SCALE FACTOR AMBIGUITY CORRECTION FOR SUBBAND BLIND MULTICHANNEL IDENTIFICATION

Nikolay D. Gaubitch, Xiang (Shawn) Lin, Patrick A. Naylor

Department of Electrical and Electronic Engineering, Imperial College London, UK

ABSTRACT

Subband system identification has proved beneficial in applications such as acoustic echo cancellation. A problem for blind multichannel system identification in subbands is that the subband systems can only be identified correctly up to an arbitrary scale factor. This scale factor ambiguity is the same across all channels but can differ between the subbands and therefore, limits the usability of such estimates. We show how the relation between the subband and the full-band estimates, together with the full-band cross-relation error, can be used to resolve the scale factor ambiguity problem. The operation of this approach is demonstrated with simulated experiments.

Index Terms— Blind system identification, Oversampled filterbanks, Scale factor ambiguity

1. INTRODUCTION

Blind system identification (BSI) is of interest in several fields of engineering including communications, exploration seismology and speech and audio processing [1]. Within the area of speech and audio processing BSI is an important component for speech dereverberation [1][2][3]. The BSI problem can be stated as follows: consider a signal $s(n)$ which is produced in a multipath environment such as a reverberant room and transmitted to an array of M sensors at a distance from the source. The observed signal at the mth sensor is then given by

$$x_m(n) = \mathbf{h}_m^T \mathbf{s}(n) + \nu_m(n), \quad (1)$$

where $\mathbf{h}_m = [h_{m,0}\ h_{m,1}\ \dots\ h_{m,L-1}]^T$ is the L-tap impulse response between the source and the mth sensor, $\mathbf{s}(n) = [s(n)\ s(n-1)\ \dots\ s(n-L+1)]^T$ is the input signal vector, $\nu_m(n)$ is additive measurement noise and $[\cdot]^T$ denotes matrix transpose. The problem of BSI is to find the impulse responses $\mathbf{h}_m$ using only the observations $x_m(n)$.

Several second order statistics BSI algorithms are based on the cross-relation (CR) between two channels [4]

$$\mathbf{x}_m^T(n)\mathbf{h}_l = \mathbf{x}_l^T(n)\mathbf{h}_m, \quad m \neq l, \quad (2)$$

where $\mathbf{x}_m(n) = [x_m(n)\ x_m(n-1)\ \dots\ x_m(n-L+1)]^T$. Taking all the different channel combinations into consideration, a system of equations can be formulated as

$$\mathbf{R}\mathbf{h} = \mathbf{0}, \quad (3)$$

where $\mathbf{h} = [\mathbf{h}_1^T\ \mathbf{h}_2^T\ \dots\ \mathbf{h}_M^T]^T$ is a vector of the concatenated impulse responses and $\mathbf{R}$ is a correlation matrix of the input signal. The channel responses $\mathbf{h}$ can then be identified exactly, up to a scale factor α, by finding the eigenvector corresponding to the smallest eigenvalue (which is zero in the noise-free case) of $\mathbf{R}$. This solution can be achieved using, for example, singular value decomposition [2] or adaptive filters [5], provided that the following identifiability conditions are satisfied [4]: (i) that the autocorrelation matrix of the source signal is of full rank and (ii) that there are no common zeros between the channels.

There are several problems associated with BSI algorithms. First, the accuracy of the estimates is degraded when the observations are noisy and in the presence of common or near-common zeros between the channels. Secondly, long impulse responses, commonly occurring in acoustic signal processing, result in large computational complexity and lower the estimation performance due to increased occurrence of common or near-common zeros [6, 7].

One approach to improve the performance of system identification algorithms is to apply them in a subband framework; this has shown benefits in terms of computational complexity and performance improvement in applications such as acoustic echo cancellation [8]. A study of a subband BSI system was presented in [2], highlighting the benefit of the shorter channels to be estimated compared to the full-band case. In general, however, blind system identification in subbands has received much less attention than the non-blind case (for example, acoustic echo cancellation) and one reason for this is the scale factor ambiguity across the different subbands which limits the use of such approaches [2]. In this paper, we propose a method to correct this scale factor ambiguity by employing a cross-relation error between the different channels together with the reconstructed full-band impulse responses from the estimated channels in the subbands.

The remainder of this paper is organized as follows: Section 2 describes the subband blind multichannel identification method using an oversampled filterbank. In Section 3, a relationship between the full-band response and the subband responses is presented and is employed to develop the scale factor correction method. Simulation results are presented in Section 4 and conclusions from this work are drawn in Section 5.

2. SUBBAND BLIND MULTICHANNEL IDENTIFICATION

For this work we employ the generalized discrete Fourier transform (GDFT) filterbank with K subbands decimated by a factor of N. Although this design results in complex subband signals, for K even, only $K/2$ subbands need to be processed since the remaining subbands are complex conjugates of these. The advantages of the GDFT filterbank include straightforward implementation of fractional oversampling ($N \leq K$) and computationally efficient implementations [8].

Within the framework of the GDFT filterbank, the analysis filters, $u_{k,i}$, are calculated from a single $L_{\rm pr}$-tap prototype filter, p_i, with bandwidth $\frac{2\pi}{K}$ according to

$$u_{k,i} = p_i e^{j\frac{2\pi}{K}(k+k_0)(i+i_0)}, \quad i = 0, 1, \ldots, L_{\rm pr} - 1, \tag{4}$$

where we set the frequency and time offset terms to $i_0 = 0$ and $k_0 = 1/2$ as in [9]. A corresponding set of synthesis filters, $v_{k,i}$ satisfying near perfect reconstruction is obtained from the time-reversed, conjugated version of the analysis filters [8], $v_{k,i} = u^*_{k,L_{\rm pr}-i-1}$. The oversampled subband structure allows aliasing between adjacent subbands to be suppressed to a very low level (around -90 dB in our implementation); this facilitates a full-band transfer function, $H_m(z)$, to be related to a set of subband filters $H'_{mk}(z),\ k = 0, 1, \ldots, K/2 - 1$ with only one filter per subband [9].

Consequently, the blind system identification from (3) can be applied in each subband as

$$\mathbf{R}_k \mathbf{h}'_k = \mathbf{0}, \quad k = 0, 1, \ldots, K/2 - 1, \tag{5}$$

where $\mathbf{h}'_k = [\mathbf{h}'^T_{1k}\ \mathbf{h}'^T_{2k}\ \ldots\ \mathbf{h}'^T_{Mk}]^T$. Now, assuming that the identifiability conditions are satisfied, each subband estimate $\hat{\mathbf{h}}'_k$ will be determined up to a complex scale factor α_k such that

$$\hat{\mathbf{h}}_k = \alpha_k \mathbf{h}_k. \tag{6}$$

It is evident from (6) that for a particular k the scale factor will be the same across the M channels but for any particular channel will be different across the $K/2$ subbands. If the subband estimates are used to equalize the observed signal, the scaling discrepancy will propagate to the reconstructed full-band signal. Alternatively, if a full-band impulse response is reconstructed from the subband estimates, the reconstructed impulse response will be incorrect; this will be demonstrated by the experiments in Section 4.

3. SCALE FACTOR AMBIGUITY CORRECTION

We concluded Section 2 with the remark that reconstructing a full-band impulse response from subband filters with different scale factors will result in an erroneous full-band estimate. We now show how this fact can be used to our advantage in order to correct for the scale factors. We here define scale factor correction in terms of a correction term β_k that can take on any value which results in $\beta_0\alpha_0 = \beta_1\alpha_1 = \ldots = \beta_{K/2-1}\alpha_{K/2-1}$.

3.1. Subband to full-band filter reconstruction

The relationship between a full-band transfer function and a set of subband transfer functions can be found using the complex subband decomposition [9]. The objective of the subband decomposition is to find a set of subband filters, $H'_{mk}(z),\ k = 0, 1, \ldots, K/2 - 1$, given the full-band filter, $H_m(z)$, such that the total transfer function of the filter bank, $F_m(z)$, is equivalent to that of the full-band filter up to an arbitrary scale factor, κ, and an arbitrary delay, τ. This can be written

$$F_m(z) = \kappa z^{-\tau} H_m(z), \quad \forall m. \tag{7}$$

The L'-tap subband filters with impulse responses $\mathbf{h}'_{mk} = [h'_{mk,0}\ h'_{mk,1}\ \ldots\ h'_{mk,L'-1}]^T$ are then estimated by solving the following optimization problem [9]

$$\hat{\mathbf{h}}'_{mk} = \arg\min_{\mathbf{h}'_{mk}} ||\mathbf{U}_{N,k}\mathbf{h}'_{mk} - \mathbf{r}_{mk}||^2, \tag{8}$$

where $\mathbf{r}_{mk} = [r_{mk,0}\ r_{mk,1}\ \ldots\ r_{mk,\lceil (L+L_{\rm pr}-1)/N \rceil}]^T$ is a vector with $r_{mk,i} = (h_{m,i} * u_{k,i})_{\downarrow N}$, $\mathbf{U}_{N,k}$ is the convolution matrix of the downsampled subband filter, $(\cdot)_{\downarrow N}$ denotes downsampling by a factor of N, '*' is the linear convolution operator and $\lceil a \rceil$ denotes the ceiling operator giving the smallest integer greater than or equal to a. The length of $\mathbf{h}'_{mk}$ is

$$L' = \left\lceil \frac{L + L_{\rm pr} - 1}{N} \right\rceil - \left\lceil \frac{L_{\rm pr}}{N} \right\rceil + 1. \tag{9}$$

The mth channel, kth subband filters are calculated in the least squares optimal sense according to

$$\hat{\mathbf{h}}'_{mk} = (\mathbf{U}^T_{N,k}\mathbf{U}_{N,k})^{-1}\mathbf{U}^T_{N,k}\mathbf{r}_{mk}. \tag{10}$$

Inversely, the full-band response can be calculated given a set of subband responses using the relation

$$\hat{\mathbf{h}}_m = \Re\left\{\sum_{k=0}^{K/2-1} \mathbf{f}_{mk}\right\}, \tag{11}$$

where $\mathbf{f}_{mk} = [f_{mk,0}\ f_{mk,1}\ \ldots\ f_{mk,L-1}]^T$ with the ith element $f_{mk,i} = \left((u_{k,i})_{\downarrow N} * h'_{mk,i}\right)_{\uparrow N} * v_{k,i}$ and $\Re\{a\}$ denotes the real part of a.

3.2. Scale factor correction

It will be demonstrated with the simulation results in Section 4 that the reconstructed full-band system response in (11) is not correct when there are different scale factors in each subband. This leads to the key idea of this paper: from (2) we can formulate an error as in [5] but using the reconstructed full-band impulse responses from (11)

$$\begin{aligned} e_{lm} &= \mathbf{x}_m^T \hat{\mathbf{h}}_l - \mathbf{x}_l^T \hat{\mathbf{h}}_m \\ &= \mathbf{x}_m^T \Re \left\{ \sum_{k=0}^{K/2-1} \mathbf{f}_{lk} \right\} - \mathbf{x}_l^T \Re \left\{ \sum_{k=0}^{K/2-1} \mathbf{f}_{mk} \right\}, \end{aligned} \quad (12)$$

for $m, l = 1, 2, \ldots, M$ and $k = 0, 1, \ldots, K/2 - 1$. We have excluded the dependency of $\mathbf{x}_m$ on the discrete time variable n as any L-sample frame from the observed signals can be used in this case.

The error in (12) will be large when the channel estimates are incorrect due to the scale factors α_k. Consequently, this can be exploited to correct for the scale factors by minimizing this error. There are generally two possible ways to incorporate this into the subband blind system identification framework:

(i) as a constraint on the subband multichannel identification;

(ii) as a post-processing step after identification.

In this paper, we consider the second of these approaches. It is assumed that the subband channels have been estimated accurately (up to a scale factor) such that they satisfy the relation in (6). Then we introduce a parameter β_k which is to correct the effect of α_k. The error can now be written

$$e_{lm} = \mathbf{x}_m^T \Re \left\{ \sum_{k=0}^{K/2-1} \beta_k \alpha_k \mathbf{f}_{lk} \right\} - \mathbf{x}_l^T \Re \left\{ \sum_{k=0}^{K/2-1} \beta_k \alpha_k \mathbf{f}_{mk} \right\}. \quad (13)$$

Parameters α_k are unknown but we can control parameters β_k. In order to find β_k we would like to minimize the error in (13) using

$$\hat{\beta}_k = \arg\min_{\beta_k} \sum_{m=1}^{M-1} \sum_{l=m+1}^{M} e_{ml}^2, \; k = 0, 1, \ldots, K/2-1. \quad (14)$$

subject to

$$||\beta_k||^2 > 0, \quad \forall k.$$

It is interesting to note that (13) can be rearranged by moving $\mathbf{x}_m$ inside the summation which causes the scale factors to cancel. This is equivalent to introducing the signal into the channel reconstruction and demonstrates that minimizing the error in the subbands also minimizes the full-band error. However, this does not result in accurate full-band channel estimates and, therefore, the optimization problem in (14) needs to be solved by first reconstructing the full-band estimates and then calculating the error. A straightforward approach that has been found suitable for solving this problem is the Simplex method [10]. The scale factors are complex and are treated by the Simplex algorithm as two parameters per subband. Therefore, there are K free parameters to optimize.

4. SIMULATIONS AND RESULTS

We now present some simulation results to demonstrate the scale factor ambiguity correction algorithm. In particular two main features are investigated: (i) the effect of the variance of the scale factors across the different subbands and (ii) the effect of misalignment in the identified channels.

We use two different metrics in our evaluation procedure. First, the Normalized Projection Misalignment (NPM) is employed to measure the misalignment between two impulse responses (disregarding the scale factor) and is defined as [11]

$$\mathrm{NPM} = 20 \log_{10} \left(\frac{1}{||\mathbf{h}||} \left\| \mathbf{h} - \frac{\mathbf{h}^T \hat{\mathbf{h}}}{\hat{\mathbf{h}}^T \hat{\mathbf{h}}} \hat{\mathbf{h}} \right\| \right) \mathrm{dB}. \quad (15)$$

Second we use the normalized variance of the corrected scale factors, defined as

$$\xi = \frac{\mathrm{var}(\mathbf{A}\boldsymbol{\beta})}{||\mathbf{A}\boldsymbol{\beta}||^2}, \quad (16)$$

where $\mathbf{A} = \mathrm{diag}\{\alpha_0 \; \alpha_1 \; \ldots \; \alpha_{K/2-1}\}$ is a diagonal matrix with the true scale factors and $\boldsymbol{\beta} = [\beta_0 \; \beta_1 \; \ldots \; \beta_{K/2-1}]^T$ are the correction coefficients. If the parameters β_k correct the scale ambiguity such that the scale factor is uniform over all subbands, then $\xi = 0$.

The filterbank used for the following experiments uses $K = 8$ subbands with a decimation factor $N = 4$. An $L_{\mathrm{pr}} = 64$-tap prototype filter was designed using the iterative least squares method [8], giving an estimated aliasing suppression of 92 dB. The source signal was white Gaussian noise with a system of $M = 3$ randomly generated channel responses of length $L = 256$ whose tap values were drawn from a zero-mean Gaussian distribution. The equivalent subband filters were calculated using (10) and scale coefficients α_k were generated randomly with varying variances and applied to the subband channel components.

In the first experiment, we introduced scale factors with different variances and measured the NPM of the full-band reconstructed impulse responses ($\mathrm{NPM_f}$) before and after the scale factor ambiguity correction algorithm. Figure 1 shows the outcome of this experiment. It can be seen that the variation of the scale factors across the subbands has little effect. We also observe that, as expected, these scale factors cause the reconstructed full-band channels to be inaccurate with $\mathrm{NPM_f}$ close to 0 dB. Applying the scale factor ambiguity correction method resolves this issue, resulting in channels with $\mathrm{NPM_f}$ in the vicinity of -80 dB.

In the second experiment, we simulated various levels of misalignment in the subband filters (measured as $\mathrm{NPM_s}$), so that the mth channel in the kth subband is given by $\hat{\mathbf{h}}_{mk} = \alpha_k(\mathbf{I} + \mathcal{E}_{mk})\mathbf{h}_{mk}$, where $\mathbf{I}$ is the identity matrix, $\mathcal{E}_{mk} = \mathrm{diag}\{\epsilon_{mk,0} \; \epsilon_{mk,1} \; \ldots \; \epsilon_{mk,L'-1}\}$ and the variance of $\epsilon_{mk,i}$ is set according to the desired $\mathrm{NPM_s}$. We then

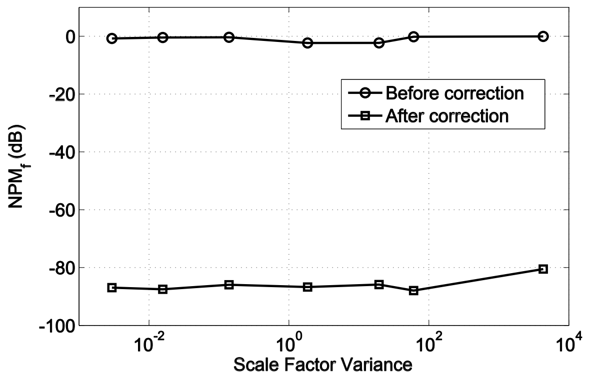

Fig. 1. Full-band NPM vs. scale factor variance before and after scale factor correction. The subband NPM is assumed to be $\mathrm{NPM_s} = -\infty$ dB.

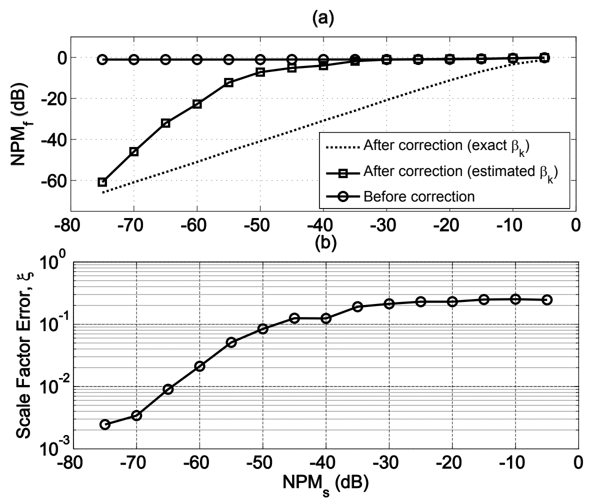

Fig. 2. Varying levels of subband NPM vs. (a) full-band NPM before correction, after correction with the proposed method and after correction with the ideal coefficients and (b) scale factor error.

used these subband channels to investigate the results obtained with the Simplex algorithm in terms of full-band misalignment, $\mathrm{NPM_f}$ and the variance of the corrected scale coefficients calculated using (16). The results, averaged over 100 scale factor realizations, are shown in Fig. 2. Figure 2a shows the subband NPM versus full-band NPM before and after scale factor correction and the 'ideal' case where α_k is known. Figure 2b shows the normalized variance of the corrected coefficients. We can deduce the following from these results: (i) the scalar factor estimation algorithm degrades with decreased $\mathrm{NPM_s}$; (ii) scalar factor correction has little effect at $\mathrm{NPM_s} < -10$ dB, even if the exact values were known; (iii) the proposed method operates with great accuracy at $\mathrm{NPM_s} \leq -50$ dB.

5. CONCLUSIONS

The scale factor ambiguity in subband multichannel blind system identification has been investigated. The relationship between a full-band transfer function and an equivalent set of subband transfer functions has been utilized to formulate a channel cross-relation error using the reconstructed full-band filters from subband estimates. It has then been shown how this error can be used to correct scale factors that differ between different subbands but not between different channels in the same subband. The error was minimized with the Simplex algorithm to find a set of scale factor correction coefficients. Simulation results showed that, although the algorithm's performance degrades with increased misalignment in the subband channel estimates, it accurately solves the scale factor ambiguity problem when the channel identification is good ($\mathrm{NPM_s} \leq -50$ dB).

6. REFERENCES

[1] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind System Identification," *Proc IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.

[2] S. Gannot and M. Monnen, "Subspace Methods for Multimicrophone Speech Dereverberation," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 11, pp. 1074–1090, 2003.

[3] 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, Sep. 2005.

[4] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.

[5] Y. Huang and J. Benesty, "A class of frequency-domain adaptive approaches to blind multichannel identification," *IEEE Trans. Signal Processing*, vol. 51, no. 1, pp. 11–24, Jan. 2003.

[6] X. Lin, N. D. Gaubitch, and P. A. Naylor, "Two-stage blind identification of SIMO systems with common zeros," in *Proc European Signal Processing Conf*, Florence, Italy, Sept. 2006.

[7] A. W. H. Khong, X. Lin, and P. A. Naylor, "Algorithms for identifying clusters of near-common zeros in multichannel blind system identification and equalization," in *Proc IEEE Intl Conf Acoustics, Speech and Signal Processing*, Apr. 2008.

[8] S. Weiss and R. W. Stewart, *On adaptive filtering in oversampled subbands*, Shaker Verlag, 1998.

[9] 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.

[10] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research*, McGraw-Hill, 8 edition, 2005.

[11] D. R. Morgan, J. Benesty, and M. M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Processing Lett.*, vol. 5, no. 7, pp. 174–176, July 1998.