Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. 17-20, 1999

MULTICHANNEL EQUALIZATION IN SUBBANDS

Stephan Weiss, Garrey W. Rice, and Robert W. Stewart

Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow G1 1XW, Scotland, UK
`{weiss,garrey,bob}@spd.eee.strath.ac.uk`

ABSTRACT

For the dereverberation of acoustic channels or the rendering of a specific sound field, the inversion of acoustics is a central problem and generally involves multichannel techniques. In this paper, we introduce a subband approach to the adaptive solution of this equalization problem. The presented method generally allows for faster convergence at lower complexity. We also address limitations of the subband technique, potential error sources, and design specifications. Simulations are presented underlining the use of our method.

1. INTRODUCTION

The inversion of multichannel acoustic environments finds applications in techniques such as dereverberation, cross-talk cancellation, or sound field rendering [1, 2, 3, 4, 5]. An example for a two-channel setup is shown in Fig. 1. In this case, the combination of two loudspeakers and two sensors creates four separate transfer paths. For dereverberation, e.g. the microphone signals $x_m[n]$, $m \in \{1, 2\}$, could be post-processed to remove the effect of the room. Cross-talk cancellation could attempt to drive the actuators with signals $y_l[n]$, $l \in \{0, 1\}$, such that the listener perceives the unmodified stereophonic signal. Similarly, sound rendering requires the creation of a particular audio impression at the listener's ears by appropriate pre-processing of the loudspeaker signals — this includes the equalization of the room's acoustics. Hence, depending on the specific application, a pre- or post-equalization is required.

The challenge for establishing an inversion lies in the properties of acoustic systems, which generally are non-minimum phase, potentially possess spectral zeroes, and exhibit very long impulse responses. Therefore, the inverse has to be of considerable length and poses a high computational burden particularly when adaptive solutions are sought. To reduce the computational complexity, in the past IIR filters have been evaluated for similar tasks [6]. However, while the maximum phase part cannot be compensated by a recursive system anyway, in modelling comparisons for similar acoustic problems IIR filters showed no particular advantage over FIR systems [7].

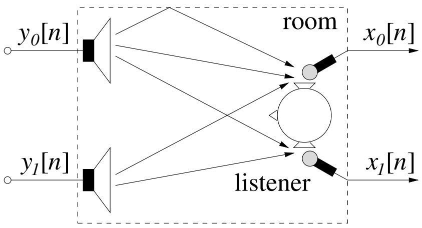

Figure 1: Cross-talk and reverberation in a stereophonic audio environment.

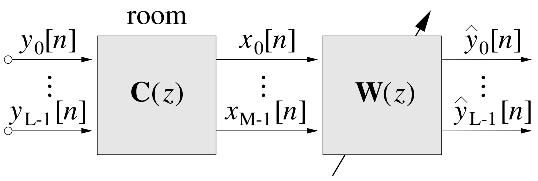

Figure 2: Inversion / dereverberation of a MIMO system $\mathbf{C}(z)$.

Here, we introduce a subband approach to adaptive multichannel equalization, whereby adaptive filtering is performed in decimated frequency bands at reduced computational cost [8, 9, 10]. For simplicity, the presentation will be restricted to a post-equalizer structure. Sec. 2 will review adaptive multichannel inversion. In Sec. 3, we discuss advantages and limitations of subband adaptive filtering and its application to the multichannel problem. Finally, simulation results are presented Sec. 4.

2. MULTICHANNEL EQUALIZATION

For an acoustic system with L loudspeakers and M microphones, the MIMO transfer function is described by a matrix $\mathbf{C}(z) \in \mathbb{R}^{M\times L}_{(z)}$,

$$\mathbf{C}(z) = \left[\begin{array}{ccc} C_{0,0}(z) & \dots & C_{0,L-1}(z) \\ \vdots & \ddots & \vdots \\ C_{M-1,0}(z) & \dots & C_{M-1,L-1}(z) \end{array} \right] . \qquad (1)$$

A first necessary condition for its invertibility is that the matrix $\mathbf{C}(z)$ does not any spectral zeroes common to all its polynomial elements [3]. Thereafter, depending on the relation between L amd M, either pre- or post-equalization can be established by a second MIMO system $\mathbf{W}(z) \in \mathbb{R}^{L\times M}_{(z)}$ defined analogously to $\mathbf{C}(z)$.

2.1. Equalization Problem

If the condition $L \leq M$ is satisfied, the configuration for the inversion problem is depicted in Fig. 2. After equalization, the outputs $\hat{y}_l[n]$ of the overall MIMO system $\mathbf{S}(z)$,

$$\mathbf{S}(z) = \mathbf{W}(z) \cdot \mathbf{C}(z) \overset{!}{=} z^{-\Delta}\mathbf{I}_{L\times L} \qquad (2)$$

should only be a delayed version of the inputs $y_l[n]$. Hence, reverberation and cross-talk effect have been removed. For the optimal inverse system $\mathbf{W}(z)$, the minimum-norm solution is provided by the left pseudo-inverse of $\mathbf{C}(z)$,

$$\mathbf{W}_{\mathrm{LS}}(z) = \left(\tilde{\mathbf{C}}(z)\mathbf{C}(z)\right)^{-1}\tilde{\mathbf{C}}(z)\ . \tag{3}$$

where $\tilde{\mathbf{C}}(z)$ is the parahermitian of $\mathbf{C}(z)$ [11]. In general, $\mathbf{W}_{\mathrm{LS}}(z)$ will be non-causal, hence the inclusion of a sufficient delay of Δ samples in (2). However, here we are interested in an adaptive solution to this problem, which will be discussed in the following.

2.2. Adaptive Multichannel Equalization

The multichannel adaptive equalization setup for the lth output channel is shown in Fig. 3. Each adaptive filter in the multichannel arrangement is fed by one of the M microphone signals, $x_m[n]$. The structure then produces an output $\hat{y}_l[n]$, which is compared to a version of the lth loudspeaker signal delayed by Δ samples, $y_l[n-\Delta]$. The difference is defined as the lth error signal:

$$\begin{aligned} e_l[n] &= y_l[n-\Delta] - \hat{y}_l[n] && (4) \\ &= y_l[n-\Delta] - \sum_{m=0}^{M-1} \mathbf{w}_{l,m}^H[n] \cdot \mathbf{x}_m[n] && (5) \end{aligned}$$

Note that using vector notation, the filter output is given as the inner product between a vector $\mathbf{x}_m[n]$ of length L_f holding the present and past input values,

$$\mathbf{x}_m^T[n] = \left[x_m[n], x_m[n-1], \cdots x_m[n-L_f+1]\right] \tag{6}$$

and the vector $\mathbf{w}_{l,m}[n]$, which holds the L_f complex conjugate coefficients of the adaptive filter at time n. The complex conjugation is for notational ease, and leads to the following one-sample gradient estimate for the mean squared error (MSE) [12]

$$\frac{\partial |e_l[n]|^2}{\partial \mathbf{w}_{l,m}^*[n]} = 2 \cdot e_l^*[n] \cdot \frac{\partial e_l[n]}{\partial \mathbf{w}_{l,m}^*[n]} = -2e_l^*[n]\mathbf{x}_m[n]. \tag{7}$$

With this gradient estimate, the update for the mth filter in Fig. 3 using the multichannel least mean square (M-LMS) update is now given by

$$\mathbf{w}_{l,m}[n+1] = \mathbf{w}_{l,m}[n] + 2\mu \cdot e_l[n] \cdot \mathbf{x}_m[n]\ . \tag{8}$$

2.3. Problems

In total, L of the filter arrangements of Fig. 3 are required to perform the task set in Sec. 2. This results in a complexity of

$$C_f = LM \cdot 2L_f \tag{9}$$

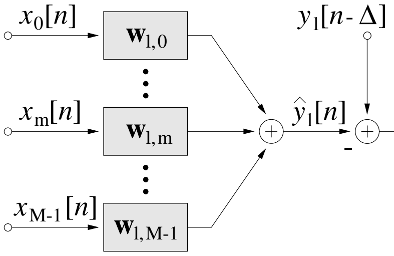

Figure 3: Multichannel adaptive filter for equalization of lth output channel, $y_l[n]$.

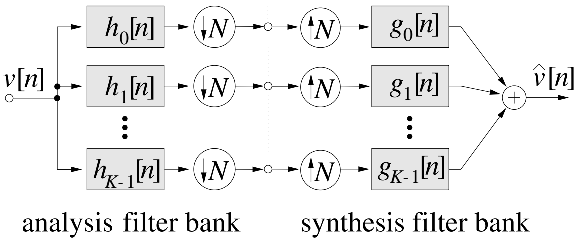

Figure 4: Decomposition of a signal $v[n]$ by an analysis bank into K decimated subbands, and reconstruction of a fullband signal $\hat{v}[n]$ by a synthesis bank.

multiply-accumulate (MAC) operations per sampling period. If the MIMO system $\mathbf{C}(z)$ is not well behaved, inverse filters of considerable length may be required, resulting in a very large computational complexity.

A large filter length L_f will also reduce the convergence speed. Additionally, the convergence is slowed down due to colouredness of the input signals to the adaptive filters, and the correlation between the inputs to the multichannel algorithm [12]. The latter is a result from the cross-talk occurring in the system $\mathbf{C}(z)$. Both the large complexity and the slow convergence set the motivation for the application of a subband adaptive filter approach.

3. SUBBAND ADAPTIVE FILTERING

This section explores some advantages and limitations of subband adaptive filtering (SAF), before it is applied to adaptive multichannel equalization.

3.1. Complex Oversampled SAFs

For SAF, all signals passed to the adaptive structure are split into K frequency bands decimated by a factor $N \leq K$ by analysis filter banks. Adaptive filters then operate on these subband signals. Due to the lower sampling rate, this leads to computational savings by a factor of $\propto K/N^2$ for LMS type algorithms [13]. Further advantages include the prewhitening property, which can bring a considerable increase in convergence speed for LMS-type algorithms, whos convergence rate depends on the eigenvalue spread of the input signal [12]. A further advantage is the parallelization of the subband processes.

From the outputs of the subband adaptive filters, fullband signals can be reconstructed by means of a synthesis bank. The two types of filter banks are shown in Fig. 4. Together, analysis and synthesis filter bank should only represent a delay, i.e. $\hat{v}[n] = v[n-\Delta]$.

Here, our filter banks are created from complex modulation of a prototype lowpass filter; this offers savings in terms of memory and can exploit a very inexpensive polyphase implementation of the filter bank operations [13]. An example for the characteristic of a $K = 16$ channel filter bank for decimation by $N = 14$ is shown in Fig. 5. The first 8 subbands cover the frequency range $\Omega = [0;\pi]$, and are sufficient to be processed if the fullband signals are real valued, since the remaining subbands are only complex conjugate versions and therefore redundant.

For the filter banks considered here, the prototype filter design affects the overall system accuracy and the achievable minimum MSE. The latter error can be linked to the aliasing present in the

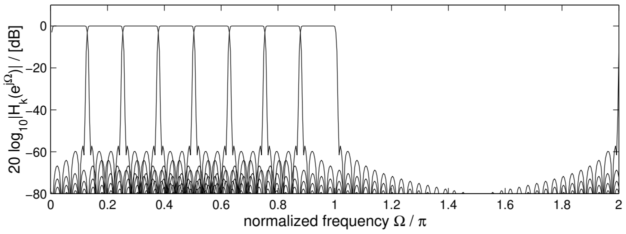

Figure 5: Example of a $K = 16$ modulated filter bank.

subband signals due to decimation and non-ideal filter banks. For modulated filter banks, a good approximation of the lower bound for the minimum MSE is given by the stopband attenuation of the prototype lowpass filter [14]. The first is limited by the reconstruction error, i.e. the deviation of the overall system in Fig. 4 from a perfect delay. Both errors can be traded off by application specific design of the prototype filter.

3.2. Subband Adaptive Equalization

Applying the subband approach to the multichannel equalization problem in Fig. 3, the structure shown in Fig. 6 results. There, the kth subband of each microphone signal $x_m[n]$ is passed to a separate multichannel adaptive filter (here the M-LMS summarized in Sec. 2.2), which tries to match the output to the kth subband of the lth desired signal, $y_l[n-\Delta]$.

The computational complexity of the subband adaptive equalizer structure comprising L of the blocks shown in Fig. 6 for each signal $\hat{y}_l[n]$ results in

$$C_s = LM\frac{4K}{N^2}L_f + (2L+M)\frac{1}{N}\big(4Klog_2K+4K+L_p\big)$$

MAC operations. The first term is the complexity of the adaptive filters, the second term describes the computations required for filter bank operations, where L_p is the length of the prototype filter. This term includes the analysis banks for the M input and L desired signals, and the L synthesis banks for the reconstruction of the equalized signals. The complexity of analysis and synthesis is identical, and is described in [13] for a very low cost implementation. However note, that the saving over (9) only takes effect if the filter length L_f of the fullband system is very large.

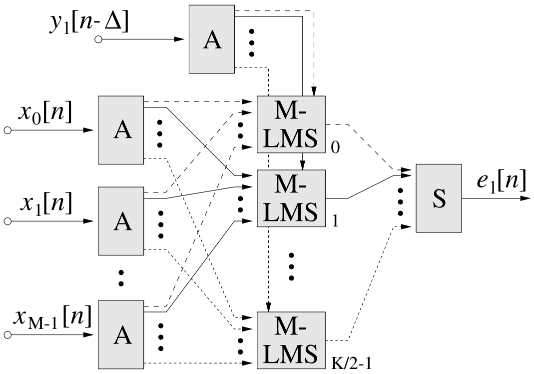

Figure 6: Multichannel subband adaptive filter for equalization of lth output.

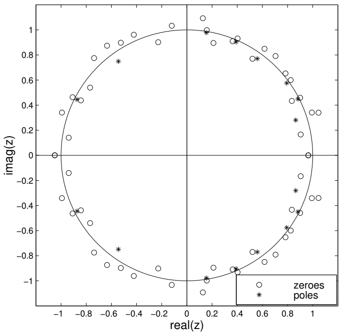

Figure 7: Pole-zero plot of the transfer path $C_{1,1}(z)$ contained in the MIMO system $\mathbf{C}(z)$.

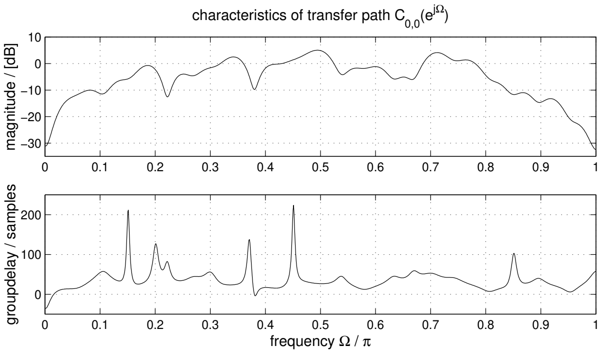
characteristics of transfer path $C_{0,0}(e^{j\Omega})$

Figure 8: Magnitude response and group delay of one transfer path $C_{1,1}(z)$ in the MIMO system $\mathbf{C}(z)$.

4. SIMULATIONS

To evaluate the potential benefits of the subband approach, we use the stereo setup in Fig. 1 for a simulation. Fig. 8 shows the characteristics of a simulated audio channel $C_{1,1}(z)$, with a general bandpass behaviour and strong dynamics in the spectrum and in the group delay. The remaining 3 transfer paths in $\mathbf{C}(z)$ exhibit characteristics of similar severity.

All 4 filters in the adaptive MIMO system $\mathbf{W}(z)$ are adapted at the same time using statistically independent loudspeaker signals $y_l[n]$. The fullband equalizer $\mathbf{W}(z)$ has a filter length of $L_f = 1120$ for each filter, while the subband system uses $L_f/N = 80$ tap filters in $K/2 = 8$ channels decimated by $N = 14$ employing the filter bank depicted in Fig. 4.

The MSE learning curves are depicted in Fig. 9. The errors of the fullband system are indicated by solid line, the reconstructed fullband error of the SAF equalizer is shown dotted. Although both systems take long time to adapt and even after $1.5 \cdot 10^4$ iterations the steady state has not been reached, the SAF implementation converges considerably faster than the fullband filter. This is due to the reduced filter length and the separation of the input spectrum, which is given in the top diagram of Fig. 8 for white noise excitation, into frequency bands with reduced eigenvalue spread [13]. Further, the SAF equalizer only requires $C_f/C_s = 12\%$ of the computations necessary for the fullband system of identical modelling capabilities.

The impulse responses of the overall MIMO system $\mathbf{S}(z) =$

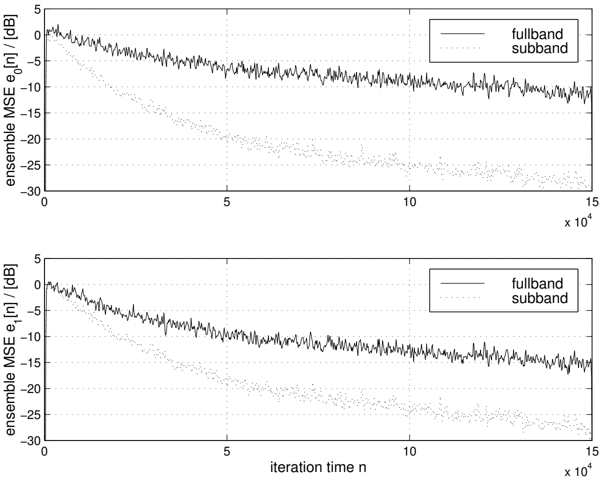

Figure 9: Ensemble MSE for the error signals $e_0[n]$ and $e_1[n]$ in both the fullband and subband simulation..