A Hardware Optimised CMOS Adaptive Noise Canceller Implementation

A.Th. Schwarzbacher[1], M. Herz[1,2], F. David[1,3] and J. Timoney[4]

[1]*Dublin Institute of Technology, Ireland*
[2]*University of Applied Sciences – Darmstadt, Germany*
[3]*ESSTIN: School of engineers in Nancy, France*
[4]*NUI Maynooth, Co. Kildare, Ireland*

andreas.schwarzbacher@dit.ie, ma_herz@gmx.de, davidfra@caramail.com, jtimoney@cs.may.ie

Abstract: Up until now, the main foci of development in mobile communication equipment have been to decrease its size and to extend its battery operation times. However, further reductions in the size of devices are physically limited by the user interface requirements and therefore, alternative aspects of these devices must be targeted for enhancement by designers. A feature of mobile communications equipment is the variety of environments within which they are used, so algorithms that can improve the quality of a transmission are highly desirable. In this paper, mobile telephony devices are being specifically considered and a CMOS implementation of the filter block of an adaptive noise canceller will be presented. Results will then be given to demonstrate how this circuit can significantly increase speech quality by suppressing interfering noise without requiring any prior assumptions on its properties.

1 Introduction

A variety of approaches are available for removing noise from speech signals. Possibly the most obvious is to use two microphones, one that will capture the speech and noise signals, and another that will attempt to capture the noise on its own, and then to subtract one from the other. In this case, the quality of the filtered signal depends on the estimation of the noise reference signal [1]. However, this approach is not be suitable for many applications, particularly mobile communications devices. Other research has tried to form a reference noise signal by assuming that the noise is stationary and that the average signal determined during periods classified as “silence” is representative of the noise [1] [2]. However, this approach is also unsatisfactory, as noise can not be assumed to be stationary in a mobile environment and the estimation of the noise signal using a finite sample could lead to gross inaccuracies. Additionally, the silence decision would not be error free. Lastly, this technique cannot be applied to quantisation noise. Bypassing all the problems associated with these approaches, the principle of noise cancelling assumes no *a priori* knowledge of the statistics of the noise and actually uses in the noisy signal to form the reference input to the filter. Figure 1 shows the principle of this approach.

The adaptive noise canceller works because speech is a quasi-periodic signal, and a section of noisy speech $x(n)$ delayed by one or two pitch periods, T, will be highly correlated with the true speech signal $s(n)$ but will be uncorrelated with the additive noise $w(n)$. The role of the adaptive filter is to minimise the energy of the system output $w'(n)$ and thus produce a signal $s'(n)$ that is, in the LMS sense, the optimum estimate of $s(n)$. Both the speech and noise are time varying and thus the filter must adapt to changes in the properties of the input, and the LMS algorithm determines the trajectory of the changing filter weights to keep the energy of its output at a minimum [4]. The dynamics of speech means its pitch does not remain constant and thus, for best performance, this has to be estimated in an iterative manner by some pitch estimation module.

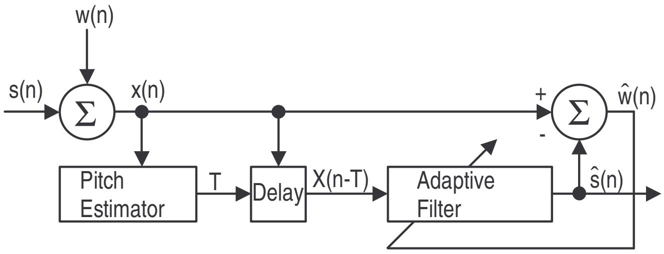

Figure 1: Adaptive Filtering Approach to Remove Noise from Speech

In the case of unvoiced speech however, no pitch value exists and the simplest solution is to freeze the filter weights at their current value until voiced speech occurs again [1]. The structure of the adaptive filter is finite impulse response (FIR), with the output of the filter being given by

$$\hat{s}(n) = \sum_{i=0}^{L} b_i x(n-i-T) \tag{1}$$

where T is the analysed pitch period for the speech, b_i denotes the filter coefficients and L is the filter order. The filter coefficients adapted using the LMS algorithm [1] whose update equation is given by

$$B_{n+1} = B_n + 2\mu\hat{w}(n)X_{n-T} \tag{2}$$

where the noise estimate

$$\hat{w}(n) = x(n) - \hat{s}(n) \tag{3}$$

$X_n = (x(n), x(n-1),...x(n-L))$, B_n denotes the coefficient vector $(b_0,b_1,\dots,b_L)$ at the time n, and μ is the step size. The stepsize factor controls stability and rate of convergence. If the Adaptive Noise Canceller starts with an arbitrary coefficient vector, the algorithm will converge and remain stable as long as the parameter μ is greater than zero but less than the reciprocal of the largest eigenvalue λ_{max} of the input correlation matrix R [3].

2 Implementation of the Adaptive Noise Canceller

The adaptive noise canceller was implemented using the ES2 ECPD07 CMOS technology. All designs were written in abstract VHDL and synthesised using Synopsys Design Compiler without any design constraints.

Figure 2 shows the block diagram of the adaptive filter part of the adaptive noise canceller. The input signals for the adaptive filter are the noisy input data *x(n)* as well as the pitch period *T*. The adaptive filter consists of four main blocks. The first block is the delay structure used to delay the input data by one pitch period. The second part is the transversal filter that filters the delayed input signal. This filter has an order of ten and a coefficient bit width of 33. The third block is the least mean square algorithm that calculates and updates the coefficients. The last block of the adaptive filter implementation is a buffer for storing the updated coefficients. Furthermore, for completeness the control structure is also shown in Figure 2.

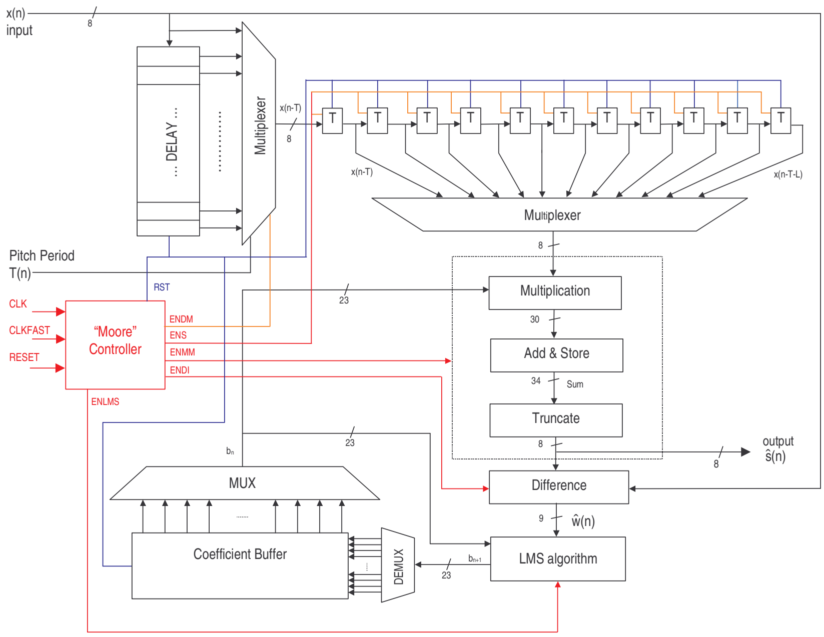

Figure 2: Block diagram of the Adaptive Filter showing the different Modules

3 Results

To demonstrate the performance of the adaptive filter, and by way of benchmarking, noisy synthetic speech signals were created and then filtered. The synthetic sound produced was the vowel 'A' with a fixed pitch frequency of 125Hz. This speech signal consists of three formant frequencies which are f_1=730Hz, f_2=1090Hz, and f_3=2440Hz [4]. The signal is distorted by white gaussian noise (WGN) and has a signal to noise ratio (SNR) of 0dB. The results are presented in Figures 3 to 6. The magnitude spectrum of the noisy and filtered signals is shown in Figures 3 and 4. It can be seen that the filtered version retains the spectral shape of the input signal with the formant frequencies remaining prominent. Hence, the perceptual characteristics of the signal are, in the main, unchanged. Furthermore, the noise component is reduced in filtered signal, being particularly noticeable in the higher frequencies from 2000Hz to 4000Hz where it is nearly completely reduced. However, in the region 0Hz to 1500Hz, although the adaptive filter manages to reduce the noise, remnants of the noise component and some attenuation of the lower harmonics can be observed.

Figures 5 and 6 shows performance from the alternative perspective of the spectrogram of the distorted signal and of the filtered signal respectively. Again, the figures show clearly that the noise component of the higher frequencies 2000Hz to 4000Hz is well reduced but that in the region of 0Hz to 1500Hz some noise remains and the harmonic attenuation persists over time. Thus, it can be concluded that the filter reduces the noise component while not significantly changing the perceptual characteristic of the speech signal. A final illustration, Figure 7 presents the results of applying the LAR distance speech quality measure [5] to the distorted and filtered signals. This measure is based on a finding a set of Linear Predictive Coefficients (LPC) for each frame of the distorted/filtered speech signals and the original clean speech, transforming them into Log Area Ratio (LAR) coefficients and then calculating the difference between them. This measure was shown to have a correlation coefficient of 0.62 with subjective speech quality assessment data [6]. Figure 7 demonstrates that the adaptive filter improves the quality of the noisy signal, with the distance in this case being shortened by 33%.

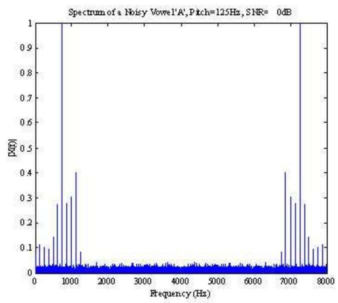

Figure 3: Spectrum of a noisy Vowel 'A' with Pitch Freq. 125Hz and SNR=0dB

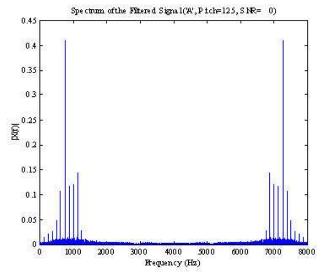

Figure 4: Spectrum of filtered Vowel 'A' with Pitch Freq. 125Hz and SNR=0dB

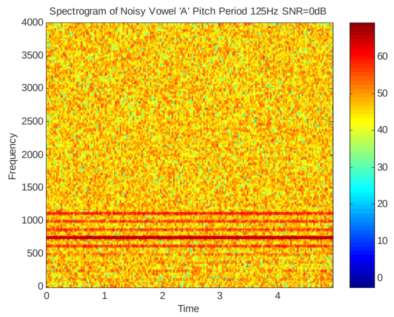

Figure 5: Spectrogram of Noisy Signal ('A' Pitch Freq. 125Hz SNR=0dB)

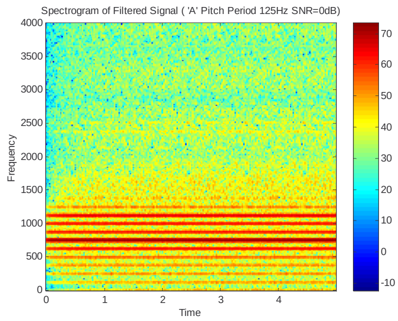

Figure 6: Spectrogram of Filtered Signal ('A' Pitch Freq. 125Hz SNR=0dB)

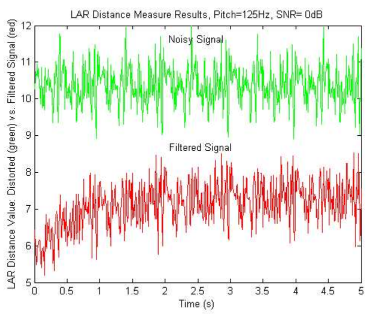

Figure 7: LAR Distance Measure Result, Pitch=125Hz, Input Signal SNR=0dB

4 Conclusion

This paper has presented a hardware optimised CMOS implementation of the adaptive filter part of an adaptive noise canceller. It was shown that the filter does not degrade the characteristics of a signal while successfully suppressing the noise, improving the speech quality by around 33% for an SNR of 0dB. This suggests that it may be possible to realise a small hardware-based filter that could be targeted specifically at the mobile communications market.

5 References

[1] R. McAulay, "Optimum classification of voiced speech, unvoiced speech and silence in the presence of noise and interference," *Lincoln Lab. Tech. Note*, 1976-7, June 1976.

[2] S. Boll, "Improving linear prediction analysis of noisy speech by predictive noise cancellation,"*1977 IEEE Proc. Int. Conf. Acoust. Speech, Signal Processing,* pp. 10-12, May 9, 1977.

[3] Marvin R. Sambur, "Adaptive Noise Cancelling for Speech Signals," *IEEE Transactions on acoustics, speech, and signal processing, vol. ASSP-26*, *No. 5*, pp. 419-423 October 1978.

[4] M. Slaney, *Auditory toolbox for Matlab*, http://rvl4.ecn.purdue.edu/~malcolm/interval/1998-010/, 1998.

[5] John H. L. Hansen and Bryan L. Pellom, "An effective quality evaluation protocol for speech enhancement algorithms," http://www.ee.duke.edu/Research/Speech (May 2002)

[6] S.R. Quackenbush, T.P. Barnwell and M.A. Clements, *Objective measures of speech quality*, Prentice Hall , Englewood Cliffs, NJ, 1988.