PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

The limitation of static transaural reproduction with two frontal loudspeakers

Lulu LIU[1]; Bosun XIE[2]

[1] Acoustic lab, School of Physics and Optoelectronics, South China University of Technology, Guangzhou, China

[2] Acoustic lab, School of Physics and Optoelectronics, South China University of Technology, Guangzhou, China

ABSTRACT

Incorporated HRTF-based binaural synthesis and cross-talk cancellation, transaural reproduction with two frontal loudspeakers is theoretically able to duplicate the binaural pressures of an actual sound source, and then is expected to be able to recreate virtual source at arbitrary direction. However, many experiments have indicated that for static transaural reproduction, in which transaural processing is fixed regardless of head turning, the perceived virtual source is usually limited in the frontal-horizontal plane. The reason may be that, in static transaural reproduction, the dynamic cue for front-back and vertical localization is omitted, while the high-frequency spectral cue is unstable against head movement. To validate this hypothesis, a psychoacoustic experiment is conducted in present work. A dynamic transaural system with two frontal loudspeakers is used in the experiment. According to the contemporary orientation of head detected by a head-tracker, the system changes the transaural processing so that binaural pressures in reproduction follow the turning of subject's head. The experimental results indicate that, by incorporating the dynamic cue caused by head turning, dynamic tranaural reproduction is able to recreate virtual source at front, back and vertical directions. Therefore, the limitation in conventional static transaural reproduction is indeed caused by the omitting of dynamic cue.

Keywords: Transaural reproduction, Vertical localization, Dynamic cue

1. INTRODUCTION

Binaural pressures include the main information for auditory localization. Based on HRTF-based binaural synthesis, binaural reproduction or virtual auditory display (VAD) duplicates the binaural pressures caused by an actual sound source and then recreates perception of virtual source localization in three dimensional space through headphone reproduction (1).

Incorporated HRTF-based binaural synthesis and cross-talk cancellation, transaural reproduction with two frontal loudspeakers is theoretically able to duplicate the binaural pressures of an actual sound source, and then is expected to be able to recreate virtual source at arbitrary direction. Some authors reported that, under a series of critical conditions (e.g., individualized HRTF processing, restriction of head movement, reproduction in anechoic rooms, etc.), a pair of frontal loudspeakers can recreate perceived virtual sources at all horizontal or three-dimensional directions for listeners, to some extent (2). However, more experimental results indicate that, in static transaural reproductions (especially these with nonindividualized HRTF processing), perceived virtual source positions are usually limited in the region of frontal-horizontal quadrants (3). The virtual sources intended for rear-horizontal quadrants or high elevations are often perceived at the frontal-horizontal quadrants with the same cone of confusion.

Actually, in the case of an actual sound source, both high-frequency spectral cue included in binaural pressures and dynamic variation of binaural pressures caused by head turning contribute to front-back and vertical localization (4,5). In static transaural reproduction, the transaural processing is fixed regardless of head turning and therefore the dynamic cue is omitted or unmatched. Moreover, the high-frequency spectral cue in loudspeaker reproduction is unstable against head movement due to the short wavelength. Therefore, it is reasonable to hypothesize or deduce that the aforementioned

[1] 348802659@qq.com
[2] Corresponding to: phbsxie@scut.edu.cn

limitation in static transaural reproduction with two front loudspeakers is caused by the unmatched dynamic cue and unstable high-frequency spectral cue. To test this hypothesis, the dynamic variation of interaural time difference (ITD) caused by head turning in transaural reproduction is analyzed in present work. Virtual source localization experiments for both static and dynamic transaural reproductions through two frontal loudspeakers are conducted. The analysis and experimental results validate the hypothesis.

2. PRINCIPLE OF TRANSAURAL REPRODUCTION

There are various methods to derive the loudspeaker signals for transaural reproduction with two frontal loudspeakers in the horizontal plane. The conventional method is to cascade the HRTF-based binaural synthesis and cross-talk processing (6). Here, an alternative but mathematically equivalent method is outlined (7).

For convenience in analyzing vertical or elevation localization, the interaural polar coordinate is used in the present work. The origin of coordinate is located at the center of the head. Source position is specified by (r, Θ, Φ). Where $0 \le r < \infty$ denotes the source distance; $-90^\circ \le \Theta \le 90^\circ$ denotes the interaural polar azimuth, that is, the angle between the directional vector of the sound source and the median plane, with Θ = -90º, 0º, and 90º being the left direction, median plane, and right direction, respectively. A constant Θ represents the cone of confusion. The $-90^\circ \le \Phi < 270^\circ$ denotes the interaural polar elevation, that is, the angle between the projection of the directional vector of the source to the median plane and the frontal axis, with Φ = -90º, 0º, 90º, and 180º being the below, front, above, and back directions, respectively.

For an actual or target virtual source at far-field distance ($r \ge 1.2$ m) and direction (Θ_S, Φ_S), the binaural pressures in the frequency domain can be calculated by filtering the input stimulus E(f) with a pair of corresponding HRTFs H_L (Θ_S, Φ_S, f) and H_R (Θ_S, Φ_S, f):

$$\begin{bmatrix} P_L \\ P_R \end{bmatrix} = \begin{bmatrix} H_L(\Theta_S, \Phi_S, f) \\ H_R(\Theta_S, \Phi_S, f) \end{bmatrix} E(f) \tag{1}$$

For transaural reproduction with two frontal loudspeakers, two loudspeakers are arranged at far-field distance ($r \ge 1.2$ m) and directions (Θ_L, $\Phi_L = 0^\circ$), (Θ_R, $\Phi_R = 0^\circ$) respectively. Let $E_L(f)$ and $E_R(f)$ be the loudspeaker signals, $H_{LL}(\Theta_L, \Phi_L, f)$, $H_{RL}(\Theta_L, \Phi_L, f)$, H_{LR} (Θ_R, Φ_R, f) and $H_{RR}(\Theta_R, \Phi_R, f)$ denote the four acoustic transfer functions from two loudspeakers to two ears. The reproduced binaural sound pressures at the two ears are given by:

$$\begin{bmatrix} P'_L \\ P'_R \end{bmatrix} = \begin{bmatrix} H_{LL}(\Theta_L, \Phi_L, f) & H_{LR}(\Theta_R, \Phi_R, f) \\ H_{RL}(\Theta_L, \Phi_L, f) & H_{RR}(\Theta_R, \Phi_R, f) \end{bmatrix} \begin{bmatrix} E_L(f) \\ E_R(f) \end{bmatrix} \tag{2}$$

Letting Eq (2) be equal to Eq.(1), that is, the binaural pressures in transaural reproduction be equal to those of target, the loudspeaker signals can be found as:

$$E'_L(\Theta_S, \Phi_S, f) = G_L(\Theta_S, \Phi_S, f) E(f) \qquad E'_R(\Theta_S, \Phi_S, f) = G_R(\Theta_S, \Phi_S, f) E(f) \tag{3}$$

Where $G_L(\Theta_S, \Phi_S, f)$ and $G_R(\Theta_S, \Phi_S, f)$ are the responses of a pair transaural filters which depend on directions of target source and loudspeakers with respective to head.

$$\begin{aligned} G_L(\Theta_{S,}\Phi_{S,}f) &= \frac{H_{RR}(\Theta_R, \Phi_R, f) H_L(\Theta_S, \Phi_S, f) - H_{LR}(\Theta_R, \Phi_R, f) H_R(\Theta_S, \Phi_S, f)}{H_{LL}(\Theta_L, \Phi_L, f) H_{RR}(\Theta_R, \Phi_R, f) - H_{LR}(\Theta_R, \Phi_R, f) H_{RL}(\Theta_L, \Phi_L, f)} \\ G_R(\Theta_{S,}\Phi_{S,}f) &= \frac{-H_{RL}(\Theta_L, \Phi_L, f) H_L(\Theta_S, \Phi_S, f) + H_{LL}(\Theta_L, \Phi_L, f) H_R(\Theta_S, \Phi_S, f)}{H_{LL}(\Theta_L, \Phi_L, f) H_{RR}(\Theta_R, \Phi_R, f) - H_{LR}(\Theta_R, \Phi_R, f) H_{RL}(\Theta_L, \Phi_L, f)} \end{aligned} \tag{4}$$

Therefore, by filtering the input stimulus with a pair of transaural filters, transaural reproduction is able to control the binaural pressures caused by two loudspeakers so that they are equal to those caused by an actual source. This is the basic principle of transaural reproduction with two front loudspeakers.

In practical uses, appropriate equalization algorithm may be supplemented to the transaural synthesis to reduce the perceived timbre coloration in reproduction. The timbre equalization is based on the fact that in two frontal loudspeaker reproduction, the perceived virtual source direction is dominated by interaural cues (especially ITD) and is limited to frontal-horizontal quadrants. The interaural cues are controlled by the relative, rather than the absolute, magnitude and phase of left and right loudspeaker signals. Scaling both loudspeaker signals with identical frequency-dependent coefficients does not alter their relative magnitude and phase, or the perceived virtual source azimuth. However, this manipulation alters the overall power spectra of loudspeaker signals and therefore equalizes timbre. Of course, equalization algorithms may alter the spectra of binaural pressures. For constant-power equalization algorithms (1), the responses of the transaural synthesis filters are equalized by their root mean square (RMS). That is, the $G_L(\Theta_S, \Phi_S, f)$ and $G_R(\Theta_S, \Phi_S, f)$ in Eq.(4) are substituted by the following $G'_L(\Theta_S, \Phi_S, f)$ and $G'_R(\Theta_S, \Phi_S, f)$:

$$G'_L(\Theta_S,\Phi_S,f)=\frac{G_L(\Theta_S,\Phi_S,f)}{\sqrt{|G_L(\Theta_S,\Phi_S,f)|^2+|G_R(\Theta_S,\Phi_S,f)|^2}}$$
$$G'_R(\Theta_S,\Phi_S,f)=\frac{G_R(\Theta_S,\Phi_S,f)}{\sqrt{|G_L(\Theta_S,\Phi_S,f)|^2+|G_R(\Theta_S,\Phi_S,f)|^2}} \tag{5}$$

After equalization, the loudspeaker signals given by Eq.(3) and Eq.(5) satisfy the constant-power spectral relationship,

$$|E'_L(\Theta_S,\Phi_S,f)|^2+|E'_R(\Theta_S,\Phi_S,f)|^2=|E(f)|^2 \tag{6}$$

Therefore, the overall power spectra of loudspeaker signals are equal to those of the input stimulus, thereby reducing reproduction coloration.

When subject's head turns, the HRTFs from target source and two loudspeakers to two ears change. In dynamic transaural reproduction, the head turning is detected by a head tracker. According to the direction of the target virtual source relative to the temporary orientation of the subject's head, the HRTFs in the two transaural filters of Eq.(4) or (5) are updated constantly. The detail of dynamic transaural reproduction is referred to (3,8).

3. ANALYSIS ON DYNAMIC LOCALIZATION CUE

To analyze dynamic localization cue, the ITD (interaural time difference) variation caused by head turning in transaural reproduction is analyzed and compared with that of an actual source.

For an actual source at direction (Θ_S, Φ_S), the binaural pressures are evaluated by Eq.(1). When the head turns, the binaural pressures are also evaluated by Eq.(1), but the HRTFs at new direction with respect to head are used.

For static transaural reproduction, the binaural pressures are evaluated by Eq.(2) with the loudspeaker signals given by Eq.(3) and (4) in the case without timbre equalization, or given by Eq.(3) and (5) in the case with timbre equalization. When the head turns, the HRTFs from the two loudspeakers to two ears should be replaced by those at new directions with respect to head, but the loudspeakers signals are unchanged.

For dynamic transaural reproduction, the binaural pressures are evaluated similar to static reproduction. But when the head turns, both the HRTFs from the two loudspeakers to two ears and the loudspeakers signals should be changed according to the new directions with respect to head.

Subject's head is able to turn in three degrees of freedom, including turning around the left-right axes (pitch), around the front-back axes (tilting or rolling), and around the up-down axes (rotation or yaw). The ITD variation caused by head rotation is supposed to provide information for front-back and vertical localization, and the ITD variation caused by head tilting also provides supplementary information for up-down discrimination (5,9,10). Therefore, the ITD variations caused by head rotation and titling are evaluated.

There are various definitions and methods for ITD calculation(1). Here, the ITDs are calculated by maximizing the normalized cross-correlation function between the pressures at the two ears. Considering ITD is an effective localization cue at low frequency, the frequency range for the ITD calculation is chosen up to 1.5 kHz.

The schemes for analysis are as follow.

1) Calculate the binaural pressures for an actual source before and after head turning, respectively.
2) Calculate the ITD for an actual source before and after the head turning, respectively, and then, evaluate the variation in ITD.
3) Calculate the binaural pressures for transaural reproduction before and after head turning, respectively.
4) Calculate the ITD for transaural reproduction before and after head turning, respectively, and then, evaluate the variation in ITD.
5) Compare the ITD variations for actual source and transaural reproduction.

As an example of analysis, suppose that two loudspeakers are arranged in the horizontal plane. The distance between the loudspeakers and the head center of subject is 1.5m, and the directions are:

$$\Theta_L = -15° \qquad \Theta_R = 15° \qquad \Phi_L = \Phi_R = 0° \tag{7}$$

The actual or target source is located in the median plane. HRTFs of KAEMAR artificial head (with DB-060/061 small pinnae but without torso) are used in analysis. The HRTFs were obtained by first scanning the images of KAEMAR using a laser scanner and then calculating via the fast boundary element method (11). The sample frequency of HRTFs is 44.1 kHz and the length is 512 points.

Figure 1 shows the ITD variation (ΔITD) for various actual or target source polar elevations in the median plane and head rotation to the left with an azimuth of 10°. The results for actual source, static and dynamic transaural reproduction with timbre equalization are plotted in the same figure. The results of static and dynamic transaural reproduction without timbre equalization are almost identical to those with timbre equalization and therefore omitted here.

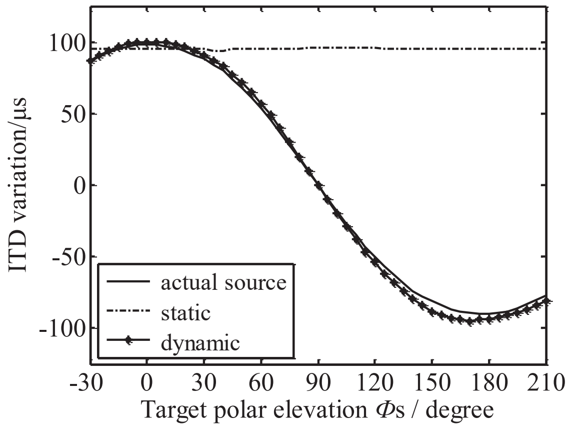


Figure 1 –Results of ITD variation for target source polar elevations in the median plane and head rotation to the left with an azimuth of 10°

In all cases, as expected, the ITD ≈ 0. In the case of actual source, the ΔITD varies with elevation. In the frontal-median plane -30º ≤ Φ_S < 90º, ΔITD > 0; and in the back-median plane, ΔITD < 0. Therefore, the sign of ΔITD provides information for front-back discrimination. In addition, the magnitude of ΔITD maximizes in the horizontal-frontal and back direction. As the source elevation departs from the horizontal plane, the magnitude of ΔITD is reduced. At the top with Φ_S = 90° (or bottom with Φ_S = -90°, not shown in the figure), the ΔITD is zero. Therefore, the ΔITD with head rotation provides information for vertical displacement from the horizontal plane. However, the ΔITD with head rotation is up-down symmetric. For example, the ΔITD for Φ_S = 30º and -30º are almost identical. In fact, the ITD variation caused by head tilting (as well as scattering by torso) provides further information for up-down discrimination.

In the case of static transaural reproduction, the ΔITD with head rotation are almost invariant against target elevation (96 μs to 97μs) and basically consistent with that of actual source near the front Φ_S = 0º. While in the case of dynamic transaural reproduction, the ΔITD with head rotation are almost consistent with that of the actual source at various elevations.

It can be proved that for head rotation with other small azimuth (for example, 20°) or for head tilting, the ITD variation exhibit a similar feather. That is, in the case of static transaural reproduction, the ITD variations with head turning are basically identical to that of actual source near the front. And in the case of dynamic transaural reproduction, the ITD variations with head turning basically match with those of actual source at target directions.

4. EXPERIMENT METHOD

A series of virtual source localization experiments were conducted to examine the perceived virtual direction of static and dynamic transaural reproduction.

The experiment was conducted in a listening room with reverberation time of 0.15 s. Static and dynamic transaural reproduction with timbre equalization were examined. Two loudspeakers (GENELEC 8010AP-5) were arranged in the horizontal plane at a distance of 1.5 m and azimuths given by Eq.(7). Nine target elevations from Φ_S = -30º to 210º at an interval of 30º in the median plane were chosen. The HRTFs used in transaural synthesis were identical to those in Sec.3. Two kinds of stimuli, pink noise with full audible bandwidth and 3 kHz low-pass pink noise, were used. The length of stimuli was 10 s.

In dynamic reproduction, an electromagnetic head tracker (Polhemus FASTRAK) was used to detect the orientation of the subject's head. It was able to detect the head turning in three degrees of freedom. The update rate and system latency time of the dynamic reproduction were 60 Hz and 25.4 ms, respectively.

The subjects judged the perceived virtual source direction and reported using an electromagnetic tracker (Polhemus FASTRAK). The tracker included two receivers. One receiver was fixed on the subject's head surface to monitor the position and orientation of the head. Another receiver was fixed at one end of a 1.0 m wooden rod. The subject pointed the rod at the position of the perceived virtual source and a computer recorded the result. The direction of the virtual source was measured relative to the head center because the data relative to the receiver on the head surface had been transferred to those relative to the head center. The subjects made the judgment and pointed at the perceived direction during the stimulus presentation. Before and after a presentation, the subjects were allowed to turn their heads to recognize the direction.

For static reproduction, the subject's heads were restricted during the judgment. The data from the head tracker indicated that the angle of head rotation and titling are less than 2º, respectively. For dynamic reproduction, the subjects were encouraged to turn their heads. For dynamic reproduction, the angle of head rotation ranged from ±5º to ±15º; and the angle of head titling ranged from ±10º to ±25º.

Eight subjects participated in the experiment. The subjects were from 22 to 30 years old and had normal hearing. Under each condition and at each target direction, each subject repeatedly judged three times. Therefore, there were 3 repetitions × 8 subjects = 24 judgments under each condition. Statistical analysis was applied to the 24 judgments.

5. RESULTS AND STATISTICS

The preliminary results from the subjects indicated that, for static reproduction with stimuli of two bandwidths and for dynamic reproduction with 3 kHz low-pass pink noise, a single virtual source was perceived. In this case, statistical analysis was applied to the perceived directions of this virtual source. However, for dynamic reproduction with pink noise of full audible bandwidth, two splitting virtual source were perceived. The high-frequency one was always perceived at the horizontal frontal directions. The low-frequency one was perceived at different elevation in the median plane, depending on the target elevation. In this case, statistical analysis was only applied to the perceived directions of the low-frequency virtual source.

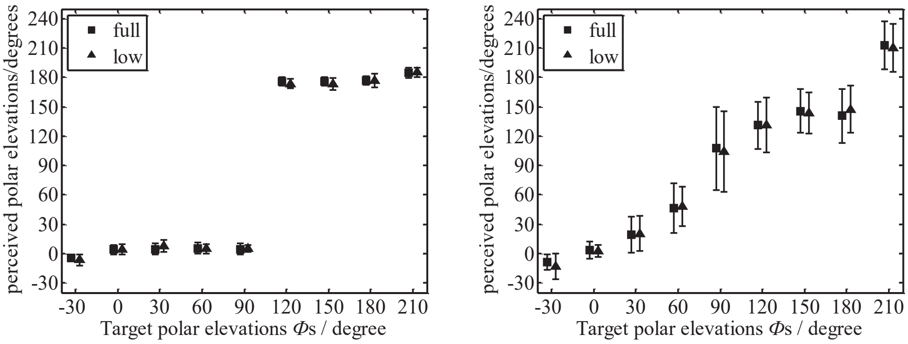
Although the localization results varied across the subjects, the overall tendency was similar. The perceived virtual source was located near the median plane with mean unassigned polar azimuth errors being less than 2.6º in all cases. For brevity, only the statistical results of all the subjects are analyzed here. Figure 2 show the mean polar elevation and corresponding standard deviation. Reversal was resolved for the raw localization results with front-back (F-B) and up-down (U-D) confusions. Table 1 lists the percentage of confusion.

Table 1 –Percentage of confusion for each target virtual source reproduction

			-30	0	30	60	90	120	150	180	210
Dynamic	Full band	F-B confusion(%)	0	0	4.2	25	/	29.1	8.3	8.3	8.3
		U-D confusion(%)	70.8	/	4.2	4.2	8.3	4.2	8.3	/	70.8
	Low-pass	F-B confusion(%)	0	0	0	16.7	/	16.7	8.3	8.3	12.5
		U-D confusion(%)	79.1	/	4.2	0	0	0	4.2	/	75
Static	Full band	F-B confusion(%)	0	0	0	0	/	100	100	100	100
		U-D confusion(%)	/	/	/	/	/	/	/	/	/
	Low-pass	F-B confusion(%)	0	0	0	0	/	100	100	100	100
		U-D confusion(%)	/	/	/	/	/	/	/	/	/

Figure 2(a) shows the results of static reproduction. The results for the stimuli with two different bandwidths are similar. For target source at the front median plane and top direction with -30º≤ Φ_S ≤ 90º, the mean perceived directions are near the Φ_I = 0º, and Table 1 indicate that no front-back confusion occurs. For target source at the back median plane with 120º ≤ Φ_S ≤ 210º , Figure 2(a) indicates the mean perceived directions are near the Φ_I = 180º. However, Table 1 indicates the percentage of front-back confusions are 100 %. Therefore, for target source at -30º ≤ Φ_S ≤ 210º, the actual perceived directions are all at the horizontal front, in spite of the target directions.

Figure 2(b) shows the results of dynamic reproduction. The results for 3 kHz low-pass pink noise and the low-frequency virtual source of pink noise with full-audible bandwidth are similar. The mean perceived elevations basically match with the target ones. Table 1 indicates that no or low front-back confusion occurs in most cases, except for some front-back confusions occur at the target directions Φ_S = 60º and 120º. In addition, only a few up-down confusions occur, expect for the low-elevation target directions of Φ_S = -30º and 210º at which the percentages of up-down confusion reach high values of 70.8 % ~ 89.1 %.



(a) Static reproduction (b) Dynamic reproduction

Figure 2 – Mean results and standard deviation of virtual source polar elevation localization in the median plane

6. DSICUSSION

The experimental results in sec.5 indicate that static transaural reproduction with two horizontal front loudspeakers is unable to recreate virtual source in back and high or low elevations in the median plane. Target virtual sources in all elevations in the median plane are perceived in front-horizontal directions. In contrast, dynamic transaural reproduction with two horizontal front loudspeakers is able to recreate virtual source in back or various high elevations in the median plane, at least for 3 kHz low-pass filtered stimuli.

Actually, both spectral cue and dynamic cue contribute to front-back and vertical localization. The spectral cue introduced by pinna is effective at high frequency above 5 kHz. It is also individual dependent. A careful designed transaural reproduction may theoretically be possible to create the high-frequency spectral cue at the ideal listening position. However，due to the short wavelength at high frequency, the spectral cue is very sensitive to small deviation from the listening position and other error in the reproduction chain. In other words, transaural reproduction is unable to create stable high-frequency spectral cue. If timbre equalization is supplemented, the spectral cue is further distorted. Therefore, front-back and vertical localization in transaural reproduction depend on dynamic cue.

Wallach hypothesized that the ITD variation caused by head rotation allows the discrimination of the front-back location as well as vertical displacement from the horizontal plane (9). Wallach’s hypothesis has been validated by some modern experiment (5,9). As shown in Sec.3, in the case of static reproduction, the information provided by ITD variation caused by head rotation is always consistent with this of a source near the horizontal frontal direction. In the case of dynamic reproduction, the information provided by ITD variation caused by head rotation matches with that of target or actual source. Of course, the incorrect spectral cue may cause another splitting high-frequency virtual source in the horizontal-frontal direction. In addition, some front-back confusion occurs at the target directions Φ_S = 60º and 120º may be due to small ΔITD magnitudes at these directions. And the large up-down confusion for low-elevation target directions of Φ_S = -30º and 210 º may be due to that the HRTFs of KEMAR without torso is used in present work. The torso-related spectral cue at low frequency contribution to up-down discrimination, although head tilting also provides supplementary information for up-down discrimination (9).

7. CONCLUSIONS

Transaural reproduction is unable to provide stable high-frequency spectral cue, therefore front-back and vertical localization depend on dynamic cue. Dynamic transaural reproduction with two horizontal-frontal loudspeakers is able to provide correct dynamic cue and therefore is able to create virtual source at back and various elevations, at least below 3 kHz. Static reproduction is unable to do so, and the perceived virtual source is usually limited in the frontal-horizontal plane.

Recently, mulichannel sound with height is developed rapidly. It is sometimes desired to downmix multichannel sound signals for reproducing with fewer loudspeakers. Transaural processing has been suggested for downmixing. The limitation discussed in present work should be considered when the downmixing scheme is designed.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (11674105) and State Key Lab of Subtropical Building Science, South China University of Technology.

REFERENCES

1. Xie BS. Head-related transfer function and virtual auditory display. J Ross Publishing, 2013.
2. Takeuchi T, Nelson P A, Kirkeby O, et al. Influence of Individual Head-Related Transfer Function on the Performance of Virtual Acoustic Imaging Systems. Audio Engineering Society Convention 104; Audio Engineering Society, 1998.
3. Gardner W G. 3-D audio using loudspeakers. Springer Science & Business Media, 1998.
4. Blauert J. Spatial hearing: the psychophysics of human sound localization. MIT Press, 1997.
5. Jiang J, Xie B, Mai H, et al. The role of dynamic cue in auditory vertical localisation. Applied Acoustics, 2019, 146: 398-408.
6. Bauck J, Cooper D H. Generalized transaural stereo and applications. Journal of the Audio Engineering Society, 1996, 44(9): 683-705.
7. Sakamoto N, Gotoh T, Kogure T, et al. Controlling sound-image localization in stereophonic reproduction. Journal of the Audio Engineering Society, 1981, 29(11): 794-799.
8. Lentz T, Assenmacher I, Vorländer M, et al. Precise near-to-head acoustics with binaural synthesis. JVRB-Journal of Virtual Reality and Broadcasting, 2006, 3(2).
9. Wallach H. The role of head movements and vestibular and visual cues in sound localization. Journal of Experimental Psychology, 1940, 27(4): 339.
10. Perrett S, Noble W. The effect of head rotations on vertical plane sound localization. The Journal of the Acoustical Society of America, 1997, 102(4): 2325-2332.
11. Rui YQ, Yu GZ, Xie BS, et al. Calculation of individualized near-field head-related transfer function database using boundary element method. Audio Engineering Society Convention 134. Audio Engineering Society, 2013.