

AN APPROACH TO LISTENING ROOM COMPENSATION WITH 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

Common room compensation algorithms are capable of dereverberating the listening room at some discrete points only. Outside of these equalization points the sound quality is often even worse compared to the unequalized case. However a new rendering technique - wave field synthesis - allows to control the wave field within the listening area. Therefore it can also be used to compensate for the reflections that the listening room causes in the complete listening area. We present a novel approach to listening room compensation which is based upon the theory of wave field synthesis. It yields improved compensation results in a large area.

1. INTRODUCTION

The reproduction of sound in enclosures is subject to certain impairments of the sound quality. One of the various causes are the reflections of the sound waves at the surfaces of the listening rooms. For mono reproduction, standing waves may result in an undesired coloration of the sound. For multichannel schemes, which have also the potential of spatial reproduction, the assessment of the listening room effects is more involved. On the one hand, room reflections may deteriorate not only the frequency response but also the spatial sound impression of the listener. On the other hand, the availability of a number of reproduction channels provides new degrees of freedom for the compensation if any unwanted listening room effects are present.

In conventional multi-channel schemes like two-channel stereo or 5.1, each reproduction channel is directly linked to the signal of a certain loudspeaker (left-right or L-C-R-LS-RS). Here we consider wave field synthesis (WFS), an advanced multi-channel reproduction approach where the number of loudspeakers is higher than any reasonable number of parallel transmission channels (typically tens or hundreds of loudspeakers). Thus the sound fields produced by WFS systems are not described in terms of loudspeaker signals but in terms of spatial directions of the reproduced sound events. The mathematical tool for such a description is the so-called wave field decomposition, a representation of sound fields in terms of plane waves closely related to the Radon transformation well-known in image processing.

WFS as a reproduction format for sound field recording, transmission, and reproduction has been established through the European research project CARROUSO [1]. This project has successfully demonstrated that sound fields can be captured by microphone-array techniques, encoded

and transmitted according to the MPEG-4 standard, and reproduced by wave field synthesis. This contribution describes a listening room compensation approach based on wave field decomposition that has been developed within the CARROUSO project.

Section 2 describes the theory and implementation of WFS rendering systems. The analysis of wave fields with wave field decomposition is presented in Section 3. A detailed review of classical and new approaches to listening room compensation is given in Section 4. The experiments conducted with WFS setup at our laboratory and the corresponding results are discussed in Section 5. Section 6 concludes the paper.

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, 3, 4, 5, 6] and is now investigated further within the CARROUSO project. In contrast to other multi-channel approaches, it is based on fundamental acoustic principles. This section gives a broad overview of the theory as well as on methods for rendering and listening room and loudspeaker compensation. In the context of WFS, rendering denotes the production of an appropriate sound field according to the use of the term rendering for computer graphics.

2.1. Theory

The theoretical basis of WFS is given by 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

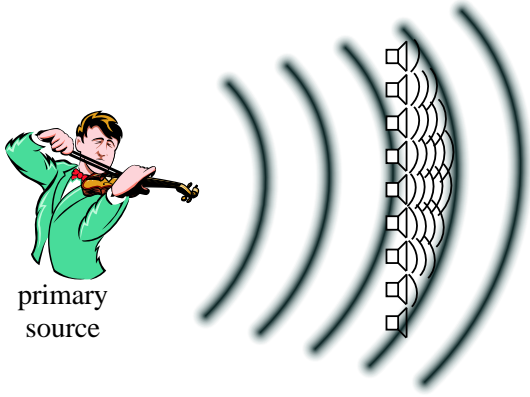


Figure 1: Basic principle of wave field synthesis.

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.

The mathematical foundation of this more illustrative description of WFS is given by the Kirchhoff-Helmholtz integral (1), which can be derived by using the wave equation and the Green's integral theorem [7]

$$P(\mathbf{r}, \omega) = \frac{1}{4\pi} \oint_S \left[P(\mathbf{r}_S, \omega) \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{-j\beta|\mathbf{r}-\mathbf{r}_S|}}{|\mathbf{r}-\mathbf{r}_S|} \right) - \frac{\partial P(\mathbf{r}_S, \omega)}{\partial \mathbf{n}} \frac{e^{-j\beta|\mathbf{r}-\mathbf{r}_S|}}{|\mathbf{r}-\mathbf{r}_S|} \right] dS. \quad (1)$$

Figure 2 illustrates the parameters used. In (1), S denotes the surface of an enclosed space V , $|\mathbf{r}-\mathbf{r}_S|$ the vector from a surface point \mathbf{r}_S to an arbitrary listener position \mathbf{r} within the volume V , $P(\mathbf{r}_S, \omega)$ the Fourier transform of the pressure distribution on S , β the wave number and \mathbf{n} the surface normal. The temporal angular frequency is denoted by ω . The Kirchhoff-Helmholtz integral states that at any listening point within the source-free volume V the sound pressure $P(\mathbf{r}, \omega)$ can be calculated if both the sound pressure and its gradient are known on the surface enclosing the volume. This can be used to synthesize a wave field within the surface S by setting the appropriate pressure distribution $P(\mathbf{r}_S, \omega)$, i.e. dipole sources and its gradient, i.e. monopole sources on the surface. This fact is used for WFS based sound reproduction. But there are several simplifications necessary to arrive at a realizable system:

1. Degeneration of the surface S to a plane between the primary sources and the listening area
2. Degeneration of the surface S to a line
3. Spatial discretization

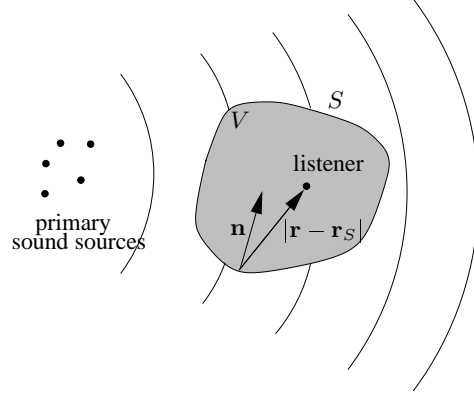


Figure 2: Definition of the parameters used for the Kirchhoff-Helmholtz integral.

The first step is to degenerate the surface S to a plane between the primary sources and the listening area. The wave field can be synthesized by either acoustic monopoles or dipoles alone. The Rayleigh I integral describes the mathematics for monopoles as follows

$$P(\mathbf{r}, \omega) = \rho c \frac{j\beta}{2\pi} \int \left[v_n(\mathbf{r}_S, \omega) \frac{e^{-j\beta|\mathbf{r}-\mathbf{r}_S|}}{|\mathbf{r}-\mathbf{r}_S|} \right] dS \quad (2)$$

where ρ denotes the static density of the air, c the speed of sound and v_n the particle velocity perpendicular to the surface. The Rayleigh II integral [6] applies for dipoles. In the second step, we observe that for our applications it is sufficient to synthesize the wave field correctly in the horizontal ear plane of the listener. For this scenario the surface further degenerates to a line Λ surrounding the listening area. As a first approximation, closed loudspeakers act as acoustic monopoles mounted on discrete positions. Equidistant spatial discretization of the line Λ with space increment $\Delta\lambda$ gives the discrete monopole positions \mathbf{r}_i . As a third step, we use this discretization and the assumption of an approximately stationary phase. Then the two dimensional Rayleigh I integral (2) can be transformed into one dimension

$$P(\mathbf{r}, \omega) = \sum_i \left[A_n(\omega) P(\mathbf{r}_i, \omega) \frac{e^{-j\beta|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|} \right] \Delta\lambda, \quad (3)$$

where $A_n(\omega)$ denotes a weighting factor. Using the equation above, WFS can be realized by mounting closed loudspeakers in a linear fashion (linear loudspeaker arrays) surrounding the listening area leveled with the listeners ears. Figure 3 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 [4]. This allows to place acoustical sources between the listener and the loudspeakers within the reproduction

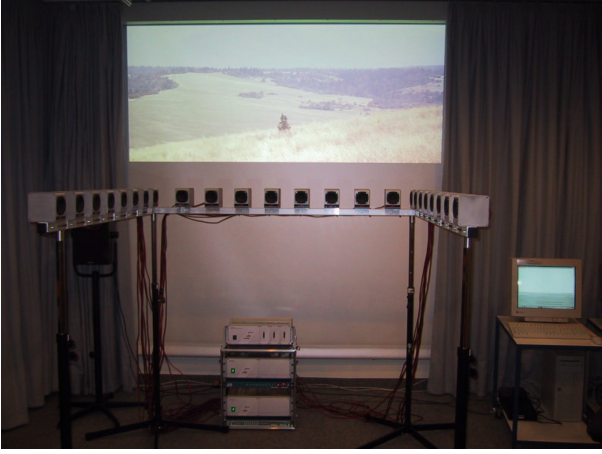


Figure 3: Typical setup of loudspeakers used for WFS.

area (focused sources). This is not possible with traditional stereo or 5.1 setups.

In practice, two effects limit the performance of real WFS systems:

1. Spatial aliasing

The discretization of the Rayleigh integral results in spatial aliasing due to spatial sampling. The cut-off frequency is given by [6]

$$f_{\text{al}} = \frac{c}{2\Delta x \sin \alpha_{\text{max}}}, \quad (4)$$

where α_{max} denotes the maximum angle of incidence of the synthesized wave field relative to the loudspeaker array. Assuming a loudspeaker spacing $\Delta x = 10$ cm, the minimum spatial aliasing frequency is $f_{\text{al}} = 1700$ 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.

2. Truncation effects

These effects are caused by wavefronts which propagate from the ends of the loudspeaker array. They can be understood as diffraction waves caused by the finite number of loudspeakers in practical implementations. Truncation effects can be minimized by filtering in the spatial domain (tapering) [6].

After this brief review of the acoustic theory, two different rendering approaches will be presented: model-based rendering and data-based rendering. Both approaches follow from suitable specialization of the discrete version of the Rayleigh I integral (3).

2.2. Model-based rendering

For model-based rendering, models for the sources are used to calculate the driving signals for the loudspeakers. Point sources and plane waves are the most common

models used here. For a point source equation (3) becomes

$$P(\mathbf{r}, \omega) = Q(\omega)K \sqrt{\frac{j\beta}{2\pi}} \times \sum_i \left[\frac{e^{-j\beta|\mathbf{r}_i - \mathbf{r}_m|}}{|\mathbf{r}_i - \mathbf{r}_m|^3} \frac{e^{-j\beta|\mathbf{r} - \mathbf{r}_i|}}{|\mathbf{r} - \mathbf{r}_i|} \right] \Delta x; \quad (5)$$

where $Q(\omega)$ denotes the spectrum of the source signal, Δx the distance between the loudspeakers, K is a geometry dependent constant and \mathbf{r}_m denotes the position of the point source. The spectrum of the loudspeaker driving signals $W_i(\omega)$ can be derived from (5) as

$$W_i(\omega) = Q(\omega)K \sqrt{\frac{j\beta}{2\pi}} \frac{e^{-j\beta|\mathbf{r}_i - \mathbf{r}_m|}}{|\mathbf{r}_i - \mathbf{r}_m|^3} \Delta x, \quad (6)$$

By transforming this equation back into the time-domain and employing time discretization the loudspeaker driving signals can be computed from the source signals by delaying, weighting and filtering,

$$w_i[k] = a_n (h[k] * q[k]) * \delta[k - \kappa], \quad (7)$$

where a_n and κ denote an appropriate weighting factor and delay respectively, $h[k]$ the inverse Fourier transform of $\sqrt{j\beta/2\pi}$. Multiple (point) sources can be synthesized by superimposing the loudspeaker signals from each source. Plane waves and point sources can be used to simulate classical loudspeaker setups, like stereo and 5.1 setups. Thus WFS is backward compatible to existing sound reproduction systems and can even improve them by optimal loudspeaker positioning in small listening rooms and listening room compensation, as discussed in this paper.

2.3. Data-based rendering

The loudspeaker driving signals $W_i(\omega)$ for arbitrary wave fields can be computed according to equation (3) as follows

$$W_i(\omega) = A_n(\omega)P(\mathbf{r}_i, \omega). \quad (8)$$

The pressure distribution $P(\mathbf{r}_i, \omega)$ contains the entire information of the sound field produced at the loudspeaker position \mathbf{r}_i by a source $Q(\omega)$. The propagation from the source to the loudspeaker position \mathbf{r}_i can be modeled by a multidimensional transfer function $H(\mathbf{r}_i, \omega)$ assuming linear wave propagation. By incorporating the weighting factors $A_n(\omega)$ into $H(\mathbf{r}_i, \omega)$ we can calculate the loudspeaker driving signals as follows

$$W_i(\omega) = H(\mathbf{r}_i, \omega)Q(\omega). \quad (9)$$

By transforming this equation back into the discrete time domain, the vector of loudspeaker driving signals $\mathbf{w}[k] = [w_1, w_2, \dots, w_M]^T$ can be expressed as a multichannel

convolution of measured or synthesized impulse responses with the source signals $\mathbf{q}[k] = [q_1, q_2, \dots, q_M]^T$

$$\mathbf{w}[k] = \mathbf{H}[k] * \mathbf{q}[k], \quad (10)$$

where $\mathbf{H}[k]$ denotes a matrix of suitable impulse responses. The impulse responses for auralization cannot be obtained the conventional way by simply measuring the response from a source to a listener position. In addition to the sound pressure also the particle velocity is required to extract the directional information. This information is necessary to take into account the direction of the traveling waves during auralization. These room impulse responses have to be recorded by special microphones and setups as shown in section 3 and extrapolated to the loudspeaker positions [7].

2.4. Practical implementation of a WFS based rendering system

The WFS setup at our laboratory consist of 24 wideband loudspeakers and a subwoofer as shown in Figure 3. The loudspeakers are driven by multichannel amplifiers that were developed at our lab for this purpose. We utilize commercial available multichannel DA-converters with ADAT interface to feed the amplifiers. The ADAT signals are provided by a digital multi-channel soundcard in a PC. We developed software for model and data based rendering. The software was developed for the LINUX operating system and is running in real-time.

Although still under development, our model-based rendering software already provides the following features:

- synthesis of point sources and plane waves
- synthesis of moving point sources with arbitrary trajectories
- interactive graphical user interface for loudspeaker and source setup
- room effects using a mirror image source model
- source input from files or ADAT/SPDIF
- simulation of a 5.1 loudspeaker setup

Figure 4 shows a snapshot of the graphical user interface of our model-based real-time rendering software. The upper half of the application window comprises the loudspeaker and source setup. Sources can be moved intuitively in real-time by clicking on the source and dragging using the computer mouse. One of the shown sources moves on a trajectory which is also displayed in the window. The lower half of the application window controls the synthesis and application parameters and the setup for the virtual room used for the mirror image model. All parameters can be changed in real-time during operation.

For data based rendering we utilize BruteFIR [8], a very fast real-time convolution software. Using a multiprocessor workstation, the computational complex convolutions can be performed in real-time by our system. To reduce the computationally complexity when rendering scenes with long reverberation times the rendering of high-quality natural audio is done typically by reproducing the direct sound as point source at the source position and the reverberation as eight plane waves as shown in [9, 10].

3. WAVE FIELD ANALYSIS

Room compensation requires that we are able to determine the influence of the listening room on the auralized wave field to take a suitable action to compensate for these effects. The influence of the listening room on the dry loudspeaker signals is measured with wave field analysis techniques in our framework. The following section will introduce the necessary tools.

Using techniques from seismic wave theory tools for the analysis of acoustic wave fields can be developed [13]. The eigensolutions of the acoustic wave equation in three dimensional space appear in different forms, depending on the type of the adopted coordinate system. For a spherical coordinate system the simplest solution of the wave equation are spherical waves, while plane waves are a simple solution for Cartesian coordinates. Of course, plane waves can be expressed as a superposition of spherical waves (see Huygen's principle) and vice versa [11]. We aim at performing a spatial transform of the acoustic wave field $P(\mathbf{r}, \omega)$ into a domain which gives more insight into the structure of the field. Accordingly two types of transformations exist for wave fields, the decomposition of the field into spherical harmonics or plane waves. We use a plane wave decomposition in our framework. For practical reasons we will assume two dimensional wave fields in the following. In this special case a cylindrical coordinate system is used. The basic idea is to transform the pressure field $P(\mathbf{r}, \omega)$ into plane waves with the incident angle θ and the intercept time τ with respect to a reference point:

$$p(\mathbf{r}, t) \xrightarrow{\mathcal{P}} s(\theta, \tau) \quad (11)$$

This technique is therefore often referred to as *plane wave decomposition*. We will denote the transformation in our paper with \mathcal{P} . The plane wave decomposition maps the pressure field into an angle, offset domain. This transformation is also well known as *Radon transformation* from image processing. The Radon transformation maps straight lines in the image domain into Dirac peaks in the Radon domain [12]. It is therefore typically used for edge detection in digital image processing.

The Radon transformation of a pressure field $p(\mathbf{r}, t)$ can be reduced to a one-dimensional integration over the pres-

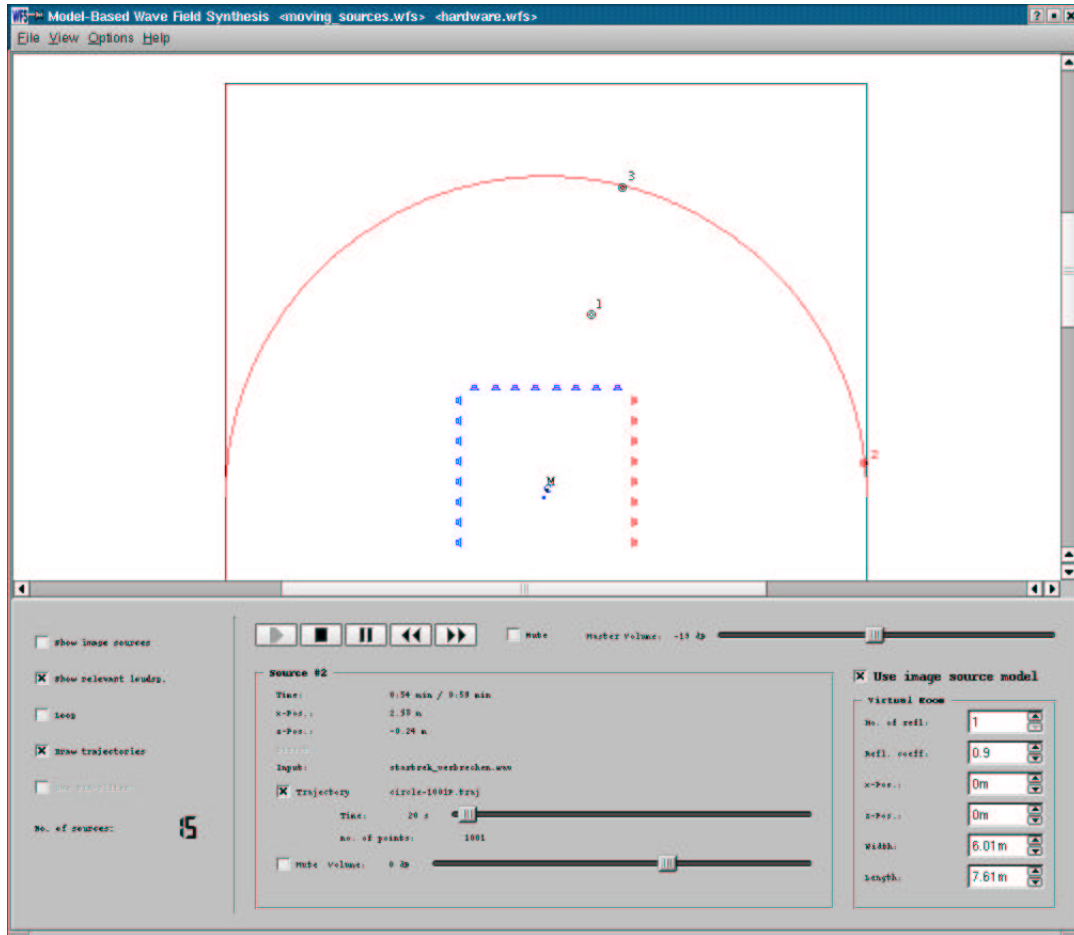


Figure 4: Screenshot of our model-based rendering application

sure distribution on a line $p(x, t)$ as follows:

$$s(\theta, \tau) = \mathcal{P}\{p(x, t)\} = \int_{-\infty}^{\infty} p(\tau \sin \theta + s \cos \theta, \tau \cos \theta - s \sin \theta) ds. \quad (12)$$

More insights into the properties of the plane wave decomposition can be derived by a multidimensional Fourier-transformation of the pressure field $P(\mathbf{r}, \omega)$ with respect to the vector \mathbf{r} of spatial coordinates. The complex amplitudes of the multidimensional Fourier transform can then be identified as the amplitudes and phases of monochromatic plane waves [13]. 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 benefits compared to working directly on the pressure field 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 fields can be easily extrapolated to other positions [7].

In general, there is no access to the whole two dimensional pressure field $P(\mathbf{r}, \omega)$ to calculate the plane wave decomposition using a multidimensional Fourier transform. However, the Kirchhoff-Helmholtz integral (1) allows to calculate the acoustic pressure field $P(\mathbf{r}, \omega)$ from the sound pressure and its gradient on the line enclosing the desired field and vice versa (see section 2.1). Therefore, the Kirchhoff-Helmholtz integral (1) can not only be used to introduce the concept of wave field synthesis, but also to derive an efficient implementation of the plane wave decomposition for acoustic wave fields.

3.1. Plane wave decomposition for a circular microphone array

The calculation of the plane wave decomposition is derived for various microphone array configurations in [14]. Because a circular microphone array has many favorable advantages over other configurations we will use this configuration for our purpose. We will shortly review the algorithm for the calculation of the plane wave decomposition for a circular array as described in [14]. Because the

Kirchhoff-Helmholtz integral would only allow to calculate the acoustic pressure field inside the circular array the plane wave decomposition is derived using cylindrical harmonics. The first step is to calculate the cylindrical harmonics from the acoustic pressure $p(\theta, t)$ and the sound velocity component in normal direction to the array $v_n(\theta, t)$ as follows

$$M^{(1)}(n_\theta, \omega) = \frac{H_{n_\theta}^{(2)}(\beta R)P(n_\theta, \omega) - H_{n_\theta}^{(2)}(\beta R)j\rho cV_n(n_\theta, \omega)}{H_{n_\theta}^{(1)}(\beta R)H_{n_\theta}^{\prime(2)}(\beta R) - H_{n_\theta}^{(2)}(\beta R)H_{n_\theta}^{\prime(1)}(\beta R)} \quad (13a)$$

$$M^{(2)}(n_\theta, \omega) = \frac{H_{n_\theta}^{\prime(1)}(\beta R)P(n_\theta, \omega) - H_{n_\theta}^{(1)}(\beta R)j\rho cV_n(n_\theta, \omega)}{H_{n_\theta}^{(2)}(\beta R)H_{n_\theta}^{\prime(1)}(\beta R) - H_{n_\theta}^{(1)}(\beta R)H_{n_\theta}^{\prime(2)}(\beta R)} \quad (13b)$$

where $P(n_\theta, \omega)$ and $V_n(n_\theta, \omega)$ denote the two dimensional Fourier transforms of the pressure and velocity measurements $p(\theta, t)$ and $v_n(\theta, t)$, R the radius of the array, n_θ the order of the cylindrical harmonic, $H_{n_\theta}^{(1)}$ and $H_{n_\theta}^{(2)}$ the Hankel functions of first and second kind and $H_{n_\theta}^{\prime(1)}$ and $H_{n_\theta}^{\prime(2)}$ their derivatives. The wave field is decomposed into incoming $M^{(1)}$ and outgoing $M^{(2)}$ cylindrical harmonics. The plane wave decomposition can then be calculated in a second step as

$$s^{(1)}(\theta, \omega) = \frac{1}{2\pi} \sum_{n_\theta} j^{(1-n_\theta)} M^{(1)}(n_\theta, \omega) e^{jn_\theta\theta} \quad (14a)$$

$$s^{(2)}(\theta, \omega) = \frac{1}{2\pi} \sum_{n_\theta} j^{(1+n_\theta)} M^{(2)}(n_\theta, \omega) e^{jn_\theta\theta} \quad (14b)$$

For practical reasons the acoustic pressure and its gradient will only be measured on a limited number of positions on the circle. This spatial sampling can result in spatial aliasing in the plane wave decomposed field if the sampling theorem is not observed. The aliasing frequency f_s can be approximated as follows

$$f_s \approx c \frac{L}{4\pi R} \quad (15)$$

where L denotes the number of different angles used for the measurement.

The incoming $s^{(1)}$ and outgoing $s^{(2)}$ parts of the plane wave decomposition can be used to distinguish between sources inside and outside the circular array. While sources outside result in a incoming part which is equal to the outgoing part, sources inside the array are only present in the outgoing part. By using only the incoming part $s^{(1)}$ sources inside the array are omitted.

Using a two dimensional wave field analysis method in a three dimensional environment causes additional problems. Wave fields emitted from sources that do not lie in

the same plane as the microphone array cause responses in the plane wave domain which can in general not be distinguished easily from sources in the same plane. This can cause artifacts when used for auralization purposes with an WFS system or for the compensation of room acoustics. Besides these drawbacks the benefit of the plane wave decomposition is that we can capture the acoustical characteristics of a whole area through a measurement on the boundary of this area. Plane wave decomposed impulse responses therefore describe the acoustical properties of the whole area surrounded by the microphone array up to the aliasing frequency.

Because the number of microphone channels that have to be captured for typical configurations of a circular microphone array exceed the number of channels that can be captured in real-time by currently available audio hardware such a circular array is typically realized by sequential measurement of the discrete positions on the circle. We use a stepper motor for this purpose as shown in figure 8.

4. 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.

4.1. Review of classical room compensation approaches

The equalization of listening rooms is a topic of past and current research. Listening room compensation aims at improving the perceived quality in sound reproduction in non anechoic environments. When listening to a recording that itself contains the reverberation of the recorded scene, which is the typical case, the reverberation caused by the listening room interferes with the recorded sound field in a potential disturbing way. Perfect listening room compensation would eliminate the effects caused by the reproduction room. Unfortunately there are a number of pitfalls when designing a room compensation system. A good overview on classical room compensation approaches and their limitations can be found in [15]. Omitting the imperfect characteristics of the loudspeaker and the microphone, the effect of the listening room can be measured by the impulse response from the loudspeaker to a microphone. In principle, perfect compensation could be achieved by prefiltering the loudspeaker signal with the inverted impulse response. Unfortunately typical room impulse responses have in general non-minimum phase [16] which prohibits to calculate an exact inverse. A number of algorithms exist to approximate the inverse filter. The situation becomes better when multiple playback channels and equalization points are used. In this situation the listening room can be modeled as multiple-input/multiple-output (MIMO) system. The multiple-input/

output inverse theorem (MINT) [17] provides an exact solution for most situations. The solution is achieved by rewriting the convolution of the loudspeaker signal with the room impulse response as matrix operation. The problem of inverse filtering can then be formulated as the algebraic problem of inverting a matrix. Because of the special structure of the matrix obtained by multichannel convolution (block Toeplitz matrix) an analytical solution and its limitations can be found. The measured impulse responses contain only the influence of the listening room at the particular position they were captured. This is the reason why these classical algorithms can only provide equalization of the listening room at these measured positions. These algorithms are therefore often termed as *multi-point equalization* algorithms [18]. Outside these equalization points the sound quality is often even worse than in the uncompensated case [15, 19].

4.2. Problem statement

The theory of WFS systems as described in section 2 was derived assuming free field propagation of the sound emitted by the loudspeakers. In real systems, however, acoustic reflections 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. 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. Of course this is only valid up to the spatial aliasing frequency (4) of the particular WFS system used. Above the aliasing frequency there is no full control over the wave field. Destructive interference to compensate for the listening room reflections will fail here. We will focus mainly on the computation of compensation filters below the aliasing frequency in this paper. Therefore all signals are low-pass filtered and downsampled to the spatial aliasing frequency of the loudspeaker array. Some indications what can be done above the aliasing frequency will be shown in section 4.5.

Figure 5 shows the signal flow diagram of room compensation with WFS. The N primary source signals $\mathbf{q}[k]$ are first fed through the WFS system operator $\mathbf{H}[k]$. For the sake of generality the WFS system is modeled as matrix of impulse responses. This covers both the model-based and data-based rendering approach. The influence of the listening room is then compensated through the compensation filters $\mathbf{C}[k]$. The resulting signals are then played back through the M loudspeakers of the WFS system. The loudspeaker and listening room characteristics are contained in the matrix $\mathbf{R}[k]$. This matrix does not only contain the temporal characteristics of the loudspeakers but also their directivity properties. After convolution

with these impulse responses the auralized wave field at the L listening position(s) is expressed through the vector $\mathbf{L}[k]$. According to figure 5, the auralized wave field $\mathbf{L}(z)$ is given as follows

$$\mathbf{L}(z) = \mathbf{R}(z) \cdot \mathbf{C}(z) \cdot \mathbf{H}(z) \cdot \mathbf{q}(z), \quad (16)$$

where e.g. $\mathbf{q}(z)$ denotes the Laplace transform of $\mathbf{q}[k]$. In principle, perfect compensation of the listening room would be obtained if

$$\mathbf{R}(z) \cdot \mathbf{C}(z) = \mathbf{F}(z), \quad (17)$$

where \mathbf{F} contains suitable impulse responses from the loudspeakers to the microphone positions for free field propagation (e. g. implying loudspeakers acting like monopoles and free field propagation). The next section will introduce our approach to calculate the compensation filters which is based on the above illustrated framework combined with the concept of WFS and wave field analysis.

4.3. Room compensation using the 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 plane wave decomposed wave fields. Instead of using the microphone signals directly we perform a plane wave decomposition of the measured wave field as described in section 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. 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 in principle valid for the complete area inside the loudspeaker array.

There are two possible strategies for the calculation of compensation filters:

1. Calculation of a full set of compensation filters

If we apply our concept straightforward to a WFS system, the desired wave field $\tilde{\mathbf{A}}$ will be determined by the wave propagation from the speakers to the listening area, as assumed in the calculation of the WFS signals. 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

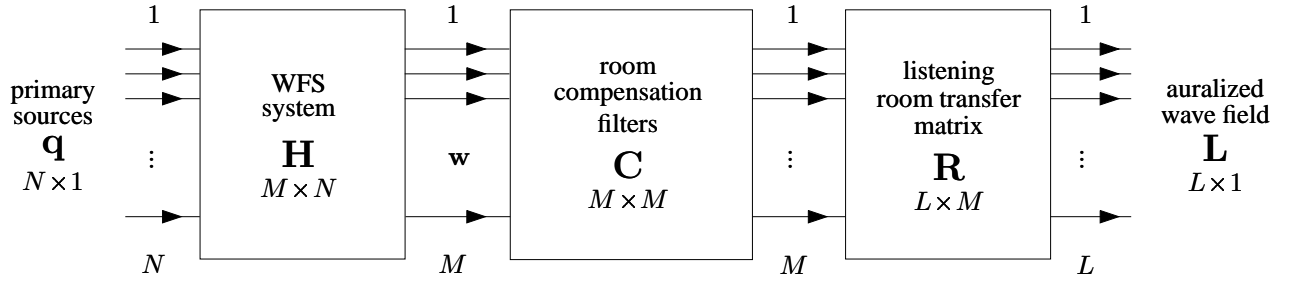


Figure 5: Block diagram of a WFS system including the influence of the listening room and the compensation filters to cope for the listening room reflections

signals of the WFS system in real time (M^2 for M loudspeakers).

2. *Incorporate WFS operator into compensation filters*
For quasi-stationary WFS operators \mathbf{H} (as used for auralization with slowly 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 N sources this results in $N \cdot M$ filters, which are in most cases significantly less than in the first approach.

In both cases it is possible to calculate the compensation filters by solving the following equation

$$\tilde{\mathbf{R}}(z) \cdot \mathbf{C}(z) = \tilde{\mathbf{A}}(z) \quad (18)$$

with an appropriate desired wave field $\tilde{\mathbf{A}}$. Using the MINT it is possible to solve the above equation exactly under certain realistic conditions as shown in [17]. This allows to have exact inverse filters. One of the basic requirements for a MINT based solution is that the number of loudspeakers M has to be higher than the number of equalization points/plane wave components L . Unfortunately this exact solution requires to have quite long filters for a typical WFS setup. The length of the compensation filters is given as follows [17]

$$N_C = \frac{M(N_R - 1)}{M - L} \quad (19)$$

where N_R and N_C denote the length of the measured impulse responses and the compensation filters, respectively. Another problem when implementing the MINT for MIMO systems with many inputs and outputs is the computational complexity required. A straightforward implementation of the MINT is therefore currently not feasible for typical WFS systems. We utilize a least-squares error (LSE) method to calculate the compensation filters in our framework for these reasons.

4.4. Practical implementation of room compensation

We chose a multichannel LSE frequency domain inversion algorithm [20] to calculate the compensation filters \mathbf{C} . Figure 6 shows a block diagram of its application to room compensation for WFS using the first approach described to calculate the compensation filters. The inversion algorithm minimizes the following cost function J derived from the error $\tilde{\mathbf{e}}$:

$$\min_{\mathbf{C}(z)} \left(J(z) = \tilde{\mathbf{e}}^H(z) \tilde{\mathbf{e}}(z) \right), \tilde{\mathbf{e}} = [\tilde{e}_1 \dots \tilde{e}_L] \quad (20)$$

In 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. This results in minimization of the mean squared error over all directions L 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. The compensation filters can then be computed as [20]

$$\mathbf{C}(z) = \left(\tilde{\mathbf{R}}^T(z^{-1}) \tilde{\mathbf{R}}(z) + \gamma B(z^{-1}) B(z) \mathbf{I} \right)^{-1} \times \tilde{\mathbf{R}}^T(z^{-1}) \tilde{\mathbf{A}}(z) z^{-m} \quad (21)$$

where γ denotes a suitable regularization weight, $B(z)$ the frequency function for the regularization weight and z^{-m} a suitable modeling delay. The modeling delay is required for causal compensation filters. The advantage of this approach is that the length of the resulting inverse filters can be chosen such that their computational complexity is bounded and audible artifacts [15] due to the truncation of the filter length compared to the exact solution of the MINT are minimized.

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 L for the plane wave decomposition.

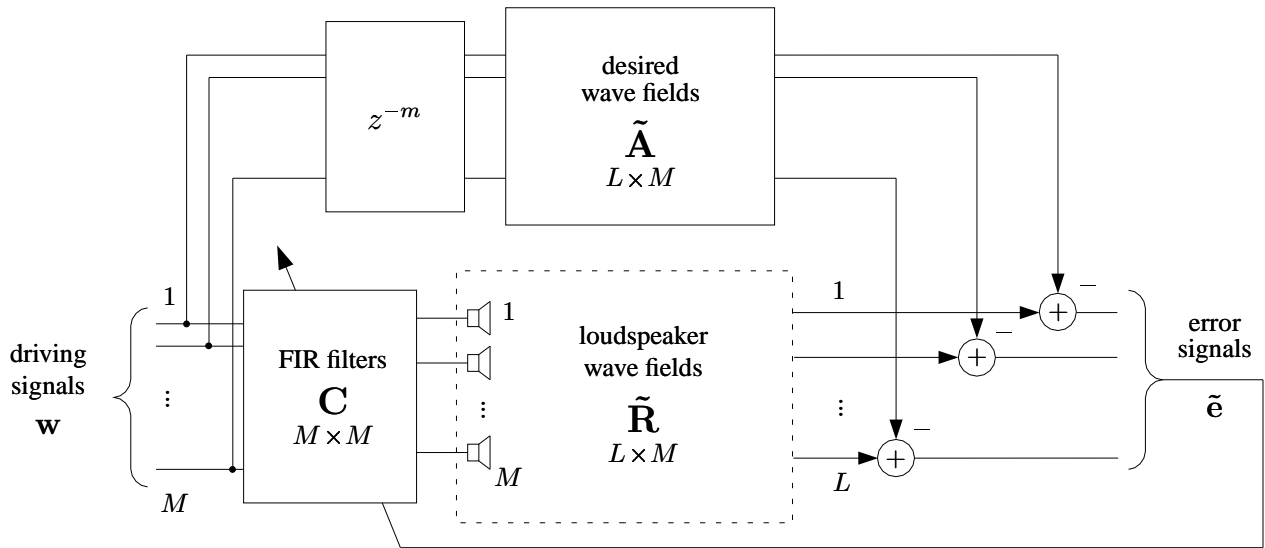


Figure 6: Block diagram of the proposed room compensation algorithm for the calculation of a full set of compensation filters

4.5. Compensation above the aliasing frequency

Above the aliasing frequency of the loudspeaker array no destructive interference can be used to compensate for the reflections caused by the listening room. True room compensation in the sense of a resulting free field wave propagation for each loudspeaker is therefore not achievable above the aliasing frequency. What can be done quite easily is to compensate the frequency response of each individual loudspeaker. An algorithm for loudspeaker compensation in the framework of WFS can be found e.g. [21]. Nevertheless the combination of loudspeaker compensation filters above the aliasing frequency with room compensation filters below the aliasing frequency is a current research topic at our lab. Other ideas include the use of psychoacoustic properties of the human auditory system to hide listening room reflections as described in [22].

5. EXPERIMENTS

The presented approach for listening room compensation with wave field synthesis relies on one major assumption: It is possible to reduce the reflections within the listening area of a three dimensional enclosure using a planar line array of loudspeakers. Unfortunately, such line arrays are only capable of controlling the sound field within a plane. A number of experiments has been conducted to investigate whether a successful listening room compensation in typical environments is still possible. This section shows the experimental setup used for the experiments and discusses results for listening room compensation.

5.1. Experimental setup

For our tests we used our 24 channel laboratory WFS system consisting of three linear loudspeaker arrays with 8 loudspeakers each as shown in figure 3. All tests were carried out in our demonstration room with the size $5.8 \times 5.9 \times 3.1$ meters (Volume about 105 m^3). Figure 7 defines the geometry used in our experiments. All walls of the room are covered by a removable acoustic absorbent curtain except behind the array on the left side ($\theta = 90^\circ$). This has two reasons: The first one is that it is only possible to compensate for room reflections coming from directions where loudspeakers are present and the second one is that the effects of room compensation can be seen much clearer in such a quite simple setup consisting of strong reflections from only one direction. All results were calculated for band limited signals. The upper frequency bound was set to the aliasing frequency $f_{al} = 900 \text{ Hz}$ corresponding to the loudspeaker spacing $\Delta x = 19 \text{ cm}$ 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.

The experimental procedure consisted of three steps

1. Measurement of the wave field produced by each loudspeaker
2. Calculation of the room compensation filters
3. Evaluation of the results

These steps are now described in detail.

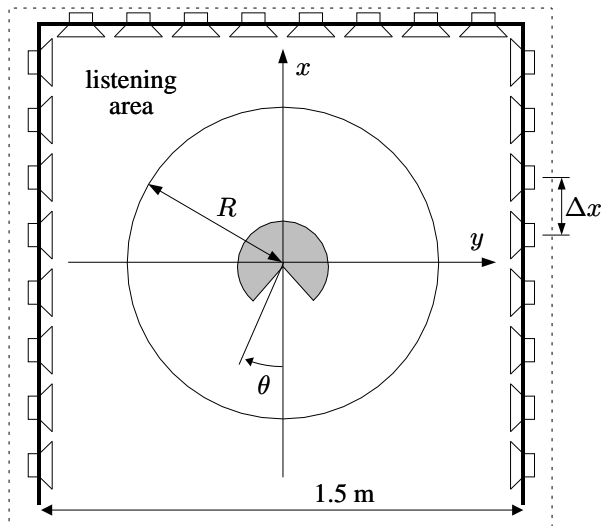


Figure 7: Definition of the geometry used for the experiments.

1. The wave field emitted by each loudspeaker was measured with a circular microphone array as described in section 3.1. The circular array was realized by mounting a pressure (omni-directional) and a gradient microphone (figure of eight) on a rod turned by a stepper drive. The radius of the array was $R = 50$ cm and 32 positions were measured on the circle. Figure 8 shows a picture of the measurement setup used for the experiments. The whole measurement was performed with a PC using a maximum length sequence (MLS) based impulse measurement method. The recorded signals for each loudspeaker were then plane wave decomposed into 32 angles. Only the incoming part of the plane wave decomposition is used assuming that no sources are present inside the array. We chose the frontmost $L = 24$ plane wave components from $\theta = 45^\circ$ to $\theta = 315^\circ$ as shown in figure 7 to cope with the angles where no compensation is possible. The result of this procedure is the matrix \mathbf{R} describing the room acoustics.
2. From the recorded room characteristics obtained in step 1, we calculated (24×1) matrix \mathbf{C} of room compensation filters including the WFS operator \mathbf{H} (as described in section 4.3). We selected a plane wave coming from the right side of the room ($\theta = 270^\circ$) as desired wave field. Using the MINT approach to calculate the inverse filters a length of 24 times the length of the recorded room impulse responses would be necessary. Using the LSE approach described in section 4.4 it is possible to calculate approximated inverse filters with shorter length with the drawback of pre- and post-ringing effects caused by the truncation.

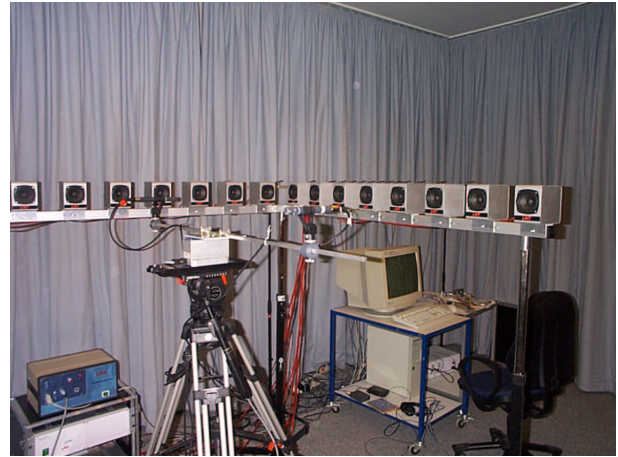


Figure 8: Measurement setup for the room compensation experiments.

An inverse filter length of four times the measured room impulse response was proven to be suitable in our experiments to calculate filters with the desired properties described in [15].

3. To show the reduction of room reflections by the proposed compensation method, the resulting wave fields have been analyzed using the measured room response. The results of this analysis are presented in detail in the following section.

5.2. Results

In the following we will present the results obtained with our proposed algorithm as plane wave decomposition of the respective fields. We used the measured wave field of each loudspeaker and the calculated compensation filters for this purpose. All results were computed for a plane wave originating from $\theta = 270^\circ$ and only for the 24 angles used for the compensation algorithm. In principle the plane wave decomposed wave fields should represent the whole area inside the circular microphone array (according to section 3). Unfortunately this assumption holds only for two dimensional wave fields.

Figure 9(a) shows the measured wave field. For better visibility of the relevant part only the part of the time axis ranging from $t = 0$ ms to $t = 100$ ms is shown. The gray levels denote the signal level in dB. The effects of the listening room on the desired dry wave field can be seen clearly. Reflections from almost every direction are present, the strongest ones from $\theta \approx 90^\circ$. This is due to the fact that the desired wave field is a plane wave coming from $\theta = 270^\circ$ and that the strongest reflection occurs at the opposite side. Figure 9(b) shows the resulting field after applying the proposed room compensation algorithm. In this case the time axis ranges from $t = 350$ ms to $t = 450$ ms because of the modeling delay intro-

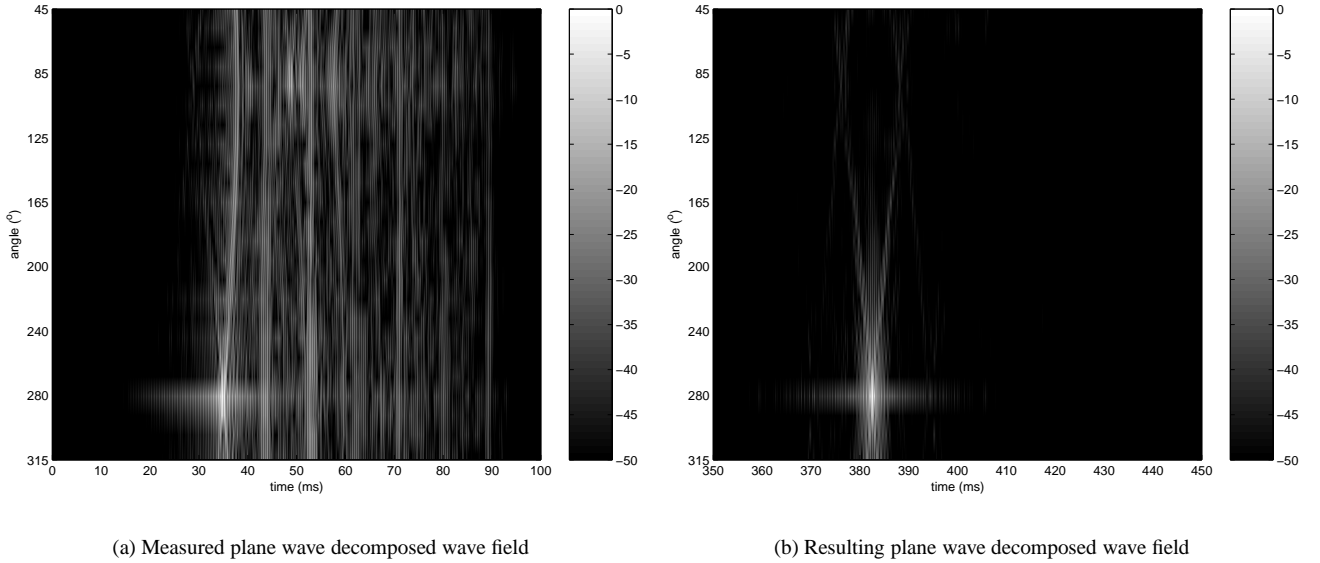


Figure 9: Results of room compensation shown as plane wave decomposition of the recorded and resulting impulse responses. The gray level scale in the plots denotes the signal level in dB.

duced when calculating the room compensation filters. It can be clearly seen that the reverberation caused by the room can be compensated by our approach. Additionally it can be seen that the room compensation filters do not produce artifacts in the resulting field. In order to better visualize the results of room compensation we also calculated the signal powers of the plane wave components. E.g. for the measured wave field the power can be calculated as follows

$$P_{\text{measured}}(\theta) = \sum_k |\tilde{L}_\theta[k]|^2 \quad (22)$$

where \tilde{L}_θ denotes the plane wave component of angle θ of the resulting wave field without applying room compensation ($\mathbf{C} = \mathbf{I}$). The same measure was calculated for the desired wave field and when the room compensation algorithm was used. Additionally we calculated the power of the difference between the desired wave field and the resulting wave field (corresponding to $\tilde{\mathbf{e}}$ in Fig. 6) after aligning them in the time domain. Figure 10 shows these results for the measured, the resulting and the target wave field. Additionally the remaining error after applying room compensation is shown. The reflection from $\theta = 90^\circ$ can be clearly seen in the measured field. It can also be seen that our room compensation algorithm yields results which are close to the desired wave field resulting in a reasonable dereverberation of the listening room. In order to further investigate the performance of our room compensation algorithm we also performed measurements at some discrete points inside the equalized area. These measurements revealed that the room compensation is not

working as ideally as indicated by the above presented results. The equalization is not optimal for the whole equalized area. But nevertheless we did not observe results worse than without room compensation, as this would be the case for multi-point equalization algorithms outside their equalization points. The degradation of the room compensation inside the equalized area indicates that the two dimensional techniques used in our framework are not capable of distinguishing between reflections coming from within the equalized plane and those from elevated angles. As a result the room compensation filters try to compensate for reflections that are not present in the plane defined by the both arrays. A detailed investigation of these effects is currently under research.

6. 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. A large compensated listening area is thus achieved. This is a result of the acoustic control WFS has over the wave field within the loudspeaker array. Unfortunately these results are therefore only valid below the spatial aliasing frequency of the loudspeaker array used. Above the aliasing frequency no destructive interference can be used to compensate for reflections caused by the listening room. Preliminary results indicate that our approach works, but could still be improved. In our experiments we obtained a gain for different loca-

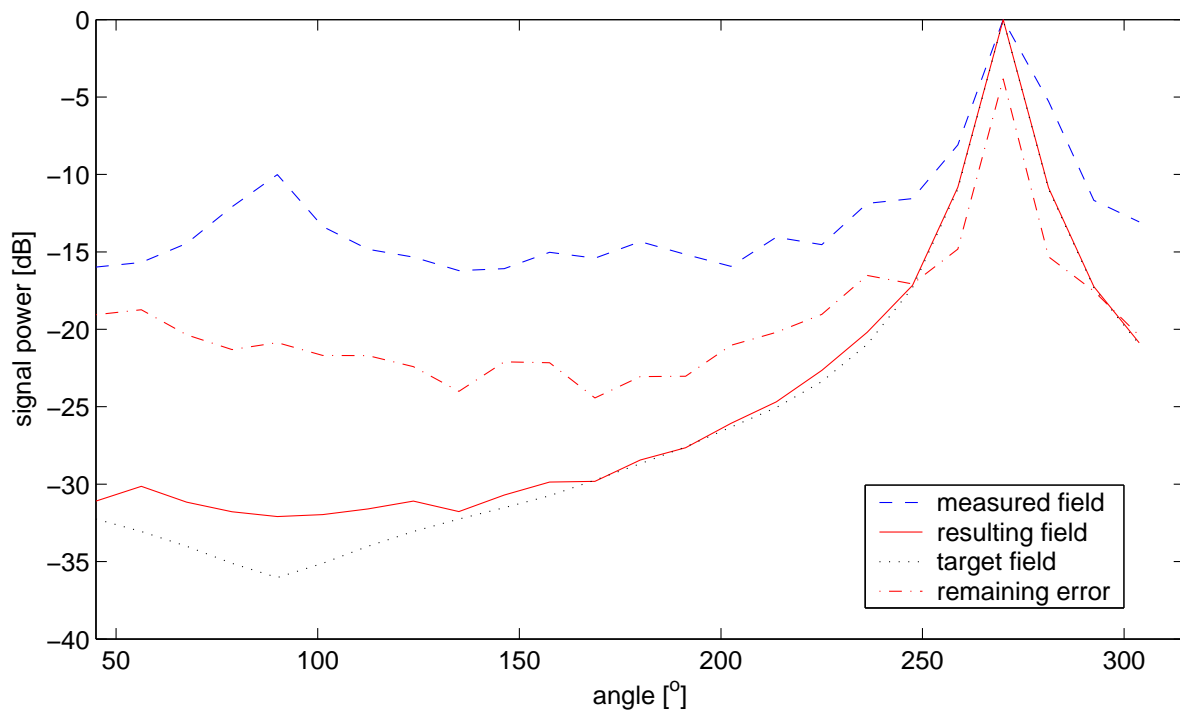


Figure 10: Results of room compensation shown as the signal power of the plane wave components.

tions which shows that we do not share the problems of common multi-point equalization systems. Further work includes further investigation of the effects caused by reflections outside the equalized plane and the combination of room compensation filters with loudspeaker compensation filters in the frequency range above the aliasing frequency.

REFERENCES

- [1] The CARROUSO project. <http://emt.iis.fhg.de/projects/carrouso>.
- [2] A.J. Berkhout. A holographic approach to acoustic control. *Journal of the Audio Engineering Society*, 36:977–995, December 1988.
- [3] E.W. Start. *Direct Sound Enhancement by Wave Field Synthesis*. PhD thesis, Delft University of Technology, 1997.
- [4] E.N.G. Verheijen. *Sound Reproduction by Wave Field Synthesis*. PhD thesis, Delft University of Technology, 1997.
- [5] P. Vogel. *Application of Wave Field Synthesis in Room Acoustics*. PhD thesis, Delft University of Technology, 1993.
- [6] D. de Vries, E.W. Start, and V.G. Valstar. The Wave Field Synthesis concept applied to sound reinforcement: Restrictions and solutions. In *96th AES Convention*, Amsterdam, Netherlands, February 1994. Audio Engineering Society (AES).
- [7] A.J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *Journal of the Acoustic Society of America*, 93(5):2764–2778, May 1993.
- [8] A. Torger. BruteFIR - an open-source general-purpose audio convolver. <http://www.ludd.luth.se/~torger/brutefir.html>.
- [9] J.-J. Sonke and D. de Vries. Generation of diffuse reverberation by plane wave synthesis. In *102nd AES Convention*, Munich, Germany, March 1997. Audio Engineering Society (AES).
- [10] E. Hulsebos and D. de Vries. Parameterization and reproduction of concert hall acoustics with a circular microphone array. In *112th AES Convention*, Munich, Germany, May 2002. Audio Engineering Society (AES).
- [11] R. Nicol and M. Emmerit. Reproducing 3D-sound for videoconferencing: A comparison between holophony and ambisonic. In *First COST-*

- G6 Workshop on Digital Audio Effects (DAFX98)*, Barcelona, Spain, Nov 1998.
- [12] P. Toft. *The Radon Transform. Theory and Implementation*. PhD thesis, Technical University of Denmark, 1996.
- [13] A.J. Berkhout. *Applied Seismic Wave Theory*. Elsevier, 1987.
- [14] 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).
- [15] L.D. Fielder. Practical limits for room equalization. In *111th AES Convention*, New York, NY, USA, September 2001. Audio Engineering Society (AES).
- [16] S.T. Neely and J.B. Allen. Invertibility of a room impulse response. *Journal of the Acoustical Society of America*, 66:165–169, July 1979.
- [17] M. Miyoshi and Y. Kaneda. Inverse filtering of room acoustics. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(2):145–152, February 1988.
- [18] J. N. Mourjopoulos. Digital equalization of room acoustics. *J. Audio Eng. Soc.*, 42(11):884–900, November 1994.
- [19] F. Talantzis and D.B. Ward. Multi-channel equalization in an acoustic reverberant environment: Establishment of robustness measures. In *Institute of Acoustics Spring Conference*, Salford, UK, March 2002.
- [20] O. Kirkeby, P. Nelson, H. Hamada, and Felipe Orduna-Bustamante. Fast deconvolution of multi-channel systems using regularization. *IEEE Transactions on Speech and Audio Processing*, 6(2):189–194, March 1998.
- [21] E. Coteel, U. Horbach, and R.S. Pellegrini. Multi-channel inverse filtering of multiexciter distributed mode loudspeakers for wave field synthesis. In *112th AES Convention*, Munich, Germany, May 2002. Audio Engineering Society (AES).
- [22] E. Coteel and R. Nicol. Listening room compensation for wave field synthesis. What can be done? In *23rd AES International Conference*, Copenhagen, Denmark, May 2003. Audio Engineering Society (AES).