WAVE FIELD CANCELLATION USING WAVE-DOMAIN ADAPTIVE FILTERING

W. Herbordt, S. Nakamura

ATR Spoken Language Translation Research
2-2-2 Hikaridai Seikacho Soraku-gun,
Kyoto 619-0288, Japan
{wolfgang.herbordt,satoshi.nakamura}@atr.jp

S. Spors, H. Buchner, W. Kellermann

Telecommunications Laboratory
Cauerstraße 7, 91058 Erlangen, Germany
{spors,buchner,wk}@LNT.de

Introduction. Wave field synthesis (WFS) [1] is an attractive and high-quality auralization technique for reproducing or synthesizing sound fields in a listening room using loudspeaker arrays. WFS allows a physically correct reproduction of sound fields and can be used for sound immersion in virtual environments or in multimedia communication environments. In reproduction systems (Fig. 1a), the wave field in the recording environment is first recorded ('wave field analysis' (WFA)) with a microphone array and, after storage or transmission, played back in the listening room using WFS. This scenario may require the playback of the wave field of only desired sound sources with the original sound realism while the wave field of undesired sources should not be reproduced. This problem may be resolved by picking up the clean source signals using close-talking microphones or steered arrays of microphones and by measuring the wave propagation in the recording room separately by impulse response measurements [2]. However, if real-time WFA/WFS is required as, e.g., in communication systems, the measurement of the room impulse responses is problematic due to the time variance of the acoustic conditions.

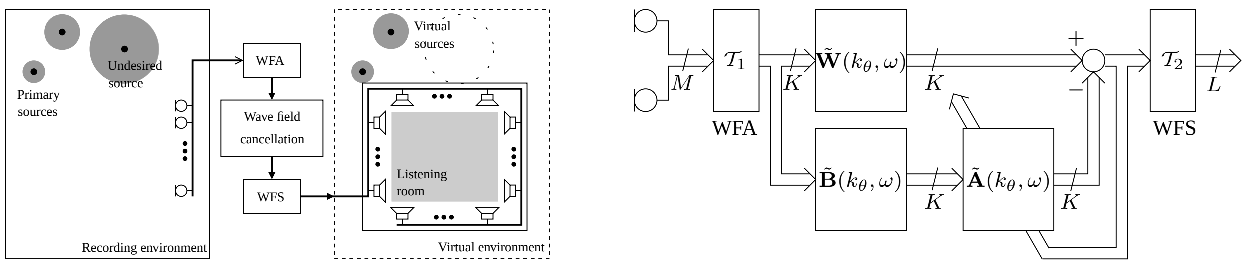

Figure 1: **(a)** Wave field analysis (WFA) and synthesis (WFS) system (after [2]) with cancellation of undesired wave fields. **(b)** Wave field cancellation using a wave-domain generalized sidelobe canceller (WGSC) and wave-domain adaptive filtering.

In this contribution, we propose a new technique called 'wave field cancellation' for suppressing undesired wave fields while preserving the wave field of desired sources. Wave field cancellation operates after WFA using wave-domain adaptive filtering (WDAF) [3] and can be efficiently realized using a 'wave-domain generalized sidelobe canceller' (WGSC). Due to the downward compatibility of WFS, the processed signal can be used for reproduction in a WFS system but also for monophonic, stereophonic, or 5.1-systems.

Wave field cancellation using wave-domain adaptive filtering. The principle of wave field cancellation using a WGSC is depicted in Fig. 1b. The recorded data is transformed by $\mathcal{T}_1$ into the wave domain in two steps: First, the sound pressure field $p(\mathbf{r}, t)$ is decomposed into plane waves using the plane wave decomposition $\bar{p}(\theta, \tau) = \mathcal{P}\{p(\mathbf{r}, t)\}$ [4].[1] The recorded data becomes independent from the microphone array geometry. Fig. 2a shows an example of a plane wave decomposed point source located in the far field ($r = 5\,\text{m}$, $\theta = \frac{3\pi}{2}$) of a circular microphone array with $M = 48$ sensors in environments with $T_{60} = 250\,\text{ms}$ and $T_{60} = 0\,\text{ms}$ reverberation time. The sensors are uniformly distributed on a circle with radius $r = 75\,\text{cm}$ with $\phi = 0$. Second, $\tilde{p}(\theta, \tau)$ is transformed into the wave domain by applying a Fourier transform on θ and τ, i.e., $\bar{p}(k_\theta, \omega) = \mathcal{F}_{\theta,\tau}\{\bar{p}(\theta, \tau)\}$, which orthogonalizes $\bar{p}(\theta, \tau)$ to angular wave number k_θ and frequency ω. The WGSC processing is similar to the GSC in the time-domain or in the frequency domain [5]. The wave-domain filters $\tilde{\mathbf{W}}(k_\theta, \omega)$ and $\tilde{\mathbf{B}}(k_\theta, \omega)$ decompose $\tilde{p}(k_\theta, \omega)$ into two orthogonal signal paths. $\tilde{\mathbf{B}}(k_\theta, \omega)$ selects the wave field of the undesired sources, which is then passed to the wave field canceller $\tilde{\mathbf{A}}(k_\theta, \omega)$. $\tilde{\mathbf{A}}(k_\theta, \omega)$ subtracts the undesired sources from the reference channels which are formed by the output of $\tilde{\mathbf{W}}(k_\theta, \omega)$. $\mathcal{T}_2$ transforms the filtered wave field to loudspeaker signals for auralization using a loudspeaker array [1]. The signal processing in the wave domain is advantageous due to the orthogonalizing property of $\mathcal{T}_1$, which decouples the input channels and reference channels of the wave field canceller in the ideal case. Essentially, for a realization using a 2-dimensional discrete Fourier transform (DFT)

This research was supported by a program with the National Institute of Information and Communications Technology (NICT), Japan, entitled "JAPAN TRUST International Research Coorporation Program".

[1]For simplicity, we consider only the two-dimensional case and neglect the angle ϕ in a spherical corrdinate system (r, θ, ϕ).

with K wavenumber bins and N frequency bins, only K decoupled filters of length N modeled in the 1-D DFT domain need to be adapted in the wave domain instead of M^2 coupled filters for a realization in the frequency domain. Assuming $K = M = 48$, this means that only 48 1-D filters in the DFT domain have to be adapted instead of 2304.

Realization. We verify out method using the sensor array described above. We consider an interference wave field $p_i(\mathbf{r}, t)$ of a point source (male speech) with coordinates ($r_i = 5\,\mathrm{m}$, $\theta_i = \frac{3\pi}{2}$) and a desired wave field $p_d(\mathbf{r}, t)$ of a point source at position ($r_d = 3\,\mathrm{m}$, $\theta_d = \frac{1\pi}{2}$). The plane wave decompositions $\bar{p}_i(\theta, \tau)$ and $\bar{p}_d(\theta, \tau)$ of $p_i(\mathbf{r}, t)$ and $p_d(\mathbf{r}, t)$ are given in Fig. 2a and Fig. 2d (dashed line), respectively. The Fourier transform is carried out using the FFT. The frequency range is $50\,\mathrm{Hz}$ to $1\,\mathrm{kHz}$.[2] $\tilde{\mathbf{B}}(k_\theta, \omega)$ extracts the plane wave component $\bar{p}(\theta_i, \tau)$ as reference for the interference and suppresses all other directions. $\tilde{\mathbf{W}}(k_\theta, \omega)$ is orthogonal to $\tilde{\mathbf{B}}(k_\theta, \omega)$ and, thus, suppresses $\bar{p}(\theta_i, \tau)$ while preserving all other directions. Note that reverberation of the desired wave field which is contained in the plane wave component $\bar{p}(\theta_i, \tau)$ is also cancelled. The effect on the sound impression during reproduction has not yet been verified, however, as long as $\bar{p}(\theta_i, \tau)$ contains no dominant reflections, we expect the sound impression to be only negligible –if not unnoticably– impaired because of the diffuse character of reverberation. Figure 2b,c show the interference suppression IR as a function of time and as a function of the plane wave components. It can be seen that the algorithm converges rapidly and that the interference wave field can be efficiently cancelled while the desired wave field is preserved except for $\bar{p}_d(\theta_i, \tau)$ (Fig. 2d, solid line). The wave field canceller $\tilde{\mathbf{A}}(k_\theta, \omega)$ is realized for each of the $K = M$ wavenumber bins by a one-channel DFT-domain adaptive filter [6] of length $N = 1024$ using a block overlap factor 4.

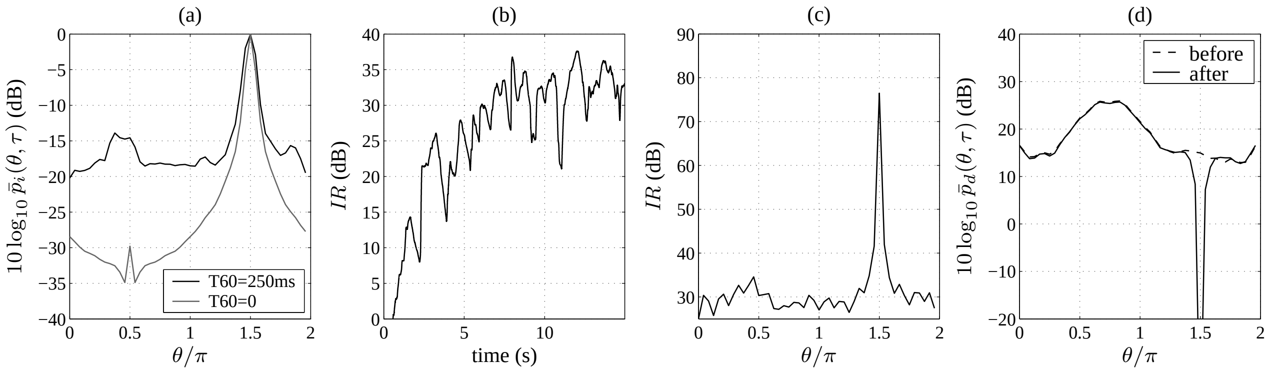

Figure 2: **(a)** Plane wave decomposition $\bar{p}_i(\theta, \tau)$ of the interference in the far field of a microphone array in a reverberant environment with $T_{60} = 250\,\mathrm{ms}$ and $T_{60} = 0$; **(b)** Interference suppression IR as a function of time after initialization averaged over all plane wave components θ and **(c)** as a function of plane wave components θ averaged over the last $5\,\mathrm{s}$. **(d)** Plane wave decomposition $\bar{p}_d(\theta, \tau)$ of the desired wave field before and after wave field cancellation.

Summary. We presented a novel framework for interference cancellation in multi-channel audio reproduction systems, which is based on a wave-domain GSC and which uses wave-domain adaptive filtering. In contrast to conventional interference cancellation, the 3-dimensional sound impression of the desired sources is preserved for most directions for high quality sound reproduction. Wave-domain adaptive filtering is also suitable for the solution of other multi-channel audio problems like acoustic echo cancellation [3] or adaptive listening room compensation [7].

References

[1] A.J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *Journal of the Acoustical Society of America*, 93(5):2764–2778, May 1993.

[2] H. Teutsch, S. Spors, W. Herbordt, W. Kellermann, and R. Rabenstein. An integrated real-time system for immersive audio applications. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, October 2003.

[3] H. Buchner, S. Spors, W. Herbordt, and W. Kellermann. Wave-domain adaptive filtering for acoustic human-machine interfaces based on wavefield analysis and synthesis. *EURASIP Proc. European Conf. on Signal Processing*, 2004.

[4] E. Hulsebos, D. de Vries, and E. Bourdillat. Improved microphone array configurations for auralization of sound fields by wave field synthesis. *110-th Convention of the Audio Engineering Society*, May 2001.

[5] L.J. Griffiths and C.W. Jim. An alternative approach to linearly constrained adaptive beamforming. *IEEE Trans. on Antennas and Propagation*, 30(1):27–34, January 1982.

[6] D. Mansour and A.H. Gray. Unconstrained frequency-domain adaptive filter. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 30(5):726–734, 1982.

[7] S. Spors, H. Buchner, and R. Rabenstein. A novel approach to active listening room compensation for wave field synthesis using wave-domain adaptive filtering. *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, 4:iv–117–iv–120, May 2004.

[2] The frequency range is limited since we use a circular WFS system with radius $r = 1.4\,\mathrm{m}$ with 48 loudspeakers to simulate the point sources, where the upper cut-off frequency is defined by spatial aliasing and the spacing of the loudspeakers [1].