IMPACT OF REGULARIZATION OF NEAR FIELD CODING FILTERS FOR 2D AND 3D HIGHER-ORDER AMBISONICS ON AUDITORY DISTANCE CUES

Sylvain Favrot, Jörg M. Buchholz

Centre for Applied Hearing Research
Department of Electrical Engineering, Technical University of Denmark
Ørsted plads 352, 2800, Kgs. Lyngby, Denmark
sf@elektro.dtu.dk

ABSTRACT

A known challenge in sound field reproduction techniques such as high-order Ambisonics (HOA) is the reproduction of nearby sound sources. In order to reproduce such nearby sound sources, the near-field compensated (NFC) method with angular weighting windows (AWWs) has been previously proposed for HOA [1]. Considering auditory distance perception, (low-frequency) interaural level differences represent the main auditory cue for nearby real sound sources outside the median plane. Simulations showed that these ILD cues can be reproduced with existing weighted NFC-HOA methods for frequencies above about 500 Hz. However, since low-frequency ILD cues seem to be important for distance perception, a novel regularization function is proposed as AWW that can reproduce natural ILDs down to about 250 Hz, even in realistic playback environments. Using this regularization function, a listening test showed an improved distance perception performance for lateral sources compared to frontal ones. This improvement was greater for 3D reproduction than for 2D but slightly lower than for real sources. The distance of virtual nearby sources reproduced by NFC-HOA with the regularization function as AWW can be perceived nearly as accurately as corresponding physical ones.

1. INTRODUCTION

The reproduction of nearby sound sources (sources located closer than the loudspeakers) with sound-field reproduction techniques such as higher-order Ambisonics (HOA) [2] is a challenging task. Since in such conditions the plane wave description of HOA is not valid anymore, HOA has been previously modified to near field compensated (NFC) HOA [1] by assuming spherical point sources.

Near field compensated HOA introduces near field coding filters to Ambisonics signals that can lead to loudspeaker signals with extremely large pressures at low frequencies. Although the increased low-frequency energy is compensated inside the sweet spot by the interference of the different loudspeaker signals, a strong low frequency boost can be observed in the sweet spot with the presence of the listeners head or small (realistic) sound path differences (from the different loudspeakers to the receiver). Therefore, in practice, angular weighting windows (AWWs) are used to restrict the large amplitude of the NFC filters at low frequencies and particularly at high orders. In the literature, angular weighting windows with high-pass filters [3] or cosine-shaped windows [4] have been proposed. Recently, a regularization window [5] has been proposed which showed a better reconstruction of the sound field than the two previously mentioned angular weighting windows, in particular at low frequencies.

Restricting HOA to the horizontal plane (2D) allows for a better directivity and requires fewer loudspeakers for a given order and is therefore often used for practical arrays. However, it can be shown that 2D NFC-HOA is not able to correctly simulate the amplitude decay ($1/r$) of a point source with distance (r) when loudspeakers are modeled as point sources. This is an important limitation and therefore 2D and 3D NFC-HOA performance are separately discussed here.

For distances below one meter, low-frequency interaural level differences (ILDs) are the most important auditory cue [6] for judging the distance of sources outside the median plane. It has been shown that low-frequency ILDs below 750 Hz significantly improve distance perception of nearby sound sources. ILDs of up to 20 dB can be measured for very close sources [7], even at very low frequencies. In particular these low-frequency ILDs (i.e., below about 750 Hz) seem to be important for auditory distance perception of very close sound sources [6], and therefore, need to be elicited by sound fields that are reproduced by NFC-HOA.

In this manuscript, ILDs for NFC-HOA reproduced nearby sources were analyzed for different angular weighting windows both in 2D and 3D. Moreover, a listening test was conducted to investigate if NFC-HOA sources can be perceived as close as corresponding physical ones.

2. NEAR FIELD COMPENSATED HIGHER-ORDER AMBISONICS

NFC-HOA modifies the plane wave description of higher-order Ambisonics by applying near field coding filters to HOA components to reproduce point sources. These filters are described by a ratio of spherical Hankel functions of the second kind h_m^- and depend on the source distance from the center of the array ρ_s and the radius of the loudspeaker array R. They are described for 2D and 3D reproduction up to an Ambisonic order M for each degree $0 \leq m \leq M$ by:

$$H_m^{NFC(\rho_s,R)}(k) = \frac{h_m^-(k\rho_s)}{h_0^-(k\rho_s)} \frac{h_0^-(kR)}{h_m^-(kR)} \quad (1)$$

where $k = 2\pi f/c$ is the wavenumber. While the use of spherical Hankel functions in 2D representations is mathematically not justified, these functions provide an amplitude decay of the sound field with distance r that better approximates the $1/r$ amplitude decay of a point source. Consequently, spherical Hankel functions are also used in 2D reproductions [3].

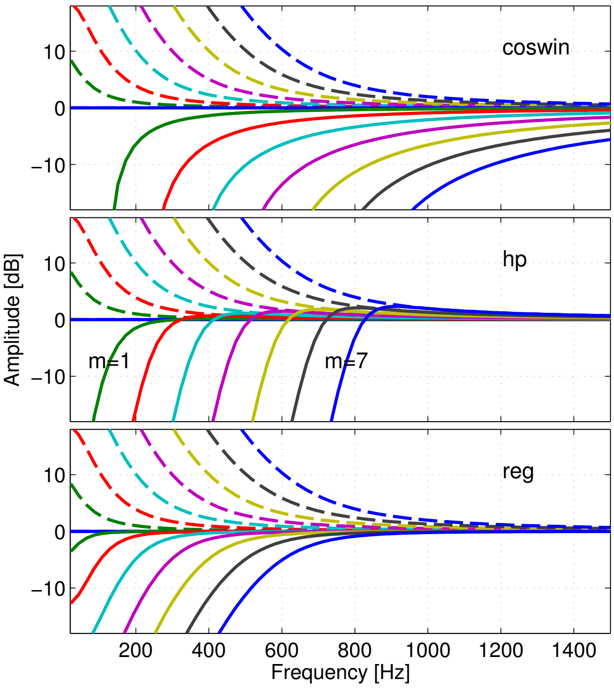

Figure 1: Amplitude response of unweighted (dashed lines) and weighted (solid lines) NFC filters for $m = 0$ to 7, for a virtual source at $\rho_s = 0.5$ m and a loudspeaker array radius of $R = 1.5$ m.

Applying the NFC filters given in Eq. 1 results in very large loudspeaker signals at low frequencies for close sources and high Ambisonic orders. Consequently, in practical realizations, angular weighting windows are introduced that limit low-frequency gain by only considering high-order Ambisonics components at high frequencies. Therefore, these weighting functions are designed to attenuate low-frequency amplitude proportionally to the component's degree m. The phase of the NFC filters is maintained to preserve the appropriate curvature of the sound field. In this paper, three angular weighting functions are investigated and the corresponding modified NFC filters amplitude responses are shown in Fig. 1 (solid lines) : (1) the cosine window ('coswin') mentioned in Ahrens and Spors (2009) [4]

$$w_m^{coswin} = \begin{cases} (cos(\frac{m\pi}{k\ \rho_s}) + 1)/2 & \text{if } m \leq k\ \rho_s \\ 0 & \text{if } m > k\ \rho_s \end{cases} \quad (2)$$

which leads to shallow high-pass filters with high cut-off frequencies, (2) the high-pass filters ('hp') from Daniel [3]

$$w_m^{hp} = \frac{\left(f/f_{lim}^{(m)}\right)^{4m}}{1 + \left(f/f_{lim}^{(m)}\right)^{4m}} \quad (3)$$

where $f_{lim}^{(m)}$ is derived from generic values listed in Table 1 in [3] which can be approximated by $f_{lim}^{(m)} = mc/(2\pi\rho_s)$ and lead to slightly positive gains in the resulting NFC filters, and (3) a new proposed regularization function ('reg') [5]

$$w_m^{reg} = \frac{2}{1 + \left|H_m^{NFC(\rho,R)}\right|^2} \quad (4)$$

which regularizes the NFC filters into sharp high-pass filters with low cut-off frequencies.

3. INTERAURAL LEVEL DIFFERENCES

Low-frequency interaural level differences are the main auditory cue for estimating distances of nearby sound sources outside the median plane. For sources at 0.125 cm from the center of the listener's head, for example, ILDs of up to 20 dB can be reached at 500 Hz [7].

ILDs obtained with NFC-HOA in an ideal setup (anechoic room, listener perfectly aligned in the center of the array) were computed using the KEMAR head-related transfer functions (HRTFs) to simulate a 2D 18-loudspeaker array and a 3D 33-loudspeaker array with a radius $R = 1.4$ m. Resulting ILDs are plotted in Fig. 2 for the three considered AWWs.

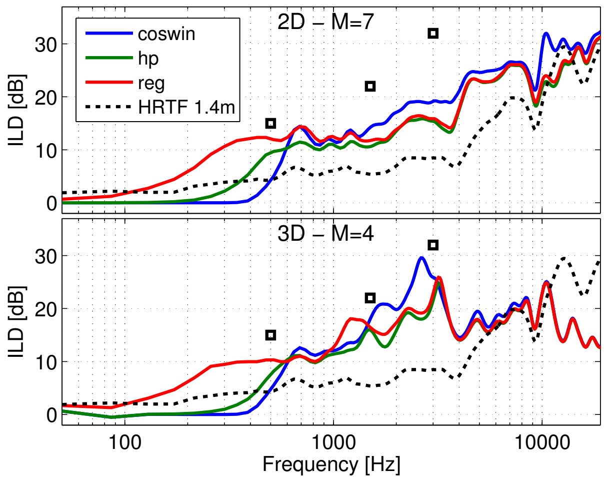

Figure 2: ILDs for NFC-HOA in ideal conditions for a virtual source at $\rho_s = 0.15$ m, $\theta = 90°$ with Ambisonic order $M = 7$ and $M = 4$ for the 2D (top) and 3D (bottom) case, respectively.

NFC-HOA with either one of the three considered AWWs provides increased ILDs for a very close source at 0.15 m compared to distant sources (1.4 m). The proposed regularization function provides significant ILDs from 250 Hz and upwards whereas the other AWWs do not show any ILD for frequencies lower than 500 Hz. For frequencies above 1.5 kHz, the cosine window provides the largest ILDs. For frequencies bellow 750 Hz, 2D NFC-HOA with the regularization function produces ILDs that are similar to ones measured for a physical source by Brungart and Rabinowitz [7] (represented by the square markers in Fig. 2). This AWW provides slightly lower ILDs for these low frequencies in the 3D case. For frequencies above 750 Hz, 3D NFC-HOA provides larger ILDs than for 2D reproductions and the cosine window provides ILDs at 1.5 and 3 kHz that are comparable to physical sources.

In practical situations, loudspeaker arrays are often used in an acoustically-damped room and the listener is never perfectly aligned in the center. Thereby, the performance of the three considered AWWs was also investigated by recording binaural signals in a realistic loudspeaker setup (i.e., within the "SpaceLab" facility at the Centre for Applied Hearing Research [8]). The same source ($\rho_s = 0.15$ m, $\theta = 90°$) was reproduced with

NFC-HOA over a 2D and 3D array of 16 and 29 loudspeakers, respectively, and the resulting ILDs are plotted in Fig. 3.

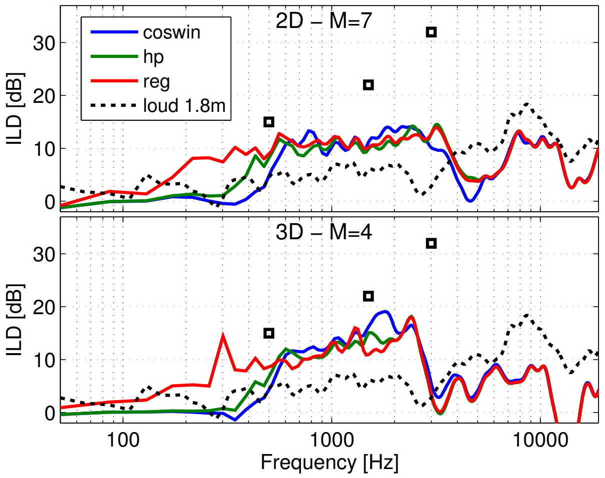

Figure 3: Recorded ILDs for NFC-HOA for a virtual source at $\rho_s = 0.15$ m, $\theta = 90°$.

As expected, due to disturbing reflections from the walls, ILDs were lower (especially from 1.5 kHz and above) than in the ideal condition but still significantly larger than for a (far) source at the loudspeaker array radius (dashed curve). The proposed regularization window is again the only one that is able to provide low-frequency ILDs from 250 Hz upwards in both 2D and 3D. The cosine windows does not provide notably larger ILDs than the other AWW for frequencies above 1.5 kHz and all AWWs fail to provide realistic ILDs above about 3 kHz.

4. LISTENING TEST

In order to evaluate the impact of the different angular weighting functions in NFC-HOA on distance perception of nearby sound sources, a listening test was conducted. The test compared the distance perception of frontal nearby sources, where no ILD cues are available, to that of sources on the "interaural axis", which naturally provide strong ILD cues. Since the high-pass filter weighting windows consistently provided lower ILDs than with the two other windows (see section 3), this function was not considered in the listening test.

4.1. Methods

Six normal hearing test-subjects participated in the experiment that took place in the "SpaceLab" facility at the Centre for Applied Hearing Research [8]. Nearby sound sources were simulated at different distances in front (0°) and on the left side (90°) of the listener with NFC-HOA, using the cosine window and the regularization function. In a 2D condition, 16 loudspeakers were used to realize 7^{th}-order NFC-HOA, and in a 3D condition, 29 loudspeakers were used to realize 4^{th}-order NFC-HOA. In each of these 8 stimulus conditions, blindfolded subjects seated at the center of the loudspeaker array were presented with 6 identical blocks of 24 stimuli at distances randomly chosen on a logarithmic axis between 0.125 and 1.7 m. In addition, 4 stimuli at a distance of R =1.8 m (corresponding to standard HOA) were included in each block. After each stimulus presentation, test-subjects were asked to orally give an estimate of the distance of the auditory event in centimeters.

The virtual sound sources played a train of 150-ms long bursts of pink noise, separated by pauses of 30 ms. In order to provide natural amplitude distance cues to the listener, the sound pressure at the center of the loudspeaker array was proportional $(1/\rho)$ to the source distance ρ. The maximum sound pressure level in the center of the loudspeaker array was 58 dB when the virtual sources was at 0.125 m.

4.2. Results

Figure 4 shows the distance estimate for an example subject ('TL') for the 8 conditions. Left and right panels represent the 2D and 3D condition and the top and bottom panels represent the cosine window ('coswin') and the regularization function ('reg') respectively. Each panel shows the logarithmic distance estimates ρ' against the logarithmic distance of the NFC-HOA sources ρ for the 0° (blue cross markers) and 90° (red plus marker) condition.

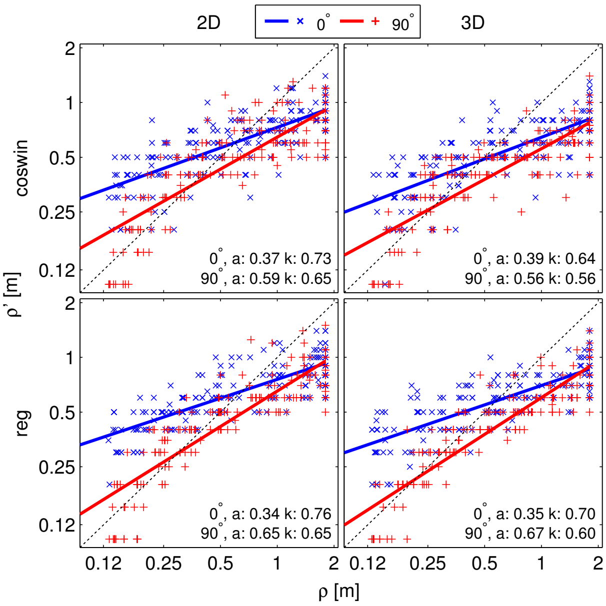

Figure 4: Distance estimations for an example subject for the 8 conditions.

For each condition, a power-law function $\rho' = k\rho^a$ was fitted to the data which corresponded to a linear fit in the double logarithmic scale. A power-law exponent a less than one indicates a compression of the estimated distances compared to the stimulus distances. The closer the exponent is to one, the better is the distance perception performance. Fitted functions are plotted in Fig. 4 on each panel for the 0° (blue line) and 90° (red line) condition. Note that this subject ('TL') systematically estimated sources at very short distances closer for left side sources (90°) than for frontal sources (0°) for both 2D and 3D conditions and both angular weighting windows.

Power-law exponents a are plotted in Fig. 5 for all subjects and conditions. The large variability of power-law exponents across the six subjects is in line with other distance perception experiments [9]. Note that power-law exponents are significantly positive for all subjects and conditions indicating that sources are perceived at different distances (inside the loudspeaker array radius R). This was expected as natural amplitude cues were provided in this experiment. In order to compare fitted power-law exponents a between the 0° and 90° condition for each subject, the parallelism of the linear fit (in the double logarithmic scale) was tested with an analysis of covariance (ANCOVA). The result of this test is indicated by the number on top of each bar-pair which either takes the value +1 or -1, when the power-law exponents are significantly increasing or decreasing, respectively, or 0 when they are not significantly different between the 0° and 90° condition.

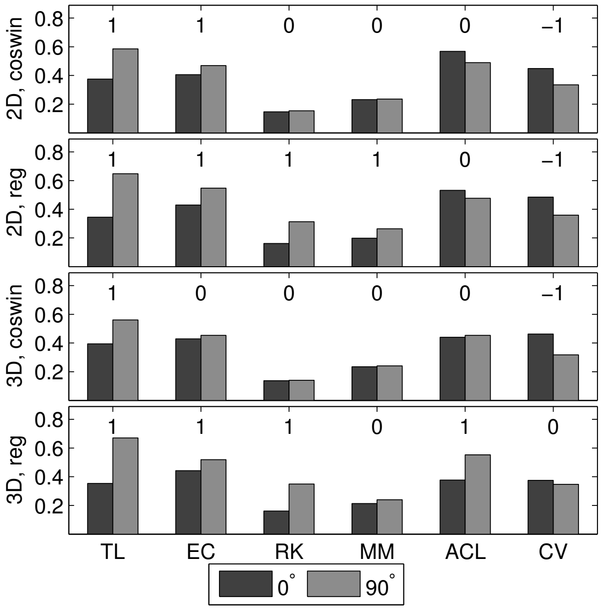

Figure 5: Fitted power law exponents a. The number on top of each bar-pair indicates if the power-law exponents are significantly different between sources on the side (90°) and on the front (0°) of the listener.

Four out of six subjects show a significant increase in power-law exponents from the 0° to 90° condition when the proposed regularization function was used (panel 2 and 4). However, only two (2D case) and one (3D case) show a similar increase when the cosine window was used (panel 1 and 3). It should be noted that subject 'CV' consistently showed a power-law exponent decrease from the 0° to 90° condition for both angular windows and 2D and 3D conditions.

In order to quantify the difference between the fitted exponents, the across-subjects mean of the difference between power-law exponents for left side and frontal virtual sources was computed. The mean differences were not significantly positive when the cosine window was used both in 2D (by 0.02) and 3D (by 0.01). This indicates that only amplitude cues, and no low-frequency ILDs, were available to the listener for sources on the interaural axis for NFC-HOA with the cosine window. When the regularization function was used, power-law exponents increased (by 0.08) from the frontal to the left side condition for the 2D condition and significantly increased (by 0.13) for the 3D condition. This improved distance perception performance indicates that NFC-HOA with regularization weighting windows provides extra auditory cues for lateral sources, which correspond to low-frequency ILDs. Since the main difference between the two considered angular weighting functions (according to section 3) is that the proposed regularization function provides significant ILD cues down to about 250 Hz (and the cosine window only down to about 500 Hz), it can be concluded that ILD cues below about 500 Hz are very important for auditory perception of very close sound sources. Hence, in NFC-HOA, such low-frequency ILD cues need to be provided, which requires a significant contribution of higher-order Ambisonics components at low frequencies. In addition, the better distance perception performance in 3D compared to 2D and the fact that larger ILDs were measured between 750 Hz and 3 kHz in the 3D condition (see section 3) indicate that the increase of ILDs in this frequency range enhances distance perception performance as shown in [6] for physical nearby sources.

4.3. Comparison with the distance perception of physical nearby sources

Brungart [6] investigated auditory distance perception of nearby sound sources using a physical point source. By analyzing the correlation coefficients between log estimates and log physical distances, he found that distance perception performance is better for lateral sources than for frontal sources. In order to compare their data to the experimental results described in section 4.2, the same analysis is applied here. The across subjects mean of the correlation coefficients for NFC-HOA nearby sources are plotted in Fig. 6 together with the values measured by Brungart. Correlation coefficients were averaged using the Fisher transform.

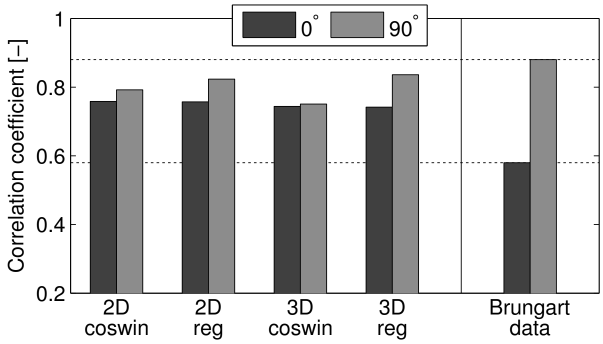

Figure 6: Mean correlation coefficients across subjects. The right bars indicates the mean correlation coefficient measured for physical nearby sound sources by Brungart [6].

Mean correlation coefficients shows similar behavior than power-law exponents described in section 4.2. Average correlation coefficients for sources on the interaural axis with the regularization window for 2D (0.82) and 3D (0.84) were slightly lower than the value of 0.88 measured by Brungart. This is in line with the lower measured ILDs with this AWW compared to ILDs measured with physical sources as indicated in section 3. For frontal sources (0°), average correlation coefficients across subjects were larger (between 0.76 and .74) for

NFC-HOA frontal sources (0°) than the one measured (0.58) for physical sources in Brungart's study. This implies that the performance improvement in distance perception for sources on the interaural axis is greatly reduced with NFC-HOA virtual sources. The relatively large correlation coefficients for frontal sources might be due to spectral artifacts introduced by NFC-HOA. As these artifacts become more salient as the distance decreases, they may have unintentionally provided an unnatural cue for distance. Further investigations are required to verify this hypothesis.

5. SUMMARY AND CONCLUSIONS

The impact of the use of angular weighting windows in NFC-HOA on nearby distance perception has been investigated in this study. A novel AWW, the regularization function, was introduced and the impact of this function on the most important cue in distance perception of nearby sources, i.e., low-frequency interaural level differences, was investigated and compared with two other available AWWs. It was shown that the proposed regularization function for NFC-HOA provides large low-frequency ILDs down to about 250 Hz, which, below about 750 Hz, are close to values measured for nearby physical sources. In contrast, the other considered angular weighting functions failed to provide ILDs bellow 500 Hz. The availability of this important low-frequency distance cue was confirmed by dummy-head recordings in a practical loudspeaker installation. In addition, it was found that 3D NFC-HOA representations provided larger ILDs than 2D representations for frequencies between 750 Hz and 3 kHz.

Results of the listening experiment showed a better distance perception performance for lateral sources than for frontal sources when the regularization window was used. This performance was greater in 3D that in 2D conditions. These results indicates that ILD cues, in addition to amplitude cues, were provided to the listener, highlighting that NFC-HOA can provide efficient ILD cues when the proposed regularization function is used. This distance perception improvement was not observed when an alternative AWW was used, which did not provide any ILD cues below about 500 Hz. By comparing the distance perception of virtual nearby sources reproduced by NFC-HOA (using the regularization function) with the perception of real sources [6], it was found that the virtual sources can be perceived nearly as accurately as corresponding physical ones.

Beside reproducing the accurate perception of distance, the accurate reproduction of the direction of the virtual nearby sound sources is also of great importance and needs to be investigated in a future experiment.

6. REFERENCES

[1] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *Proc. of the 23rd International Conference of the Audio Engineering Society*, 2003.

[2] ——, "Representation de champs acoustiques, application a la transmission et a la reproduction de scenes sonores complexes dans un contexte multimedia (in french)," Ph.D. dissertation, 1996-2000 Universite Paris 6, 2000.

[3] J. Daniel and S. Moreau, "Further study of sound field coding with higher order ambisonics," in *Proc. of the 116th Audio Engineering Society Convention*, 2004.

[4] J. Ahrens and S. Spors, "Spatial encoding and decoding of focused virtual sound sources," in *Proc. of the 1st Ambisonics Symposium*, 2009.

[5] S. Favrot and J. Buchholz, "Reproduction of nearby sound sources using high-order ambisonics: Implementation and evaluation," in *Proc. of DAGA2010*, 2010.

[6] D. Brungart, "Auditory localization of nearby sources. III. Stimulus effects," *The Journal of the Acoustical Society of America*, vol. 106, p. 3589, 1999.

[7] D. Brungart and W. Rabinowitz, "Auditory localization of nearby sources. Head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 106, p. 1465, 1999.

[8] S. Favrot and J. M. Buchholz, "LoRA - A loudspeaker-based room auralization system," *Acta Acustica united with Acustica*, In press.

[9] P. Zahorik, D. Brungart, and A. Bronkhorst, "Auditory distance perception in humans: A summary of past and present research," *Acta Acustica united with Acustica*, vol. 91, no. 3, pp. 409–420, 2005.