

A NOVEL APPROACH TO ACTIVE LISTENING ROOM COMPENSATION FOR WAVE FIELD SYNTHESIS USING WAVE-DOMAIN ADAPTIVE FILTERING

S.Spors, H.Buchner and R.Rabenstein

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

ABSTRACT

Wave field synthesis is an auralization technique which allows to control the entire wave field within the entire listening area. However, reflections in the listening room interfere with the auralized wave field and may impair the spatial reproduction. Active listening room compensation aims at reducing these impairments by using the playback system. Due to the high number of playback channels used for wave field synthesis, the existing approaches to room compensation are not applicable. A novel approach to active room compensation overcomes these problems by a transformation from the space-time to the wave domain and application of wave-domain adaptive filtering.

1. INTRODUCTION

Spatial audio systems aim at reproducing the sound of fixed or moving sources which are distributed in space. The room where these sources are originally located is called the recording room. It may be a real room or a virtual one. The room where the sound sources are reproduced is called the listening room.

An ad hoc approach to the realization of a spatial audio system would be to place a number of microphones within the recording room and to transmit the recorded microphone signals to an identical arrangement of loudspeakers in the listening room [1]. More advanced methods decouple the setup used for the recording process from the one used for the playback process. Existing methods for this purpose include advanced panning methods, Ambisonics, and wave field synthesis [2].

However, all these methods assume an anechoic listening room. This assumption is only realistic if the listening room conforms to certain acoustic requirements. Otherwise, the reflections at the walls of the listening room impair the spatial reproduction. In these cases, a compensation of listening room reflections is required. Ideally, this compensation applies not only to certain selected positions, but to a large area within the listening room. Two different approaches for room compensation exist: passive and active methods. Passive compensation methods include mainly acoustic damping of the listening room. They do not work well for low frequencies. Active listening room compensation uses active methods to compensate for the listening room reflections. For practical reasons, an existing multichannel playback system should be utilized to avoid additional expense.

Active room compensation requires the input from reference microphones within the listening room to derive the compensation signals for the playback loudspeakers by an adaptive multichannel

system.

Unfortunately the traditional approaches to room compensation have severe drawbacks: In most cases it is only possible to compensate for the room reflections at the measured points (microphone positions). Outside these compensated points the result after applying room compensation is often worse than without [3]. The main reasons for this failure is that typical playback systems, like the 5-channel surround system, have very limited control over the wave field inside the desired listening area. This is a result of the spatial sampling performed by the loudspeakers. On the other hand the few equalization points used in typical compensation systems provide no detailed information on the wave field inside the listening area.

The problems of the traditional approaches could be overcome with a system that is able to

1. control the wave field inside the entire listening area,
2. analyze the wave field inside the entire listening area.

This paper proposes a novel approach to listening room compensation using wave field synthesis (WFS), wave field analysis (WFA) and wave-domain adaptive filtering (WDAF) as a solution to these requirements.

2. WAVE FIELD SYNTHESIS

The theory of WFS has been initially developed at the Technical University of Delft over the past decade [1]. In contrast to other multi-channel approaches, it is based on fundamental acoustic principles. WFS allows a physically correct reproduction of wave fields in contrary to the traditional multichannel systems, like 5-channel surround. This section gives a short overview of the theory.

WFS is based on the Huygens' principle. Its mathematical foundation is given by the Kirchhoff-Helmholtz integral, which can be derived by using the wave equation and the Green's integral theorem [4]. 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. 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 [1]. The Rayleigh I integral states that a pressure

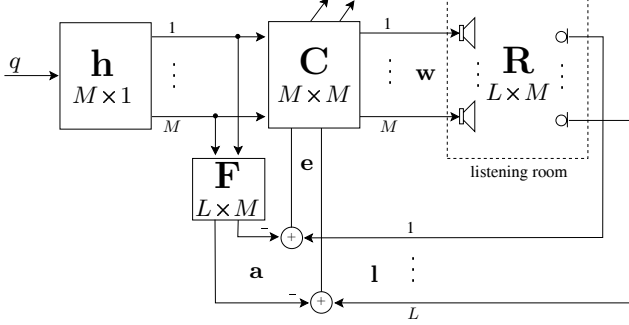


Fig. 1. Block diagram of traditional approaches to room compensation

field may be synthesized by means of a monopole distribution on a line. 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. The vector of loudspeaker driving signals $\mathbf{w}[k] = [w_1, \dots, w_M]^T$ can be derived by a multichannel convolution of measured or synthesized impulse responses with a source signal $q[k]$.

$$\mathbf{w}[k] = \mathbf{h}[k] * q[k], \quad (1)$$

where the symbol $*$ denotes the time-domain convolution operator. These impulse responses can be derived either from measurements or from suitable source models as shown in [2]. 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 [1].

The fact that loudspeakers can only be mounted at discrete positions results in spatial aliasing due to spatial sampling. The aliasing frequency is given in [1]. Assuming a typical 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 seems not to be very sensitive to these aliasing artifacts.

3. CONVENTIONAL APPROACH TO ROOM COMPENSATION

Typical listening rooms can be modeled as a multidimensional linear system [4, 5]. The basic idea of active listening room compensation for WFS is to utilize the playback system to compensate for the listening room reflections. A common approach is to pre-filter the loudspeaker driving signals with suitable compensation filters. Most of the existing room compensation methods can be described as illustrated in Figure 1. The M loudspeaker driving signals are derived by multichannel convolution with the WFS operator \mathbf{h} . They are then pre-filtered by a matrix of compensation filters \mathbf{C} resulting in the filtered loudspeaker signals \mathbf{w} . These are then played back through the loudspeakers in the listening room. The influence of the listening room can be characterized by impulse responses from the M speaker positions to L microphone positions. These measured impulse responses are stacked into the matrix \mathbf{R} . The actual sound field \mathbf{l} at the microphone positions is then compared to the desired free-field propagation \mathbf{a} from each loudspeaker to each microphone position. The difference between these two is

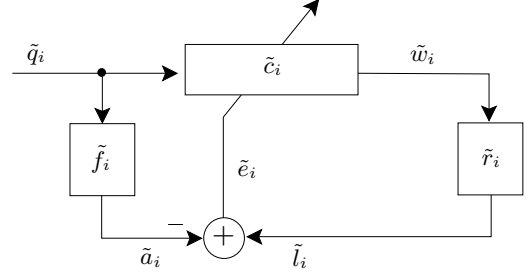


Fig. 3. Block diagram of the filter adaption process for one spatial component

denoted as \mathbf{e} , the matrix of free-field impulse responses as \mathbf{F} . The compensation filters \mathbf{C} are derived by minimizing the error \mathbf{e} . The adaption of the compensation filters can be seen as multidimensional inverse modeling problem [6] of the room response matrix \mathbf{R} . The multiple-input/output inverse theorem (MINT) [5] allows to calculate an exact inverse of the room transfer matrix in most practical cases. A typical WFS system consists of tens to hundreds of loudspeakers. A room compensation system in this structure would have to adapt a very high number of compensation filters, which is a very complex task. In order to overcome this problem we will introduce our proposed framework to efficient wave-domain adaptive filtering for room compensation.

4. A NOVEL APPROACH TO ROOM COMPENSATION

Figure 2 shows a block diagram of our proposed approach. The basic idea is to introduce a set of spatial transformations \mathcal{T}_1 to \mathcal{T}_3 that orthogonalize the listening room response \mathbf{R} . As a consequence, the matrix of compensation filters \mathbf{C} is decomposed into a set of compensation filters \tilde{c}_i , each acting on only one spatial signal component. Transformation \mathcal{T}_1 transforms a virtual source q into its transformation \tilde{q} . After filtering with the compensation filters \tilde{c}_i , transformation \mathcal{T}_2 transforms the signals back to the speaker signals. Transformation \mathcal{T}_3 transforms the microphone signals into the same domain as transformation \mathcal{T}_1 . The transfer functions $\tilde{\mathbf{F}}$ model free-field propagation from each transformed component for the given loudspeaker setup to the given microphone positions. Transformation \mathcal{T}_2 , the room transfer matrix \mathbf{R} and transformation \mathcal{T}_3 can be combined into the diagonal room transfer matrix $\tilde{\mathbf{R}}$ in the transformed domain. The adaption of the compensation filters is then performed efficiently in this transformed domain as illustrated in Figure 3. The error between the desired wave field $\tilde{\mathbf{a}}$ (free-field propagation) and the measured wave field $\tilde{\mathbf{l}}$ is given for the i -th component in the time domain as follows

$$\tilde{e}_i[k] = \tilde{r}_i[k] * \tilde{c}_i[k] * \tilde{q}_i[k] - \tilde{f}_i[k] * \tilde{q}_i[k] \quad (2)$$

where e.g. \tilde{e}_i denotes the i -th component of $\tilde{\mathbf{e}}$. Minimization of the error \tilde{e}_i yields the filters \tilde{c}_i . Applying the above presented approach transforms the multiple-input multiple-output (MIMO) inverse filtering problem into N one-dimensional inverse filters. Thus, adaptive room compensation becomes feasible for a high number of loudspeaker channels.

For an implementation a method is required to represent the wave field inside the desired listening area. This representation also defines the spatial transformations \mathcal{T}_1 , \mathcal{T}_2 and \mathcal{T}_3 . The next section introduces wave field analysis as a suitable solution.

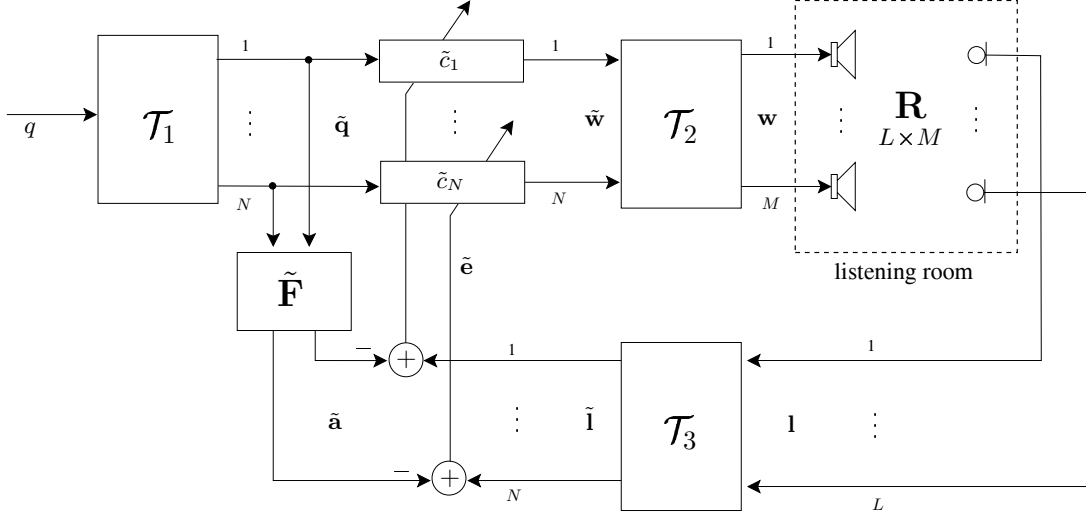


Fig. 2. Block diagram of proposed approach

4.1. Wave Field Analysis

Active listening room compensation as introduced in the previous sections requires to gain knowledge about the listening room influence in the entire listening area. The following section introduces the wave field analysis (WFA) techniques used in the context of WFS and room compensation. These techniques are mainly derived from seismic wave theory [4].

A first approach to derive the desired decoupling would be to analyze the entire two-dimensional listening area with microphones. Performing a two-dimensional spatial Fourier transform on the pressure field $P(\mathbf{r}, \omega)$

$$\tilde{p}(\mathbf{k}, t) = \mathcal{F}_{\mathbf{r}}\{p(\mathbf{r}, t)\} \quad (3)$$

would then provide the desired orthogonal signal components. The same principle applies for frequency-domain adaptive algorithms [7]. In general, there will be no access to the entire two-dimensional pressure field $p(\mathbf{r}, t)$ for analysis. However, the Kirchhoff-Helmholtz integral allows to calculate the acoustic pressure field $p(\mathbf{r}, t)$ from the sound pressure and its gradient on a line enclosing the desired field and vice versa [4]. Measurements of these quantities on a closed contour are sufficient to describe the entire area.

The next step is to transform the pressure field $p(\mathbf{r}, t)$ into plane waves with the incident angle θ and the temporal offset τ with respect to an arbitrary reference point. The recorded data $p(\mathbf{r}, t)$ become independent from the geometry used for recording. This transformation is called *plane wave decomposition* [8, 4] and is given by

$$\tilde{p}(\theta, \tau) = \mathcal{P}\{p(\mathbf{r}, t)\}, \quad (4)$$

where $\tilde{p}(\theta, \tau)$ denotes the plane wave decomposed wave field and \mathcal{P} the plane wave decomposition operator. This transformation is also well known as the *Radon transformation* from image processing. The Radon transformation maps straight lines in the image domain into Dirac peaks in the Radon domain. It is therefore typically used for edge detection in digital image processing. The same principle applies for acoustic fields: A plane wave can be understood as an 'edge' in the pressure field $p(\mathbf{r}, t)$. It can be shown

that the complex amplitudes and phases of the two-dimensional Fourier transform (3) represent the phases and amplitudes of plane waves [4]. Applying a Fourier transformation on θ and τ

$$\tilde{P}(k_{\theta}, \omega) = \mathcal{F}_{\theta\tau}\{\tilde{p}(\theta, \tau)\} \quad (5)$$

provides orthogonal signal components as desired.

One of the benefits of using a spatial transform of the wave field is, that plane waves can be extrapolated easily to other positions [4]

$$\begin{aligned} P(r, \theta, \omega) &= \mathcal{P}^{-1}\{\tilde{P}(\theta, \omega)\} = \\ &= \int_0^{2\pi} \tilde{P}(\theta', \omega) e^{-j\beta r \cos(\theta - \theta')} d\theta', \end{aligned} \quad (6)$$

where r and θ denote the position in cylindrical coordinates with respect to the origin of the plane wave decomposition and $\beta = \omega/c$ the wave number. This allows, in principle, to extrapolate a recorded wave field to arbitrary points without loss of information. Wave field extrapolation can be used to extrapolate a measured field to the loudspeaker positions for reproduction purposes or to create a complete image of the captured sound field. As for WFS spatial sampling of the microphone positions results in spatial aliasing.

4.2. Application to room compensation

As stated before, WFS allows to control the wave field inside the listening area while WFA allows to analyze this wave field. Because of the quite high number of channels used for typical WFS and WFA systems (up to several hundred) the application of traditional room compensation methods is not applicable. Our proposed framework allows an efficient realization of room compensation in the context of WFS. Unfortunately this approach is limited by the aliasing frequency but it can be complemented with passive methods for higher frequencies.

In the following we will specify the transformations \mathcal{T}_1 to \mathcal{T}_3 . The transformation \mathcal{T}_1 transforms the virtual source q into the frequency components of the plane wave domain

$$\mathcal{T}_1\{q\} = \mathcal{F}_{\theta}\{\mathcal{P}\{q\}\} \quad (7)$$

After filtering with the compensation filters the transformation \mathcal{T}_2 extrapolates the filtered loudspeaker signals in the transformed domain back to the loudspeaker positions

$$\mathcal{T}_2\{\tilde{\mathbf{w}}\} = \mathcal{P}^{-1}\{\mathcal{F}_\theta^{-1}\{\tilde{\mathbf{w}}\}\}. \quad (8)$$

Equation (6) can be used for this purpose. The transformation \mathcal{T}_3 is a similar transformation as \mathcal{T}_1 corresponding to the microphone array geometry. Implementations of the plane wave decomposition for different microphone array geometries can be found in [8].

5. RESULTS

The performance of the proposed approach is shown by numerical simulation of a WFS system using two-dimensional models. The simulated setup consists of a circular loudspeaker array with radius $R_{LS} = 1$ m consisting of 32 loudspeakers. A circular microphone array with 32 microphones and a radius $R_{mic} = 0.98$ m is located concentric inside the loudspeaker array. It consists of pressure and pressure gradient microphones. The plane wave decomposition for the circular microphone array was calculated as shown in [8].

The aliasing frequency of the setup is $f_{al} \approx 900$ Hz. All signals were downsampled accordingly. The wall at one side of the array at $\theta = 270^\circ$ is assumed to be reflective, the other walls of the room are ideally damped. The reflections of the wall were calculated using the mirror-image method. The input signal q is a music signal that is played back as virtual point source located at 4m distance to the array centers at an angle of $\theta = 90^\circ$. The compensation filters were calculated using the frequency domain adaptive filtering algorithm described in [7]. This rather simple setup allows to monitor the performance of room compensation. Ideally the reflection at $\theta = 270^\circ$ should be eliminated by the compensation filters. Figure 4 shows the mean squared error of all components

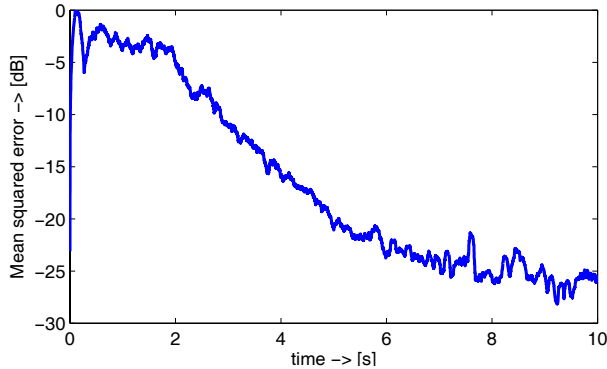


Fig. 4. Mean squared error between the desired wave field and the compensated wave field.

in $\tilde{\mathbf{e}}$ over the time axis. This equals a omni-directional pickup of the error signal. The adaption capabilities of our approach can be seen clearly.

We also calculated the impulse responses of the target field, the measured and the compensated field in the plane wave domain after the adaption has converged. Figure 5 shows the signal energies of these impulse responses. The desired maximum energy at $\theta = 90^\circ$, resulting from the virtual source, as well as the reflection from $\theta = 270^\circ$ can be seen clearly. The energy of the compensated wave field is nearly identical to the desired wave field. This illustrates the successful application of our proposed method.

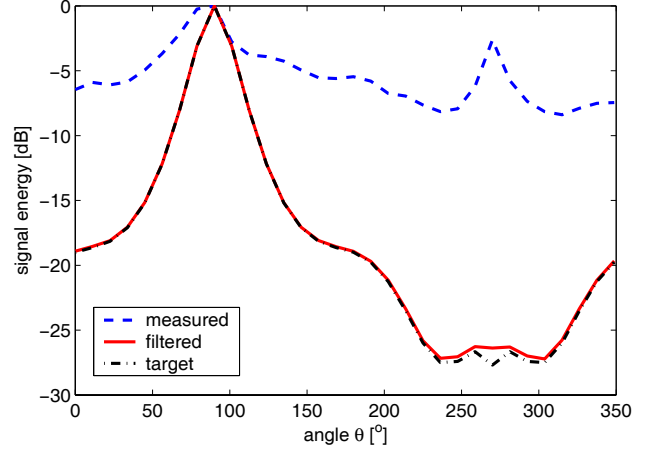


Fig. 5. Results of room compensation shown as the signal energy of the plane wave components.

6. CONCLUSION

A novel framework for the listening room compensation problem in multichannel audio reproduction has been presented. It is based on wave-domain adaptive filtering. The application to wave field synthesis has been shown. First results prove the applicability of our approach. The presented framework is also suitable for the solution of other multichannel audio problems like echo cancellation and adaptive beamforming.

7. REFERENCES

- [1] 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.
- [2] S.Spors, H.Teutsch, A.Kuntz, and R.Rabenstein, "Sound field synthesis," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Y.Huang and J.Benesty, Eds. Springer, 2004.
- [3] L.D. Fielder, "Practical limits for room equalization," in *111th AES Convention*, New York, NY, USA, September 2001, Audio Engineering Society (AES).
- [4] A.J. Berkhout, *Applied Seismic Wave Theory*, Elsevier, 1987.
- [5] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 2, pp. 145–152, February 1988.
- [6] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1996.
- [7] H.Buchner, J.Benesty, and W.Kellermann, "Multichannel frequency-domain adaptive algorithms with application to acoustic echo cancellation," in *Adaptive signal processing: Application to real-world problems*, J. Benesty and Y. Huang, Eds. Springer, 2003.
- [8] 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).