

Distributed Echo Cancellation in Multimedia Conferencing System

Balan Sinniah¹, Sureswaran Ramadass²

¹KDU College Sdn.Bhd, A Paramount Corporation Company , 32, Jalan Anson, 10400 Penang, Malaysia.

sbalan@kdupg.edu.my
<http://www.kdupg.edu.my>

²Network Research Group, School Of Computer Science, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia.

sures@cs.usm.my
<http://nrg.cs.usm.my>

Abstract. *As quality of video and audio frames transmission via internet/LAN is vital, numerous methods and techniques are employed to sustain a better performance of multimedia streaming. Yet, echo cancellation for speech and audio at software level still under research. The prime objective of echo cancellation is to improve clarity of audio/speech signal. Echo cancellation is a digital signal processing techniques for removing unwanted signal from speech/audio. Many techniques have been implemented to reduce the echo during conferencing; however there is more space to refine and enhance the existing techniques. This proposal suggests a software approach to achieve echo-cancellation in point-to-point multimedia conferencing system. The proposed technique use phase shifting method to eliminate any existing echo in the audio data received from the network. The solution will also enhance incoming audio quality by reducing background noise. The filtered or process audio data will provide high quality speech and hopefully will improve the quality of multimedia conferencing system one step further.*

1 Introduction

An echo is the repetition of a sound caused by the reflection of sound waves [1]. In multimedia conferencing system, particularly audio conferencing echoes are problematic if the speakers hear a delayed version of the same signal (voice). Numerous researches have been conducted by the telecommunications industry to control/eliminate unwanted signal (echo) few decades ago. Echo is becoming the factor of producing low audio quality signal when the Round Trip Delay (the time taken to reflect an echo) is more than 30 milliseconds [1].

Echo or Acoustic echo is caused by acoustic coupling problems between an audio conferencing speaker and its microphone. The tendency to produce echoes is roughly inversely proportional to the distance between the speaker and microphone. Usually in video conferencing systems, the use of earphones eliminates this problem. However, it is only applicable for desktop video conferencing units. The problem still exists in all the boardroom video conferencing units.

This paper will focus on the software based *phase shifter* technique to remove the echo in audio conferencing tool. Phase shifting is a technique where the original signal (noised signal) would be “added” with the shifted signal, which needs to be removed from the original signal by using the *180° phase shifter*. The paper will not address the implementation procedure or techniques for the phase shifter method but rather discuss more on theoretical aspects of designing approach to eliminate this problem.

This would be an overall discussion of the proposed method to remove echo in audio conferencing. It is believed that this method would be able to provide a hardware independent solution for acoustic echo cancellation.

2 Literature Review

The process that is described by the proposed echo cancellation technique involved Digital Signal Processing (DSP) approach, which is concerned with the digital representation of analogue signals and the use of digital processors to analyse, modify or extract information from signals. Analogue signals are sampled at regular intervals and converted to digital form. Processing a digital signal would be able to guarantee the accuracy (number of bits used), reproduce the signal perfectly and give a space for reprogramming.

It is known that human brain does very well in speech recognition while computers failed to compete with the human brain. Even computers are able to store and recall vast amounts of data, perform mathematical calculations at high speeds and do repetitive task without failing or getting bored, yet they perform very poorly when faced with raw sensory data [2]. In speech recognition, each word in the incoming audio signal is isolated and then analyzed to identify the type of excitation and resonate frequencies.

The Phase Shifting method, which has identified to eliminate the echo indeed, would be using digital filters to accomplish the task. Digital filters were designed to provide high performance in DSP. Digital filters’ main uses would be *signal separation* and *signal restoration* [2]. A *digital* filter is just a filter that operates on digital signals, such as sound represented inside a computer. It is a *computation*, which takes one sequence of numbers (the input signal) and produces a new sequence of numbers (the filtered output signal) [3].

A real digital filter T_n is defined as any real-valued function of a signal for each integer $n \in \mathbf{Z}$. Thus, a real digital filter maps every real, discrete-time signal to a real, discrete-time signal. A *complex* filter, on the other hand, may produce a complex output signal even when its input signal is real [3].

Phase is always measured relative to a reference, which if known, permits absolute phase measurement and if not known, permits only relative phase measurement [4]. The following diagrams illustrate this concept further.

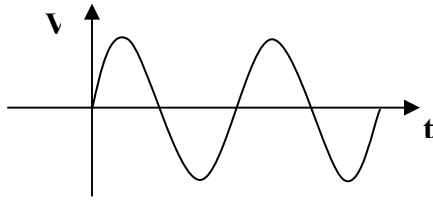


Figure 1.0 Signal with 0° phase which is used as a reference for other phase measurement.

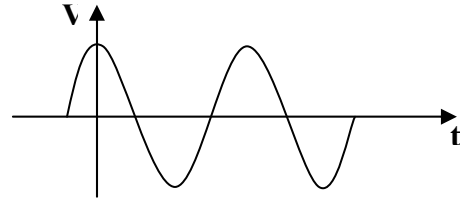


Figure 2.0 Signal with $+90^\circ$ phase with reference to the signal from figure 1.0. It is also known as “phase lead”.

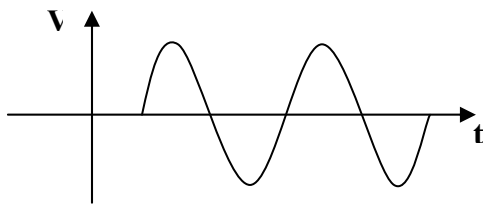


Figure 3.0 Signal with -90° phase with reference to the signal from figure 1.0. It is also known as “phase delay”.

2.1 Previous Work

A very less software based approaches have been taken to reduce the echo cancellation problem. Moreover, most of the techniques do not concentrating any specific application area such as multimedia conferencing area. Octastic’s Advance processors (Octasic OCT6100) use a deterministic technique to deliver good quality in a reliable and deterministic manner [5]. The OCT6100 series employed use of inexpensive memory and high processing power. It uses their own predefined “Least Squares” algorithm, which works based on minimizing the energy of the echo. The algorithm ensures the improved handling of double talk and background noise.

Speaker identification technology using Least Mean Square (LMS) algorithm is another effort in handling the echo problem [7]. The approach is for canceling the echo in long distance telephone conversation due to the irregularities of the analog telephone network. The implementation of the echo canceller has been optimized for two-way telephone conversation and has been tested on the SWITCHBOARD corpus.

3 The Proposed System Architecture

The proposed design would be purely based on a software approach. The **180° Phase Shifter** is a software component that would work on audio signal cancellation in a machine, which receives the acoustic echo. Figure 4.0 shows the illustrated diagram of the proposed design.

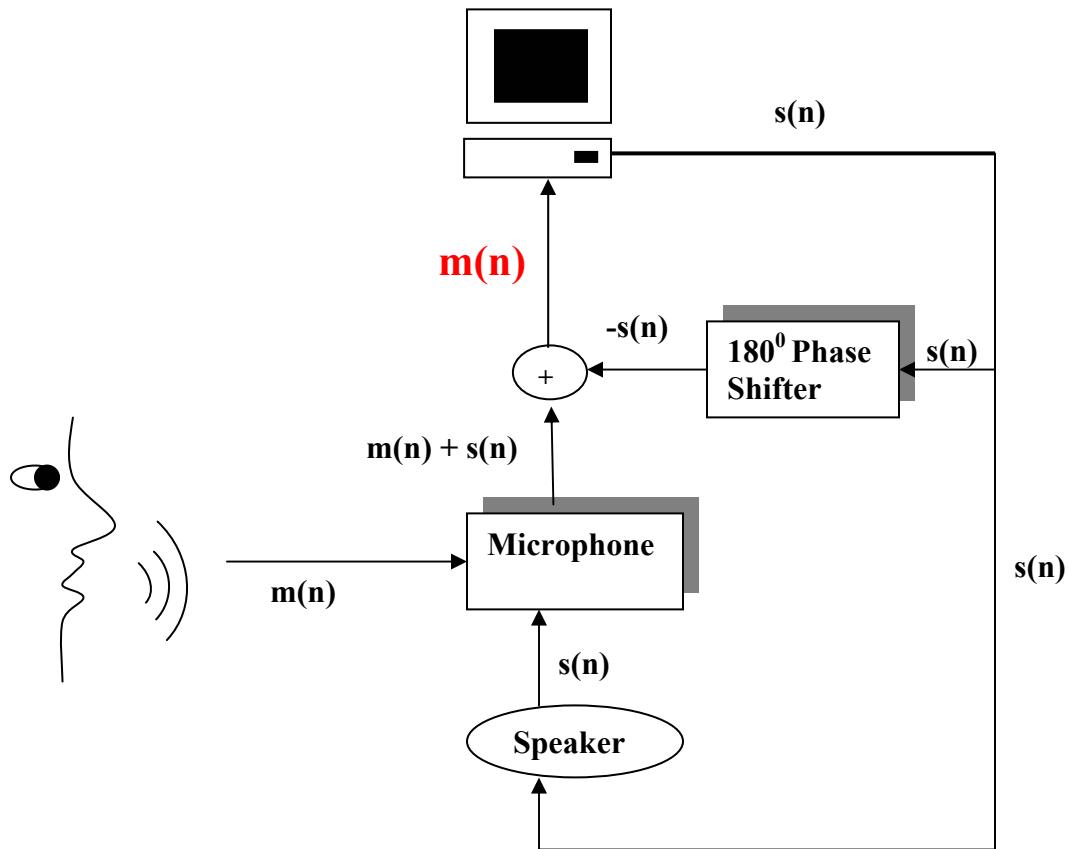


Figure 4.0

In audio (boardroom) conferencing system, when a near-end speaker speaks the voice signal would be added to the far-end speaker's original voice and it produces acoustic echo for the near-end speaker. The process is similar to the far-end speaker. The above diagram shows the near-end speaker's overall echo cancellation architecture using **180° Phase Shifter algorithm**.

In a normal (without echo cancellation) boardroom conferencing environment, when a user speaks through microphone the signal is captured, $m(n)$. While this signal is transmitting to the far-end user, the original signal (voice) is added with the signal from the speaker $s(n)$. Thus the overall signal received by the far-end speaker would be $m(n) + s(n)$ which in turn produces the echo for the far-end speaker.

Considering the acoustic echo which is solely produced by the speaker signal $s(n)$, the **180° Phase Shifter** approach is introduced to eliminate the additional signal or echo. This shifter would synthesis a signal that has a different phase (180°) as compare to the signal produced by the speaker. Thus the newly synthesized signal would be $-s(n)$. In order to remove the echo produced by the speaker, this signal ($-s(n)$) would be added to the echoed signal $m(n) + s(n)$. This addition operation will produce only the signal from near-end speaker voice ($m(n)$), that is :-

$$\underbrace{m(n) + s(n)}_{\text{Echoed signal}} + \underbrace{(-s(n))}_{\text{Newly synthesized signal}} = m(n)$$

The signal addition process would involve the implementation of digital filters. The Finite Impulse Response (FIR) filter is chosen to complete the above-mentioned task. FIR is a filter structure that can be used to implement almost any sort of frequency response digitally.

3.1 The 180° Phase Shifter Algorithm

Below is the brief description of the proposed phase shifter algorithm.

```

capture_audio_data
send_audio_data
receive_audio_data
do_phase_analysis
if phase NOT Constant then
    do_phase_averaging
else
    get_phase
end if
synthesis_new_audio_data_with_phase_shifter
while(capture_audio_data)
    do
        add_echoed_audio_data_with_synthesized_data_using_FIR_Filter
        send_audio_data
        clear_echoed_audio_data_buffer
        clear_synthesis_audio_data_buffer
    done
end while

```

It is known that for fast and high bandwidth transmission, User Datagram Protocol (UDP) would be the best option as the mode of communication. However, the route taken by the UDP packets are not consistent and therefore the phase of audio data that arrives most likely would be varied. Thus, it is vital to do a phase analysis on the received audio data before synthesizing new audio data for canceling the echo.

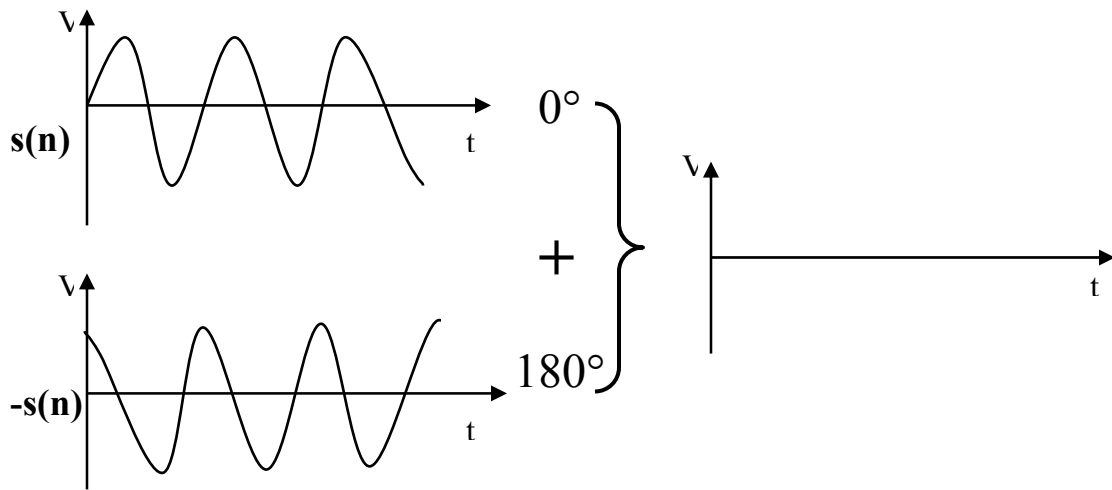


Figure 5.0

Figure 5.0 shows two analog audio signals with same amplitude but different phases. These two signals would be able to cancel each other when they are added with a digital filter. Thus, produce a silence audio data, which eventually cancel the echo.

4 Problems and Constraints

The proposed method would not be an ideal solution for pure echo cancellation in multimedia conferencing. There are some other external elements could caused the accomplishment of the proposed method being disrupted.

a) External Environmental Noises

The proposed phase shifter method would not be able to eliminate the external environmental noises due to the fact that these distortions are negligible as its amplitude usually is low. The impact of these elements would be serious to the solution if the amplitude of these noises is high. Thus, additional unnecessary data is sent. So inclusion of proper elimination method for these noises would be essential.

b) Electrical Echo

Electrical echo is caused by impedance mismatches in the analog local loop [5]. Electrical echo would not be an obvious problem in an intranet audio conferencing system where the distance is relatively short and does not produce significant delays. However, the electrical echo has to be controlled especially in dial up connection, which is using analog line as a medium of communication.

c) UDP Packets' Variable Route

UDP is an ideal protocol for audio conferencing since they place a higher priority on continuous media streaming [6]. However the routes taken by UDP packets are always varied and not predictable. Consequently, the audio packets may arrive in

different phase. Due to the inconsistent routes, which have caused the different phases of audio data, it would be much complicated in predicting the phase of incoming audio data. The averaging process of different phases is believed to give a solution for this problem, yet it does not guaranteed an exceptional quality in delivering the echoless audio data. Implementation of audio conferencing in IPv6 is able to eliminate the occurrences of variable phase problem as the flow label and traffic class in IPv6 prioritized the UDP packets at the router level.

5 Conclusion and Future Works

The proposed idea for echo cancellation in multimedia conferencing system uses an 180° phase shifter technique. The approach is more focusing on the aspect of digital signal processing as well as the phase manipulation. Practical implementation of this technique is yet to complete but the theoretical idea is believed to give a reasonable result in canceling the echo. Besides, it is a purely hardware independent approach to eliminate the critical echo problem in multimedia conferencing. Even the proposed idea is lacking in purifying the audio signal, it is able to provide a better solution for the primary problem. The idea can be improved with the implementation of multimedia conferencing system in IPv6 as well as introducing external noise reduction technique.

References

- 1 Octasic Semiconductor, *Echo Cancellation and Voice Quality in Today's Circuit Switched, Digital Wireless, and Voice over Packet Networks*, August 2002
- 2 Steven W. Smith, *The Scientist and Engineers Guide to Digital Signal Processing*, Second Edition, 1997
- 3 ``Introduction to Digital Filters'', by Julius O. Smith III, 2003 DRAFT.
- 4 Andrew Bateman, Iain Paterson-Stephens, *The DSP Handbook*, Prentice Hall, 2002.
- 5 Octasic, *Echo Cancellation and Voice Quality in Today's Circuit Switched, Digital Wireless, and Voice Over Packet Networks*, 2002.
- 6 Josh Beggs & Dylan Thede, *Designing Web Audio* , 2000
- 7 Aravind Ganapathiraju & Joseph Picone, *Echo Cancellation for Evaluating Speaker Identification Technology*,