ACTIVE ACOUSTIC ECHO CANCELLATION IN SPATIAL SOUNDFIELD REPRODUCTION

Dumidu S. Talagala, Wen Zhang and Thushara D. Abhayapala

Research School of Engineering, CECS, Australian National University. Canberra. Australia.
Email: {dumidu.talagala, wen.zhang, thushara.abhayapala}@anu.edu.au

ABSTRACT

The equalization of reverberation effects is essential for spatial soundfield reproduction, but estimation of the reverberant channel presents several challenges to existing equalization techniques. This paper presents a method of active acoustic echo cancellation (AEC) for soundfield reproduction applications, using a modal description of the reverberant soundfield. We describe how individual modes of the measured soundfield can be equalized adaptively, thus reducing the complexity of the channel estimation process. AEC and reproduction performance is compared with existing adaptive and non-adaptive equalization techniques through simulation examples. Equalization performance is comparable to existing methods, achieving a normalized region reproduction error of 1% and echo return loss enhancement of 15 - 30 dB at 50 dB SNR. The results suggest that the proposed model can be used to obtain a parallel implementation of a room equalizer for active AEC.

Index Terms— Acoustic echo cancellation, room equalization, reverberation, spatial filtering, soundfield reproduction

1. INTRODUCTION

Reproduction of complex soundfields within a region has become a topic of interest in signal processing during the past decade [1–6]. Applications such as virtual reality teleconferencing and gaming involve full duplex communication and spatial soundfield reproduction; hence, acoustic echo cancellation (AEC) of the reproduced soundfield becomes essential. However, the large number of loudspeakers and microphones used by various reproduction techniques introduce new challenges to traditional mechanisms of echo cancellation [7, 8]. These are mainly related to the correlation between channels, which results in ill conditioning of matrices during channel estimation. Spatial transformation of the received signals has been proposed to overcome the inter-channel correlation problem in multi-channel AEC [9–13], and receiver-side adaptive channel estimation has been proven effective in both reverberant and non-reverberant conditions.

Unlike receiver-side AEC, equalization of the reverberation effects is an active control problem in spatial soundfield reproduction. Given the accurate reproduction of the desired soundfield, echo cancellation at any point within that region is achieved by computing the difference between the measured and desired soundfields. In Higher-order Ambisonics (HOA) [4, 14, 15] and Wave Field Synthesis (WFS) [16, 17] a desired soundfield is recreated by incorporating estimates of the reverberant channels. Both require fine sampling of the soundfield at the edge of the desired region, which translates to a large number of loudspeakers and microphones. In practise, a limited number of reproduction channels are used to recreate the desired soundfield, aided by various models that characterise the reverberant room behaviour [18–21]. Modelling errors will naturally affect the echo cancellation performance, but better approximation of the reverberant soundfield at the receiver array can improve the AEC performance. Wave Domain Adaptive Filtering (WDAF) [3,6,12,22–24] used in soundfield reproduction provides some physical insights into the underlying structure of the reverberant soundfield, and has inspired the channel model presented in this work.

In this paper, we consider the problem of active AEC in spatial soundfield reproduction using a circular microphone array. First, we use the wave domain representation of signals to describe the desired and reverberant soundfields in a region, where the desired soundfield modes can be computed assuming free field propagation. Next, we describe a method of characterising the reverberation component of each mode as a linear transformation of the corresponding desired soundfield mode. This parameterization further reduces the complexity in the wave domain adaptation process of estimating the reverberant channel and computing the required compensation signals. Finally, simulation results of AEC and region reproduction performance are presented for the reproduction of virtual narrowband plane waves in a reverberant environment.

2. CHARACTERISATION OF REVERBERANT SOUNDFIELDS

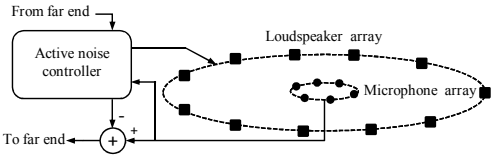


Fig. 1. Overview of the active noise canceller including the loudspeaker and microphone array configurations.

The structure of the proposed active echo canceller is shown in Fig. 1 above. It consists of concentric circular arrays of P loudspeakers and Q microphones, where the annulus forms the region of interest. The loudspeakers driving signals are preconditioned to reproduce a desired soundfield, while the microphone array is required to record the soundfield generated by the other sources within the room.

2.1. Modal characterisation of soundfields

The received signals at the microphone array can be expressed in matrix form as

$$\mathbf{Y_q}(\omega, t) = \mathbf{H_{pq}}(\omega)\mathbf{X_p}(\omega, t), \tag{1}$$

This work was supported under the Australian Research Councils Discovery Projects funding scheme (project no. DP110103369).

where $\mathbf{X_p}(\omega,t) = \left[\, X_1(\omega,t)\ X_2(\omega,t)\ \cdots\ X_P(\omega,t) \,\right]^T$ are the loudspeaker driving signals, $\mathbf{Y_q}(\omega,t) = \left[\, Y_1(\omega,t)\ Y_2(\omega,t)\ \cdots\ Y_Q(\omega,t) \,\right]^T$ are the received signals at the microphone array, and $\mathbf{H_{pq}}(\omega)$ represents the channel between the $(p,q)^{th}$ loudspeaker - microphone pair at an angular frequency ω. Applying the modal-based approach, the received signals in (1) can now be characterised in terms of their soundfield modes as

$$Y_q(\omega,t) = \sum_{n=-N}^{N} \alpha_n(\omega,t) e^{in\phi_q}, \tag{2}$$

where N is the truncation length [2,25,26] (number of active modes) for a specified error, ϕ_q is the direction of the q^{th} microphone and $\alpha_n(\omega,t)$ are the measured soundfield coefficients. The values of $\alpha_n(\omega,t)$ can be calculated using the spatial Discrete Fourier Transform (DFT) [26] as

$$\alpha_n(\omega,t) = \frac{1}{Q} \sum_{q=0}^{Q-1} Y_q(\omega,t) e^{-i(2\pi qn/Q)}, \tag{3}$$

for equally spaced microphones in the receiver array. Hence, the calculation of the $2N+1$ soundfield coefficients can be expressed in the matrix form

$$\boldsymbol{\alpha_n}(\omega,t) = \mathbf{T_{CH}} \mathbf{H_{pq}}(\omega) \mathbf{X_p}(\omega,t), \tag{4}$$

where $\boldsymbol{\alpha_n}(\omega,t) = \left[\, \alpha_{-N}(\omega,t) \cdots \alpha_N(\omega,t) \,\right]^T$ and $\mathbf{T_{CH}}$ is a DFT matrix, an expression of (3) as a matrix multiplication.

Given the soundfield coefficients $\alpha_n^D(\omega,t)$ of a desired soundfield, the soundfield produced by other sources within the current room can be recreated from $\alpha_n^T(\omega,t) = \alpha_n(\omega,t) - \alpha_n^D(\omega,t)$. Hence, the actions of the active noise controller in Fig. 1 can simply be described as the rectification of the reverberation effects caused by the room (i.e., minimizing the sum of square errors of $\alpha_n^T(\omega,t)$ for all $n = -N \ldots N$), when all other sources are inactive.

2.2. Modelling the reverberant channel

Consider the measured soundfield coefficients in (4); a result of the superposition of the desired and reverberant soundfields, and that of other sources within the room. If we now consider other sound sources within the room to be inactive,

$$\boldsymbol{\alpha_n}(\omega,t) \triangleq \boldsymbol{\alpha_n^D}(\omega,t) + \boldsymbol{\alpha_n^R}(\omega,t), \tag{5}$$

where $\boldsymbol{\alpha_n^D}(\omega,t)$ and $\boldsymbol{\alpha_n^R}(\omega,t)$ represent the desired and reverberant soundfield coefficients respectively.

Since reverberation is a result of our attempt to reproduce the desired soundfield, each reverberant soundfield mode can be considered a transformation of its desired soundfield mode. Exploiting the orthogonality and independence of the soundfield modes with each other [26], this can be expressed by a transformation of the soundfield coefficients

$$\boldsymbol{\alpha_n^R}(\omega,t) = \mathbf{H_n^R}(\omega) \boldsymbol{\alpha_n^D}(\omega,t), \tag{6}$$

where $\mathbf{H_n^R}(\omega) = \mathrm{diag}\left[H_{-N}^R(\omega) \ldots H_N^R(\omega)\right]$ represents the transformation applied to each mode. The diagonal structure of $\mathbf{H_n^R}(\omega)$ can be visualized intuitively, by considering each desired soundfield mode as an independent source, where $H_n^R(\omega)$ represents the effects of reverberation (i.e., the resultant transformation of multiple scaled and delayed versions of the desired mode). A hint of this underlying structure has been shown in the works related to WDAF in [6, 24]. $\mathbf{H_n^R}(\omega)$ now represents the reverberant channel effects at a frequency ω, and can be used to derive the necessary loudspeaker compensation signals, as described next.

3. REVERBERATION COMPENSATION

3.1. Compensation signals at the loudspeakers

Consider expressing (5) by separating the direct and reverberant path effects of $\mathbf{H_{pq}}(\omega)$. Thus,

$$\boldsymbol{\alpha_n}(\omega,t) = \mathbf{T_{CH}} \left[\mathbf{H_{pq}^D}(\omega) + \mathbf{H_{pq}^R}(\omega)\right] \mathbf{X_p}(\omega,t), \tag{7}$$

where $\mathbf{H_{pq}^D}(\omega)$ and $\mathbf{H_{pq}^R}(\omega)$ represent the direct and reverberant channels between the $(p,q)^{th}$ loudspeaker - microphone pair. Incorporating the reverberation model in (6)

$$\boldsymbol{\alpha_n}(\omega,t) = \left[\mathbf{I} + \mathbf{H_n^R}(\omega)\right] \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \mathbf{X_p}(\omega,t), \tag{8}$$

where $\mathbf{I}$ is the identity matrix.

Next, we introduce reverberation compensation signals to each loudspeaker, in order to maximize the echo cancellation performance. This is achieved when $\widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}(\omega,t) = \boldsymbol{\alpha_n^D}(\omega,t)$, where $\widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}(\omega,t)$ is the measured soundfield coefficient after compensation. Hence, (8) now becomes

$$\widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}(\omega,t) = \left[\mathbf{I} + \mathbf{H_n^R}(\omega)\right] \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \left[\mathbf{X_p}(\omega,t) + \boldsymbol{\delta} \mathbf{X_p}(\omega,t)\right], \tag{9}$$

where $\boldsymbol{\delta}\mathbf{X_p}(\omega,t)$ are the loudspeaker compensation signals. Expanding (9) further using $\boldsymbol{\alpha_n^D}(\omega,t) = \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \mathbf{X_p}(\omega,t)$, from (4), we find

$$\begin{aligned} \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \boldsymbol{\delta} \mathbf{X_p}(\omega,t) = &-\mathbf{H_n^R}(\omega) \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \mathbf{X_p}(\omega,t) \\ &- \mathbf{H_n^R}(\omega) \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \boldsymbol{\delta} \mathbf{X_p}(\omega,t). \end{aligned} \tag{10}$$

Note that the second term of (10) is the reverberation of the compensation signals, which we assume is negligible for small $\boldsymbol{\delta}\mathbf{X_p}(\omega,t)$, or moderate to high direct to reverberation path ratios. Thus,

$$\mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \boldsymbol{\delta} \mathbf{X_p}(\omega,t) = -\mathbf{H_n^R}(\omega) \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \mathbf{X_p}(\omega,t), \tag{11}$$

and loudspeaker compensation signals are given by

$$\boldsymbol{\delta} \mathbf{X_p}(\omega,t) = -\left[\mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega)\right]^{\dagger} \mathbf{H_n^R}(\omega) \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \mathbf{X_p}(\omega,t), \tag{12}$$

where $\left[\mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega)\right]^{\dagger}$ is the Moore-Penrose pseudoinverse of $\mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega)$. Since $\mathbf{H_{pq}^D}(\omega)$ is known for a given loudspeaker - microphone configuration, the loudspeaker compensation signals can be calculated, provided an estimate of the reverberant channel transformation matrix $\mathbf{H_n^R}(\omega)$.

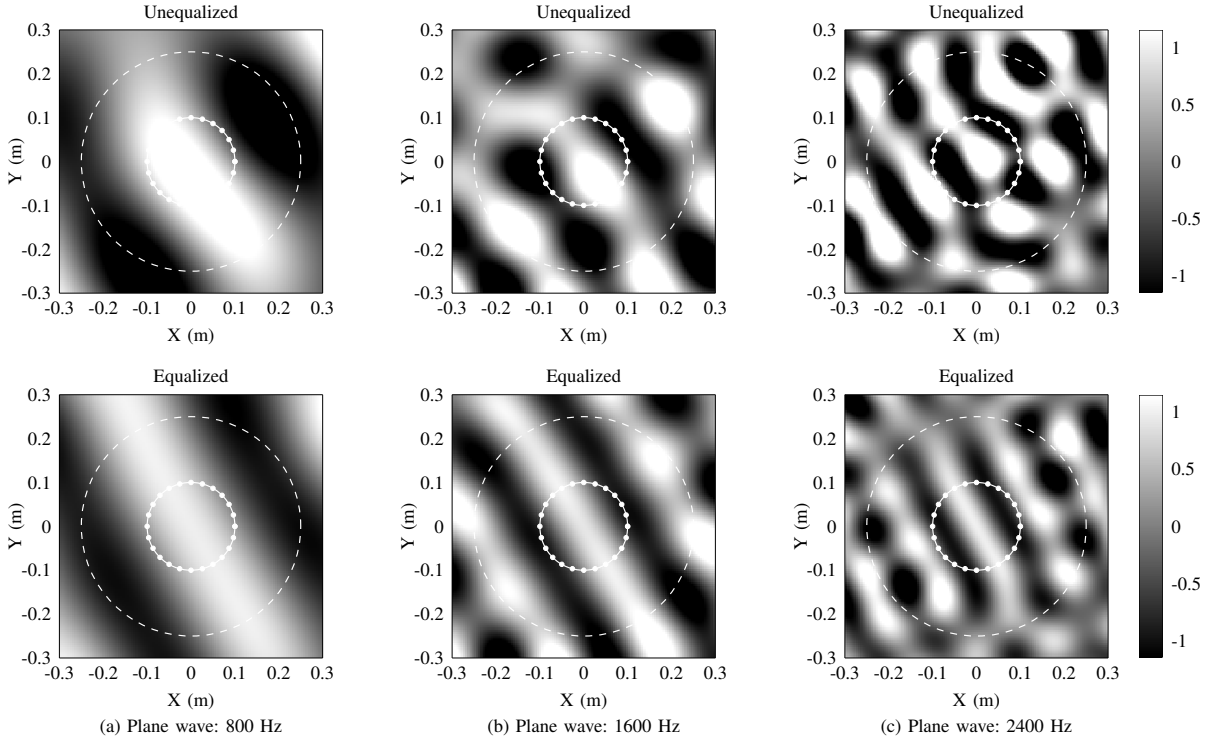
3.2. Estimation of the reverberant channel

Using the reverberation model in (5) and (6), the measured soundfield coefficients can now be expressed as

$$\begin{aligned} \widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}(\omega,t) = \boldsymbol{\alpha_n^D}(\omega,t) + \mathbf{H_n^R}(\omega) \boldsymbol{\alpha_n^D}(\omega,t) + \\ \mathbf{T_{CH}} \left[\mathbf{H_{pq}^D}(\omega) + \mathbf{H_{pq}^R}(\omega)\right] \boldsymbol{\delta} \mathbf{X_p}(\omega,t). \end{aligned} \tag{13}$$

Suppose $\widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega)$ is an estimate of $\mathbf{H_n^R}(\omega)$. Equation (13) can now be simplified further by using the results of (11) and by neglecting the reverberation effects of the compensation signals. Thus, the soundfield to be transmitted becomes

$$\begin{aligned} \widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}^{\boldsymbol{T}}(\omega,t) &= \mathbf{H_n^R}(\omega) \boldsymbol{\alpha_n^D}(\omega,t) - \widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega) \mathbf{T_{CH}} \mathbf{H_{pq}^D}(\omega) \mathbf{X_p}(\omega,t) \\ \widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}^{\boldsymbol{T}}(\omega,t) &= \left[\mathbf{H_n^R}(\omega) - \widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega)\right] \boldsymbol{\alpha_n^D}(\omega,t), \end{aligned} \tag{14}$$



(a) Plane wave: 800 Hz (b) Plane wave: 1600 Hz (c) Plane wave: 2400 Hz

Fig. 2. Unequalized and equalized soundfields of a plane wave source reproduced in the direction $\pi/3$ at (a) 800 Hz, (b) 1600 Hz and (c) 2400 Hz at 50 dB SNR. The dotted white inner circle indicates the 0.1 m radius microphone array and their locations, while the outer dashed circle indicates the reproduction region of 0.25 m radius.

where $\widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}^{\boldsymbol{T}}(\omega,t) = \widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}(\omega,t) - \boldsymbol{\alpha}_{\boldsymbol{n}}^{\boldsymbol{D}}(\omega,t)$. Estimating $\widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega)$ is now a matter of minimizing the square error of $\widetilde{\boldsymbol{\alpha}}_{\boldsymbol{n}}^{\boldsymbol{T}}(\omega,t)$, a classical adaptive filtering problem.

The adaptation equation

$$\widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega,t_k)^H = \widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega,t_{k-1})^H + \boldsymbol{\Phi}\boldsymbol{\alpha}_{\boldsymbol{n}}^{\boldsymbol{D}}(\omega,t_k)\boldsymbol{\alpha}_{\boldsymbol{n}}^{\boldsymbol{T}}(\omega,t_k)^H \quad (15)$$

can now be used to obtain an iterative estimate of $\widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega)$ at a time $t = t_k$, where $\boldsymbol{\Phi}$ is a technique dependent adaptation gain [27]. Although (15) typically involves the calculation of the $(2N+1)^2$ unknown coefficients of $\widehat{\mathbf{H}}_{\mathbf{n}}^{\mathbf{R}}(\omega)$, the diagonal structure of $\mathbf{H}_{\mathbf{n}}^{\mathbf{R}}(\omega)$ can be used to further simplify the adaptation process. For example, calculating each $\widehat{H}_n^R(\omega)$ can be implemented as a low complexity single tap adaptive filter, where

$$\widehat{H}_n^R(\omega,t_k)^H = \widehat{H}_n^R(\omega,t_{k-1})^H + \phi_n(t_k)\alpha_n^D(\omega,t_k)\alpha_n^T(\omega,t_k)^H. \quad (16)$$

$\phi_n(t_k)$ is the gain factor of the adaptive technique, where

$$\phi_n(t_k) = \left[\lambda\sigma^2(t_{k-1}) + \left|\alpha_n^D(\omega,t_k)\right|^2\right]^{-1},$$

$$\sigma^2(t_k) = \lambda\sigma^2(t_{k-1}) + \left|\alpha_n^D(\omega,t_k)\right|^2$$

and λ is the forgetting factor for the Recursive Least Squares (RLS) algorithm [27]. In this paper, we use the RLS adaptation method given above, but any appropriate adaptive technique can be applied.

4. EVALUATION

We evaluate the model performance through simulation, using a 5 m $\times$ 6.4 m reverberant room with wall absorption coefficients of 0.36. The floors and ceilings are assumed to be non-reflective and the image source method [28] is used to simulate the reverberant soundfield up to a depth of 5 image sources. Two circular arrays of 27 loudspeakers and 24 microphones are concentrically located with their origins at $\{2.4$ m, 3.8 m$\}$ and radii $r_l = 2$ m and $r_m = 0.1$ m respectively. Both arrays are designed for frequencies up to 4 kHz for soundfield reproduction at the microphone array. A narrowband virtual plane wave source is reproduced every 100 Hz in the [100, 3000] Hz frequency range with an angle of incidence of $\pi/3$ and 0 dB power at the centre of the reproduction region. The mode truncation length $N = \lceil e\omega r_m/2c \rceil$, where the speed of sound in air $c =$ 343 m/s. Ambient noise is white Gaussian with a signal to noise ratio (SNR) of 50 dB at the centre of the reproduction region. This configuration yields an average direct to reverberant power ratio of 1.1 dB across frequencies up to 4 kHz. All sound sources are sampled at 44.1 kHz, while the RLS algorithm is used for adaptation at a rate of 14.7 kHz and $\lambda = 0.75$.

Performance is compared with the fixed multi-point equalization method [29] using perfect channel information, and the Filtered-X Recursive Least Squares (FxRLS) [3, 30] adaptive algorithm using channel information at 99% accuracy. Echo return loss enhancement (ERLE) and normalized region reproduction error (RRE) given below, are used as metrics for comparison.

$$ERLE(\omega,t) = 10\log_{10}\left|\frac{\boldsymbol{\alpha}_{\boldsymbol{n}}(\omega,t)^H\boldsymbol{\alpha}_{\boldsymbol{n}}(\omega,t)}{\boldsymbol{\alpha}_{\boldsymbol{n}}^{\boldsymbol{T}}(\omega,t)^H\boldsymbol{\alpha}_{\boldsymbol{n}}^{\boldsymbol{T}}(\omega,t)}\right| \quad (17)$$

$$NRRE(\omega,t) = 10\log_{10}\frac{\int_{\mathcal{S}}\left|Y(\omega,t) - Y^D(\omega,t)\right|^2 \mathrm{d}S}{\int_{\mathcal{S}}\left|Y^D(\omega,t)\right|^2 \mathrm{d}S}, \quad (18)$$

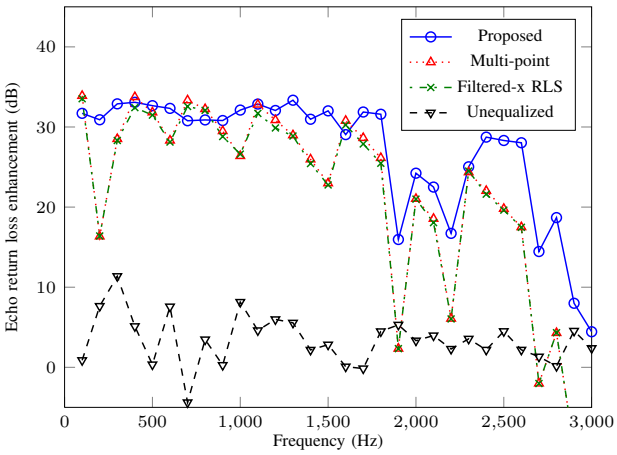


Fig. 3. ERLE of a reproduced plane wave source in the direction $\pi/3$, averaged over 10 trial runs at 50 dB SNR.

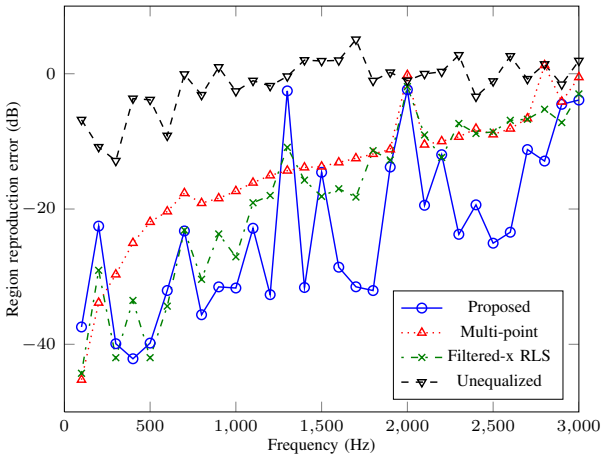


Fig. 4. Reproduction error within a region of radius 0.25 m, averaged over 10 trial runs at 50 dB SNR.

where $\mathrm{d}S$ is the differential area in the region of interest S and $Y(\omega,t)$, $Y^D(\omega,t)$ are the measured and desired soundfields at each location.

Fig. 2 illustrates the narrowband reproduced soundfield for a virtual plane wave source located in the direction $\pi/3$ after 500 adaptation steps. The selected frequencies correspond to three reproduction scenarios, where the number of reproduced modes is (a) greater, (b) equal to and (c) less than the design limit (corresponding to $2N+1$) at a radial distance of 0.25 m. The white dotted circle indicates the microphone array and microphone locations. Comparison of the unequalized and equalized soundfields suggests that good equalization can be achieved at the microphone array, within the design limits of the loudspeakers.

Acoustic echo cancellation performance at the microphone array is presented in Fig. 3, averaged over 10 trial runs after 3000 adaptation steps. The proposed reverberation model and active controller achieves gains in ERLE between 15 and 30 dB up to 2.5 kHz, limited by the 50 dB noise floor of the system. Performance is comparable with multi-point equalization and adaptive FxRLS, where the multi-point equalization assumes complete knowledge of the reverberant channel. Degradation of performance at frequencies above 2.5 kHz suggests that finer sampling maybe required at frequencies closer to the design limits of the arrays.

Normalized region reproduction error within a reproduction region of 0.25 m radius is shown in Fig. 4. Reproduction performance is comparable to other equalization method and the results suggest that reproduction errors below 1% is achievable. The sudden spikes in the reproduction error curves in Fig. 4 can be attributed to near zero valued soundfield coefficients at the measurement locations. This is caused by the zero crossings of the Bessel function ($J_n(\omega r_m/c)$) at the microphone array, which describes the radial change of the soundfield mode in the solutions to the wave equation [2, 26]. Appropriately designed dual microphone arrays can be used to mitigate this problem as described in [2].

These results suggest that the proposed reverberation model is an accurate representation of the reverberant soundfield. Additionally, reproduction errors below 1% up to 1.5 kHz beyond the microphone array suggests that the model may also be used in place of the reverberant room models used in soundfield reproduction applications.

5. CONCLUSION

In this paper, we present a model of the reverberant soundfield for acoustic echo cancellation in soundfield reproduction applications. We show that the mode domain can be used to characterise the reverberant soundfield, where the reverberant soundfield modes are transformations of individual desired soundfield modes. Next, we derive the reverberation compensation signals at the loudspeakers using these transformations. Finally, we show how the structure of the reverberation model itself simplifies the channel estimation problem, where single tap adaptive filters are used to estimate each transformation. Acoustic echo cancellation performance is comparable to other methods based on perfect channel information, which suggests that the proposed reverberation model is an accurate representation of the reverberant soundfield. Future work will investigate improving echo cancellation performance at higher frequencies. In conclusion, our results suggest mode domain channel models can be used to obtain a parallel implementation of a soundfield equalizer.

6. RELATION TO PRIOR WORK

The work presents a reverberant soundfield, modelled as a transformation of the desired soundfield, using a modal representation. This approach is well suited for echo cancellation in applications of soundfield reproduction within a region [2, 4]. Wave domain adaptive filtering [10, 12] originally proposed for AEC in multi-channel systems introduces the concepts of adaptive channel estimation in the spatial domain, but use a space-time approach to estimate the reverberant channel. This does not fully exploit the properties of the spatial representation, a hint of which is shown in [6, 23, 24]. The present study considers reverberation modelling as transformations of uncoupled, independent soundfield modes in space-frequency domain, which has not been considered in earlier studies.

7. REFERENCES

[1] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. on Speech and Audio Processing*, vol. 9, no. 6, pp. 697 –707, Sep. 2001.

[2] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2100–2111, 2005.

[3] S. Spors, H. Buchner, R. Rabenstein, and W. Herbordt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering," *J. Acoust. Soc. Am.*, vol. 122, no. 1, pp. 354–369, 2007.

[4] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 17, no. 1, pp. 107 –116, Jan. 2009.

[5] N. Radmanesh and I.S. Burnett, "Reproduction of independent narrowband soundfields in a multizone surround system and its extension to speech signal sources," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011*, May 2011, pp. 461 –464.

[6] M. Schneider and W. Kellermann, "Adaptive listening room equalization using a scalable filtering structure in the wave domain," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012*, Mar. 2012, pp. 13–16.

[7] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Trans. on Speech and Audio Process.*, vol. 6, no. 2, pp. 156 –165, Mar. 1998.

[8] T. Gaensler and J. Benesty, "Multichannel acoustic echo cancellation: what's new?," in *Proc. IEEE Workshop on Acoustic Echo and Noise Control (WAENC), 2001*, 2001.

[9] H. Buchner, W. Kellermann, and J. Benesty, "An extended multidelay filter: fast low-delay algorithms for very high-order adaptive systems," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2003*, Apr. 2003, vol. 5, pp. 385–388.

[10] H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2004*, May 2004, vol. 4, pp. 117 – 120.

[11] S. Spors, H. Buchner, and R. Rabenstein, "Eigenspace adaptive filtering for efficient pre-equalization of acoustic MIMO systems," in *Proc. European Signal Processing Conference (EUSIPCO)*, Sep. 2006.

[12] H. Buchner and S. Spors, "A general derivation of wave-domain adaptive filtering and application to acoustic echo cancellation," in *Asilomar Conference on Signals, Systems and Computers, 2008*, Oct. 2008, pp. 816 –823.

[13] K. Helwani, S. Spors, and H. Buchner, "Spatio-temporal signal preprocessing for multichannel acoustic echo cancellation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011*, May 2011, pp. 93 –96.

[14] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc*, vol. 53, no. 11, pp. 1004–1025, 2005.

[15] A. Gupta and T. D. Abhayapala, "Three-dimensional sound field reproduction using multiple circular loudspeaker arrays," *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 19, no. 5, pp. 1149 –1159, Jul. 2011.

[16] A. J. Berkhout, "A holographic approach to acoustic control," *J. Audio Eng. Soc*, vol. 36, no. 12, pp. 977–995, 1988.

[17] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, 1993.

[18] A. Krokstad, S. Strom, and S. Sørsdal, "Calculating the acoustical room response by the use of a ray tracing technique," *Journal of Sound and Vibration*, vol. 8, no. 1, pp. 118 – 125, 1968.

[19] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. E. West, G. Pingali, P. Min, and A. Ngan, "A beam tracing method for interactive architectural acoustics," *J. Acoust. Soc. Am.*, vol. 115, no. 2, pp. 739–756, 2004.

[20] E. A. Lehmann and A. M. Johansson, "Diffuse reverberation model for efficient image-source simulation of room impulse responses," *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 18, no. 6, pp. 1429 –1439, Aug. 2010.

[21] C. Lauterbach, A. Chandak, and D. Manocha, "Interactive sound rendering in complex and dynamic scenes using frustum tracing," *IEEE Trans. on Vis. Comput. Graphics*, vol. 13, no. 6, pp. 1672 –1679, Nov.-Dec. 2007.

[22] S. Spors and H. Buchner, "An approach to massive multichannel broadband feedforward active noise control using wave-domain adaptive filtering," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2007*, Oct. 2007, pp. 171 –174.

[23] M. Schneider and W. Kellermann, "A wave-domain model for acoustic MIMO systems with reduced complexity," in *Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA), 2011*, Jun. 2011, pp. 133 –138.

[24] M. Schneider and W. Kellermann, "A direct derivation of transforms for wave-domain adaptive filtering based on circular harmonics," in *Proc. European Signal Processing Conference (EUSIPCO)*, Aug. 2012, pp. 1034–1038.

[25] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. on Signal Processing*, vol. 55, no. 6, pp. 2542 –2556, Jun. 2007.

[26] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, 1999.

[27] S. O. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1996.

[28] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, 1979.

[29] S. J. Elliott and P. A. Nelson, "Multiple-point equalization in a room using adaptive digital filters," *J. Audio Eng. Soc*, vol. 37, no. 11, pp. 899–907, 1989.

[30] M. Bouchard and S. Quednau, "Multichannel recursive-least-square algorithms and fast-transversal-filter algorithms for active noise control and sound reproduction systems," *IEEE Trans. on Speech and Audio Process.*, vol. 8, no. 5, pp. 606 –618, Sep. 2000.