Audio Engineering Society
Convention Paper
Presented at the 116th Convention
2004 May 8–11 Berlin, Germany

Efficient Active Listening Room Compensation for Wave Field Synthesis

Sascha Spors, Herbert Buchner, and Rudolf Rabenstein [1]

[1] *Telecommunications Laboratory, University of Erlangen-Nuremberg, Germany*

Correspondence should be addressed to Sascha Spors (`spors@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

Wave field synthesis (WFS) is an established method for spatial sound reproduction in large spaces. It employs arrays of loudspeakers with some ten to several hundred independent channels. WFS has the capability the generate the virtual spatial impression of a possibly large recording room inside a given listening room. This spatial impression is stable within the whole area enclosed by the loudspeaker arrays, the so-called listening area. The lis-

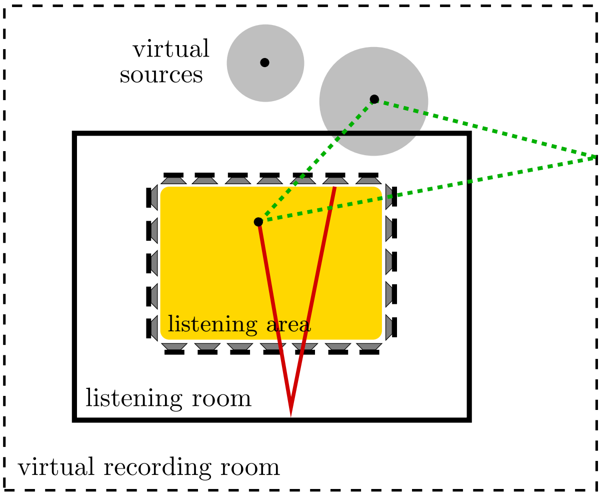


Fig. 1: Typical setup for WFS reproduction. Green dashed line: reflections in the virtual recording room. Red solid line: reflections in the listening room.

teners are allowed to move freely inside the listening area and are not restricted in their activities by headphones, head tracking systems, or alike. They perceive virtual sound sources which may be located inside or outside the listening area. Figure 1 shows a typical setup for WFS reproduction.

The driving signals for the loudspeakers are generated from the source signals, from geometrical information on the source locations, and from information on the room acoustics of the recording room. This information may have been recorded in an existing recording room (e.g. a concert hall) or it may have been artificially created to render a virtual acoustic scene. The signal processing algorithms for the generation of the loudspeaker signals are derived from fundamental acoustic principles. They describe the propagation of acoustic waves in the recording room and project it into the listening area.

This projection of the recording room to the listening room correctly describes reflections in the recording room (green dashed line in Figure 1) which create the desired spatial impression. However, typical systems cannot take into account any reflections within the listening room itself (red solid line in Fig. 1) which impair the effects of the WFS system. Of course, listening room reflections are diminished by acoustic insulation materials (so called passive room compensation). But the effect of this countermeasure is limited by cost and room design considerations. Especially for low frequencies, passive room compensation does not provide effective attenuation of listening room reflections.

Active room compensation methods, on the other hand, require extensive use of microphones to record the acoustic situation within the listening room, and of control equipment, amplifiers, and loudspeakers to produce acoustical waves which counteract the undesired listening room reflections. However, most of this equipment is necessary for WFS reproduction anyway. By clever design of an active compensation system, the WFS installation can be used for room compensation as well. So, apparently, WFS is the predetermined spatial reproduction system for the combination with active listening room compensation.

This paper describes an efficient approach to active listening room compensation for WFS. Its system requirements and existing design approaches are discussed in Section 2. Section 3 describes the theoretical foundations of wave field synthesis and analysis. Section 4 presents the proposed approach to active listening room compensation. Simulation results based on room measurements are shown in Section 5.

2. SYSTEM REQUIREMENTS AND DESIGN APPROACHES

The following requirements for an ideal room compensation system can be deduced from the description of the problem given above:

An ideal room compensation system should be capable to

1. (perfectly) analyze the wave field inside the listening area and
2. (perfectly) control the wave field inside the listening area.

The first requirement ensures that all necessary information about the acoustic properties of the listening room are available. The second requirement ensures that any undesired reflections can be eliminated.

Unfortunately, it is never possible to fully meet both requirements. Existing active room compensation

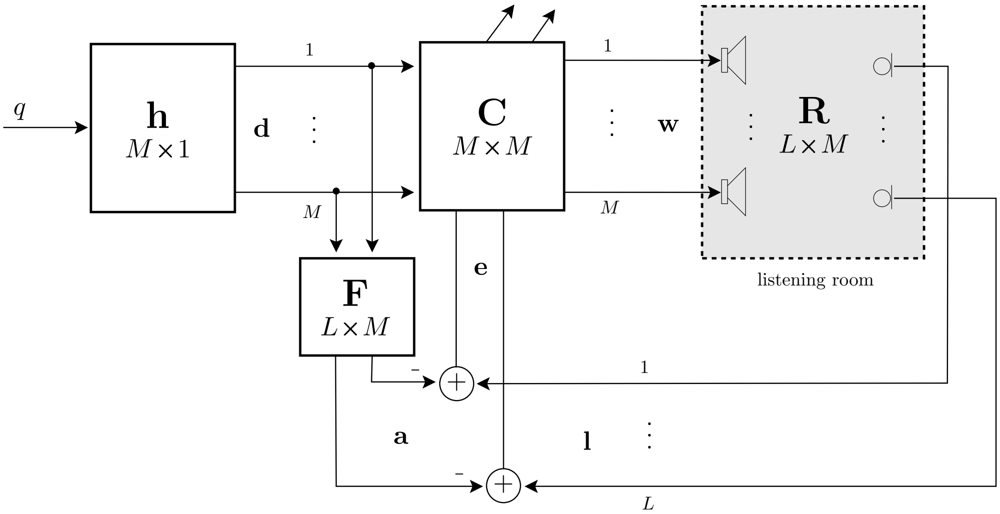


Fig. 2: Block diagram of traditional multi-point approaches to room compensation

methods differ in their approaches to meet these requirements at least partially. The basic approach of these methods and some refinements are discussed here briefly. Fig. 2 shows a block diagram of traditional approaches to active room compensation. For simplicity only a single virtual source signal q is considered. The vector of M spatial impulse responses $\mathbf{h}$ computes from q a set of M uncompensated loudspeaker signals $\mathbf{d}$. The matrix $\mathbf{C}$ of compensation filters turns $\mathbf{d}$ into M compensated loudspeaker signals $\mathbf{w}$. The loudspeakers generate a sound field in the listening room which is subject to interference with unavoidable reflections. Ideally, this sound field corresponds to the desired sound field for proper spatial rendering.

To control the compensation filters $\mathbf{C}$ accordingly, the sound field in the listening room is recorded by L microphones. The room impulse responses from the loudspeakers to the microphones are collected in the room transfer matrix $\mathbf{R}$. The microphone signals $\mathbf{l}$ are compared with the free-field propagation signals $\mathbf{a}$. They are obtained from the uncompensated signals $\mathbf{d}$ after filtering with the entries of the free-field propagation matrix $\mathbf{F}$. Theses filters can be computed from the acoustical wave equation. Finally, the difference $\mathbf{e}$ between the computed free-field propagation signals $\mathbf{a}$ and the recorded microphone signals $\mathbf{l}$ controls the adaptive compensation filters $\mathbf{C}$. Different choices of the spatial impulse responses $\mathbf{h}$, the adaptive filtering method for updating the compensation filters $\mathbf{C}$ and the setup of the loudspeakers and microphones in the listening room determine different compensation methods [1].

However, these traditional approaches have a number of drawbacks: (1) many loudspeakers and microphones are required for a satisfactory analysis and control of the wave field, (2) consequently the number of compensation filters M^2 is quite large, (3) the loudspeaker and microphone signals are correlated, such that the adaptation of $\mathbf{C}$ is poor [2].

Since the error signal $\mathbf{e}$ is derived from the recordings $\mathbf{l}$ at many points inside the listening room, these approaches are called *multi-point methods*. They are, in most cases, only capable to provide room compensation at the measured points. The results between the compensated points are often worse than without applying room compensation.

To overcome the first drawback, the multi-point approach has to be given up in favor of a method which analyzes the wave field within the whole listening area and not just at selected points. The mathematical tool for this purpose is the Kirchhoff-Helmholtz integral [3]. It allows to express the sound field within an enclosure by the the sound pressure and the particle velocity on the boundary of the enclosure. As a practical consequence, it is sufficient to record the sound field at the boundary of the listening area rather than at densely spaced points within

the room. The error signal is now expressed in terms of a plane wave decomposition (see Section 3.2.1) of the wave field, rather than in terms of multi-point measurements.
However, some problems still remain: The number of compensation filters is proportional to the squared number of loudspeakers (M^2) and adaption is slow due to correlation between the channels. The number of compensation filters can be reduced to be proportional to M by combining the compensation filters $\mathbf{C}$ with the spatial impulse responses $\mathbf{h}$. Since the spatial impulse responses depend on the source position, this method is only feasible for static sources [4].
Another approach to reduce the number of compensation filters is to apply a Fourier transformation to the plane waves. Then the error signal is expressed in the wave domain and correspondingly, the compensation filters are approximately decoupled [5]. However, while wave domain filtering is successful in decoupling the compensation filters, it does not guarantee decorrelation of the channels and thus fast convergence of the adaptation.
To explore the trade-off between the minimal number of compensation filters and fast convergence, this contribution analyzes the decoupling process more thoroughly. It is shown, how faster convergence of the compensation filters can be achieved at the expense of using slightly more than the their minimal number.

3. THEORY

In the following we will introduce the necessary theoretical background for our proposed approach. The next section will introduce WFS as a suitable reproduction technique for active room compensation. A section on listening room analysis follows.

3.1. Wave Field Synthesis

The theory of wave field synthesis (WFS) has been initially developed at the Technical University of Delft over the past decade [6, 7, 8, 9, 10]. In contrast to other multi-channel approaches, it is based on fundamental acoustic principles. This section gives a broad overview of the theory.

The theoretical basis of WFS is given by Huygens' principle as illustrated in Figure 3. The wave field of the primary source can also be generated by a number of point sources with properly weighted and delayed acoustic signals. These point sources are in-

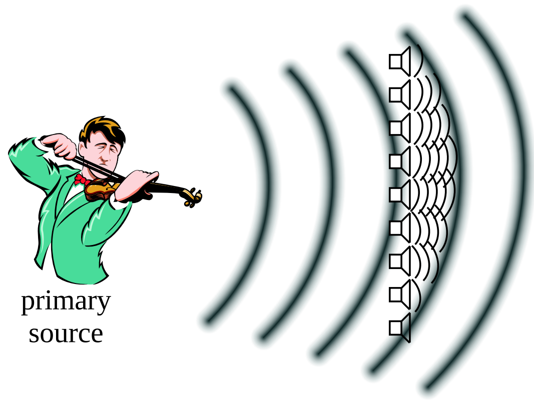


Fig. 3: Basic principle of wave field synthesis.

dicated by loudspeaker symbols in Figure 3 and are also realized by small loudspeakers in reality. The mathematical foundation of WFS is described by the Kirchhoff-Helmholtz integral [3].
Straightforward implementation of this principle would require to place acoustic sources on the entire surface surrounding the listening volume. As this requires an unfeasible high number of reproduction channels, a practical system is limited to a listening area instead of an listening volume. A typical WFS system is realized by mounting closed loudspeakers on discrete positions of an arbitrary shaped array surrounding the listening area leveled with the listeners ears.
WFS is also capable to reproduce sources that lie inside the listening area [8]. Thus, acoustical sources can be placed between the listener and the loudspeakers within the reproduction area (focused sources). This is not possible with most traditional reproduction systems.
The discretization of the underlying physical and mathematical relations results in spatial aliasing due to spatial sampling. The minimal cut-off frequency is given by [10]

$$f_{\mathrm{al}} = \frac{c}{2\Delta x}, \tag{1}$$

where Δx denotes the loudspeaker distance. Assuming a loudspeaker spacing $\Delta x = 19$ cm, the minimum spatial aliasing frequency is $f_{\mathrm{al}} \approx 900$ Hz. Fortunately, the human auditory seems not to be very sensitive to the artifacts introduced by spatial aliasing in wide-band reproduction.

However, active room compensation requires full control over the reproduced wave field and thus is only applicable below the spatial aliasing frequency of the WFS system. The restriction to two-dimensional reproduction for practical WFS systems poses additional limitations to active room compensation: (1) it is only possible to control the wave field in the plane where the loudspeakers are located and (2) system inherent amplitude errors.
The first effect allows only to compensate for reflections that origin from the plane where the reproduction is performed in. Thus, it is not possible to compensate reflections originating from the ceiling or the floor of the listening room. The second effect limits the achievable amount of room compensation.

3.2. Listening Room Analysis

As stated in Section 2 successful application of active room compensation requires the analysis of the listening room influence on the reproduced wave field inside the listening area. The following section will introduce the tools required for this task in our context.

3.2.1. Plane wave decomposition

Let us first assume that we have access to the two-dimensional wave field $P(x, y, \omega)$ inside the entire listening area. Discarding the effects of spatial sampling the wave field could be recorded by placing microphones inside the entire listening area. We want to decompose the wave field into its plane wave components with respect to an arbitrary chosen origin. The transformation into plane waves for arbitrary wave fields is given as follows [11]

$$\bar{P}(\theta, \omega) = \mathcal{P}\{\, P(r, \alpha, \omega)\,\} = \\ = \int_0^{\infty} \int_0^{2\pi} P(r, \alpha, \omega) e^{jkr\cos(\theta-\alpha)} r \, d\alpha \, dr \quad (2)$$

where r and α denote the position in cylindrical coordinates with respect to the origin, k and θ the wave number and incidence angle of the plane waves. This transformation is termed as *plane wave decomposition* and is well known from seismic imaging. One benefit of this approach is that this signal representation is independent of the particular geometry used for analysis. Plane wave impulse responses efficiently describe the acoustics of the captured area. The inverse transformation to the plane wave decomposition is given as follows

$$P(r, \alpha, \omega) = \mathcal{P}^{-1}\{\, \bar{P}(\theta, \omega)\,\} = \\ = \int_0^{2\pi} \bar{P}(\theta', \omega) e^{-jkr\cos(\alpha-\theta')} d\theta' \quad (3)$$

This operation is often also termed as *plane wave extrapolation* or *wave field extrapolation* [12]. In principle, it is possible 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.
However, recording $P(x, y, \omega)$ for the entire listening area with microphones is not feasible in our context. On the one hand this would require a quite high number of microphones, on the other hand the microphones would occupy the listening positions. A solution to this problem is provided by the 2D Kirchoff-Helmholtz (KH) integral [12]. The KH integral states that at any listening point within a source-free area the sound pressure can be calculated if both the sound pressure and its gradient are known on the line surrounding the volume. However, as for WFS, the degeneration to two-dimensions has some major drawbacks: (1) amplitude errors and (2) interference of sources not located in the microphone array plane into the plane wave decomposed signals of this plane.
The first effect is caused by the different amplitude decays between a 2D line and a 3D point source. The second effect is related to the 2D geometry of the analysis array. As a consequence it is not possible to distinguish between reflections coming from boundaries at the analyzed plane and reflections caused by boundaries outside this plane. The fact that microphones can only be mounted at discrete positions can result in spatial aliasing due to spatial sampling. Thus principles already known from WFS apply also to listening room analysis.

3.2.2. Wave domain transformation

As the frequency domain is an efficient tool for the description of time domain signals, we will introduce an analogous spatial transformation of the plane wave decomposed signals. This transformation consists of a Fourier transformation of the plane wave decomposed signals with respect to the angle

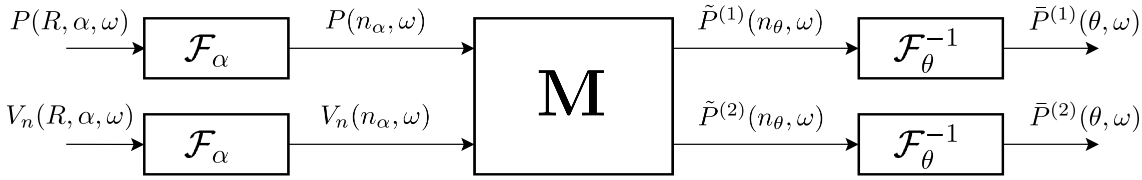


Fig. 4: Block diagram of the plane wave decomposition for a circular microphone array.

θ of the plane waves

$$\tilde{P}(n_\theta,\omega) = \mathcal{F}_\theta\{\ \bar{P}(\theta,\omega)\ \} \tag{4}$$

where the angular frequency is denoted as n_θ. This transformation will be denoted as *wave domain transformation* and the transformed signal representation as *wave domain* in the following.
The wave domain transformation has a link to a decomposition into cylindrical harmonics. The only difference between the components of these two transformations is the factor $j^{(1+n_\theta)}$ as shown in [13].

3.2.3. Plane wave decomposition for circular microphone arrays

There exist numerous algorithms to calculate the plane wave decomposition for various microphone array geometries. 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 calculation of the plane wave decomposition for a circular array given in [13] in the following section. Figure 4 shows a block diagram of the algorithm. The basic idea is the same as for filtering time signals with the fast convolution. This analogy applies because the plane wave decomposition from (2) contains a convolution of the wave field $P(r,\alpha,\omega)$ and the term $e^{jkr\cos\alpha}$ with respect to the angle α. This convolution can be realized more efficiently by operations in an angular frequency domain associated with the angles α and θ from (2).
The starting point are acoustic pressure P and velocity measurements V_n on a circle with radius R. The acoustic velocity is measured in radial direction. The microphone position on the circle is denoted by the angle α. The first step is to perform a Fourier transformation on the angle α of the microphone signals $P(R,\alpha,\omega)$ and $V_n(R,\alpha,\omega)$. The angular frequency is denoted as n_α. The incoming and outgoing wave domain signals are derived by a two-dimensional filtering operation $\mathbf{M}$. An inverse Fourier transformation with respect to the angular frequency n_θ derives then the incoming $\bar{P}^{(1)}$ and outgoing $\bar{P}^{(2)}$ parts of the plane wave decomposed wave field. The decomposition into incoming and outgoing plane waves can be used to distinguish between sources inside and outside the circular array. While sources outside result in an 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 $\bar{P}^{(1)}$ sources inside the array are omitted for room compensation purposes. In the following we will only consider the incoming part of the plane wave decomposed wave field.

4. ROOM COMPENSATION BASED ON WAVE DOMAIN ADAPTIVE FILTERING

In order to overcome the limitations of traditional room compensation algorithms in the context of massive multi-channel reproduction systems, we propose to introduce a set of linear spatial transformations. These transformations should decouple and decorrelate the room response of the loudspeaker system. Figure 5 shows a block diagram of the proposed approach. The signal representations in the transformed domain will be termed as wave domain signals in the following. Transformation $\mathcal{T}_1$ transforms the virtual source signal q into its spatial representation $\tilde{\mathbf{q}}$ in the wave domain. Simple source models (point source, plane wave) as well as complex wave fields can be used for the spatial characteristics of the source. The number of components N used for the signal representations in the wave domain is derived from the spatial aliasing frequency of the WFS system. The application of the room compensation filters $\tilde{\mathbf{C}}$ is then performed in the wave domain. Transformation $\mathcal{T}_2$ creates suitable loudspeaker signals from the wave field representation in the wave domain. It has to be chosen such that a given wave field is reproduced in the best

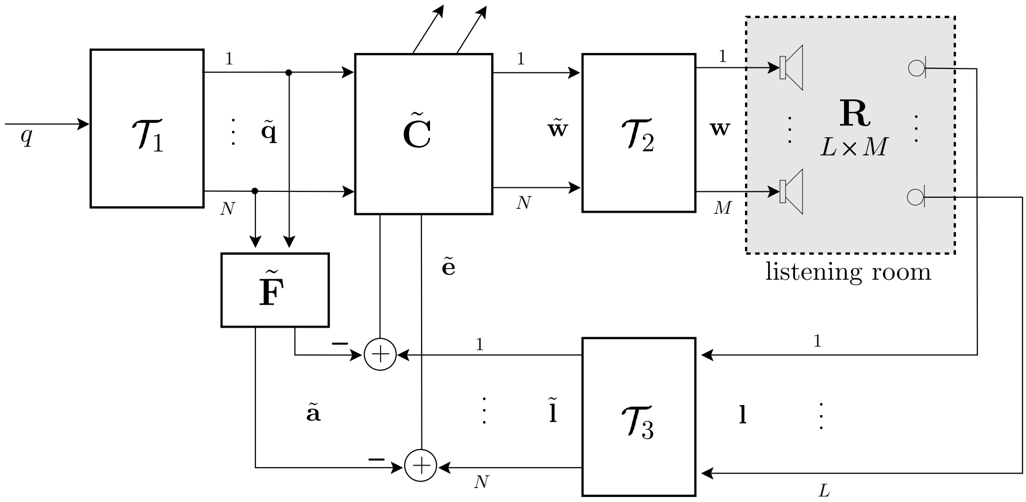


Fig. 5: Block diagram of active room compensation algorithm based on wave domain adaptive filtering.

possible way by a given loudspeaker array. Transformation $\mathcal{T}_3$ transforms the microphone signals into the wave domain. The entire adaption process of the room compensation filters is then performed in the wave domain. However, this approach relies on the assumption that we have a suitable transformation at hand that fulfills the requirements stated in Section 2. The following section will introduce the proposed choice for those transformations.

4.1. Decoupling of the room response

The optimal transformation for perfect decoupling of the room transfer matrix $\mathbf{R}$ would be a singular value decomposition (SVD) [14] for every discrete time step. As a result the filter matrix $\tilde{\mathbf{C}}$ is decomposed into N filters $\tilde{c}_i$ acting only on one wave domain signal. Figure 6 illustrates the filter configuration. Unfortunately, this optimal solution is dependent from the geometry of the problem and no efficient algorithm exists in general. Thus it is not realizable in a practical system. An efficient realizable transformation will have suboptimal performance in terms of decoupling. As a fully occupied filter matrix $\tilde{\mathbf{C}}$ is not realizable we need a transformation which packs the relevant spatial information of the room transfer matrix $\mathbf{R}$ into a few components. This should result in a sparsely occupied filter matrix $\tilde{\mathbf{C}}$. Since correlations between the filtered channels would result in a poor convergence of the adaptive filters, this suboptimal solution should additionally result in decorrelated channels for optimal adaption.

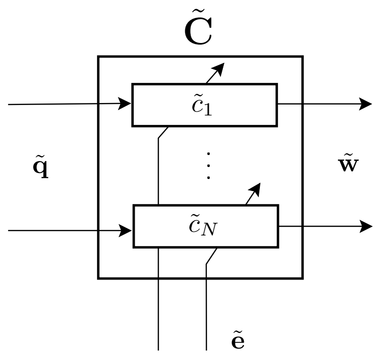
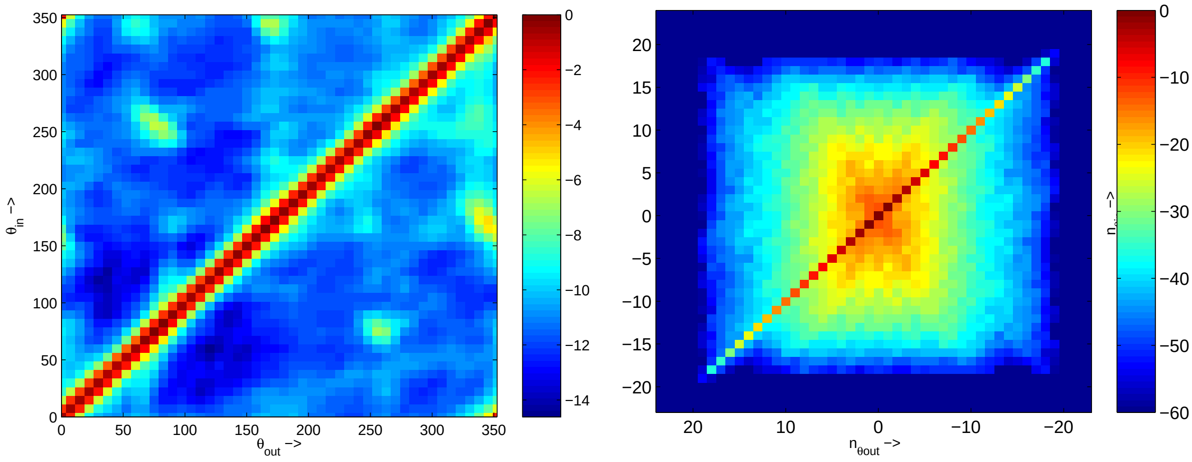


Fig. 6: Room compensation filters after perfect decoupling of the room response

The optimal transformation in this case is the Karhunen-Loève transformation (KLT) [14]. It performs optimally in the sense of decorrelation and energy compaction but does not, in general, result in a diagonalization of the room transfer matrix $\mathbf{R}$. Unfortunately, the KLT is also data dependent and therefore not applicable in our scenario. However, the KLT is well known from image coding and the problem of data dependency is solved in this application by using analytical transformations which have approximately the same performance as the KLT on a given dataset. For natural images the



(a) Plane wave decomposition

(b) Wave domain transformation

Fig. 7: Energy of the elements of the room transfer matrix for different signal representations for a typical listening room. The colors denote the energy in dB.

discrete cosine transform (DCT) has proven to be suitable choice [14]. As cylindrical harmonics are free-field solutions of the acoustic wave equation for a cylindrical coordinate system we propose to decompose the room response using the wave domain transformation. As stated in Section 3.2.2 the wave domain signals differ only by a simple factor from the cylindrical harmonics. The benefit of using this transformation is that it can be calculated directly from plane wave decomposed signals which can be derived for nearly all practical microphone array geometries.

To illustrate the applicability of the wave domain transformation in our framework we simulated and measured the room transfer matrix $\mathbf{R}$ for different geometries. The energy of the elements of the room transfer matrix was calculated for each signal representation (pressure signals, plane wave decomposed signals, wave domain signals, KLT) and each scenario, resulting in a two-dimensional matrix. However, the elements of $\mathbf{R}$ have different meanings for different signal representations. For pressure signals, the room transfer matrix is a $L \times M$ matrix describing the relations between the loudspeakers and the microphones at discrete positions. For plane wave decomposed signals, the room transfer matrix consists of the relations between incoming and outgoing plane waves at discrete angle θ. In the wave domain, the room transfer matrix describes the same relations with respect to the angular frequency n_θ. Although θ and n_θ are continuous variables, the room transfer matrix represents the spatial information only at discrete angles and discrete angular frequencies n_θ respectively. To unify the the terminology, the entries into the room transfer matrix are now called its *spatial components* irrespective of the specific signal representation at hand.

In the following we will show the properties of these transformations for one measured typical scenario also presented in Section 5 (see this section for a description of the measurement setup).

Figure 7(a) and 7(b) show the spatial energy distribution for the plane wave decomposed room response and the wave domain room response. It can be seen clearly that in the wave domain the energy is packed on the main diagonal (please note the different scales). The spatial signal components of the different representations were sorted by descending energy in order to analyze the energy compaction performance. Figure 8 shows the energy distribution

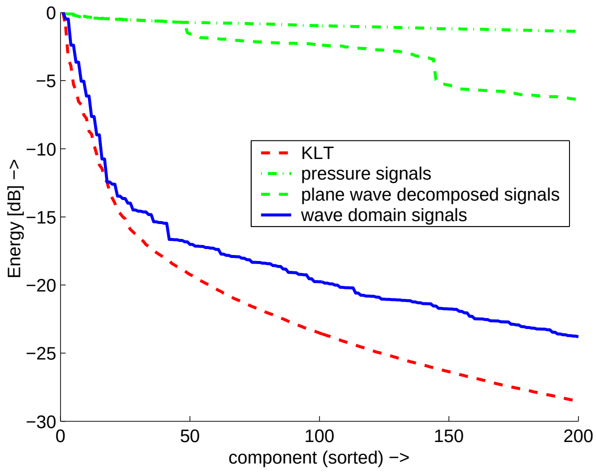


Fig. 8: Energy of the spatial signal components for a measured room response in different domains. The components are sorted by their energy in descending order.

of the different signal representations. As stated before the KLT performs best in the sense of energy compaction. However, the wave domain components perform only slightly worse. The plane wave decomposed signals and the pressure microphone signals show very poor energy compaction performance in this example. These results show that the wave domain decomposition is a suitable choice in our context.

4.2. Application to WFS

We will now introduce our novel approach based on wave domain adaptive filtering to the problem of room compensation for a WFS system. The basic idea, as outlined in the previous section, is to use the Fourier transformed incoming plane wave components to perform the adaptive filtering (see Figure 5). We will specify the transformations used in the context of WFS now.

Closed-form solutions for typical acoustic sources, e.g. point sources and plane waves, can be used to calculate transformation $\mathcal{T}_1$ analytically [11]. This allows an efficient implementation of transformation $\mathcal{T}_1$ by a plane wave decomposition followed by a Fourier transformation with respect to θ. Transformation $\mathcal{T}_2$ make use of plane wave extrapolation (3) and an inverse Fourier transformation. The loudspeaker signals can be calculated by extrapolation of the wave field to the loudspeaker positions according to the underlying principle of WFS (Huygens' principle). However, not all plane wave components can be used for a particular loudspeaker. Only those plane wave components which fit to the main emission axis of the particular loudspeaker are taken into concern. This requires only an appropriate modification of the integration limits in Equation (3). Transformation $\mathcal{T}_3$ is implemented using the algorithm outlined in Section 3.2.3.

The free-field transfer matrix $\tilde{\mathbf{F}}$ is calculated assuming free-field propagation from each loudspeaker to each microphone. It has only contributions on the main diagonal. In general N filters operating on each of the N transformed signal components will not deliver the best possible result. This results from imperfect decoupling as stated in the previous section. However, this can be taken into consideration by using not a diagonal but a sparsely occupied filter matrix $\tilde{\mathbf{C}}$. The additional filter elements are chosen according to the energy in the spatial components of the room transfer matrix $\tilde{\mathbf{R}}$. However, the room transfer matrix has to be pre-measured by a suitable measurement procedure. This procedure relies on the assumption that the energy distribution does not change dramatically for reasonable small changes in the listening room acoustics. Mirror image based room simulations performed by the authors have indicated that this assumption holds in typical environments. A large change in listening room acoustics could e.g. be caused by a broad variation in the number of listeners. A solution to this problem could be to measure a set of room transfer matrices for typical scenarios and use the most appropriate one during the adaption of the compensation filters.

5. RESULTS

The following section presents application results of our proposed algorithm. The results were simulated on the basis of measured data. For this purpose the listening room transfer matrix $\tilde{\mathbf{R}}$ of a real WFS system was measured in a typical listening room. The following section will introduce the measurement setup used for the experiments.

5.1. Measurement Setup

All experiments were carried out in our demonstration room with the size 5.8×5.9×3.1 meters (Volume about 105 m^3). The walls of the room are covered

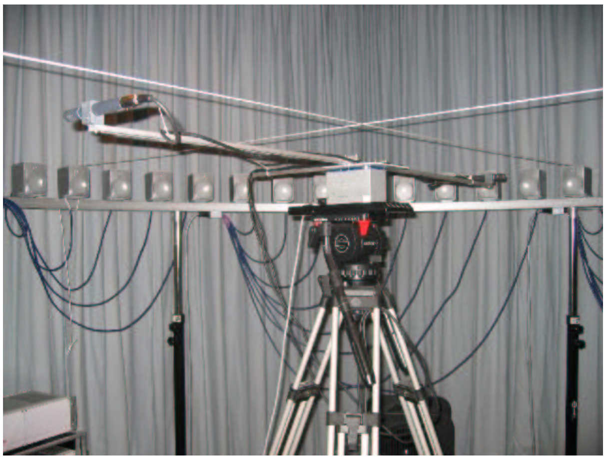


Fig. 9: Virtual microphone array used for the experiments.

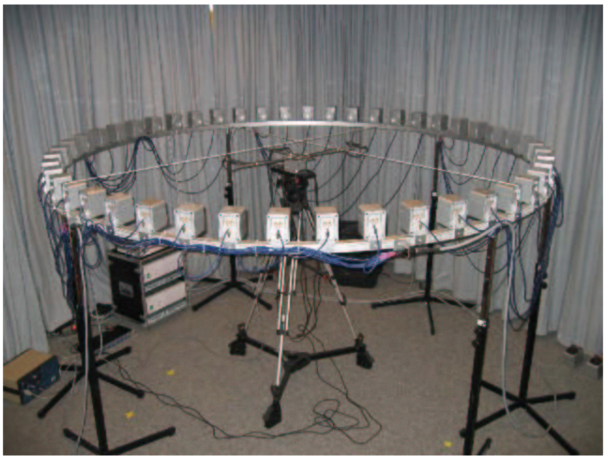


Fig. 10: Loudspeaker and microphone setup used for the experiments.

by a removable acoustic absorbent curtain that allows to create different acoustic scenarios. All curtains where removed for the results presented here, resulting in $T_{60} \approx 350$ ms. Similar experiments with partially closed curtains revealed the same performance. We used our laboratory circular WFS array for the measurements. It consists of 48 equidistant placed high-quality loudspeakers. The radius of the loudspeaker array was $R_{LS} = 1.50$ m. A circular microphone array located concentric inside the loudspeaker array was used to capture the room acoustics. The radius of the array was $R_{MIC} = 1.00$ m. The circular array is realized by sequential measurement of the discrete positions on the circle. We use a stepper motor for this purpose as shown in Figure 9. The pressure and pressure gradient microphones are mounted on the opposite sides of a rod. Using this setup for the microphone array has several benefits: (1) the geometry can be changed easily, (2) we only need two high-quality microphones and (3) the number of channels that has to be captured is limited. However, this setup is not suitable for real-time implementation of room compensation. We measured 48 microphone positions for the experiments. Figure 10 shows an overview of the loudspeaker and microphone array. The curtains in the room are closed here.

The impulse responses from each loudspeaker to each microphone position where measured resulting in the room transfer matrix $\mathbf{R}$. The measurements where downsampled according to the spatial aliasing frequency (see Eq. (1)) of the loudspeaker array. The entire algorithm was implemented as shown in Figure 5. According to the sampling theorem we used 48 wave domain components. The compensation filters where calculated adaptively using the frequency domain adaptive filtering algorithm described in [15]. The filter adaption is thus performed in the spatial and temporal frequency domain. We used two filter configurations for the presented results: (1) one filter for each wave domain component (48 filters in total) as shown in Figure 6 and (2) the first configuration plus 30 additional filters operating between different wave domain components. The additional filters where chosen accordingly to the energy distribution in the wave domain transformed room transfer matrix $\tilde{\mathbf{R}}$ as shown in Figure 7(b).

5.2. Results

The performance of our proposed algorithm is now demonstrated by simulations of a moving source using the described measurements. The simulation setup involves one point source located at a distance of 5 meters from the WFS array center. The source is moving with constant distance and varying angle. The source changed its position every 10 seconds from $\alpha = 0^o$ to $\alpha = 30^o$ in 10 degree steps. We used the error signal $\tilde{\mathbf{e}}$ to illustrate the adaption process. The error signal was transformed back into the plane wave domain for this purpose. We calculated the mean squared error of all plane wave compo-

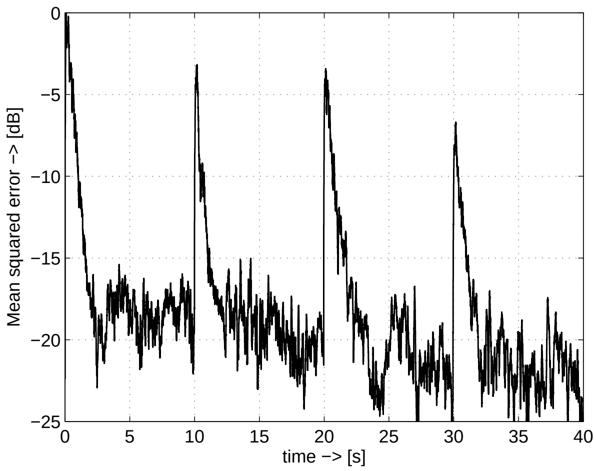


Fig. 11: Resulting mean squared error between the compensated and the desired wave field when using 48 compensation filters.

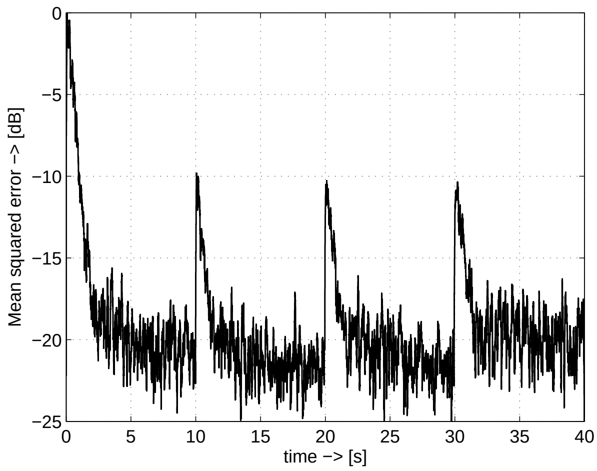


Fig. 12: Resulting mean squared error between the compensated and the desired wave field when using 78 compensation filters.

nents in $\bar{\mathbf{e}}$ over the time axis. This corresponds to a omni-directional pickup of the error signal at the origin. Figure 11 shows the results when using only 48 compensation filters. The varying source positions can be seen clearly at the error peaks. This is a result of the imperfect decoupling of the room response. Nevertheless, the filters converge fast after the source position has changed. To improve these results we also included some of the off-diagonal elements of $\tilde{\mathbf{R}}$. Figure 12 shows the results when using 78 compensation filters. It can be seen clearly that the error after the source position has changed is now much smaller. As the error after convergence is in the same range as for the 48 filter case, the filtered channels seem to exhibit very low correlations in-between. Figure 13 shows the plane wave decomposed measured wave field for the last source position $\alpha = 30^o$. 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. Figure 14 shows the resulting field after applying the proposed room compensation algorithm with 78 filters. In this case the time axis ranges from $t = 200$ ms to $t = 700$ ms because of a delay introduced when calculating the room compensation filters that insures causality of the solution. It can be seen clearly that the reverberation caused by the room is compensated by our approach. Additionally it can be seen that the room compensation filters do not produce artifacts in the resulting field.

6. CONCLUSIONS

The approach to active listening room compensation presented in this paper overcomes the limitations of classic room compensation approaches in the context of massive multi-channels reproduction systems. Its benefits follow from the features of the underlying principles:

1. WFS allows full control over the entire wave field inside the listening area. As a result it is possible to perform compensation in an extended area instead of multiple compensated points.

2. Listening room analysis based on the plane wave decompositions efficiently analyzes the entire wave field inside the listening area. As a result it is possible to calculate compensation filters which are valid for the entire listening area.

3. Transformation of the problem into the wave domain allows an efficient implementation of the filtering process. As a result it is possible to perform adaptive room compensation for typical WFS systems.

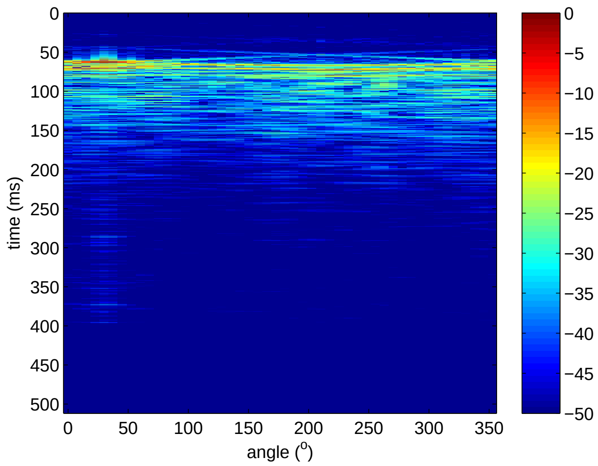


Fig. 13: Plane wave decomposed measured wave field for a point source at an angle of 30^o and 5m distance. The colors denote the signal energy in dB.

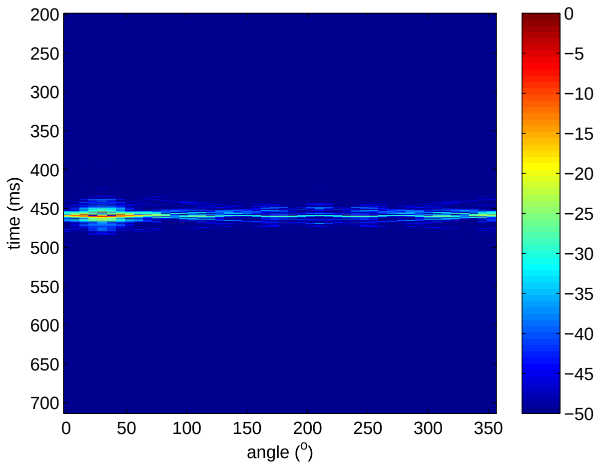


Fig. 14: Resulting (plane wave decomposed) wave field after compensation with 78 compensation filters for a point source at an angle of 30^o and 5m distance. The colors denote the signal energy in dB.

However, there are still a number of open questions that have to be investigated further. Typical WFS systems are currently limited to two-dimensions. This restriction poses problems if the room resembles strong reflections from the ceiling or the floor. Because these reflections cannot be separated from reflections coming from the reproduction or analysis plane, the algorithm generates compensation signals for them. Research is currently going on in the field of wave field analysis to overcome this problem. However, first simulations showed that the influence of the ceiling/floor reflections is not too severe. Another problem related to the restriction to two-dimensions are the amplitude errors introduced by the analysis and reproduction system. Compensation of these artifacts is another research topic at our lab. The algorithm presented here relies on the fact that the energy distribution of room transfer matrix $\tilde{\mathbf{R}}$ in the wave domain changes only slightly for typical variations of the listening room acoustics. Simulations and measurements revealed that this assumption holds in first approximation for typical scenarios. However, this topic is further analyzed.

The presented approach to active listening room compensation can be generalized to adaptive filtering in the wave domain. Successful application to the problem of acoustic echo cancellation (AEC) for WFS systems is presented in [16].

The research reported here has been funded by the German “Bundesministerium für Wirtschaft und Arbeit” and has been performed in collaboration with Airbus Deutschland GmbH.

7. REFERENCES

[1] L.D. Fielder. Practical limits for room equalization. In *111th AES Convention*, New York, NY, USA, September 2001. Audio Engineering Society (AES).

[2] M.M. Sondhi, D.R. Morgan, and J.L. Hall. Stereophonic acoustic echo cancellation – an overview of the fundamental problem. *IEEE Signal Processing Letters*, 2(8):148–151, August 1995.

[3] 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.

[4] S. Spors, A. Kuntz, and R. Rabenstein. An approach to listening room compensation with wave field synthesis. In *AES 24th International Conference on Multichannel Audio*, pages 49–52, Banff, Canada, June 2003. Audio Engineering Society (AES).

[5] S.Spors, H.Buchner, and R.Rabenstein. Adaptive listening room compensation for spatial audio systems. In *European Signal Processing Conference (EUSIPCO)*, 2004.

[6] A.J. Berkhout. A holographic approach to acoustic control. *Journal of the Audio Engineering Society*, 36:977–995, December 1988.

[7] E.W. Start. *Direct Sound Enhancement by Wave Field Synthesis*. PhD thesis, Delft University of Technology, 1997.

[8] E.N.G. Verheijen. *Sound Reproduction by Wave Field Synthesis*. PhD thesis, Delft University of Technology, 1997.

[9] P. Vogel. *Application of Wave Field Synthesis in Room Acoustics*. PhD thesis, Delft University of Technology, 1993.

[10] 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).

[11] J.A. Scales. *Theory of acoustic imaging*. Samizdat Press, 1997.

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

[13] 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).

[14] A.K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, 1989.

[15] H.Buchner, J.Benesty, and W.Kellermann. Multichannel frequency-domain adaptive algorithms with application to acoustic echo cancellation. In J. Benesty and Y. Huang, editors, *Adaptive signal processing: Application to real-world problems*. Springer, 2003.

[16] H.Buchner, S.Spors, and W.Kellermann. Full-duplex systems for sound field recording and auralization based on Wave Field Synthesis. In *116th AES Convention*, Berlin, Germany, 2004. Audio Engineering Society (AES).