MONITORING DISTANCE EFFECT WITH WAVE FIELD SYNTHESIS

Noguès M., Corteel E., Warusfel O.

IRCAM, 1 place Igor Stravinsky, 75004 PARIS
etienne.corteel@ircam.fr

ABSTRACT

Wave Field Synthesis (WFS) [1] rendering allows the reproduction of virtual point sources. Depending on source positioning, the wave front synthesized in the listening area exhibits a given curvature that is responsible for a spatial perspective sensation. It is then possible to monitor the distance of a source with a “holophonic distance” parameter concurrently with conventional distance cues based on the control of direct/reverberation ratio. Presentation of this holophonic distance is made and then discussed in the context of authoring sound scenes in WFS installations. Three goals to this work:

- Introducing WFS to sound engineers in an active listening test where they manipulate different parameters for the construction of a sound scene.
- Assessing the perceptual relevance of the *holophonic distance* modifications.
- Studying the possible link between the *holophonic distance* parameter and conventional *subjective distance parameters* traditionally used by sound engineers.

1. INTRODUCTION

A listening and monitoring test has been organised at IRCAM to evaluate the distance effect in the context of WFS installations. As a holophonic sound reproduction technique, WFS gives an independent access to two characteristics related to the distance of virtual sources:

- the acoustic field associated to the direct sound of the virtual source
- the energy distribution of the room response linked to the source

Sound engineers have long experience on monitoring the latter in order to create an apparent source distance. A common situation is to adjust the direct sound fader while keeping the auxiliary output of the artificial reverberator constant. Psychoacoustic studies have been dedicated to the analysis of the influence of the time and spatial distribution of the early reflections on the perceptual distance effect. These studies lead to the proposition of perceptually relevant rules in order to govern the apparent distance in audio spatialization processors or software [2][3].

WFS rendering allows the reproduction of the acoustical properties of a sound field in an extended listening area, especially in terms of wave fronts curvature associated to the position of the virtual sound source. It is known that the curvature of wave fronts doesn’t provide, by itself, any reliable cue for the auditory perception of distance [4]. However, the immediate consequence of the curvature is the sensation of the source direction together with the creation of an attenuation law in the reproduction area according to the position of the virtual source. If the sound installation is large enough to allow wandering around the reproduction area, the listeners will experience dynamic impressions of distance, linked to the variations of the direction and level of the sources according to their path.

The authoring of the sound scene may then be thought as creating a spatial perspective. A listener may want to change his or her standpoint in order to get close to one of the sound events. On the contrary, reproducing the room effect from a series of virtual plane waves surrounding the listening area creates a sensation of a nearly constant sound field in the whole reproduction area.

These examples show that it is possible to control the apparent distance of a source by both parameters: the wave field structure that governs the virtual position of the direct sound event and the energy distribution of the room effect. In a real situation, both parameters are obviously linked together according to the acoustic properties of the room. For synthesized sound fields it is possible to control them independently. For example, all sources, from a close monopole to a plane wave (monopole at infinite distance) may be reproduced by the WFS system with a normalized energy using a multichannel equalization algorithm [5].

The aim of the present study is to evaluate the respective role of the two parameter categories available for monitoring the distance effects in a WFS installation. We first introduce the following terminology in order to avoid confusion between these two categories:

- the *holophonic distance* will refer to the virtual source position created by the Wave Field Synthesis. Its perceptual efficiency may depend on the accuracy of the rendering system (size and characteristics of the loudspeaker arrays, spatial aliasing frequency, etc…)
- the *source presence* will refer to the subjective distance impression linked to the time and spatial distribution of the room response. The formulation and name of this perceptual attribute is derived from MPEG-4 standard [2].

The *presence* parameter is linked to an objective criterion Es calculated as follows:

$$\mathrm{Es} = \int_0^{40} h^2(t)dt + 0.18\int_{40}^{80} h^2(t)dt,$$

where $h(t)$ is the impulse response of the room. Refinements of this attribute can be found in [3].

2. TEST PREPARATION

2.1. Context of the experience: realistic perspective

The general framework of the test consists in the authoring of a "realistic perspective", i.e. a perspective that could be experienced in real life situation. The musical example is a small ensemble of instruments with guitar trio and a lead female singer. Each source was recorded with close microphone, allowing complete freedom in the spatial organization of the sound scene.

2.2. Rendering setup

The rendering setup used was composed of six MAP (Multi Actuator Panels) loudspeakers, each made of 8 transducers (cf Figure 1). The frontal array is made of 4 adjacent MAPs that forms a line of 5.3 m and allows the reproduction of sound event with point sources as well as plane waves. Seven room effect channels are synthesized using the IRCAM Spatialization software *Spat~*. These room effect channels are rendered by 3 plane waves in front, 2 side loudspeakers and 2 virtual loudspeakers synthesized by the rear panels.

For each source, a room effect can be synthesized according to the MPEG 4 perceptual parameters [2]. An interface was designed under MAX/MSP in order to monitor the sound scenes from dry recordings on a WFS rendering system.

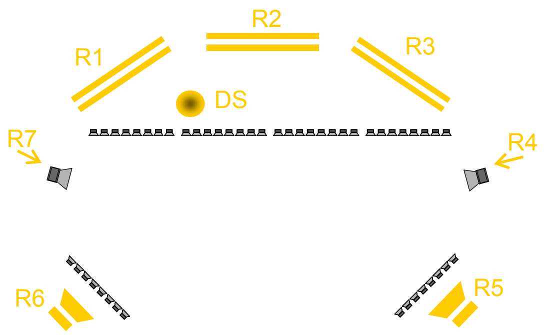

Figure 1 : *WFS rendering setup with 4 MAPs at the front, 2 loudspeakers on the sides, 2 MAPs at the back. DS represents the virtual sound source location, R1-R7 refer to the reproduction of the 7 room effect channels*

2.3. Test procedure

As an introducing example, a single source was reproduced, centred toward the frontal array, and the sound engineers could manipulate the *holophonic distance* in order to experience the changes in terms of localization variations when moving in the listening area (cf. Figure 2). A small portable MIDI fader is given to the sound engineers such that they can manipulate the target parameter while wandering in the listening area. 14 sound engineers participated to the different test experiments.

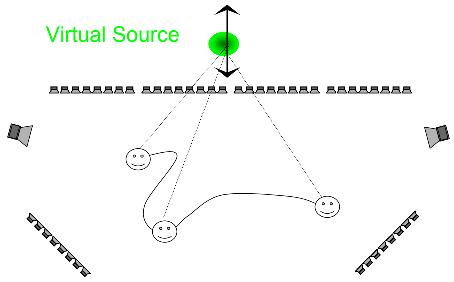

Figure 2: Training experiment: monitoring the holophonic distance of the virtual source, while exploring the listening area.

3. EXPERIENCE 1

3.1. Description

The goal of this experience is to check the perception of the *holophonic distance* in terms of perspective coherence. Two guitars are reproduced with a given holophonic distance. The third one is under control of the sound engineer who is asked to locate it between the two others at the same depth and using the holophonic distance parameter (cf. Figure 3). The sound engineer is invited to control the perspective walking around the listening area. Nine different test situations (distance, aperture and direction of the guitars pair) are proposed in random order to the sound engineer (cf. Table 1).

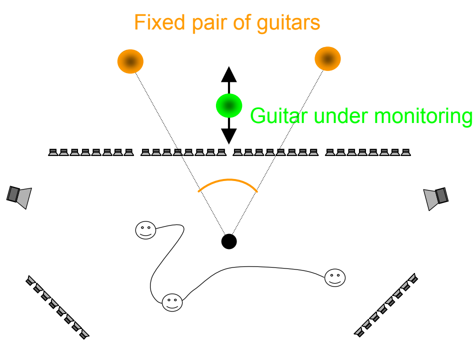

Figure 3: *Experience 1 : the sound engineer is asked to put the three guitars at the same depth, by monitoring the holophonic distance of the central guitar.*

3.2. Results

Figure 4 shows for each configuration the mean values and quartiles repartition of the *holophonic distance* adjusted by subjects. These results show a strong correlation between holophonic distance tuned for the central guitar and the imposed holophonic distance of the side guitars (ANOVA Analysis returns p=3.79e^{-10} for centred / 30° situations; p=3.37e^{-8} for off centred / 30° situations and p=0.0003 for centred /60° situations). No significant correlation could be found with the other configuration type (direction, aperture).

	Pair Distance (m)	*Pair aperture*	*Pair direction*
Situation 1	1,5	30°	Centred
Situation 2	1,5	60°	Centred
Situation 3	1,5	30°	off centred
Situation 4	4	30°	Centred
Situation 5	4	60°	Centred
Situation 6	4	30°	off centred
Situation 7	9	30°	Centred
Situation 8	9	60°	Centred
Situation 9	9	30°	off centred

Table 1: Configurations for experience 1

The range of the distance chosen by subjects is smaller than the range of distance imposed for the fixed guitars. This may come from some limitations of the rendering system. For distances shorter than 1,5 m, little spectral variations can be noticed and might have been rejected by the subjects. On the contrary, distances larger than 9 m are submitted to diffraction. This may have influenced subjects in their appreciation of the scene organization, especially if they explored very off-centred listening positions.

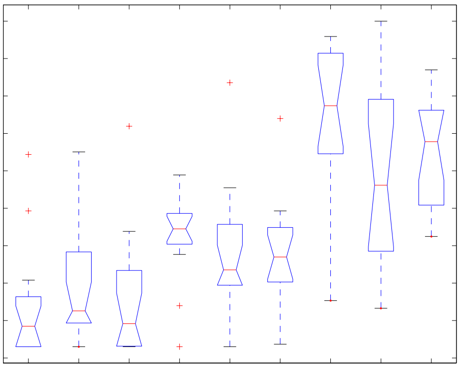

Figure 4: Experience 1. Mean values and quartiles of the holophonic distance adjusted by subjects.

3.3. Conclusion on experience 1

Results of the first experience show that sound engineers had a reliable perception of the holophonic distance. They were able to use this parameter for monitoring the distance when authoring the sound scene in the WFS installations.

4. EXPERIENCE 2

4.1. Description

In natural situation, *holophonic distance* and *source presence* are linked together, according to the room behavior. The goal of this second experience is to check whether an authoring tool dedicated to WFS installation should link these two parameters in order to form a single one or if it should keep them independent.

The guitar trio is now reproduced with a specified holophonic distance and source presence, forming a reference plane in the virtual sound scene. The sound engineer is proposed to monitor the lead female voice in order to obtain a coherent sound scene (cf.Figure 5).

The global organization consists in:

- imposing different values of the *source presence* for the voice and allow the adjustment of its *holophonic distance*.
- imposing different *holophonic distances* for the voice and allow the adjustment of the *source presence*.
- presenting these situations in random order, without informing the subject which of the two types of distance parameter is under control (the subject manipulates a small portable midi fader which can be automatically remapped to the chosen parameter).

Twelve configurations are tested (cf. Table 2 and Table 3). Two reference planes are proposed for the guitar trio defined by the couple *holophonic distance* / *source presence*.

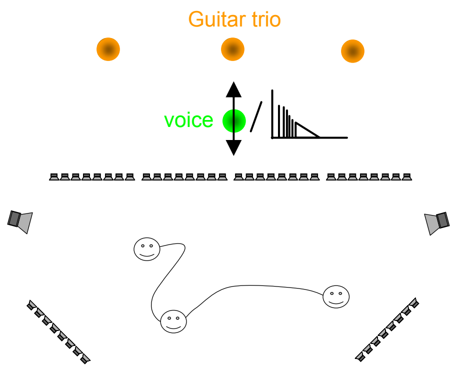

Figure 5: *Experience 2: the sound engineer has to produce a coherent sound scene by monitoring a fader that is randomly mapped on the holophonic distance or the source presence of the lead voice.*

For best understanding, the results are analysed by group of configurations according to the parameter under control. It is reminded that all configurations were presented in a single test and with random order. The first part corresponds to the monitoring of the *source presence* and the second one corresponds to the monitoring of the *holophonic distance*.

	Voice distance (m)	*guitars distance (m)*	*guitars presence (spec.units)*
Situation 1	4	4	78 (-14 dB)
Situation 2	8	4	78 (-14 dB)
Situation 3	1,5	4	78 (-14 dB)
Situation 4	7	8	63 (-19 dB)
Situation 5	10	8	63 (-19 dB)
Situation 6	3	8	63 (-19 dB)

Table 2: *Experience 2: configurations with imposed holophonic distance. The parameter under monitoring is the source presence of the voice.*

	Voice presence	*guitars distance (m)*	*guitars presence (spec.units)*
Situation 1	78 (-14.0 dB)	4	78 (-14 dB)
Situation 2	86 (-11.3 dB)	4	78 (-14 dB)
Situation 3	69 (-17.0 dB)	4	78 (-14 dB)
Situation 4	63 (-19.0 dB)	8	63 (-19 dB)
Situation 5	71 (-16.3 dB)	8	63 (-19 dB)
Situation 6	54 (-22.0 dB)	8	63 (-19 dB)

Table 3: Experience 2: configurations with imposed source presence. The parameter under monitoring is the holophonic distance of the voice

4.2. Source presence monitoring

Figure 6 shows the mean values and quartiles repartition of voice *source presence* adjusted by subjects. The results show no correlation between the proposed holophonic distance of the voice and the source presence adjusted by subjects (ANOVA returns p=0.9734). On the contrary, the chosen values of the presence parameter for the voice are strongly correlated to the imposed presence values of the guitars (ANOVA returns p≈0).

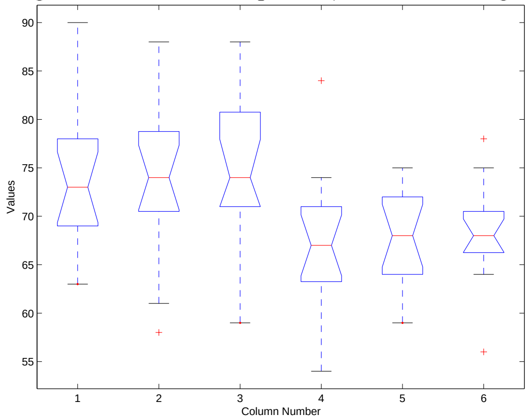

Figure 6: *Experience 2: source presence monitoring*

4.3. Holophonic distance monitoring

Figure 7 shows the mean values and quartiles repartition of voice holophonic distance (meters) adjusted by subjects. The results show no significant correlation between the proposed *source presence* of the voice and the holophonic distance adjusted by subjects (ANOVA returns p=0.6405). On the contrary, they show a strong correlation between the holophonic distance of the guitar trio and the holophonic distance adjusted by the subjects for the voice (ANOVA returns p=0.0043).

4.4. Conclusion

This second experiment shows that the two kinds of distance parameter (holophonic distance and source presence) are perceived as two independent parameters. When asked to mind the coherence of the scene, sound engineers try to match the value of the controlled parameter with the corresponding values of the other instruments, with no inference of the other parameter.

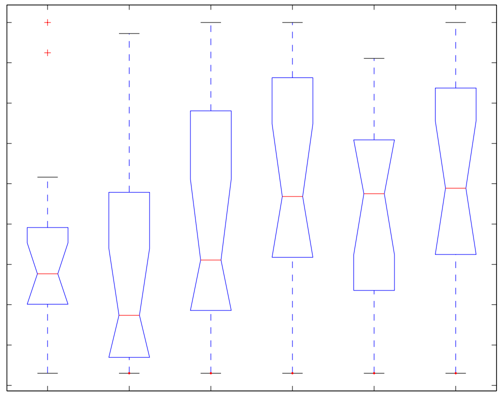

Figure 7: Results of experience 2 (second part)

5. CONCLUSION

A panel of sound engineers has been introduced to authoring sound scenes for WFS installations. In this context, it has been shown that distance monitoring partly relies on the control of a specific *holophonic distance*. The perceptual relevance of the holophonic distance has been verified and was shown to be independent from the *source presence* attribute linked to the energy distribution of the room effect.

This research has been conducted in the context of the IST CARROUSO project [6].

REFERENCES

[1] A. J. Berkhout, "A Holographic Approach to Acoustic Control", *Journal of the Audio Engineering Society*, Vol. 36, No. 12, December 1988, pp 977-995.

[2] ISO/IEC 14496-1, Amendment 1, "*Coding of Audiovisual Objects - Part 1: Systems*", 2000 (Output document w3850 of the 55th MPEG meeting, Pisa, Italy, Dec. 2000). § 9.4.2.78.2

[3] R. Pellegrini, "*A virtual reference listening room as an application of auditory virtual environments",* Ph. D. Thesis to be published VDI Verlag, Germany, 2000.

[4] J. Blauert , « *Spatial hearing : the psychophysics of human sound localization* »,revised ed. Cambridge, MA :MIT Press, 1997.

[5] E. Corteel, U. Horbach, R. S. Pellegrini, « Multichannel inverse filtering of Multiexciter Distributed Mode Loudspeaker for WFS », *AES convention paper*, Munich Mai 2002.

[6] S. Brix, Th. Sporer, J. Plogsties, "CARROUSO – A European Approach to 3D-Audio", presented at the *110th Convention of the Audio Engineering Society*, Amsterdam, May 2001.