MULTI-MICROPHONE ACOUSTIC ECHO CANCELLATION USING MULTI-CHANNEL WARPED LINEAR PREDICTION OF COMMON ACOUSTICAL POLES

Pepe Gil-Cacho[1], *Toon van Waterschoot*[1], *Marc Moonen*[1] *and Søren Holdt Jensen*[2]

[1]Katholieke Univeristeit Leuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
[2]Aalborg Univeristy, Dept. Electronic Systems, Niels Jernes Vej 12, DK-9220 Aalborg, Denmark

ABSTRACT

A computationally cheap extension from single-microphone acoustic echo cancellation (AEC) to multi-microphone AEC is presented for the case of a single loudspeaker. It employs the idea of common-acoustical-pole and zero modeling of room transfer functions (RTFs). The RTF models used for multi-microphone AEC share a fixed common denominator polynomial, which is calculated off-line by means of a multi-channel warped linear prediction. By using the common denominator polynomial as a prefilter, only the numerator polynomial has to be estimated recursively for each microphone, hence adapting to changes in the RTFs. This approach allows to decrease the number of numerator coefficients by one order of magnitude for each microphone compared with all-zero modeling. In a first configuration, the prefiltering is done on the adaptive filter signal, hence achieving a pole-zero model of the RTF in the AEC. In a second configuration, the (inverse) prefiltering is done on the loudspeaker signal, hence achieving a dereverberation effect, in addition to AEC, on the microphone signals.

1. INTRODUCTION

Acoustic echo cancellation (AEC) is used in speech communication applications where the existence of echoes degrades the intelligibility and listening comfort, such as in mobile and hands-free telephony and in teleconferencing. An acoustic echo canceller seeks to cancel the echo signal component $y(t)$ in a microphone signal $d(t)$, ideally leading to an *echo-free* error signal $e(t)$. This is done by subtracting an estimate of the echo signal $\hat{y}(t)$ from the microphone signal as shown in Fig. 1 (a) and (b). Therefore, an adaptive filter is used to provide a model that represents the best fit to the echo path H_{room} or room impulse response (RIR) [6]. This model is used to filter the input signal $x(t)$ to obtain the estimated echo signal. There are situations in which several microphones are employed simultaneously and therefore multi-microphone techniques are called for.

The most common model for the RIR is the finite impulse response (FIR) model corresponding to an all-zero model of the room transfer function (RTF). This is due to well-known behaviour and guaranteed stability of the adaptive FIR filters [5]. It is also connected to the physics of room acoustics where a microphone signal is a weighted sum of discrete reflections of the loudspeaker signal. Its drawback is that the number of FIR filter coefficients required to model a RIR increases dramatically if the RIR is long which is typical in room acoustics applications. Besides, any change of loudspeaker-microphone or obstacle position inside the room will change the coefficients of the model and therefore a recalculation of every coefficient will be needed. The problem of calculating a large amount of coefficients becomes more apparent in the case of a multi-microphone scenario where one adaptive FIR filter is used for each microphone.

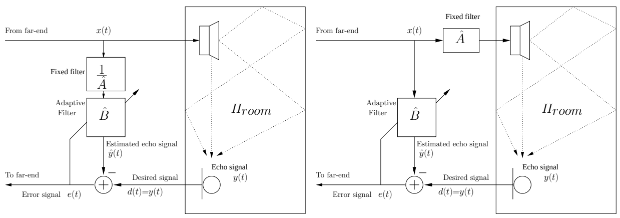
One of the alternatives is to use an infinite impulse response (IIR) model which corresponds to a pole-zero model of the RTF. This model reduces the number of coefficients to be estimated for the same modelling capabilities [5]. Moreover, this model again represents the physics of room acoustics now also including the modelling of the acoustic resonances by means of the poles of the transfer function. Poles can represent long impulse responses caused by resonances, while zeros represent time delays and antiresonances [8]. The well known drawbacks of the adaptive IIR filter are the nonlinear shape of the cost function for the filter coefficients estimation and its potential instability [5]. Although different algorithms have been proposed to overcome these limitations with more or less success, their use, especially the use of high order filters is very limited in practice [5].

In [8] the concept of a common-acoustical-pole and zero model is introduced. The underlying idea is that the acoustic resonances in a room depend on the dimensions and shape of the enclosure and not on the loudspeaker-microphone position. Each RTF may be expressed using a common set of poles and different zero functions. The RTF models used for the proposed multi-microphone AEC share a fixed common denominator polynomial which is calculated off-line, from a set of measured RIRs, by means of a multi-channel warped linear prediction. By using the fixed common denominator polynomial as a prefilter (see Fig. 1 (a) and (b)) only the numerator polynomial has to be estimated for each microphone. Therefore, the prefiltering has two positive impacts: first, it avoids the problems of adaptive IIR filters and second, it reduces the number of filter coefficients to be estimated. Moreover, as reducing the number of filter coefficients is a major concern in the proposed multi-microphone AEC, the calculation of the common poles will be performed in the frequency-warped domain. By frequency-warping the RIR, one is able to focus the computational resources in specific frequency regions of interest [2], [1], relaxing the modelling effort in those frequency regions that are of less interest [4].

This paper is organized as follows. In Section 2, the proposed model is shown. In Section 2.1 the concept of the common-acoustical-pole and zero model is further explained. In Section 2.2 the standard procedure for frequency warping is explained. In Section 2.3 the equations for multi-channel warped linear prediction are presented. In Section 3 it is shown the adaptive algorithm employed in the proposed AEC and the signals participating which depend on whether the prefilter is located in the adaptive filter signal path or

This research work was carried out at the ESAT laboratory of Katholieke Univeristeit Leuven, in the frame of the EC-FP6 project SIGNAL: 'Core Signal Processing Training Program' Marie-Curie Fellowship program (http://est-signal.i3s.unice.fr) under contract No. MEST-CT-2005-021175, the Concerted Research Action GOA-MaNet and the the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011). The scientific responsibility is assumed by its authors

in the loudspeaker signal path. In Section 4 simulation results are provided that demonstrate the performance of our proposed model in terms of echo reduction. Finally Section 5 concludes the paper.



(a) Prefilter in adaptive filter signal path (b) Prefilter in loudspeaker signal path

Figure 1: Acoustic echo canceler set-up with prefilter using fixed coefficients polynomial

2. PROPOSED MODEL

2.1 Common-acoustical-pole and zero model

Although the RTFs are different for each loudspeaker-microphone position, all RTFs in a room share the same resonance frequencies. These resonance frequencies may be visible as spectral peaks in the RTFs [5]. If only the zeros cause RTF variation then the RTFs can be expressed using a common denominator for all and a different numerator for each of them (i.e. $H_i(q,t) = B_i(q,t)/A_c(q)$) as depicted in Fig.3. This can be represented by either common poles ($p_c(k)$) and distinct zeros ($z_i(k,t)$) or in polynomial form using common autoregressive (AR) ($a_c(k)$) and distinct moving average (MA) ($b_i(k,t)$) coefficients [8],

$$H_i(q,t) = \frac{\prod_{k=1}^{Q}(1 - z_i(k,t)q^{-1})}{\prod_{k=1}^{P}(1 - p_c(k)q^{-1})} = \frac{\sum_{k=1}^{Q} b_i(k,t)q^{-k}}{1 - \sum_{k=1}^{P} a_c(k)q^{-k}} \tag{1}$$

where Q and P are the order of numerator and denominator respectively, $i = 1, ..., M$ the number of microphones and q denotes the time shift operator, i.e., $q^{-k}u(t) = u(t-k)$

2.2 Warped linear prediction

To obtain the AR coefficients of an impulse response a set of simultaneous equations must be solved (see also section 2.3). This set of simultaneous equations in our case takes the same mathematical form as the Wiener-Hopf equations for linear prediction, the Yule-Walker equations for an autoregressive model [6] and is equivalent to the equation error between a measured impulse response and the estimated all-pole impulse response [7]. Here, a warped linear prediction (WLP) will be used as in [3]. The standard procedure in producing frequency-warped impulse responses involves replacing the unit delay operator, q^{-1} of the original RTF, $H_i(q) = \sum_{t=0}^{N} h_i(t)q^{-t}$, by a first-order all-pass filter [1] $D(q,\lambda)$, i.e.,

$$H_i^w(q) = \sum_{t=0}^{N} h_i^w(t)q^{-t} = \sum_{t=0}^{N} h_i(t)D^t(q,\lambda) \tag{2}$$

where N denotes the impulse response length, the warping parameter $\lambda \in (-1,1)$ and

$$D(q,\lambda) = \frac{q^{-1} - \lambda}{1 - \lambda q^{-1}} \tag{3}$$

The superscript w means that the impulse response or RTF is transformed to the warped domain. The inverse mapping or dewarping (i.e., from the warped domain to the original domain) follows directly by applying again the same mapping with the sign of λ changed [1].

$$H_i(q) = \sum_{t=0}^{N} h_i^w(t)D^t(q,-\lambda) \tag{4}$$

Fig. 2 shows the RTF of a 2001 samples long measured room impulse response (sampling frequency $f_s = 16$ kHz) together with the spectrum of the estimated denominator polynomial (AR coefficients) calculated with order $P = 200$ and by warping the impulse response with $\lambda = 0$, $\lambda = 0.5$ and $\lambda = 0.9$ respectively. With $\lambda = 0$ the modelling effort is uniformly spread over the frequency axis. With $\lambda = 0.5$, and especially so with $\lambda = 0.9$, the modelling effort is put in lower frequency regions. As it can be seen, the low frequency region contains the main resonant peaks. Therefore it seems obvious to employ the limited computational resources in this area.

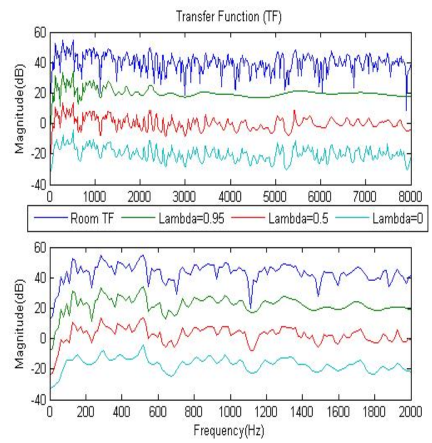


Figure 2: Frequency spectra of an impulse response and its estimated AR coefficients

2.3 Warped multi-channel linear prediction of common acoustical poles

From a set of warped impulse responses a set of common AR coefficients can be calculated and then transformed back to the original domain. This set of AR coefficients corresponds to the main resonances of the RTFs. Assuming that $h_i^w(t)$ is the time-domain version of the warped RTF $H_i^w(q)$ (2), the warped all-pole estimate of the impulse response coefficients is [1]

$$\hat{h}_i^w(t) = \sum_{k=1}^{P} a_c^w(k)h_i^w(t-k) \tag{5}$$

The warped common AR coefficients can be calculated as those that minimize the cost function

$$\min_{a_c^w(k)} \sum_{i=1}^{M}\sum_{t=0}^{\infty} e_i^2(t) \tag{6}$$

with

$$\begin{aligned} e_i(t) &= h_i^w(t) - \hat{h}_i^w(t) \\ &= h_i^w(t) - \sum_{k=1}^{P} a_c^w(k)h_i^w(t-k) \end{aligned} \tag{7}$$

The warped multi-channel linear prediction of the common AR coefficients $a_c^w(k)$ that minimize (6) are calculated by solving the normal equations [7], i.e.

$$\mathbf{a} = \left(\mathbf{W}^T\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{v} \tag{8}$$

where

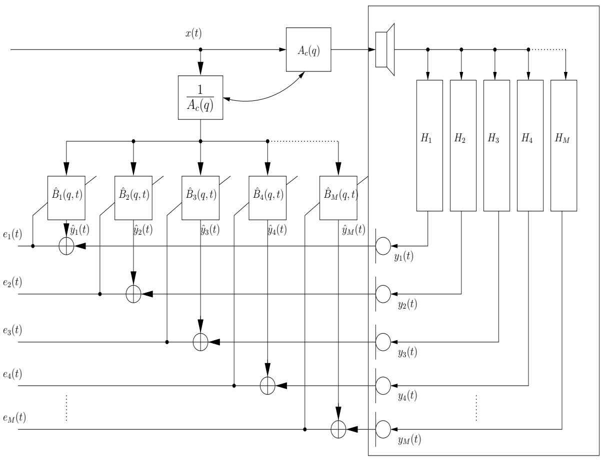


Figure 3: Multi-microphone acoustic echo canceler set-up

$$\mathbf{a} = [a_c^w(1), a_c^w(2), ..., a_c^w(P)]^T \tag{9a}$$

$$\mathbf{W} = [\mathbf{H}_1^T, \mathbf{H}_2^T, ..., \mathbf{H}_M^T]^T \tag{9b}$$

$$\mathbf{v} = [\mathbf{h}_1^{wT}, \mathbf{h}_2^{wT}, ..., \mathbf{h}_M^{wT}]^T \tag{9c}$$

$$\mathbf{h}_i^w = [h_i^w(1), h_i^w(2), .., h_i^w(N-1), 0, 0, .., 0]^T \tag{9d}$$

$$\mathbf{H}_i = \begin{bmatrix} h_i^w(0) & 0 & \cdots & 0 \\ h_i^w(1) & h_i^w(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & h_i^w(0) \\ h_i^w(N-1) & 0 & \cdots & \vdots \\ 0 & h_i^w(N-1) & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & h_i^w(N-1) \end{bmatrix} \tag{9e}$$

$$\Rightarrow \text{size } [(N+P-1) \times P]$$

3. AEC ALGORITHM

Vector $\mathbf{a}$ (9a) contains the warped common AR coefficients that are mapped back to the original domain to provide the fixed prefilter polynomial

$$A_c(q) = 1 - a_c(1)q^{-1} - ... - a_c(P)q^{-P} \tag{10}$$

The prefiltering is applied to either the adaptive filter signal or the loudspeaker signal as depicted in Fig. 3. In a first configuration, the prefiltering is applied to the adaptive filter signal, hence achieving a pole-zero model of the RTF in the AEC. In a second configuration, the (inverse) prefiltering is applied to the loudspeaker signal, hence achieving a dereverberation effect, in addition to AEC, on the microphone signals.

For every microphone the numerator coefficients

$$B_i(q,t) = b_i(0,t) + b_i(1,t)q^{-1} + ... + b_i(Q,t)q^{-Q} \tag{11}$$

are adapted using the well-known Normalized Least Mean Squares (NLMS) algorithm [6], i.e.

$$e_i(t) = y_i(t) - \hat{y}_i(t) \tag{12a}$$

$$= y_i(t) - \mathbf{b}_i(t)\mathbf{u}^T(t) \tag{12b}$$

$$\mathbf{b}_i^T(t+1) = \mathbf{b}_i^T(t) + \mu\frac{1}{\mathbf{u}(t)\mathbf{u}^T(t)}\mathbf{u}(t)^T e_i(t) \tag{12c}$$

where $i = 1, ..., M$, the vector $\mathbf{u}(t) = [u(t), u(t-1), ..., u(t-Q)]$ is the input to the adaptive filter, $\mathbf{b}_i^T(t) = [b_i(0,t), b_i(1,t), ..., b_i(Q,t)]$ are the adaptive filter coefficients, $y_i(t)$ are the microphone signals, $e_i(t)$ are the error signals, $\hat{y}_i(t)$ are the estimated echo signals and μ is the step size. Signals $u(t)$ and $y_i(t)$ will depend on whether the prefilter is in the adaptive filter signal path or in the loudspeaker signal path as follows:

- Prefilter in the adaptive filter signal path

$$u(t) = \frac{1}{A_c(q)}x(t) \tag{13}$$

$$= x(t) - a_c(1)u(t-1) - ... - a_c(P)u(t-P) \tag{14}$$

$$y_i(t) = H_i(q,t)x(t) \tag{15}$$

- Prefilter in the loudspeaker signal path

$$u(t) = x(t) \tag{16}$$

$$y_i(t) = A_c(q)H_i(q,t)x(t) \tag{17}$$

$$= A_c(q)\frac{B_i(q,t)}{A_i(q,t)}x(t) \approx B_i(q,t)x(t) \tag{18}$$

4. SIMULATION RESULTS

Matlab computer simulations were performed at $f_s = 16$ kHz. Five room impulse responses, (h_i with $i = 1, ..., 5$) of length $N = 2001$ samples were measured in a rectangular room of about $5 \times 3 \times 3$ m. In every simulation the NLMS step size $\mu = 1$ as it offered the best results. The input signal was speech (female voice) recorded at $f_s = 16$ kHz of 6.7 seconds duration (i.e., length $L = 1340876$ samples). The warping parameter $\lambda = 0.7$ was found to be optimal. The performance measures were: *Attenuation*,

$$Attenuation = 10\log_{10}\frac{\sum_{t=0}^{L} y(t)^2}{\sum_{t=0}^{L} e(t)^2} \quad (dB) \tag{19}$$

which is a scalar that measures the difference in dB between the power of the error and microphone signals; and *Echo Return Loss Enhancement* (*ERLE*),

$$ERLE(n) = 10\log_{10}\frac{\sum_{k=1}^{p} y^2\big((n-1)p+k\big)}{\sum_{k=1}^{p} e^2\big((n-1)p+k\big)} \quad (dB) \tag{20}$$

for $n = 1, ..., \frac{L}{p}$, which is (19) averaged over time frames of length p. The order of the numerator, Q, and denominator P are the same in every simulation.

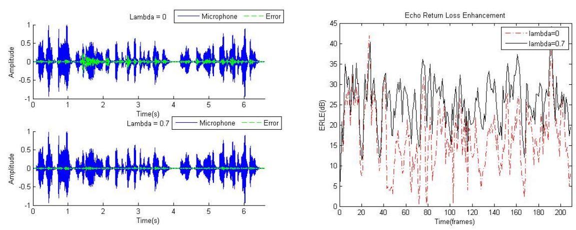
The performance in terms of *Attenuation* and *ERLE* of the 5 acoustic echo cancellers (AEC_i with $i = 1, ..., 5$) is shown in two different scenarios: first, when the AR coefficients of the prefilter are calculated from a single impulse response and second, when they are calculated using the set of impulse responses (i.e., common-acoustical-poles).

Table 1: Attenuation(dB). AR Coefficients obtained from h_1. 2.Left) Prefilter in adaptive filter signal path. 2.Right) Prefilter in loudspeaker signal path

LPC (h_1) (Adap. sig.)	AEC_1	AEC_4	LPC (h_1) (Loud. sig.)	AEC_1	AEC_4
$P = 200$	28.0	16.3	$P = 200$	19.4	14.0
$P = 500$	37.8	18.3	$P = 500$	29.9	14.2

4.1 Comparison of ERLE with and without warping

The effect of warping an impulse response prior to calculating the AR coefficients is shown in this section. Fig. 4 shows the differences in the AEC performance when using different values of the warping parameter λ. The optimal value $\lambda = 0.7$ was applied and compared with $\lambda = 0$ which means that no warping is applied to the impulse response. The WLPC (AR) order was $P = 200$. The prefilter was applied to the adaptive filter signal path. AR coefficients were extracted from impulse response h_1. Fig. 4(a) shows the time evolution (L samples) of the microphone y_1 and error signal e_1 and Fig. 4(b) shows the ERLE. The value of the attenuation of the AEC_1 was 16.6 dB with $\lambda = 0$ and 28 dB with $\lambda = 0.7$, which shows that, with warping, an improvement of more than 11 dB can be achieved.



(a) Error signal evolution with different λ (b) ERLE of AEC_1 with different λ

Figure 4: Prefilter in the adaptive filter signal path with WLPC (AR) order $P = 200$ with optimal $\lambda = 0.7$ and $\lambda = 0$ which means no warping

4.2 Prefilter in the adaptive signal path

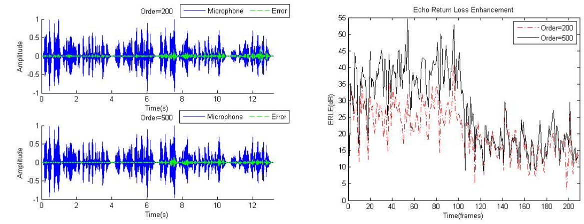
This section shows the performance of the AEC_i, with $i = 1, ..., 5$, when the prefilter is applied to the adaptive filter signal. In such case a pole-zero model of the RTF, with fixed denominator and variable numerator, is used in the AEC. The warping parameter was set to $\lambda = 0.7$. The AR coefficients were calculated with two different WLPC (AR) orders ($P = 200$ and $P = 500$).

- AR coefficients calculated from one impulse response (h_1)

Fig. 5(a) and Fig. 5(b) show the time evolution of the microphone and error signal and ERLE respectively, of AEC_1 and AEC_4. Table 1.(Left) shows the values of the attenuation achieved by AEC_1 and AEC_4 where the difference between them is 11.7 dB ($P = 200$) and 19.5 dB ($P = 500$).

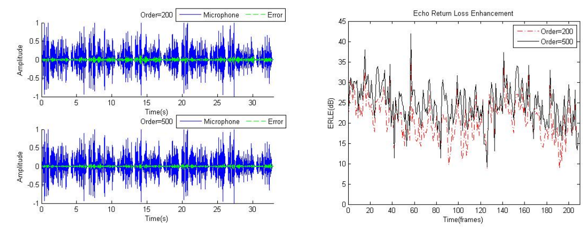
- Common-acoustical-pole AR coefficients

In this case the fixed filter coefficients are calculated from the set of impulse responses using (8). Fig. 6(a) and Fig. 6(b) show the time evolution of the microphone and error signal and ERLE respectively, of every AEC_i, with $i = 1, ..5$. Table 2.(Top) shows that the difference among the *Attenuation* values is highly reduced with satisfactory individual results (i.e., average *Attenuation* $= 20.4$ and 24.1 dB with $P = 200$ and 500 respectively).



(a) Error signal evolution with different WLPC (AR) orders (b) ERLE with different WLPC (AR) orders

Figure 5: Prefilter in the adaptive filter signal path. AR coeffients are calculated from impulse response h_1 with optimal λ=0.7 and $P = 200$ and $P = 500$. Performance of AEC_1 and AEC_4 are shown consecutively



(a) Error signal evolution with different WLPC (AR) orders (b) ERLE with different WLPC (AR) orders

Figure 6: Prefilter in the adaptive filter signal path. AR coeffients are calculated from common acoustical poles with optimal λ=0.7 and $P = 200$ and $P = 500$. Performance of every AEC_i are shown consecutively

Table 2: Attenuation(dB). AR coefficients obtained from common acoustical poles. Top) Prefilter in the Adaptive filter signal path. Middle) Prefilter in the loudspeaker signal path. Bottom) No prefilter

LPC_{common} (Adap. sig.)	AEC_1	AEC_2	AEC_3	AEC_4	AEC_5
$P = 200$	24.8	20.3	19.4	24.2	19.4
$P = 500$	27.9	24.3	24.4	26.0	22.5
LPC_{common} (Loud. sig.)					
$P = 200$	15.4	15.7	13.8	15.2	13.4
$P = 500$	24.0	19.9	19.7	22.5	18.3
All-zero (No prefilter)					
$Q = 2000$	36.4	31.6	33.3	28.7	32.1

4.3 Prefilter in the loudspeaker signal path

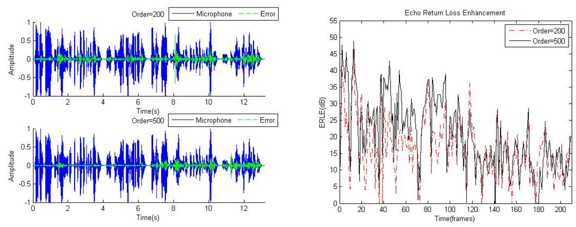
This section shows the performance of the AEC_i, with $i = 1, ..., 5$, when the prefilter is applied in the loudspeaker signal path. In such case a dereverberating effect is achieved by cancelling the main resonances. The warping parameter was set to $\lambda = 0.7$. The AR coefficients were calculated with two WLPC (AR) orders ($P = 200$ and $P = 500$).

- AR coefficients calculated from one impulse response (h_1)

Fig. 7(a) and Fig. 7(b) show the time evolution of the microphone and error signal and ERLE respectively, of AEC_1 and AEC_4. Table 1.(Right) shows the values of the attenuation achieved by AEC_1 and AEC_4 where the difference between them is 5.4 dB ($P = 200$) and 15.7 dB ($P = 500$).

- Common-acoustical-pole AR coefficients

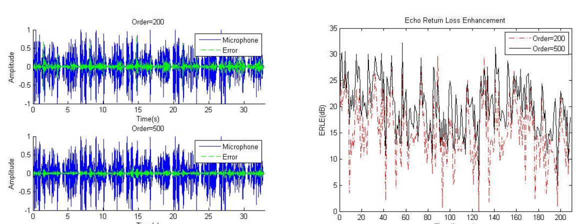
In this case the fixed filter coefficients are calculated from the set of impulse responses using (8). Fig. 8(a) and Fig. 8(b) show the time evolution of the microphone and error signal and ERLE respectively, of every AEC_i, with $i = 1, ..5$. Table 2.(Middle) shows that the difference among the *Attenuation* values is reduced with satisfactory individual results only with $P = 500$ (i.e., average *Attenuation* = 14.2 and 19.9 dB with $P = 200$ and 500 respectively).



(a) Error signal evolution with different WLPC (AR) orders (200 Top, 500 Bottom)
(b) ERLE with different WLPC (AR) orders

Figure 7: Prefilter in the loudspeaker signal path. AR coeffients are calculated from impulse response h_1 with optimal λ=0.7 and $P = 200$ and $P = 500$. Performance of AEC_1 and AEC_4 are shown consecutively

Table 2.(Bottom) also shows the values of the *Attenuation* in the case no prefilter is applied (i.e., all-zero case) with order $Q = 2000$ which is the total length of the impulse response. It can be seen that although the achieved attenuation is higher compared with our proposed model, the number of coefficients increases dramatically. Assuming that NLMS requires $2Q + P$ multiply-add operations for each update (12a)-(12c) [5], the number of operations in the all-zero case is $M \cdot 2Q = 5 \cdot 4000 = 20000$ whereas with fixed prefiltering, the number of operations is $M \cdot 2Q + P = 5 \cdot 400 + 200 = 2200$ (i.e., 91% saving) or $M \cdot 2Q + P = 5 \cdot 1000 + 500 = 5500$ (i.e., 36% saving). The better results obtained in the case of prefiltering in the adaptive filter signal path are due to the IIR nature of this configuration that leads to a better modelling capability.



(a) Error signal evolution with different WLPC (AR) orders
(b) ERLE with different WLPC (AR) orders

Figure 8: Prefilter in the loudspeaker signal path. AR coeffients are calculated from common acoustical poles with optimal λ=0.7 and $P = 200$ and $P = 500$. Performance of every AEC_i are shown consecutively

5. CONCLUSION

In this paper a model has been proposed for multi-microphone AEC which employs the idea of common-acoustical-pole and zero modelling of RTFs using warped linear prediction of the impulse responses. The common acoustical poles are calculated off-line from a set of measured impulse responses. In RTFs the predominant spectral peaks are located in the low frequency region. Warping allows us to focus the modelling effort in this frequency region to obtain better modelling results. This leads to a higher echo reduction for the same number of filter coefficients. Moreover these predominant spectral peaks are common to every RTF which allows to have a fixed common denominator polynomial for every channel. Hence only the numerator polynomial has to be estimated recursively. This approach allows to reduce the number of numerator coefficients up to one order of magnitude for each microphone compared with all-zero modelling, and yet provides satisfactory results (about 22 dB of echo attenuation). Better results are obtained in the case of prefiltering in the adaptive filter signal path.

REFERENCES

[1] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications", *J. Audio Eng. Soc.*, vol. 48, no. 11, pp. 1011-1031, 2000.

[2] T. Paatero and M. Karjalainen, "Kautz Filters and Generalized Frequency Resolution - Theory and Audio Applications", *J. Audio Eng. Soc.*,vol. 51, no. 1/2, pp. 27-44, January/February 2003.

[3] H. W. Strube, "Linear prediction on a warped frequency scale", *J. Acoust. Soc. Amer.*, vol. 68, no. 4, pp. 1071-1076, Oct. 1980.

[4] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio signals using a warped all-pole near-end model," in *Proc. IEEE Int. Conf. on Acoust. Speech Signal Process. (ICASSP '08)*, Las Vegas, Nevada, USA, Apr. 2008, pp. 269-272.

[5] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementations*. Springer, Boston, MA., 2008.

[6] S. Haykin, *Adaptive Filter Theory.* Prentice Hall, 2001.

[7] Y. Haneda, S. Makino and Y. Kaneda, "Multiple-Point Equalization of Room Transfer Functions by Using Common Acoustical Poles," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 4, pp. 325-333, July 1997.

[8] Y. Haneda, S. Makino and Y. Kaneda, "Common Acoustical Pole and Zero Modeling of Room Transfer Function," *IEEE Trans. Speech Audio Process.*, vol. 2, no. 2, pp. 320-328, Apr. 1994.