Proceedings of the Institute of Acoustics

WAVE FIELD SYNTHESIS AND ANALYSIS: THE STATE OF THE ART

Diemer de Vries and Marinus M. Boone

Laboratory of Acoustical Imaging and Sound Control
Delft University of Technology
P.O.B.5046, 2600 GA Delft, The Netherlands
diemer@akst.tn.tudelft.nl

ABSTRACT

The concept of wave field synthesis (WFS) was introduced by Berkhout in 1988 [1]. It enables the generation of sound fields with natural temporal and spatial properties within a volume or area bounded by arrays of loudspeakers. Applications are found in real time performances as well as in reproduction of multi-track recordings. A logic next step was the formulation of a new wave field analysis (WFA) concept by Berkhout et al. in 1997 [2], where sound fields in enclosures are recorded with arrays of microphones and analyzed with postprocessing techniques commonly used in acoustical imaging. This way, both the temporal and spatial properties of the sound field can be investigated and understood. WFS and WFA meet in auralization applications: sound fields measured (or modeled) along arrays of microphone positions can be generated by arrays of loudspeakers for perceptual evaluation.

1 INTRODUCTION

In traditional sound enhancement and sound reproduction practice, individual (groups of) loudspeakers are used to generate a replica of the recorded sound. Using high-quality systems in the appropriate manner, the temporal properties of this replica may be correct. Spatially, however, its properties are fully determined by the interfering directivity patterns of the loudspeakers. In figure 1, this is illustrated for a monochromatic source signal

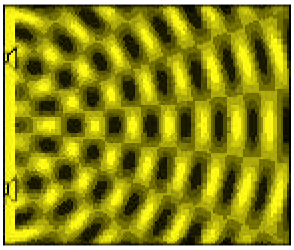

Figure 1. *Monochromatic source signal reproduced by two loudspeakers.*

reproduced by two omnidirectional loudspeakers. Only in a limited listener area, often only at one 'sweet spot', the perceived spatial image is correct. If, for instance, the two loudspeakers enhance the signal of a primary point source positioned behind them, most listeners receive the loudspeaker signal earlier than the primary signal which leads to mislocalization: the first arriving wave front determines the direction from which the sound is heard.

Traditionally, when measuring the sound field in an auditorium in order to analyze its properties in a physical or perceptual way, impulse responses are recorded with microphones placed at a limited number of rather arbitrarily chosen 'representative' positions. Each measured response, and the acoustic parameters (predictors of perception) derived from it, are supposed to be valid for some area around the corresponding microphone position. In practice it appears that (1) it is often difficult to physically interpret a single impulse response, and that (2) significant differences between neighboring microphone positions are observed which are also difficult to explain. This is not surprising, since this way only local temporal information about the sound field is obtained. Information on spatial properties related to interference and diffraction cannot be taken into account.

In order to overcome the above mentioned drawbacks in reproduction and analysis of sound fields, array technology should be applied based on the concepts of wave field synthesis (WFS) and wave field analysis (WFA) which will be explained in the following sections.

2 THE WAVE FIELD SYNTHESIS CONCEPT

In this section the underlying theory of WFS is summarized. For an extensive treatment of the theory and the mathematical formulation, the reader is referred to [3].

According to the Huygens principle, the propagation of a wave through a medium can be qualitatively described by adding the contributions of all secondary sources positioned along a wave front. This implies that, when the wave field on the boundary surface *S* of a closed, source-free volume *V* is known in terms of pressure and normal particle velocity, the sound pressure at any point within that volume can be determined. It appears that surface *S* can be interpreted as covered with a continuous distribution of secondary monopole sources driven by the local normal velocity generated by the primary sources, plus a distribution of dipole sources driven by the local primary pressure. Together, all these virtual secondary sources can be seen as generating a field within *V* which is identical to the field that the primary sources would have generated there - or, when *S* is a virtual boundary, really generates there. (This can be mathematically described by the Kirchhoff representation theorem.) When the closed boundary is replaced by a (pseudo)infinite plane with the primary sources at only one adjacent half-space, the sound pressure at any point in the other half-space can be determined from whether the normal velocity distribution in that plane, or from the pressure distribution. (This can be mathematically described by the Rayleigh representation theorems.) In the first case, the plane can be interpreted as being covered with secondary monopole sources driven by the local primary normal velocity distribution, in the second case as covered with secondary dipole sources driven by the local primary pressure distribution.

As described by Berkhout et al. [3], this concept is a strong base for application in audio and acoustics technology. When a plane is covered with loudspeakers having monopole source (i.e., omnidirectional) characteristics, being driven with signals corresponding to the normal velocity distribution in that plane generated by a real or virtual source or sources in one half-space, a replica (in case of real sources) or a simulation (in case of virtual sources) is generated in the entire other half-space. The same holds for a plane covered with dipole-type ('figure-of-eight') loudspeakers driven with signals corresponding to the primary pressure distribution. It has been shown [3] that with the use, instead of planar arrays as prescribed by the theory above, of *linear* arrays of loudspeakers - which are much more appropriate for practical use from a visual point of view and with respect to the hardware and computational power required - good results can be obtained in a horizontal plane, e.g., the earplane of an audience. Also, it has been shown [4] that the concept holds for any type of loudspeaker, by adapting the driving signal operator to its properties.

This means that, instead of the spatially erroneous wave field shown in figure 1, now a spatially (and temporally) correct wave field of a point source positioned behind the array is synthesized by all loudspeakers together, as illustrated in figure 2.

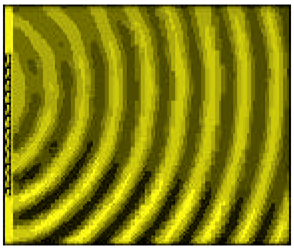

Figure 2. *Monochromatic signal of a point source placed behind an array of loudspeakers, reproduced by that array. Note the correct spatial and temporal properties.*

In the next section, applications of wave field synthesis will be discussed.

3 APPLICATIONS OF WAVE FIELD SYNTHESIS

When generating or reproducing sound with traditional loudspeaker configurations, the sound field is spatially correct at only one or a few 'sweet spots': local solutions are obtained. Systems based on WFS, however, generate a wave field with natural properties in time and space in an extensive audience area, yielding a volume solution. This also holds for moving sources.

3.1 Sound Enhancement in Theatres

When the instantaneous positions of actors and singers in theatre performances are known, their direct sound fields at the position of a loudspeaker array addressing the audience area can be calculated - usually, omnidirectivity of the primary sources is assumend as an approximation - and used to drive the individual elements of the array. This way, replicas of the primary field are generated which can be amplified with full preservation of the original properties in time and space. In order to know the instantaneous source positions, directional microphones, focussed microphone arrays or source tracking systems (in case of 'on body'-miking) have to be used.

3.2 Variable Acoustics in Multipurpose Auditoria

In multipurpose auditoria, the acoustical conditions should be optimally adapted to the type of performance. Often, such venues have acoustic properties that are acceptable for speech productions (lectures, drama), but insufficient for musical performances due to a lack of reflections and reverberation. A discrete reflection at the boundary of a hall can be seen as the direct signal of a mirror image source positioned behind that boundary. An early (i.e., up to about 100 ms after direct sound arrival) reflection pattern being optimal for a certain type of performance can be realized by generating the field of optimally distributed mirror image sources by WFS, using arrays of loudspeakers around the audience area. Perceptual experiments [5] have shown that an optimal reverberant field, i.e., the highly dense reflection pattern following the direct sound and the early reflections, can be realized by synthesizing a distribution of about 8 uncorrelated plane waves. The temporal properties of the reverberant waves to be generated strongly depend on the type of performance.

For a certain structure of an acoustic field generated with WFS as described above, a specific filter is required for each loudspeaker of the synthesizing array. During a performance, the direct sound of the sources has to be convolved in real-time with these filters. Applying dedicated block-partitioned FFT algorithms [6], this can nowadays be realized with normal PCs containing an advanced soundcard.

3.3. Reproduction of Multi-channel Recordings

When multi-track recordings of sources (voices, instruments) at known positions are available, they can be replayed with preservation of the original spatial properties by means of WFS: within the total area within a (e.g., rectangular) configuration of loudspeaker arrays, a wave field is created that matches the wave field on the recording location [7]. This is especially the case when the acoustics of the recording venue (reflections,

Figure 3. *With WFS, not only the sound of sources outside the listeners area (a), but also within that area (b) can be reproduced.*

reverberation) has been recorded separately and is regenerated in the same way as described in subsection 3.2.

It is also possible to position the sources on other positions than their original ones ("build your favorite orchestra arrangement"), even within the listeners area between the loudspeaker arrays using the focussing principle as illustrated in figure 3. Also, the sound of moving sources can be reproduced in a natural way.

WFS-based reproduction can well be combined with visual information with applications in cinema's, home theatres and virtual reality theatres. A real-time application in this context is the WFS-based reproduction of speech in teleconferencing systems, in order to improve the correspondence of visual and acoustical perception.

Much progress in this field has been made in the European IST 5th framework project “CARROUSO’’ (2001-2003), where 10 partners successfully cooperated on the realization of real-time recording, transmission and WFS rendering of multi-channel sound [8]. The result of the project was a prototype demonstrator of a system capable of doing all this. During the project, several demonstrations of the WFS reproduction facility have been given at AES and VDI conferences. Besides arrays of traditional loudspeakers, also multi-actuator distributed mode loudspeaker (DML) flat panels were used here, simultaneously as speaker arrays and projection screens. As a dissemination of CARROUSO, a first WFS cinema surround system for the reproduction of common 5.1 material and specially produced multi-channel trailers is in daily use [9].

3.4. Auralization

As discussed in the next section, wave fields in halls should be recorded along arrays of microphone positions for fruitful analysis and understanding. When these recordings are regenerated by arrays of loudspeakers, the acoustics of the hall is auralized, i.e., made audible in spaces other than where it has been recorded. Convolution of the wave field with ‘dry’ music enables to accurately reconstruct a concert in its original acoustic environment [10], or to simulate a concert in a hall where it never has been performed. In commonly used auralization techniques, the (often binaural) sound field at one listener position is reproduced, to be perceived with headphones or a pair of near-field loudspeakers. This way, the spatial properties of the sound field are quite difficult to assess, since often the full acoustic image is localized within the listener's head. In the WFS approach of auralization, however, listeners can 'walk around' within the wave field generated by the arrays.

Note that this auralization can also be done for wave fields, simulated by some modeling algorithm along an array of microphone positions. This way, the differences between measured and modeled data can be perceptually evaluated. When modeling is done properly, architects and consultants can acquire a realistic impression of the acoustic properties of a space under design, and of the acoustic effects of possible modifications.

A special application of auralization is the simulation of acoustic events in flight simulators. Because of the complexity of modern airplanes, high quality simulation of the real environment, including the sound field, is of high importance.

4. WAVE FIELD ANALYSIS

Berkhout et al. [2] have shown that multi-channel recording or calculation of impulse responses in an enclosed space along an array of microphone positions introduces a new concept of wave field analysis (WFA), yielding much insight in the temporal and spatial structure of the wavefield. An example is given in figure 4, showing the impulse responses measured in a 50-seat, rectangular lecture room at Delft University, along a linear array of microphone positions with 0.05m interspacing, over the full width of the hall at a distance of 5m from an omnidirectional source placed on the usual lecturer position centrally at the front side of the room. The vertical axis represents the traveltime coordinate *t* which equals zero when the pulse leaves the source. The horizontal axis gives the lateral microphone position *x*, the so-called offset, *re* the center of the array which in this case coincides with the center of the room.

Already without any further processing the dataset clearly shows the wave character of the sound field. In spite of the complex structure of the field due to interference and diffraction, with array-based WFA many reflection and diffraction events can be easily discriminated since - other than when displaying individual, isolated impulse responses - now the spatial correlation between neighboring responses is revealed. By taking the hall geometry into account, the origins of many reflected or diffracted wave fronts can be identified, as indicated in figure 4. By applying a spatial Fourier transform or a Radon transform to the dataset, the wave field is decomposed into plane wave components [2] which enables further study to properties as diffusivity, lateral energy content, etc as a function of time.

Since the sound pressure was measured with an omnidirectional microphone, the dataset of figure 4 allows no discrimination in the elevation plane around the microphone array: wave fronts from front, back, above and below are all projected in the same offset-traveltime plane.

Title:
intro.eps
Creator:
MATLAB, The Mathworks, Inc.
Preview:
This EPS picture was not saved
with a preview included in it.
Comment:
This EPS picture will print to a
PostScript printer, but not to
other types of printers.

Figure 4. *Impulse responses measured along a microphone array in a rectangular lecture room.*

When, however, not only the sound pressure is recorded, but also the three components of the particle velocity vector on each array position, a directional microphone can be *simulated* by post-processing [11]. This simulated microphone can be rotated around the microphone array under each azimuthal angle with the array between -90 and +90 degrees, such that wave components incident on the array under different elevation angles can now be discriminated.

Using wave field extrapolation techniques as developed for seismic extrapolation purposes [12], from a multi-channel recording along one array the responses at the position of any other array in the hall can be estimated. In principle, by combining the techniques of wave field extrapolation and directional microphone synthesis described above, one array measurement gives ample 3D information on the acoustics of the hall.

Taking aspects of spatial resolution into account it becomes clear that, in addition to recording or calculation along an array over the width of the hall, data acquisition along an array with front-to-back orientation improves this resolution. A much faster procedure is the measurement of impulse responses along a circular array of microphone positions by slowly (e.g., 720 seconds per cycle) rotating a microphone on a rod mounted on a turntable. This way, it is possible to making response recordings for several source positions within the usually restricted time available. By decomposing the recorded data into cylindrical harmonics and transforming these into plane waves, the same features for analysis and extrapolation are basically available as there are for linear arrays [13].

Circular array impulse response measurements have been made, within the framework of the CARROUSO project, in several concert halls such as the Amsterdam Concertgebouw and the Berlin Philharmonie. Besides for room acoustic research purposes they were also used to auralize and compare the different environments during the conference demonstrations mentioned above.

One of the results of the physical analysis of multi-trace impulse responses is that the traditional roomacoustical parameters such as Clarity Index, Lateral Fraction, etc show such significant fluctuations on a small spatial scale - e.g., within the 12 microphone array positions at one and the same seat - that their relevance to predict perceptual cues is quite doubtful [14]. Apparently, these parameters are sensitive to local interference where the human perception is not. New, spatially more stable versions of the parameters should be defined based on perceptual evaluation of the measured datasets.

Recently, planar microphone array impulse response measurements have been done in a vertical plane on a certain distance of a wall configuration. Extrapolation to the wall enables to image the configuration. Acoustical modification (e.g., removal, or changing the shape or the absorption coefficient) of the objects on the wall and inverse extrapolation to the array position enables to quantify the influence of these modifications in physical and/or perceptual terms [15]. This may answer the question in howfar it makes sense to include details in room acoustic modeling.

5. CONCLUSIONS

- The wave field synthesis (WFS) concept, according to which sound waves are generated using arrays of loudspeakers, enables the realization of sound fields with natural or prespecified virtual properties in time *and* space, not on a few 'sweet spots', but in an entire audience area, for a variety of applications.
- Wave field analysis (WFA) based on the measurement of impulse responses along a closely spaced array of microphone positions reveals the spatial coherence of neighboring responses, leading to a far better insight in the complex wave fields in enclosed spaces than the analysis of individual impulse responses. Besides it enables spatial auralization of acoustic environments.

6. REFERENCES

[1] Berkhout, A.J., "A holographic approach to acoustic control", *J. Audio Eng. Soc., Vol. 36, 1988, pp 977-995.*

[2] Berkhout, A.J., de Vries, D., and Sonke, J.J., "Array technology for acoustic wave field analysis in enclosures", *J. Acoust. Soc. Am., Vol.102, 1997, pp 2757-2770.*

[3] Berkhout, A.J., de Vries, D., and Vogel, P., "Acoustic control by wave field synthesis", *J. Acoust. Soc.Am.*, *Vol. 93, 1993, pp 2764-2778;*
Start, E.W., *Direct Sound Enhancement by Wave Field Synthesis*, dissertation Delft Univ. of Technology, 1997.

[4] de Vries, D., "Sound reinforcement by wavefield synthesis: adaptation of the synthesis operator to the loudspeaker directivity characteristics", *J. Audio Eng. Soc., Vol. 44, 1996, pp 1120-1131.*

[5] Sonke, J.J., and de Vries, D., "Generation of diffuse reverberation by plane wave synthesis", *preprint 4455, 102nd Convention Audio Eng. Soc., Munich, 1997;*
Sonke, J.J., *Variable acoustics by wave field Synthesis,* dissertation Delft Univ. of Technology, 2000.

[6] Sonke, J.J., "Design of an adaptive filter for electro-acoustic reverberation", *M.Sc. thesis, Delft Univ. of Technology, 1995.*

[7] Boone, M.M, Verheijen, E.N.G., and van Tol, P.F., "Spatial sound field reproduction by wave field synthesis", *J. Audio Eng. Soc., Vol. 43, 1995, pp 1003 - 1012;*
Verheijen, E.N.G., *Sound Reproduction by Wave Field Synthesis*, dissertation Delft Univ. of Technology, 1998;
Boone, M.M., and Verheijen, E.N.G., "Sound reproduction applications with wave field synthesis", *preprint 4689, 104th Convention Audio Eng. Soc., Amsterdam, 1998.*

[8] Brix, S., Plogsties, J., and Sporer, T., "CARROUSO- A European approach to 3D-audio", *preprint 5314, 110th Convention Audio Eng. Soc., Amsterdam, 2001.*

[9] This system has been designed and built by CARROUSO-partner Fraunhofer-AEMT and was inaugurated February 19th, 2003 in the "Lindenlichtspiele"-cinema in Ilmenau, Germany.

[10] De Vries, D., Brix, S. and Hulsebos, E.M., "Measurements of church impulse responses using a circular microphone array for natural spatial reproduction of a choir concert recording", *proceedings of the 21st AES Conference on Architectural Acoustics and Sound Reinforcement, St.Petersburg, 2002.*

[11] Baan, J., and de Vries, D., "Array technology for discrimination between specular and non-specular reflections in enclosed spaces", *preprint 4713, 104th Convention Audio Eng. Soc., Amsterdam, 1998.*

[12] Berkhout, A.J., *Applied Seismic Wave Theory*, Elsevier, Amsterdam etc, 1987.

[13] Hulsebos, E.M., de Vries, D., and Bourdillat, E., "Improved microphone array configurations for auralization of sound fields by wave-field synthesis", *J. Audio Eng. Soc., Vol. 50, 2002, pp 779 – 790.*

[14] de Vries, D., Hulsebos, E.M., and Baan, J., "Spatial fluctuation in measures for spaciousness", *J. Acoust. Soc.Am.*, *Vol. 110, 2001, pp 947-954.*

[15] Kuster, M., de vries, D., and Hulsebos, E.M., "Acoustic reflection imaging in enclosed spaces", *paper in preparation to be submitted to J. Acoust. Soc.Am.*, *2003.*

Note:
This paper is a significantly updated version of a contribution that the authors earlier presented at the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. 17-20, 1999.