SUB–BAND, DUAL–CHANNEL ADAPTIVE NOISE CANCELLATION USING NORMALISED LMS

David J. Darlington, Douglas R. Campbell

Department of Electrical and Electronic Engineering, University of Paisley,
High Street, Paisley PA1 2BE, Scotland, UNITED KINGDOM.

Tel: **+44 141 848 3428** *Fax*: **+44 141 848 3404**
E–mail: `darl_ee0@helios21.paisley.ac.uk`

Abstract
An adaptive noise cancellation scheme for speech processing is proposed. In this, the adaptive filters are implemented in frequency–limited sub–bands, based on a simplified model of the human cochlea. A modification to the basic LMS structure is introduced which guarantees stability and convergence at all times. This modification, a non–recursive normalisation, is used both in a wideband and in a sub–band implementation of the scheme. The effect of this normalisation on the quality of the processed speech is to cause the speech to be distorted, showing that there is no benefit to using normalised LMS in a sub–band scheme, whether the application uses classical or intermittent noise cancellation.

1. INTRODUCTION

1.1 Speech enhancement system

A multi–channel, sub–band adaptive system for enhancement of speech signals corrupted by background noise is being developed. Enhancement in this context means improvement of the quality or intelligibility of the speech signal, by reduction of background noise or speech distortion, and hence improvement of the signal to noise ratio of the contaminated speech. This may be desired to render the speech more intelligible either to a human listener (e.g. in a hearing aid system or hands–free mobile telephone), or an automatic speech recogniser.

1.2 Sub–band scheme based on cochlear model

A sub–band system decomposes the wideband input signals into a number of frequency–limited signals, in similar fashion to how the human ear treats incoming signals. A significant advantage of using sub–band processing for speech enhancement is that it allows for diverse processing in each sub–band depending on signal power, noise power and level of coherence between signal and noise in the two channels. Implementing a classical adaptive noise cancellation scheme in a number of frequency–limited sub–bands also permits faster convergence of the filter coefficients due to the reduction of signal power and adaptive filter length in each sub–band. The system under development is shown in Fig.1 below. In this work, all sub–bands are linearly spaced in the frequency domain. The effect of modifying the distribution of filters has also been investigated by the authors **[Darlington and Campbell 1996]**.

1.3 Modification to scheme

In the work reported here, a modified algorithm for updating the filter coefficients is investigated. This modification has been widely reported as being useful for guaranteed convergence and stability in a non–stationary environment such as speech processing, but its performance in a sub–band environment had not been reported.

2. ADAPTIVE NOISE CANCELLATION

The sub–band speech enhancement scheme described here is an extension of that of Toner and Campbell **[1993]** and Campbell **[1994]**.

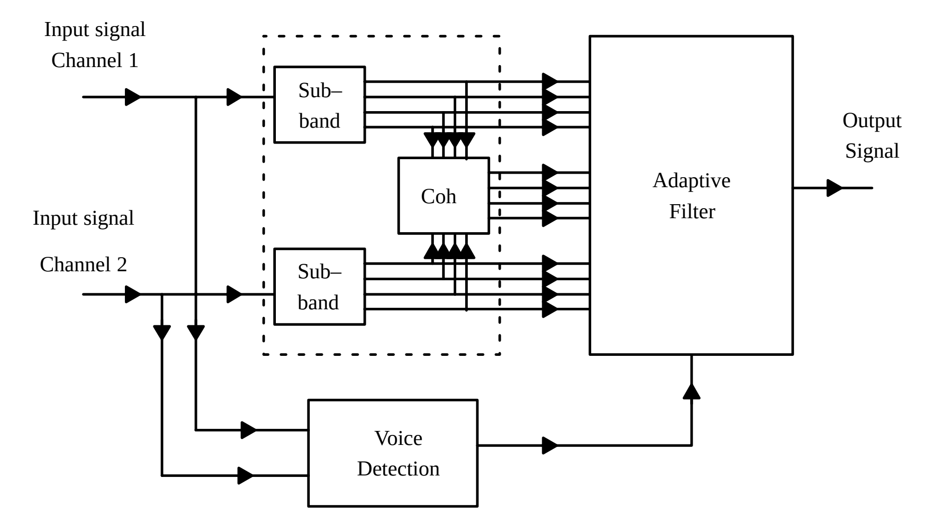

Fig.1: Sub–band adaptive speech enhancement system under development.

It uses the least mean squares (LMS) algorithm in an adaptive noise cancellation scheme **[Widrow and Stearns 1985]** to model the differential transfer function between noise signals in a number of sub–bands. In the classical noise canceller, shown in Fig.2 below, it is assumed that desired speech (***s***) is present only in one of the two channels (the primary) and that the noise signal (n_2) at the reference input is highly correlated with the noise signal (n_1) at the primary. The adaptive filter weights will converge to the differential transfer function between the two inputs, resulting in filter output ***y*** being an estimate of only the noise present in signal ***d***. The output ***e*** will therefore be an estimate of speech signal ***s***.

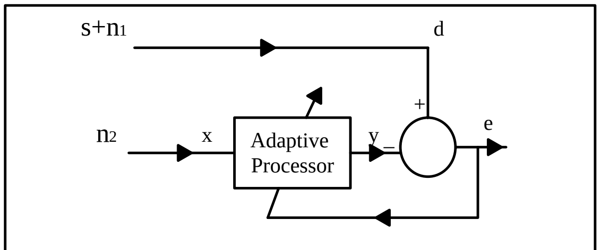

Fig.2: Classical Adaptive Noise Canceller

The sub–band approach reduces the problem of identifying a single, lengthy impulse response between the two channels to one of identifying a set of shorter, parallel filters, with approximately the same computational complexity as conventional LMS (CLMS). Toner and Campbell reported that the sub–band LMS (SBLMS) approach considerably improved the mean square error (MSE) convergence rate. The improvement in algorithm convergence has been investigated by Mahalanobis et al **[1993]** and the authors.

3. WEIGHT UPDATE ALGORITHMS

3.1 The LMS algorithm

The sub–band adaptive noise cancellation scheme using the conventional LMS algorithm requires an adaptation constant (stepsize)μ, which controls the rate of convergence and stability of the weight update (W).

$$W_{k+1} = W_k + 2\mu e_k x_k \tag{1}$$

where ***e*** and ***x*** have the same meaning as in Fig.2 and ***k*** is the iteration index. **Tarrab and Feuer [1988]** asserted that users of the LMS tend to be overly conservative in their choice of stepsize, causing unnecessarily slow convergence. A recent suggestion for an optimal stepsize is given in **Moir [1994]**.

In practical applications, CLMS uses an estimate of the maximum stepsize for stability, which is inversely proportional to both signal power and adaptive filter length. In speech processing work, the limit on the stepsize tends to be lower than this, due to the non–stationarity of the signals. In this work, CLMS is defined as LMS using a recursive, long–term estimate of signal power, updated at each iteration.

3.2 Normalised LMS

In the normalised LMS (NLMS) algorithm **[Bershad 1986]**, a non–recursive, short–term variance estimation is used:

$$W_{k+1} = W_k + \frac{\mu e_k x_k}{\sum_{i=1}^{N} x^2_{k-i+1}} \tag{2}$$

where ***N*** is the adaptive filter length. The normalisation provides protection against power level fluctuation, guaranteeing convergence and stability at each iteration for any stepsize in the range 0–2. **Clarkson [1993]** has claimed than in applications such as speech enhancement where the input signals exhibit widely fluctuating power levels, normalisation is essential. CLMS as defined above has been reported **[Mahalanobis et al 1993]** as a form of normalisation, but does not exhibit the convergence and stability characteristics of NLMS. Previous work by the authors showed that the improvement in performance of the sub–band implementation of CLMS (SBCLMS) was due to the increase of stepsize permitted by the reduced power and filter order in each band. It was of interest to determine whether implementing NLMS in a sub–band scheme (SBNLMS) would show a similar improvement.

4. MSE CONVERGENCE

An adaptive modelling experiment **[Widrow and Stearns 1985]** was performed using white noise input signals of unity variance decomposed into 1, 4, 16 and 64 sub–bands. Identical signals were fed to the two input channels i.e. the transfer function being modelled by the filters was unity. The results of 50 experiments were ensemble averaged and are presented as log MSE plots in Fig.3. This shows that

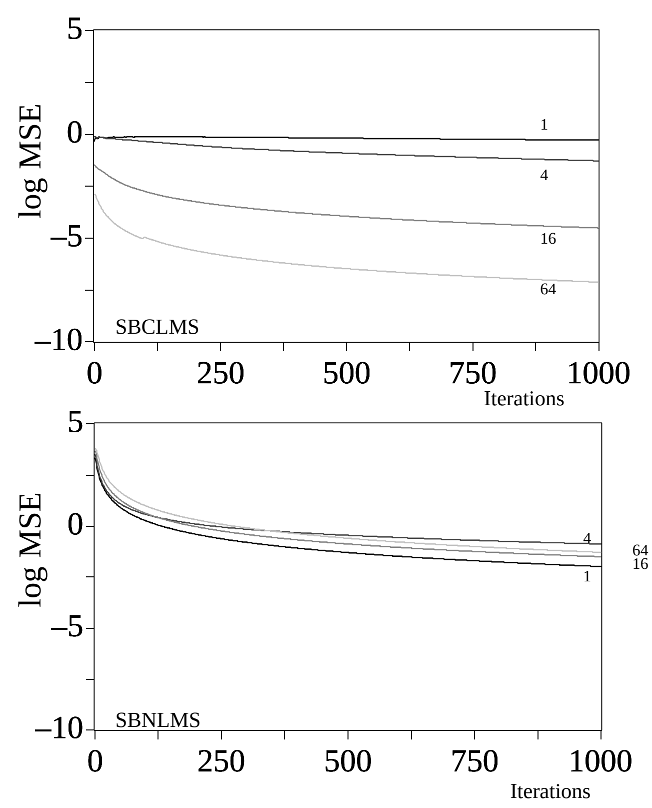

Fig.3: MSE convergence of CLMS and NLMS algorithms wideband and in sub–bands.

while the convergence speed of CLMS is improved by implementation in sub–bands, that of NLMS is marginally degraded. This is because the convergence speed of CLMS is improved by the reduction of power and filter length in each sub–band, whereas that of NLMS is unaffected by changes in signal power. However, there would appear still to be an advantage in using NLMS, even when the performance of that algorithm is degraded by a sub–band implementation, since the performance is unaffected by long or short–term changes in signal power, both of which affect CLMS.

5. RESULTS

5.1 Classical noise canceller

The steady–state response of the NLMS algorithm is much noisier than that of CLMS, due to the short–term normalisation. The implications of this become apparent when NLMS is used in a classical noise cancellation scheme to remove noise added to a clean, anechoic speech signal sampled at 10kHz (Fig.5(a)). For this first test, it was desired to check the convergence and steady–state characteristics of the modified algorithm without the added complications of echoes and signal leakage induced by a simulated reverberant environment. A white noise signal of comparable power and the same length (Fig.5(b)) was generated. This white noise formed the reference input to the adaptive noise canceller. The noise was also convolved with a simple linear transfer function of length 22, and added to the clean, anechoic speech to produce the primary input signal (Fig.5(c)). The signal–to–noise ratio in this channel was approximately –10dB. The adaptive filter weights in the wideband case should converge to those of the linear transfer function, and the filter act to cancel the noise.

Tests were performed using two different wideband adaptive filter lengths, 256 and 2048. The control test using CLMS used an adaptation stepsize of 1/50th of the maximum, which was necessary to avoid instability problems. In this case an intermittent, non–stationary signal (speech) was being used at the primary input only, and so the variance estimator would not adjust for short–term changes in speech power.

The SNR of the signal could be improved by up to 30dB using both CLMS and NLMS, in both wideband and sub–band implementation. However, on listening to the processed speech it is evident that using NLMS in the sub–band implementation results in highly distorted speech. Using NLMS, convergence to the desired weight values was achieved, but the variations in the filter coefficients around the asymptotic values resulted in distortion of the desired speech signal. In the single–band implementation, the normalised LMS approach did remove the noise efficiently with little distortion apparent, but distortion of the speech became more pronounced with an increase in the number of sub–bands (M). This can be attributed to (i) reduction in magnitude of the x_k^2 term at each iteration k and (ii) reduction of the summation term in the denominator by a factor of M. The denominator of the weight update is therefore considerably reduced in value compared with the wideband case, resulting in much larger changes in weight values with each iteration. The output of the 16–band CLMS scheme is given in Fig.5(d), with the corresponding NLMS implementation in Fig.5(e). Fig.4(a) shows the SNR improvement (SNRI) achieved using NLMS, compared with the CLMS baseline. SNRs were estimated by discounting an initial convergence period of 2s, and then calculating average signal and noise power during speech and speech–free periods of the signal respectively.

Implementing NLMS in a classical sub–band architecture did not, therefore, seem beneficial. Use of the NLMS algorithm in an intermittent ANC, wherein adaptation is paused when the presence of speech in both channels is detected, was also investigated.

5.2 Intermittent noise canceller

A similar experiment was set up, but in this case the clean anechoic speech was also added to the reference input signal. This loosely simulates two microphones being placed close to a talker but more distant from the unwanted noise. The start and endpoints of the speech were manually labelled, and it was ensured that no speech energy was present during the denoted 'silence' periods. During these periods, the filters should attempt to converge to model the differential noise transfer function. The filter weights are frozen during speech periods.

However, in this case although distortion of the signal during the actual speech periods is less evident, a new problem has been introduced when using NLMS. Once initial convergence has occured, the filter will cancel the noise efficiently during speech periods. However, the variation in steady–state response of the adaptive filter means that the noise previously apparent at all times is now only present during speech–free periods while adaptation is ongoing. This affects the SNR adversely (Fig.4(b)) and is more distracting due to being 'switched on and off'. A plot of the 16–band output is shown in Fig.5(g), wherein noise during the speech–free periods is evident.

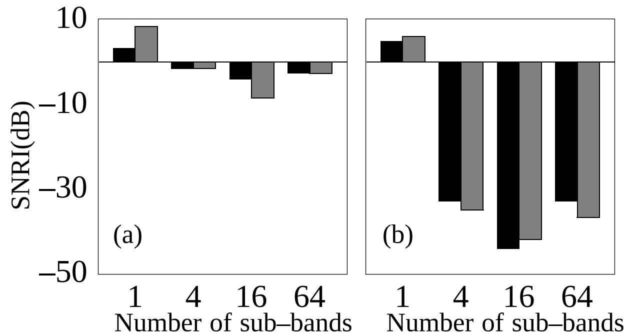

Fig.4:
Signal–to–noise ratio improvement vs number of sub–bands (a): anechoic speech, classical noise cancellation, (b): anechoic speech, intermittent noise cancellation.

FILTER LENGTH ■ 256 ▨ 2048

5.3 Simulated reverberant environment

Additional tests, similar to those of section 5, are being carried out within a simulated reverberant environment, transfer functions for which are calculated using the image modelling method of Allen and Berkley **[1979]**, still in use by researchers such as Culling et al **[1994]**. However, at present the results of this work are less conclusive.

6. CONCLUSION

There appears to be no advantage in using NLMS in a classical sub–band ANC. Despite the guaranteed stability and convergence of the algorithm, the convergence speed of NLMS cannot be improved further by splitting into sub–bands, unlike CLMS, and considerable distortion may ensue. In an intermittent application, the distortion is less apparent but induced noise is very obvious during speech–free periods.

REFERENCES

J. B. Allen, D. A. Berkley, *Image method for efficiently simulating small room acoustics*, Journal of the Acoustical Society of America vol. 65 no. 4, April 1979, pp.943–950.

N. J. Bershad, *Analysis of the normalized LMS algorithm with Gaussian inputs*, IEEE Transactions on Acoustics, Speech and Signal Processing vol.34 no.4, August 1986 pp.793–806.

D.R Campbell, *Adaptive speech enhancement with diverse sub–band processing,* EUSIPCO–94 Signal Processing VII: Theories and Applications.

P. M. Clarkson, *Optimal and Adaptive Signal Processing*, CRC Press 1993.
J. F. Culling, Q. Summerfield, D. H. Marshall, *Effects of simulated reverberation on the use of binaural cues and fundamental–frequency differences for separating concurrent vowels*, Speech Communication 14 (1994), pp.71–95.
D. J. Darlington, D. R. Campbell, *Sub–band adaptive filtering applied to speech enhancement*, ESCA Tutorial and research workshop on the Auditory Basis of Speech Perception, Keele University, U.K. 15th – 19th July 1996.
A. Mahalanobis, S. Song, S.K. Mitra and M. R Petraglia, *Adaptive FIR filters based on structural subband decomposition for system identification problems*, IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing vol.40 no.6, June 1993.
T. J. Moir, *The application of servo theory to the LMS algorithm*, EUSIPCO–94 Signal Processing VII: Theories and Applications.
M. Tarrab and E. Feuer, *Convergence and performance analysis of the normalized LMS algorithm with uncorrelated Gaussian data*, IEEE Transactions on Information Theory vol.34 no.4, July 1988 pp.680–691.
E. Toner and D. R. Campbell, *Speech enhancement using sub–band intermittent adaption*, Speech Communication 12 (1993), pp.253–259.
B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice Hall, 1985.

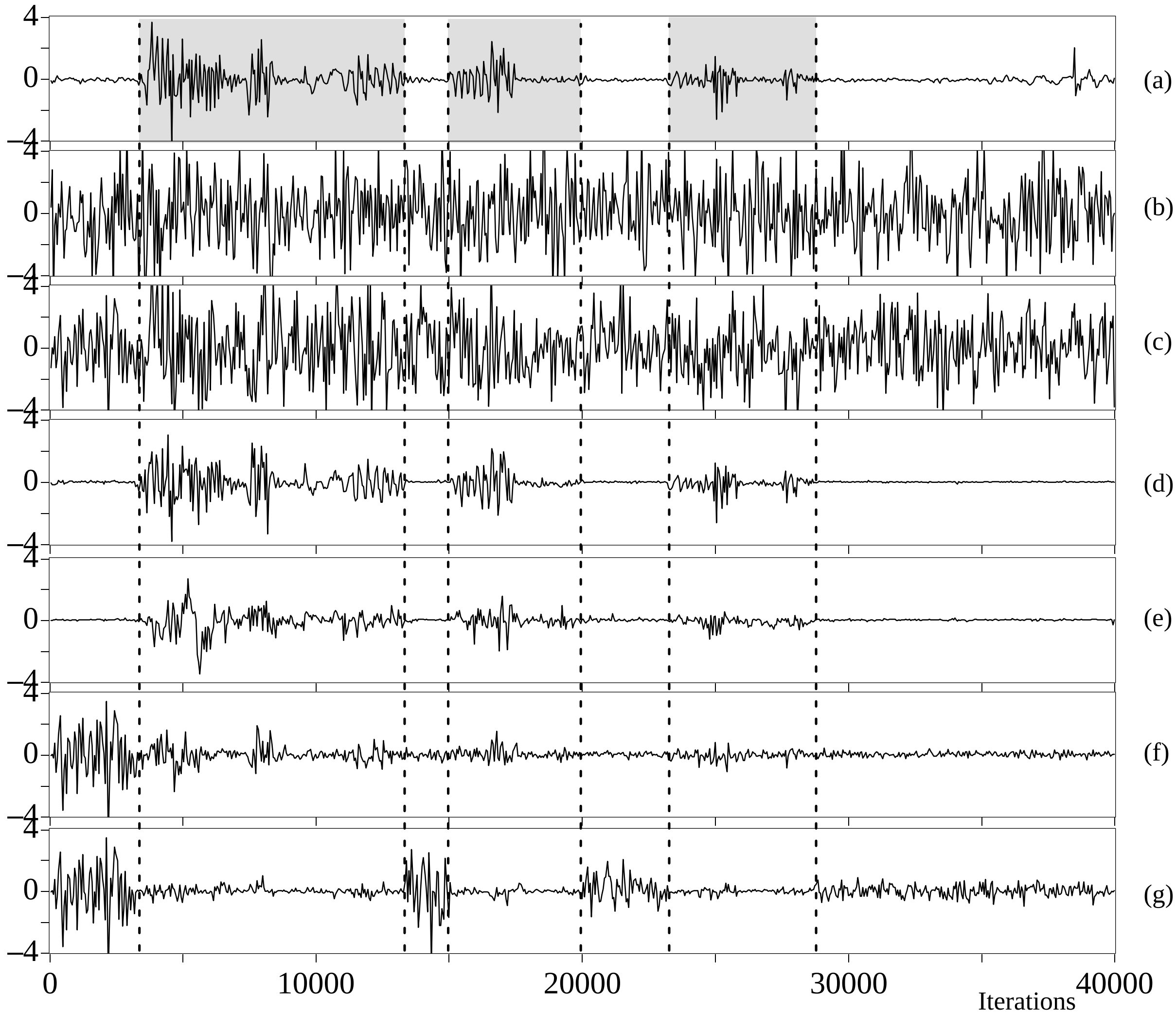

Fig.5: (a) Clean speech signal. Shaded area shows sections containing speech energy.
(b) Simulated white noise signal (reference input to classical ANC).
(c) Speech plus filtered white noise (primary input to ANCs)
(d) Processed signal using 16–band CLMS in classical ANC (cf (a)).
(e) Processed signal using 16–band NLMS in classical ANC.
(f) Processed signal using 16–band CLMS in intermittent ANC.
(g) Processed signal using 16–band NLMS in intermittent ANC.