International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.2, Issue No.10, pp : 1017-1021 1 Oct. 2013

Implementation of AEC(Acoustic Echo Cancellation) on Cortex-M4

[1] D. Vishnu Vardhan, [2] Vimala Jane Prathima P.
[1]Assistant Professor, J.N.T.U.A College Of Engineering Pulivendula, A.P, INDIA-516390
[2]P.G student, J.N.T.U.A College Of Engineering Pulivendula, A.P, INDIA-516390
[1]vishnushreyan@yahoo.co.in
[2]jane.prathima565@gmail.com

Abstract–***This paper presents the Acoustic Echo cancellation technique implemented on Cortex M4 processor. Acoustic echo is seen in full-duplex communication applications. The solution for the above problem is Acoustic echo canceller (AEC). The most important component of AEC is the adaptive filter. The performance of AEC are closely related to the underlying adaptive filter (adaptive algorithm + adjustable filter). The adaptive filter used here is the Multi delay block frequency domain adaptive filter (MDF), which is a frequency domain adaptive filter.In terms of performance the MDF adaptive filter introduces smaller block delay and is faster, best suited for a time varying system. Here the AEC is realized by implementing MDF algorithm on the advanced ARM processors like Cortex-M4 processor.***

Keywords–Acoustic Echo Cancellation(AEC); Multi Delay block frequency domain adaptive Filter(MDF);Cortex-M4;Adaptive Filter; Double Talk Detector; Non Linear Processor(NLP);

I.INTRODUCTION

In this new age of global communications, wireless phones are regarded as Essential communications tools and have a direct impact on people's day-to-day personal and business communications. The new network infrastructures are implemented. The competition between wireless carriers is increased, the digital wireless subscribers are becoming even more critical of the service and voice quality they receive from network providers. Subscriber demand for enhanced voice quality over wireless networks has driven a new and key technology termed echo cancellation, which can provide near wire line voice quality across a wireless network. Today's subscribers use speech quality as a standard for assessing the overall quality of a network. For subscriber's loyalty, the effective removal of hybrid and acoustic echoes seen within the telecommunications network infrastructure is the key to maintain and to improve the perceived voice quality of a call. Ultimately, the search for improved voice quality has led to intensive research into the area of echo cancellation. A research is conducted with the aim of providing solutions that can reduce background noise and remove hybrid and acoustic echoes before any trans-coder processing occurs. By doing echo cancellation process, the quality of speech can be improved significantly. A definition of echo gave rise to the discussion of the fundamentals of echo cancellation and the voice quality challenges encountered in today's networks.

II. AEC PROCESS

BASICS OF ECHOES AND ITS TYPES

Echo is a phenomenon where a delayed and distorted version of an original sound or electrical signal is reflected back to the source. With rare exceptions, conversations take place in the presence of echoes. Echoes of our speech are heard as they are reflected from the floor, walls and other neighboring objects. If a reflected wave arrives after a very short time of direct sound, it is called a spectral distortion or reverberation.

When the leading edge of the reflected wave arrives a few tens of milliseconds after the direct sound, it is heard as a distinct echo. Since the advent of telephony echoes have been a problem in communication networks. Echoes can be generated electrically due to impedance mismatches at various points along the transmission medium. The most important factor in echoes is latency ,also called end-to-end delay. Latency is the time between the generation of the sound at one end of the call and its reception at the other end. Round trip delay, which is the time taken to reflect an echo, is approximately twice the end-to-end delay. Echoes become annoying when the round trip delay exceeds 30 ms. Such an echo is typically heard as a hollow sound. Echoes must be loud enough to be heard. Those less than 30 decibels (dB) are unlikely to be noticed. However, not all echoes reduce voice quality.

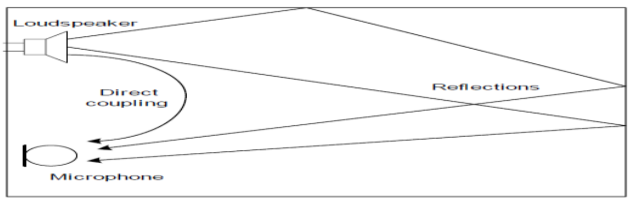

Fig 1: Acoustic Echo

In telecommunication networks there are two types of echoes. One source for an echo is **electrical** and the other echo source is **acoustic**. The electrical echo is due to the impedance mismatch at the hybrids of a Public Switched Telephony Network, (PSTN), exchange where the subscriber two-wire lines are connected to four-wire lines. If a communication is simply between two fixed telephones, then only the electrical echo occurs. However, the development of hands-free teleconferencing systems gave rise to another kind of echo known as an acoustic echo. The acoustic echo occurs due to the coupling between the loudspeaker and microphone.

ACOUSTIC ECHO CANCELLATION

An echo canceller is basically a device that detects and removes the echo of the signal from the far end after it has echoed on the local end's equipment. In the case of circuit switched long distance networks, echo cancellers reside in the Central Offices of Metropolitan that connect to the long distance network. An echo canceller consists of three main functional components:

1. Adaptive filter
2. Doubletalk detector
3. Non-linear processor

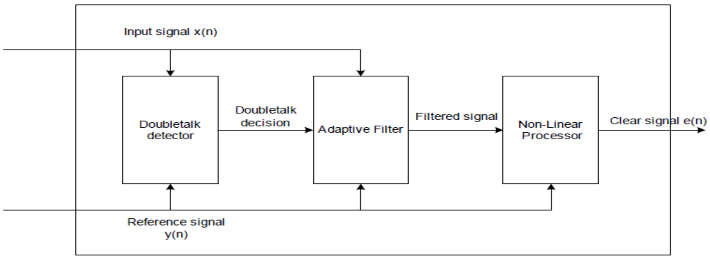

Fig 2: Block Diagram of Echo Canceller

A. Adaptive Filter:

The adaptive filter is made up of an echo estimator and a sub-tractor. The echo estimator monitors the received path and dynamically builds a mathematical model of the line that creates the returning echo. The model of the line is convolved with the voice stream on the receive path. This yields an estimate of the echo, which is applied to the sub-tractor. The sub-tractor eliminates the linear part of the echo from the line in the send path. The echo canceller is said to converge on the echo as an estimate of the line is built through the adaptive filter.

B. Double-Talk Detector:

A doubletalk detector is used with an echo canceller to sense when far-end speech is corrupted by near-end speech. The role of this important function is to freeze adaptation of the model filter when near-end speech is present. This action prevents divergence of the adaptive algorithm.

C. Non linear Processor:

The non-linear processor evaluates the residual echo, which is nothing but the amount of echo left over after the signal has passed through the adaptive filter. The nonlinear processor removes all signals below a certain threshold and replaces them with simulated background noise which sounds like the original background noise without the echo.

ADAPTIVE FILTERING

An adaptive filter is no more than a digital filter, which can adjust its characteristics. It adapts to changes in its input signals automatically according to a given algorithm. The algorithm will vary the coefficients according to a given criteria, typically an error signal to improve it's performance. An adaptive filter is a digital filter combined with an adaptive algorithm, which is used to modify the coefficients of the filter. Adaptive filters are used in many diverse applications in today's world in telephone echo canceling, radar signal processing, equalization of communication channels etc.

Adaptive filters are useful

1. when the characteristics of a given filter must be variable
2. when the spectrum of a given signal overlaps with the noise spectrum
3. If the frequency band occupied by noise is unknown or may vary with time.

In most real world scenarios adaptive filters are realized using **Finite Impulse Response** (FIR) filters, since they are guaranteed to be stable and are simple to use. It is often the case that signals are represented in the frequency domain to enable the use of discrete transforms that reduce the processing required in signal processing applications. Of all the transforms, the Fourier Transform is the most widespread.

The advantages of the Fourier Transform over other transforms include:

- The efficiency of the Fast Fourier transform
- Adequate representations of data for even short data lengths
- More faithful representation of data
- Components are sinusoidal and are not distorted when transmitted over linear systems
- A high degree of familiarity and thus a lot of development

D. Frequency domain Transforms

1. Discrete Fourier Transform:

The Discrete Fourier Transform is a form of the Fourier Transform that can be used to transform discrete time signals to the frequency domain. The equation used to calculate the Discrete Fourier Transform is shown below

$$X(k) = \sum_{n=0}^{N-1} x(n)\exp(-j2\prod kn/N)$$

For k=0, 1, 2….N-1 , Where k is the harmonic number of the transform component.

2. The Fast Fourier Transform

The Fast Fourier Transform (FFT) is an algorithm used to compute the Discrete Fourier Transform (DFT) of a vector ***x*** or in other words to convert the vector ***x*** to the frequency domain. It utilizes the built in redundancy in the DFT to decrease the number of calculations necessary and thus make the algorithm more efficient. The Inverse FFT (IFFT) is a version of the FFT that converts signals back to the time domain.

3. The Computational Complexity

To compute the DFT directly N^2 complex multiplies is required, if the FFT algorithm is used the number of operations for N a power of 2 reduces to ($N/2$) $\log_2$ ($2N$) complex multiplies + N complex adds**.** To maximize the efficiency of the FFT it is important that the block length N is an integer power of two. The block length being the length of the input block for which the FFT will be calculated.

4. The Overlap-Save Algorithm:

The Overlap-Add and the Overlap-Save are the two main fixed frequency domain algorithms implementing block convolution and as we know that convolution and digital filtering are the same. In this project we will focus on the Overlap-Save method since it is the most computationally efficient. In the Overlap-Save algorithm the input sequences overlap, a 50% overlap is generally considered the most efficient. The current input sequence values are composed of the current input block concatenated with the previous input block.

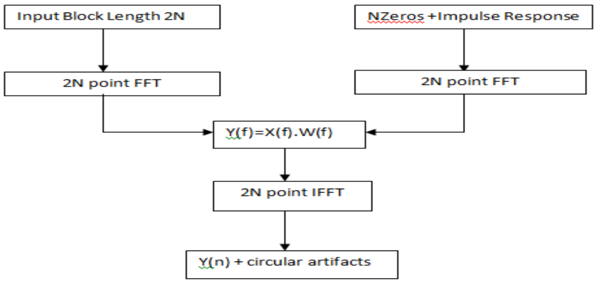

Fig 3: Block Diagram for the overlap-save algorithm

This means that the last N samples of the current input sequence are saved for concatenating with the next input block. The circular convolution computed results in the rotation of circular artifacts, which appear as the last N samples of the output of the IFFT. These samples are simply discarded and the remaining samples are concatenated to give the output of the overlap save algorithm. The algorithm is implemented as illustrated in Fig 3. Let N be the length of the impulse response of the system. The length of the input sequence is then twice this at $2N$.N zeros are added to the impulse response if necessary so that the result of the FFT will be the same length as that of the FFTs of the input sections. The FFT of the impulse response is calculated and stored in memory because it will remain unchanged.

$$W(k) = FFT\{w(k)\}$$

Next the current block of the input is taken and the FFT of it is calculated, for the first block there are N zeros in front of it for subsequent blocks the previous input block precedes the current input block.

$$X(k) = FFT\{x(n)\}$$

The two FFTs are now multiplied, that is to say that each element in one of the arrays will be multiplied by the corresponding element in the other. This procedure corresponds to convolution in the time domain.

$$Y(k) = X(k).W(k)$$

The IFFT of $Y(k)$ must now be calculated to bring the results back to the time domain.

$$y(n) = IFFT\{Y(k)\}$$

The second half of this result is dumped[1] for each convolution. The first half is added to an array as the output of the filter for the given input block. It is a fine algorithm so long as the filter coefficients are known.

MDF ALGORITHM

When the filter length, N, is large and a block length, L, much smaller than N is used, an efficient implementation of the fast-LMS (FLMS) algorithm can be derived by dividing (partitioning) the convolution sum into a number of smaller sums and proceeding as discussed below. The resulting implementation is called the partitioned FLMS or multi-delay frequency domain (MDF) algorithm.

Let us assume that N = P.M, where p and m are integers, and note that the convolution sum can be written as

$$y(n) = \sum_{l=o}^{p-1} yl(n)$$

where $$yl(n) = \sum_{i=0}^{M-1} wi + lM.x(n-lM-i)$$

To develop a frequency domain implementation of the convolutions , we can choose block length L=M and divide the input data into blocks of length 2M samples such that the last m samples of say, the k th block are same as first M samples of the (k+1) th block. Then, the convolution sum in can be evaluated using circular convolution of these data blocks with the appropriate weight vectors having been padded with M zeros. Using x(kM+M-1) to represent the newest sample in the input , we can define the vectors.

xF,l**(**k**)** = **FFT**([x((k-l)M-M) x((k-l)M-M+1)…….x((k-l)M+M-1)]) (1)

wF,l(k) **= FFT**([wlM(k) wlM+1(k)……..wlM+M-1(k) 0 0…….0])

yl(k) = [yl(kM) yl(kM+1) ……………yl(kM+M-1)]

and note that

yl(k) = the last M elements of **IFFT**(**w**F,l(k) . **xF,l**(k)) (2)

where . denotes multiplication on an element by element basis, k, block index, and l is the partition index. We also define

y(k) = [y(kM) y(kM+1)……………..y(kM+M-1)]

and y(k) = $$\sum_{l=0}^{P-1} yl(k) \qquad (3)$$

Furthermore, from (1) we note that

xF,l**(**k**) = x**F,0**(**k-l**)** (4)

International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.2, Issue No.10, pp : 1017-1021 1 Oct. 2013

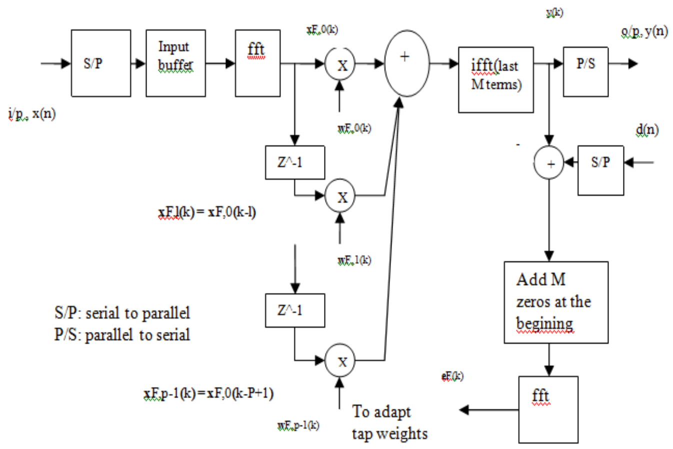

Fig 4: BLOCK DIAGRAM OF MDF ALGORITHM

Substituting (2)in(3), interchanging the order of summation and IFFT, and using(4), we obtain

y(k) = the last M elements of IFFT

$$\left(\sum_{l=0}^{p-1} wF,l(k) \otimes xF,0(k-l)\right)$$

Based on this result, the block diagram of the MDF algorithm may be proposed as depicted in above Fig . Here the delays, the z^-1 s, are in the unit of block size and the thick lines represents frequency domain vectors. Implementation of the summation on the right-hand side of equation can also be considered as parallel bank of 2M transversal filters, each of length P, with the j the filter processing the frequency domain samples belonging to the jth frequency bin,

For j = 0,1,....,2M -1 ,The adaptation of the filter tap weights is done according to the recursions

$$\mathbf{w}F,l(k) = \mathbf{w}F,l(k)+\mu(k)\mathbf{.x}F,0(k\text{-}l)\mathbf{.}\ \mathbf{e}F(k),$$

for l = 0,1,2,……,P-1, where μ(k) is the vector of the associated step-size parameters.

$$\mathbf{e}F(k) = \begin{bmatrix} d(k)-y(k) \\ 0 \end{bmatrix}$$

d(k) = [d(kM) d(kM + 1) …………..d(kM + M -1)] and 0 is the length zero column vector.

Recursion corresponds to the unconstrained MDF algorithm.

Advantages of MDF algorithm

1) Efficient Use of Hardware:

For an adaptive filter, a 2N-point FFT is generally used for an N-point weight factor. Most of the available FFT or DSP chips are designed and optimized for small size FFT, typically 256 point. To implement an acoustic echo canceller of a few thousands taps, several FFT chips are cascaded together with external memory to form a larger FFT configuration, which is rather inefficient and expensive.

2) Long Block Delay:

Since the FLMS algorithm implements block processing, if the weight size N = 1024, the first output y(n + 1) needs to wait after the last output y(n + 1024) of the same block is processed or a delay of 128 ms for an 8 kHz sampling rate. Such a long delay would make the echo more annoying.

3) Large Quantization Error in FFT:

As the size of an FFT becomes larger, the number of multiplications and scaling increases. This causes extra quantization error.

CORTEX-M4

The ARM Cortex-M4 processor is the latest embedded processor. It has a 32-bit core with built-in integer DSP, and an optional floating point unit. It is an efficient, easy –to-use blend of control and developed to address digital signal control markets that demand an efficient, easy-to-use blend of control and signal processing capabilities. The combination of high-efficiency signal processing functionality with the low-power, low cost and ease-of-use benefits of the Cortex-M family of processors is designed to satisfy the emerging category of flexible solutions specifically targeting the applications in motor control, automotive, power management, embedded audio and industrial automation markets. The Cortex-M4F is a processor with the same capability as the Cortex-M4 processor; it includes floating point arithmetic functionality.

ARM Cortex-M4 Specification

1. It has Thumb/Thumb-2 ISA support
2. It has DSP Extensions like Single cycle 16,32-bit MAC, Single cycle dual 16-bit MAC,8,16-bit SIMD arithmetic, Hardware Divide (2-12 Cycles)
3. It has single precision floating point unit IEEE 754 (std).
4. It has a 3-stage + branch speculation pipeline
5. It has a performance efficiency of 2.19 Core Mark/MHz - 1.25 DMIPS/MHz
6. It has Optional 8 region MPU with sub regions and background region
7. It has Non-mask-able Interrupt (NMI) + 1 to 240 physical interrupts and 8 to 256 interrupt priority levels.

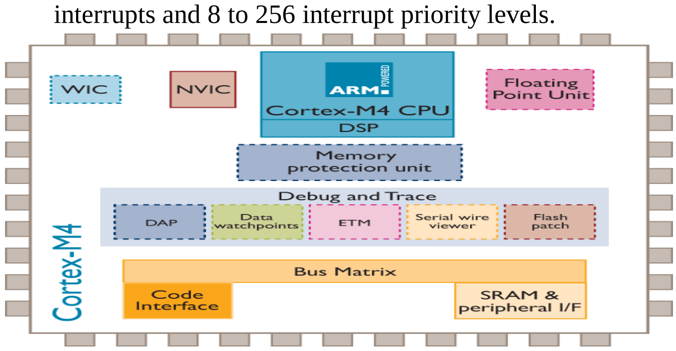

Fig 5: Cortex-M4 Block Diagram

Cortex-M4 processor provides a highly efficient solution for digital signal control (DSC) applications, while maintaining the industry leading capabilities of the ARM Cortex-M family of processors for advanced microcontroller (MCU) applications. It has a single-cycle(MAC) unit with optimized single instruction multiple data (SIMD) instructions and it has a target frequency of 150MHz. The processor consumes

International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.2, Issue No.10, pp : 1017-1021 1 Oct. 2013

40µW/MHz of power and has 65K gates with the optional floating-point unit adding an additional 25K gates.

III. RESULTS

To test an Acoustic echo canceller the required inputs are

- far-end speech signal
- microphone signal

And the output of AEC is echo free signal.For the illustration purpose short duration speeches were considered Far-end speech signal is 'far_end_speech_wav' is given by

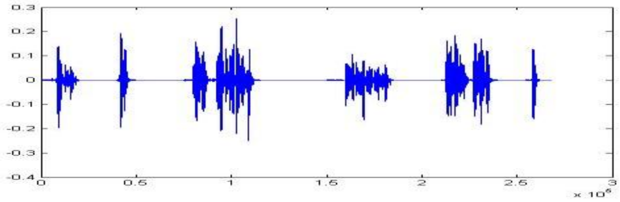

Fig 6: Far-end speech signal

The output of micro phone consists of near-end speech added with echo of far end speech.

Microphone signal is 'micro_phone_out.wav' is given by

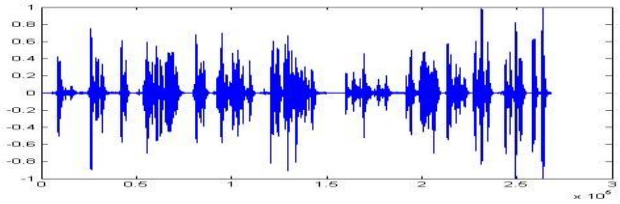

Fig 7: Micro-phone output

These two speech signals are given to AEC as inputs and AEC will produce echo free speech of the near-end speaker. The output of AEC 'AEC_output.wav' is given by

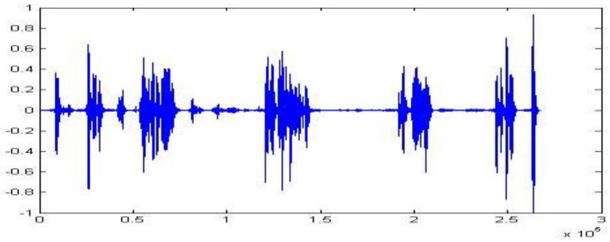

Fig 8: AEC output

IV.CONCLUSION

In the present work, the MDF adaptive filter is implemented on Cortex-M4 processor to eliminate the acoustic echo of the far-end speaker. It requires less memory storage, small FFT size. In performance, the MDF adaptive filter has a smaller block delay and is faster. This is achieved by updating the weight vectors more often and reducing the total execution time in most of the processor. Furthermore, the total number of blocks needed can be changed dynamically without interrupting the normal operation. The MDF adaptive filter is most suitable for real-time applications implemented on the hardware.

REFERENCES

i. J. S. Soo, K. K. Pang " Multidelay block frequency adaptive filter", IEEE Trans. Acoust. Speech Signal Process., Vol. ASSP-38, No. 2, February 1990.

ii. N.Fumya, Y. Itoh, Y. Maruyama, and Araseki, "Audio conference equipment with acoustic echo canceller," NEC Res. Develop., pp. 18-23, Jan. 1985

iii. John G. Proakis, Dimitris G.Manolakis, "Digital signal processing: principles, algorithms and applications", Prentice Hall, March 2002.

iv. S. Haykins , "Adaptive Filter Theory" , Prentice Hall ,New Jersey, 1996.

v. Farhang-Boroujeny, "Adaptive Filters, Theory and Applications" John Wiley and Sons, New York, 1999

vi. A. Feuer, "Performance analysis of the block least mean square algorithm" IEEE Trans. Circuits Syst., vol. CAS-32, pp. 960-963,Sept. 1985.

vii. http://www.arm.com/products/processors/cortex-m/cortex-m4-processor.php

viii. http://ics.nxp.com/products/mcus/cortex-m4/

ix. S Raghavendran, Implementation of an acoustic echo canceller Using Matlab.

x. W.-P. Ang and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size lms algorithms," IEEE Transactions on Signal Processing, vol. 49, no. 4, pp.s 805–810, 2001.