Weiss, Stephan and Rice, Garrey and Stewart, Robert (1999) 

This version is available at https://strathprints.strath.ac.uk/32640/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
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 modeling comparisons for similar acoustic problems IIR filters showed no particular advantage over FIR systems [7].

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 \( C(z) \in \mathbb{R}^{M \times L} \),

\[
C(z) = \begin{bmatrix}
C_{0,0}(z) & \ldots & C_{0,L-1}(z) \\
\vdots & \ddots & \vdots \\
C_{M-1,0}(z) & \ldots & C_{M-1,L-1}(z)
\end{bmatrix}
\]

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

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 \( S(z) \),

\[
S(z) = W(z) \cdot C(z) = z^{-\Delta} I_{L \times L}
\]

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 $W(z)$, the minimum-norm solution is provided by the left pseudo-inverse of $C(z)$,

$$W_{LS}(z) = (\hat{C}(z)C(z))^{-1}\hat{C}(z).$$

(3)

where $\hat{C}(z)$ is the parahermhid of $C(z)$ [11]. In general, $W_{LS}(z)$ will be non-causal, hence the inclusion of a sufficient delay of $\Delta$ 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 $i$th 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 $\hat{y}_i[n]$, which is compared to a version of the $i$th loudspeaker signal delayed by $\Delta$ samples, $y_i[n-\Delta]$. The difference is defined as the $i$th error signal:

$$e_i[n] = y_i[n-\Delta] - \hat{y}_i[n]$$

(4)

and the vector $w_{i,m}[n]$, 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]

$$\frac{\partial |e_i[n]|^2}{\partial w_{i,m}^*[n]} = 2 \cdot e_i[n] \cdot \frac{\partial e_i[n]}{\partial w_{i,m}^*[n]} = -2e_i[n]x_m[n].$$

(7)

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

$$w_{i,m}[n+1] = w_{i,m}[n] + 2\mu \cdot e_i[n] \cdot x_m[n].$$

(8)

2.3. Problems

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

$$C_f = LM \cdot 2L_f$$

(9)

![Figure 3: Multichannel adaptive filter for equalization of $i$th output channel, $y_i[n]$.](image)

![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.](image)
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 recon-
struction 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 sepa-
rate 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, $y_d[n-L]$. The computational complexity of the subband adaptive equal-
izer structure comprising $L$ of the blocks shown in Fig. 6 for each
signal $y_d[n]$ results in

$$C_s = LM \frac{4K}{N^2} I_T + (2L+M) \frac{1}{N} (4K \log_2 K + 4K + L_p)$$

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 $L$ de-
sired 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 implementa-
tion. However, note, that the saving over (9) only takes effect if the
filter length $L_T$ of the fullband system is very large.

![Figure 5: Example of a K = 16 modulated filter bank.](image5)

![Figure 7: Pole-zero plot of the transfer path $C_{1,1}(z)$ contained in the MIMO system $C(z)$.](image7)

![Figure 8: Magnitude response and group delay of one transfer path $C_{1,1}(z)$ in the MIMO system $C(z)$.](image8)

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 char-
acteristics 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 $C(z)$ exhibit
characteristics of similar severity.

All 4 filters in the adaptive MIMO system $W(z)$ are adapted at
the same time using statistically independent loudspeaker signals
$y_d[n]$. The fullband equalizer $W(z)$ has a filter length of $L_T =
1120$ for each filter, while the subband system uses $L_T / N = 80$
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 $1.5 \cdot 10^6$ 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 spec-
trum, 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 $C_T/C_s = 12\%$
of the computations necessary for the fullband system of identical
modeling capabilities.

The impulse responses of the overall MIMO system $S(z) =
by placing the inverse system \( W(z) \) in front of the system \( C(z) \), which requires the use of at least as many loudspeakers as microphones, \( L \geq M \) for a viable solution and offers a number of interesting applications [5]. Adaptive methods for this inverse problem are discussed in e.g. [4], and require filtered-X type LMS algorithms. Due to these filter terms in the algorithm, the computational complexity of the fullband system is much larger than for the case presented in this paper. Hence the development of pre-equalizing SAF techniques also appears very attractive.

6. ACKNOWLEDGMENT

The authors would like to gratefully acknowledge Dr. Ian Proudlcr of DERA, Malvern, England, who partially supported this work.

7. REFERENCES