

# A DSP BASED LONG DISTANCE ECHO CANCELLER USING SHORT LENGTH CENTERED ADAPTIVE FILTERS

Paulo A. C. Marques\*

Fernando M. G. Sousa\*

José M. N. Leitão\*\*

\*Instituto Superior de Engenharia de Lisboa, Portugal  
E-mail: pmarques@cc.isel.pt, fsousa@cc.isel.pt

\*\*Instituto Superior Técnico, Portugal  
E-mail: jleitao@red.lx.it.pt

## ABSTRACT

This paper describes an implementation of a long distance echo canceller which copes with double talking situations and exceeds the CCITT G.165 recommendation. The proposed solution is based on short length adaptive 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 greater echo return loss enhancement and shows faster convergence speed as compared with results reported in recent literature.

## 1. INTRODUCTION

With the development of digital technologies like ISDN, mobile digital telephones and teleconferencing systems, the involved delays due to temporal multiplexing, channel coding and speech coding are becoming very significant, and comparable with those of long-distance communications. Echo cancellation in this kind of systems is presently a problem. For satellite communications the far-end echo can have a delay greater than half a second. The typical solution, as stated in [1], would need an adaptive FIR filter with more than 4000 taps (with 8KHz sampling rate) and the same number of associated coefficients. The computational effort needed is such that it is virtually impossible to implement it in real-time. Even if it was possible, the adaptation step would have to be so small [1] that the convergence speed would be unacceptable, and, due to finite precision effects, the achieved solution after convergence would be inadmissible. However, as stressed in [2], the echo-path impulse response is characterised mainly by two active regions, corresponding to the near-end and the far-end signal echoes (see Figure 1). Each active region has a length that is usually much shorter than the total supported echo-path length. We present a system based on time-delay

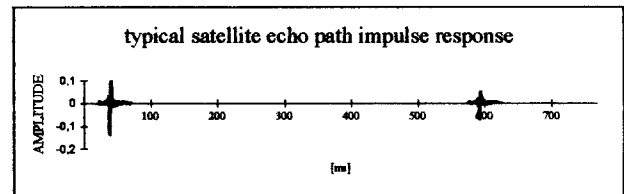


Figure 1. Typical echo path impulse response.

estimators to track the position of these active regions, where short-length adaptive filters have to be centered.

The resulting algorithm has reduced computational effort due to the lower number of coefficients that need adjustment, when compared with the total echo-path length, and achieves greater convergence speed, since the adaptation step can now be larger [1,2]. The residual error is reduced because the coefficients, which would otherwise converge to zero, now take precisely that null value.

In Figure 2 the traditional full-tap FIR and short-length centered filter solution are compared by simulation (due to the huge requirements of the full-tap structure) in an echo path with half a second delay. Although the simulation precision is 32 bit floating point, the conventional structure converges to a solution where the echo return loss enhancement (ERLE) is less than 10dB, while the centered filter achieves approximately 80dB.

## 2. PROPOSED APPROACH

As shown in Figure 3, the proposed system is a combined echo cancellation structure including two echo cancellers, one for each communication direction, and a speech detector. Each echo-canceller 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 speech detector is very important in echo cancellation systems where double talking may

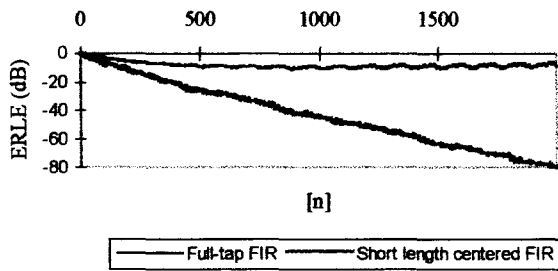


Figure 2. For a delay of 4000 taps the traditional structure converge for an unsatisfactory ERLE.

occur[3,4,5,6] 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[3,5,6]. Additionally, in the proposed system, erroneous time-delay estimation may happen[6]. The strategy is to inhibit the filters adjustment and the delay estimation when double talking is detected.

### 3. IMPLEMENTATION ASPECTS

The system is based on the Texas Instruments Digital Signal Processor TMS320C50; it can however be supported on other DSP architectures.

Reference [7] deals with basic aspects of time delay estimation (TDE) based on sampled signals. The direct cross correlation method (DC) is analysed and compared with the average square difference function (ASDF) and the average magnitude difference function (AMDF). Supported on [7] and on our simulation results[6,8], we have chosen the DC method, as it outperforms the others (AMDF and ASDF) for low signal-to-noise ratios (SNR).

The following model is considered for signals  $x(t)$  and  $y(t)$  received by two sensors:

$$\begin{cases} x(t) = s(t) + n_1(t) \\ y(t) = As(t - D) + n_2(t) \end{cases} \quad (1)$$

They correspond to two differently delayed and scaled versions of the same signal  $s(t)$ , with measurements noises  $n_1(t)$  and  $n_2(t)$  that are mutually uncorrelated and with  $s(t)$ . Both signal and noises are assumed realisations of zero-mean stationary processes characterised by their autocorrelation functions. The problem is to find an estimate  $\hat{D}$  of the true delay  $D$  using a finite set of samples of  $x(t)$  and  $y(t)$ .

The cross-correlation function between  $x$  and  $y$  is estimated according to:

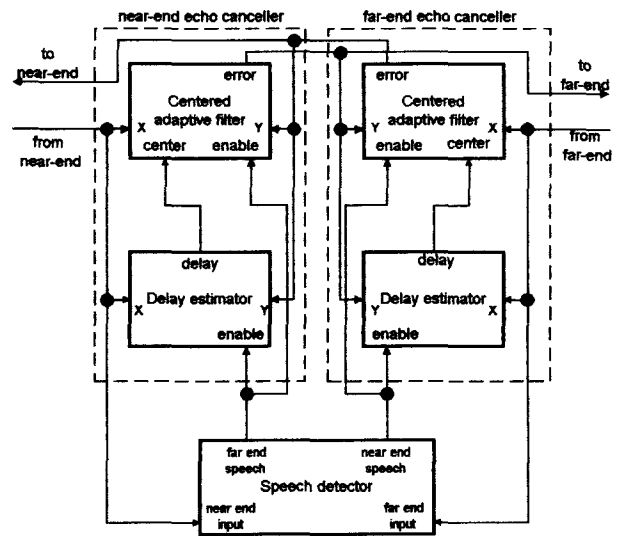


Figure 3. Proposed echo canceller structure.

$$\hat{R}_{xy}(\tau) = \frac{1}{M} \sum_{k=1}^M x(kT)y(kT + \tau) \quad (2)$$

The estimated delay  $\hat{D}$  is the absolute maximum of the previously estimated function:

$$\hat{D} = \arg \max_{\tau} \hat{R}_{xy}(\tau) \quad (3)$$

Usually an accurate delay estimate is needed. Our system uses a sample period resolution, thus avoiding interpolation. The technique consists in searching the absolute extreme of the cross-correlation function corresponding to each direction. As shown in Table 1, the delay estimation module is the one that requires more computational power, as it is concerned with all the echo path. 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*. These operations are performed over several sampling periods, thus leading to longer delay estimation times. The resulting degradation can be seen in Figure 4, for an echo path delay of half a second.

The short length adaptive filters are designed according to the normalised least mean squares algorithm (NLMS)[1]. They track the previously determined (active) regions, leading to an accurate echo path estimation. The filter output  $o$  at iteration  $i$  is given by:

$$o(i) = \sum_{k=0}^{ncoeffs-1} W_k(i)y(i - k - center + ncoeffs/2) \quad (4)$$

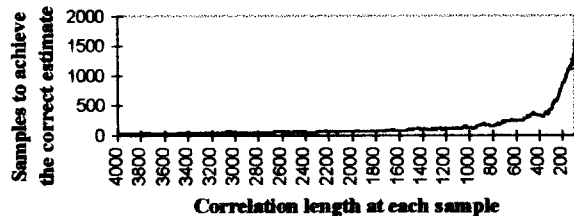


Figure 4. The cross-correlation computation by blocks leads to slower time-delay estimation.

where *center* is the active region center, *ncoefs* the supported region length, and  $w_k$  the filter coefficients. The supported length was chosen to double that of the typical active region. In this way fluctuations on the estimated delay are supported without requiring the adaptive filter repositioning. Re-centering occurs only if the distance between  $\hat{D}$  and *center* implies that the short length filter is no longer adapting to the active region.

The speech detector is a modified version of the algorithm presented in [3]. Assuming that the echo path return loss is at least 6dB, near-end or far-end speech are declared when (5) or (6), respectively, are verified:

$$|r_n(i)| \geq \frac{1}{2} \max\{|r_f(i)|, |r_f(i-1)|, \dots, |r_f(i-N_f)|\} \quad (5)$$

$$|r_f(i)| \geq \frac{1}{2} \max\{|r_n(i)|, |r_n(i-1)|, \dots, |r_n(i-N_n)|\} \quad (6)$$

where  $r_n$  and  $r_f$  are the received signal from the near-end and the far-end signal, respectively;  $N_f$  and  $N_n$  are the corresponding number of memorised elements.

If the supported delays are significant, the comparisons in (5) and (6) can be computationally demanding. Alternatively, comparisons are made with the recursively computed power estimates given by:

$$\hat{r}_n(i+1) = (1-\alpha)\hat{r}_n(i) + \alpha|r_n(i)| \quad (7)$$

$$\hat{r}_f(i+1) = (1-\alpha)\hat{r}_f(i) + \alpha|r_f(i)| \quad (8)$$

with  $\alpha=2^{-5}$  as in [3]. Since the detector is based on short-term power peaks, it is maintained for 75ms[3] after initial detection.

Table 1 shows the computational requirements for a TMS320C50 DSP. Considering a unidirectional configuration and an active region of 4ms, the maximum supported echo delay is very significant (greater than 2.5 seconds).

simulation versus real situations

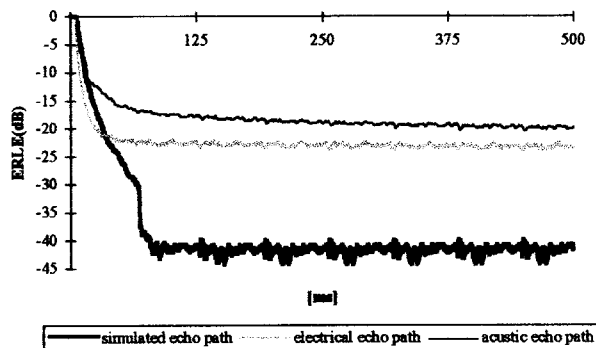


Figure 5. Echo canceller ERLE in simulated and real echo paths.

Module function	Processor clock-cycles	%
Speech detector	65	2.6
Time-delay estimator	$82+18*\text{corr1}$	$3.28+0.72*\text{corr1}$
Centered adaptive filter	$114+6*n\text{coefs}$	$4.56+0.24*n\text{coefs}$
Main cycle	31	1.2

Table 1. Module requirements for a TMS320C50 DSP.

#### 4. PERFORMANCE EVALUATION

Evaluation has been made by simulation and in real situations. The system exceeds the CCITT G.165 recommendation in simulated echo paths (Table 2). The ERLE is greater than 41dB in just 80ms for a simulated echo path (Figure 5). With real electrical and acoustic echo paths, 24.5dB and 19.2dB have been measured (respectively).

Test description	CCITT G.165 performance requirement	System performance
Final echo return loss after convergence, singletalk mode	-40dbm0	- 51 dbm0
Convergence rate, singletalk mode	$\geq 27\text{dB}$	41 dB
Leak rate	$\leq 10\text{dB}$	$\cong 0\text{ dB}$
Infinite return loss convergence	-40dBm0	- 51 dbm0

Table 2. Echo canceller performance.

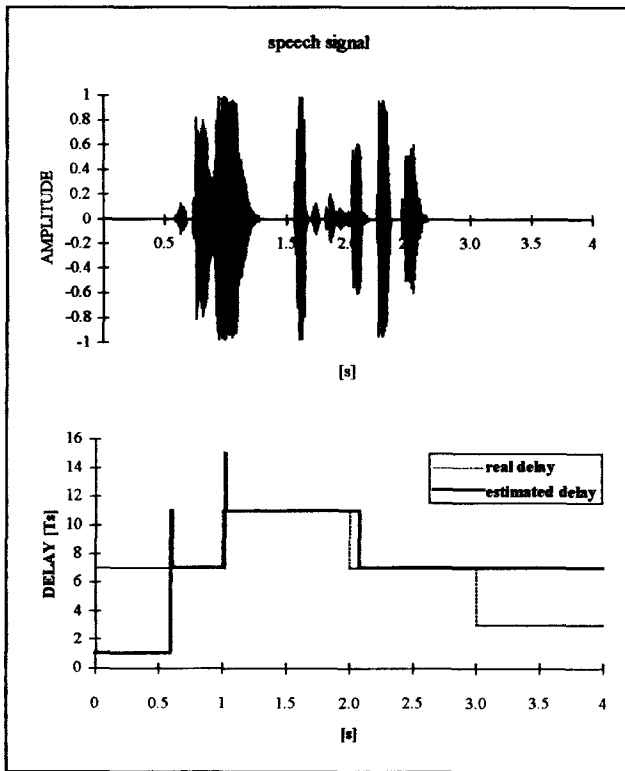


Figure 6. Time-Delay estimation in a time-varying delay using a real speech signal.

Figure 6 shows the time delay estimator ability to track time varying delays in the presence of real speech signals. In Figure 7 the usefulness of the speech detector to prevent the filter coefficients drift is emphasised (for clarity only the central weight is shown).

## 5. CONCLUSIONS

In this work we presented and evaluated an echo canceller system exploiting the fact that the typical impulse response of an echo path has a set of active regions which are only a small fraction of the total length. Slowly time-varying echoes are tracked by time-delay estimators. Short length adaptive filters, centered on those positions, assure an accurate echo-path estimation. A speech detector is employed to disable adjustments in double talk situations.

The developed structure exceeds the CCITT G.165 recommendation and enables the support of long echo paths, with low computational requirements. Additionally it achieves faster convergence speed and greater ERLE, when compared with other systems reported in recent literature[6].

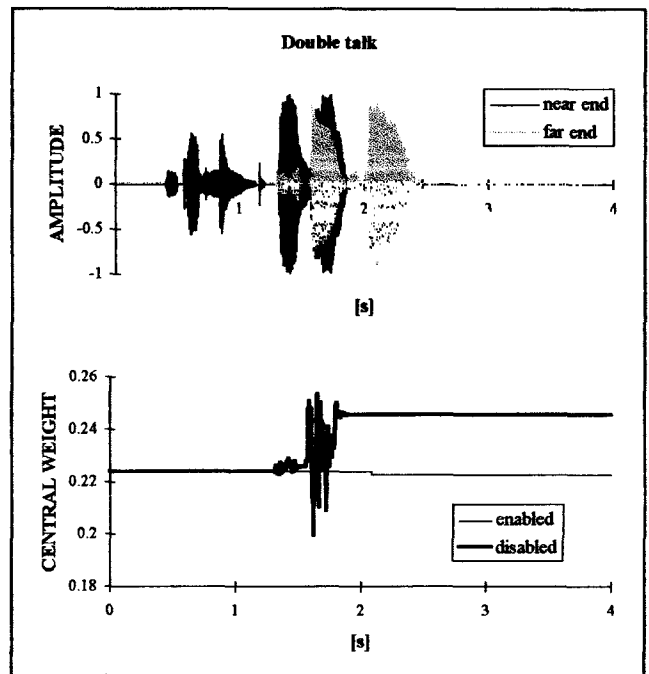


Figure 7. With the speech detector disabled the coefficient drift happens in the presence of double talk.

## REFERENCES

- [1] S. Haykin, Adaptive Filter Theory, 2nd Edition, Prentice-Hall, 1991.
- [2] V. A. Margo, D. M. Etter, "Multiple Short-Length Adaptive Filters for Time-Varying Echo Cancellation", Proc. of ICASSP 93, Vol 1, pp 161-164, 1993.
- [3] D. Messerschmitt, D. Hedberg, C. Cole, A. Haoui, P. Winship, "Digital Voice Echo Canceller with a TMS32020", Digital Signal Processing Applications, Volume I, Texas Instruments, pp 415-454, 1986.
- [4] C. Carlemalm, F. Gustafsson, B. Wahlberg, "On the Problem of Detection and Discrimination of Double Talk and Change in the Echo Path", Proc. of ICASSP 96, Vol 5, pp 1295-1298, 1996.
- [5] C. Johnson Jr., "Yet Still More On the Interaction of Adaptive Filtering, Identification, and Control", IEEE Signal Processing Magazine, pp 22-37, March 1995.
- [6] P. Marques, Long Distance Echo Cancellation Using Centered Short-Length Transversal Filters, Master Thesis (in Portuguese), Instituto Superior Técnico, Lisbon, 1996.
- [7] G. Jacovitti and G. Scarano, "Discrete Time Techniques for Time Delay Estimation", IEEE Transactions on Signal Processing, Vol 41, N 2, pp 525-533, February 1993.
- [8] P. Marques, F. Sousa, "TMS320C50 Echo Canceller Based on Low Resolution Time-Delay Estimation", Proc. of The First European DSP Education and Research Conference, Texas Instruments, Paris, 1996.