Spatial perception in WFS rendered sound fields

Günther Theile

Institut für Rundfunktechnik, D-80939 München, Email: theile@irt.de

Introduction

Modern convolution technologies offer possibilities to overcome principle shortcomings of loudspeaker stereophony by exploiting the Wave Field Synthesis (WFS) concept for rendering virtual spatial characteristics of sound events. Based on the Huygens principle loudspeaker arrays are reproducing a synthetic sound field around the listener, whereby the dry audio signal is combined with measured or modelled information about the room and the source's position to enable the accurate reproduction of the source within its acoustical environment. Recent developments have shown that a number of applications could be possible in the near future. However, they have also shown that basic and practical constraints of WFS systems limit the rendering accurateness and the perceived spatial audio quality to a certain degree, dependent on characteristic features and technical parameters of the sound field synthesis.

Wave field synthesis concept

Principles of this technology have been presented in a number of papers, see e.g. [1] – [3]. The signal of the source ("Gestalt") and its spatial properties are transmitted separately (e.g. using the MPEG-4 format [4]). For reproduction, the dry source signal is convolved with the measured or modelled set of impulse responses (containing the spatial information), and emitted by a loudspeaker array. In contrast to stereophony WFS could be able to

- produce virtual stable sources localized at the same position throughout the entire listening area,
- produce virtual sources in front of the loudspeaker array ("focused sources")
- produce plane waves that are localized in the same direction throughout the entire listening area,
- enhance the sense of depth, spatial impression and envelopment through a realistic reproduction of the original room response

Figure 1 illustrates the basic WFS concept by means of music recording. Step one is always the room response measurement in the recording room, done e.g. with a stepwise rotating microphone. This measured spatial information is stored in the WFS processor.

For recording of orchestra and soloist closely spaced spot microphones are used. The stereophonic orchestra mix should be composed in a way that it contains as little room information (reverb, reflections, etc.) as possible; but it should contain the adequate spatial distribution of source elements [5]. This three channel stereophonic mix signals and the soloist signal are being convolved with the appropriate spatial impulse responses. As a result, the rendered WFS sound field represents stable virtual sources located in the concert hall. Listeners within the listening area perceive a three-channel stereophonic image of the orchestra and a point source image of the soloist, whereby the reproduced characteristics of the concert hall can give a new sense of realism.

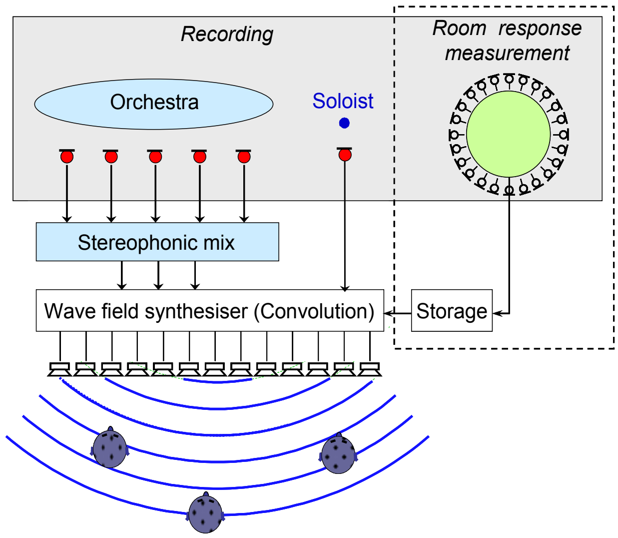

Figure 1: WFS: Separate handling of sources and measured (or modelled) spatial information

Practical constraints

Not surprisingly, in practice it is not possible to match all theoretical requirements for a perfect result. The rendered WFS sound field differs from the desired sound field to some degree for a number of reasons (for details see [6]):

Discreteness of the array (spatial aliasing)

This effect produces spatial and spectral errors of the synthesized sound field due to the discretisation of a continuous secondary source distribution. Above the spatial aliasing frequency f_{alias} the time difference between two successive loudspeaker signals interferes at the listener's position, depending on the spatial sampling interval, i.e. the loudspeaker / microphone inter-spacing.

Reflections of the reproduction room (spatial interference)

A WFS array can not render the desired sound field perfectly if reflections of the reproduction room produce interference in spatial perception. In particular, perception of distance, depth and spatial impression are affected, because fragile distance cues of synthesised sources can be dominated by the stronger distance cues generated by the array speakers. They interfere with the desired reflection pattern of the synthesised source. Special room compensation algorithms being under investigation [7], [8] will perhaps be able to minimize this effect.

Restriction to the horizontal plane

Theory does not restrict WFS to the horizontal plane. However, the reduction of the array dimension to the horizontal plane is the practical approach, having a number of consequences. First, virtual sources can be synthesized only within the horizontal plane. This includes virtual reflections affecting the completeness of a natural reflection pattern and thus possibly resulting in impairments of perception of distance, depth, spatial impression and envelopment. Another aspect is related to the measuring techniques used for capturing the room response. In practice there is some mismatch with respect to elevated reflections, because the measured room response includes elevated reflections although they are reproduced only in the horizontal plane. The effects of these types of inaccurateness on spatial perception parameters are not well-known yet. - Furthermore, horizontal arrays do not generate real spherical waves, but cylindrical waves. In the case of imaging a plane wave for example there results an error with respect to the roll-off level (3dB/doubling of distance), in comparison with the ideal plane wave (no roll-off) [6], [9].

Limitation of array dimensions (diffraction)

In practical applications the loudspeaker array will have a finite length. Due to a finite array so-called diffraction waves originate from the edges of the loudspeaker array [6], [9]. These contributions appear as after-echoes (and pre-echoes respectively for focussed sources), and – depending on their level and time-offset at the receiver's location – may give rise to colouration. - However, methods to reduce these truncation effects are known, e.g. by applying a tapering window to the array signals. This means that a decreasing weight is given to the loudspeakers near the edges of the array. In this way the amount of diffraction effects can substantially be reduced at the cost of a limitation of the listening area [9].

Effects on perception

Current and future studies	Discreteness of array (spatial aliasing)	Reflections of the reproduction room (spatial interference)	Restriction to the horizontal plane	Limitation of array dimensions (diffraction)
Direction				
Distance				
Spatial depth				
Spatial perspective				
Spatial impression				
Reverberance				
Envelopment				
Sound colour				

Table 1: Perceptual attributes vs. physical constraints

Several effects of the practical constraints on specific perceptual attributes are not known yet in detail (Table 1). However, this knowledge is important for further developments of spatial audio systems in view of future applications. Current studies are concentrating on both the physical improvement of rendering techniques as well as the subjective evaluation of principle characteristics of WFS systems in comparison with stereophonic or binaural systems. It seems attractive to exploit advantages or avoid disadvantages of specific systems by combining parts of them, dependent on the intended field of application.

Summary

Practical constraints of WFS systems can limit the rendering accurateness and the perceived spatial audio quality to a certain degree, dependent on characteristic features and technical parameters of the sound field synthesis, such as the discreteness of the array (spatial aliasing effects), impact of reflections of the reproduction room (spatial interference), restriction to the horizontal plane, limitation of array dimensions (diffraction effects), etc. Psychoacoustic studies are necessary to evaluate the resulting impacts on attributes of spatial perception not only with respect to the development of WFS systems for different applications but also in view of scientific knowledge. Particular attention should be turned to the perception of direction, distance, spatial depth, spatial perspective, spatial impression, reverberance, and envelopment, as well as sound colour.

Acknowledgment

I would like to thank my colleague Helmut Wittek for his contributions to this paper.

References

[1] Boone, M.; de Bruijn, W.; Piccolo, T.: „Sound Recording Techniques for Wavefield Synthesis and Other Multichannel Systems". 104th AES Convention, 1998, Preprint 4690

[2] Boone, M.; de Bruijn, W.: „On the Applicability of Distributed Mode Loudspeaker Panels for Wave Field Synthesis Based on Sound Reproduction". 108th AES Convention, 2000, Preprint 5165

[3] Hulsebos, E., de Vries D., Bourdillat, E.: "Improved Microphone Array for Auralization of Sound Fields by Wave Field Synthesis", J. Audio Eng. Soc., Vol.50, No.10, pp.779-790, 2002

[4] Brix, S.; et al: „CARROUSO – An European Approach to 3D-Audio".110th AES Conv., 2001, Preprint 5800

[5] Theile, G.; Wittek, H.; Reisinger, M.: „Potential Wavefield Synthesis Applications in the Multichannel Stereophonic World". AES 24th International Conference, 2003. Proceedings pp. 43 – 57

[6] Wittek, H., 2003: "Perception of Spatially synthesized sound fields – Literature review about WFS", available online at: www.hauptmikrofon.de/wfs.htm

[7] Spors, S., Kuntz, A., Rabenstein, R., 2003: „An Approach to Listening Room Compensation with Wavefield Synthesis", Proc. AES 24th Int. Conference on Multichannel Audio, Banff, Canada

[8] Corteel E., Nicol R., "Listening room compensation for Wave Field Synthesis. What can be done?" AES 23rd Int. Conf., Copenhagen, Denmark, May 2003.

[9] Sonke, J.-J.: "Variable Acoustics by wave field synthesis", Thela Thesis, Amsterdam, Netherlands, 2000, ISBN 90-9014138-3