# ACOUSTIC ECHO CANCELLATION WITH REDUCED MEMORY REQUIREMENT USING PARTIAL ROOM CHARACTERISTIC CLONING

*Nenad Bojic*
DaudioR Laboratories / Victoria University of Technology
Signal Processing Research Group, Australia
ned@cabsav.vut.edu.au

## ABSTRACT

Developing low-cost Acoustic Echo Cancellation (AEC) systems can be a difficult task due to the limited internal memory available on low cost Digital Signal Processors (DSP). This paper introduces the Partial Room Characteristic Cloning (PRCC) AEC that has an inherently low memory requirement and computational rate. Real-time experimental results indicate that the PRCC AEC can achieve an ERLE of 15 – 20 dB with 200 words of memory and a computational rate of 0.8 MIPS, making it ideal for low cost AEC applications.

## 1. INTRODUCTION

After many years of both academic and commercial research, AECs still fail to achieve rapid convergence when confronted with highly coloured signals. In addition to this, many AECs suffer from a high memory requirement and/or a high computational rate. In order to increase the performance of their AECs, many companies are taking advantage of newer, more powerful DSPs to implement computationally expensive algorithms. Unfortunately, this type of AEC hardware is costly.

This paper introduces the PRCC AEC for acoustic echo cancellation in conferencing systems. The PRCC AEC takes advantage of several relationships in order to reduce the memory requirement and computational rate of an AEC system, allowing for implementation on low cost DSPs.

The PRCC AEC is introduced in section 2. The test set-up is described in section 3. The experimental results are presented in section 4. The experimental results and their implications are discussed in section 5.

## 2. PARTIAL ROOM CHARACTERISTIC CLONING

It is hypothesised by the author of this paper that the loudspeaker to room to microphone impulse response of a conferencing system is influenced by two distinct phenomena. The first phenomenon, the early room environment, determines the initial shape of the impulse response, otherwise know as the early room characteristic. In the case of a desktop conferencing system, the early room environment is the desktop itself. The second phenomenon, the late room environment, determines the latter shape of the impulse response, otherwise know as the late room characteristic. The late room environment includes the physical shape of the room, as well as any furniture, users, etc.

It is further hypothesised that for a given system, if the loudspeaker to room to microphone transfer function is fixed, the shape of the early room characteristic will be approximately constant, despite the late room environment. The loudspeaker to room to microphone transfer function of a particular system is fixed if the following conditions are observed.

a) The loudspeaker and microphone characteristics are constant.

b) The relative position of the microphone and loudspeaker is fixed. This property is inherent to many existing conferencing systems, some of which are depicted in Figure 1. These include audio conferencing phones, videoconferencing systems with multimedia loudspeakers and hands free phones. In Figure 1, the solid black dots indicate the location of the microphones. The loudspeakers are mounted behind the grills.

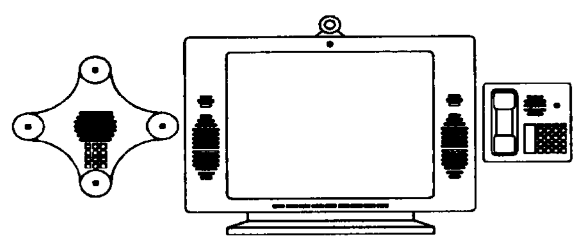

Figure 1 Conferencing systems suitable for PRCC AEC

c) The early room environment for which the system was designed is maintained. For instance, conferencing system designed for desktop use must be used on a desktop.

Based on the above hypothesis, once a set of coefficients representing the loudspeaker to room to microphone transfer function is determined, the early room characteristics can be cloned and stored for later use. These coefficients, known as the PRCC coefficients, are subsequently used within the PRCC AEC to directly model the early room characteristics. Since only the early room characteristics are modelled, the memory requirement of the PRCC AEC will be low.

This paper presents two separate PRCC AEC implementations, adaptive and non-adaptive.

### 2.1 Non-adaptive PRCC AEC

In low cost AEC implementations, the PRCC AEC can be used in conjunction with an echo suppresser to form a non-adaptive PRCC AEC, as illustrated in Figure 2.

The PRCC AEC works by modelling the early room characteristics $h_{erc}$ using a set of cloned coefficients $h'(k)$. The cloned coefficients $h'(k)$ and the input $x'(k)$ are convolved to generate an approximation of the echo $y'(k)$. The approximation of the echo is then subtracted from the real echo $y(k)$ to produce an error term $e(k)$.

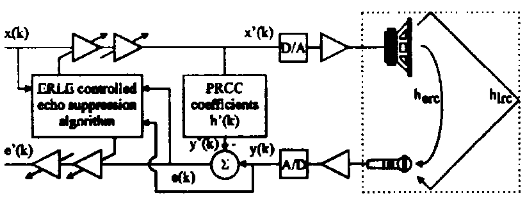


Figure 2 Non-adaptive PRCC AEC

Since the PRCC AEC does not model the late room characteristics $h_{lrc}$, the error term $e(k)$ will mainly consist of the $x'(k)*h_{lrc}$ terms, along with some low-level echo due to the PRCC approximation error. In the non-adaptive PRCC AEC, the residual echo is suppressed by an Echo Return Loss Enhancement (ERLE) controlled echo suppresser. Unlike threshold switching suppressers, the ERLE controlled echo suppresser does not gate low-level signals. Instead, the ERLE controlled echo suppresser uses the level of ERLE to distinguish between echo and low-level speech. It then acts to suppress the echo while allowing low-level speech to pass, thereby allowing full-duplex operation.

Software controlled volume levels have been incorporated into the design to maintain a constant gain between $x'(k)$ and $y(k)$. Consequently the PRCC coefficients are independent of the input and output volume levels.

### 2.2 Adaptive PRCC AEC

The PRCC AEC can also be used as a pre-echo canceller for an adaptive system as shown in Figure 3. The PRCC AEC cancels the echo due to early room characteristics, while the adaptive filter cancels the low-level echo due to late room characteristics and the approximation error of the PRCC AEC.

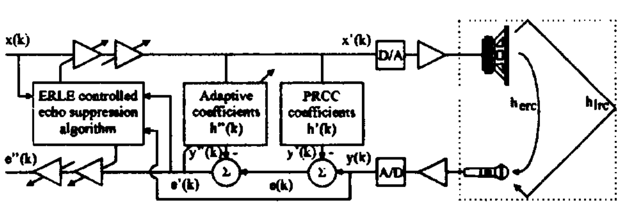


Figure 3 Adaptive PRCC AEC

## 3. TEST SETUP

In order to evaluate the effectiveness of the PRCC AEC it was necessary to test an audio conferencing system in different acoustic environments.

### 3.1 Acoustic Environments

Two rooms were used to generate seven different acoustic environments, as shown in Figure 4. Desks are represented by dashed lines.

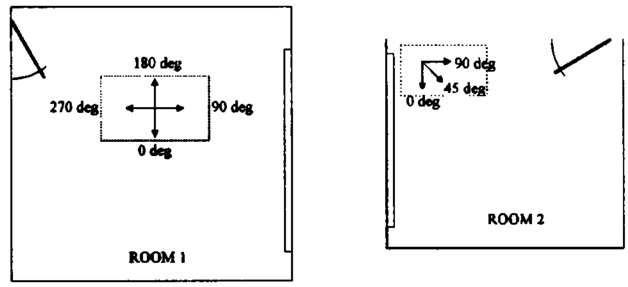


Figure 4 Test Rooms Layout

Room 1 was used to generate four different acoustic environments by placing the audio conferencing system on the desk and rotating it at different angles for each test. Room 2 was used to generate three different acoustic environments. It should be noted that the desk in room 1 was cleared of any objects while the desk in room 2 had several books scattered around the conferencing system. The specifications of the two rooms are given in Table 1.

Table 1 Room Specifications

| | *Room 1* | *Room 2* |
|---|---|---|
| ***Size(W x D x H)*** | 3.9 x 3.8 x 2.4 m | 3.3 x 2.9 x 2.4 m |
| ***Wall material*** | Plaster board | Plaster Board |
| ***Floor material*** | Carpet | Polished Wood |

### 3.2 Hardware

The audio conferencing system used in the tests consisted of a multimedia loudspeaker with a built-in omni-directional condenser microphone. The speaker enclosure contained a 4" woofer and 2" tweeter. An omni-directional microphone was chosen for its ability to detect voices from all users, both near and far. Figure 5 illustrates the conferencing system used in the tests.

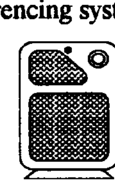

Figure 5 Audio-conferencing system used for testing

The soundcard used to sample the voice contained a 16-bit codec. The sampling frequency was set to 8 kHz. A sound file with a male voice was used as an excitation signal. The file was played back from CD-ROM. The loudspeaker volume was adjusted until the output of the loudspeaker was around 60 - 65 dB SPL (1m), A-weighted. This sound pressure level is typical of normal to above normal speaking levels.

## 4. TEST RESULTS

In order to illustrate the cancellation properties of the PRCC AEC, the following test were carried out with the echo suppresser disabled.

### 4.1 Room Characteristic Comparisons

The impulse responses of seven acoustic environments are illustrated superimposed in Figure 6.

By comparing the individual impulse responses in Figure 6, it can be observed that the initial coefficients in each of the seven acoustic environments are almost identical, while the latter have some variation with respect to each other. This is consistent with the initial hypothesis that under a set of predefined conditions, the shape of the early room characteristic is approximately constant for a given system. Therefore, if we clone the truncated impulse response of one of the acoustic environments, we can use it to approximate the initial impulse response of the other acoustic environments. For the purpose of the following tests, the first of the seven impulse responses was cloned.

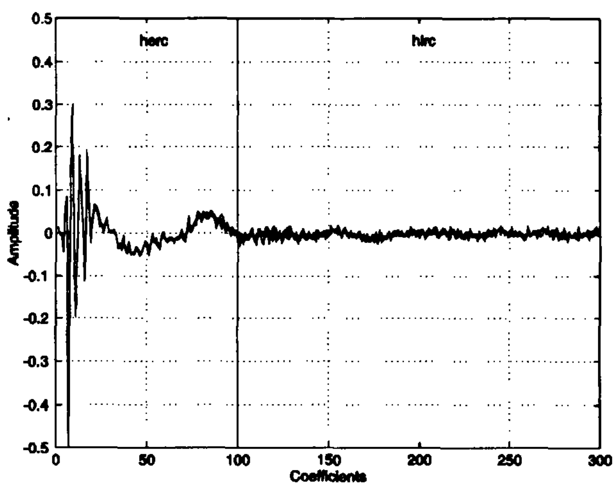


Figure 6 Impulse responses of various acoustic environments

### 4.2 Non-adaptive PRCC performance

The performance of a 64-coefficient clone was investigated in all seven acoustic environments. The result of this investigation is presented superimposed (in terms of ERLE) in Figure 7 (a). From Figure 7 (a) it is obvious that the cloned impulse response produces very similar levels of cancellation in all the acoustic environments. However, the ERLE achieved is small because an insufficient number of coefficients are used to model the echo. The nominal ERLE is around 8 dB while the peaks are about 22 dB.

The performance of a 100-coefficient clone was also investigated in all seven acoustic environments. The result of this investigation is presented in Figure 7 (b). From Figure 7 (b) it can be seen that there is still very little variation in cancellation between the acoustic environments. However, the nominal ERLE has increased to 15 dB.

Finally the performance of a 256-coefficient clone was investigated in all seven acoustic environments. The result of this investigation is presented in Figure 7 (c). Although increasing the number of coefficients improved the ERLE in some acoustic environments, there was little or no improvement observed in others. This suggests that an excessive number of coefficients are used to model the echo. In this case, the clone begins to model the late room characteristics, in addition to the early room characteristics. Given that the late room characteristics are environment dependent, it is inappropriate to include the late room characteristics among the cloned coefficients.

The real echo *y(k)* is overlaid onto the base of Figure 7 (c) to aid in the interpretation of the ERLE graphs.

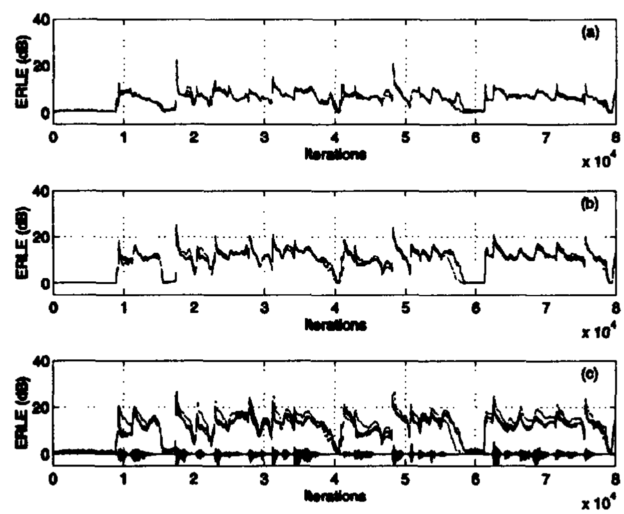


Figure 7 ERLE graphs for the non-adaptive PRCC AEC

### 4.3 Adaptive PRCC AEC performance

In order to investigate the effectiveness of the PRCC AEC as a pre-echo canceller, the performance of a 100-coefficient clone in conjunction with a 768-coefficient adaptive filter based on the Normalised Least Mean Square (NLMS) algorithm was studied.

For the purposes of comparison, the performance of the 100-coefficient PRCC AEC (without any adaptive filtering) and the 768-coefficient NLMS AEC (without pre-echo cancellation) is illustrated in Figure 8 (a) and (b) respectively. As illustrated in Figure 8 (b), the NLMS AEC suffers from slow convergence. Figure 8 (c) illustrates the performance of the adaptive PRCC AEC. From Figure 8 (c) it is obvious that the first syllable of the adaptive PRCC AEC has a 15 dB improvement over the NLMS AEC. It can also be seen that the NLMS AEC can not sustain 20 dB of ERLE until the 72000$^{th}$ iteration. The adaptive PRCC AEC passes the 20 dB ERLE after the 12000$^{th}$ iteration and sustains 25 dB of cancellation after the 17000$^{th}$ iteration.

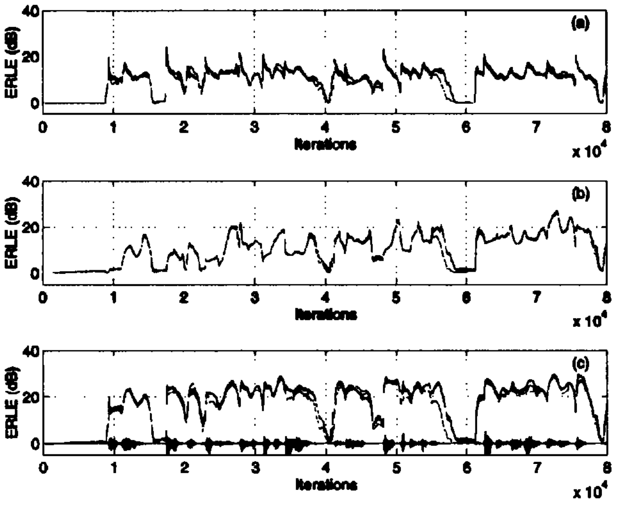


Figure 8 ERLE graphs for the adaptive PRCC AEC

## 5. DISCUSSION

Section 4.2 demonstrated the performance of the non-adaptive PRCC AEC (with the ERLE controlled echo suppresser disabled). It was shown that in all tested acoustic environments, a 100-coefficient PRCC AEC could achieve between 15 to 20 dB of ERLE. However, by enabling the echo suppresser, the performance of the non-adaptive PRCC AEC can be improved further.

Section 4.2 also demonstrated that the PRCC AEC approximation error is small when an appropriate number of coefficients is used to model the loudspeaker to room to microphone impulse response for a particular conferencing system. As shown, when a 100-coefficient PRCC AEC was used, the PRCC AEC performed equally well in all tested acoustic environments.

The specification of the non-adaptive PRCC AEC (with the ERLE controlled echo suppresser disabled) is shown in Table 2. From these results, it is clear that both the memory requirement and the computational rate of the PRCC AEC are very low. (The DSP used for the MIPS calculation processes one multifunction instruction per clock cycle at a sampling rate of 8 kHz.)

Table 2 Specifications for the non-adaptive PRCC AEC

| | |
|---|---|
| *Memory requirements* | 100 + 100 = 200 words |
| *MIPS* | 0.8 MIPS |
| *ERLE* | 15 – 20 dB |

Section 4.3 demonstrated the performance of the adaptive PRCC AEC (with the ERLE controlled echo suppresser disabled), where the PRCC AEC was used as a pre-echo canceller for an adaptive filter. It was shown that by using the PRCC AEC as a pre-echo canceller the performance of the NLMS algorithm could be significantly improved. Aurally, a sense of immediate convergence was perceived when the PRCC AEC was used as a pre-echo canceller.

In its standard form, the NLMS algorithm would never be used in a real time system due to its slow convergence and high computation requirement (as compared to the sub-band NLMS). However, in these tests it was shown that even a poorly converging algorithm like the NLMS algorithm can be improved with very little extra computational complexity by using the PRCC AEC as a pre-echo canceller.

## 6. CONCLUSION

The PRCC AEC has been presented. It has been shown that PRCC AEC is well suited for low-cost conferencing systems due to its low computational rate and inherently low memory requirement. Real-time experimental results have indicated that the non-adaptive PRCC AEC is able to achieve an ERLE of 15 to 20 dB with 200 words of memory and a computational rate of 0.8 MIPS.

It was also shown that by using the PRCC AEC as a pre-echo canceller for an adaptive system, the performance of the adaptive system could be significantly improved.

It was also demonstrated that for a given system, if the loudspeaker to room to microphone transfer function is fixed, the shape of the early room characteristic is approximately constant, despite the late room environment. This makes the PRCC AEC an attractive alternative to conventional adaptive filters in high noise environments where conventional adaptive filters fail to converge. Since the PRCC coefficients are fixed in all environments, both high and low noise, the PRCC AEC's ability to cancel echo is not limited by the noise floor. Additionally, since the PRCC coefficients are fixed, there is also no training time associated with determining the coefficients or possibility of divergence.