Analysis on the interaural level difference in near-field-compensated higher order Ambisonics reproduction

Bosun Xie

Acoustic Lab, School of Physics and Optoelectronics, South China University of Technology, Guangzhou 510641, China. phbsxie@scut.edu.cn

Abstract

Near-field-compensated higher order Ambisonics (NFC-HOA) is a spatial sound reproduction technique based on spherical harmonics decomposition and each order approximation of sound field. The aim of NFC-HOA is to reconstruct the curve wavefront of spherical wave cause by sound source at different distances. It is desired that NFC-HOA is capable of recreating appropriate interaural level difference (ILD) which is considered to be an auditory distance localization cue for nearby sound source within a distance of 1 m relative to head center and outside the median plane. The present work analyzes theILD in NFC-HOA reproduction by using head-related transfer functions and compares with the case of a real point source. The results indicate that, due to the requirement of excessive low-frequency boost in the distance-compensated filters, it is difficult for NFC-HOA to recreate appropriate ILD for nearby target virtual source below the frequency of 0.7 to 1kHz. On the other hand, in order to recreate appropriate high-frequency ILD for nearby target virtual source, a much higher order NFC-HOA is needed. An illustrative example indicates that, even for the center listening position, a 12-order NFC-HOA with not less than 169 loudspeakers is needed for recreating appropriate ILD of a lateral virtual source at 0.25m and up to the frequency of 5kHz. Therefore, in practice, NFC-HOA is workable in certain mid-frequency range.

Keywords: Ambisonics, Near-field, Interaural Level difference
Analysis on the interaural level difference in near-field-compensated higher order Ambisonics reproduction

1 Introduction

In the case of free-field, the sound field created by a point sound source is spherical wave with curve wavefront. When a listener enters the sound field, the sound wave is scattered and diffracted by listener's anatomical structures including head, pinnae and torso etc. and then transferred into binaural pressures. The pressures at ears include major cues for auditory localization, such as interaural time difference (ITD), interaural level difference (ILD), spectral cue and dynamic cue [1]. Especially, for a source at nearby distance (within 1.0 m relative to head center) and outside the median plane, the distance-dependent ILD is considered to be an auditory distance localization cue in the free-field [2].

Ambisonics is a series of spatial sound reproduction systems. Based on spherical Bessel and harmonics decomposition as well as each order truncation of the sound field, it aims at reconstructing the target sound field below some frequency limit and within a spherical region. The upper frequency limit and radius of region increase with the order of Ambisonics, which is a consequence of Shannon-Nyquist spatial sampling theorem of sound field. The conventional far-field-high-order Ambisonics was designed to reconstruct a target plane wave field caused by a source at far distance. The near-field-compensated higher order Ambisonics (NFC-HOA) is designed to reconstruct a target spherical wave caused by a nearby source, and at the same time, to compensate for the unwanted influences of finite loudspeaker distance in reproduction [3].

NFC-HOA is desired to be able to reconstruct the spherical wavefront of a nearby source, and accordingly to recreate distance-dependent ILD accurately for auditory distance localization. In order to evaluate the ability of NFC-HOA, the present work analyzes the ILD in NFC-HOA reproduction by using head-related transfer functions and compares with the case of a real point source.

2 Principle of NFC-HOA

A clockwise spherical coordinate system is used. The origin of coordinate is located at the center of the head. Spatial position is specified by distance $0 \leq r \leq \infty$, azimuth $-180^\circ < \theta \leq 180^\circ$ and elevation $-90^\circ \leq \phi \leq 90^\circ$. Where $\phi = -90^\circ$, $0^\circ$ and $90^\circ$ represent the bottom, horizontal and top direction, respectively. In the horizontal plane, $\theta = 0^\circ$, $90^\circ$ and $180^\circ$ represent the front, left and back direction, respectively.

Let $r_s = (r_s, \theta_s, \phi_s) = (r_s, \Omega_s)$ denote the source position, $r = (r, \theta, \phi) = (r, \Omega)$ denote the position of field point. The pressure at arbitrary field point caused by a harmonic point source with unit intensity can be decomposed by spherical harmonics functions:
\[ P(r, r_s, f) = \frac{1}{4\pi} \left| \frac{1}{r - r_s} \right| \exp(-jk \mid r - r_s \mid) \]  

\[ = -jk \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{\sigma=1}^{2} j_l(kr)h_l(kr_s)Y_{lm}^{(\sigma)}(\Omega)Y_{lm}^{(\sigma)\dagger}(\Omega) \]  

Where \( f \) is the frequency and \( k \) is the wave number; \( j_l(kr) \) and \( h_l(kr_s) \) are the \( l \)th-order spherical Bessel and spherical Hankel function of the second kind, respectively; \( Y_{lm}^{(\sigma)}(\Omega) \) is real-value spherical harmonics function of the order \( l \) and degree \( m \):

\[ Y_{lm}^{(\sigma)}(\Omega) = \begin{cases} N_{lm}P_l^m(\sin \phi) \cos m\phi & \sigma = 1 \\ N_{lm}P_l^m(\sin \phi) \sin m\phi & \sigma = 2 \end{cases} \]  

\( l = 0, 1, 2, \ldots; \quad m = 0, 1, 2, \ldots, l. \)

\( P_l^m(\sin \phi) \) is the associated Legendre polynomial;

\[ N_{lm} = \sqrt{(l-m)! (2l+1) \over (l+m)! 2\pi \Delta_m} \quad \Delta_m = \begin{cases} 2 & m = 0 \\ 1 & m \neq 0 \end{cases} \]  

\( Y_{i0}^{(2)}(\Omega) = 0 \), retaining these terms in Eq.(1) is for convenience in written.

In NFC-HOA, suppose that \( M \) loudspeakers are arranged uniformly or nearly uniformly on a spherical surface with radius \( r_0 \). The position of the \( i \)th loudspeaker is \( r_i = (r_0, \theta_i, \phi_i) = (r_0, \Omega_i) \), corresponding signal amplitude is \( A_i = A_i(r_s, \Omega_s, r_0, \Omega_i, f) \). The pressure in NFC-HOA reproduction is the summing of these created by \( M \) loudspeakers:

\[ P(r, f) = \sum_{i=0}^{M-1} A_i \left| \frac{1}{r - r_i} \right| \exp(-jk \mid r - r_i \mid) \]  

Similar to Eq.(1), the pressure created by each loudspeaker can also be decomposed by spherical harmonics functions. In order to reconstruct the target sound pressure caused by a point source, the pressure in Eq.(4) should match with that in Eq.(1). Let Eq.(4) equal to Eq.(1) and truncate to the order of \( (L-1) \), we obtain a set of linear equations with respect to \( A_i \):

\[ h_i(kr_0) \sum_{i=0}^{M-1} A_i(r_s, \Omega_s, r_0, \Omega_i, f)Y_{lm}^{(\sigma)}(\Omega_i) = h_i(kr_s)Y_{lm}^{(\sigma)}(\Omega_s) \]  

\( l = 0, 1, 2, \ldots, (L-1) \quad m = 0, 1, \ldots, l. \quad \text{if} \quad m \neq 0, \quad \sigma = 1, 2 \quad \text{if} \quad m = 0, \quad \sigma = 1 \)

In order to find the loudspeaker signals, \( A_i(r_s, \Omega_s, r_0, \Omega_i, f) \) are also decomposed by spherical harmonics functions and truncated to the order \( (L-1) \):

\[ A_i(r_s, \Omega_s, r_0, \Omega_i, f) = \sum_{l=0}^{L-1} \sum_{m=0}^{l} \sum_{\sigma=1}^{2} D_{lm}^{(\sigma)}(r_s, r_0, \Omega_s, f)Y_{lm}^{(\sigma)}(\Omega_s) \]  

Where \( D_{lm}^{(\sigma)}(r_s, r_0, \Omega_s, f) \) are a set of coefficients to be solved. Substituting Eq.(6) into Eq.(5) yields a matrix equation:

\[ [\Xi] = [Y][D] \]
Where \( [Y] \) is an \( L^2 \times M \) matrix consisting of spherical harmonics functions \( Y_{lm}^{(\sigma)}(\Omega) \) of loudspeaker directions, with elements in each row corresponding to a given \((l, m, \sigma)\) and elements in each column corresponding to a loudspeaker direction \( \Omega \). \([D]\) is an \( M \times L^2 \) matrix consisting of coefficients \( D_{lm}(r_s, r_0, \Omega, f) \), with elements in each row corresponding to a loudspeaker direction \( \Omega \) and elements in each column corresponding to a given \((l, m, \sigma)\). \([\Xi]\) is an \( L^2 \times L^2 \) diagonal matrix, its elements corresponding to the \( l \)-order spherical harmonics term are given by

\[
\Xi_l(kr_s, kr_0) = \frac{h_l(kr_s)}{h_l(kr_0)}
\]

When

\[
M \geq L^2
\]

The decoding matrix or coefficients are solved by following pseudoinverse method

\[
[D] = \text{pinv}[Y][\Xi] = [Y]^T \{[Y][Y]^T\}^{-1}[\Xi]
\]

(10)

Substituting the coefficients given by \([D]\) into Eq.(6) yields the loudspeaker signals for NFC-HOA.

In Eq.(5), the spherical harmonics expansion is truncated up to the order of \((L-1)\). In NFC-HOA, the contribution of the \( l \)-order spherical harmonics component to the reconstructed pressure in Eq.(4) is directly proportional to \( j(kr) \Xi_l(kr_s, kr_0) \). Calculation indicates that, given \( r < r_s < r_0 \), when the wave number \( k \) or frequency \( f \) exceeds some value, the contribution of \( j(kr) \Xi_l(kr_s, kr_0) \) descends as the order \( l \) increasing. Therefore, the required order \((L-1)\) for accurate reconstruction of pressure can be numerically evaluated from \( j(kr) \Xi_l(kr_s, kr_0) \). A general tendency is that, for given \( r_s \) and \( r_0 \), \((L-1)\) increases with \( kr \), or equally, increases with the product of frequency and the radius of region. This is the consequence of spatial sampling theorem of sound field.

Eq.(9) is the minimal number of loudspeakers required for the \((L-1)\) order HFC-HOA reproduction. Therefore, the required number of loudspeaker increases quickly with the order.

### 3 Limitation by the NFC filters

Eq.(8) represents a set of NFC filters which code the target source distance information into loudspeaker signals, and at the same time, compensate for the finite loudspeaker distance. For \( r_s < r_0 \), i.e., target source locating within the array of loudspeakers, the filters in Eq.(8) exhibit a boost at low frequency. Actually, from the asymptotic equation of spherical Hankel function of the second kind,

\[
\lim_{\zeta \rightarrow \infty} h_i(\zeta) = \frac{j^{i+1}}{\zeta} \exp(-j\zeta)
\]

(11)

The low-frequency magnitude response of Eq.(8) are approximated by
\[
\lim_{k \to 0} | \Xi_k(kr_S, kr_0) | \approx \left( \frac{r_0}{r_S} \right)^{l+1} \quad (12)
\]

The low-frequency boost of filter responses increases with the ratio between loudspeaker and target distances, as well as increases with the order \( l \). In the case of \( r_S = 1.5 \text{m} \) and \( r_0 = 0.25 \text{m} \), the low-frequency boosts estimated from Eq.\((12)\) are 15.6dB, 31.1dB and 46.7dB for \( l = 1, 2 \) and 3, respectively.

In order to accurately reconstruct the target sound field up to a higher frequency limit and within a reasonable region, a higher order NFC-HOA is preferred. However, a higher order NFC-HOA requires distance filters with large boost at low frequency. An excessive boost at low frequency is infeasible for any practical electroacoustic system. A compromise and practical method is to use adaptive or varying order NFC-HOA for different frequency ranges. That is, using lower order NFC-HOA for lower frequency and higher order NFC-HOA for high frequency range. This inevitably causes error in reconstructed sound field and ILD at low frequency, as will be seen in the following sections.

### 4 Method for analyzing the ILD

As stated in Section 1, distance-dependent ILD is a cue for auditory distance localization. In order to evaluate the performance of NFC-HOA, the ILD created by NFC-HOA is calculated and then compared with this created by a target point source.

For a target point source with unit intensity and at position \((r_S, \Omega_S)\) relatives to listener’s head center, the binaural pressures can be calculated using a pair of head-related transfer functions (HRTFs) at corresponding position [4]:

\[
P_L(r_S, \Omega_S, f) = \frac{1}{4\pi r_S} H_L(r_S, \Omega_S, f) \exp(-jkr_S)
\]

\[
P_R(r_S, \Omega_S, f) = \frac{1}{4\pi r_S} H_R(r_S, \Omega_S, f) \exp(-jkr_S)
\]

For NFC-HOA, binaural pressures are the combination of those caused by all loudspeakers and can also be calculated using the HRTFs at \( M \) loudspeakers positions:

\[
P'_L(f) = \frac{1}{4\pi r_0} \sum_{i=0}^{M-1} A_i H_L(r_0, \Omega_i, f) \exp(-jkr_0)
\]

\[
P'_R(f) = \frac{1}{4\pi r_0} \sum_{i=0}^{M-1} A_i H_R(r_0, \Omega_i, f) \exp(-jkr_0)
\]

The ILD can be calculated from binaural pressures. The ILD for a target point source is calculated from Eq.\((13)\):

\[
ILD(r_S, \Omega_S, f) = 20\log_{10} \left\{ \frac{P_L(r_S, \Omega_S, f)}{P_R(r_S, \Omega_S, f)} \right\}
\]

Similarly, the ILD for NFC-HOA is calculated from Eq.\((14)\).
In the following section, the calculated HRTFs from a rigid-spherical head model with a mean radius of \( a = 0.0875 \) m are used in calculating the binaural pressures. Of course, more accurate but complicated HRTFs, for example those measured or numerically calculated from KEMAR artificial head, can also be used. But the results are similar.

## 5 Results and Discussion

Using the aforementioned procedures, the ILD for target source and NFC-HOA are analyzed. As mentioned above, adaptive order NFC-HOA is chosen for different frequency ranges. The order \((L-1)\) at each frequency range is chosen by comprehensively considering the accuracy of NFC-HOA reconstruction and limiting the maximal low-frequency gain of distance filters in the Eq.(9) to 20dB. Table 1 lists the chosen order \((L-1)\) for the case of target source distance \( r_S = 0.25 \) m, loudspeaker distance \( r_0 = 1.50 \) m. Below the frequency of 5kHz, the chosen order increases with frequency. Above that frequency, the chosen order is fixed to a constant of \((L-1) = 20\).

\( M = 900 \) loudspeakers are arranged nearly-uniformly in the spherical surface. Fig.1 shows the resulted ILD for target source and NFC-HOA. The target direction is \((\theta_S, \phi_S) = (90^\circ, 0^\circ)\), i.e., at the lateral direction in horizontal plane. It is observed that the ILD for target source increases with frequency as whole but with some fluctuation. Even at low frequency, the ILD exhibits a large value of near 10 dB. This low-frequency ILD for nearby source, which is caused by head shadow and the ratio between the source distances relative to left and right ears, is considered to be a distance localization cue [2].

It is also observed that the ILD for NFC-HOA approximately matches with this of target source only within the frequency range from about 1kHz to 10kHz. The deviation of ILD at low frequency is due to that lower-order NFC-HOA is unable to accurately create binaural pressures. As stated in Section 3, limited by the gain or dynamic range of practical NFC filter in Eq.(8), increasing the order of NFC-HOA to improve the performance at low frequency is infeasible.

The deviation of ILD at high frequency is due to the high-frequency limit of \((L-1) = 20\) order NFC-HOA, which is imposed by spatial sampling theorem. Increasing the order of NFC-HOA heightens the upper frequency limit for accurate reconstruction of ILD, but makes the NFC-HOA complicated.

The results for other target source directions are similar to above example. However, as target source departs from the lateral direction toward the front or departs from the horizontal plane, the ILD for target source decreases. Accordingly, the low-frequency ILD for NFC-HOA begins to decrease.

### Table 1: Chosen order for the case of target source distance \( r_S = 0.25 \) m, loudspeaker distance \( r_0 = 1.50 \) m.

<table>
<thead>
<tr>
<th>( f ) (kHz)</th>
<th>0.15</th>
<th>0.25</th>
<th>0.35</th>
<th>0.45</th>
<th>0.65</th>
<th>0.95</th>
<th>1.25</th>
<th>1.70</th>
<th>3.5</th>
<th>5.0</th>
<th>&gt;5.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>((L-1))</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>20</td>
</tr>
</tbody>
</table>
match with that of target source at lower frequency. Fig. 2 shows the case of target source direction \((\theta_s, \phi_s) = (30^\circ, 0^\circ)\). In this case, the target source distance, loudspeaker distance, and number of loudspeakers are identical to above case, i.e., \(r_s = 0.25\)m, \(r_0 = 1.5\)m, \(M = 900\). The order of NFC-HOA is chosen according to Table 1.

Similarly, as target source distance increases or loudspeakers distance reduces, the low-frequency ILD for NFC-HOA begins to match with that of target source at lower frequency. Fig. 3 shows the case of target source direction \((\theta_s, \phi_s) = (90^\circ, 0^\circ)\) and target source distance \(r_s = 0.5\)m. The loudspeaker distance and number of loudspeakers are identical to above case. The order of NFC-HOA is also chosen according to above adaptive method.
6 Discussion and conclusion

Source-distance-dependent ILD is considered to be an auditory distance localization cue for nearby source outside the median plane. It is desired that NFC-HOA is capable of recreating appropriate ILD so as to recreate auditory distance perception. However, the present results indicate that, due to the requirement of excessive low-frequency boost in the NFC filters, it is difficult for NFC-HOA to recreate appropriate ILD for nearby target virtual source below the frequency of 0.7 to 1 kHz. On the other hand, in order to recreate appropriate high-frequency ILD for nearby target virtual source, a much higher order NFC-HOA is needed. Above examples indicate that, even for the central listening position, a 12-order NFC-HOA (requiring not less than 169 loudspeakers) is needed for recreating appropriate ILD of a lateral virtual source at 0.25m and up to the frequency of 5 kHz. Therefore, in practice, NFC-HOA is workable in certain mid-frequency range. Future works include evaluate the relative contribution of ILD in different frequency range to distance localization, which requires further psychoacoustic experiment.

Acknowledgments

This work was supported by the Guangzhou Science and Technology Plan Projects (No. 2014Y2-00021).

References