Audio Engineering Society

Convention Paper

Presented at the 124th Convention
2008 May 17–20 Amsterdam, The Netherlands

Wave Field Synthesis Evaluation using the Minimum Audible Angle in a Concert Hall

Georgios Marentakis[1], Etienne Corteel[2], Stephen Mc Adams[1]

[1] *CIRMMT, Schulich School of Music, McGill University, 527 Sherbrooke St. West Montreal, Quebec, Canada H3A 1E3*

[2] *sonic emotion ag, Eichweg 6, CH-8154 Oberglatt, Switzerland*

Correspondence should be addressed to Georgios Marentakis (gmarentakis@music.mcgill.ca)

ABSTRACT

Localization accuracy with Wave Field Synthesis (WFS) was estimated in a variable-acoustics concert hall. Contrary to previous studies, we employed a Minimum Audible Angle (MAA) paradigm as a measure of localization performance. The MAA was estimated for three different listening positions, three orientations of the listeners (0°,60°, 90°) and two acoustical conditions. WFS was found to produce satisfying localization cues that depend little on the reverberation time of the room and only weakly on the position of the listener.

INTRODUCTION

Wave Field Synthesis (WFS) is an advanced rendering technique that enables the synthesis of sound fields within an extended listening area [1]. Virtual sources may thus be localized accurately within an extended listening area, typically for the entire audience in a concert hall. However, some artifacts may alter localization in such a situation [2]. For example, the spatial sampling of the loudspeaker distribution used in WFS introduces aliasing artifacts that distort the reproduced sound field above the so-called aliasing frequency. The natural room effect (of the early reflections and reverberation) is also a known factor that may alter sound localization. For all these reasons, evaluation methodologies for WFS are of particular interest. Perceptual evaluation of 3D audio systems is a large field where a quantitative approach is undertaken to assess the perceptual qualities that are reproduced by a system. These include localization, spaciousness, listener envelopment and coloration. In this work, we are interested in evaluating wave field synthesis by estimating the Minimum Audible Angle (MAA) under certain combinations of parameters of interest in the deployment of a Wave Field synthesis system in a concert hall.

0.1. Localization accuracy and MAA

Auditory localization performance is evaluated by either estimating the Localization Error or the Minimum Audible Angle. Localization Error is estimated by asking listeners to indicate the direction of sound incidence. This is done either using a visual reference system, or in more recent studies aligning an acoustical pointer with the perceived direction of sound incidence [4] [5]. Localization error studies are useful because they reveal biases in the perception of sound incidence. It is therefore a measure of how well the synthesized sound location is perceived by listeners. However, when a visual reference system is used or participants are not blindfolded they are prone to sensory bias.

The Minimum Audible Angle (MAA) is a measure of the smallest perceivable angular displacement. Mills [Mills58], studied MAAs in an anechoic chamber with real sounds. He showed that they depend strongly on the frequency of the stimulus and the direction of sound incidence. For frontal incidence, MAA remains relatively constant around 1° up to 1500 Hz. Between 1500 Hz and 3000 Hz localization ability diminishes to improve again above 3000 Hz. This behavior is also observed for directions other than frontal. MAAs also depend on the interstimulus interval. For inter-stimulus intervals larger then 150 msec they remain essentially constant [14].

MAA s are estimated by presenting a stimulus in the direction of interest followed by a displaced stimulus and listeners indicate the perceived direction of sound displacement. When the effect of the frequency of the sound is not of direct interest, it is usual to perform the estimation using white noise bursts. Saberi et al. [9], estimated MAAs to be 1°, 1.6° and about 5° for directions of incidence of 0°, 60° and 90° respectively.

The Minimum Audible Angle and the Localization error are two complementary ways to estimate auditory localization abilities. Hartmann [12] indicates that in the absence of sensory bias it can be thought to relate to the standard deviation of the localization judgments.

0.2. Localization studies for Wave Field Synthesis

Localization in WFS systems has been studied mainly through absolute localization studies. Verheijen [10] estimated localization error for virtual sources synthesized using a 24-channel monopole electrodynamic array and compared it to real sounds. Measurements were carried out both in an anechoic and a medium-sized listening room (RT_{60} = 0.5s). Two loudspeaker spacings of $\Delta x = 22$ and $\Delta x = 11$ cm were used. Virtual sources were positioned 2 m behind the loudspeaker array. The stimuli were 1.5 s white-noise bursts and bandlimited versions of them. The bandlimited noise stimuli did not contain frequency components above the respective aliasing frequency for each loudspeaker spacing. This was 775 Hz for $\Delta x = 22$ cm and 1550 Hz for $\Delta x = 11$ cm. Seven directions of sound incidence were evaluated, spaced equally between $-25°$ and 25° degrees, with 0° corresponding to the middle of the loudspeaker array. Participants indicated which of the labeled loudspeakers was closest to the perceived location of the stimulus. There were two repetitions for each combination of the experimental factors. The authors found that localization judgments for white noise were not significantly different from the true location of the sound. The deviation was increasing however as sounds were displaced from the middle of the loudspeaker array. In the anechoic room, mean localization error for white noise was 1.9° (1.1°) and 2.2° (1.1°) degrees for virtual and real sources respectively. In the listening room, the error became 2.6° (1.1°) and 2.5° (1.1°) respectively. There was no clear effect to delimiting the bandwidth to frequencies less than the aliasing frequency. There was also no significant effect of increasing the spacing between the loudspeakers.

Start [11] performed in situ localization error measurements on a Wave Field Synthesis system in an anechoic room, an auditorium (RT_{60}=1.1 sec.) and a concert hall (RT_{60}=2.3 sec.) for noise stimuli of variable bandwidth. Different loudspeaker arrays where used in the three conditions leading to an aliasing frequency of 1.4 kHz in the anechoic chamber, 1.2 kHz in the auditorium and 750 Hz in the concert hall. Participants indicated the perceived direction of sound incidence by indicating which of the labeled loudspeakers was closest to the perceived direction of sound incidence. There were five repetitions for each experimental condition. Localization performance for broadband stimuli was quite similar in the cases of the auditorium and the con-

cert hall but deteriorated compared to the anechoic chamber. Localization accuracy was best for speech stimuli. In general synthesized sources were on average less well localized than real ones in all acoustic conditions, the discrepancy increasing for broadband noise stimuli. However, for stimuli below the spatial aliasing frequency and under anechoic conditions, localization performance was as good as with real sounds. In contrast, for stimuli bandlimited above the spatial aliasing frequency, localization performance was markedly worse than for real sources. For white noise stimuli, the localization bias was on average 3.0°, 5.9° and 4.2° and the standard deviation 1.4°, 3.0° and 2.2° in the anechoic chamber, the auditorium and the concert hall respectively. It is however difficult to compare these results between each other since too many parameters changed (array shape and size, aliasing frequency, listening and source positions, listening room acoustics).
Furthermore, in his PhD thesis [2], Start performed Minimum Audible Angle estimation for sounds in front of a user for a Wave Field Synthesis system. Real and synthesized noise stimuli were recorded in an anechoic room using a KEMAR artificial head and then played back to the listeners through headphones. The estimation was done for frontal sound incidence for stimuli bandwidths of 100-1500 Hz and 100-8000 Hz bandwidths respectively. Mean MAAs were 1.15° (0.16°) and 0.79° (0.09°) versus 1.1° (0.11°) and 0.77° (0.11°) for real and virtual sources and 100-1500 Hz and 100-8000 Hz respectively. Participants judged the direction of sound displacement vs. a reference sound stimulus in the center of the array. Stimuli were 1 s. in duration and the interstimulus interval was 300 ms.

0.3. Motivation for the present study

The literature review shows that the Minimum Audible Angle for Wave Field Synthesis has not been estimated for directions of sound incidence other than frontal and in a real world setting such as the concert hall. In addition, it has not been estimated at different listening positions in a concert hall.
The minimum audible angle also provides interesting data from a system design point of view. Multichannel equalization techniques may be used to compensate for rendering artefacts in Wave Field Synthesis for an ensemble of target sources that should span all possibly synthesized virtual sources on a given loudspeaker configuration [6]. Finite Impulse Response filters are calculated for each loudspeaker and each source and stored in a database. The calculation of the database is thus linearly linked to the number of sources. Studies on the Minimum Audible Angle enable thus to optimize both database size and calculation time by providing guidelines for the definition of the required source ensemble.
For these reasons, we undertook the following experimental study to investigate how the aforementioned parameters affect the Minimum Audible Angle for the particular Wave Field Synthesis system.

1. DESCRIPTION OF THE SYSTEM AND THE CONCERT HALL

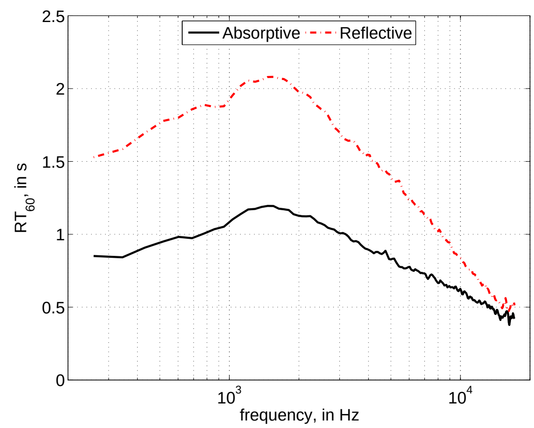

Fig. 1: Reverberation time (RT_{60}) for the absorptive and the reflective configurations of the ESPRO

The experiment took place at IRCAM in the Espace de Projection (ESPRO). The ESPRO is a 375 m^2 variable acoustics concert hall (23.5m(Length) × 15.5m(Width) × 11m(Height)). The prismatic modules inserted into the walls, referred to as peri-actes in the latter, have 3 surfaces with different acoustical properties (absorptive, diffusive, reflective). By choosing different configurations and modifying ceiling height, the reverberation time can be adjusted from 0.4 to 4 s.
Two different configurations of the concert hall are used. The ceiling height was not changed so that the volume of the concert hall remained the same (c.a. 3700 m^3). In the "absorptive" configuration, all peri-actes are arranged to present their absorptive

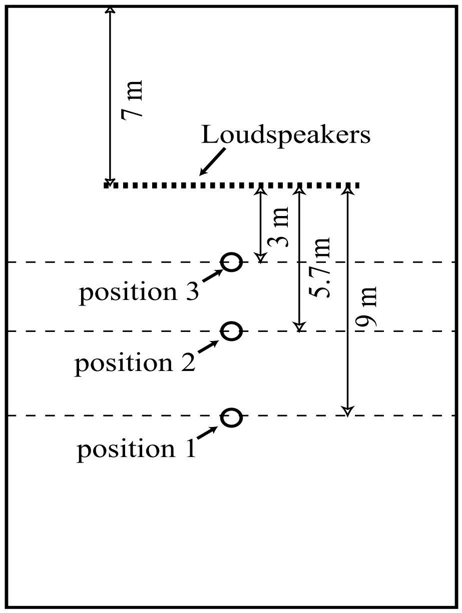

Fig. 2: Position of loudspeakers and test positions in the concert hall

surface. In the "reflective" configuration, the lower part of the side walls was set to reflective and the ceiling was half absorptive and half diffusive. This configuration increases the energy of side reflections. The corresponding reverberation times are displayed in Figure 1.

A 48-channel loudspeaker array was installed in the concert hall at 7 m from the back wall (see Figure 2 for details). The array was positioned at 4 m height which can be regarded as a realistic configuration for a front stage-loudspeaker array position in the context of a mixed-music performance involving real and synthetic instruments.

The total length of the array was 8.8 m. We used a double logarithmic spacing as displayed in Figure 3. Loudspeakers have generally larger spacing for center loudspeakers (24 cm) than for side loudspeakers (13 cm). Figure 4 displays the aliasing frequency calculated for each source angle and listening position with the method proposed in [7]. It shows that the aliasing frequency is not the same for all sources and listening positions ranging from 1.1 to 1.65 kHz: The aliasing frequency is generally increasing with listening position distance

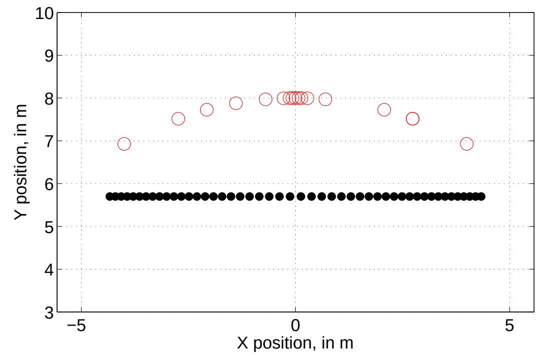

Fig. 3: Loudspeaker (black dots) and sources (red circle) positions

and decreasing with absolute value of the source angle. The aliasing frequency was also estimated for the same ensemble of sources and listening positions considering an array of same length (8.8 m) but with a regular spacing (18.3 cm) of the 48 loudspeakers. This additional estimation gives an increase of 20 % on average of the aliasing frequency due to the double logarithmic spacing compared to the regularly spaced array. This is consistent to the results of [7] in which a 3.6 m long array of 24 loudspeakers was used.

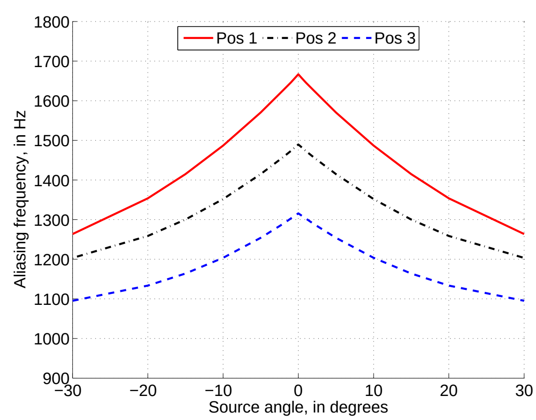

Fig. 4: Aliasing frequency calculated for the three listening positions

2. METHOD

The estimation of the Minimum Audible Angle was performed for two acoustic conditions and three directions of sound incidence at three different positions in the concert hall as shown in Figure 2. Independent variables were listening seat, direction of sound incidence and acoustic condition.

2.1. Participants

Nine participants, ranging from 24 to 45 with an averaged age of 29, participated in the experiment. All participants reported having normal hearing.

2.2. Apparatus & Materials

A computer was used to control the WFS system and present the stimuli in the desired order. Participants indicated the perceived direction of sound displacement using two keys on a keyboard. Labels were installed at 0°, 60° and 90° relative to the frontal direction of each listening seat. Participants were asked to turn and look at the desired label to estimate corresponding MAA s. The stimulus was a 250 ms burst of white noise with 5 ms cosine rise and decay. Virtual sources were located 2 m behind the loudspeaker array (8 m from listening seat 2) at an angular range of ($-30°$ to $30°$). The reference source position (0°) was centered in the middle of the loudspeaker array. Figure 3 displays the position of the virtual sources. The reverberation time for the two acoustic conditions, was $RT_{60} = 1.0$ s and $RT_{60} = 2.0$ s. Three listening positions were chosen, centered towards the hall and the loudspeaker array. Seat locations were chosen such that the direct sound attenuated by 3 dB from one position to the next. Assuming that the mean ear level of the participant was at 1.2 m from the floor, the elevation of the array was at 17°, 26° and 43° respectively for positions 1, 2 and 3.

2.3. Procedure

Three experimental sessions were performed for each acoustic condition, each with three listeners. Listeners were seated at three locations along an axis perpendicular to the center of the loudspeaker system. MAA was estimated using a 2AFC procedure in which participants listened to a reference sound coming from the middle of the loudspeaker array and a displaced version of the same sound with a delay of 350 ms. They then had to infer the direction of sound displacement (clockwise vs. counterclockwise). A variation of the constant stimuli method was used where sound displacement was varied randomly within each trial set in one of four predetermined values. These were: 0.5°, 1°, 5° and 10° for frontal, 2°, 5°, 10°, 15° for oblique and 5°, 10°, 20°, 30° for lateral incidence. Both left-right and right-left sequences were presented, with seven repetitions each. MAAs were estimated separately for each direction of sound incidence in a random order for each experimental session. The same group of nine people participated in both acoustic conditions.

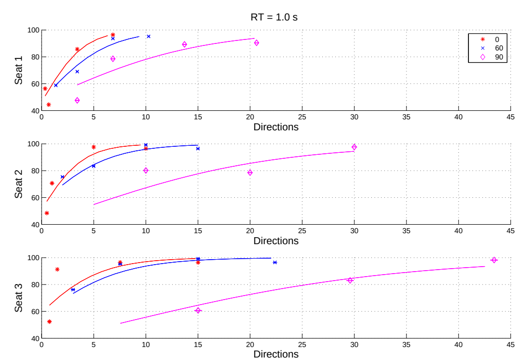

Fig. 5: Percent Correct Identification and fitted Logistic Functions for $RT_{60} = 1.0$ s

3. RESULTS

Percent correct identifications based on pooled data and the associated logistic regression curve fit are presented in Figures 5 & 6. The angular displacement for each data point has been scaled to correspond to the apparent sound displacement at the position of the listener. There is obviously a strong effect of the direction of the sound event and a lesser effect of the listening seat.

Based on the logistic function curve fit thresholds corresponding to 75% correct identification performance were calculated. Thresholds for each individual subject and median values are presented in Table 1.

In the cases where the estimation of the threshold was not possible, data were replaced with the median value and an Analysis of Variance was performed. There was a significant main effect of the room, $F(1,8) = 6.85$, $p < 0.05$, the listening seat, $F(2,16) = 28.796$, $p < 0.001$ and the direction of

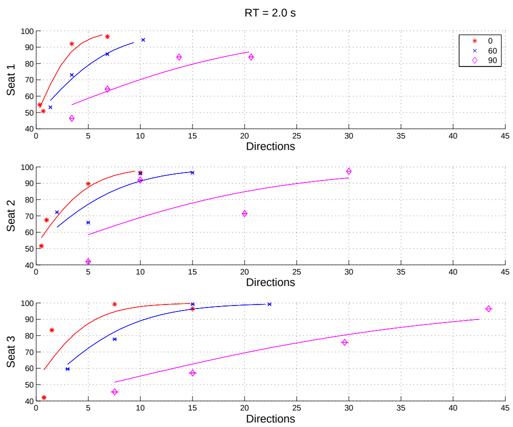

Fig. 6: Percent Correct Identification and fitted Logistic Functions for RT_{60} = 2.0 s

sound incidence, F(2,16) = 138.82, $p < 0.001$. Furthermore, there were significant two-way interaction between the listening seat and the direction of sound event, F(4,32) = 29.925, $p < 0.001$. Post-Hoc Tukey HSD comparisons showed all directions of sound incidence to differ significantly among them. Post-Hoc Tukey HSD comparisons showed listening seat 3 to differ significantly from both other seats but seats 1 and 2 were not significantly different.

4. DISCUSSION

The results indicate that localization in Wave Field Synthesis systems follows similar tendencies as localization in the real world. Localization ability diminishes for oblique and lateral incidence. The extent to which this happens depends on the listening seat, as revealed by the interaction between listening seat and the direction of sound event.

It is evident from Table 2 that MAAs for 60° are about 2 times those for 0° and MAAs for 90° are about 5.5 times those for 0° for seats 1 and 2. However, for Seat 3 there is significant degradation in localization acuity for oblique and lateral direction of sound incidence. This is attributed to the fact that Seat 3 was too close to the (elevated) loudspeaker array. We are currently undertaking acoustical analyses to investigate the effect. It is evident though that there is a limit to how close to the loudspeaker listeners can be placed without significant degradation of the localization cues.

		Absorptive			Reflective		
Seat	Subject	0°	60°	90°	0°	60°	90°
	S1	3.9	3.3	NaN	2.7	5.3	NaN
	S2	1.3	3.1	14.7	1.2	9.7	NaN
	S3	3.4	4.8	15.1	2.0	1.8	NaN
	S4	1.9	5.7	NaN	1.9	4.4	20.0
1	S5	2.2	3.6	8.6	2.0	4.1	NaN
	S6	3.9	3.5	NaN	3.8	4.1	NaN
	S7	1.7	3.3	6.0	2.1	4.0	10.3
	S8	2.2	3.0	4.8	1.5	3.7	17.7
	S9	1.9	5.1	7.9	1.9	4.8	5.5
	Median	2.2	3.5	8.2	2.0	4.1	14.0
	MAD	(0.8)	(0.9)	(3.6)	(0.6)	(1.3)	(5.4)
	S1	1.5	5.1	28.5	0.8	8.2	25.7
	S2	2.8	2.8	16.6	4.6	3.1	15.3
	S3	1.8	7.8	15.7	1.4	3.8	26.4
	S4	1.2	1.3	15.5	2.4	4.8	13.6
2	S5	1.7	2.4	12.0	2.4	2.9	10.7
	S6	1.8	3.2	11.2	4.9	6.7	13.6
	S7	2.7	1.4	8.2	2.6	4.9	11.2
	S8	2.9	2.2	9.7	3.0	5.0	20.6
	S9	2.5	2.8	6.8	2.4	2.1	3.4
	Median	1.8	2.8	12.0	2.4	4.8	13.6
	MAD	(0.7)	(1.4)	(4.7)	(0.9)	(1.5)	(4.8)
	S1	2.4	5.7	26.9	2.6	8.3	28
	S2	1.0	3.4	16.8	3.0	4.2	18.8
	S3	2.2	3.7	23.2	2.5	6.4	38.1
	S4	2.7	4.4	NaN	4.0	9.0	NaN
3	S5	2.7	4.9	29.4	3.3	6.8	21.6
	S6	2.2	2.3	16.5	3.8	3.9	28.1
	S7	2.9	2.5	12.8	2.5	5.8	18.2
	S8	1.8	3.4	18.1	3.0	7.7	30.0
	S9	2.4	3.4	28.6	2.7	7.2	20.0
	Median	2.4	3.4	20.7	3.0	6.8	24.8
	MAD	(0.4)	(0.9)	(5.5)	(0.4)	(1.3)	(5.7)

Table 1: MAA estimations for each participant as well as median and median absolute deviation (MAD) for the conditions in the experiment

Localization acuity was greatest in Seat 2. Localization was also satisfactory in Seat 1, however, a number of participants (1 & 6) produced high MAAs for frontal incidence. This was examined informally after the experiment and was found to depend on a certain discontinuity of the reproduced sound field that could nonetheless be resolved through head movements. Participants 1 & 6 however, did not take advantage of this and for this reason their thresholds are higher than the rest of the population.
Increasing the reverberation time resulted in higher MAAs. Averaged over sound incidence and listening seat MAAs were about 20% higher in the reverberant room. This is not surprising given that the room was transformed so as to increase the lateral reflec-

Seat	60°	90°
1	1.9	5.5
2	1.8	6.3
3	1.9	8.6

Table 2: Degradation in MAA as a function of the angle of sound incidence averaged over the two acoustic settings.

tions, a factor that is known to decrease localization acuity. The extent of the degradation is relatively small in terms of absolute values. It should be noted, however that localization acuity is relatively insensitive to early reflections for stimuli of substantial temporal variation as those used in the study (white noise). For sounds with harmonic structure the effect may be bigger.
Our results for frontal incidence are higher than those obtained by Start [2]. This can be attributed to the effect of the reverberation. In the study by Start, participants listened to anechoic stimuli over headphones, whereas this study was performed in a concert hall. In addition, the accuracy of the MAA estimators would be improved by using more data points and repetitions for the estimation of the psychometric function. The estimated values are close but higher than those obtained for real sounds in an acoustically damped room by a factor that depends on the direction of sound event and the listening seat, a consequence of the deployment of the system in a real world setting.

5. CONCLUSION

In conclusion, we evaluated localization accuracy with Wave Field Synthesis (WFS) in a variable acoustics concert hall, using the Minimum Audible Angle (MAA) paradigm as a measure of localization performance. The MAA was estimated for three different listening positions, three orientations of the listeners (0°, 60°, 90°) and two acoustical conditions. WFS was found to produce satisfying localization cues that depend little on the reverberation time of the room and only weakly on the position of the listener.

ACKNOWLEDGEMENTS

This work was partially supported by IRCAM and a grant by the National Council of sciences and Engineering of Canada. The authors would like to thank to Olivier Warusfel for having provided the room and assistance to prepare the test. Thanks to all participants of the test.

6. REFERENCES

[1] Berkhout, A. J.; de Vries, D.; Vogel, P., Acoustic Control by Wave Field Synthesis Journal of the Acoustical Society of America, 1993, 93, 2764-2778.

[2] Start, E. W. Direct Sound Enhancement by Wave Field Synthesis, PhD thesis, TU Delft, Delft, Pays Bas, 1997.

[3] Verheijen, E. N. G. Sound Reproduction by Wave Field Synthesis TU Delft, Delft, Pays Bas, 1997.

[4] V. Pulkki and T. Hirvonen. Localization of virtual sources in multichannel audio reproduction. IEEE Trans.on Speech and Audio Proc., 13(1), January 2005.

[5] Sanson J.; Corteel E. and Warusfel O. Objective and subjective analysis of localization accuracy in Wave Field Synthesis. 124th Convention of the Audio Engineering Society, Amsterdam, the Netherlands, 2008.

[6] Corteel, E., Equalization in extended area using multichannel inversion and Wave Field Synthesis, Journal of the Audio Engineering Society, 2006, 54.

[7] Corteel, E., On the use of irregularly spaced loudspeaker arrays for Wave Field Synthesis, potential impact on spatial aliasing frequency, 9th Int. Conference on Digital Audio Effects (DAFx-06), 2006.

[8] Mills A. W., On the Minimum Audible Angle, The Journal of the Acoustical Society of America, 30(4), pp. 237 - 236, 1958.

[9] Saberi. K., Dostal, L, Sadralodabai, T. Perrott, D.R., Minimum Audible angles for Horizontal, Vertical and Oblique Orientations: Lateral and Dorsal Planes, Acustica, 75, 1991.

[10] Verheijen, E., van Tol, P. and Boone, M., Evaluation of Loudspeaker Arrays for Wave Field Synthesis in Audio Reproduction, 98th Convention of the Audio Engineering Society, 1995.

[11] Start, E., Roovers, M. and de Vries, D., In Situ Measurements on a Wave Field Synthesis System For Sound Enhancement. 102nd Convention of the Audio Engineering Society, 1997.

[12] Hartmann, W, On the minimum audible angle - A decision theory approach, The Journal of the Acoustical Society of America, 85(5), pp. 2031-2041.

[13] Hartmann, W, Localization of sound in rooms, The Journal of the Acoustical Society of America, 74 (5), 1380-1391.

[14] Perrott, D. and Pacheco, S., Minimum audible angle thresholds for broadband noise as a function of the delay between the onset of th elead and lag signals. The Journal of the Acoustical Society of America, 85(6), pp. 2669-2672.