A NOISE-ROBUST DUAL FILTER APPROACH TO MULTICHANNEL BLIND SYSTEM IDENTIFICATION

Rehan Ahmad, Nikolay D. Gaubitch and Patrick A. Naylor

Department of Electrical and Electronic Engineering, Imperial College London
Exhibition Road, London, SW7 2AZ, United Kingdom
phone: + (44) 20759 46235, fax: + (44) 20759 46302, email: {rehan.ahmad02, nikolay.gaubitch, p.naylor}@imperial.ac.uk
web: www.commsp.ee.ic.ac.uk/sap

ABSTRACT

We present a novel frequency domain blind multichannel identification approach. This is a dual filter based approach comprising a background and a foreground filter. The foreground filter tracks the changes in the cost function of the background filter and employs an update decision criterion making it robust to observation noise and facilitating tracking of changes in the unknown system. Simulation results for both speech and white Gaussian noise input signals are presented to illustrate the algorithm’s performance for acoustic systems.

1. INTRODUCTION

Blind channel identification (BCI) for single-input multiple-output (SIMO) systems is an important technique with a variety of potential applications in acoustics, wireless communications and other signal processing systems. Techniques for BCI based upon second order statistics [1] [2] and higher order statistics [3] have been studied. In addition, multichannel BCI techniques are becoming feasible for applications. The normalized multichannel frequency domain LMS (NMCFLMS) algorithm [4] has been shown to be effective in identifying room impulse responses which are of particular interest in acoustic dereverberation. However, NMCFLMS lacks robustness to additive observation noise and can suffer from divergence even in moderate noise conditions. This has been studied in [5] [6].

We propose a novel approach to blind multichannel identification which is robust to the presence of additive noise in the system. Dual-filter methods have been used successfully to achieve double-talk robustness in echo cancellation systems [7]. The dual-filter (DF) approach for multichannel BCI involves implementation of a dual-filter structure comprising a background and a foreground filter. A measure is defined on the basis of the cost function of the adaptive algorithm in the background filter. The value of the measure is decided on the basis of the cost function of the adaptive estimation algorithm that in turn determines when to copy the estimates from the background filter into the foreground filter. Simulations with white Gaussian noise (WGN) and speech signal input display the noise robustness of the DF approach over the NMCFLMS algorithm for multichannel BCI.

Consider a speech signal recorded inside a non-anechoic room using an array of microphones. The microphone signals represent the convolution of the speech signal and the impulse responses of the acoustic paths between source and microphones. With reference to Fig. 1 and defining $s(n)$ and $v_i(n)$ as the source signal and background noise respectively,

the $i$th channel output signal $x_i(n)$ is given by

$$x_i(n) = y_i(n) + v_i(n), \quad i = 1, 2, \ldots, M; \quad (1a)$$

$$y_i(n) = h_i^T(n)s(n), \quad (1b)$$

where $M$ is the number of channels, $h_i(n) = [h_{i,0}(n) \ h_{i,1}(n) \ \ldots \ h_{i,M-1}(n)]^T$, $s(n) = [s(n) \ s(n-1) \ \ldots \ s(n-L+1)]^T$, $h_i(n)$ is the $i$th channel impulse response, $L$ is the length of the longest channel impulse response and the superscript $T$ denotes vector transposition. Defining $E\{\cdot\}$ as the expectation operator, we assume that the additive noise $v_i(n)$ on $M$ channels is uncorrelated, i.e., $E\{v_i(n)v_j(n)\} = 0$ for $i \neq j$ and $E\{v_i(n)v_i(n-n')\} = 0$ for $n \neq n'$ while additive noise $v_i(n)$ is uncorrelated with the input signal $s(n)$. For channel identifiability [4], we also assume that

1. The channel transfer functions do not contain any common zeros and
2. The autocorrelation matrix of the source signal, $R_{ss} = E\{s(n)s^T(n)\}$, is of full rank.

A blind multichannel system can be identified adaptively by minimizing the cross-relation error given, for $i \neq j$, by $e_{ij}(n) = x_i^T(n)\tilde{h}_j(n-1) - x_j^T(n)\tilde{h}_i(n-1), \quad i, j = 1, \ldots, M; \quad (2)$

where $\tilde{h}_i(n)$ is the estimated $i$th channel impulse response. Using (2), BCI algorithms such as NMCFLMS are derived by minimizing the cost function $J(n) = \sum_{i=1}^{M} \sum_{j=i+1}^{M} e_{ij}(n)^2$ with respect to the estimated impulse response $\tilde{h}_i(n)$ for $i = 1, \ldots, M$. The NMCFLMS [4] algorithm is a frame-based frequency-domain BCI algorithm given for each $m$th
frame, by:

\[
\hat{h}_i^0(m) = \sum_{j=1}^{M} \mathbb{P}_j(m) \mathbb{Q}_j^0(m), \quad i = 1, \ldots, M,
\]

\[
\mathbb{P}_j(m) = \lambda \mathbb{P}_j(m-1) + (1 - \lambda) \sum_{j=1, j \neq i}^{M} \mathbb{P}_j(m) \mathbb{Q}_j(m),
\]

\[
\mathbb{Q}_j^0(m) = \mathbb{F}_L \mathbb{W}_{2L \times 2L} \mathbb{Q}_j(m),
\]

\[
e_i^j(m) = \varepsilon_i^j(mL - L) \ldots \varepsilon_i^j(mL + L - 1)^T,
\]

where \( * \) denotes complex conjugate, \( \rho \) is the step-size, \( \lambda \) is the forgetting factor and \( \sigma \) is the regularization constant. Defining \( \mathbb{I}_{L \times L}, \mathbb{O}_{L \times L} \), and \( \mathbb{F}_L \) as the identity, null and Fourier matrices of dimension \( L \times L \) respectively, \( \mathbb{W}_{2L \times 2L} = [\mathbb{I}_{L \times L} \mathbb{O}_{L \times L}]^T, \mathbb{W}_{2L 	imes 2L}^0 = [\mathbb{O}_{L \times L} \mathbb{I}_{L \times L}]^T \),

\[
\hat{h}_i(m) = \mathbb{F}_L \mathbb{h}_i(m), \quad \hat{h}_i(m) = \mathbb{F}_L \left[ \hat{h}_i(m) \right] = \mathbb{F}_L \left[ \mathbb{h}_i(m) \right], \quad \text{and} \quad \mathbb{P}_j(m) = \text{diag} \left\{ \mathbb{F}_L \mathbb{h}_j(mL - L) \ldots \mathbb{h}_j(mL + L - 1)^T \right\}.
\]

The accuracy of system identification can be quantified using the normalized projection misalignment (NPM) [8] given, for frame \( m \), by

\[
\text{NPM}(m) = 20 \log_{10} \left( \frac{\| \hat{h}_i(m) - \mathbb{h}_i \|}{\| \mathbb{h}_i \|} \right) \text{dB},
\]

where \( \| \cdot \| \) is the \( l_2 \) norm and \( \mathbb{h}_i(m) = [\hat{h}_i^T(m) \hat{h}_i^2(m) \ldots \hat{h}_i^M(m)]^T \).

2. NMCFLMS UNDER NOISY CONDITIONS

The NMCFLMS algorithm lacks robustness to additive noise [5]. It can be seen from Fig. 2 that the NMCFLMS algorithm misconverges after achieving NPM of \( -20 \text{ dB}, -23 \text{ dB}, -25 \text{ dB}, \) and \( -30 \text{ dB} \) for corresponding signal-to-noise ratios (SNR) \( 25 \text{ dB}, 30 \text{ dB}, 35 \text{ dB}, \) and \( 40 \text{ dB} \), respectively for a constant adaptation gain of \( \mu = 0.6 \). Hence, the problem of misconvergence in the NMCFLMS algorithms [4] increases with noise level.

Using (1a) and (1b), we can write

\[
x_i(n) = y_i(n) + v_i(n),
\]

where \( y_i(n) = [y_i(n) y_i(n-1) \ldots y_i(n-L + 1)]^T \), and \( v_i(n) = [v_i(n) v_i(n-1) \ldots v_i(n-L + 1)]^T \).

Using (2) and (8), the error to be minimized can then be divided into a noise-free component \( e_i^nf \) and a noise-only component \( e_i^nf \) as

\[
e_i^f(n) = [y_i^f(n) \hat{y}_i(n-1) - y_i^f(n) \hat{h}_i(n-1)] + [v_i^f(n) \hat{h}_i(n-1) - v_i^f(n) \hat{h}_i(n-1)]
\]

\[
e_i^f(n) + e_i^f(n).
\]

In this noisy case, the cost function is described by

\[
J(n) = \frac{1}{\| \hat{h}_i(n) \|^2} \sum_{j=1}^{M-1} \sum_{j+1}^{M} \left\{ [e_i^f(n)]^2 + [e_i^f(n)]^2 \right\}
\]

\[
= J_f(n) + J_s(n).
\]

Hence the cost function \( J(n) \) is a sum of a noise-free element \( J_f(n) \) and a noise-only element \( J_s(n) \). It can be deduced that the minimization of the cost function \( J(n) \) under noisy conditions does not necessarily minimize the noise-free cost function element \( J_f(n) \), which is desired so as to identify the unknown system correctly. This is illustrated for two different SNR conditions in Fig. 3, showing that minimization of \( J(n) \) does not result in correct estimates of the unknown system in the presence of observation noise.

3. THE DUAL FILTER APPROACH FOR MULTICHANNEL BCI

We now develop the DF approach for multichannel BCI. The approach comprises two sets of filters; background filters \( \mathbb{B}(z) \) and foreground filters \( \mathbb{F}(z) \) as shown in the block diagram in Fig. 4. The \( \mathbb{B}(z) \) are adapted by the NMCFLMS algorithm. \( \mathbb{B}(z) \) estimates are copied into \( \mathbb{F}(z) \) at different times during the adaptation following an update decision criterion.
3.1 Estimation of the critical point

The critical point of $B(z)$ in the DF approach is determined by smoothing the cost function $J(n)$ using a 1024 tap rectangular window moving average filter. Smoothed $J(n)$ initially decreases then flattens to a certain value which is specific to the SNR of the system. Denoting the initial value of the smoothed $J(n)$ in dB as $J_1$ and its value when it flattens in dB as $J_f$, we define

$$\delta_f = |J_f - J_1|$$.

The value of $\delta_f$ has been found to be proportional to the SNR of the system but independent of the number of channels, $M$, as well as the length, $L$, of the system and the specific value of $h(n)$ which is shown in the Fig. 5. The flattening time, $t_f$, is the instant in time when $J(n)$ decreases to $J_f$ as shown in Fig. 3. The critical point $t_c$ is found empirically from the flattening time $t_f$ using the following relationship

$$t_c = t_f \times 2^{\gamma}$$.

3.2 The Dual filter Approach

The DF approach aims to avoid misconvergence in the presence of observation noise. The procedure is to copy the coefficients of $B(z)$ to $F(z)$ according to a decision criterion. $\gamma$ is used as a flag to implement the decision criterion. $B(z)$ is initialized after every time interval $T_{BG}$ and adapts using the NMCFLMS algorithm during the interval. Interaction between $B(z)$ and $F(z)$ is controlled by the value of $\gamma$ as shown in Fig. 6 with $\gamma$ initialized to 1. When $\gamma = 1$, $B(z)$ estimates are copied into $F(z)$ after each iteration until $B(z)$ reach the critical point $t_c$ then estimates are not copied. At the end of $T_{BG}$ interval, $\gamma$ is set to 0. In the next $T_{BG}$ interval $\gamma = 0$ so $B(z)$ estimates are not copied after each iteration, instead just once after reaching the critical point $t_c$. During $B(z)$ adaptation $J(n)$ is analyzed at each iteration for a change equal to $\delta_f$ after its flattening. This $\delta_f$ change in $J(n)$ reflects a change in the system. In this case, $\gamma$ is reset to 1 and the new value of $\gamma$ is used in the following $T_{BG}$. Hence the DF approach always keeps the best estimates in $F(z)$, avoiding the misconvergence which is otherwise observed in multichannel BCI in the presence of observation noise. Fast update of $F(z)$ could be achieved by adapting multiple instances of $B(z)$ with a small $T_{BG}$. Hence $F(z)$ is more frequently updated but with a price of higher computation.

4. SIMULATIONS AND RESULTS

We now present simulation results to compare the performance of the DF approach against the NMCFLMS algorithm [4] in the context of acoustic room impulse response identification. The dimensions of the room are $(5 \times 4 \times 3) m$ and impulses are generated using the method of images [9] with reverberation time $T_{60} = 0.1 s$ which are then truncated to length $L = 128$. A linear microphone array containing $M = 5$ microphones with uniform separation $d = 0.2 m$ is used. The source and the first microphone are placed at $(1.0, 1.5, 1.6) m$ and $(2.0, 1.2, 1.6) m$, respectively and a source at $(1.0, 3.0, 1.6) m$ is used to simulate change in the system. Input signals are white Gaussian noise (WGN) and a male speech while uncorrelated zero-mean WGN is added to achieve the SNR specified for each experiment. The sampling frequency is 8 kHz and the SNR is 20 dB unless otherwise specified. Defining $h = [h_1^T \ h_2^T \ \ldots \ h_M^T]^T$, the SNR for this BCI application is given [4] as $SNR = 10 \log_{10}[\sum_{i=1}^M |h_i|^2/(M \sigma_n^2)]$ where $\sigma_n^2$ and $\sigma_n^2$ are the signal and noise powers, respectively, while the following parameters are chosen for all simulations: $\gamma = 1$ at the start, $\lambda = |1 - 1/3L|^{1/2}$, $h_i(0) = [0 \ldots 0]^T / \sqrt{M}$ for $B(z)$ filters.

Figure 7 shows the comparison of convergence of the DF approach with the NMCFLMS algorithm using a WGN input signal at an SNR = 20 dB. The step-size for both algorithms is $\mu = 0.6$. $T_{BG}$ for $B(z)$ of the DF approach is 0.1 s. In Fig. 7(a) NMCFLMS clearly misconverges after initial convergence up to $t_c$. The NPM of $B(z)$ is shown with a dotted
The NMCFLMS is an effective algorithm for adaptive blind system identification but suffers from misconvergence in the presence of noise. To alleviate this problem, we proposed a dual filter approach. An expression, which is proportional to the SNR but invariant to channel coefficient values, channel length and number of channels was given, relating the smoothed cost function of an adaptive background filter with the critical point of misconvergence. We then employed this relation in a decision criterion for the update of a foreground filter. In this way, misconvergence is avoided and system variations can be tracked. Finally, simulation results demonstrated the performance improvement over the standard NMCFLMS using the proposed approach.

5. CONCLUSION

The NMCFLMS is an effective algorithm for adaptive blind system identification but suffers from misconvergence in the

Figure 6: Flow Diagram of the DF approach for multichannel BCI.

Figure 7: Comparison of convergence of NMCFLMS, $B(z)$ and $F(z)$ filters with WGN input for (a) Single system (b) Step change in the system.

line and misconverge within the $T_{BG}$ interval. $B(z)$ are re-initialized at the end of every $T_{BG}$ interval. $F(z)$ initially converge with $B(z)$ but after $t$, they keep the best estimates of the unknown channels following the update decision criterion hence do not misconverge. Fig. 7(b) shows the comparison of convergence with a step change in system after 50 s. $F(z)$ has the best estimate of the first system for 50 s but has wrong estimate after system change until the following $T_{BG}$ starts. Then it has the best estimate of second system copied from $B(z)$ following the update decision criterion.

Figure 8 shows the comparison of convergence of the DF approach with the NMCFLMS algorithm with speech signal input at an SNR = 40 dB. The step-size for the algorithms is $\mu = 0.08$ and $T_{BG}$ for $B(z)$ is 32.88 s. Fig. 8(a) shows the DF approach does not misconverge while NMCFLMS does and Fig. 8(b) shows the DF approach tracks changes in the system successfully.

REFERENCES


