(f_{cj}, t) = at^{n-1} \exp(-2\pi b \text{ERB}(f_{cj})t) \cos(2\pi f_{cj}t + \phi),$$

where f_{cj} is the center frequency, and parameters in $g(f_{cj}, t)$ were set as follows:

$$\begin{aligned} a &= 1 \\ b &= 0.9826 \\ \phi &= 0 \\ n &= 10 \log_{10} 2. \end{aligned}$$

Fig. 2 shows an example of a filtered median plane HRTF. Although the frequency resolution is reduced due to the use of band-passed filters, frequency notches can still be observed.

3. CORRELATION ANALYSIS

From the output levels of each ERB band, the level change of individual ERB bands was obtained as a function of elevation angle: $P(\theta_k, f_{ci})$, ($\theta_k = 0, 10, \dots, 180^\circ$). The correlation across frequency bands for the level change as a function of angle was then calculated. The correlation coefficient value $C(f_{ci}, f_{cj})$ is expressed as:

$$C(f_{ci}, f_{cj}) = \frac{\sum_{k=1}^N \left(P(\theta_k, f_{ci}) - \overline{P(\theta, f_{ci})} \right) \left(P(\theta_k, f_{cj}) - \overline{P(\theta, f_{cj})} \right)}{\sqrt{\sum_{k=1}^N \left(P(\theta_k, f_{ci}) - \overline{P(\theta, f_{ci})} \right)^2 \sum_{k=1}^N \left(P(\theta_k, f_{cj}) - \overline{P(\theta, f_{cj})} \right)^2}}$$

This produced 39 cross-correlation values and one auto-correlation for each band with a correlation matrix of 40 bands \times 40 bands for each listener. Fig. 3 shows an

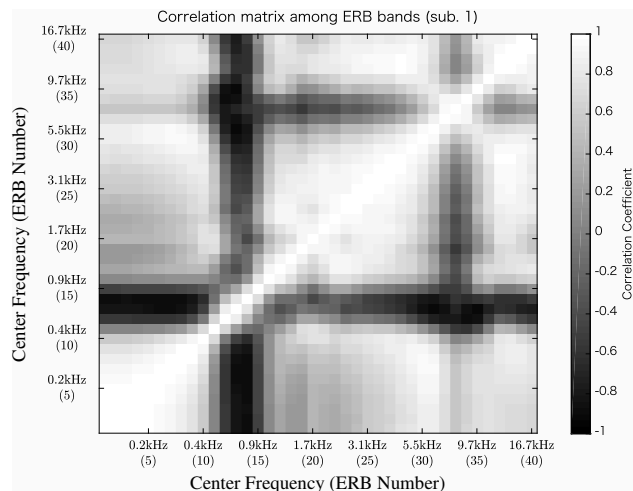


Figure 3. An example ERB band correlation matrix (HRTF 1).

example correlation matrix for an HRTF. The diagonal expresses the auto-correlation coefficient of the band, equal to unity. Fig. 3 is symmetric with respect to the diagonal. It can be observed that there are ERB-pairs with negative correlations. Level changes of these ERB-pairs are opposite each other. These tendencies were observed in all 150 HRTF sets. There are two frequency band regions which have negative correlations with respect to the remaining frequency bands. These are the frequency ranges of $\approx 0.4 - 1$ kHz and $\approx 6 - 10$ kHz.

4. ANALYSIS OF ERB BAND LEVEL CHANGES

The correlation coefficient cannot represent the absolute value of the level change. Therefore small change could be emphasized in previous analysis in Sec 3. We confirmed concrete level changes of all bands. Fig. 4 shows the actual level change of each band as a function angle. Observed level changes for low bands (ERB bands #1-#15) in the first sub-figure are small (2-3 dB), less than observed in the other bands. We can find the levels (ERB bands #11-#15) are slightly increasing in the first sub-figure of Fig.4. In contrast, the level changes in second and fourth sub-figures tends to decrease. Therefore, negative correlation would be obtained and emphasized nevertheless level change is very small in low frequency bands.

In the second sub-figure (ERB bands #16-#30), level changes tend to decrease. In the fourth sub-figure (ERB bands #37-#40) shows similar tendencies. Level change in the third sub-figure, however, relative level rises then falls. This is due to the frequency shift of the characteristic notches.

From these results, since the change of the lower frequency bands having a strong negative correlation is small, the influence on perception is considered to be small. In contrast, the level change in the bands in the third sub-figure, which have a negative correlation, is characteristic because their level changes are large and dynamic.

In the next section, based on these analysis results, the

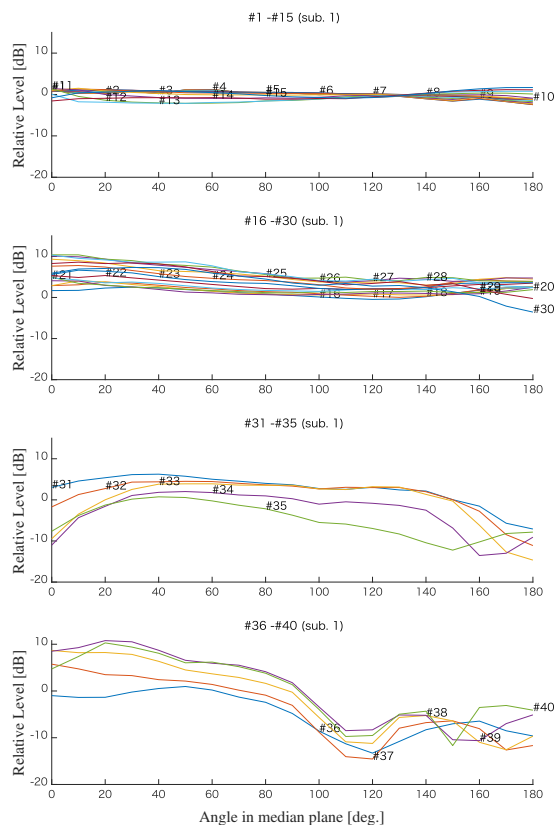


Figure 4. Level change of each ERB band as a function of median plane angle in Fig. 2. ID# indicates the ERB number (HRTF 1).

ERB bands are grouped into three regions and their relative changes are analyzed to clarify the negative correlation bands, including the characteristic notches.

5. ORIGIN OF NEGATIVE CORRELATION IN FREQUENCY REGION INCLUDING CHARACTERISTIC NOTCHES

Although there are two frequency regions with negative correlation values, the absolute level changes for the lower region (ERB bands #11-#15) is small and therefore ignored in subsequent analysis. In terms of the auditory system, it is assumed that it is reasonable to group frequency bands, which have similar level changes in median plane. Therefore, we divide the audible frequency bands into three large frequency region bands:

1. Band 1 (ERB bands #1-#30), which includes first and second sub-figures of Fig 4.
2. Band 2 (ERB bands #31-#36), in which the characteristic notches are included
3. Band 3 (ERB bands #37-#40), in which level change tends to decrease in median plane as you can see in fourth sub-figure of Fig. 4.

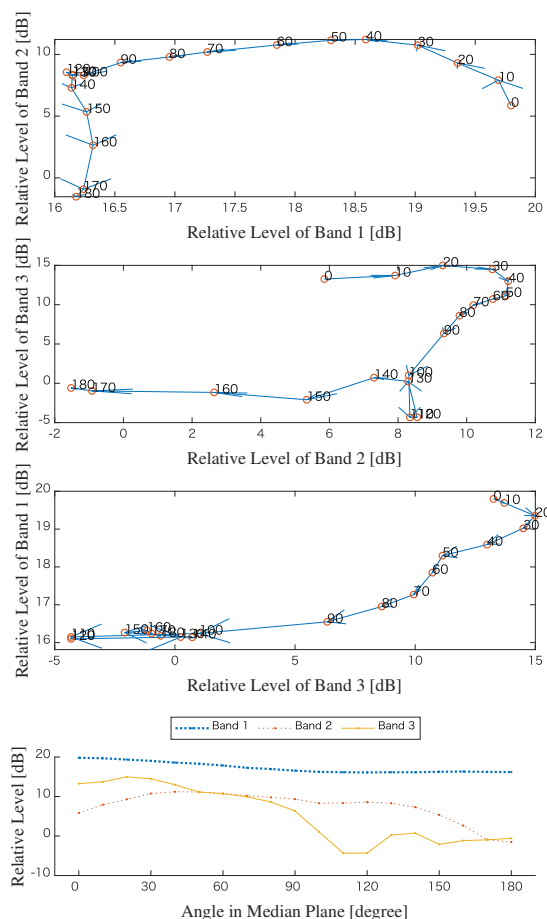


Figure 5. Scatter plot of progression of relative level between grouped bands. ID#'s along plot line indicate median plane angle. Data points are connected by arrows according to median plane angle progression. The bottom sub-figure shows relative level of grouped bands according to median plane angle (HRTF 1).

Figure 5 shows the progression of relative level as a function of angle between the three region bands. The fourth figure shows relative level of three grouped bands along to angle in median plane.

From Fig. 5, considering Band 1 & Band 2, it can be seen that there are opposite changes (levels decreasing in one band while increasing for the other) from 0° to 40° in front, and from 140° to 180° at the rear. The negative correlation is caused from these tendency. In contrast, from 50° to 130°, the level progression is positive.

From the second sub-figure (Band 2 & Band 3), the relation from 150° to 180° is slightly negative. In front, this negative relation can be partially observed from 20° to 50°. From the third sub-figure (Band 1 & Band 3), the relationship is almost always positive, because relative levels in both Band 1 and Band 3 decrease with median plane angle. The negative correlation observed for Band 2 is due to the frequency shifting of the characteristic notches.

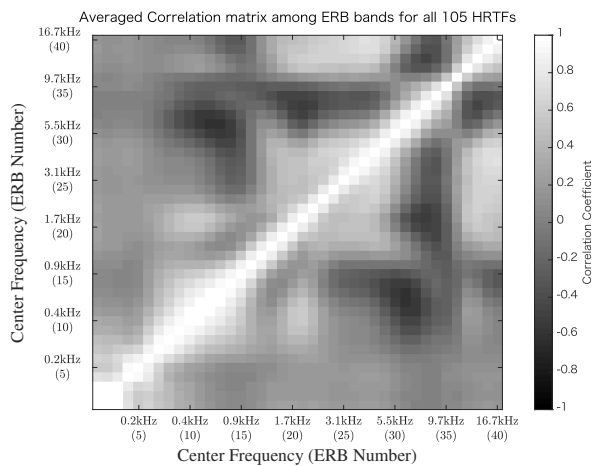


Figure 6. Correlation matrix averaged across all 105 HRTFs.

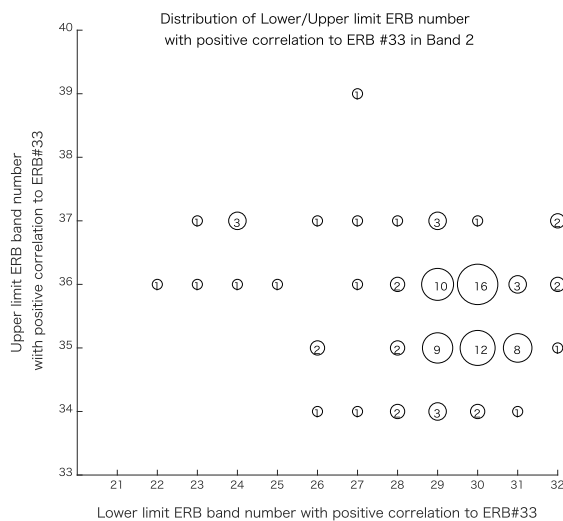


Figure 7. Distribution of upper/lower ERB band number with positive correlation to ERB #33 for all 150 HRTF sets

Fig. 6 shows the correlation matrix averaged across all 105 HRTFs. From Fig. 6, the negative correlation of level change in the grouped Band 2, in which the characteristic notch exists, was still observed. Therefore, this negative correlation in Band 2 is common in HRTF spectrum.

Furthermore, we investigated details of band width of Band 2. Fig. 7 shows distribution of upper and lower limit ERB band number of Band 2 with positive correlation to ERB #33, which is included in the grouped Band 2, for all HRTFs. Number in each circle indicates number of HRTF sets in the condition. We can see that the spread of positive correlations is different among HRTFs. In other words, the band width of Band 2 differs among individuals. These are caused from the fact that the frequency shift of characteristic notches differs among individuals.

6. CONCLUSION

In this paper, upper median plan HRTFs were analyzed and inter-band correlations were investigated to examine relationships between spectral cues associated with sound localization in the median plane. Results showed that the audible frequency could be divided into three frequency bands. The middle band, comprising the characteristic pinnae notches, showed a progression of level as a function of elevation angle which had a negative correlation relative to the upper and lower bands.

It is hypothesized that this relative information could be exploited by the auditory system to provide elevation information. If shown in perceptual studies, this could indicate that we do not need to detect frequency notches directly. Namely, we may compare relative level across three large frequency bands. Therefore, listening test are the next step in our work to investigate this hypothesis.

7. ACKNOWLEDGEMENTS

This study was partially supported by JSPS KAKENHI (19K12286, 19001004).

8. REFERENCES

- [1] K. Iida, M. Itoh, A. Itagaki, and M. Morimoto, "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," *Applied Acoustics*, vol. 68, pp. 835–850, 2007.
- [2] M. Morimoto and H. Aokata, "Localization cues of sound sources in the upper hemisphere," *J Acoust Soc Jpn (E)*, vol. 5, no. 3, pp. 165–173, 1984.
- [3] J. Hebrank and D. Wright, "Spectral cues used in the localisation of sound sources on the median plane," *Journal of the Acoustical Society of America*, vol. 56, pp. 1829–1834, 1974.
- [4] R. Greff and B. Katz, "Perceptual evaluation of HRTF notches versus peaks for vertical localisation," in *Intl. Cong. on Acoustics 19*, (Madrid, Spain), pp. 1–6, 2007.
- [5] J. Blauert, *Spatial hearing: The Psychophysics of Human Sound Localization*. Cambridge, Massachusetts, US: The MIT Press, 1996.
- [6] P. Majdak, Y. Iwaya, *et al.*, "Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions," in *roc. 2013 International AES Convention*, (Rome, Italy), pp. 12–13, 2013.
- [7] R. Patterson, M. Allerhand, and C. Giguere, "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.*, vol. 98, pp. 1890–1894, 1995.
- [8] B. Glasberg and B. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing Research*, vol. 47, pp. 103–138, 1990.