INFLUENCE OF THE NUMBER OF LOUDSPEAKERS ON THE TIMBRE IN MIXED-ORDER AMBISONICS REPRODUTION

Haiming Mai[1], **Bosun Xie*[1, 2], *Jianliang Jiang*[1]

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

[2] State Key Laboratory of Subtropical Building Science, South China University of Technology, Guangzhou, China

*E-mail:phbsxie@scut.edu.cn

ABSTRACT

Ambisonics is a series of flexible spatial sound systems based on spatial harmonics decomposition and each order approximation of sound field. Accuracy and complexity of system increase with order. Considering that the horizontal localization resolution of human hearing is higher than vertical resolution, mixed-order Ambisonics (MOA) reconstructs horizontal sound field with higher order spatial harmonics, while reconstructs vertical sound field with lower order spatial harmonics, and thereby reaches a compromise between the perceptual performance and the complexity of system. For a given order MOA, the number of horizontal loudspeakers is flexible, providing that it exceeds some low limit. By using Moore's revised loudness model, the present work analyzes the influence of the number of horizontal loudspeakers on timbre in MOA reproduction. The results indicate that below the Shannon-Nyquist limit of spatial sampling, increasing the number of horizontal loudspeakers influence little on binaural loudness level spectra (BLLS) then timber. Above the limit, increasing the number of horizontal loudspeakers may increase the change of BLLS in some case, but reduce the change in some other case. The influence of the number of horizontal loudspeakers on BLLS and timbre reduces when virtual source departs from horizontal plane to the high or low elevation.

Index Terms— mixed-order Ambisonics, timbre, binaural loudness level spectra, the number of loudspeakers

1. INTRODUCTION

Ambisonics is a series of spatial sound systems based on spatial harmonics decomposition and each order approximation of sound field [1]. The region size and high-frequency limit of accurate sound field reconstruction increases with increasing order of Ambisonics, while the complexity of system also increases with order. According to Shannon-Nyquist spatial sampling theorem, an L_{2D}-order horizontal Ambisonics reproduction requires $K = (2L_{2D}+1)$ independent signals and $M \geq K$ loudspeakers. For L_{3D}-order spatial Ambisonics reproduction, K is equal to $(L_{3D}+1)^2$. Considering that the horizontal localization resolution of human hearing is higher than vertical resolution [2], mixed-order Ambisonics (MOA) reconstructs horizontal sound field with higher order spatial harmonics, while reconstructs vertical sound field with lower order spatial harmonics [3], and thereby reaches a compromise between the perceptual performance and the complexity of system.

The high-frequency limit and size of region for accurate reconstructing sound field in Ambisonics are determined by the order of Ambisonics. This is the consequence of Shannon-Nyquist spatial sampling theorem. Beyond the condition of spatial sampling theorem, spatial aliasing error occurs in the reproduction sound field, resulting in various audible artifacts, including timbre change. On the other hand, for a given order MOA, the number of horizontal loudspeakers is flexible, providing that it exceeds some low limit. Therefore, it is necessary to evaluate the influence of number of horizontal loudspeakers on the perceptual quality in MOA reproduction, especially beyond the condition of spatial sampling theorem.

Timbre is an important perceptual quality which contributes more to the overall perceptual quality than the spatial attributes [4]. Although there have been some researches of timbre change on spatial sound reproduction [5-6], these researches are mainly based on psychoacoustics experiments which are time-consuming. For convenience and efficiency, timbre change in spatial sound reproduction can be analyzed by using Moore's revised loudness model. Liu Yang et al. analyzed the timbre change in conventional Ambisonics reproduction by calculating binaural loudness level spectra (BLLS) [7], and indicated that timbre change reduces with the increasing order of Ambisonics. In our previous work, the BLLS and binaural pressures error were used to analyze the influence of the number of loudspeakers on horizontal and spatial Ambisonics reproduction, respectively [8-9].

For the flexibility of MOA, the various mixed-order combination reproductions can be implemented by separately changing the number of horizontal loudspeakers. By using the Moore's revised loudness model, the present work further analyzes the influence of the number of loudspeakers on the timbre in MOA reproduction. An MOA reproduction with layer-wise layout of loudspeakers is taken as an example. The BLLS and then timbre for various number of horizontal loudspeakers are analyzed.

2. MIXED-ORDER AMBISONICS

A clockwise spherical coordinate system is used in analysis. The origin of coordinate is located at the center of head. The spatial position is specified by distance $0 \le r < +\infty$, elevation $-90° \le \phi \le 90°$ and azimuth $0° \le \theta \le 360°$. $\phi = 0°$ and $90°$ denote horizontal and above direction, respectively. In horizontal plane, $\theta = 0°$ and $90°$ denote the front and right direction, respectively.

Suppose that M loudspeakers are arranged on a spherical surface with far-field radius r around head center and thus the incident wave cause by each loudspeaker can be approximated as plane wave. Let $\Omega_S = (\theta_S, \phi_S)$ and $\Omega_i = (\theta_i, \phi_i)$ denote the target direction of incident sound and the direction of i th loudspeaker, respectively. For MOA reproduction, the signal of the i th loudspeaker is a linear combination of spatial harmonics of various orders, with the horizontal harmonics being truncated to a higher order L_{2D} and vertical harmonics being truncated to a lower order L_{3D} [3,10].

$$E_i(\Omega_S) = \sum_{l=0}^{L_{3D}} \sum_{m=0}^{l} \sum_{\sigma=\pm 1} D_{lm}^{\sigma}(\Omega_i) S_{lm}^{\sigma}(\Omega_S) + \sum_{l=L_{3D}+1}^{L_{2D}} \sum_{\sigma=\pm 1} D_{ll}^{\sigma}(\Omega_i) S_{ll}^{\sigma}(\Omega_S) \tag{1}$$

Where $S_{lm}^{\sigma}(\Omega_S) = S_0 Y_{lm}^{\sigma}(\Omega_S)$ is a set of spatial harmonics or independent signals with various orders. S_0 is a constant related to overall signal amplitude. $Y_{lm}^{\sigma}(\Omega)$ are the normalized real-valued spherical harmonic functions:

$$Y_{lm}^{\sigma}(\Omega) = \begin{cases} N_{lm}^{\sigma} P_l^m \left[\cos(90^{\circ} - \phi)\right] \cos(m\theta) & \sigma = +1 \\ N_{lm}^{\sigma} P_l^m \left[\cos(90^{\circ} - \phi)\right] \sin(m\theta) & \sigma = -1 \end{cases} \tag{2}$$

Where $P_l^m\left[\cos(90^{\circ} - \phi)\right]$ is the associated Legendre polynomial, the normalized factor is:

$$N_{lm}^{\sigma} = \sqrt{\frac{(l-m)!(2l+1)}{(l+m)!2\pi\Delta_m}} \qquad \Delta_m = \begin{cases} 2 & m = 0 \\ 1 & m \neq 0 \end{cases} \tag{3}$$

$D_{lm}^{\sigma}(\Omega_i)$ is the decoding coefficient which depends on loudspeaker direction, and it can be solved from spherical harmonics function matrix of loudspeaker position by using pseudo-inverse method [11-12].

In addition, according to spatial sampling theorem, the high-frequency limit $f_{max,H}$ of accurate sound field reconstruction within a spherical region with radius a in L-order Ambisonics reproduction is evaluated by:

$$f_{max,H} = \frac{Lc}{2\pi a} \tag{4}$$

Where c is the speed of sound. When Eq.(4) is satisfied, the average reconstructed sound pressure error is less than -10 dB [10].

3. ANALYSIS ON TIMBRE BY USING MOORE'S REVISED LOUDNESS MODEL

Moore's revised loudness model can be used to predict the perceived loudness of sound field in various frequency bands, which is an index for timbre change in spatial sound reproduction. Through Moore's revised loudness model, BLLS can be obtained. Fig.1 shows the block diagram of Moore's revised loudness model. BLLS represent the perceived loudness in various frequency band of equivalent rectangular bandwidth (ERB). The ERB approximates frequency resolution of human hearing. The relationship between the ERB (in Hz) and the center frequency of auditory filter (in kHz) is [13]:

$$ERB = 24.7(4.37f + 1) \tag{5}$$

And a new scale of frequency is used in BLLS calculation—the number of equivalent rectangular bandwidth (ERBN):

$$ERBN = 21.4\log_{10}(4.37f + 1) \tag{6}$$

The detail of Moore's loudness model is referred to [14].

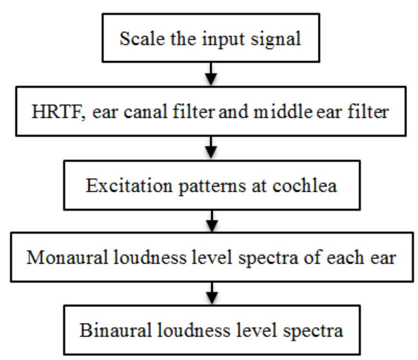


Fig.1. The block diagram of BLLS calculation by Moore's revised loudness model

The procedures for analyzing the timbre change in MOA reproduction are:

I. Scale the magnitude of input stimuli to a value corresponding predetermined free-field pressure level.
II. The binaural pressures of target sound field, $P_{\alpha}(\Omega_S, f)$ are calculated by filtering the scaled stimuli $S_0(f)$ with a pair of HRTFs (head-related transfer functions) at direction Ω_S:

$$P_{\alpha}(\Omega_S, f) = H_{\alpha}(\Omega_S, f) S_0(f) \tag{7}$$

Where α = L and R denotes left and right ear, respectively. Then the BLLS for target sound field is calculated from binaural pressures by using Moore's revised loudness model.

III. Similar to step II, the reconstructed binaural pressures in MOA, $P'_\alpha(\Omega_S, f)$ are calculated by filtering each loudspeaker signal with corresponding pair of HRTFs and then summing:

$$P'_\alpha(\Omega_S, f) = \sum_{i=1}^{M} H_\alpha(\Omega_i, f) E_i(\Omega_S, f) \tag{8}$$

Then the BLLS for MOA reproduction is calculated from binaural pressures by using Moore's revised loudness model.

IV. The BLLS for MOA reproduction and target sound field are compared. If they match well, no perceived timbre change occurs in MOA reproduction. Otherwise, if the deviation between them (denoted by BLLSD) exceeds 1 Phon/ERB, which is just noticeable difference (JND) of BLLS, perceivable timbre change occurs in MOA reproduction. The larger is the difference, the more timbre change occurs.

By using the above procedures, timbre change in certain order MOA with various number of horizontal loudspeakers is analyzed in the following section.

4. RESULTS AND DISSCUSSION

Ambisonics reproduction with 28+1 layer-wise loudspeaker layout is taken as a reference. 28 loudspeakers are arranged in three elevation layers on a spherical surface with far-field radius. There are 8, 12 and 8 loudspeakers in the -45°, 0° and 45° elevation layers, with a uniform azimuthal interval of 45°, 30° and 45°, respectively. An additional loudspeaker is arranged on the top with (θ, ϕ) = (0°, 90°). MOA reproduction with the increase number of horizontal loudspeakers are evaluated and compared with the reference.

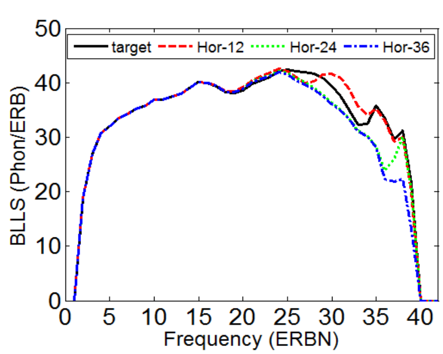
According to Ref.[12], the reference loudspeaker layout is able to reproduce conventional Ambisonics up to L = 3 order, and MOA up to L_{3D} = 3 and L_{2D} = 5 order (denoted by 3/5 order). For comparison, when the number of horizontal loudspeakers increases to 24, 36 and 72 respectively while the number of loudspeakers in other elevation layers are intact (corresponding, the total number of loudspeakers are 41, 53 and 89, respectively), it is still able to reproduce conventional Ambisonics up to 3 order but stability of reproduction deteriorates; on the other hand, it is able to reproduce MOA up to 3/11, 3/17 and 3/35 order, respectively.

Table.1. The condition number of loudspeaker position matrix

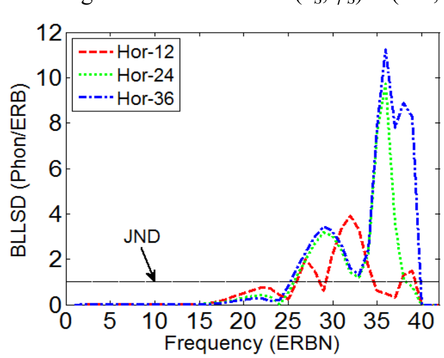
M / L_{3D}/L_{2D}	29 (Hor-12)	41 (Hor-24)	53 (Hor-36)	89 (Hor-72)
3/3	2.51	3.24	3.92	5.50
3/5	2.51	3.24	3.92	5.50
3/11	10^{16}	3.77	4.58	6.44
3/17	10^{16}	10^{16}	5.08	7.13
3/35	10^{16}	10^{16}	10^{15}	8.50

In following analysis, the input stimuli is pink noise, which is scaled to a value corresponding pre-determined free-field pressure level of 70 dB. The KEMAR-HRTFs obtained from optical scanning and BEM-based calculation are used in the analysis. The directional resolution of HRTFs is 1°.

The BLLS in horizontal target directions is first analyzed. Fig.2 shows the results for 3/5 order MOA for target incident direction (θ_S, ϕ_S) = (15°, 0°) with M = 29, 41 and 53 loudspeakers, respectively. The corresponding horizontal loudspeaker number is 12, 24 and 36, respectively. Fig.2(a) plots the BLLS and Fig.2(b) plots the BLLSD. It is observed that below the frequency of 25 ERBN, the BLLS for MOA reproduction with various number of loudspeakers match well with that of target plane wave and no perceivable timbre change occurs. However, above this frequency, the BLLSD increases and is larger than the JND of 1 Phon/ERB, resulting in perceivable timbre change. In this case, increasing number of loudspeakers results in more timbre change. It should be noticed that the frequency of 25 ERBN approximately corresponds to that $f_{max,H}$ evaluated from Eq.(4) for L = 5 order Ambisonics reproduction and with average radius a = 0.0875 m of human head. Fig.3 shows the results for target incident direction (θ_S, ϕ_S) = (75°, 0°), other

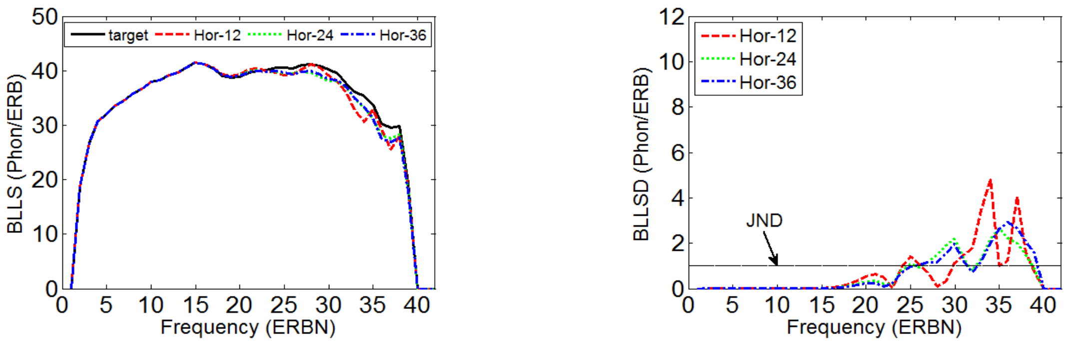


(a) The binaural loudness level spectra



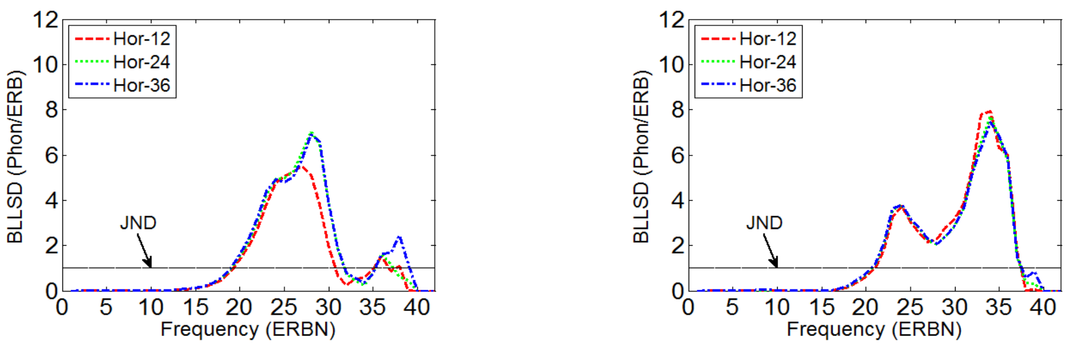
(b) The deviation between the BLLS of reproduction and target

Fig.2. 3/5 order MOA reproduction, (θ_S, ϕ_S) = (15°, 0°)



(a) The binaural loudness level spectra (b) The deviation between the BLLS of reproduction and target

Fig.3. 3/5 order MOA reproduction, (θ_S, ϕ_S) = (75°, 0°)



(a) Target direction (θ_S, ϕ_S) = (15°, 45°) (b) Target direction (θ_S, ϕ_S) = (75°, 45°)

Fig.4. 3/5 order MOA reproduction, BLLSD

conditions are identical to those in Fig.2. The results are similar to those of Fig.2 below the frequency of 25 ERB. However, above that frequency, the BLLSD reduces with the increase of number of loudspeakers.

Fig.4(a) and (b) show the results for target incident direction (θ_S, ϕ_S) = (15°, 45°) and (θ_S, ϕ_S) = (75°, 45°) respectively, other conditions are identical to those in Fig.2. It is observed that the BLLSD is small below 20 ERBN which corresponds to $f_{max,H}$ for L = 3 order Ambisonics reproduction evaluated from Eq.(4). Above that frequency, the influence of number of loudspeakers is similar to that in horizontal plane but with smaller BLLSD.

Similar analysis can be applied to MOA reproduction with various mixed-orders, number of loudspeakers and target incident directions. The results are similar to the cases mentioned above. Overall, below the high-frequency limit $f_{max,H}$, the BLLSD is less than JND of 1 Phon/ERB in all the cases, then no perceivable timbre change occurs. Above that frequency, the BLLSD depends on both target directions and the number of loudspeakers. The BLLSD reduces with increasing the number of loudspeakers at lateral directions with θ_S from 70° to 110°, but increases at other frontal and back directions. The influence of the number of horizontal loudspeakers on BLLS reduces when the target direction departs from horizontal plane to high or low elevation.

5. CONCLUSIONS

Increasing the horizontal order L_{2D} in MOA increases the high-frequency limit for horizontal virtual source. Below this limit, no perceivable timbre change occurs. This limit is consistent with the result of Shannon-Nyquist spatial sampling theorem. For each given order MOA, increasing the number of horizontal loudspeakers to more than the minimal requirement influences little on the BLLS and then timbre below the high-frequency limit. Above the high-frequency limit, however, the change of BLLS and then timbre depends on both the target virtual source direction and the number of loudspeakers. Increasing the number of horizontal loudspeakers in MOA reproduction increases the change of BLLS at frontal and back directions, but reduces the change of BLLS at other directions. The influence of the number of horizontal loudspeakers on BLLS and timbre reduces when the target direction departs from horizontal plane to high or low elevation.

6. ACKNOWLEDGEMENT

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.

7. REFERENCES

[1] J. Daniel, S. Moreau, R. Nicol, “Further investigations of high-order Ambisonics and wavefield synthesis for holophonic sound imaging,” *The 114th Convention of the Audio Engineering Society*, Amsterdam, Netherlands, Paper: 5788,2003.

[2] J. Blauert, *Spatial hearing (revised edition)*. MIT Press, Cambridge, MA, England, 1997.

[3] S. Favrot, M. Marschall, J. Käsbach, et al., “Mixed-order Ambisonics recording and playback for improving horizontal directionality,” *The 131th Convention of the Audio Engineering Society*, New York, USA, Paper: 8528, 2011.

[4] F. Rumsey, S. Zieliński, R. Kassier, S. Bech, “On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality,” *The Journal of the Acoustical Society of America*, vol. 118, pp. 968-976, 2005.

[5] C. Mori, T. Nishiguchi, K. Ono, “Measurement and generation method of many impulse responses for 22.2 multichannel sound production,” *The 2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pp. 1-4, 2016.

[6] Z. Ben-Hur, F. Brinkmann, J. Sheaffer, et al., “Spectral equalization in binaural signals represented by order-truncated spherical harmonics,” *The Journal of the Acoustical Society of America*, vol.141, pp. 4087-4096, 2017.

[7] Y. Liu, B. S. Xie, “Subjective evaluation on the timbre of horizontal ambisonics reproduction,” *The 2014 International Conference on Audio, Language and Image Processing (ICALIP)*, pp. 11-15, 2014.

[8] H. M. Mai, J. L. Jiang, B. S. Xie, et al., “Influence of the number of loudspeakers on the timbre in horizontal Ambisonics reproduction,” *The 2016 International Conference on Audio, Language and Image Processing (ICALIP)*, pp. 95-99, 2016.

[9] J. L. Jiang, H. M. Mai, B. S. Xie, et al., “Analysis on the influence of number of loudspeakers on the error in binarual pressure spectra for spatial ambisonics reproduction,” *The 2016 International Conference on Audio, Language and Image Processing (ICALIP)*, pp. 197-201, 2016.

[10] D. Ward, T. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” *IEEE Transactions on speech and audio processing*, vol. 9, pp. 697-707, 2001.

[11] J. M. Batke, F. Keiler, “Using VBAP-derived panning functions for 3D ambisonics decoding,” *The 2nd Ambisonics Symposium*, Paris, France, 2010.

[12] H. M. Mai, B. S. Xie, “Analysis on the stability of Mixed-order Ambisonics reproduction system,” *Technical Acoustics*, vol. 35, pp. 569-572, 2016.

[13] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, Emerald Group Publishing, England, 2012.

[14] B. C. J. Moore, B. R. Glasberg, “Modeling binaural loudness,” *The Journal of the Acoustical Society of America*, vol. 121, pp. 1604-1612, 2007.