

SUBBAND METHOD FOR MULTICHANNEL LEAST SQUARES EQUALIZATION OF ROOM TRANSFER FUNCTIONS

Nikolay D. Gaubitch, Mark R. P. Thomas and Patrick A. Naylor

Imperial College London
Exhibition Road, SW7 2AZ, London, UK
{ndg,mrt102,p.naylor}@imperial.ac.uk

ABSTRACT

Equalization of room transfer functions (RTFs) is important in many speech and audio processing applications. It is a challenging problem because RTFs are several thousand taps long and non-minimum phase and in practice only approximate measurements of the RTFs are available. In this paper, we present a subband multichannel least squares method for equalization of RTFs which is computationally efficient and less sensitive to inaccuracies in the measured RTFs compared to its fullband counterpart. Experimental results using simulated impulse responses demonstrate the performance of the algorithm.

1. INTRODUCTION

Equalization of room transfer functions (RTFs) is essential in several applications in acoustic signal processing, including speech dereverberation [1] and sound reproduction [2]. Although, in theory, exact equalization is possible when multiple observations are available [2], there are many obstacles for RTF equalization in practice [3].

Consider the L -tap room impulse response of the acoustic path between a source and the m th microphone in an M -element microphone array, $\mathbf{h}_m = [h_{m,0} \ h_{m,1} \ \dots \ h_{m,L-1}]$, with a z -transform $H_m(z)$ constituting the RTF. Equalization can be achieved, in principle, by an inverse system with transfer function $G_m(z)$ satisfying

$$G_m(z)H_m(z) = \kappa z^{-\tau}, \quad m = 1, 2, \dots, M \quad (1)$$

where τ and κ are arbitrary delay and scale factors respectively. Equivalently, considering the L_i tap impulse response of $G_m(z)$, $\mathbf{g}_m = [g_{m,0} \ g_{m,1} \ \dots \ g_{m,L_i-1}]^T$, (1) can be written in the time domain as

$$\mathbf{H}_m \mathbf{g}_m = \mathbf{d}, \quad (2)$$

where \mathbf{H}_m is a $(L + L_i - 1) \times L_i$ convolution matrix, and $\mathbf{d} = [\underbrace{0 \ \dots \ 0}_{\tau} \ \kappa \ 0 \ \dots \ 0]^T$ is the $(L + L_i - 1) \times 1$ output vector of the equalized RTFs. The problem of equalization is to find $G_m(z)$.

In practice, RTF equalization is not straightforward since: (i) RTFs are non-minimum phase [4] and do not give a stable causal solution for $G_m(z)$; (ii) the average difference between maxima and minima in RTFs are in excess of 10 dB [3, 5] and therefore RTFs typically contain spectral nulls that, after equalization, give strong peaks in the spectrum causing narrow band noise amplification; (iii) equalization filters designed from inaccurate estimates of $H_m(z)$ will cause distortion in the equalized signal [3]; (iv) the length L of \mathbf{h}_m can be several thousand taps in length [5].

Several methods for RTF equalization have been proposed. Single channel methods [4, 6, 7] typically result in large processing delay, which is problematic for many communications applications, extremely long and non-causal inverse filters, and provide only approximate equalization [2]; inherently these only partially equalize deep spectral nulls, which makes them less sensitive to noise and inexact RTF estimates [1]. In the multichannel case, the non-minimum phase problem is eliminated and exact inversion can be achieved [2, 8]. However, it has been observed that exact equalization is of limited value in practice, when the RTF estimates contains even moderate errors [1, 3, 9]. Various alternatives have been proposed for improving robustness to RTF inaccuracies [10, 11, 12].

In this paper, we introduce a new method for equalization filter design. Given a set of multichannel RTFs, we decompose the RTFs into their subband equivalent filters. These are then used to design the subband inverse filters and the equalization is performed in each subband before a fullband equalized signal is reconstructed. It is shown that this approach not only reduces the computational load, but also reduces the sensitivity to estimation errors and the effect of measurement noise in the RTFs. An important result is that this method accommodates multichannel equalization of large order systems, taking advantage of the shorter length of multichannel equalization filters and the low sensitivity to RTF inaccuracies of single channel methods.

The remainder of the paper is organized as follows. Fullband multichannel least squares (LS) equalization is reviewed in Section 2. The subband multichannel LS method is described in Section 3. In Section 4, experimental results are given and conclusions are drawn in Section 5.

2. MULTICHANNEL LS EQUALIZATION

In the multichannel case exact inversion can be achieved using Bezout's theorem [2, 8]: given a set of M RTFs, $H_m(z)$, and assuming that these do not have any common zeros, a set of filters, $G_m(z)$, can be found such that [2, 8]

$$\sum_{m=1}^M H_m(z)G_m(z) = 1. \quad (3)$$

The relation in (3) can be written in the time domain using (2) with $\tau = 0$ and $\kappa = 1$ as

$$\sum_{m=1}^M \mathbf{H}_m \mathbf{g}_m = \mathbf{H} \mathbf{g} = \mathbf{d}, \quad (4)$$

where $\mathbf{H} = [\mathbf{H}_1 \ \mathbf{H}_2 \ \dots \ \mathbf{H}_M]$, and $\mathbf{g} = [\mathbf{g}_1^T \ \mathbf{g}_2^T \ \dots \ \mathbf{g}_M^T]^T$.

An optimization problem can be formulated then as

$$\hat{\mathbf{g}} = \arg \min_{\mathbf{g}} \|\mathbf{H}\mathbf{g} - \mathbf{d}\|^2, \quad (5)$$

and the multichannel least squares (MCLS) equalization filters can be calculated according to [8]

$$\hat{\mathbf{g}} = \mathbf{H}^+ \mathbf{d}, \quad (6)$$

where $\mathbf{H}^+ = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ is the matrix pseudo-inverse. The length of the inverse filters is given here by

$$L_i = \left\lceil \frac{L-1}{M-1} \right\rceil, \quad (7)$$

where $\lceil \cdot \rceil$ denotes rounding to the next highest integer. In the special case when (7) gives an integer result, the solution is equivalent to that of MINT [2].

3. SUBBAND MULTICHANNEL LS EQUALIZATION

Multirate processing [13] has been applied successfully in many signal processing areas including acoustic echo cancellation [14, 15], but has been explored much less for acoustic equalization. A subband version of MINT was first investigated in [16]; this was an experimental approach where the subband filters were determined using a reference signal and the relation between fullband and subband impulse responses was not specifically considered. An adaptive method for multichannel equalization in oversampled subbands was demonstrated in [17] and was shown to provide significant improvement over the fullband counterpart.

The relation between fullband and subband filtering has been studied for filtering of MPEG audio signals [18], acoustic echo cancellation [15] and an AR model of the room impulse response [19]. Reilly et al [15] show that good approximations can be obtained with a diagonal filtering matrix, involving only one filter per subband for complex oversampled filterbanks. We now extend this approach to the multichannel case with application to RTF equalization.

3.1. Oversampled Filterbanks

The generalized discrete Fourier transform (GDFT) filterbank [14] is employed in the subsequent development work. The advantages of this filterbank include straightforward implementation of fractional oversampling and computationally efficient implementations [14]. The filterbank used for the illustrative experiments in this paper uses $K = 32$ subbands and decimation factor, $N = 24$. An $L_{\text{pr}} = 512$ -tap prototype filter was designed using the iterative least squares method [14], giving an estimated aliasing suppression of 82 dB. From the characteristics of the oversampled GDFT, the following two properties can be assumed to be valid:

P1: Aliasing is sufficiently suppressed in the subbands

$$U_k(zW_N^i)V_k(z) \approx 0, \quad i > 1, \forall k, \quad (8)$$

where $W_N = e^{-j2\pi/N}$.

P2: Magnitude distortion of the filterbank is negligible

$$\sum_{k=0}^{K/2-1} U_k(z)V_k(z) \approx \kappa z^{-\tau}, \quad (9)$$

where $U_k(z)$ and $V_k(z)$ are the z -transforms of the subband analysis and synthesis filters respectively.

3.2. Subband Decomposition

Consider a K subband, M microphone system. In order to design the subband equalizers $G'_{km}(z)$, the subband RTFs $H'_{km}(z)$ must be found using, for example, complex subband decomposition [15]. The objective of the subband decomposition is to find a set of subband filters, $H'_{km}(z)$, $k = 0, 1, \dots, K/2 - 1$, given the fullband filter $H_m(z)$, such that the total transfer function of the filter bank, $F_m(z)$, is equivalent to the that of the fullband filter up to an arbitrary scale factor, κ , and an arbitrary delay, τ . This can be written

$$F_m(z) = \kappa z^{-\tau} H_m(z), \quad \forall m. \quad (10)$$

The total transfer function of the filterbank for the m th channel is given by

$$F_m(z) = \frac{1}{N} \sum_{k=0}^{K/2-1} \sum_{i=0}^{N-1} U_k(zW_N^i)H'_{mk}(z^N)V_k(z), \quad (11)$$

Evoking property P1 in (8), the filterbank transfer function reduces to

$$F_m(z) \approx \frac{1}{N} \sum_{k=0}^{K/2-1} U_k(z)H'_{mk}(z^N)V_k(z), \quad (12)$$

which allows for a single filter per subband.

Next, following the approach in [15], we choose the filters in each subband, $H'_{mk}(z)$, such that they satisfy the relation

$$U_k(z)H'_{mk}(z^N) = U_k(z)H_m(z), \quad \forall k. \quad (13)$$

Substituting, (13) into (12) and due to property P2 in (9), the overall filterbank transfer function is

$$F_m(z) \approx \frac{\kappa}{N} z^{-\tau} H_m(z), \quad (14)$$

which is the desired result. Thus, the remaining problem is to solve for $H'_{mk}(z)$ in (13). Decimating (13) by a factor of N , the following approximation can be formed

$$\sum_{i=0}^{N-1} U_k(z^{1/N}W_N^i)H'_{mk}(z) \approx \sum_{i=0}^{N-1} U_k(z^{1/N}W_N^i)H_m(z^{1/N}W_N^i), \quad (15)$$

The estimates of the L' -tap subband filters $\mathbf{h}'_{mk} = [h'_{mk,0} \ h'_{mk,1} \ \dots \ h'_{mk,L'-1}]^T$ are then found by solving the following optimization problem [15]

$$\hat{\mathbf{h}}'_{km} = \arg \min_{\mathbf{h}'_{km}} \|\mathbf{U}_{N,k} \mathbf{h}'_{km} - \mathbf{r}_{N,km}\|^2. \quad (16)$$

where $\mathbf{r}_{N,km} = [r_{km,0} \ r_{km,N} \ \dots \ r_{km,N(L-1)}]^T$ is a vector with $r_{km,i} = h_{m,i} * u_{k,i}$ and $\mathbf{U}_{N,k}$ is the convolution matrix of the decimated subband filter. The length of \mathbf{h}'_{km} is

$$L' = \left\lceil \frac{L + L_{\text{pr}} - 1}{N} \right\rceil - \left\lceil \frac{L_{\text{pr}}}{N} \right\rceil + 1 \quad (17)$$

The k th subband, m th channel filters are calculated in the least squares optimal sense according to

$$\hat{\mathbf{h}}'_{km} = \mathbf{U}_{N,k}^+ \mathbf{r}_{N,km}. \quad (18)$$

Thus, given a fullband RTF, $H_m(z)$, and $K/2$ subband filters satisfying near perfect reconstruction and aliasing suppression in the subbands, a set of subband filters, $H'_{mk}(z)$, of the order L/N , can be found such that the overall subband transfer function is equivalent to the fullband filter response, which can result in significant order reduction of the very long room impulse responses.

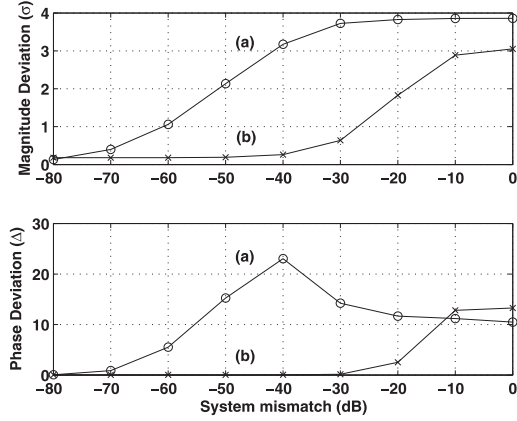


Figure 1: Magnitude and phase deviation vs. system mismatch for (a) inverse filtering with the fullband MCLS equalization and (b) SB-MCLS equalization of random impulse responses.

3.3. Subband equalization

The multichannel equalization filters, $\hat{G}'_{km}(z)$, are calculated for each subband using the filters $\hat{H}'_{km}(z)$ obtained from (18). Here, this is done utilizing the multichannel LS filter design from (6), which now becomes

$$\hat{\mathbf{g}}'_k = \hat{\mathbf{H}}_k^+ \mathbf{d}, \quad k = 0, 1, \dots, \frac{K}{2} - 1. \quad (19)$$

Equalization is then achieved by applying the inverse filters, $\hat{\mathbf{g}}'_k$, to the subband signals of the reverberant observations in each subband k , $\forall k$ and an equalized fullband signal is constructed. Assuming that exact equalization is achieved in each subband, the accuracy of the final result will depend on the reconstruction properties of the filterbank, the accuracy of aliasing suppression and, consequently, on the design of the prototype filter.

4. EXPERIMENTAL RESULTS

Simulation results are presented to demonstrate the performance of the proposed SB-MCLS equalization method. The experiments aim to show: (i) a comparative performance evaluation with the fullband MCLS using randomly generated channels and (ii) the applications of the SB-MCLS to simulated room impulse responses.

To study the design of an equalization filter for \mathbf{h}_m using inexact RTF estimates, we define an inexact system impulse response, $\tilde{\mathbf{h}}_m$, as an impulse response with system mismatch $\mathcal{M}_m > -\infty$ dB, with

$$\mathcal{M}_m = 20 \log_{10} \left(\frac{\|\mathbf{h}_m - \tilde{\mathbf{h}}_m\|}{\|\mathbf{h}_m\|} \right) \text{ dB}, \quad (20)$$

where $\|\cdot\|$ denotes Euclidean distance. We then model system mismatch, as in [20], according to

$$\tilde{\mathbf{h}}_m = (\mathbf{I} + \mathcal{E}_m) \mathbf{h}_m, \quad (21)$$

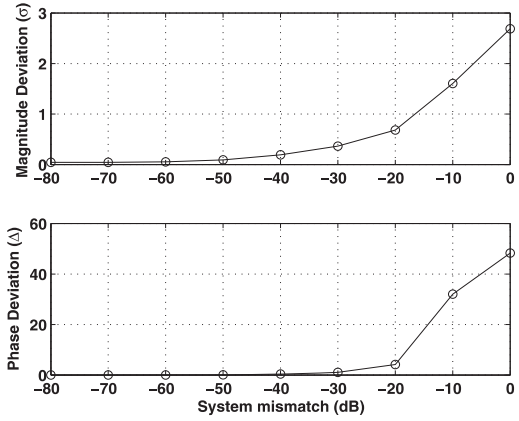


Figure 2: Magnitude and phase deviation vs. system mismatch for SB-MCLS equalization of simulated room impulse responses.

where $\mathcal{E}_m = \text{diag}\{\epsilon_{m,0} \epsilon_{m,1} \dots \epsilon_{m,L-1}\}$, \mathbf{I} is the identity matrix, and $\epsilon_{m,i}$ is a zero mean Gaussian variable with the variance set to the desired system mismatch.

For evaluation purposes we consider the magnitude and the phase separately. We define the equalized system $\hat{\mathbf{d}} = \mathbf{H}\hat{\mathbf{g}}$ with I -point discrete Fourier transform $\hat{D}(f) = |\hat{D}(f)|e^{j\theta(f)}$. The following measures are then used:

(i) *Magnitude deviation* is defined here as the standard deviation of the equalized magnitude response [3]

$$\sigma = \sqrt{\frac{1}{I} \sum_{f=0}^{I-1} \left(10 \log_{10} |\hat{D}(f)| - \bar{D} \right)^2}, \quad (22)$$

with $\bar{D} = \frac{1}{I} \sum_{f=0}^{I-1} 10 \log_{10} |\hat{D}(f)|$. This measure is scaling independent and equals zero for exact equalization.

(ii) *Linear phase deviation* is defined as the deviation of the unwrapped phase from a linear fit to its values and is defined here as

$$\Delta = \sqrt{\frac{1}{I} \sum_{f=0}^{N-1} (\theta(f) - \bar{\theta}(f))^2}, \quad (23)$$

where $\bar{\theta}(f)$ is the least squares linear approximation to the phase at frequency bin f .

Experiment 1: This experiment demonstrates the performance of the SB-MCLS equalizer, compared with the fullband MCLS. A system with $M = 5$ randomly generated $L = 512$ -tap channels was used and system misalignment, \mathcal{M}_m varying between 0 and -80 dB was simulated with (21). The results averaged over 100 different channel realizations are shown in Fig. 1 for (a) the fullband MCLS and (b) for the proposed subband implementation. Notably, the SB-MCLS exhibits much gentler performance degradation with increased misalignment in comparison with the fullband MCLS; the SB-MCLS method is shown in these results to be less sensitive to inexact impulse responses, while benefiting from the shorter filters of multichannel equalization. This improvement is a consequence of the reduced filter length in the subbands.

Experiment 2: We now demonstrate the performance of the SB-MCLS equalizer for simulated RTFs. We have not compared these outcomes with the fullband case since equalization of channels of the length considered is not feasible with current known methods. A linear array of $M = 5$ uniformly distributed microphones with 0.1 m separation between adjacent sensors was simulated using the source-image method [21] for a room with dimensions $6.4 \times 5 \times 4$. The channel lengths are $L = 4800$ taps, which is equivalent to $T_{60} = 0.3$ s at $f_s = 16$ kHz sampling frequency. Moreover, keeping the source-microphone configuration fixed, RTFs were simulated at 10 different locations in the room. Figure 2 shows the results in terms of magnitude and phase distortion, as an average of the 10 measurement locations. This again shows a similar performance pattern as in the previous experiment. In addition, nearly perfect equalization is achieved with the SB-MCLS method for $\mathcal{M}_m \leq -40$ dB.

5. CONCLUSIONS

Exact equalization with no delay and with inverse filters of similar order to the room impulse responses is possible in the multi-microphone case, even for non-minimum phase RTFs. However, multichannel least squares methods are very sensitive to inaccuracies in the estimated systems to be equalized, causing significant distortions to the equalized signal. Consequently, a new method for equalization of RTFs was derived operating on decimated oversampled subband signals, where the fullband impulse response is decomposed into equivalent filters in the subbands and multichannel least squares equalization is applied to each subband. Simulation results were presented to evaluate the performance of this method and equalization of channels of several thousand taps was demonstrated; the experimental results indicated that the new method is more robust to errors in the impulse responses, which is due to a combination of shorter filters and approximation of the filtering in the subbands. Thus, the proposed subband multichannel least squares equalization benefits from the reduced sensitivity to channel estimation errors, shorter equalization filters, no delay due to the equalization, giving significant advantages over existing single and multichannel techniques.

6. REFERENCES

- [1] P. A. Naylor and N. D. Gaubitch, "Speech dereverberation," in *Proc. Int. Workshop Acoust. Echo Noise Control*, Eindhoven, The Netherlands, Sept. 2005.
- [2] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 2, pp. 145–152, Feb. 1988.
- [3] B. D. Radlović, R. C. Williamson, and R. A. Kennedy, "Equalization in an acoustic reverberant environment: Robustness results," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 8, no. 3, pp. 311–319, May 2000.
- [4] S. T. Neely and J. B. Allen, "Invertibility of a room impulse response," *J. Acoust. Soc. Amer.*, vol. 66, no. 1, pp. 165–169, July 1979.
- [5] H. Kuttruff, *Room Acoustics*, 4th ed. Taylor & Francis, Oct. 2000.
- [6] J. Mourjopoulos, P. Clarkson, and J. Hammond, "A comparative study of least-squares and homomorphic techniques for the inversion of mixed phase signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 7, May 1982, pp. 1858–1861.
- [7] B. D. Radlović and R. A. Kennedy, "Nonminimum-phase equalization and its subjective importance in room acoustics," *IEEE Trans. Speech Audio Processing*, vol. 8, no. 6, pp. 728–737, Nov. 2000.
- [8] Y. Huang, J. Benesty, and J. Chen, "A blind channel identification-based two-stage approach to separation and dereverberation of speech signals in a reverberant environment," *IEEE Trans. Speech Audio Processing*, vol. 13, no. 5, pp. 882–895, Sept. 2005.
- [9] F. Talantzis, D. B. Ward, and P. A. Naylor, "Performance analysis of dynamic acoustic source separation in reverberant rooms," *IEEE Trans. Audio, Speech, Language Processing*, vol. 14, no. 4, pp. 1378–1390, July 2006.
- [10] S. Bharitkar, P. Hilmes, and C. Kyriakakis, "Robustness of spatial average equalization: A statistical reverberation model approach," *J. Acoust. Soc. Amer.*, vol. 116, no. 6, pp. 3491–3497, Dec. 2004.
- [11] Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Trans. Speech Audio Processing*, vol. 2, no. 2, pp. 320–328, Apr 1994.
- [12] T. Hikichi, M. Delcroix, and M. Miyoshi, "On robust inverse filter design for room transfer function fluctuations," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Sept. 2006.
- [13] P. P. Vaidyanathan, *Multirate systems and filter banks*. Prentice Hall, 1993.
- [14] S. Weiss and R. W. Stewart, *On adaptive filtering in oversampled subbands*. Shaker Verlag, 1998.
- [15] J. P. Reilly, M. Wilbur, M. Seibert, and N. Ahmadvand, "The complex subband decomposition and its application to the decimation of large adaptive filtering problems," *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2730–2743, Nov. 2002.
- [16] K. Yamada, J. Wang, and F. Itakura, "Recovering of broad band reverberant speech signal by sub-band MINT method," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1991, pp. 969–972.
- [17] S. Weiss, G. W. Rice, and R. W. Stewart, "Multichannel equalization in subbands," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct. 1999, pp. 203–206.
- [18] C. A. Lanciani and R. W. Schafer, "Subband-domain filtering of MPEG audio signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 2, Mar. 1999, pp. 917–920.
- [19] J. R. Hopgood and P. J. W. Rayner, "A probabilistic framework for subband autoregressive models applied to room acoustics," in *Proc. IEEE Workshop Statistical Signal Processing*, Aug 2001, pp. 492 – 495.
- [20] J. H. Cho, D. R. Morgan, and J. Benesty, "An objective technique for evaluating doubletalk detectors in acoustic echo cancelers," *IEEE Trans. Speech Audio Processing*, vol. 7, no. 7, pp. 718–724, Nov. 1999.
- [21] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.