Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

Visualization and dereverberation of head-related transfer function based on spatio-temporal frequency analysis

Yasuko Morimoto (1), Takanori Nishino (2) and Kazuya Takeda (1)

(1) Graduate School of Information Science, Nagoya University, Nagoya, Aichi, Japan
(2) Graduate School of Engineering, Mie University, Tsu, Mie, Japan

PACS: 43.66.Pn, 43.60.Gk

ABSTRACT

A new method of visualizing characteristics of head-related transfer function (HRTF) is proposed. The proposed visualization method can illustrate the HRTFs and other factors such as the reverberations separately. The HRTF is an acoustic transfer function between a sound source and the ear canal entrance, and it is defined as a function on time and direction of the sound source. Since the HRTF depends on sound source direction and subject, HRTFs are usually measured with a dummy head or a human. Measured HRTFs are generally visualized by a figure whose axes correspond to the angle of sound source and the temporal frequency. The conventional figure can illustrate the difference in HRTFs among the directions of sound source, and most previous works employed frequency analysis in the time domain while emphasizing the time variation. In this paper, the measured HRTFs are analyzed with spatio-temporal frequency analysis and used to examine the efficiency of the proposed visualization method. The spatio-temporal frequency analysis can visualize and analyze the characteristics of HRTFs using the spectrum calculated by two-dimensional Fourier transform in time and space. In our experiments, the theoretical property of the spatio-temporal frequency characteristic is investigated and the influence of reverberation in the measurement environment is also examined. Moreover, a dereverberation method is proposed. From the results, the characteristics of HRTFs are mostly concentrated in a specific frequency band, and the proposed visualization method is efficient for illustrating the HRTF and other factors such as reverberation and reflection waves. The dereverberation method decreases the average reverberation time from 382.4 to 316.0 ms in a reverberant condition.

INTRODUCTION

A head-related transfer function (HRTF) is an acoustic transfer function that includes an interaural level difference, an interaural time difference, and a frequency characteristic. The HRTFs can control a sound image by convolving with the source signal. There have been many studies on HRTF, and its many applications include 3D sound reproduction systems and examinations of the relationship between HRTF and a human's sound localization ability.

Since the HRTF depends on sound source direction and subject, HRTFs are usually measured with a dummy head or a human (Gardner and Martin 1995; Nishino 1999; Algazi et al. 2001; Sutou 2003). Measured HRTFs are generally visualized by a figure whose axes correspond to the angle of sound source and the temporal frequency. The conventional figure can illustrate the difference in HRTFs among the directions of sound source, and most previous works employed frequency analysis in the time domain while emphasizing the time variation. However, the acoustic transfer function defined in the wave equation is a function of time and space. Based on this definition, not only the time variation but also the space variation should be analyzed. Space variation is caused in HRTFs when the sound source or the listener moves. By analyzing HRTFs over both variations, the relation between HRTF and space can be clarified to simplify the extraction of the factors attributed to space and to process the signals by moving sound sources. Moreover, HRTF measurement should be conducted in a free sound field environment; however, this is difficult because special equipment and facilities such as an anechoic room are needed. When the HRTF measurement was conducted in a non-free sound field, the measured HRTFs contained various acoustic factors besides HRTFs. Therefore, the data must be analyzed to separate and extract the factors, and it is difficult to distinguish the HRTF and other extracts by the conventional method.

In this paper, we analyzed the measured HRTFs with spatio-temporal frequency analysis (Morimoto, Nishino, and Takeda 2009) and examined the efficiency of the proposed visualization method. The spatio-temporal frequency analysis is done in time and space. While such analysis is mostly used in the visual neurology field, some studies have applied it to acoustic waves. Ajdler *et al.* calculated the specific frequency band concentrated by HRTF spectra in the spatio-temporal frequency domain (Ajdler et al. 2008). They determined the number of measurement positions and investigated the interpolation of HRTFs. We applied the same analysis as the previous work to measured HRTFs under different measurement environments. From the similarities and differences obtained by analyses, we investigated whether the environmental factors affect the spatio-temporal frequency characteristics. Moreover, a dereverberation method that suppresses the late reverberation is proposed and evaluated. The outline of our paper is as follows. First, the method of the spatio-temporal frequency analysis is described and applied to measured HRTFs. Then, the reverberation components, the difference among the public HRTF databases, and the influence of pinnae are examined. The dereverberation method is proposed and evaluated based on observations of the influence of reverberation components.

SPATIO-TEMPORAL FREQUENCY ANALYSIS

An HRTF, which is an acoustic transfer function between a sound source and an ear, can be defined by sound source direction θ and time t. Spatio-temporal frequency analysis for HRTFs $h(\theta_m, t_n)$ is calculated by a two-dimensional Fourier transform on time and space:

$$\underline{H}(\phi, f) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} h(\theta_m, t_n) e^{\left\{-j2\pi\left(\frac{\phi m}{M} + \frac{fn}{N}\right)\right\}}, \quad (1)$$

$$\underline{H}(M-\phi, N-f) = \underline{H}^*(\phi, f) \quad (0 < \phi \leq M/2,\ 0 < f \leq N/2), \quad (2)$$

where f is temporal frequency. ϕ denotes the periodicity of HRTFs in each temporal frequency when the sound source direction is changed. ϕ is called "spatial frequency" or "angular wave number" (Buchner, Spors, and Kellermann 2004). In this paper, ϕ is called azimuthal frequency because HRTFs on the horizontal plane were used. ϕ is related to the periodicity when the sound source makes one rotation around a subject and its unit is rad^{-1} based on the definition of the Fourier transform. M is the number of angular sampling points used for the sound source direction, and N is the length of HRTFs. A complex conjugate is denoted by $*$ in eq. (2). These equations indicate that the spatio-temporal frequency characteristics are the origin symmetry. Therefore, the following figures in this paper show only the positive temporal frequency domain.

Figure 1 shows the spatio-temporal frequency characteristics of the impulse responses that were calculated with the spherical model, which is the simplest model of a head. The horizontal and vertical axes represent azimuthal frequency ϕ rad^{-1} and temporal frequency f kHz, respectively. In this calculation, a sound source and received point are on the horizontal plane, and the distance between them is 1.0 m. The radius is 0.09 m. Most of the components are concentrated in a triangular region. Because a sound wavelength of the low temporal frequency is long, the variability of its characteristics is small when the received point is changed. Therefore, there are fewer high-azimuthal frequency components in the low temporal frequency. However, since a sound wavelength is short in the high temporal frequency, azimuthal frequency components exist in a wide bandwidth. The spatio-temporal frequency characteristics of HRTFs are expected to be similar to those in this figure.

If the HRTF measured at sound source direction θ_m is considered the m-th measurement, the spatio-temporal frequency analysis shows us periodicity related to not only a change of sound source direction but also a repeat of the measurement. Therefore, the spatio-temporal frequency analysis could extract a component of the diffused noise and variations of the measurement system such as a loudspeaker.

ANALYSIS OF IMPULSE RESPONSES

Comparison of public HRTF databases

There are many public HRTF databases measured under different measurement conditions; however, there has been little research comparing them. We used the HRTF databases constructed by four research groups (Gardner and Martin 1995; Nishino 1999; Algazi et al. 2001; Sutou 2003). The measurement conditions are shown in Table 1. An important difference between B-1 and B-2 is the measurement room. Acoustic conditions of B-2 were better than B-1. C-1 and C-2 were measured with the same dummy head, but C-1 used large pinnae while C-2 used small ones. D-2 was measured when the room temperature varied from 14.0 to 21.5 centigrade.

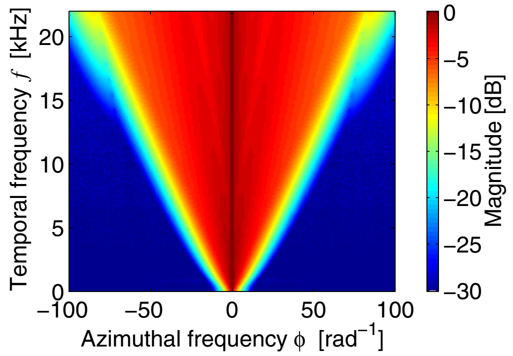


Figure 1: Two-dimensional spectrum using spherical model in spatio-temporal frequency domain.

Figure 2 shows the two-dimensional spectra obtained by the right ear's HRTFs of these databases. The horizontal and vertical axes represent azimuthal frequency ϕ rad^{-1} and temporal frequency f kHz, respectively. Colors indicate the magnitude in dB. Spectra were normalized by the maximum value in each condition. Most of the components are concentrated in a triangular region. Although a triangle appears in all figures, they are not all the same. The figures have relatively clear triangle shapes except for Figs. 2(b) and (g). From these results, it seems that A, B-2, C-1, C-2, and D-1 were measured in well-conditioned environments.

The HRTFs of B-1 and B-2 were measured with the same equipment at the same institute, except for the measurement room. Since the room for B-2 is larger than that for B-1, their differences may be attributed to the ceiling and wall reflections and to background noise. Although D-2 was measured in an anechoic room, the spatio-temporal frequency characteristics are deformed. This indicates that the two-dimensional spectra are influenced by a change in room temperature.

The HRTF should be measured in a free sound field and under rigid conditions. However, it is difficult to achieve these conditions. From the results, the spatio-temporal frequency analysis can display extra factors that appear in an area opposite the triangles. This clarifies the investigated method's ability to detect measurement errors. This method is also useful for comparing and confirming the HRTFs in projects of designed to compare measurement systems (Katz and Begault 2007).

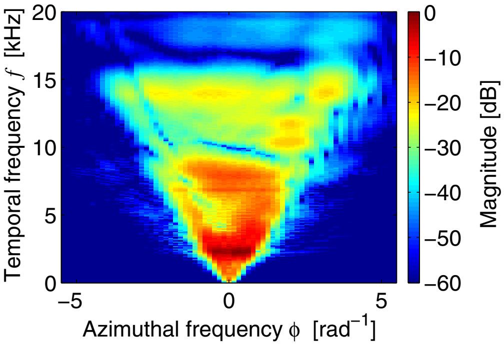
Reverberation

We investigated the influence of reverberation on the spatio-temporal frequency characteristics. Two binaural room impulse responses (BRIRs), measured with different reverberation times in a variable reverberation room, were compared. The distance from a sound source (BOSE Acoustimass) to a dummy head (B&K 4128) was 1 m. The sampling frequency was 48 kHz, and the measurement direction intervals were 5°. $h_{R1}(\theta, t)$ and $h_{R2}(\theta, t)$ denote the impulse responses measured when the reverberation time was set to 151 ms and 459 ms, respectively. The less reverberant condition is called R-1, and the other is called R-2. $\underline{H}_{R1}(\phi, f)$ and $\underline{H}_{R2}(\phi, f)$ are two-dimensional spectra obtained by applying the spatio-temporal frequency analysis to $h_{R1}(\theta, t)$ and $h_{R2}(\theta, t)$, respectively.

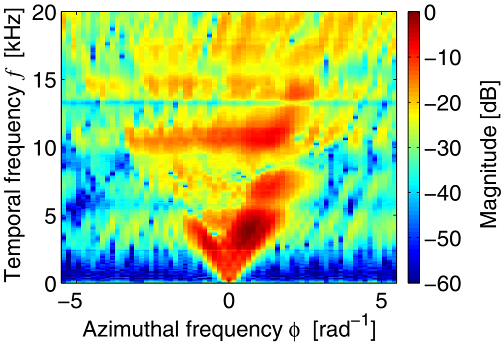
To compare the two characteristics, their spectral differences (SD) were calculated by:

$$\underline{\mathrm{SD}}(\phi, f) = 20\log_{10}\frac{|\underline{H}_{R2}(\phi, f)|}{|\underline{H}_{R1}(\phi, f)|} \quad [\mathrm{dB}]. \quad (3)$$

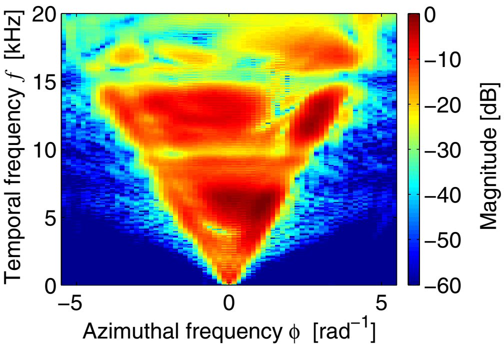
The spectral difference obtained by the conventional method



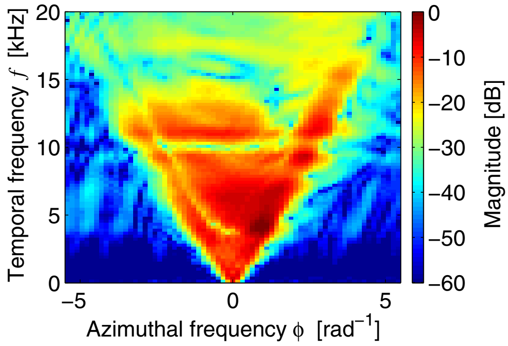
(a) A (MIT, KEMAR) (Gardner and Martin 1995)



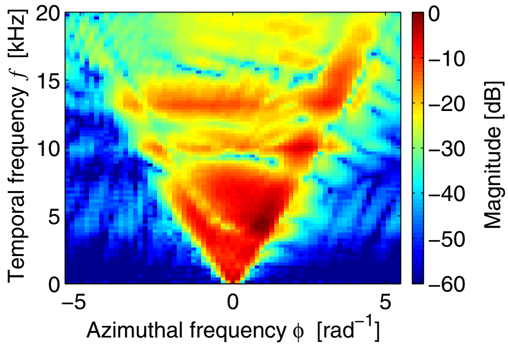
(b) B-1 (Nagoya Univ., B&K) (Nishino 1999)



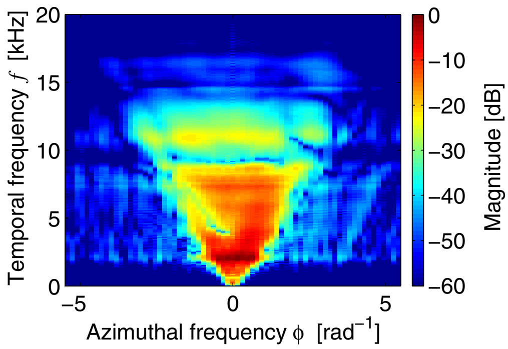
(c) B-2 (Nagoya Univ., B&K) (Nishino 1999)



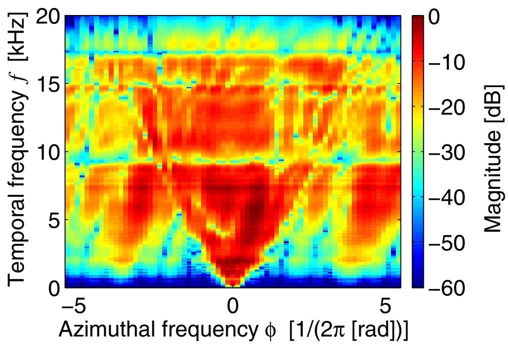
(d) C-1 (CIPIC, KEMAR, large pinna) (Algazi et al. 2001)



(e) C-2 (CIPIC, KEMAR, small pinna) (Algazi et al. 2001)



(f) D-1 (Nagaoka Univ. of Tech., KOKEN) (Sutou 2003)



(g) D-2 (Nagaoka Univ. of Tech., KOKEN) (Sutou 2003)

Figure 2: Two-dimensional spectrum of measured HRTFs with dummy head in spatio-temporal frequency domain.

Table 1: Measurement conditions of HRTF databases. Database A is described in (Gardner and Martin 1995), B-1 and B-2 are (Nishino 1999), C-1 and C-2 are (Algazi et al. 2001), and D-1 and D-2 are (Sutou 2003).

	A	B-1	B-2	C-1	C-2	D-1	D-2
Dummy head	KEMAR	B&K	B&K	KEMAR	KEMAR	KOKEN	KOKEN
Distance of sound source [m]	1.4	1.52	1.2	1.0	1.0	1.5	1.5
Length of HRTF [point]	512	512	512	200	200	600	600
Sampling frequency [kHz]	44.1	44.1	48	44.1	44.1	48	48
Measurement room	anechoic	soundproof	soundproof	soundproof	soundproof	anechoic	anechoic

using one-dimensional Fourier transform was also calculated:

$$\mathrm{SD_c}(\theta,f) = 20\log_{10}\frac{|\,H_{\mathrm{R2}}(\theta,f)\,|}{|\,H_{\mathrm{R1}}(\theta,f)\,|} \quad [\mathrm{dB}], \qquad (4)$$

where H_{R1} and H_{R2} are magnitude responses of h_{R1} and h_{R2} obtained by the traditional frequency analysis. If an absolute value of an SD score is a large, this indicates the influence of reverberation.

In this experiment, the impulse responses were originally measured with a 65,536-point length (about 1.365 s), and then this was truncated to a 512-point length (about 0.010 s). The results are shown in Figures 3 and 4, where the warm colors represent the large spectral difference. In Fig. 3, these are concentrated on the opposite areas from the triangle in the result of the spherical model. However, it is difficult to distinguish the HRTFs and the reverberations in Fig. 4. Since the HRTFs depend on the sound source direction, the HRTF is a directional acoustic transfer function. On the other hand, other acoustic components such as reflection waves and reverberation are diffusive. Therefore, the proposed visualization method can illustrate the HRTFs and extra features such as the reverberations separately, and the results suggest that we can easily verify whether HRTFs were measured in a well-conditioned environment.

Influence of pinnae

This experiment examined how pinnae influence the spatio-temporal frequency characteristic. Figures 2(d) and 2(e) show that the difference that is related to size of pinna is small. Therefore, the HRTFs were measured using a dummy head (KEMAR) with or without pinnae and then evaluated. In the case of not having pinnae, the gaps made by removing the pinnae were filled with clay. Other measurement conditions were the same as those in the experiments of the previous section. Figures 5 and 6 show two-dimensional spectra of measured HRTFs in the cases of the dummy head with and without pinnae, respectively. The horizontal and vertical axes represent azimuthal frequency ϕ rad^{-1} and temporal frequency f kHz, respectively. Colors indicate the magnitude in dB. Spectra were normalized by the maximum value in each condition. In Fig. 5, the components are especially concentrated in the triangle area of absolute temporal frequency lower than about 8 kHz. In contrast, in Fig. 6 the same feature is not shown and the spectrum is wavy regardless of temporal frequency. This phenomenon seems to be caused by a sound gathering effect of the pinnae, and it appears clearly on the two-dimensional spectrum in Fig. 5.

DEREVERBERATION

In this section, the dereverberation method using two-dimensional filtering is proposed, since the spatio-temporal frequency analysis can distinguish HRTFs and other extra factors.

Method

Components related to the direct wave were concentrated in the triangles in the spatio-temporal frequency domain; however, extra factors such as reflection waves and reverberation appeared in other regions. This suggests that the reverberation could be suppressed by reducing the spectra in the area opposite the triangle. Therefore, a suppression method using two-dimensional filtering is proposed and evaluated.

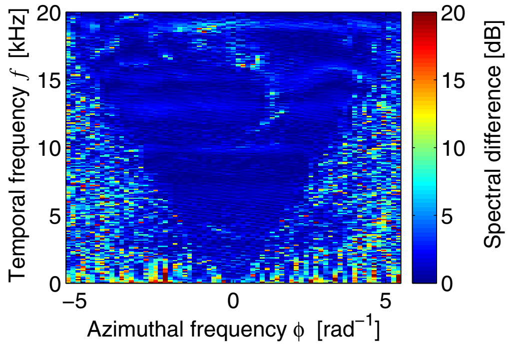


Figure 3: Absolute values of spectral difference between $\underline{H}_{\mathrm{R1}}(\phi,f)$ and $\underline{H}_{\mathrm{R2}}(\phi,f)$, whose reverberation times differ (right ear).

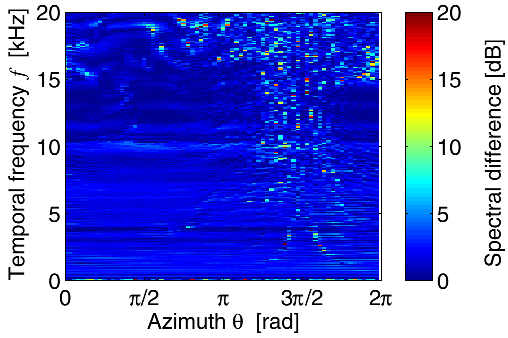


Figure 4: Absolute values of spectral difference between $H_{\mathrm{R1}}(\theta,f)$ and $H_{\mathrm{R2}}(\theta,f)$ obtained by conventional one-dimensional Fourier transform (right ear).

A two-dimensional filter for reducing the spectra related to extra factors was designed using R-1 and R-2, which were measured in different reverberant conditions. The suppressed area is determined by the spectral differences $\underline{\mathrm{SD}}(\phi,f)$ calculated with equation (4). In an area with a small SD score, we can assume that there are components involving direct waves. On the other hand, an area with a large SD score shows a greater influence of reverberations. Therefore, boundaries can be fixed by SD scores. In this method, the pass-band is 1 and the stop-band is β $(0 \le \beta < 1)$. The transition band between the pass-band and stop-band is

$$\arctan\left(\frac{f/1000}{2\pi\phi}\right) = 20^{\circ}. \qquad (5)$$

A coefficient of the transition band changes linearly. The designed two-dimensional filter is applied in the complex spec-

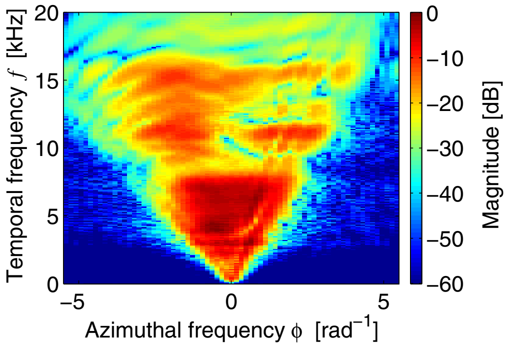


Figure 5: Two-dimensional spectrum of HRTFs when dummy head with pinnae were used for measurement (right ear).

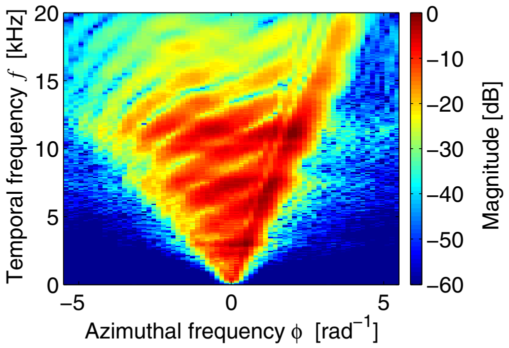


Figure 6: Two-dimensional spectrum of HRTFs when dummy head without pinnae were used for measurement (right ear).

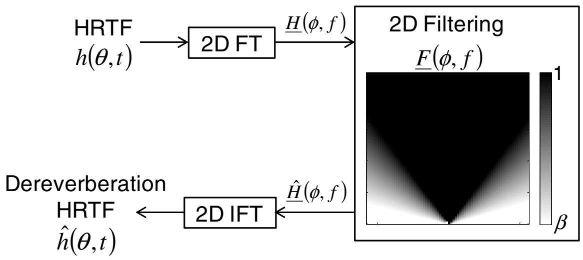


Figure 7: Diagram of proposed suppression method.

trum domain. Impulse responses that reduce the reverberation $\hat{h}(\theta,t)$ are obtained by inverse two-dimensional Fourier transform of the filtered spectrum $\underline{\hat{H}}(\phi,f)$. Figure 7 shows a diagram of the proposed method.

Experimental conditions

In our experiments, BRIRs were used instead of HRTFs because they have many reverberation components. Experimental results are evaluated by the reverberation time and signal-to-deviation ratio (SDR). The reverberation time evaluates the degree of suppression, and the SDR evaluates the structural distortion of impulse responses. The reverberation time is calculated at every azimuth θ_m by Schroeder integration (Schroeder 1965). In this experiment, the length of BRIRs was 1.365 s; however, signals after 1 s were quite small. Therefore, we used signals of 1 s ($N = 48000$) length to calculate the reverberation time.

SDR is also calculated in every azimuth θ_m by equation (6):

$$\mathrm{SDR}(\theta_m) = 10\log_{10}\frac{\sum_{n=0}^{N-1}\{h(\theta_m,t_n)\}^2}{\sum_{n=0}^{N-1}\{h(\theta_m,t_n)-\hat{h}(\theta_m,t_n)\}^2}\ \text{[dB]},\qquad(6)$$

where $h(\theta_m,t_n)$ is the original impulse response and $\hat{h}(\theta_m,t_n)$ is the processed impulse response. A large SDR indicates that the processed impulse response has a small time-structural distortion compared with the original impulse response. In this experiment, SDRs were calculated for three intervals: The first interval related to direct waves is from 0 to 11 ms, the second related to the early reflections is from 50 to 61 ms, and the third related to the late reverberation is 150 to 161 ms.

Results

In our experiments, the stop-band coefficient β is 0.1, as determined by the results of a preliminary experiment. Table 2 shows the reverberation time of the original and processed impulse responses. Table 3 shows average SDR.

The reverberation time of the processed signals was shorter than that of the original signals. The resulting SDRs indicate that the proposed method has less influence on the direct waves and reduces the late reverberation.

Table 2: Results of dereverberation evaluated by reverberation time.

		Average	Standard Deviation
R-1	Original	86.2	10.5
	Processed	71.8	5.8
R-2	Original	382.4	53.8
	Processed	316.0	36.5

Table 3: Results of dereverberation evaluated by signal-to-deviation ratio.

		Average	Standard Deviation
R-1	Direct wave	36.1	8.4
	Early refrection	11.4	2.7
	Late reverberation	0.2	1.3
R-2	Direct wave	34.7	8.7
	Early refrection	16.4	1.2
	Late reverberation	5.8	2.4

CONCLUSION

In this paper, we proposed a new visualization method of HRTF using the spatio-temporal frequency characteristics. We examined the differences in reverberation time in the measurement environment. Measured HRTFs from databases were also analyzed and compared. The reverberation influences were distributed on different areas from the results of a spherical model. The results show that the two-dimensional spectra varied with the HRTF databases. Consequently, the spatio-temporal frequency analysis can represent not only components of HRTF but also other extra factors such as reverberation and changes in room temperature. A dereverberation method was also proposed. The results indicate that the proposed method can suppress the late reverberation without influencing the direct wave components. Future works include an investigation of individuality and how to controll the reverberation.

Acknowledgment This work was supported by a Grants-in-Aid for Scientific Research (No. 22700171).

REFERENCES

Ajdler, T. et al. (2008). "Sound field analysis along a circle and its applications to HRTF interpolation". *J. Audio Eng. Soc.* 56.3, pp. 156–175.

Algazi, V.R. et al. (2001). "The CIPIC HRTF Database". *Proc. WASPAA'01*, pp. 99–102. URL: http://interface.cipic.ucdavis.edu/.

Buchner, H., S. Spors, and W. Kellermann (2004). "Wave-domain adaptive filtering: acoustic echo cancellation for full-duplex systems based on wave-field synthesis". *Proc. ICASSP 2004* IV, pp. 117–120.

Gardner, B. and K. Martin (1995). "HRTF measurements of a KEMAR". *J.A.S.A.* 97(6), pp. 3907–3908. URL: http://sound.media.mit.edu/resources/KEMAR.html.

Katz, B. and D. Begault (2007). "Round robin comparison of HRTF measurement systems: preliminary results". *Proc. ICA 2007*.

Morimoto, Y., T. Nishino, and K. Takeda (2009). "Analysis of head related transfer functions based on the spatio-temporal frequency characteristic". *Proc. AES Tokyo convention*.

Nishino, T. (1999). *Head related transfer function database*. URL: http://www.sp.m.is.nagoya-u.ac.jp/HRTF/.

Schroeder, M. R. (1965). "New method of measuring reverberation time". *J. Acoust. Soc. Am.* 37.3, pp. 409–412.

Sutou, H. (2003). *Shimada Laboratory HRTF Database*. URL: http://audio.nagaokaut.ac.jp/hrtf/.