LISTENING ROOM COMPENSATION FOR WAVE FIELD SYNTHESIS

S. Spors, A. Kuntz and R. Rabenstein

Telecommunications Laboratory
University of Erlangen-Nuremberg
Cauerstrasse 7, 91058 Erlangen, Germany
E-mail: {spors, kuntz, rabe}@LNT.de

ABSTRACT

Common room compensation algorithms are capable of dereverberating the listening room at some discrete points only. Outside these equalization points the sound quality is often even worse compared to the unequalized case. As wave field synthesis in principle allows to control the wave field within the listening area it can also be used to compensate for the reflections caused by the listening room in the complete listening area. We present a novel approach to listening room compensation which is based upon the theory of wave field synthesis and that results in a large compensated area.

1. INTRODUCTION

Modern multimedia systems include multichannel sound reproduction systems that aim at reproducing the spatial properties of scenes as well. The sound reproduction systems in current use rely on free field propagation of the sound emitted by the loudspeakers. However, the typical listening room produces disturbing reflections. Since advanced multichannel reproduction systems, like wave field synthesis, allow to control the wave field within the listening area to some degree they can also be used to compensate for the disturbing effects of the listening room. This contribution describes a compensation approach developed within the EC project CARROUSO [1].

2. WAVE FIELD SYNTHESIS

The theory of wave field synthesis (WFS) has been initially developed at the Technical University of Delft over the past decade [2]. In contrast to other multi-channel approaches, it is based on fundamental acoustic principles. This section gives a short overview of the theory as well as on rendering methods and wave field analysis.

2.1. Theory

WFS is based on the Huygens' principle. Huygens stated that any point of a wave front of a propagating wave at any instant conforms to the envelope of spherical waves emanating from every point on the wavefront at the prior instant. This principle can be used to synthesize acoustic wavefronts of an arbitrary shape. Of course, it is not very practical to position the acoustic sources on the wavefronts for synthesis. By placing the loudspeakers on an arbitrary fixed curve and by weighting and delaying the driving signals, an acoustic wavefront can be synthesized with a loudspeaker array. Figure 1 illustrates this principle.
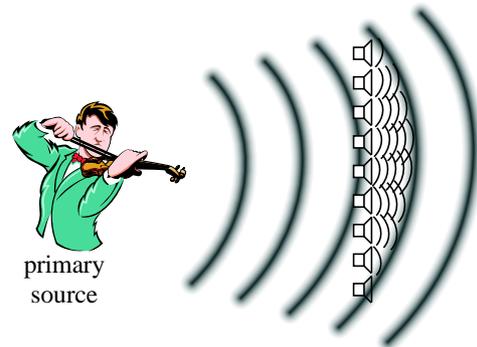

Figure 1: Basic principle of wave field synthesis

The mathematical foundation of this more illustrative description of WFS is given by the Kirchhoff-Helmholtz integral, which can be derived by using the wave equation and the Green's integral theorem [3]. The Kirchhoff-Helmholtz integral states that at any listening point within a source-free volume V the sound pressure can be calculated if both the sound pressure and its gradient are known on the surface S enclosing the volume. This principle can be used to synthesize a wave field within a volume V by setting the appropriate pressure distribution and its gradient on the surface. This fact is used for WFS based sound reproduction. However, two essential simplifications are necessary to arrive at a realizable system: Degeneration of the surface S to a line and spatial discretization. Performing these steps the so called Rayleigh integrals can be derived [2]. The Rayleigh I integral states that a pressure field may be synthesized by means of a monopole distribution on a plane. Using this result a WFS system can be realized by mounting closed loudspeakers in a linear fashion (linear loudspeaker arrays) surrounding the listening area leveled with the listeners ears. Figure 2 shows a typical setup. Up to now we assumed that no acoustic sources lie inside the volume V. The theory presented above can also be extended to the case that sources lie inside the volume V [2].

The fact that loudspeakers can only be mounted at discrete positions results in spatial aliasing due to spatial sampling. The cut-off frequency is given by [2]

$$f_{al} = \frac{c}{2\Delta x \sin \alpha_{max}}, \tag{1}$$

where α_{max} denotes the maximum angle of incidence of the synthesized wave field relative to the loudspeaker array, c the speed of
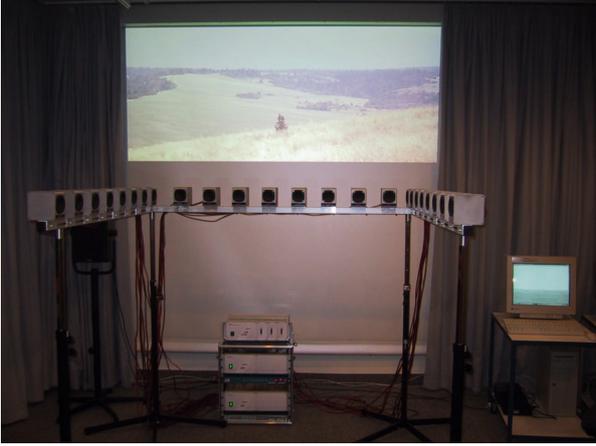
Figure 2: Typical setup of loudspeakers for WFS

sound and Δx the loudspeaker spacing. Assuming a loudspeaker spacing of $\Delta x = 19$ cm, the minimum spatial aliasing frequency is $f_{al} \approx 900$ Hz. Regarding the standard audio bandwidth of 20 kHz spatial aliasing seems to be a problem for practical WFS systems. Fortunately, the human auditory system is not very sensitive to these aliasing artifacts.

PSfrag replacements

2.2. Rendering Techniques

In general, the loudspeaker driving signals can be expressed as a convolution of measured or synthesized impulse responses $\mathbf{W}[k]$ with the source signals:

$$\mathbf{q}[k] = \mathbf{W}[k] * \mathbf{s}[k], \qquad (2)$$

where k denotes the discrete time index, $\mathbf{s}[k]$ the vector of M source signals and $\mathbf{q}[k]$ the vector of L loudspeaker driving signals. The impulse responses $\mathbf{W}[k]$ for auralization cannot be obtained the conventional way by simply measuring or simulating the impulse responses from a source to a listener position. The wave £eld has to be captured in a way that yields information on the traveling direction of the sound waves. There are two different approaches to compute the WFS matrix $\mathbf{W}[k]$ often referred to as rendering techniques:

1. *Data-based rendering*
   The impulse responses $\mathbf{W}[k]$ can be derived after recording with special microphones and post processing by wave £eld analysis techniques in order to extract the wave £eld information [4].

2. *Model-Based Rendering*
   Models for the spatial source characteristics are used to calculate the impulse response matrix $\mathbf{W}[k]$. Point sources and plane waves are the most common models used here.

2.3. Wave Field Analysis

Using techniques from seismic wave theory an acoustic wave £eld can be analyzed with special microphone arrays. The basic idea is to transform the pressure £eld $P(\mathbf{r}, \omega)$ into the spatial frequency domain by a spatial multidimensional Fourier transform with respect to the vector $\mathbf{r}$ of spatial coordinates. The temporal angular

frequency is denoted by ω. The complex amplitudes of the multidimensional Fourier transform can then by identi£ed as the amplitudes and phases of monochromatic plane waves [3]. This technique is therefore often referred to as *plane wave decomposition*. Because the spatial Fourier transform uses the same orthogonal basis functions as the well known temporal Fourier transform it also shares its properties. The plane wave decomposition has several bene£ts compared to working directly on the pressure £eld in our application: Information about the direction of the traveling waves is included, spatial properties of sources and receivers can be easily included into algorithms and plane wave decomposed wave £elds can be easily extrapolated to other positions.

In general, we will not have access to the whole three dimensional pressure £eld $P(\mathbf{r}, \omega)$ to calculate the plane wave decomposition using a multidimensional Fourier transform. Utilizing the Kirchhoff-Helmholtz integral, not only the concept of wave £eld synthesis can be derived, but also some tools for analyzing wave £elds. With the help of arrays consisting of pressure and velocity microphones, it is possible to decompose wave £elds into plane wave components using measurements on the boundary of the region of interest. In [4] the calculation of the plane wave decomposition is explained for various microphone array geometries.

3. LISTENING ROOM COMPENSATION

In this section, we point out the problem of compensating large listening areas and introduce our approach to overcome the drawbacks of common multi-point compensation systems.

3.1. Problem Statement

The theory of WFS systems as described above was derived assuming free £eld propagation of the sound emitted by the loudspeakers. In real systems, however, acoustic re¤ections at the walls of the listening room can degrade the sound quality, especially the perceptibility of the spatial properties of the auralized acoustic scene. Common room compensation algorithms are capable of dereverberating the listening room at some discrete points only (*multi-point equalization*) [5]. Outside these equalization points the sound quality is often even worse compared to the unequalized case. As wave £eld synthesis in principle allows to control the wave £eld within the listening area it can also be used to compensate for the re¤ections caused by the listening room. Of course this is only valid up to the spatial aliasing frequency (1) of the particular WFS system used. Figure 3 shows the signal ¤ow diagram of a WFS system including the in¤uence of the listening room. The
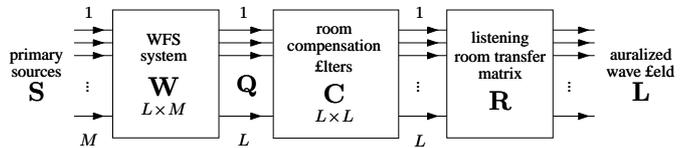

Figure 3: Block diagram of a WFS system including the in¤uence of the listening room and the compensation £lters

listening room characteristics are contained in the matrix $\mathbf{R}$, the set of compensation £lters in the matrix $\mathbf{C}$. According to £gure 3, the auralized wave £eld $\mathbf{L}(z)$ is given as follows

$$\mathbf{L}(z) = \mathbf{R}(z) \cdot \mathbf{C}(z) \cdot \mathbf{W}(z) \cdot \mathbf{S}(z), \qquad (3)$$

where e.g. $\mathbf{S}(z)$ denotes the Laplace transform of $\mathbf{s}[k]$. Perfect compensation of the listening room would be obtained if $\mathbf{R}(z) \cdot \mathbf{C}(z) = \mathbf{F}(z)$, where $\mathbf{F}(z)$ denotes the free field propagation matrix. In practice, however, it is not possible to fulfill this constraint in general. The next section will introduce our approach to calculate the compensation filters.

3.2. Room Compensation using Plane Wave Decomposition

One reason for multi-point equalization systems' failure in dereverberating large areas is the lack of information about the traveling directions of the reflected sound waves. Compensation signals traveling in other directions cancel out the reflections at the microphone positions only. Therefore, our approach is a novel compensation algorithm which takes into account directional information about the sound waves by utilizing the plane wave decomposed wave fields.

Our new room compensation system works as depicted in figure 4: First, we measure the wave field $\mathbf{R}$ produced by each loudspeaker
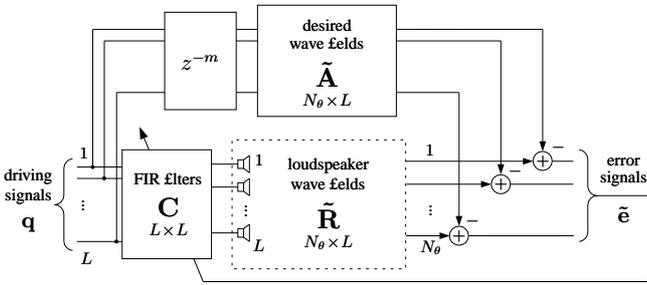

Figure 4: Block diagram of the proposed room compensation algorithm

inside the listening area using microphone arrays. Instead of using the microphone signals directly we perform a plane wave decomposition of the measured wave field as described in section 2.3. The transformed wave field is denoted as $\tilde{\mathbf{R}}$. We then adapt the compensation filters $\mathbf{C}$ of this MIMO system so that a given desired wave field $\tilde{\mathbf{A}}$ is met. For this purpose the cost function J derived from the error $\tilde{\mathbf{e}}$ is minimized:

$$\min_{\mathbf{C}(z)} \left(J(z) = \tilde{\mathbf{e}}^{\mathrm{H}}(z)\tilde{\mathbf{e}}(z) \right), \tilde{\mathbf{e}} = [\tilde{e}_1 \ldots \tilde{e}_{N_\theta}] \qquad (4)$$

Contrary to multi-point equalization algorithms, the error is not measured at several points but for several directions θ of the plane wave decomposed signals. Using the plane wave decomposed wave fields instead of the microphone signals has the advantage that the complete spatial information about the listening room influence is included. This allows to calculate compensation filters which are valid for the complete area inside the loudspeaker array. We choose a multichannel least-squares error (*LSE*) frequency domain inversion algorithm [6] to calculate the compensation filters $\mathbf{C}$. It minimizes the mean squared error over all directions N_θ of the plane wave decomposition for every frequency. As each plane wave component describes the wave field inside the whole listening area for one direction θ, minimizing the error for all directions results in filters compensating the whole listening area. Because in general the aliasing frequency of the measured wave field and the WFS system do not have to be the same, it has to be taken care to

select an appropriate number of directions N_θ for the plane wave decomposition.

3.3. Desired Wave Fields

If we apply this concept to a WFS system, we have to take the WFS driving signals $\mathbf{q}$ as the input signals for the listening room compensation. The desired wave fields $\tilde{\mathbf{A}}$ will be determined by the wave propagation from the speakers to the listening area, as assumed in the calculation of the WFS signals (e.g. implying loudspeakers acting like monopoles and free field propagation). This concept has the advantage of the room compensation filters being independent from the WFS operator. The drawback is the high number of compensation filters that have to be applied to the output signals of the WFS system in real time (L^2 for L loudspeakers). For stationary WFS operators $\mathbf{W}$ (as used for auralization without moving virtual sources), the WFS system is a linear time invariant (*LTI*) system. Therefore, the WFS operator can be integrated into the room compensation filters. Models for point sources and plane waves, as described in section 2.2, can be used as desired wave fields in this case. For M sources this results in $M \cdot L$ filters, which are in most cases significantly less than in the non-stationary case.

4. RESULTS

4.1. Experimental Setup

For our tests we used 16 channels of our 24 channel laboratory WFS system consisting of three linear loudspeaker arrays with 8 loudspeakers each as shown in figure 2. All tests were carried out in a low-reverberant room (reverberation time $T_{60} \approx 60$ ms) with one wall covered by a reflective material to get defined reflections from one direction only. Figure 5 shows the experimental setup used. The wave field produced by each loudspeaker was measured.
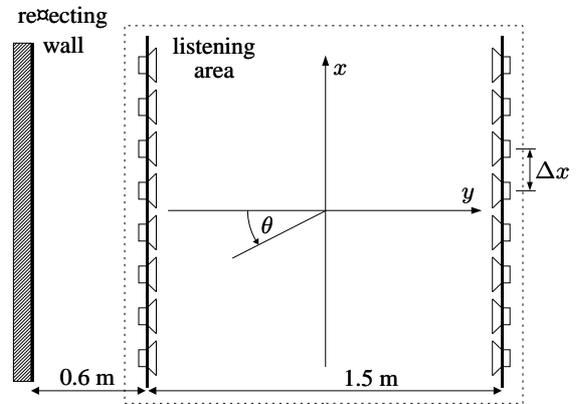

Figure 5: Experimental setup

This was done with a pressure and a pressure gradient microphone moved along the two axes of a cross shaped array centered in the listening area and performing a plane wave decomposition as described in [4]. Additionally some measurements were done with a linear microphone array that was moved 20 cm towards the reflecting wall to check if the compensation system works correctly for different areas of the listening room. As desired wave fields $\tilde{\mathbf{A}}$ we selected non-moving point sources and plane waves from the back of the room ($\theta = 180°$). Thus the wave field operator $\mathbf{W}$ is

included into the compensation filters. For each of these stationary scenarios, we calculated the (16×1) matrix $\mathbf{C}$ of room compensation filters including the WFS operator $\mathbf{W}$. A filter length of 8192 coefficients was suitable at a sampling frequency of 48 kHz.

4.2. Results

All results were calculated for band limited signals. The upper frequency bound was set to the aliasing frequency $f_{al} = 900$ Hz corresponding to the loudspeaker spacing $\Delta x = 19$ cm of our WFS system. A lower frequency bound of 100 Hz was chosen because the small WFS speakers are not designed to reproduce lower frequencies. The WFS system uses an additional subwoofer speaker for this task.

Results are shown for a plane wave as desired wave field. Results for point sources do not differ fundamentally. In order to visualize the three dimensional wave fields we calculated the signal power of the measured impulse responses in the plane wave domain. Additionally we calculated the error power between the desired wave field and the resulting wave field (corresponding to $\tilde{\mathbf{e}}$ in Fig. 4) compared to the power of the desired wave field. Figure 6 shows the results if no compensation is used. Apparently, the un-
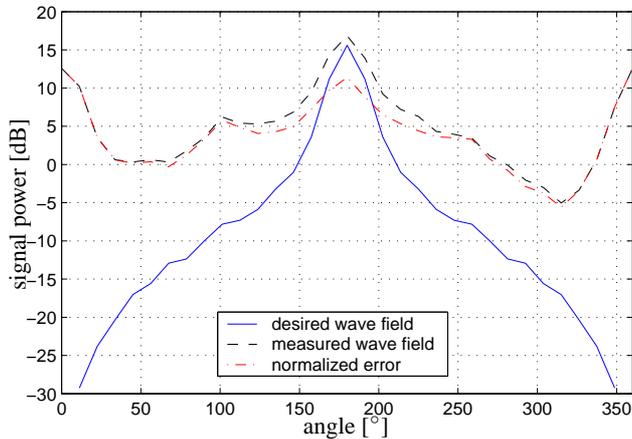

Figure 6: Resulting signal power of plane wave decomposed wave field without room compensation

compensated wave field exhibits the largest errors for plane wave components from $\theta = 0°$ which originate from the reflecting wall. As the desired wave field shows, there should be no signal power from this direction. Figure 7 shows the results if our compensation algorithm is used. In contrast to the uncompensated wave field the results from applying the compensation filters shows that the error power could be significantly reduced with our filter design approach. The largest gain of 18 dB is obtained at $\theta = 0°$. The error power over all directions was decreased by 12.9 dB compared to the uncompensated case. Experiments with other microphone positions show a moderate degradation of the gain obtained at the exact microphone positions used for the compensation. However, as we do not get a loss in gain compared to measured wave field, this shows that our room compensation system does not degrade the sound quality like multi-point equalization algorithms do outside their equalization points. There are several possible causes for this gain degradation: Inaccurate acquisition of the physical reality with the microphone array used, errors caused by the inver-
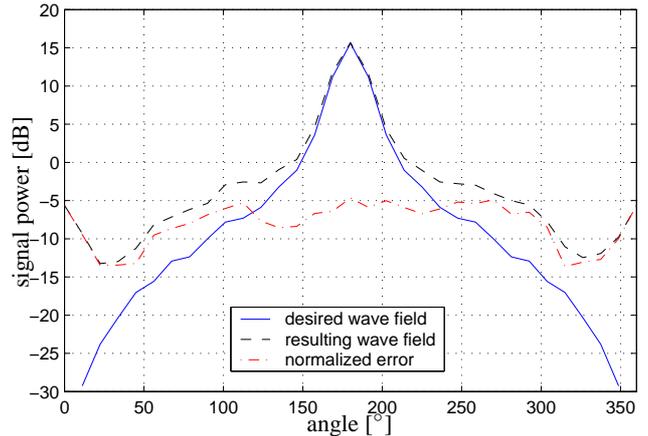

Figure 7: Resulting signal power of plane wave decomposed wave field with room compensation

sion algorithm and nonlinearities caused by the loudspeakers and microphones.

5. CONCLUSION

We have proposed a new approach for dereverberating listening rooms, especially for the application with WFS systems. Using wave field analysis and WFS our algorithm allows to compensate for listening room reflections in a large area. This results in a large compensated listening area. Primarily results indicate that our approach works, but could be improved. In our experiments we obtained a gain for different locations which shows that we do not share the problems of common multi-point equalization systems. Further work includes the use of circular microphone arrays as suggested in [4] and the combination of room compensation filters with loudspeaker compensation filters in the frequency range above the aliasing frequency.

6. REFERENCES

[1] "The CARROUSO project," http://emt.iis.fhg.de/projects/carrouso.

[2] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *Journal of the Acoustic Society of America*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[3] A.J. Berkhout, *Applied Seismic Wave Theory*, Elsevier, 1987.

[4] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis," in *110th AES Convention*, Amsterdam, Netherlands, May 2001, Audio Engineering Society (AES).

[5] J. N. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.*, vol. 42, no. 11, pp. 884–900, November 1994.

[6] O. Kirkeby, P. Nelson, H. Hamada, and Felipe Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Transactions on Speech and Audio Processing*, vol. 6, no. 2, pp. 189–194, March 1998.