

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Echo Cancellation for Hands-Free Systems

Artur Ferreira and Paulo Marques
Instituto de Telecomunicações
Instituto Superior de Engenharia de Lisboa
Portugal

1. Introduction

Echo is defined as the delayed and attenuated version of the original signal produced by some device, such as a loudspeaker. As a consequence a person listens to a delayed replica of its own voice signal. This is an undesired effect that appears whenever the output signal is fed back into the system's input and it can be quite disturbing on voice conversations.

Echo arises in long distance communication scenarios such as hands-free systems Hänsler (1994); Jeannès et al. (2001); Liu (1994), *voice over internet protocol* (VoIP) Witowsky (1999), teleconferencing Kuo & Pan (1994), mobile phone conversation, and satellite communications among others.

In order to minimize or even remove the presence of echo in communications, echo suppression and echo cancellation techniques have been proposed in the last three decades Sondhi (2006). An echo suppressor is a voice-operated switch that disconnects the communication path (or introduces a very large attenuation) whenever some decision mechanism indicates that we are in the presence of echo. The emitting circuit is disconnected whenever we have signal on the reception part of the circuit; the reception circuit is disconnected whenever we have signal emission. Their behavior is not adequate for cross conversation (full duplex) scenarios. Echo suppressors were the first approach to this problem. In the last decade, due to their unsatisfactory results, they have been replaced by digital echo cancelers. An echo canceler device, as opposed to an echo suppressor, does not interrupt the echo path; it operates by removing (subtracting) the detected echo replicas from the information signal. The term usually coined for the cancellation of echoes with acoustic coupling is *acoustic echo cancellation* (AEC) Gilloire & Hänsler (1994).

In the past years, adaptive filtering techniques Haykin (2002); Sayed (2003); Widrow & Stearns (1985) have been employed for the purpose of AEC Breining (1999); Widrow et al. (1975). Typically, these techniques rely on the use of *finite impulse response* (FIR) filters Oppenheim & Schaffer (1999); Veen & Haykin (1999) whose coefficients are updated along the time by an efficient rule guided by some statistical criterion. Usually, one employs a gradient descent technique in order to minimize some cost (error) function. The most popular of these techniques is the Widrow-Hoff *least mean squares* (LMS) algorithm as well as its variants, that minimize the *mean square error* (MSE) between two signals. Moreover, in many cases such as real-time conversations over mobile phones, AEC algorithms must run in real-time to be useful. We thus have the need for efficient implementations of echo cancellation techniques on digital embedded devices like *field programmable gate array* (FPGA) and/or *digital signal processor* (DSP), to fulfill real-time requirements of many applications, these days.

This chapter reviews and compares existing solutions for AEC based on adaptive filtering algorithms. We also focus on real-time solutions for this problem on DSP platforms. Section 2 states the echo cancellation problem. Section 3 reviews some basic concepts of adaptive filtering techniques and algorithms. Section 4 describes some existing solutions for AEC. Section 5 details real-time implementations of AEC systems with DSP from Texas Instruments. Section 6 presents some experimental results and Section 7 ends the chapter with some concluding remarks and future directions and challenges for AEC techniques.

2. The echo cancellation problem

2.1 What is echo?

Echo signal is defined as the delayed and attenuated version of the original signal produced by some device, such as a loudspeaker. Let's consider some signal $x(t)$ and its attenuated and delayed version

$$x_d(t) = \alpha x(t - t_d), \quad (1)$$

where α is the attenuation factor and t_d is time delay of the echo replica. Whenever we have a delay (t_d) larger than 35 ms, echo becomes perceptible to the listener Oh et al. (1994). As t_d increases, the more annoying is the echo effect. Larger values of t_d imply that we should have larger attenuation on the echo signal to minimize the undesired echo effect. For satellite communications, a typical delay value between two end-points is about 270 ms. This leads to a total round trip delay of at least 540 ms, including the terrestrial circuits Texas Instruments (1986). In these situations, we must apply a large attenuation to the echo signal, to make conversation possible. The worst case happens when the feedback is sustained (non-decaying); this makes the most annoying effect named as howling and it happens whenever α is not small enough.

The existence of undesired echo is a well-known problem that frequently arises in telecommunications, being most probable to happen in long distance communications and most problematic in voice conversations. Therefore, echo cancellation is needed for long-distance communications which have shown a growing use in the last decade. For instance, in the VoIP application the network load changes the transmission time and the time delay of the echo(es). Another challenging problem is that the echo path is not static because the channel characteristics change over time, due to many factors such as the distance between the loudspeaker and the microphone. Fig. 1 shows the existence of acoustic echo. The received signal from the far end talker, $r[n]$, is transmitted through the loudspeaker to the near end. A version of this signal is received by the microphone (due, for example, to direct coupling between the loudspeaker and the microphone), together with the near end speech, constituting the received signal from the acoustic channel, $r_A[n]$. The echo path is defined by functions f and g on both ends; these functions represent the linear or non-linear behavior of the echo path.

2.2 Sources of echo

In the telecommunications field, we can find many sources of echo caused by long distance communications and/or the hands-free voice conversation setup Sondhi (2006). For instance we have echoes:

- on the hands-free VoIP setup, in which we have a microphone and a loudspeaker;
- on a (satellite) mobile telephone connection;

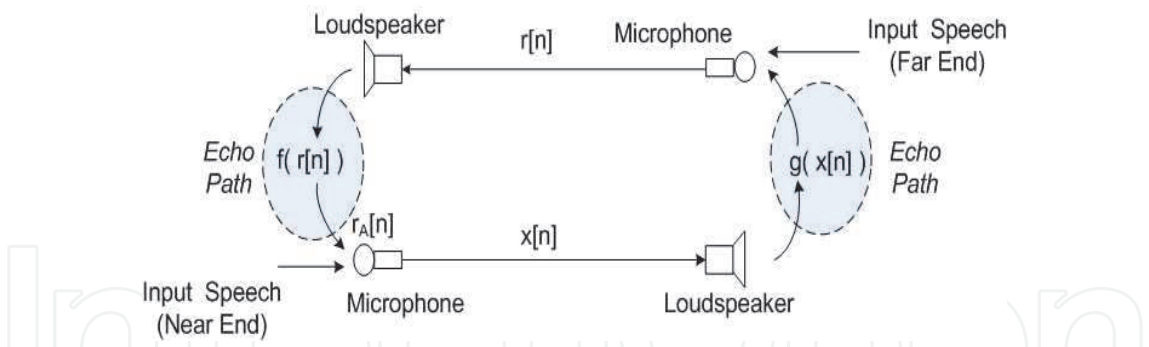


Fig. 1. The acoustic echo scenario. A version of the received signal trough the loudspeaker is fed back into the microphone re-entering the system leading to undesired feedback. The worst case happens when the feedback is sustained (non-decaying); this makes the most annoying effect named as howling. Functions f and g represent the echo path on both ends.

- whenever someone makes a speech on some room; each room has its own acoustic conditions leading to echo Antweller & Symanzik (1995);
- the two-wire/four-wire conversion Sondhi & Berkeley (1980) in telephony carried out by a hybrid circuit on a telephone line, as depicted in Fig. 2, in which we consider a simplified connection between two subscribers S_1 and S_2 .

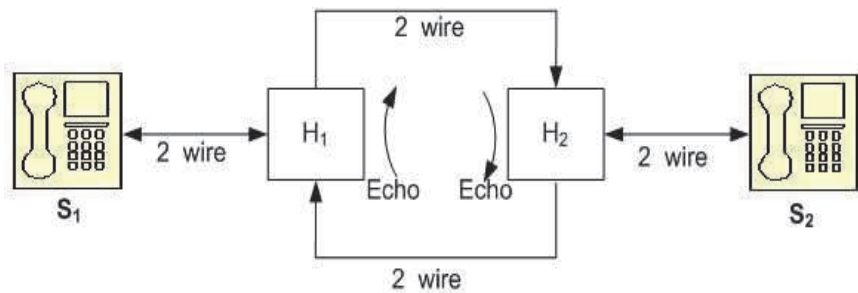


Fig. 2. Simplified long distance connection between two subscribers, using a hybrid circuit for the two-wire/four-wire conversion. The impedance mismatch on the two-wire/four-wire conversion originates the return of a portion of the emitted signal. This causes echo on both ends of the conversation system.

The subscriber loop connects the analog telephone with a two-wire line. In order to establish a connection, the central office must connect the two-wire line from one subscriber to another. This way, long distance telephone connections are four-wire connections with two-wires for transmission and the other two for reception. The hybrid circuit is a device that establishes the connection and conversion between the two-wire and the four-wire circuits. The connection between S_1 and H_1 (or between S_2 and H_2) is a two-wire connection between the subscriber and the central office. Between central offices, we have four-wire connections (between H_1 and H_2). Each two-wire connection is usually designated as *the subscriber loop*; in this portion of the circuit, both directions of communication are supported in the same pair of wires. The hybrid circuit H_1 converts the signal from S_2 to the two-wire connection to S_1 , whitout back reflection of energy from this signal. In practice, this is not possible to accomplish because due to many varying characteristics such as the length of the subscriber loop, the individual subscriber devices and line impedance; these factors altogether inhibit the perfect separation between emission and reception signals. Since there is some energy reflection from

the emitted signal, as a consequence S_2 (or S_1) receives a delayed version of its own voice signal with some attenuation and distortion.

Applications such as hands-free telephony, tele-conferencing and video-conferencing require the use of *acoustic echo cancellation* (AEC) techniques to eliminate acoustic feedback from the loudspeaker to the microphone Gay & J.Benesty (2000).

Echo cancellation is usually achieved by using an adaptive filter which attempts to synthesize a replica of the echo signal and subtract it from the returned signal. An adaptive filter changes its coefficients along the time; as a consequence it changes its frequency response in order to satisfy the adaptation criterion. This is the principle illustrated in Fig. 3 in the AEC context. The adaptive filter operates on the voice signal and tries to replicate the echo signal, which is

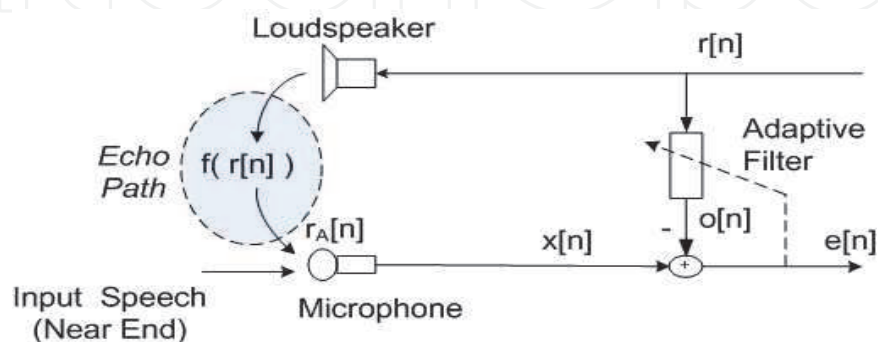


Fig. 3. The acoustic echo cancellation scenario for the setup of Fig. 1. Echo cancellation is achieved by using an adaptive filter which attempts to synthesize a replica of the echo signal and subtract it from the returned signal.

subtracted from the emitted signal. The adaptive filter imitates the echo path thus canceling its effects. In the case of the two-wire/four-wire conversion, depicted in Fig. 2, AEC is performed with the block diagram of Fig. 4.

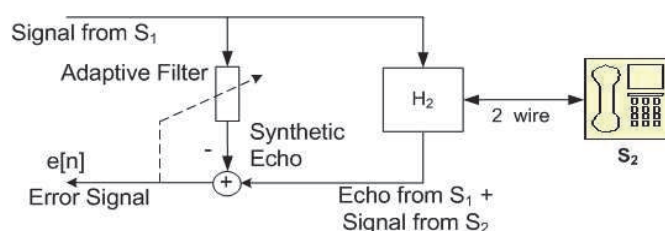


Fig. 4. Echo cancellation on the telephone line using an adaptive filter. The adaptive filter compensates the non-ideal behavior of the hybrid circuit.

The adaptive filter synthesizes a replica of the echo, which is subtracted from the returned signal. This removes/minimizes the echo without interrupting the echo path. The adaptive filter compensates the undesired effects of the non-ideal hybrid circuit.

3. Adaptive filtering techniques

In order to model the referred time-changing echo characteristics and to cancel its undesired effects on the conversation, adaptive filtering Haykin (2002); Sayed (2003); Widrow & Stearns (1985) has been used extensively in the last three decades (see for instance, Benesty et al. (2001); Greenberg (1998); J.Ni & Li (2010); Krishna et al. (2010); Marques et al. (1997); Sondhi & Berkeley (1980)). The main reasons for the success of adaptive filtering to solve the AEC problem are:

- its efficiency, allowing real-time implementations;
- its ability to cope with statistically changing environments;
- its adequate results in noisy environments.

Due to the time varying characteristics of the echo path and the devices in the circuit, it is not possible to cancel echo with (static) classic filtering which removes some frequency band. The filter coefficients and thus its frequency response must change along the time to efficiently model (imitate) the behavior of the echo path, thus leading to echo cancellation.

3.1 FIR and IIR filtering structures

A *finite impulse response* (FIR) filter Oppenheim & Schaffer (1999); Veen & Haykin (1999) has difference equation given by

$$o[n] = \sum_{k=0}^{M-1} w_k x[n-k], \quad (2)$$

where w_k with $k \in \{0, \dots, M-1\}$ are the filter coefficients and $x[n-k]$ are the (past) samples on the input of the filter. Using vector notation, we compute each output sample of the filter by the inner product between row vectors $\mathbf{w} = [w_0, w_1, \dots, w_{M-1}]$ and $\mathbf{x} = [x[n], x[n-1], \dots, x[n-(M-1)]]$,

$$o[n] = \mathbf{w}\mathbf{x}^T = \mathbf{x}\mathbf{w}^T. \quad (3)$$

The *infinite impulse response* (IIR) Oppenheim & Schaffer (1999); Veen & Haykin (1999) difference equation is

$$o[n] = \sum_{k=0}^{M-1} w_k x[n-k] + \sum_{k=1}^{N-1} v_k o[n-k], \quad (4)$$

with feedback coefficients v_k . Depending on the value of the feedback coefficients, the filter can become an unstable system. In order to prevent this situation to happen, adaptive filtering algorithms with IIR filters must take additional measures to assure the stability of the filter. Fig. 5 depicts the FIR and IIR structures for digital filtering, corresponding to (2) and (4), respectively. In the case in which these structures are applied on adaptive filtering scenarios, the filter coefficients are periodically updated.

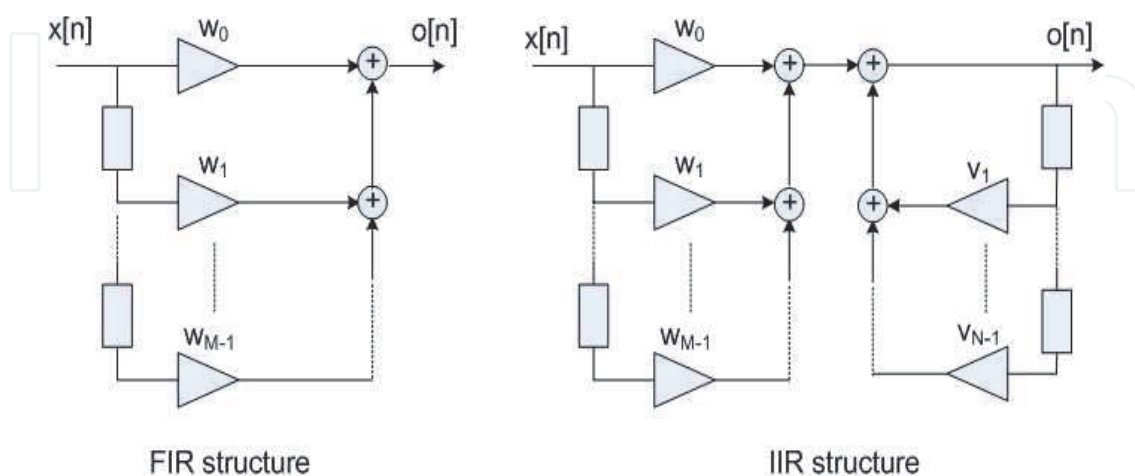


Fig. 5. FIR and IIR structures for digital filtering. When these structures are applied on adaptive filtering problems, their coefficients are time-changing, being updated by some rule.

3.1.1 Analysis on the z-plane

In the z-transform domain, the FIR filter has a rational transfer function

$$H(z) = \sum_{k=0}^{M-1} w_k z^{-k} = w_0 \prod_{k=1}^M (1 - q_k z^{-1}), \quad (5)$$

with M zeros represented by q_k , whereas a IIR filter has a transfer function

$$H(z) = \frac{\sum_{k=0}^{M-1} w_k z^{-k}}{1 - \sum_{k=1}^{N-1} v_k z^{-k}} = \frac{w_0 \prod_{k=1}^M (1 - q_k z^{-1})}{\prod_{k=1}^N (1 - r_k z^{-1})}, \quad (6)$$

with M zeros (given by q_k) and N poles (given by r_k). The zeros correspond to the direct connections between input and output whereas the poles indicate the feedback connections. It is well-known that the use of poles can lead to accomplish a given filter specification more easily than using only zeros. However, for causal filters poles cause instability whenever placed outside the unit circle in the z-plane; the adaptation algorithm has to assure stability. IIR filters are, by definition, systems with infinite impulse response and thus we can theoretically accommodate any echo path of largest length. In the case of IIR filters, the adaptive filtering algorithm must take additional steps in order to keep the N poles of the filter inside the unit circle on the z-plane. Since a FIR filter is stable independently of its coefficients, adaptive filtering algorithms usually employ FIR filtering instead of IIR filtering.

3.2 Statistical and adaptive filtering

Fig. 6 shows the block diagram of a typical statistical filtering problem. The adaptive filter operates on the sequence $x[n]$ which is made up by the desired signal with uncorrelated additive white noise. The impulse response of the adaptive filter is given by $w[n]$. At each time instant n , the filter outputs $o[n]$ which is an estimate of the desired response $d[n]$. Since both these signals are instances of stochastic processes, the error estimate $e[n]$ has specific statistical properties. The goal of statistical filtering is to minimize the estimation error, according to some statistical criterion, like the *mean square error* (MSE) between the output and the desired signal.

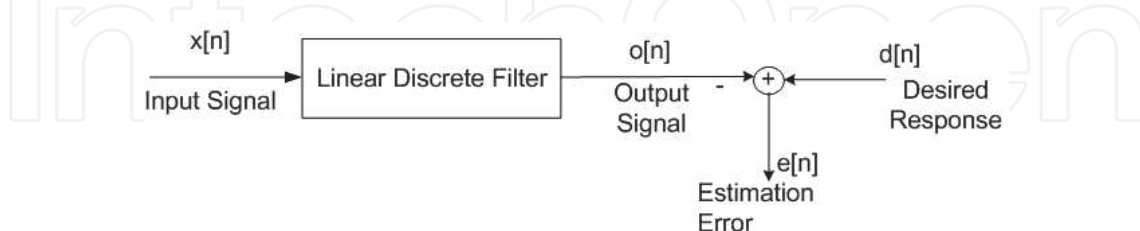


Fig. 6. Block diagram of statistical filtering. The discrete filter is applied to the input signal $x[n]$ and produces output $o[n]$, which is compared against the desired signal $d[n]$. The error signal (to be minimized) is the difference between $o[n]$ and $d[n]$. The discrete filter is learned with the statistics from the input signal.

Fig. 7 shows the block diagram of a typical adaptive filtering application. The error signal is used as input to the coefficient update algorithm of the adaptive (FIR or IIR) filter. As time goes by and samples move along vector \mathbf{x} , the coefficients in vector \mathbf{w} are updated by some

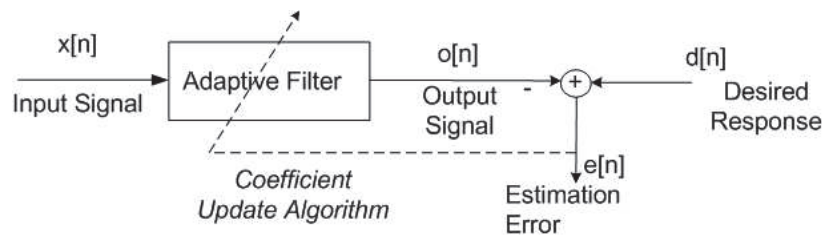


Fig. 7. Block diagram of adaptive filtering. The discrete filter is applied to the input signal $x[n]$ and produces output $o[n]$, which is compared against the desired signal $d[n]$. The error signal is used as input to the coefficient update algorithm of the adaptive filter. The adaptive filter coefficients are updated according to the error signal produced at each time instant.

rule using a statistical criterion. The minimization of the *mean square error* (MSE) between the output and the desired signal, leads to minimization of the energy of the error signal. If we define the error signal as the difference between the desired signal $d[n]$ and the output signal $o[n]$, we get

$$e[n] = d[n] - o[n] = d[n] - \mathbf{x}\mathbf{w}^T = d[n] - \mathbf{w}\mathbf{x}^T. \quad (7)$$

The energy of the error signal is defined as

$$\begin{aligned} E_e &= \sum_{n=-\infty}^{\infty} e^2[n] = \sum_{n=-\infty}^{\infty} (d[n] - o[n])^2 \\ &= \sum_{n=-\infty}^{\infty} (d^2[n] - 2d[n]o[n] + o^2[n]). \end{aligned} \quad (8)$$

Each term of the sum in (8) is given by

$$\begin{aligned} e^2[n] &= d^2[n] - 2d[n]o[n] + o^2[n] \\ &= d^2[n] - 2d[n]\mathbf{x}\mathbf{w}^T + \mathbf{w}\mathbf{x}^T\mathbf{x}\mathbf{w}^T. \end{aligned} \quad (9)$$

Taking the expectation operator \mathcal{E} on (9) and since $\mathcal{E}[\mathbf{w}] = \mathbf{w}$ we get

$$\begin{aligned} \mathcal{E}[e^2[n]] &= \mathcal{E}[d^2[n]] - 2\mathcal{E}[d[n]\mathbf{x}\mathbf{w}^T] + \mathcal{E}[\mathbf{w}\mathbf{x}^T\mathbf{x}\mathbf{w}^T] \\ &= \mathcal{E}[d^2[n]] - 2\mathcal{E}[d[n]\mathbf{x}]\mathbf{w}^T + \mathbf{w}\mathcal{E}[\mathbf{x}^T\mathbf{x}]\mathbf{w}^T \\ &= \mathcal{E}[d^2[n]] - 2\mathbf{p}\mathbf{w}^T + \mathbf{w}\mathbf{R}\mathbf{w}^T, \end{aligned} \quad (10)$$

where \mathbf{R} is the square auto-correlation matrix of the input vector

$$\begin{aligned} \mathbf{R} &= \mathcal{E}[\mathbf{x}^T\mathbf{x}] \\ &= \begin{bmatrix} x^2[n] & x[n]x[n-1] & \dots & x[n]x[n-(M-1)] \\ x[n-1]x[n] & x^2[n-1] & \dots & x[n-1]x[n-(M-1)] \\ \vdots & \vdots & \ddots & \vdots \\ x[n-(M-1)]x[n] & x[n-(M-1)]x[n-1] & \dots & x^2[n-(M-1)] \end{bmatrix}, \end{aligned} \quad (11)$$

and \mathbf{p} is the vector with the cross-correlation between the desired and the input signal

$$\mathbf{p} = \mathcal{E}[d[n]\mathbf{x}] = [d[n]x[0], d[n]x[n-1], \dots, d[n]x[n-(M-1)]]. \quad (12)$$

The error surface is a multidimensional paraboloid with concavity aligned along the positive axis, leading to a surface with a single minimum. In order to find this minimum, we search for the optimal weight vector \mathbf{w}^* that leads to a null gradient

$$\nabla \frac{\delta \mathcal{E}[e^2[n]]}{\delta \mathbf{w}} = 0. \quad (13)$$

For stationary input signals, the minimization of (13) leads to the Wiener filter Orfanidis (2007); Widrow & Stearns (1985) solution that computes the optimal weight vector given by

$$\mathbf{w}^* = \mathbf{R}^{-1} \mathbf{p}, \quad (14)$$

leading to the optimal solution in the MSE sense. However, Wiener filter is not adequate for non stationary situations, because it requires prior knowledge of the statistical properties of the input signal. This way, Wiener filter is optimal only if these statistics match with the ones considered for the design of the filter. In cases which we do not have such information, it may not be possible to design the Wiener filter, or at least it can not be optimal. We can overcome this problem with a two step approach:

1. estimation of the statistical parameters of those signals;
2. compute the filter parameters.

For real-time applications, the computational complexity of this approach is prohibitive. An efficient method to solve this problem consists on the use of an adaptive filter. These filters exhibit a satisfactory behavior in environments in which there is no prior knowledge of the statistical properties of the signals under consideration. The adaptive filter designs itself automatically, with some recursive learning algorithm to update its parameters, minimizing some cost function. It starts with a set of parameters which reflect the absence of knowledge of the environment; if we have a stationary environment the adaptive filter parameters converge, at each iteration step, to the Wiener filter solution given by (14) Haykin (2002); Sayed (2003); Widrow & Stearns (1985). On a non-stationary environment, the adaptive filtering algorithm provides a way to follow the variation of the statistical properties of the signal, along the time, provided that this variation is slow enough.

Adaptive filtering has been used in a variety of applications such as communications, radar, sonar, seismology, and biomedical engineering. Although these applications come from such different fields, they all share a common characteristic: using an input vector and a desired response, we compute an estimation error; this error is then used to update the adaptive filter coefficients. Table 1 summarizes classes and some applications of adaptive filtering. The echo cancellation problem belongs to the class IV-Interference Cancellation.

3.2.1 Echo cancellation

The block diagram for AEC depicted in Fig. 2, operates by synthesizing an echo replica subtracting it from the received signal. This synthetic echo is generated from the speech signal. The result of this subtraction is named acoustic error signal, which can be used to adjust the filter coefficients. In the ideal case, the adaptive filter has a transfer function which is equal to the echo path, leading to a total echo cancellation. The principle shown in Fig. 6 is applied in Fig. 2. Notice that if the echo paths given by functions f and g in Fig. 2 are non-linear, it is not possible to achieve full echo cancellation because we are using linear filters to imitate the echo path. In practice, linear filters achieve adequate results.

Class	Applications
I. Identification	System Identification
	Terrestrial Layer Modelization
II. Inverse Modelization	Predictive Deconvolution
	Adaptive Equalization
III. Prediction	Linear Predictive Coding
	Auto-Regressive Spectral Analysis
	Signal Detection
IV. Interference Cancellation	Noise Cancellation
	Echo Cancellation
	Adaptive Beamforming

Table 1. Classes and applications of adaptive filtering techniques.

The adaptive filter coefficients adjustment has been addressed using the *least mean squares* (LMS) adaptive filtering algorithm and its variants Haykin (2002); Widrow & Stearns (1985). The choice of the adaptive filter adjustment coefficients must be made taking into account some important characteristics of the algorithm such as:

- rate of convergence and precision;
- numerical complexity;
- filter structure and stability.

The rate of convergence is defined as the number of iterations that the algorithm requires, operating on stationary signals, to approximate the optimal Wiener solution. The higher the rate of convergence, the faster the algorithm adapts to non stationary environments with unknown characteristics.

The numerical complexity is the number of operations needed to complete an iteration of the algorithm. Depending on the adaptation algorithm and on the processor where the algorithm runs, it is possible to have numerical problems. The most common source of problems of this kind is the so-called *finite precision effects*, due to limited number of bits used to represent data types for coefficients and samples. As the algorithm performs its computations, it is possible to accumulate quantization errors. If this situation is not monitored, the adaptive filter coefficients may enter into an overflow (or underflow) problem. These factors prevent that the adaptive algorithm converges to an acceptable solution. In the worst case of overflow and underflow, we say that the algorithm is numerically unstable.

The filter stability is assured by choosing a FIR filter; whatever are its coefficients, a FIR filter is always stable since it has no feedback from the output to the input. If the filter input becomes zero at a given time instant, then its output will be zero after M/F_s , where M is the order of the filter and F_s is the sampling frequency. The longest echo path (time delay) that we can accommodate is given by the order of the filter; long echo paths lead to higher order filters that are computationally more demanding to store in memory and to update in real-time. Thus, the computational complexity that we can put into the adaptive filter limits the longest echo path.

On the other hand, if we use an adaptive IIR filter care must be taken to assure stability. In any case (FIR or IIR filter) we have a time varying impulse response, due to the adaptation of the filter coefficients.

3.3 Adaptive filter update: Least mean squares algorithm and variants

In this section, we review the most common adaptive filtering techniques. For the purpose of explanation we consider adaptive FIR filters. The *least mean square* (LMS) algorithm operates by updating the filter coefficients \mathbf{w} according to a gradient descent rule given by

$$\mathbf{w}_{(i+1)} \leftarrow \mathbf{w}_{(i)} + 2\mu e_{(i)} \mathbf{x}_{(i)}, \quad (15)$$

where $\mathbf{w}_{(i)}$ is the value of the coefficient vector \mathbf{w} at time instant i , μ is the step size, $e_{(i)}$ is the error signal and $\mathbf{x}_{(i)}$ represents the present and previous samples at the filter taps, at time instant i .

The LMS step μ must be chosen carefully in order to assure proper convergence. A small μ will assure convergence but the rate of convergence can be quite slow that is not adequate for a real-time conversation. A large μ can cause that the LMS algorithm does not find an adequate local minimum, thus leading to unsatisfactory echo cancellation. It can be shown Haykin (2002); Widrow & Stearns (1985) that the choice

$$0 < \mu < \frac{1}{(M+1)P_{(i)}}, \quad (16)$$

assures adequate convergence rate; M is the order of the adaptive filter and $P_{(i)}$ is the average power of the signal present in the input of the filter. In Özbay & Kavsaoglu (2010), the choice of the optimal value for the LMS adaptation step is discussed.

3.3.1 Variants of LMS

Some acceleration techniques have been proposed to improve LMS convergence. Moreover, in the presence of speech signals, one must be careful in the coefficient update because speech signals have a power which exhibit a large dynamic range, making ineffective the use of a constant step size. To overcome this difficulty, the *normalized least mean squares* (NLMS) Birkett & Goubran (1995); Kuo & Lee (2001) algorithm (a variant of LMS) uses a variable step size, $\mu_{(i)}$, computed by

$$\mu_{(i)} = \frac{\eta}{a + P_{(i)}}, \quad (17)$$

in which $\eta > 0$, $P_{(i)}$ is the instantaneous power of signal \mathbf{x} at time index i and $a > 0$ is a suitably chosen value to avoid numerical problems caused by zero division. This way, we have an adaptation step given by

$$\mathbf{w}_{(i+1)} \leftarrow \mathbf{w}_{(i)} + 2\mu_{(i)} e_{(i)} \mathbf{x}_{(i)}, \quad (18)$$

for NLMS, which has a relatively low convergence speed but it is quite stable and has low complexity. The *leaky least mean squares* (LLMS) algorithm is another variant of LMS Kuo & Lee (2001), which introduces a leakage factor β such that ($0 < \beta < 1$) into the coefficient update

$$\mathbf{w}_{(i+1)} \leftarrow \beta \mathbf{w}_{(i)} + 2\mu e_{(i)} \mathbf{x}_{(i)}. \quad (19)$$

Thus, for the following iteration we use a portion of the coefficient value (they have leaks caused by β). The NLMS and LLMS can be used simultaneously in the coefficient update process, leading to

$$\mathbf{w}_{(i+1)} \leftarrow \beta \mathbf{w}_{(i)} + 2\mu_{(i)} e_{(i)} \mathbf{x}_{(i)}. \quad (20)$$

The *recursive least squares* algorithm (RLS) Haykin (2002) recursively computes the filter coefficients such that minimize a weighted linear least squares cost function. Notice that the LMS algorithm aims at minimizing the mean square error. RLS algorithm considers the input signals as deterministic, while for the LMS and variants these are considered as instances of stochastic processes. RLS has extremely fast convergence at the expense of high computational complexity, which makes it uninteresting for real-time implementations. The main drawback of RLS is its poor performance when the filter to be estimated changes its statistical properties. A LMS variant named *frequency-response-shaped least mean squares* (FRS-LMS) was proposed in Kukrera & Hocanin (2006) and shown to have good convergence properties. The FRS-LMS algorithm has improved performance when a sinusoidal input signal is corrupted by correlated noise. The algorithm shapes the frequency response of the transversal filter. This shaping action is performed on-line using an additional term similar to leakage β in LLMS shown in (19). This term involves the multiplication of the filter coefficient vector by a matrix, and it is computed efficiently with the *fast Fourier Transform* (FFT) Oppenheim & Schafer (1999). The authors show analytically that the adaptive filter converges in both the mean and mean-square senses. They also analyze the filter in the steady state in order to show its frequency-response-shaping capability. The experimental results show that FRS-LMS is very effective even for highly correlated noise.

The *decoupled partitioned block frequency domain adaptive filter* (DPBFAD) approach is quite demanding regarding computational requirements in the number of operations as well as the required memory, when compared to LMS, NLMS, and LLMS. The coefficients are updated in blocks, to allow real-time implementation, on the frequency domain. It requires the computation of the DFT (using FFT) and its inverse at each iteration.

Table 2 compares three algorithms that have proven effective for adaptive filtering, namely the *least mean squares* (LMS), *recursive least squares* (RLS), and fast RLS (also known as fast Kalman) Orfanidis (2007). This list is by no means exhaustive, since there are many different algorithms for this purpose.

Algorithm	Rate of Convergence	Complexity	Stability
LMS	Slow	Low	Stable
RLS	Fast	High	Stable
Fast RLS	Fast	Low	Unstable

Table 2. Rate of convergence, complexity, and stability of well-known adaptive filtering algorithms.

4. Solutions for echo cancellation

This section presents some solutions for the AEC problem. We focus on real-time solutions carried out over DSP platforms.

4.1 Short-length centered adaptive filter approach

In Marques (1997); Marques & Sousa (1996); Marques et al. (1997) we have an implementation of a long distance echo canceler which copes with double talking situations and exceeds the CCITT (now ITU-T, International Telecommunication Union-Telecommunications) G.165 recommendation levels for echo cancellation. The proposed solution is based on short length adaptive FIR filters centered on the positions of the most significant echoes, which are tracked by time-delay estimators. To deal with double talking situations a speech detector

is employed. The resulting algorithm enables long-distance echo cancellation with low computational requirements. It reaches high *echo return loss enhancement* (ERLE) and shows fast convergence. The key issue to use centered adaptive filters is that the echo-path impulse response is characterized mainly by two active regions, corresponding to the near-end and the far-end signal echo respectively, as shown in Fig. 8.

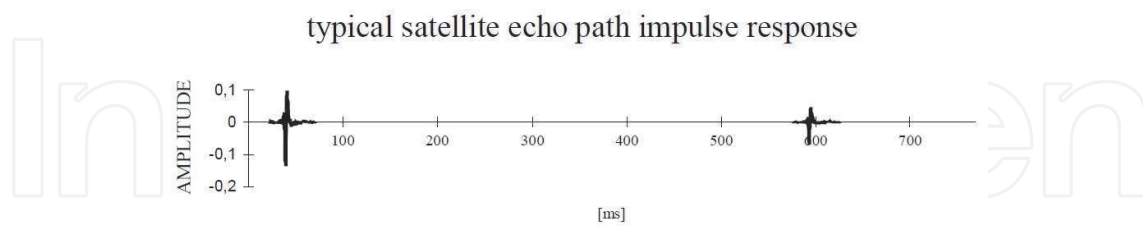


Fig. 8. Typical echo path impulse response (adapted from Marques et al. (1997)). We have two active regions corresponding to the near-end and far-end echo, respectively.

Each active region has a length usually much shorter than the total supported echo-path length. The proposed system is based on time-delay estimators to track the position of these active regions, where short-length adaptive filters have to be centered.

Fig. 9 shows the impulse response of an acoustic echo path, resulting from the direct coupling between the speaker and the microphone of an IRISTEL telephone. Although the supported echo path length is 64 delay elements, only a small region is active. Knowing its position and length, the adaptive filter has to adjust only the corresponding coefficients.

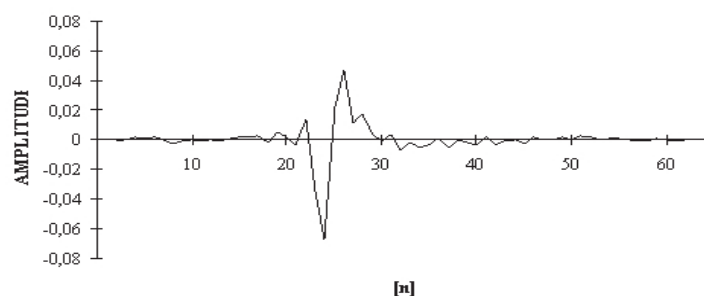


Fig. 9. Acoustic echo path impulse response for an IRISTEL telephone. Most of the coefficients are near zero and only a small subset of the coefficients has a significant value (adapted from Marques et al. (1997)).

In Marques et al. (1997) the authors compare the traditional full-tap FIR and short-length centered filter solution in an echo path with a delay of half a second. The conventional structure converges to a solution where the ERLE is less than 10 dB, while the centered filter achieves approximately 80 dB, as depicted in Fig. 10.

4.1.1 System architecture

Fig. 11 shows the proposed system architecture which is a combined echo cancellation structure including two echo cancelers, one for each communication direction, and a speech detector. Each echo canceler is composed by a centered adaptive filter and a time delay estimator. The delay estimator tracks the corresponding main signal reflection position where the short length adaptive filter is to be centered. The near-end and far-end speech detectors inhibit the adaptation of the filter whenever speech is present.

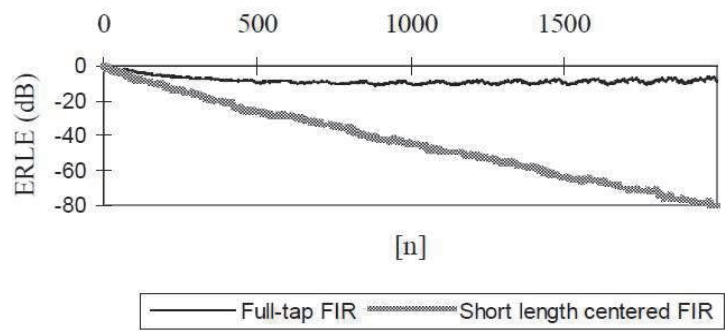


Fig. 10. Convergence of the traditional FIR structure compared to the centered adaptive filter, for a delay of 4000 taps (adapted from Marques et al. (1997)).

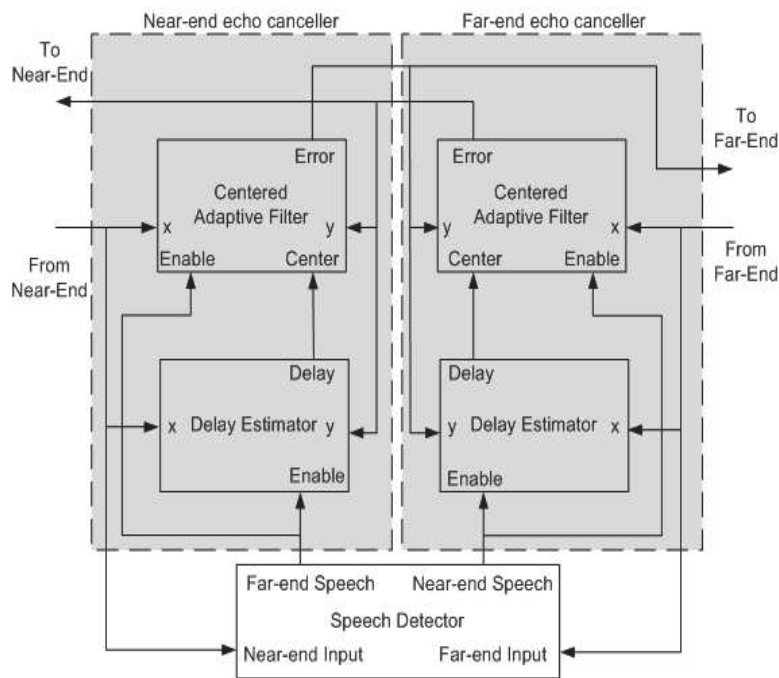


Fig. 11. The block diagram of the proposed system with two echo cancelers and a speech detector. Each echo canceler has a short-length centered adaptive filter and a time delay estimator (from Marques et al. (1997)).

The speech detector is very important in echo cancellation systems where double talking may occur (full duplex mode) as this situation originates the abrupt increase of the adjustment error. The common solution of using adaptive FIR filters to approximate the echo-path impulse response becomes insufficient; if this situation occurs and no action is taken, drift of the adaptive filter coefficients is possible Johnson (1995). Additionally, in this system, erroneous *time-delay estimation* (TDE) may happen. To overcome this problem, the strategy is to inhibit the filters adjustment and the delay estimation when double talking is detected. In Fig. 12 a centered adaptive filter example is shown. The supported echo path length is M taps, the position of the active region is $(M - 1)T_s$ and for illustration purposes only, the considered length is 3 taps. $T_s = 1/F_s$ is the sampling time, that is, the time between two consecutive samples. The main advantages of the centered adaptive filter, as compared to the typical full-length FIR solution, are:

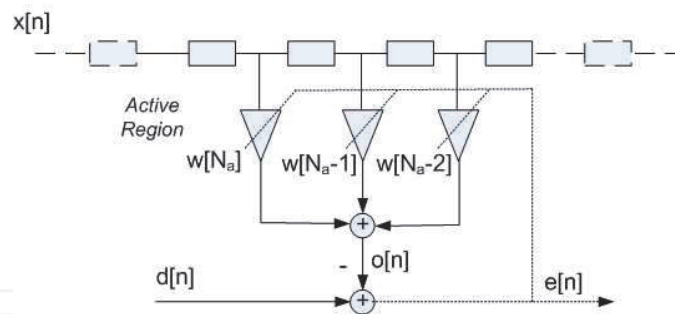


Fig. 12. Centered adaptive filter. The supported echo path length is N_a taps, but considering an active region of 3 taps, only the corresponding 3 coefficients need adjustment.

- reduced computational cost, due to the lower number of coefficients that need adjustment, when compared with the total supported echo path length;
- greater convergence speed, since the adaptation step can now be larger;
- reduced residual error because the coefficients which would otherwise converge to zero, now take precisely that value.

To adjust the short length centered adaptive filter coefficients, the NLMS algorithm was employed, due to its adequacy in the presence of speech signals Haykin (2002).

4.1.2 Time delay estimation

The TDE block estimates the presence of the echo as well as its delay, given by parameter t_d in (1). The basic aspects of TDE are discussed in Jacovitti & Scarano (1993). The *direct cross correlation* (DCC) method is analyzed and compared with the *average square difference function* (ASDF) and the *average magnitude difference function* (AMDF).

The DCC method computes the cross-correlation between signals $x[n]$ and $y[n]$ given by

$$DCC_{xy}[k] = \sum_{k=0}^{L-1} x[n]y[n-k] = \langle x[n], y[n-k] \rangle, \quad (21)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product between its two arguments. Essentially, DCC is the inner product between $x[n]$ and shifted versions of $y[n]$. The DCC can be computed efficiently on the frequency domain using the *fast Fourier transform* (FFT) and its inverse Oppenheim & Schaffer (1999) by

$$DCC_{xy}[k] = \text{IFFT}[\text{FFT}[x[n]] \cdot \text{FFT}[y[-n]]], \quad (22)$$

using L -point FFT/IFFT and \cdot being the dot product between vectors. The maximum value of $DCC_{xy}[k]$ corresponds to the estimated location of the delay $\hat{k} = \arg \max_k DCC_{xy}[k]$; the value of t_d as in (1) is given by $t_d = \hat{k}T_s$. The ASDF estimator is given by

$$ASDF_{xy}[k] = \frac{1}{L} \sum_{k=0}^{L-1} (x[n] - y[n-k])^2, \quad (23)$$

which is similar to the Euclidian distance between two signals. Finally, the AMDF estimator computes

$$AMDF_{xy}[k] = \frac{1}{L} \sum_{k=0}^{L-1} |x[n] - y[n-k]|, \quad (24)$$

with the advantage, over the previous two measures, that it requires no multiplications to measure the similarity between two signals. For ASDF and AMDF we are interested in finding indexes $\hat{k} = \arg \min_k ASDF_{xy}[k]$ and $\hat{k} = \arg \min_k AMDF_{xy}[k]$.

Supported on many experimental tests, the authors in Marques et al. (1997) chose the DCC method, because it outperforms the other two (AMDF and ASDF) for low *signal-to-noise ratio* (SNR) scenarios. The TDE component is the most demanding block on the entire system; it takes about 90 % of the total processing time.

4.1.3 Speech detector

The speech detector copes with double talking situations. When both speakers are talking, the received signal is a composition of the received echo and the other speaker's signal. If no action is taken, adaptive filter coefficient drift may happen as well as erroneous time delay estimation, so the common solution of using adaptive FIR filters to approximate the echo path impulse response becomes insufficient. Thus, whenever speech is present we have to inhibit the coefficients adjustment as well as delay estimation. The speech detector criterion, based on Messerschmidt et al. (1986), is as follows:

- the far-end speaker is considered present when the power of the original signal is above a given threshold, established by the noise average power; we consider that the speaker is still present for more 75 ms, after the power level gets below this threshold;
- the near-end speaker is considered present when the power of the echoed signal is 6 dB below the power of the original signal.

Each filter is adapted only when the corresponding speaker is detected. This way, when speech is detected on both directions, echo cancellation is performed using the coefficients that were updated just before the beginning of the presence of speech.

4.2 Improved short-length centered adaptive filter approach

In Ferreira & Marques (2008) we have an improvement on the approach detailed in Subsection 4.1, and depicted in Fig. 11. The system architecture is the same, but some modifications were placed on the centered adaptive filter as well as on the TDE block. The DSP platform and the hands-free conversation setup used in the experimental tests are also different.

The centered adaptive FIR filter has a small number of coefficients, corresponding to the length of the active area of the echo, with 32 and 64 coefficients for a sampling rate of 8 kHz; this assures fast convergence. The experimental tests showed that is preferable to set to (absolute) zero all the filter coefficients that are near zero and to keep with a non-zero value, only the coefficients on the active area.

On the TDE block, both DCC and ASDF methods were considered. The authors also considered a new approach with a *maximum delay FIR filter*, that has a (large) number of coefficients $\mathbf{c} = [c_0, c_1, \dots, c_{M-1}]$ corresponding to the maximum expected delay. These coefficients are updated by

$$\mathbf{c}_{(i+1)} \leftarrow \mathbf{c}_{(i)} + 2\mu e_{(i)} \mathbf{x}_{(i)}, \quad (25)$$

but only a small fraction ($t \ll M$) of these coefficients is updated between two consecutive sampling times ($125 \mu s$), in order to meet real-time requirements. For the typical scenario, there is no need to update the entire set of coefficients in order to get an accurate estimation of the time delay. The main component of the echo is given by the coefficient with the highest (absolute) amplitude value. Using the DSP parallel instructions Kehtarnavaz (2004), this

update is carried out simultaneously with the update of the centered filter coefficients. In the experimental tests, this delay estimator with low complexity, has obtained good results even in situations of low signal-to-noise ratio. The number of coefficients that need adjustment is small when compared with the total number of elements in the supported delay line, thus enabling a larger step.

4.3 Normalized subband adaptive filter

In J.Ni & Li (2010), the authors address two well-known problems of AEC:

- high correlation between both speech signals (on the input of the adaptive filter);
- the large length of the impulse response of the echo path.

These characteristics slow down the convergence rate of the adaptive filter whenever NLMS algorithm is used. The *normalized subband adaptive filter* (NSAF) offers a good solution to this problem due to its decorrelating property. However, similar to the NLMS-based adaptive filter, the NSAF requires a tradeoff between fast convergence rate and small steady-state MSE. In J.Ni & Li (2010), the authors propose an adaptive combination scheme to address this tradeoff. The combination is carried out in subband domain and the mixing parameter is adapted by means of a stochastic gradient algorithm which employs the sum of squared subband errors as the cost function. The proposed combination scheme obtains both fast convergence rate and small steady-state MSE.

4.4 Hirschman optimal transform

In Krishna et al. (2010), the transform domain adaptive filter approach finds an adequate solution for AEC as it results in a significant reduction in the computational burden, as compared to the traditional approaches. There are some different orthogonal transform based adaptive filters for echo cancellation. In Krishna et al. (2010), the authors present *Hirschman optimal transform* (HOT) based adaptive filter for elimination of echo from audio signals. Simulations and analysis show that HOT based LMS adaptive filter is computationally efficient and has fast convergence compared to LMS, NLMS and Discrete Fourier Transform-LMS (DFT-LMS). The spectral flatness measure shows a significant improvement in canceling the acoustic echo.

4.5 An optimal step approach

The optimal step approach described in Özbay & Kavsaoglu (2010) deals with the AEC problem in the context of practical applications in which the positions of talker, microphone and loudspeaker are not stationary. They explore the fact that the amplitude of the speech signals depend on the acoustic properties of the environment. Choosing a constant step μ does not always yield an optimal result as echo power spectrum density depends on the distance between speaker microphone and loudspeaker which can not be known a priori. The authors show the need for a dynamic choice of the LMS step value. They provide an algorithm for obtaining an optimal step value for any AEC problem.

The authors show that for the optimal μ value in the case that the microphone of the speaker is far from its loudspeaker, adaptive filter output power spectrum density is approximately equal to that of the one obtained for a small μ value. Therefore, in AEC application, the optimal μ value is computed as a function of the echo power spectrum density. By using the proposed algorithm, the step is adaptively set depending on the distance between speaker microphone and loudspeaker and the acoustic properties of the environment.

5. Real-time DSP implementations

This section addresses real-time implementations of long-distance AEC systems on DSP from Texas Instruments with code optimized for real-time processing. We consider AEC on a hands-free system developed on the TMS320C50 and TMS320C6713 DSP platforms.

The TMS320C50 operates at 41 MHz, has a on-chip RAM of 10 K word and a 32 kB PROM with communication kernel Texas Instruments (1993). The *analog interface circuit* (AIC) has 14 bit resolution. The TMS320C6713 Kehtarnavaz (2004) has a master clock of 225 MHz, delivering 1800 MIPS and 1350 MFLOPS. The analog stereo interface is carried out by the AIC 23 codec, with sampling rates from 8 to 96 kHz, with 16, 20 and 24 bits per sample. Fig. 13 shows the block diagram of the *development starter kit* (DSK) C6713 developed by Spectrum Digital¹ for Texas Instruments; the DSK has 192 kB of internal RAM and 16 MB of external RAM.

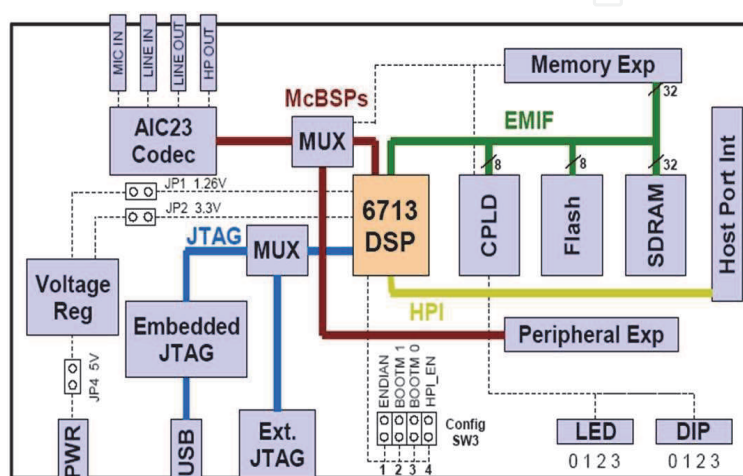


Fig. 13. The block diagram of the Development Starter Kit C6713 developed by Spectrum Digital for Texas Instruments (adapted from Spectrum Digital Inc., DSP Development Systems (2003)).

5.1 The (improved) short-length centered adaptive filter approach

The first approach of AEC system based on centered adaptive filter reported in Marques et al. (1997) and described in Subsection 4.1 was implemented on the TMS320C50 DSP. The approach described in Subsection 4.2 and reported in Ferreira & Marques (2008) was implemented with TMS320C6713 using the DSKC6713. The code, written in C++ programming language, is located on the 192 kB internal RAM, along with the data. The code was compiled with level-3 optimization Kehtarnavaz (2004), for faster execution:

- using allocation of variables to registers;
- elimination of unused code, unused assignments and local common expressions;
- simplification of expressions and statements;
- software pipelining;
- loop optimizations and loop unrolling;
- removal of functions that are never called; simplification of functions with return values that are never used.

¹ www.c6000.spectrumdigital.com

The filters are managed as circular buffers and inline functions are used whenever possible. The sampling rate is 8 kHz, and the number of bits per sample is 16 (the minimum allowed by the AIC23 codec), suited for speech signals. This way, we have 125 μ s between two consecutive samples, and this is the maximum processing time to meet real-time requirements (28125 instructions, under a 225 MHz clock). The time delay estimator has the largest amount of total processing time, being not possible to completely update the time delay estimation, within 125 μ s. Between two consecutive samples, we update only a small portion of the filter coefficients.

5.2 An optimal step approach

The optimal step approach of Özbay & Kavsaoğlu (2010) also uses the Texas Instruments TMS320C6713 with DSKC6713, because it is an up to date architecture. The authors established an experimental setup including the DSKC6713 board, a laptop computer, an amplifier, a loudspeaker, and two microphones. They have considered two scenarios of application:

- in the first scenario, two microphones were placed close to the loudspeaker;
- in the second scenario one microphone was placed close to the loudspeaker and speech trial was implemented in the far-end microphone.

The experimental results show the adequacy of the proposed solution.

6. Experimental results

This section presents some experimental results obtained with the AEC systems described in Subsections 4.1 and 4.2, respectively.

6.1 The short-length centered adaptive filter approach

Using a single TM5320C50 DSP with no external memory, the system detects and cancels an echo with a delay of more than 380 ms. Considering a configuration with 64 Kwords of data memory, the maximum supported delay is larger than 2.5 seconds.

Table 3 shows the computational requirements for a TMS320C50 DSP. Considering an unidirectional configuration and an active region of 4 miliseconds, the maximum supported echo delay is very significant (greater than 2.5 seconds).

Module function	Processor clock-cycles	Percentage
Speech detector	65	2.6
Time-delay estimator	$82+18*corrl$	$3.28+0.72*corrl$
Centered adaptive filter	$114+6*M$	$4.56+0.24*M$
Main Cycle	31	1.2

Table 3. Computational requirements for a TMS320C50 DSP with the AEC approach described in Subsection 4.1. *M* is the supported echo region length (order of FIR filter). The number of computations per sampling period has been reduced by dividing the computation of the cross-correlation function into blocks, each with length *corrl*.

Table 4 describes the main features of the developed AEC system. The maximum length of the echo path is proportional to the available amount of memory. We have two values for this parameter, corresponding to the internal memory of the DSP and the external memory available on the DSK (64 kB), respectively.

Feature	Value
Maximum supported echo delay	381 ms // 2643 ms
Maximum length of dispersion area	4 ms
Absolute delay	0.375 ms
Minimum attenuation on the returned echo	6 dB
Convergence	Improvement of 41 dB in 80 ms
Residual echo level	-51 dBm0
Speech detector level	6 dB below emission level
Hold time after speech detection	75 ms

Table 4. Main features of the AEC approach with TMS320C50 DSP described in Subsection 4.1. The maximum supported echo delay depends on the amount of internal//external memory.

Fig. 14 shows the ERLE (in dB) obtained by the AEC system with simulated, electric, and real echo paths, as a function of time. As expected, we get the best results on the simulated echo path, due to the adequacy of the adaptive filter to this path. The electric echo path is easier to cancel than the acoustic echo path, in which due to its non-linearities, the experimental results show less attenuation than for the other two paths. Even on the acoustic echo path which is the most difficult, we rapidly get 10 dB of attenuation, in less than 30 ms (which is roughly the delay time that a human user perceives the echo); this attenuation stops about -20 dB which is a very interesting value. In summary, ERLE is greater than 41 dB in just 80 ms in a simulated echo path; with real electrical and acoustic echo paths, 24.5 dB and 19.2 dB have been measured, respectively.

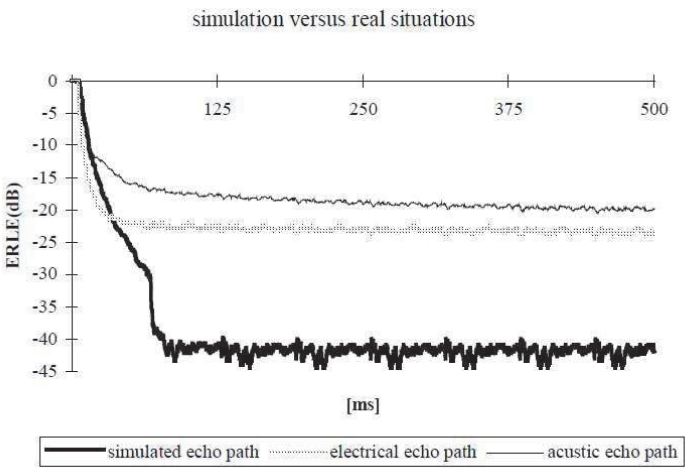


Fig. 14. Echo canceller ERLE in simulated, electric and acoustic paths. On the acoustic path, which is the most difficult we get about 10 dB of attenuation in less than 30 ms.

Table 5 compares this system with the CCITT G.165 recommendation, for a real situation, on the following tests:

- CR - Convergence Rate;
- FERLAC - Final Echo Return Loss After Convergence;
- IRLC - Infinite Return Loss Convergence;
- LR - Leak Rate.

Test Description	CCITT G.165 Requirement	System Performance
CR	≥ 27 dB (500 ms)	41 dB
FERLAC	-40 dBm0	-51 dBm0
IRLC	-40 dBm0	-51 dBm0
LR	≤ 10 dB	≈ 0 dB

Table 5. System performance - comparison with CCITT G.165 recommendation levels for AEC.

We conclude that the system exceeds the recommendation levels for all these tests. The CR and FERLAC measures are taken on the single-talk scenario. Fig. 15 shows the time delay estimator ability to track time varying delays in the presence of real speech signals. On the voiced parts of the speech signals the TDE block tracks the delays accurately.

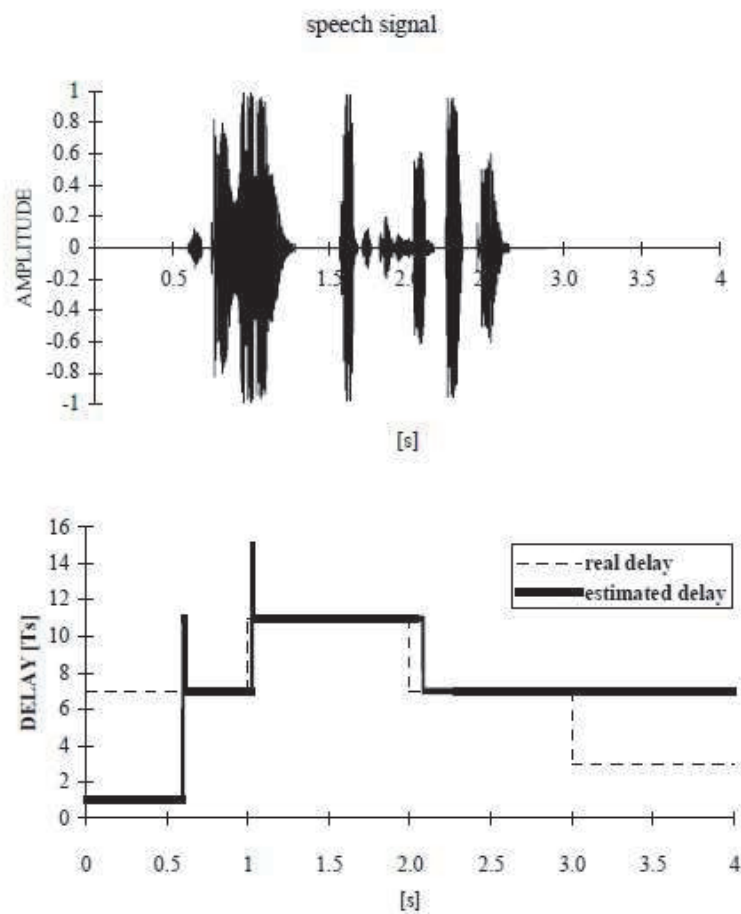


Fig. 15. Real speech signal (top) and real/estimated delay obtained by the TDE module. The TDE block has a good performance on the presence of real speech. Adapted from Marques et al. (1997).

In Fig. 16 the usefulness of the speech detector to prevent the filter coefficient drift is emphasized. In the presence of double talk, with the speech detector disabled the coefficient drift happens.

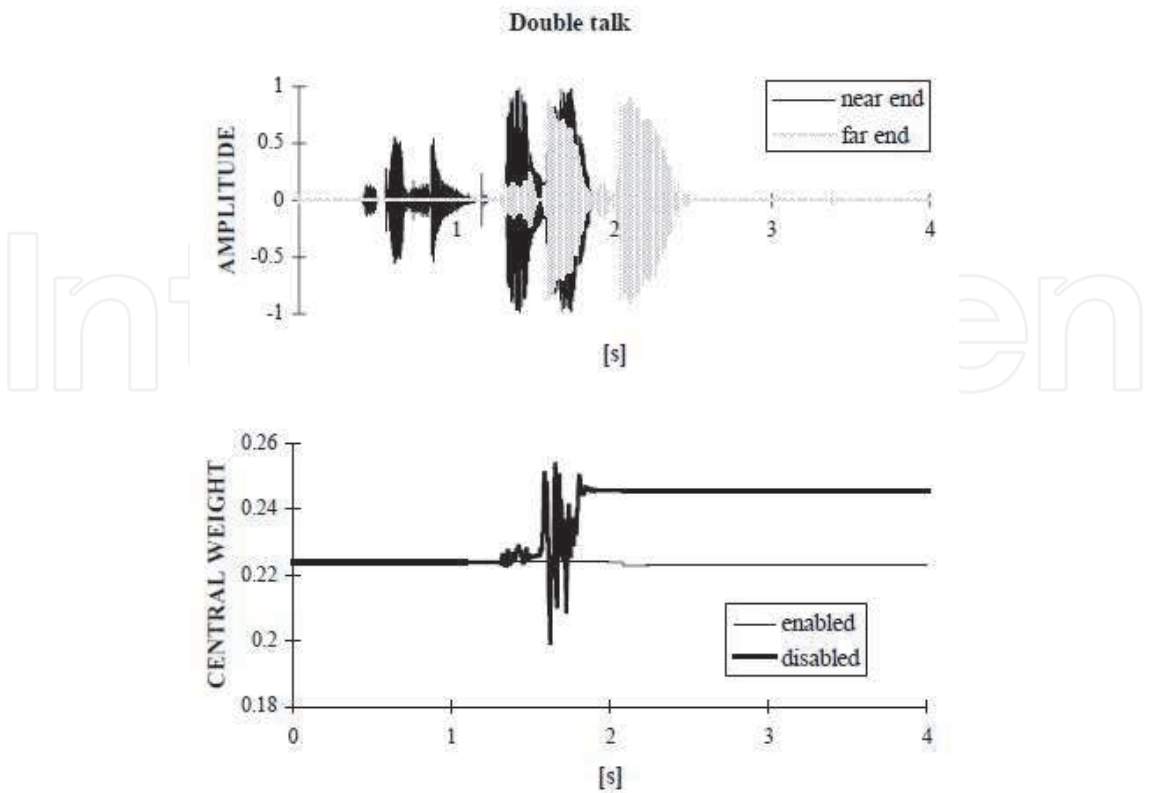


Fig. 16. The speech detector prevents filter coefficient drift in the case of double talk. With the speech detector disabled, coefficient drift happens. Adapted from Marques et al. (1997).

Feature	Value
Absolute delay	0.375 ms
Convergence speed	27 dB (125 ms)
Digitalization	$F_s = 8000H_z$ $n = 16$ bit/sample
Hold time after speech	75 ms
Max. length	256 ms
Max. length of dispersion area	4 ms
Max. memory (data + code)	< 192 kB
Residual echo	-42.26 dBm0
Returned echo minimum loss	6 dB
Speech detector	6 dB below threshold

Table 6. Main features of the AEC approach with TMS320C6713 DSP described in Subsection 4.2.

6.2 The improved short-length centered adaptive filter approach

The tests were carried out in DSP Code Composer Studio (CCS) environment, with code written in C++, using real signals. Table 6 summarizes the developed system features. The total amount of memory needed for the echo canceler data and code is low (and proportional to the maximum expected delay) making it suited for an embedded system. The total amount of memory required can be reduced, using a fixed-point DSP. The adaptive centered filters

have 32 or 64 coefficients, while FIR-based time delay estimator uses up to $M=4000$ coefficients (delays up to 0.5 seconds), giving a reasonable range of delays, suited for several applications. Fig. 17 shows the (typical) centered adaptive filter coefficients (with 32 active coefficients), for a speech signal. Only a small subset of coefficients is far from zero according to the echo path characteristics, as expected; this is a typical test situation. Fig. 18 displays the echo and error

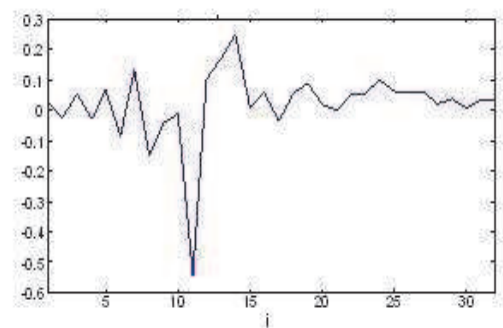


Fig. 17. Centered adaptive filter coefficients. Only a small subset of coefficients is far from zero.

signals for a speech signal, while Fig. 19 displays the achieved attenuation, of about 20 dB, for the speech signal on its voiced parts. It is interesting to note that how attenuation increases on the voiced parts of the speech signal, according to the AEC fundamental concepts presented in Subsections 2.1 and 2.2.

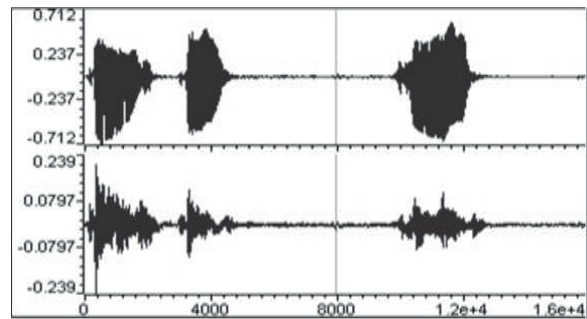


Fig. 18. Echo (top) and error (bottom) signal. Whenever there is echo with higher energy the adaptive filter error signal follows it. On its portions with higher energy, the error signal shows a decaying behavior that corresponds to the convergence of the adaptive filter.

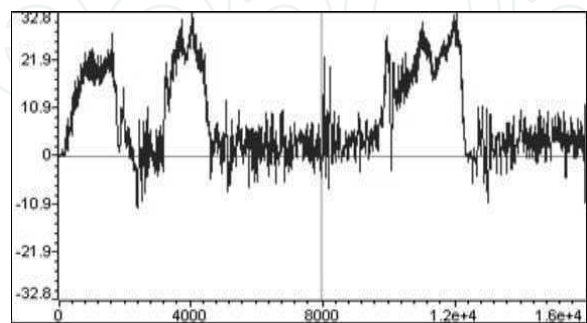


Fig. 19. Attenuation obtained for the speech signal of Fig. 18. We have increased attenuation on the voiced parts of the speech signal.

Table 7 compares our system with the CCITT G.165 recommendation levels, for a real conversation. We conclude that this system approaches the recommendation levels for

Test Description	CCITT G.165 Requirement	System Performance
CR	≥ 27 dB (500 ms)	27 dB (125 ms)
FERLAC	-40 dBm0	-37.39 dBm0
IRLC	-40 dBm0	-37.39 dBm0
LR	≤ 10 dB	≈ 0 dB

Table 7. System performance - comparison with CCITT G.165 recommendation levels for AEC.

FERLAC and IRLC measures, matches for CR and exceeds it for the LR measure. The CR and FERLAC measures are taken on the single-talk scenario.

Fig. 20 displays the attenuation obtained for several electric and acoustic echo paths, using the average power of the received echo as the reference value, because the attenuation on the acoustic channel is not constant along these tests. The attenuation for the simulated echo path

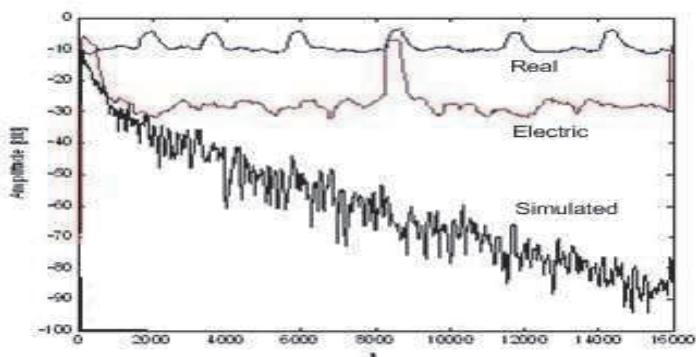


Fig. 20. Attenuation for the echo paths real (acoustic), electric and simulated (real-time on CCS).

is much larger than the other two, as expected. On the other hand, the attenuation for the electric echo path is around 30 dB, which is a considerable value. Finally, for the acoustic path we get about 10 dB of attenuation, which is also an acceptable practical value. This result was expected due to the strong non-linearities in the acoustic echo path, caused by the loudspeakers and microphone.

7. Conclusions

In this chapter, we have addressed the problem of acoustic echo cancellation. Echo being a delayed and attenuated version of the original signal produced by some device, such as a loudspeaker, causes disturbing effects on a conversation. This problem arises in many telecommunication applications such those as hands-free systems, leading to need of its cancellation in real-time. The echo cancellation techniques have better performance than the oldest and simpler echo suppression techniques.

We have reviewed some concepts of digital, statistical, and adaptive filtering. Some solutions for real-time acoustic echo cancellation based on adaptive filtering techniques, on digital signal processors were described in detail.

We have addressed some implementation issues of adaptive filtering techniques in real-time. After the description of the acoustic echo cancellation solutions in some detail, we have focused on their real-time implementations on well known digital signal processor platforms,

discussing its main characteristics as well as their experimental results according to standard measures.

7.1 Open challenges: future work

Although adaptive filtering techniques have been proved efficient for the echo cancellation problem, there are some open challenges that lead to directions of future work. One of the most important directions of current and future research, due to its importance and difficulty is to model the non-linear echo path. Since we use linear filters to model the echo path, it is not possible to guarantee a complete echo cancellation when the echo path is non-linear. In these situations, better results can be obtained with non-linear filters or with linear filters complemented by non-linear functions. The challenge is thus positioned at choosing adequate non-linear filters and non-linear functions that accurately model the echo path, being able to achieve even better and faster cancellation results.

8. Acknowledgments

This work has been partly supported by the Portuguese *Fundação para a Ciência e Tecnologia* (FCT) Project FCT PTDC/EEA-TEL/71996/2006.

9. References

- Antweller, C. & Symanzik, H. (1995). Simulation of time variant room impulse responses, *IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP'95*, Vol. 5, Detroit, USA, pp. 3031–3034.
- Benesty, J., Gaensler, T., Morgan, D., Sondhi, M. & Gay, S. (2001). *Advances in Network and Acoustic Echo Cancellation*, Springer-Verlag.
- Birkett, A. & Goubiran, R. (1995). Acoustic echo cancellation using NLMS-neural network structures, *IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP'95*, Detroit, USA.
- Breining, C. (1999). Acoustic echo control. an application of very-high-order adaptive filters, *Digital Signal Processing* 16(6): 42–69.
- Ferreira, A. & Marques, P. (2008). An efficient long distance echo canceler, *International Conference on Signals and Electronic Systems, ICSES'08*, Krakow, pp. 331–334.
- Gay, S. & J. Benesty (2000). *Acoustic signal processing for telecommunications*, Kluwer Academic Publishers.
- Gilloire, A. & Hänsler, E. (1994). Acoustic echo control, *Annals of Telecommunications* 49: 359–359. 10.1007/BF02999422.
URL: <http://dx.doi.org/10.1007/BF02999422>
- Greenberg, J. (1998). Modified LMS algorithms for speech processing with an adaptive noise canceller, *IEEE Transactions on Signal Processing* 6(4): 338–351.
- Hänsler, E. (1994). The hands-free telephone problem: an annotated bibliography update, *Annals of Telecommunications* 49: 360–367. 10.1007/BF02999423.
URL: <http://dx.doi.org/10.1007/BF02999423>
- Haykin, S. (2002). *Adaptive Filter Theory*, 4th edn, Prentice-Hall.
- Instruments, T. (1986). Digital voice echo canceler with a TMS32020, I: 415–454.
- Instruments, T. (1993). TMS 320C5X users guide.
- Jacovitti, G. & Scarano, G. (1993). Discrete time techniques for time delay estimation, *IEEE Transactions on Signal Processing* 41(2): 525–533.

- Jeannès, R., Scalart, P., Faucon, G. & Beaugeant, C. (2001). Combined noise and echo reduction in hands-free systems: A survey, *IEEE Transactions on Speech And Audio Processing* 9(8): 808–820.
- J.Ni & Li, F. (2010). Adaptive combination of subband adaptive filters for acoustic echo cancellation, *IEEE Transactions on Consumer Electronics* 56(3): 1549 – 1555.
- Johnson, C. (1995). Yet still more on the interaction of adaptive filtering, identification, and control, *IEEE Signal Processing Magazine* pp. 22 – 37.
- Kehtarnavaz, N. (2004). *Real-Time Digital Signal Processing*, Newnes.
- Krishna, E., Raghuram, M., Madhav, K. & Reddy, K. (2010). Acoustic echo cancellation using a computationally efficient transform domain lms adaptive filter, *10th International Conference on Information Sciences Signal Processing and their Applications (ISSPA)*, Kuala Lumpur, Malaysia, pp. 409–412.
- Kukrera, O. & Hocanin, A. (2006). Frequency-response-shaped LMS adaptive filter, *Digital Signal Processing* 16(6): 855–869.
- Kuo, S. & Lee, B. (2001). *Real-Time Digital Signal Processing*, John Wiley & Sons.
- Kuo, S. & Pan, Z. (1994). Acoustic echo cancellation microphone system for large-scale videoconferencing, *IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP'94*, Vol. 1, Adelaide, Australia, pp. 7–12.
- Liu, Z. (1994). A combined filtering structure for echo cancellation in hand-free mobile phone applications, *International Conference on Signal Processing Applications & Technology*, Vol. 1, Dallas, USA, pp. 319–322.
- Marques, P. (1997). *Long distance echo cancellation using centered short length transversal filters*, Master's thesis, Instituto Superior Técnico, Lisbon, Portugal.
- Marques, P. & Sousa, F. (1996). TMS320C50 echo canceller based on low resolution time delay estimation, *The First European DSP Education and Research Conference*, Paris.
- Marques, P., Sousa, F. & Leitaó, J. (1997). A DSP based long distance echo canceller using short length centered adaptive filters, *IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP'97*, Munich.
- Messerschmidt, D., Hedberg, D., Cole, C., Haoui, A. & Winship, P. (1986). Digital voice echo canceller with a TMS320C30, *Digital Signal Processing Applications* 1: 415–454.
- Oh, D., Shin, D., Kim, S. & Lee, D. (1994). Implementation of multi-channel echo canceler for mobile communication with TMS320C50, *International Conference on Signal Processing Applications & Technology*, Vol. 1, Dallas, USA, pp. 259–263.
- Oppenheim, A. & Schaffer, R. (1999). *Discrete-Time Signal Processing*, 2nd edn, Prentice Hall International, Inc.
- Orfanidis, S. (ed.) (2007). *Optimum Signal Processing*, 2nd edn, McGraw-Hill.
- Özbay, Y. & Kavsaoglu, A. (2010). An optimum algorithm for adaptive filtering on acoustic echo cancellation using tms320c6713 dsp, *Digit. Signal Process.* 20: 133–148.
URL: <http://portal.acm.org/citation.cfm?id=1663653.1663859>
- Sayed, A. (2003). *Fundamentals of Adaptive Filtering*, Wiley-IEEE Press.
- Sondhi, M. (2006). The history of echo cancellation, *IEEE Signal Processing Magazine* 23(5): 95 – 102.
- Sondhi, M. & Berkeley, D. (1980). Silencing echoes on the telephone network, *Proceedings of the IEEE* 68(8): 948–963.
- Spectrum Digital Inc., DSP Development Systems (2003). TMS320C6713 DSK technical reference.
- Veen, B. & Haykin, S. (1999). *Signals and Systems*, John Wiley & Sons.

- Widrow, B., Glover, J., McCool, J., Kaunitz, J., Williams, C., Hearn, R., Zeidler, J., Dong, E. & Goodlin, R. (1975). Adaptive noise cancelling: Principles and applications, *Proceedings of the IEEE* 63(12): 1692 – 1716.
- Widrow, B. & Stearns, S. (1985). *Adaptive Signal Processing*, Prentice Hall.
- Witowsky, W. (1999). Understanding echo cancellation in voice over IP networks, *International Conference on Signal Processing Applications and Technology (ICSPAT)*, Orlando.

IntechOpen

IntechOpen



Adaptive Filtering

Edited by Dr Lino Garcia

ISBN 978-953-307-158-9

Hard cover, 398 pages

Publisher InTech

Published online 06, September, 2011

Published in print edition September, 2011

Adaptive filtering is useful in any application where the signals or the modeled system vary over time. The configuration of the system and, in particular, the position where the adaptive processor is placed generate different areas or application fields such as prediction, system identification and modeling, equalization, cancellation of interference, etc., which are very important in many disciplines such as control systems, communications, signal processing, acoustics, voice, sound and image, etc. The book consists of noise and echo cancellation, medical applications, communications systems and others hardly joined by their heterogeneity. Each application is a case study with rigor that shows weakness/strength of the method used, assesses its suitability and suggests new forms and areas of use. The problems are becoming increasingly complex and applications must be adapted to solve them. The adaptive filters have proven to be useful in these environments of multiple input/output, variant-time behaviors, and long and complex transfer functions effectively, but fundamentally they still have to evolve. This book is a demonstration of this and a small illustration of everything that is to come.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Artur Ferreira and Paulo Marques (2011). Echo Cancellation for Hands-Free Systems, Adaptive Filtering, Dr Lino Garcia (Ed.), ISBN: 978-953-307-158-9, InTech, Available from:

<http://www.intechopen.com/books/adaptive-filtering/echo-cancellation-for-hands-free-systems>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen