Objective and subjective comparison of electrodynamic and MAP loudspeakers for Wave Field Synthesis

Etienne Corteel[1,2], Khoa-Van NGuyen[1], Olivier Warusfel[1], Terence Caulkins[1], Renato Pellegrini[2]
[1] *IRCAM, 1. Pl. Igor Stravinsky, 75004 Paris, France, Email: etienne.corteel@ircam.fr*
[2] *sonic emotion, Eichweg 6, CH-8154 Oberglatt, Switzerland, Email: renato.pellegrini@sonicemotion.com*

Introduction

Wave Field Synthesis (WFS) is a multichannel sound rendering technique that allows for the synthesis of physical properties of sound fields within an extended listening area [1]. It relies on a large number of closely spaced loudspeakers (typically 15-20 *cm*) forming an acoustic aperture through which the target sound field (as emanating from a target sound source) propagates into the listening environment.
Practical implementation of WFS requires simplifications to the underlying physical principles (Kirchhoff-Helmholtz and Rayleigh integrals). Real loudspeakers radiation characteristics may also contribute to alter the synthesized sound field compared to the target one. Two types of loudspeakers are used nowadays for Wave Field Synthesis (see figure 1):

- array-mounted electrodynamic loudspeakers,
- Multi-Actuator Panels (MAP).

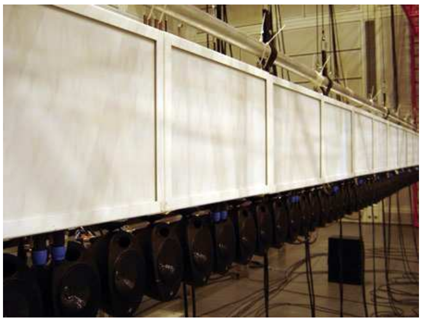
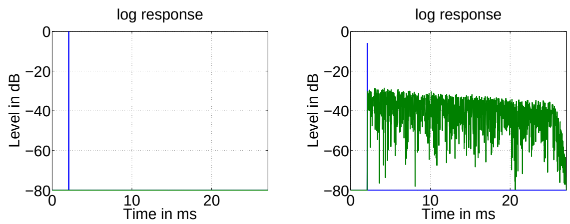


Figure 1: MAP and electrodynamic loudspeakers

MAP loudspeakers have been recently proposed as an alternative to electrodynamic loudspeakers for WFS [2]. Thanks to their low visual profile, MAP loudspeakers were originally thought as a way to facilitate the integration of tens to hundreds of loudspeakers in an existing environment. They are known to exhibit non-ideal radiation properties but can be partly compensated for using dedicated multichannel equalization techniques [3]. Contrary to electrodynamic loudspeakers, they also exhibit "diffuse" radiation properties [4] particularly at high frequencies.
The goal of current studies is to compare, at an objective and a subjective level, the transparency of Wave Field Synthesis rendering using electrodynamic or MAP loudspeakers. A companion paper [4] focuses on the characteristics of the free field radiation of loudspeaker arrays and its influence on perceptual dimensions such as coloration, Auditory Source Width (ASW) and angular localisation. The goal of this paper is to perform an objective analysis and a listening test dedicated to distance perception.

Diffuse filtering for Wave Field Synthesis

Figures 2(a) and 2(b) display impulse responses of filters that may be used for WFS rendering. The impulse response of figure 2(a) is a "classical" WFS filter (a delayed, possibly attenuated dirac pulse). Figure 2(b) displays the impulse response of a partly "diffuse" filter that will be referred to as Discrete-Diffuse (DD) filter. The diffuse part of the filter is obtained from a time limited white noise that is generated independently for each loudspeakers in order to obtain uncorrelated outputs. The DD filter is designed so as to reproduce the known diffuse properties of MAP loudspeakers on electrodynamic loudspeakers. For more details on the implementation, see [4].
Below the spatial aliasing frequency, the filters are designed using the multichannel equalization (MEQ) technique from [3] which enables to appropriately synthesize the sound field within an extended listening area. Therefore, the DD filter is only applied above the spatial aliasing frequency which is approximately at 1200 Hz for the considered configuration. In [4], the DD filter is also shown to reduce spatial color variations that can be experienced while wandering in a WFS sound installation.



(a) Discrete filter (MEQ)

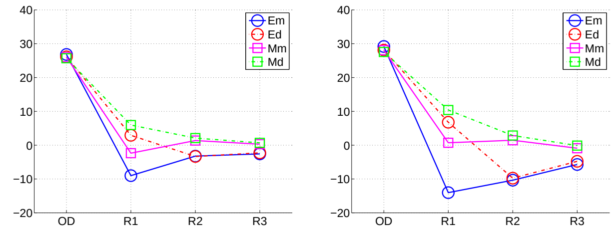
(b) Discrete-diffuse filter (DD)

Figure 2: Various diffusing filters used at high frequencies

Objective analysis

Two "identical" 48 channel, 15 *cm* spacing (i.e. 7.2 *m* long), loudspeaker arrays were installed in the Espace de Projection (see figure 1), a 22(l)×15(w)×11(h) m^3 variable acoustic concert hall at IRCAM. All surfaces (periactes) are set to absorptive (walls and ceiling). The obtained reverberation time is therefore below 1 s at all frequencies.

Room impulse measurements were achieved using an omnidirectional microphone located at 3 closely spaced (5 cm) positions at 3 m from the loudspeaker array, 1.5 m to the left of the center of the loudspeaker array. Figures 3(a), and 3(b) display the energy calculated in 4 temporal sections (0D: 0-20 ms, R1: 20-40 ms, R2:40-100 ms, R3: after 100 ms) and 7 octave bands (125/250/500 Hz, 1/2/4/8 kHz). This analysis is inspired from the low-level model used in the IRCAM's virtual room processor (le spatialisateur [5]). Figure 3(a) displays mean values averaged on the three microphone positions and all octave bands. It can be seen that MAP loudspeakers have more energy in later temporal sections (R1/R2/R3/R4) which may indicate an increased distance compared to electrodynamic loudspeakers. Diffusion only modifies energy in R1 section (20-40 ms) on both loudspeaker types. However, both observations are particularly noticeable at high frequencies (2/4/8 kHz, see figure 3(b)).



(a) Mean value on all octave bands

(b) High frequencies (2/4/8 kHz octave bands)

Figure 3: Energy in temporal section (0D: 0-20 ms, R1: 20-40 ms, R2:40-100 ms, R3: after 100 ms) and 7 octave bands (125/250/500 Hz, 1/2/4/8 kHz) (Em: electrodynamic MEQ filtering, Ed: electrodynamic DD filtering, Mm: MAP MEQ filtering, Md: MAP DD filtering).

Subjective test

A pair comparison direct scaling method is used. The task of the subjects is to indicate on a continuous scale if the second stimulus is closer, at the same distance, further than the first stimulus. The loudspeakers are hidden by an acoustically transparent curtain and all lights are dimmed so as to limit biasing visual cues. The subjects were positioned at 1.5 m to the left and 4.5 m distance from the loudspeakers so as to avoid discrimination based on elevation difference (less than 5 degrees in this setup). A virtual room processor is used in order to elicit three levels of distance (close: no additional room effect, mid distance, far distance) [6]. The room effect is rendered using three virtual loudspeakers on the WFS array and 6 side and rear loudspeakers. The chosen sound material is a guitar excerpt that was used in [6] to validate the virtual room model on a binaural setup. Two virtual sources positions were used (3 m behind the loudspeaker array and 3 m to the left, 8 m behind the loudspeaker array and 3 m to the right). 4 stimuli, thus 6 pairs, are randomly presented in both orders (12 and 21) in 6 conditions (elicited distance*virtual source position) forming a total of 72 pairs with no possible repetition. 11 subjects completed the test in an average of 20 minutes.
The results of the test are simply analyzed by extracting

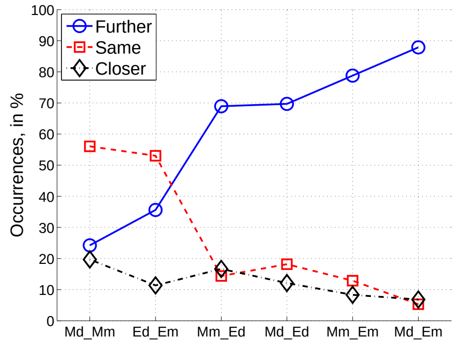


Figure 4: Results for distance test. Occurrences of responses indicating further, same, or closer distance of stimulus 1 versus stimulus 2 (Em: electrodynamic MEQ filtering, Ed: electrodynamic DD filtering, Mm: MAP MEQ filtering, Md: MAP DD filtering).

the number of occurrences indicating a greater, smaller, or same distance for each of the 6 pairs. These are presented in figure 4. MAP loudspeakers are shown to significantly increase perceived distance compared to electrodynamic loudspeakers. Diffusion has a similar but less pronounced effect, especially for electrodynamic loudspeakers.
Single way analysis of variance did not show any significant influence of either pair order ($p > 0.9$) or elicited distance with virtual room effect ($p > 0.7$). Only a loose influence of source position can be noticed ($p \sim 0.2$).

References

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

[2] Boone, M.M.: Multi-Actuator Panels (MAPs) as Loudspeaker Arrays for Wave Field Synthesis. Journal of the Audio Engineering Society (52) 2004, 712-723

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

[4] Corteel, E., NGuyen, K., Warusfel, O., Caulkins, T. & Pellegrini, R.S.: Objective and subjective comparison of electrodynamic and MAP loudspeakers for Wave Field Synthesis. 30^th conference of the Audio Engineering Society, 2007

[5] Jot, J.-M., Warusfel, O., Kahle, E., & Mein, M.: Binaural Concert Hall Simulation in Real Time, IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics, 1993, Mohonk, USA

[6] Pellegrini, R. & Horbach, U.: Perception-based Design of Virtual Rooms for Sound Reproduction. 112^{th} convention of the Audio Engineering Society, 2002